

621.391(075.8)

В48

В. П. Вінницький
В. Г. Поліщук

**Термінальне
устаткування
та передавання
інформації
в телекомунікаційних
СИСТЕМАХ**

ПОЛІТЕХНІКА
НТУУ «КПІ»



621.391(075.8)

В48

Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний інститут»

В. П. Вінницький

В. Г. Поліщук

Термінальне устаткування та передавання інформації в телекомунікаційних системах

За загальною редакцією члена-кореспондента
НАН України М. Ю. Ільченка

*Затверджено Міністерством освіти і науки України
як підручник для студентів вищих технічних навчальних закладів,
які навчаються за напрямом «Телекомунікації»*



478646

621.391(075.8) В48 2004

Вінницький В. П. Термінальне устаткування та п

Київ

«Політехніка»

2004

ДЕПЕРІРОВАНО
2012
КНИГОСХОВИЩЕ

УДК 621.397.6 + 621.398(075.8)

ББК 32.88я73

В48

*Гриф надано Міністерством освіти і науки України
(Лист № 14/18.2-2441 від 18 листопада 2004 р.)*

Рецензенти:

*В. В. Петров, чл.-кор. НАН України
(ІПРІ НАН України)*

*Л. Н. Беркман, д-р техн. наук, проф.
(ДУІКТ)*

Вінницький В. П.

В48 Термінальне устаткування та передавання інформації в телекомунікаційних системах : Підруч. для студ. вищ. техн. навч. закл., які навчаються на напрямом «Телекомунікації» / В. П. Вінницький, В. Г. Поліщук. – К.: ІВЦ “Видавництво «Політехніка»”, 2004. – 436 с.: іл. – Бібліогр.: с. 425–426.

ISBN 966–622–179–9

Розглянуто в систематизованому вигляді телекомунікаційні системи (ТС), які утворюються на ресурсах мереж термінальними кінцевими устаткуваннями даних (КУД) абонентів: ТС телефонного, вокодерного, комп’ютерного та факсимільного зв’язку. Усі види ТС розглянуто з єдиної позиції класичного визначення системи зв’язку. Теорію утворення аналогової (мовні сигнали) і дискретної форм повідомлень, передавання їх трактами (каналами) у разі протокольного керування обміном інформаційними повідомленнями між КУД абонентів застосовано до розроблення, проектування та технічної експлуатації ТС. Висвітлено методи підвищення достовірності та розрахунків надійності під час передавання інформації трактами (каналами) мереж, а також методи системного аналізу динамічних процесів функціонування ТС.

Для студентів вищих технічних навчальних закладів, які навчаються за напрямом «Телекомунікації».

УДК 621.397.6 + 621.398(075.8)

ББК 32.88я73 + 32.94я73

478646

ISBN 966–622–179–9

© В. П. Вінницький,
В. Г. Поліщук, 2004



Вступ

Життя і діяльність людства нерозривно пов'язані з інформацією, її зберіганням, передаванням й обробленням. Оскільки інформація частіше з'являється не в тих місцях, де її використовують, то одна з найважливіших проблем на всіх стадіях розвитку людського суспільства – це її передавання. Це завдання вирішують за допомогою ресурсів телекомунікаційних мереж (ТМ).

Сучасним телекомунікаційним системам (ТС), створюваним на середовищах ТМ, надається пріоритетність в їх розвитку та поширенні завдяки тій ролі, яку вони відіграють у різноманітних сферах людської діяльності – економіці та промисловості, науці та культурі, будівництві, транспорті та ін. Ці системи утворюють інфраструктуру суспільства, яка об'єднує людей всього світу та яка надає можливість спілкуватись людям у будь-якому місці та в будь-який час.

Основу створення ТС становлять розподілені ТМ як середовище до-
ставляння інформації (повідомлень) між об'єктами, що спілкуються, та кінцевим обладнанням – термінальним устаткуванням (ТУ), яке ще називають кінцевим устаткуванням даних (КУД), DTE – Data Terminal Equipment.

Вивчаючи дисципліну «Термінальне устаткування та передача інформації в телекомунікаційних системах», будемо користуватись як терміном «термінальне устаткування», так і «абонентський пункт» (АП). Коли йтиметься про структуру та принципи конструктивного виконання складових частин, здебільшого в частині перетворення сигналів, будемо користуватись терміном ТУ, а якщо про функціональні можливості, яким користується абонент, – терміном АП.

Телекомунікаційні системи утворюються на каналах (трактах) ТМ, що дозволяє розглядати їх з позиції класичного вивчення системи зв'язку *джерело інформації – канал зв'язку – одержувач інформації*, у якій інформація у вигляді повідомлень або сукупності сигналів є об'єктом транспортування технічними пристроями мереж по каналах зв'язку.

Підручник складається з трьох частин. У першій частині розглянуто ТС телефонного зв'язку і термінальні пристрої перетворення мовних звукових повідомлень. У цій частині детально вивчаються процес формування мовних звукових сигналів та їх характеристики, які відображаються фізико-математичними моделями фізичної акустики. Використання таких модельних апроксимацій відкрило широкі можливості глибокого за-
своєння функціональних залежностей подання мови людини та її перетворення в другу форму сигналів – електричних, що називають електроакустичним перетворенням. Наведено системи акустикоелектричних та

електроакустичних перетворень і аналіз відхилень від адекватних значень цих перетворювань.

Якість ТУ, до складу якого входять розглядувані перетворювачі, оцінюють суб'єктивними й об'єктивними методами. На підставі детального вивчення системи телефонного зв'язку аналоговими (неперервними) електричними сигналами в заключному (шостому) розділі цієї частини розглянуто методи перетворення електричних сигналів аналогового типу, які зберігають форму мовних звукових сигналів, у сигнали дискретної форми. Вокодерний зв'язок утворюється за принципом частотного компандування мови для реалізації ідеї безпосереднього і параметричного компандування. Детально проаналізовано принцип дії загального частотного вокодера та його реалізацію у вигляді смушкового вокодера.

Перехід до цифровізації мовних звукових сигналів описано у вигляді дискретизації мовних сигналів шляхом використання амплітудно-імпульсної модуляції (АІМ) та імпульсно-кодової модуляції (ІКМ). Це і є шляхом переходу до ТС, в яких інформація передається у дискретній формі.

У другій частині підручника розглянуто ТС, обмін даними (повідомленнями), в яких здійснюється передаванням дискретної інформації як широкосмуговими цифровими каналами, так і каналами телефонного зв'язку (каналами тональної частоти – ТЧ каналами).

Основу ТС дискретної дії становить система передавання дискретної інформації (СПДІ). Детальне вивчення процесів поетапного перетворення сигналів у СПДІ дозволяє розглядати кожний складовий процес як окремий процес у вигляді задач і методів їх розв'язання, при цьому послідовність кожного визначається математичним описом загального алгоритму роботи СПДІ.

У восьмому розділі наведено методи перетворення дискретних сигналів у форму, зручну для передавання каналом зв'язку, тобто методи модуляції, а також процедури визначення спектрів перетворених сигналів. Під дією зовнішніх завад та характеристик сигналів перетворені сигнали під час поширення їх електричними ланцюгами з розподіленими параметрами (каналами зв'язку) спотворюються.

Перший крок боротьби із спотворенням сигналів – упровадження засобів ресстрації пристроями СПДІ (дев'ятий розділ). Найважливіші питання забезпечення узгодженої (синхронної) роботи електрокомунікаційних пристроїв та їх складових частин розглянуто в десятому розділі: методи підтримки синхронізації та фазування. У цьому розділі наведено приклад побудови схеми для виділення синхросигналу і даних із сигналу, який перетворюється каналом зв'язку. Здебільшого схеми виконано на основі генераторів із фазовим автопідстроюванням частоти. Варіанти схем генераторів різні, але розв'язувана за їх допомогою задача одна і та

ж – правильна прив'язка синхросигналу до бітових інтервалів приймальних даних.

Друга частина закінчується одинадцятим розділом. З'ясовуються закономірності виникнення помилок через діяння завад, яким неможливо протидіяти розглянутими в попередніх розділах методами і засобами. З методичного погляду наведено основні процедури дії статистичної обробки результатів вимірювання випадковості (розділ математичної статистики) як інструментальний засіб установлення закономірності виникнення помилок у каналі передавання даних і побудови математичної моделі дискретного каналу.

Третю частину присвячено питанням керування обміном повідомленнями між взаємодіючими станціями (абонентськими пунктами). Ця частина є продовженням другої, у якій на підставі знань про установлену закономірність виникнення помилок – розподіл вектора \vec{e} – розкриваються методи підвищення достовірності передавання дискретної інформації каналами (трактами) мережі. В одинадцятому та дванадцятому розділах розглянуто теорію передавання дискретної інформації. Коло розглядуваних у них питань дуже широке, оскільки мають бути охоплені всі аспекти призначення СПДІ – забезпечити максимальну швидкість передавання за високої якості (правильності передавання) і економічності системи. На жаль, завдання оптимізації у такій постановці ще не вирішено. Тому зазвичай прагнуть оптимізувати систему частинами.

Керування обміном даними між взаємодіючими АП протокольними процедурами розглянуто на механізмах роботи синхронних байтзорієнтованих і бітзорієнтованих протоколів. Детально описано можливості контрольних функцій протоколу двійкового синхронного зв'язку HDLC, а також порівняльний аналіз його з протоколами розвинутих технологій Frame Relay (FR) та асинхронним доставленням повідомлень АТМ.

В окремому розділі (чотирнадцятому) вивчаються принципи утворення ТС факсимільного передавання зображень ТМ, принципи сканування і передавання зображень. Під час розгляду системи факсимільного зв'язку наводиться структура факсимільних апаратів та СПДІ, яка забезпечує передавання сигнальних конструкцій (послідовності). На основі знань формування факсимільних сигналів приділено увагу способам скорочення надмірності, яка виникає за відповідних способів сканування оригіналу.

Створена ТС на середовищі мереж знаходиться в складних умовах, які в загальному вигляді зводяться до ненадійності складових трактів обміну інформацією та забезпечення вимоги високої надійності. У п'ятнадцятому розділі із системних позицій розглянуто питання надійності систем передавання дискретної інформації та характеристики надійності трактів передавання даних. На підставі цієї інформації за класичним

принципом виконують розрахунки надійності тракту передавання інформації. Виходячи з визначення надійності як властивості СПДІ забезпечують нормальне виконання усіх заданих функцій протягом регламентованого проміжку часу зі збереженням у заданих границях усіх якісних показників. Для цього з урахуванням впливу ненадійності трактів та витрат на відновлення введено підрозділи визначення характеристики доставлення повідомлень в умовах ненадійності та відновлення.

В останньому, шістнадцятому розділі проведено системний аналіз обміну даними організованою ТС, яку розглянуто не тільки як взаємодію КУД користувачів утворювальних ТС, а і як взаємодію сусідніх маршрутизаторів мережі. Це дає можливість з узагальнювальних позицій досліджувати основну складову частину ТМ, яка може функціонувати в дуплексному та піддуплексному режимах. Причому піддуплексний режим роботи ТС застосовують у системах взаємодії абонентських КУД, а дуплексний – у системах передавання даних магістральними каналами мереж. Математичні моделі, які описують процеси організації зв'язку і обміну інформацією між абонентськими станціями та маршрутизаторами магістральних каналів мереж, зводяться до моделей спеціальних випадкових процесів масового обслуговування – циклічного періодичного обслуговування черг заявок. Повний аналіз процесів функціонування ТС включає урахування впливу ненадійності каналів передавання на системні характеристики, для чого в попередньому розділі отримано збільшені функції розподілу часу обслуговування повідомлення – часу зайняття каналу передавання повідомлення, або їх ще називають *процедурами збільшення модельованих параметрів*.

Підручник написано на основі лекцій, що читаються протягом останніх десяти років на факультеті авіаційних і космічних систем та в Інституті телекомунікаційних систем Національного технічного університету України (НТУУ «КПІ»), а також на кафедрі телекомунікаційних систем Національного авіаційного університету для студентів, які навчаються за напрямом «Телекомунікації».

Частина 1

Телекомунікаційні системи телефонного та вокодерного зв'язку. Термінальне устаткування неперервної дії

1. Основні положення фізичної акустики

Фізична акустика займається вивченням звукових коливань, а за три останніх десятиріччя охопила і коливання, які лежать за межами чутності (ультраакустика). Вона широко використовує різноманітні методи перетворення механічних коливань в електричні і навпаки; такі методи називають методами електроакустики.

Щодо звукових коливань, то фізична акустика вивчає і властивості фізичних явищ, які зумовлюють ті чи ті якості звука, що розрізняються слухом.

1.1. Звук і його основні властивості

Звуком називають коливання частинок пружного середовища, частоти яких лежать у просторі сприймання слухом людини.

Цей простір обмежений границями від 20 до 20 000 Гц.

Колівання, частоти яких менше 20 Гц, називають *інфразвуками*, а коливання з частотами понад 20 000 Гц – *ультразвуками*.

Утворення звуку пов'язано з наявністю деякого коливального тіла або системи, які порушують статичний стан середовища і являють собою випромінювач звуку.

Під час роботи випромінювача звуку виникають звукові хвилі, зумовлені кінцевою швидкістю поширення звуку.

Звуковими хвилями називають переміжні стиснення та розрідження частинок середовища, які поширюються у просторі від випромінювача.

Простір, в якому поширюються звукові хвилі, називають звуковим полем.

Сукупність точок звукової хвилі, які знаходяться в однаковій фазі коливань, називають *фронтом звукової хвилі*.

Фронт хвилі може мати вигляд сфери, площини або поверхні неправильної форми, що залежить як від розмірів і форми випромінювача, так і від розмірів, форми і відображувальних або поглинальних властивостей поверхні, що обмежують звукове поле.

Найбільше значення для використання в техніці телекомунікацій мають сферичні звукові хвилі.

Найпростіший приклад утворення сферичних звукових хвиль показано на рис. 1.1.

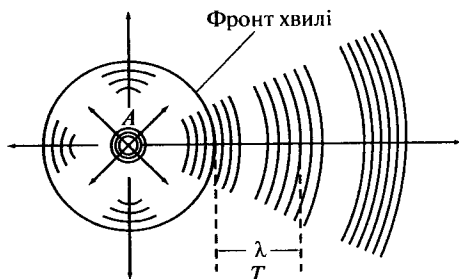


Рис. 1.1. Сферична звукова хвиля

Тут випромінювачем є пульсуюча куля A . Якщо її діаметр буде змінюватись за гармонічним законом (наприклад, $D = D_m \sin(\omega t)$), тоді з кожним його збільшенням виникатимуть згущення частинок ближчої до нього повітряної кулі. Тиск їх в цій кулі зросте і буде діяти на частинки наступної, більш віддаленої кулі і т. д. За однакових умов поширення процесу стиснення буде симетрично поширюватися на всі боки від випромінювача. Зменшення об'єму спричинить розрідження частинок повітря в прилеглому до кулі шарі, яке також буде поширюватися у всі боки від випромінювача.

Отже, у звуковому полі виникнуть сферичні звукові хвилі, які в цьому разі відповідно до закону змінювання діаметра кулі також будуть мати гармонічний характер. При цьому частинки повітря періодично зміщуватимуться на деяку величину x від свого спокійного положення, яке змінюється за законом $x = x_m \sin(\omega t)$. Напрямок цього зміщення буде нормальним до поверхні випромінювача.

Звукові хвилі, які утворюються під час розмови людини, з достатньою для практики точністю можна вважати сферичними на відстані до 10 см від рота людини.

Як і кожний хвильовий процес, звукові хвилі характеризуються такими основними параметрами, як довжина хвилі λ , період T , частота f і швидкість поширення c .

Довжину звукової хвилі можна визначати як відстань λ між двома найбільш близькими стисненнями або розрідженнями. Час, протягом якого звукова хвиля поширюється на відстань λ , дорівнює її періоду T , а кількість звукових хвиль, які проходять через деяку точку звукового поля за секунду, – частоті f звукового коливання.

Швидкість звуку c – це відстань, на яку звукова хвиля поширюється за одну секунду.

Ці параметри зв'язані один з одним очевидним співвідношенням

$$c = \frac{\lambda}{T} = \lambda f.$$

Швидкість звуку залежить від фізичних властивостей середовища, у якому поширюється звук, і не залежить від його частоти.

У рідких і твердих тілах швидкість звуку значно вища, ніж у повітрі.

Звукові хвилі мають здатність відбиватися від твердих поверхонь і вбиратися м'якими та пористими матеріалами.

Звукове поле, у якому поширюються звукові хвилі без завад, називають *вільним звуковим полем*.

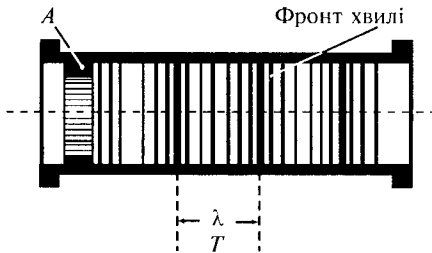


Рис. 1.2. Плоска звукова хвиля

Плоску звукову хвилю можна отримати (рис. 1.2) під час коливально-го руху поршня у трубі з абсолютно жорсткими та гладенькими стінками. Плоскі хвилі виникають також близько від великих коливальних поверхонь.

1.2. Енергетична характеристика звуку

Енергетичну характеристику звуку визначають за допомогою двох взаємозв'язаних параметрів звукового поля – звукового тиску та інтенсивності звуку.

Звуковим тиском називається надмірний тиск порівняно з атмосферним тиском у певній точці звукового поля, яке створює джерело звуку.

Виникнення такого надмірного тиску можна пояснити за допомогою графіка (рис. 1.3).

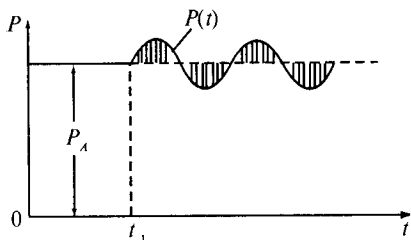


Рис. 1.3. Звуковий тиск

На цьому графіку показано часову залежність тиску P у будь-якій точці простору. У проміжку часу від $t = 0$ до $t = t_1$ джерело звуку не діє, тому тиск, який розглядається в цій точці, дорівнює атмосферному тиску $P = P_A$. У момент t_1 вмикається джерело звуку. Тоді на постійно діючий атмосферний тиск P_A буде накладатись додатковий тиск $p(t)$ і результуючий тиск у цій точці дорівнюватиме $P = P_A + p(t)$. Змінна складова цього тиску, що дорівнює $p(t)$, і є звуковим тиском.

Закон, за яким змінюється звуковий тиск за часом, може бути будь-яким. Він визначається характером коливань випромінювача. В окремому випадку гармонічних коливань (див. рис. 1.3), його можна описати виразом

$$p(t) = p_m \sin(\omega t),$$

де $p(t)$ – миттєві значення звукового тиску; p_m – його амплітуда. В акустичних розрахунках користуються зазвичай ефективним значенням звукового тиску p , за своїм змістом еквівалентним поняттю про діюче значення змінної електричної напруги.

Для розглядуваного випадку гармонічного звуку

$$p = \frac{p_m}{\sqrt{2}}.$$

Звуковий тиск у Міжнародній системі одиниць SI вимірюють у ньютонах на квадратний метр (Н/м^2), а в системі CGS у барах (бар); $1 \text{ бар} = 1 \text{ дН/см}^2$. Щоб перевести з однієї системи в другу, користуються співвідношенням $1 \text{ Н/м}^2 = 10 \text{ бар}$.

Інтенсивністю звуку (I) називають кількість звукової енергії, яка проходить за одну секунду через одиницю площини, перпендикулярної до напряму поширення звукової хвилі.

Отже, за своїм змістом інтенсивність звуку являє собою питому потужність (потужність, яка припадає на одиницю площі), яка переноситься звуковою хвилею і вимірюється у ватах на квадратний метр ($\text{Вт}/\text{м}^2$); в системі CGS одиниця виміру – $\text{ерг}/(\text{с}\cdot\text{см}^2)$.

Між інтенсивністю звуку і звуковим тиском існує певна залежність. У загальному випадку ця залежність досить складна; вона залежить від характеру коливального процесу та форми фронту звукової хвилі. Для сферичних звукових хвиль у разі гармонічних коливань така залежність описується простою формулою, яку можна одержати, використавши акустичний закон Ома, згідно з яким відношення звукового тиску p до коливальної швидкості зміщення частинок середовища $\vartheta = \frac{\partial x}{\partial t}$ – постійна величина, яка дорівнює добутку густини середовища ρ на швидкість поширення звуку в середовищі c :

$$\frac{P}{v} = \rho c. \quad (1.1)$$

Добуток ρc за аналогією із законом Ома для електричного кола називають питомим акустичним опором середовища $Z_a = \rho c$, який характеризує рівень протидії середовища поширенню в ній звукової хвилі.

Проте, виходячи з означення інтенсивності звуку як його питомої потужності, можна записати

$$I = p v. \quad (1.2)$$

Із рівнянь (1.1) та (1.2) одержимо шукане співвідношення:

$$I = \frac{p^2}{\rho c}. \quad (1.3)$$

Підставивши у формулу (1.3) відомі константи для повітря при температурі 20°C $\rho = 1,205 \text{ кг}/\text{м}^3$ і $c = 344 \text{ м}/\text{с}$, одержимо дві практичні формули:

$$I = 2,41 \cdot 10^{-3} \rho^2;$$
$$p = 20,4 \sqrt{I}.$$

Зазначимо, що в сферичній хвилі звуковий тиск з віддаленням від випромінювача зменшується обернено пропорційно відстані, а інтенсивність звуку — обернено пропорційно квадрату відстані:

$$\frac{P_1}{P_2} = \frac{r_2}{r_1}; \quad \frac{I_1}{I_2} = \frac{r_2^2}{r_1^2},$$

де r_1 і r_2 – відстані розглядуваних точок звукового поля до випромінювача.

1.3. Часові та спектральні характеристики звуків. Класифікація звуків

Досі ми розглядали властивості звуку на прикладах гармонічних звуків, оскільки всі закономірності для звуків цього класу найпростіші й очевидні. На практиці, зокрема у техніці зв'язку, частіше трапляються звуки більш складної структури.

Під *часовою характеристикою звуку* розуміють залежність звукового тиску від часу, тоді функція $p = f(t)$.

Під *частотним спектром звуку* розуміють закон розподілу звукового тиску або звукової енергії на частотному діапазоні.

Розподіл амплітуд складових частин звукового тиску за частотами являє собою функцію вигляду $p_m = \varphi(f)$; його ще називають *амплітудним спектром*.

У загальному випадку поняття спектра включає спектр амплітуд і спектр фаз окремих гармонічних складових, але останній не відіграє значної ролі у передаванні мови по телефонному зв'язку, оскільки тривалість мовних звуків досить велика і вухо не реагує на фазові співвідношення між складовими гармоніками. Тому спектр фаз зазвичай не розглядають в акустиці та телефонній техніці.

Розподіл енергії звуку являє собою розподіл інтенсивності звуку за частотами $I = \varphi(f)$. Цей вид спектральної характеристики дуже зручний, оскільки він не мав інформації про фазові співвідношення складових звуку.

Часові характеристики і частотні спектри звуків взаємопов'язані: певному виду часової характеристики однозначно відповідає спектральний розподіл.

Усі звуки, які є в природі, можна поділити на три основні групи: прості або гармонічні, складні періодичні, складні неперіодичні звуки.

Приклади часових характеристик звуків цих трьох груп і відповідні їм амплітудно-частотні спектри показано на рис. 1.4.

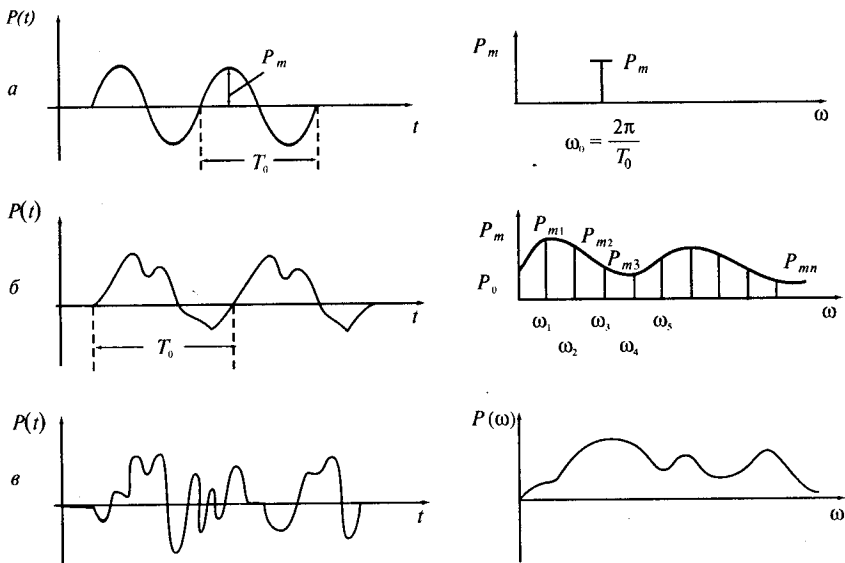


Рис. 1.4. Часові характеристики й амплітудно-частотні спектри звуків:
 а – простий гармонічний звук; б – складний періодичний звук;
 в – складний неперіодичний звук

Простий або гармонічний звук – це такий звук, у якого звуковий тиск змінюється за синусоїдним або косинусоїдним законом:

$$p(t) = p_m \sin \omega t.$$

Періоду T_0 такого коливання відповідає колова частота $\omega_0 = \frac{2\pi}{T_0}$ і амплітуда p_m . Амплітудний спектр гармонічного звуку показано однією вертикальною лінією, яка зображує амплітуду p_m у точці, що відповідає $\omega = \omega_0$ (рис. 1.4, а).

Прикладом гармонічного звуку є звуки, які створюються камертоном або коливанням струни музичального інструмента.

Складний періодичний звук – це такий звук, звуковий тиск якого змінюється за будь-яким періодичним законом, крім синусоїдного та косинусоїдного. Можливий вигляд часової характеристики та амплітудно-частотного спектра такого роду звуку показано на рис. 1.4, б.

Поклавши, що функція $p(t)$ задовольняє відомі умови Діріхле (ці умови зазвичай задовольняють часові характеристики періодичних звуків, які

трапляються в техніці зв'язку), можна подати її у вигляді ряду Фур'є, з якого й отримуємо амплітуди гармонік для побудови спектра:

$$p(t) = \frac{P_0}{2} + \sum_{k=1}^{\infty} p_{mk} \cos(k\omega_0 t - \varphi_k),$$

де $p_{mk} = \frac{1}{\pi} \int_{-\infty}^{\infty} p(t) \cos k\omega_0 t dt.$

Умови Діріхле. Функція $p(t)$ повинна мати в інтервалі $(-\pi, \pi)$ лише кінцеву кількість максимумів і мінімумів і бути неперервною, за винятком кінцевої кількості точок розриву першого роду.

Слід зазначити, що функція $p(t)$ є парною функцією, тому в розв'язанні в ряд Фур'є залишаються косинусоїдні складові.

Частотний спектр складного періодичного звуку зображується у вигляді дискретного ряду амплітуд гармонічних складових, частоти яких кратні $\omega_0 = \frac{2\pi}{T_0}$. Прикладом складних періодичних звуків є окремі звуки

людської мови, які допускають протяжне вимовляння, у процесі якого вони будуть багатократно повторюватись.

Складні неперіодичні звуки – це звуки, в яких звуковий тиск змінюється у часі неперіодично. Такого типу часову характеристику і відповідний їй спектр показано на рис. 1.4, в.

Зв'язок між неперіодичною функцією $P(t)$, яка задовольняє вимоги Діріхле, та її поданням у вигляді нескінченної суми найпростіших гармонічних складових, установлюється інтегралом Фур'є:

$$P(t) = \int_0^{\infty} P(\omega) \cos[\omega t - \varphi(\omega)] d\omega,$$

де $P(\omega)$ – спектральна щільність амплітуди; $P(\omega)d\omega$ – нескінченно мала амплітуда складової.

Амплітудно-частотний спектр неперіодичних звуків – неперервний, складові якого є на всіх частотах діапазону.

Звуками цього класу можна вважати мовний процес в цілому, який розглядається як послідовність звуків мови, що змінюють один одного, та різного роду шумів.

Для з'ясування спектральних характеристик звуків, розглянутих вище трьох видів, були використані амплітудні спектри. Енергетичні спектри (спектри інтенсивностей) цих звуків для випадків сферичних і звукових хвиль можна отримати, використовуючи залежність інтенсивності звуку від його тиску, яку наведено у формулі (1.1).

1.4. Рівні інтенсивності звуку та звукового тиску

1.4.1. Загальні положення

Для виконання розрахунків і кількісного аналізу акустичних явищ частіше зручніше замість абсолютних величин інтенсивності та тиску користуватись їх рівнями, які виражаються у відносних логарифмічних одиницях.

Рівнем інтенсивності звуку B називають 10 десяткових логарифмів відношення інтенсивності певного звуку I до інтенсивності звуку, який умовно взято за звук нульового рівня I_0 :

$$B = 10 \lg \frac{I}{I_0}. \quad (1.4)$$

Рівень інтенсивності звуку вимірюють у децибелах (дБ). Один децибел – це такий рівень, за якого інтенсивність звуку перевищує інтенсивність нульового рівня в 1,26 разу.

За нульовий рівень взято величину $I_0 = 10^{-12}$ (Вт/м²), що приблизно відповідає мінімальній інтенсивності, яка сприймається слухом людини.

Рівнем звукового тиску B називають 20 десяткових логарифмів відношення звукового тиску певного звуку P до звукового тиску, який умовно взято за звук нульового рівня $T_0 = 2 \cdot 10^{-5}$ (Нп/м²).

Формулу для обчислення рівня звукового тиску можна отримати, якщо у рівнянні (1.4) замінити I і I_0 їх виразами через звукові тиски згідно з рівнянням (1.3):

$$B = 10 \lg \frac{I}{I_0} = 10 \lg \frac{P^2}{P_0^2} = 20 \lg \frac{P}{P_0}.$$

Рівень звукового тиску 1 дБ відповідно є перевищенням заданого звукового тиску над звуковим тиском нульового рівня в $\sqrt{1,26} = 1,12$ разу.

Кількісно рівень звукового тиску дорівнює рівню інтенсивності. Тому величину B називають інколи просто рівнем звуку, який можна визначити або через інтенсивність, або через звуковий тиск.

Зручність користування рівнями замість абсолютних значень відповідних параметрів звуку визначають, керуючись таким міркуванням. Границі змінювання звукових тисків та інтенсивності у разі вираження їх в абсолютних одиницях виявляються надто великими, що приводить до непорівнянних за значенням чисел і незручності у разі їх графічного зображення.

Для прикладу розглянемо зміни звукового тиску в телефонній акустиці. Мінімальний звуковий тиск, який сприймається вухом людини, $P_{\min} = 2 \cdot 10^{-5}$ Нп/м², максимальне його значення, яке витримує вухо людини, $P_{\max} = 20$ Нп/м².

Отже, максимальні та мінімальні значення відрізняються один від одного значенням $\frac{P_{\max}}{P_{\min}} = \frac{20}{20 \cdot 10^{-5}} = 10^6$ у мільйон разів. Водночас логарифмічний діапазон змін цих тисків, визначуваних їх рівнями, становить лише $\Delta B = 20 \lg \frac{P_{\max}}{P_{\min}} = 20 \lg \frac{20}{2 \cdot 10^{-5}} = 120$ дБ.

Інколи виникає потреба в узгодженні результатів акустичних розрахунків з даними електричних розрахунків, коли для визначення рівнів (у неперах) можна застосовувати систему натуральних логарифмів.

Рівні інтенсивності та звукового тиску також можна виражати в неперах (Нп) та обчислювати за формулою

$$B = \frac{1}{2} \ln \frac{I}{I_0} = \ln \frac{P}{P_0}.$$

Для переведення одних одиниць в інші користуються співвідношеннями між ними: 1 дБ = 0,115 Нп; 1 Нп = 8,69 дБ.

1.4.2. Рівні спектрів складних звуків

Для складного періодичного звуку загальна інтенсивність дорівнює простій арифметичній сумі інтенсивності його складових. Тому рівень інтенсивності такого звуку визначають так:

$$B_{\Sigma} = 10 \lg \frac{\sum_{k=1}^{\infty} I_k}{I_0}.$$

Для складного неперіодичного звуку здебільшого користуються одним з основних видів їх спектральних параметрів, які виражають у вигляді відповідних рівнів: рівня у смузі $B_{\Delta f}$, спектрального рівня B_1 та сумарного рівня B_{Σ} .

Зміст цих понять легко зрозуміти, розглядаючи спектр інтенсивності складного неперіодичного звуку, приблизний вигляд якого показано на рис. 1.5. Складові неперіодичного звуку розміщено в межах f_a і f_b ; $I(f)$ – спектральна щільність інтенсивності.

Рівнем звуку у певній смузі шириною Δf називають рівень $B_{\Delta f}$ звукової енергії (інтенсивності), взятої суми в межах цієї смуги. Так, якщо f_1 і f_2 – межі цієї смуги, то сумарна інтенсивність у цієї смуги $I_{\Delta f} = \int_{f_1}^{f_2} I(f) df$, а її рівень

$$B_{\Delta f} = 10 \lg \frac{I_{\Delta f}}{I_0}.$$

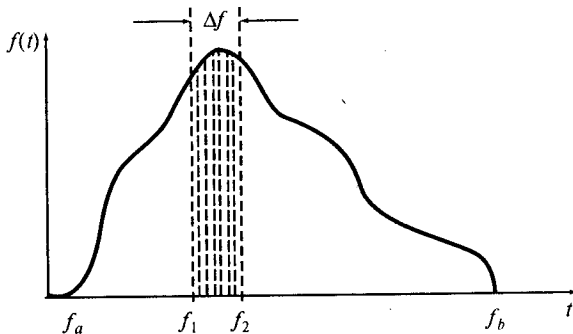


Рис. 1.5. Спектр інтенсивності складного неперіодичного звуку

Якщо взяти смугу частот шириною 1 Гц, то рівень енергії (інтенсивності), який припадає на цю смугу, називають рівнем B_1 .

Очевидно, що спектральну інтенсивність, яка припадає на 1 Гц, можна визначити як $I_1 = \frac{I_{\Delta f}}{\Delta f} \approx I(f)$, тоді $B_1 = 10 \lg \frac{I_{\Delta f}}{\Delta f I_0} = 10 \lg \frac{I_{\Delta f}}{I_0} - 10 \lg \Delta f$, звідки

$$B_1 = B_{\Delta f} - 10 \lg \Delta f.$$

Сумарна інтенсивність у цьому разі

$$I_{\Sigma} = \int_{f_a}^{f_b} I(f) df,$$

тоді

$$B_{\Sigma} = 10 \lg \frac{\int_{f_a}^{f_b} I(f) df}{I_0}.$$

1.4.3. Загальний рівень звуку, який виникає від дії декількох джерел

Під час розв'язання практичних задач виникає потреба у визначенні результуючого звуку, який створюється у разі одночасної дії декількох джерел (наприклад, за одночасної дії декількох звуків у кімнатах операторів телефонного зв'язку, що перешкоджає передачі корисного мовного звуку в канал зв'язку); ці звуки називають ще *шумовими завадами*.

Знайдемо загальний рівень звуку, створений дією декількох джерел. Нехай задано спектральні рівні звуку від кожного джерела – B_1, B_2, \dots, B_n .

Потрібно визначити результуючий спектральний рівень B . У символічній формі підсумовування цих рівнів записують у такому вигляді:

$$B = B_1(+)B_2(+)\dots(+)B_n.$$

Підсумовування в такому вигляді символічне, оскільки величина B не є сумою величин B_i , тому що ці величини відносні. Для виконання дії складання рівнів звуків слід перейти до початкових параметрів, які підлягають вимірюванню та складанню як арифметичні величини.

Цю операцію називають логарифмічним складанням, а знак $(+)$ – символом логарифмічного складання.

Визначивши спектральний рівень, знайдемо спектральні інтенсивності складових:

$$B_1 = 10 \lg \frac{I_1}{I_0}, \text{ звідси } I_1 = I_0 \cdot 10^{0,1B_1};$$

$$B_2 = 10 \lg \frac{I_2}{I_0}, \text{ звідси } I_2 = I_0 \cdot 10^{0,1B_2};$$

$$\dots\dots\dots$$

$$B_n = 10 \lg \frac{I_n}{I_0}, \text{ звідси } I_n = I_0 \cdot 10^{0,1B_n}.$$

Визначимо результуючу інтенсивність I як

$$I = I_1 + I_2 + \dots + I_n = I_0 10^{0,1B_1} + I_0 10^{0,1B_2} + \dots + I_0 10^{0,1B_n} =$$

$$= I_0 (10^{0,1B_1} + 10^{0,1B_2} + \dots + 10^{0,1B_n}).$$

Тоді

$$B = 10 \lg \frac{I}{I_0} = 10 \lg \frac{1}{I_0} I_0 (10^{0,1B_1} + 10^{0,1B_2} + \dots + 10^{0,1B_n}) =$$

$$= 10 \lg (10^{0,1B_1} + 10^{0,1B_2} + \dots + 10^{0,1B_n}). \quad (1.5)$$

Для практичного виконання логарифмічного складання використовують графік-номограму, побудований за формулою (1.5). Його показано на рис. 1.5.

За допомогою цього графіка знаходять логарифмічну суму двох рівнів $B_1(+)B_2$. Якщо потрібно скласти декілька рівнів, тоді підсумовування виконують послідовно: перший складають з другим, потім результуючий – з третім і т. д. Порядок користування графіком показано на прикладі (рис. 1.6).

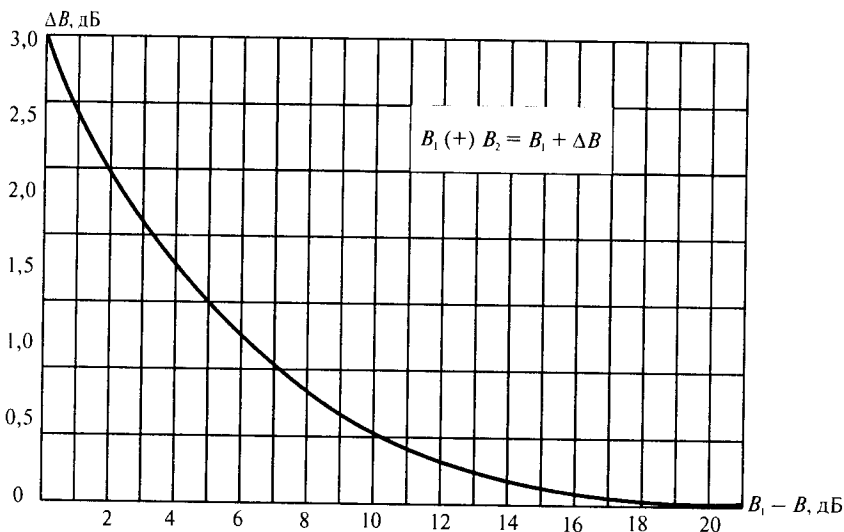


Рис.1.6. Графік для логарифмічного складання

Нехай потрібно знайти результуючий спектральний рівень B , який створюється двома складовими з рівнями $B_1 = 80$ дБ і $B_2 = 74$ дБ.

Для цього від більшого рівня віднімемо менший та визначимо їх різницю: $B_1 - B_2 = 80 - 74 = 6$ дБ. Цій різниці на осі абсцис відповідає на графіку ордината $\Delta B = 1$ дБ, яку треба додати до більшого рівня. Тоді пошуковий результуючий рівень $B = B_1 + \Delta B = 80 + 1 = 81$ дБ.

Зазначимо, що максимально можливим значенням добавки ΔB є 3 дБ, якщо складові рівні однакові, а якщо різниця становить більше ніж 10 дБ, тоді меншим можна знехтувати ($\Delta B < 0,4$ дБ) і вважати, що результуючий рівень дорівнює найбільшому серед тих, які складаються.

1.4.4. Шуми та їх характеристики

Сигнали, які несуть корисну інформацію в системах зв'язку засобами телекомунікацій, завжди передаються і приймаються в умовах впливу на ці системи різноманітних завад. На телефонних трактах, призначених для передавання мови, основним видом завад є шуми, котрі прослуховуються на приймальному кінці і заважають тим самим правильному прийманню. Механізм їх заважальних дій та заходи захисту від них пояснюються протягом вивчення всього курсу. Тут обмежимося розглядом власних характеристик основних типів шумів, хоча якісно вже і тепер зрозуміло, що чим ширший спектр, який займають частотні складові шуму, та чим ви-

щий рівень мають ці складові, тим більший заважальний вплив буде чинити цей шум за інших однакових умов на якість приймальної мови.

Залежно від виникнення (джерела шуму) розрізняють шуми за видами, які виникають у мовних трактах:

- мовні завади або шуми від сторонніх розмов;
- атмосферні завади;
- індустриальні завади або шуми від різних видів електромагнітних процесів, які виникають під час роботи промислових та транспортних агрегатів;
- шуми від роботи двигунів й інших механізмів (наприклад, літакові і танкові шуми);
- власні шуми апаратури зв'язку, які створюють мікрофони та підсилювальні пристрої.

Залежно від шляху проникнення до вуха того, хто слухає всі види шумів можна поділити на два класи:

- акустичні, які надходять безпосередньо зі звукового поля;
- електричні, які проходять через електричний тракт та приймальний перетворювач.

У кінцевому підсумку всі чутні шуми перетворюються в акустичні.

Типові спектральні характеристики шумів та розподіл імпульсної послідовності шумів показано на рис. 1.7–1.9.

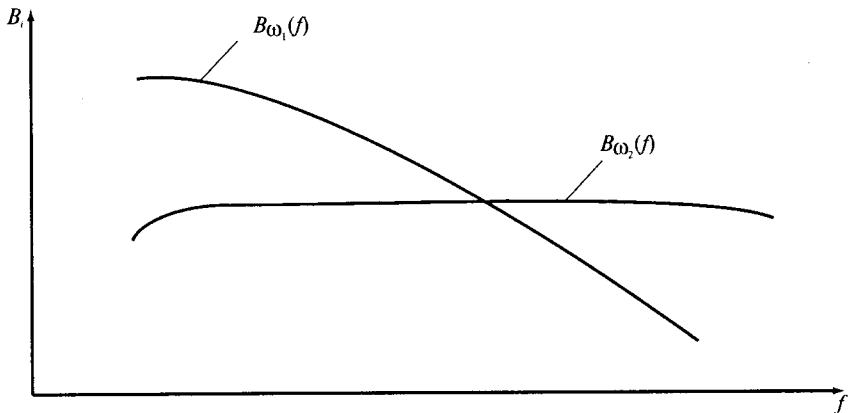


Рис. 1.7. Основні види спектральних характеристик шумів

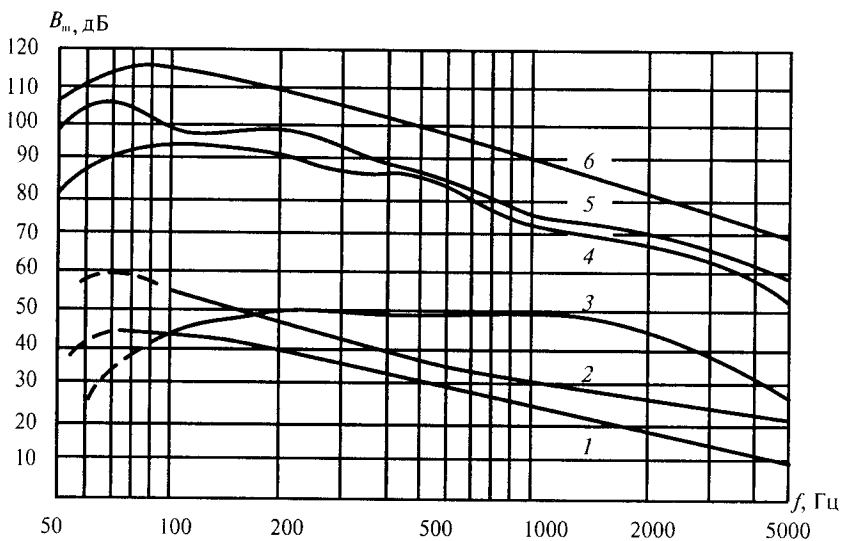


Рис. 1.8. Частотні спектри деяких типів шумів:
 1 – кімнатний шум $B_{m\Sigma} = 60$ дБ; 2 – шум вулиці $B_{m\Sigma} = 70$ дБ;
 3 – шум від роботи телеграфних апаратів $B_{m\Sigma} = 80$ дБ;
 4 – корабельний шум $B_{m\Sigma} = 115$ дБ; 5 – танковий шум $B_{m\Sigma} = 125$ дБ;
 6 – літаковий шум $B_{m\Sigma} = 1300$ дБ

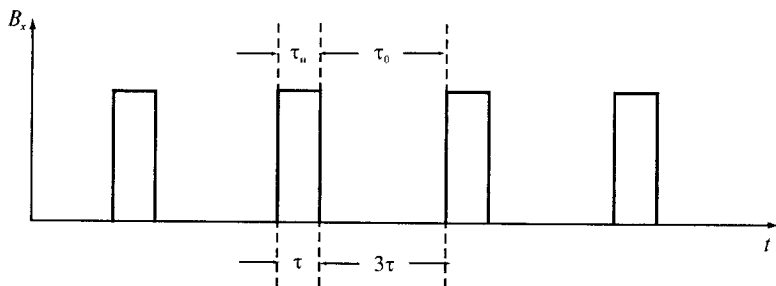


Рис. 1.9. Приклад визначення щільності імпульсної завади

За одночасної дії корисного мовного сигналу із спектральними характеристиками, наведеними у підрозд. 1.3, та шумів з їх характеристиками утворюється загальний сигнал. Із цього сигналу приймач має виділити корисний сигнал. Функцію виділення мовного сигналу визначають за методами боротьби із завадами.

2. Мовний сигнал і його сприйняття

2.1. Механізм утворення мовних звуків та їх основні властивості

Звуки мови утворюються під час роботи голосового апарата людини, схематичне зображення якого показано на рис. 2.1.

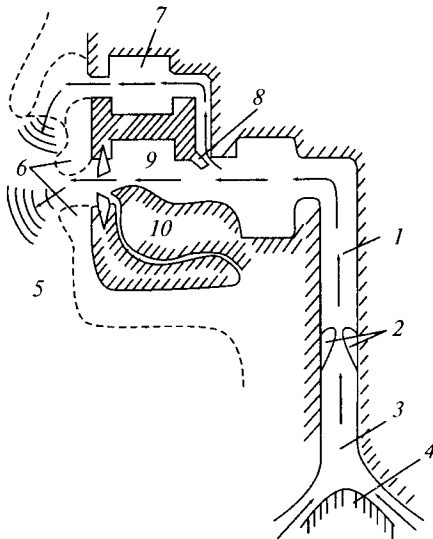


Рис. 2.1. Схематичне зображення голосового апарата людини

У мовотворенні беруть участь легені, які спілкуються з гортанню 1 через бронхи 4 і трахею 3, голосові зв'язки 2, порожнину рота 9 та пов'язані з нею язик 10, зуби 5 і губи 6, а також носову порожнину 7, яка за допомогою піднебінної завіски 7 може включатися в мовний тракт.

Залежно від способу утворення і залежних від нього акустичних властивостей звуки мови поділяють на голосні та приголосні. Голосні та деякі приголосні, які ще називають «вокалізованими» (м, н, р, л, й), складають групу *сонорних* звуків. Ці звуки створюються завдяки квазіперіодичним коливанням голосових зв'язків.

Решту приголосних звуків поділяють на дзвінкі та глухі. В утворенні глухих приголосних голосові зв'язки не беруть участі. Джерелом утворення глухих приголосних є перепади тиску повітря внаслідок звуження тих чи тих ланок голосового тракту (шумове збудження). Дзвінкі приго-

лосні утворюються у результаті сумісної дії як шумового збудження, так і голосових зв'язок.

Під час дії на слуховий апарат звуки мови відрізняються один від одного спектральним розподілом і його змінами з плином часу. Частіше ці характеристики однозначно збігаються з натуральними звуками людини, що дозволяє розпізнавати звуки мови за достатнього для ряду практичних задач наближення зв'язати їх із середньостатистичними спектрами. Таке наближення покладено в основу формантної теорії сприйняття мови – основи більшості технічних додатків теорії розбірливості.

Раніше було згадано про високу інформативність мовного сигналу, який несе в собі дані про тип голосу мовця, зміст інформації, тембральні ознаки, що дозволяють розрізнити голоси мовців, а також інтонаційні особливості передаваних повідомлень.

Формування звуків мови та їх інформаційні характеристики розглянемо на прикладі вимовляння сонорних звуків (рис. 2.1, 2.2). У процесі вимовляння легені через бронхи та трахею створюють потік повітря, який поступає через вібруючі голосові зв'язки до порожнини рота та носової порожнини. Голосові зв'язки то відкриваються, то стуляються, тим самим створюють послідовність імпульсів звукового тиску, частота яких майже постійна. Часову характеристику цього процесу показано на рис. 2.2, *a*.

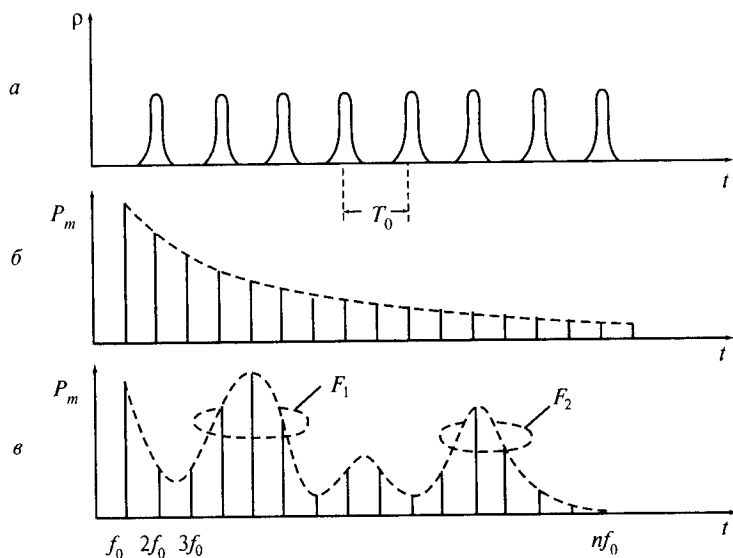


Рис. 2.2. Формування звуків мови

Одержаний при цьому частотний спектр містить велику кількість гармонічних складових, амплітуда яких зменшується зі зростанням частоти, як це показано на рис. 2.2, б. Далі повітряний потік зіштовхується на своєму шляху із системою резонаторів, які створюються об'ємом порожнини рота і носової порожнини і видозмінюються в процесі вимовляння різних звуків положенням язика, зубів і губ. Проходячи через цю систему резонаторів, деякі гармонічні складові підсилюються, інші навпаки заглушуються. Тоді картина спектра звуку, який формується ротом, набуває вигляду, який приблизно зображено на рис. 2.2, в.

У цьому спектрі видно основний тон (ОТ) f_0 (перша гармоніка, яка відповідає частоті коливань голосових зв'язок), значно підсилені діапазони частот F_1 , F_2 , F_3 і дещо менше підсилені інші частотні складові.

Частота ОТ визначає тип голосу мовця (бас, баритон, тенор, альт, контральто, сопрано). Діапазон основного тону – 70...350 Гц. Мова з відфільтрованим ОТ зберігає розбірливість, але губить енергію та частково природність.

Значно підсилені амплітуди тиску частотних діапазонів, які характерні для спектра певного звуку, що дозволяють відрізнити його від інших звуків, називають *формантами*.

Оскільки форманти значно потужніші від інших складових, то вони гзовним чином і визначають слухове сприймання того чи того звуку. Отже, звуки мови відрізняються один від одного кількістю формант та їх розміщенням у частотному спектрі.

Значно менше підсилення (підсилення неформантних складових) отримують діапазони частот, які зв'язані з індивідуальними особливостями голосового апарата мовця. Вони визначають тембр мови і разом з ОТ дозволяють впізнавати мовця по голосу.

Спектри глухих приголосних звуків, які виникають унаслідок шумового збудження, суцільні, а спектри дзвінких приголосних можуть мати в своєму складі як дискретні складові, так і ділянки суцільного спектра, оскільки в їх утворенні беруть участь як шумові збудження, так і голосові. Ті й ті так само, як і сонорні звуки, відрізняються підсиленими формантними ділянками (рис. 2.3).

Різні звуки мови можуть мати в своєму складі від однієї до трьох формант. Формантні складові знаходяться в діапазоні 200...10000 Гц. У цьому діапазоні вони розподіляються нерівномірно: в діапазоні низьких і високих частот їх менше, а в середині смуги – більше.

Тривалість окремих звуків мови в нормальному мовному потоці становить 20...350 мс. Голосні звуки мають більшу тривалість (найбільш тривалий звук «а»), приголосні – меншу тривалість (найкоротший – звук «п»).

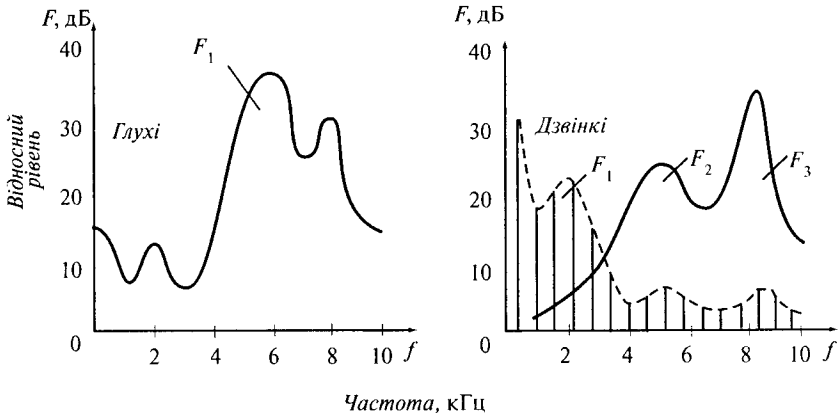


Рис. 2.3. Амплітудно-частотні спектри приголосних звуків

Окремі звуки мови в реальних словах вимовляються по-різному. Так, наприклад, голосний звук може бути наголошеним, або в сполученнях з різними звуками. При цьому змінюються тривалість, інтенсивність, а для деяких звуків і спектральний розподіл. Через це кількість реалізацій звуків у мові значно більша від кількості звукотипів, що відповідають номінальним найменуванням звуків. Ці звукотипи, які об'єднують по декілька можливих форм існування звуків у мові, називають *фонемами*.

2.2. Інтегральні характеристики мови

Мова людини, як відомо, складається з окремих мовних звуків. Для засобів телефонного зв'язку для передавання мовних звуків як неперервним електричними сигналами, так і дискретними (цифровими) потрібно вмійти визначати параметри не тільки окремих звуків, а і їх вплив на загальну характеристику мови в цілому. Мову формує багато людей, які мають різні особливості; крім того, спектр використання окремих звуків дуже широкий і формування мови залежить загалом від випадковості їх використання. Це потребує узагальнення характеристики мови. Людей у всьому світі дуже багато, і кожна людина має свої особисті якості вимовлення мовних сигналів. Користуючись поняттям з математичної статистики «генеральна сукупність», можна подати множину мовних сигналів як генеральну сукупність. Тоді, щоб знайти закономірність подання характеристик мови, достатньо взяти вибіркочну сукупність і обробити сигнали згідно з діями механізму статистичного аналізу й отримати узагальнену статистичну характеристику, яку називають *інтегральною характеристикою мови*.

Загальна інтегральна характеристика мови, або характеристики її складових потрібні передусім для виконання розрахунків можливості та якості передавання мови різними каналами зв'язку та створення спеціальних перетворювачів мовних сигналів типу «вокодер» чи аналого-цифрового перетворювача. На практиці використовують декілька інтегральних характеристик: спектр мови, формантний спектр, амплітудний склад мови, динамічний та частотний діапазони і формантний розподіл.

Спектром мови називають залежність середнього значення протягом тривалого часу спектрального рівня мови B_p від частоти f . Тут під середнім розуміють середнє арифметичне значення спектральних інтенсивностей, які спостерігаються тривалий час. Залежність середнього рівня мови від частоти показано на рис. 2.4.

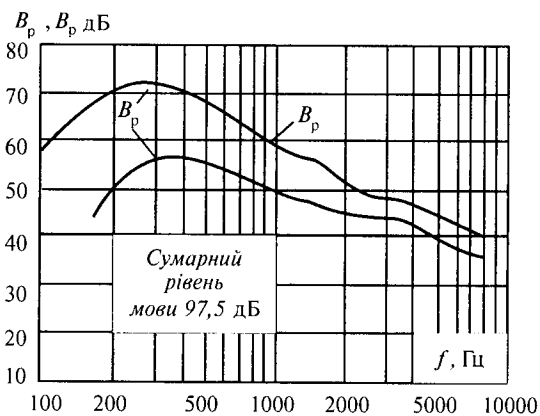


Рис. 2.4. Спектр мови (B_p) і формант (B_p')

Формантним спектром називається залежність середнього значення протягом тривалого часу спектрального рівня формант B_p' від частоти f .

Під середнім формантним рівнем розуміють такий рівень, який протягом тривалого часу перевищується в 50 % випадків, а в 50 % випадків формантні складові мають рівні, менші від наявного.

Формантний спектр мови не можна отримати безпосереднім вимірюванням інтенсивності формантів в обмежених частотних смугах, оскільки ніяк не можливо відокремити формантні складові від неформантних. Таку характеристику розраховують аналітично.

У деяких випадках потрібно знати не тільки закон розподілу середньої енергії складових мови від частоти, а й статистичний розподіл їх за рівнями. Таким розподілом є характеристика амплітудного складу мови (рис. 2.5).

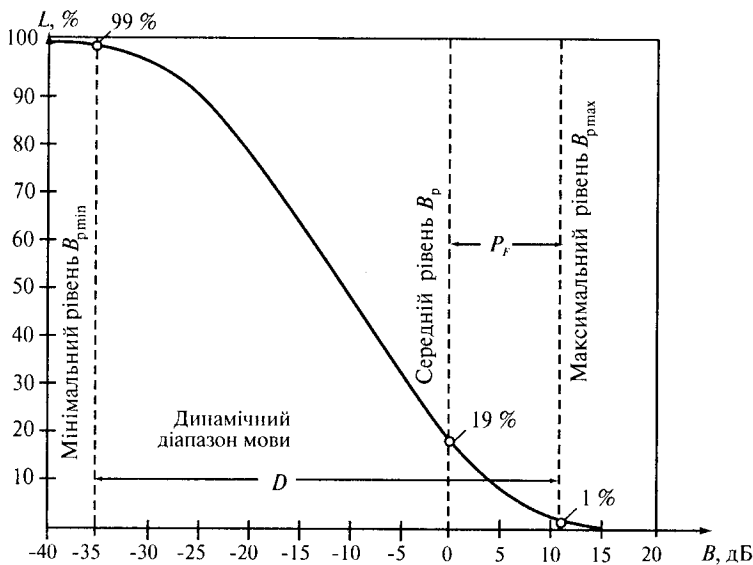


Рис. 2.5. Характеристика амплітудного складу мови

Амплітудний склад мови – це інтегральний розподіл відносної кількості випадків β , у яких той чи той рівень B перевищує величину цього рівня.

На рис. 2.5 усі рівні B відраховано від середнього рівня мови, B_p взятого за 0. Із цього видно, наприклад, що рівні, які дорівнюють середньому рівню, перевищуються в мовному процесі в 19% інтервалів за тривалий час спостереження; рівні, які перевищують середній на 12 дБ – усього лише в 1% інтервалів; рівні, менші від середнього на 25 дБ, – в 90% інтервалів.

Динамічний діапазон мови – це різниця між максимальними і мінімальними рівнями:

$$D = B_{p_{\max}} - B_{p_{\min}}$$

Пояснення динамічного діапазону ілюструє рис. 2.6, на якому показано залежності від частоти максимального, середнього та мінімального спектральних рівнів мови.

Оскільки основна інформація мовного складу міститься в формантах, то потрібно знати кількість формантних складових, які передаються на різних частотних смугах. Для цього використовують характеристику, яку називають *формантним розподілом*.

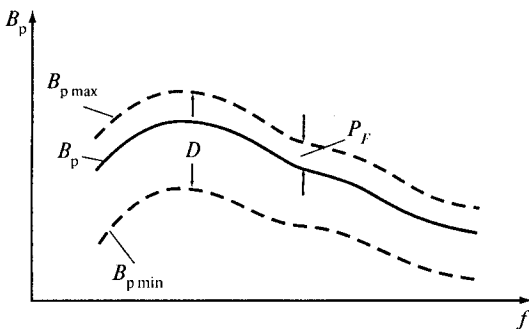


Рис. 2.6. Динамічний діапазон і пікфактор мови

Формантний розподіл в інтегральній формі – це залежність кількості мовних формант A від ширини, яку пропускає фільтр низькочастотної смуги частот $0..f$, припускаючи, що в межах цієї смуги немає ніяких втрат формантних складових. Цю залежність показано на рис. 2.7.

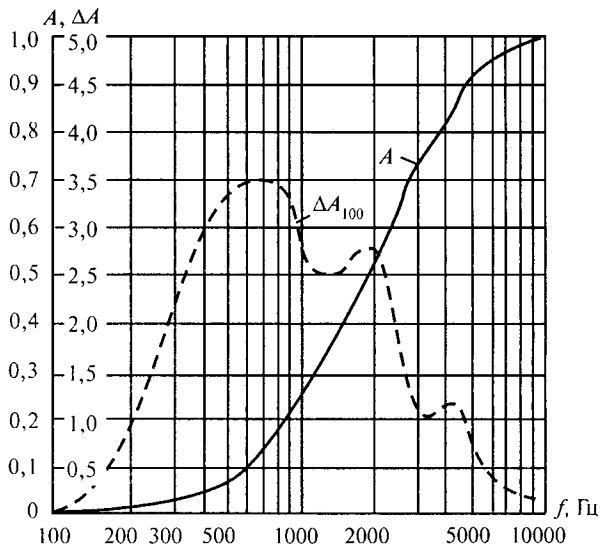


Рис. 2.7. Формантний розподіл в інтегральній (A) і диференціальній (ΔA_{100}) формі

Аналіз цієї залежності дозволяє зробити практичний висновок: основна частина формантних складових розміщується в смузі частот $300..3400$ Гц, достатній для забезпечення високої розбірливості мови. Тому така смуга регламентується як смуга частот нормального телефонного каналу, який ще називають каналом тональної частоти.

2.3. Механізм слухового сприйняття

Механізм слухового сприйняття, або принцип формування певних слухових образів під час дії на вхід слухового апарата звуків з тими чи тими фізичними характеристиками – складний процес, який до останнього часу являє собою об'єкт біопсихологічних досліджень.

Слуховий апарат – це комплексний аналізуючий пристрій, що працює як частотний, часовий та просторовий аналізатори звуків. Знання принципів роботи цього пристрою потрібні для аналізу характеристик слуху, які мають певне значення в техніці мовного зв'язку.

Слуховий аналізатор людини складається з периферичної та центральної частин. Периферична частина складається із зовнішнього, середнього та внутрішнього вуха і виконує переважно роль первинного приймача та амплітудно-частотного аналізатора звуку. Центральна частина, до якої надходять слухові нерви та слухові центри головного мозку (правий і лівий), виконує часовий та просторовий аналіз і синтез результуючого слухового образу.

Спрощену модель периферичної частини слухового апарата показано на рис. 2.8.

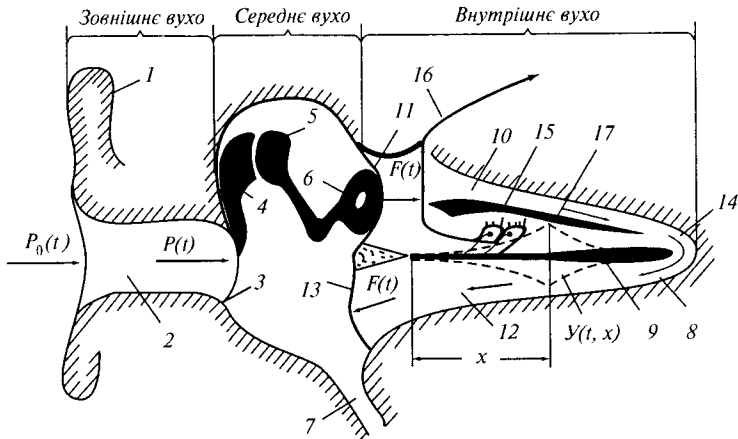


Рис. 2.8. Модель периферичної частини слухового апарата

До складу зовнішнього вуха входять слухова раковина 1 і слуховий канал 2, який виконує роль резонатора, що забезпечує невелике підсилення приймального звуку. Слуховий канал закінчується барабанною перетинкою 3, яка відділяє зовнішнє вухо від середнього.

У повітряній порожнині, яка обмежує середнє вухо, знаходиться система кісточок: молоточок 4, ковадло 5 та стріменце 6.

Ця система діє за принципом важеля і забезпечує трансформацію зусиль, які передаються від коливань барабанної перетинки на вхід внутрішнього вуха таким чином, що його амплітуда зменшується, а тиск збільшується приблизно в 15 разів. Середнє вухо спілкується із зовнішнім простором через евстахієву трубку 7 та порожнину носа, що потрібно для вирівнювання атмосферного тиску на барабанну перетинку.

Внутрішнє вухо являє собою кісткову трубку, згорнуту в спіраль з 2,5 витка, що називається завитком – 8. На рисунку завиток умовно показано в розпрямленому стані. У середині завиток по всій довжині розділено на дві частини перегородкою, основу якої складає базиллярна мембрана 9. У спрощеному вигляді базиллярну мембрану можна розглядати у вигляді напрямних, які йдуть по всій довжині завитка та поперечних пружних волокон, власна резонансна частота яких спадає від основи, де мембрана тонша та жорсткіша, до її кінця, де вона більш масивна та податлива.

Перегородка поділяє завиток на два канали, заповнені в'язкою рідиною – перелімфою. Верхній канал 10 із вхідним овальним вікном 11, затягнутим пружною перетинкою, називають вестибулярним ходом, а нижній канал 12, обмежений круглим вікном 13, що також має пружну перетинку, – барабанним ходом. На базиллярній мембрані розміщений кортіїв орган, який забезпечує спілкування периферичної та центральної частин слухового аналізатора. До його складу входять близько 30 000 чутливих волосяних клітинок 15, до яких підходять закінчення волокон слухового нерва 16, і текторіальна мембрана 17.

Функціонування внутрішнього вуха розглянемо за рис. 2.8. Звукові коливання, перетворені в середньому вусі в механічні коливання підвищеної сили, через стременце діють на пружну перетинку овального вікна. Столпчик рідини в сполучених вестибулярному та барабанному каналах також набуває коливального руху і діє на волокна базиллярної мембрани. Найбільшу амплітуду коливань $y(t)$ отримують волокна на відстані x від її основи, власна резонансна частота яких збігається з частотою зовнішнього коливання $p(t)$. За певного значення цієї амплітуди волоскові клітини кортієвого органа, які знаходяться над ділянкою з координатою x , дотикаються до текторіальної мембрани. У результаті цього в них виникає нервовий імпульс, який через відповідне волокно слухового нерва, що зв'язане з цією клітиною, надходить у мозок. Якщо звук має складний спектр, то одночасно через інші нервові закінчення передаються сигнали і від інших нервових клітин, які знаходяться над резонуючими ділянками базиллярної мембрани.

Природа нервових імпульсів має електричний характер. Комбінація збуджень, які надходять до відповідних ділянок у слухових центрах кори головного мозку, формує в свідомості той чи той звуковий образ.

Ідеалізовану еквівалентну схему центральної частини слухового аналізатора показано на рис. 2.9.

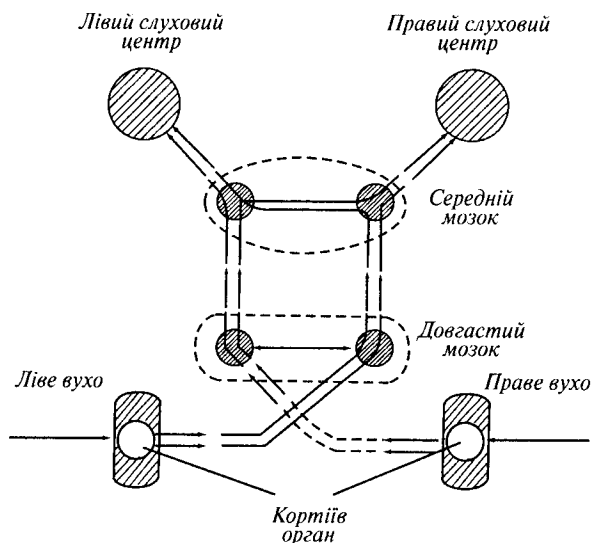


Рис. 2.9. Еквівалентна схема центральної частини слухового апарата

Кожний із слухових нервів, що йдуть від лівого та правого вуха, складається приблизно із 30 000 нервових волокон, які називаються *нейронами*.

Від кожного вуха є провідні шляхи до правого та лівого слухових центрів мозку. Як видно із схеми, сигнали, які надходять до слухових центрів, взаємодіють між собою також на проміжних рівнях довгастого і середнього мозку. Інтенсивність нервових імпульсів і швидкість їх поширення від кортієвого органа до слухового центра через вузлову систему з проміжними центрами залежить від величини збудження.

Сильні звукові імпульси мають більшу швидкість поширення, ніж слабкі. Інформація про частотний склад приймального звуку, яка передавалась по цій розгалуженій мережі від периферичної частини слухового органа до слухових центрів і доповнена даними часового аналізу, забезпечує розпізнавання цього звуку. Наявність двох синтезуючих центрів дозволяє здійснити просторову локалізацію звуку.

Фізичну природу роботи центральної частини слухового аналізатора вивчено не достатньо, але схематичне еквівалентне зображення дозволяє в загальному вигляді розглянути механізм сприйняття мовних звуків та їх розпізнавання.

2.4. Основні властивості та характеристики слуху

2.4.1. Поріг чутності та больовий поріг

Чутливість слухового апарата різна на різних частотах і характеризується порогом чутності.

Порогом чутності для чистого тону β_0 називають рівень мінімальної інтенсивності (звукового тиску), за якого виникає слухове відчуття цього тону:

$$\beta_0 = 10 \lg \frac{I_{\min}}{I_0} = 20 \lg \frac{P_{\min}}{P_0}$$

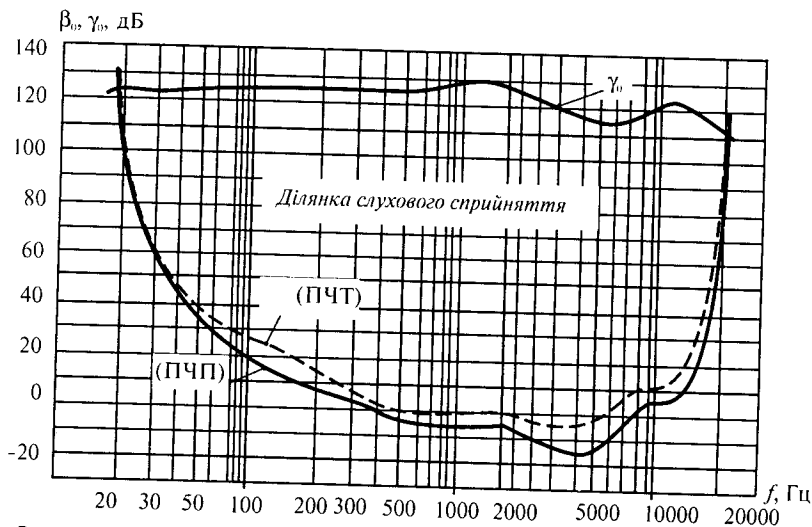


Рис. 2.10. Усереднений (50 %) поріг чутності (β_0) і поріг больового відчуття (γ_0) для нормального слуху: ПЧП – за вільним звуковим полем; ПЧТ – за тиском на вході в слуховий канал

Пороги чутності у людей, слух яких вважається нормальним, різняться дуже сильно (розкид до 12 дБ).

На поріг чутності впливає характер джерела звуку — звукове поле, або притиснутий до вуха телефон. В останньому випадку фізичні властивості вуха дещо змінюються і чутливість слуху знижується на всіх частотах приблизно на 6 дБ.

Якщо збільшувати інтенсивність тону після перевищення порогового рівня, то спочатку слухове відчуття буде поліпшуватися, потім з'являється подразлива дія, а після досягнення деякого максимального рівня слуховий апарат перестає сприймати цей тон як звук, а сприймає його як больове відчуття. Цей максимальний рівень γ_0 називається больовим порогом, або порогом больового відчуття. Перевищення больового порога може викликати травматичні пошкодження слухового апарата. У більшості людей больовий поріг визначається в діапазоні 120...130 дБ.

2.4.2. Рівень відчуття

Перевищення рівня інтенсивності звуку визначеної частоти над рівнем порога чутності на тій самій частоті називають рівнем відчуття цього звуку.

Рівень відчуття чистого тону E визначають різницею

$$E = B - \beta_0,$$

де B – рівень інтенсивності; β_0 – пороговий рівень тону певної частоти, або, якщо виразити B і β_0 через абсолютні значення інтенсивності,

$$B = 10 \lg \frac{I}{I_0} \text{ і } \beta_0 = 10 \lg \frac{I_{\min}}{I_0}, \quad E = 10 \lg \frac{I}{I_{\min}}.$$

Поняття про рівні відчуття ілюструє рис. 2.11. Тут рівні двох чистих тонів B_1 і B_2 зіставлено з кривою $\beta_0 = \varphi(f)$. Рівні відчуття цих тонів зображено відповідно відрізками E_1 і E_2 , причому тон частоти f_1 з рівнем B_1 – це позитивний рівень відчуття, а тон частоти f_2 з рівнем B_2 – негативний. Загальне правило, що ілюструє цей приклад: усі чутні звуки мають позитивні рівні, а всі нечутні – негативні рівні відчуття.

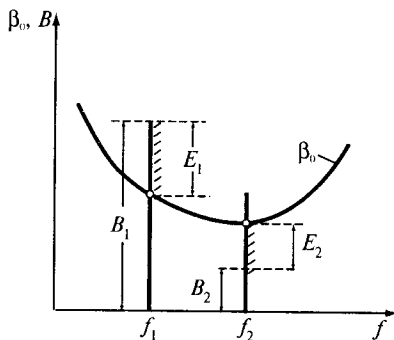


Рис. 2.11. Рівні відчуття чистих тонів

Рівні звуків, які мають суцільний спектр, обчислюють складнішим способом, який потребує врахування ще однієї властивості слуху – його інтегровальної здатності.

2.4.3. Критична смуга слуху

Якщо вухо сприймає не чистий тон, а звук із суцільним спектром, то на кожне волокно базиллярної мембрани діє не тільки коливання частотою, яка відповідає резонансній частоті цього волокна, але і група сусідніх складових із частотами в деяких границях, вищих і нижчих від неї. Ці границі визначаються гостротою резонансу, властивою базиллярній мембрані.

Природно, що ці сусідні складові будуть підсилювати слухове відчуття на кожній частоті спектра.

Властивість слуху підсумовувати збудження в деякому частотному діапазоні називають його *інтегровальною здатністю* за частотою.

Для кількісної оцінки ефекту цієї властивості слуху розглянемо рис. 2.12.

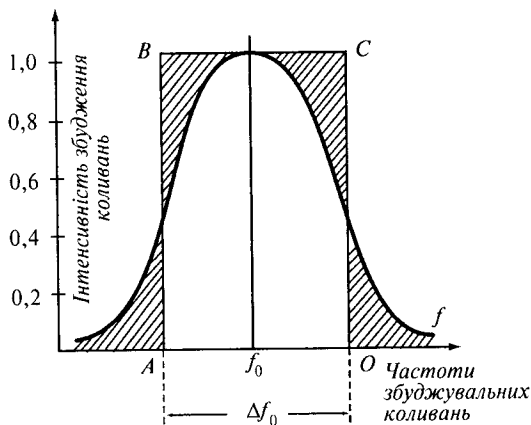


Рис. 2.12. Визначення ширини критичної смуги слуху

Тут зображено залежність інтенсивності збудження нейрона (в умовних одиницях), який зв'язаний через відповідну нервову клітину з точкою базиллярної мембрани (відповідним волоском), який має особисту резонансну частоту f_0 , від частоти збуджувальних коливальних частот.

Якщо у спектрі коливальних частот наявні складові всіх цих частот, тоді результуюче слухове відчуття буде визначатися сумою збуджень від усіх складових, значення яких зменшуватимуться у міру віддалення від частоти f_0 .

Для досить рівномірного спектра можна вважати, що результуюче збудження буде таким, яке було б отримане від підсумовування однакових збуджень у границях деякої смуги частот шириною Δf_k , для якої площа $ABCD$ дорівнює підінтегральній площі резонансної кривої.

Смуга частот, за межами якої вважають, що складові звуку з рівномірним суцільним спектром не впливають на величину слухового збудження на частоті, відповідній центру цієї смуги, називають *критичною смугою слуху*. Вона кількісно характеризує інтегровальну здатність слуху за частотою.

2.4.4. Ефективний рівень відчуття

Ефективним рівнем відчуття звуку суцільного спектра на заданій частоті називають рівень відчуття його складових у межах критичної смуги слуху, середня частота якої збігається із заданою частотою.

Якщо середня спектральна інтенсивність звуку суцільного спектра в межах критичної смуги дорівнює I_1 , то його загальна інтенсивність в цій смузі буде дорівнювати $I_1 \Delta f_k$.

Тоді згідно з визначенням ефективного рівня відчуття його визначають як

$$E = B_{\Delta f_k} - \beta_0 = 10 \lg \frac{I_1 \Delta f_k}{I_0} - \beta_0 = 10 \lg \frac{I_1}{I_0} + 10 \lg \Delta f_k - \beta_0,$$

де $B_{\Delta f_k}$ – рівень інтенсивності в смузі.

Беручи до уваги, що $B_1 = 10 \lg \frac{I_1}{I_0}$ (спектральний рівень інтенсивності), а $10 \lg \Delta f_k = k$ – логарифмічна ширина критичної смуги слуху, то формула для визначення ефективного рівня відчуття звуку суцільного спектра для заданої частоти має вигляд

$$E = B_1 + k - \beta_0,$$

де $B_1 + k$ – енергія, яка створює результуючий слуховий ефект на частоті f_1 , якій відповідає спектральний рівень та логарифмічна ширина критичної смуги слуху k .

2.4.5. Гучність та адаптація слуху

Оскільки чутливість слуху у різних людей різна, один і той самий звук для одних може здаватися гучним, а для інших – тихим. Тому гучність звуку можна визначити як суб'єктивні відчуття інтенсивності дії звуку на слух певного індивідуума. Для цього індивідуума у зв'язку з час-

тотною залежністю порога чутності звуки різних частот мають однакові інтенсивності, але за гучністю будуть різними.

Слух людини під час сприймання тихих звуків змінюється, поріг чутності зменшується, і навпаки, коли сильні звуки (дуже чутні), поріг чутності збільшується, чутливість слуху спадає. Цей процес пристосування слуху людини до інтенсивності сприйманого звуку одержав назву «адаптація слуху до звуку».

Фізична природа цього процесу полягає в тому, що біомеханічний опір слухового органа або зменшується, або збільшується.

Процес адаптації виникає не миттєво, а з деяким запізненням. Перехід від одного стану до другого потребує часу на пристосування до відповідного рівня інтенсивності. Тому в процесі телефонного зв'язку перші слова (або звуки) сприймаються не розбірливо, що викликає перепиткування.

2.4.6. Маскування звуків

Якщо одночасно з корисним сигналом до вуха людини, що слухає, надходить сторонній звук (наприклад, шум або інша передача, яка через свої електромагнітні параметри впливає на сусідню розмову), тоді за певних відношень рівнів і частот цих звуків сторонній звук може повністю або частково заглушити корисний сигнал.

Явище повного пропадань або ослаблення чутності корисного сигналу на фоні заважального звуку називають *маскуванням*. Кожний звук при цьому називають *маскованим*, а заважальний – *маскувальним*.

Фізіологічний механізм маскування залежить від властивості адаптації слуху, яка виявляється в тому частотному діапазоні, в якому діє маскувальний звук. У цьому діапазоні виникає підвищення порога чутності, в результаті якого ті складові корисного сигналу, що раніше мали вищий від порогового рівень, тепер виявляються або нижчими за поріг чутності, або «вибиваються» із сприйняття, або сприймаються з пониженим рівнем відчуття.

Наочно процес маскування можна показати на прикладі маскування чистим тоном (рис. 2.13).

Тут сильний маскувальний тон частоти f_m з рівнем B_m зумовлює підвищення порога чутності в околі частоти f_m і її гармонік $2f_m$ і $3f_m$.

Поріг чутності β_0 набуває нового положення β_m . У результаті цього виявляється, що складова корисного сигналу частотою f_1 взагалі не буде сприйматися, оскільки буде повністю замаскована, а складова частотою f_2 прослуховуватиметься зі значним зниженням рівня відчуття, тобто вона буде частково маскована.

Маскуванням називають різницю між рівнем порога чутності під час дії шуму на слуховий апарат β_m і рівнем порога чутності в тишині β_0 для фіксованої частоти, тобто

$$M = \beta_m - \beta_0.$$

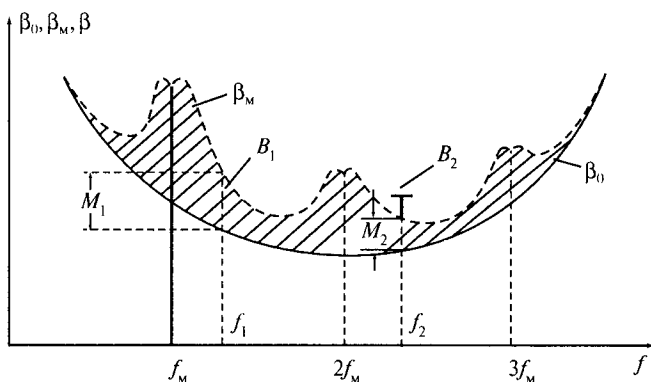


Рис. 2.13. Маскування чистим тоном

Отже, виявляється, що вплив шуму на процес слухового сприйняття еквівалентний включенню в тракт додаткового частотно-залежного затування, величина якого на кожній частоті f дорівнює $M(f)$ децибелів.

Якщо маскування виконується не чистим тоном, а шумом із суцільним спектром, що головним чином і трапляється в практиці, то зсув порога чутності на кожній частоті f через інтегровальну здатність слуху буде створюватися усіма складовими маскувального спектра, який знаходиться в межах критичної смуги слуху Δf_k , центром якої є ця частота f .

Величина маскування M від шуму суцільного спектра на кожній частоті f – це повністю визначена функція від ефективного рівня відчуття шуму Z . Цю функцію зображено на рис. 2.14.

Щоб розрахувати маскування M від шуму суцільного спектра, треба мати спектральну характеристику цього шуму, тобто залежність $B_{ш} = \varphi(f)$, де $B_{ш}$ – спектральний рівень шуму.

Розрахунок виконують у цьому разі для кожної частоти так, що спочатку визначають ефективний рівень відчуття шуму за формулою

$$Z = B_{ш} + k - \beta_0,$$

а потім за графіком $M = f(Z)$ шукають маскування M .

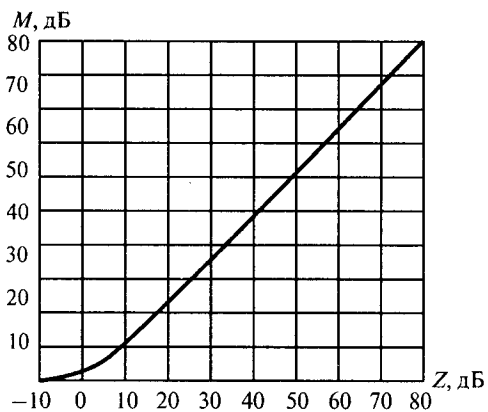


Рис. 2.14. Залежність маскування звуком суцільного спектра від ефективного рівня відчуття шуму

Отже, в другому розділі виявлено частотні звукові характеристики у процесі формування мовного сигналу та вплив на них характеристик шумів.

3. Перетворення акустичної енергії мови в електричну і навпаки

3.1. Призначення і класифікація перетворювачів

Перетворювачі в трактах телефонного передавання мови призначені для перетворення акустичної енергії мови в електричну енергію на передавальному кінці тракту та оберненого перетворення електричної енергії в акустичну на приймальному кінці тракту. Згідно з цим розрізняють акустично-електричні та електроакустичні перетворення відповідно до таких етапів:

1) перетворення енергії акустичного збудження в механічну енергію руху коливальної системи перетворювача, яка, в свою чергу, перетворюється у звукову енергію.

2) перетворення електричної енергії в механічну, що перетворюється у звукову.

3) перетворення, використовувані в ларингофонах, в яких відбувається неповний цикл першого етапу, тобто без акустичного збудження, коли виконується безпосереднє перетворення механічної енергії в електричну.

Перетворювачі такого типу належать до класу механіко-електричних перетворювачів.

Винахід телефону 1876 року по суті став винаходом електроакустичного перетворювача, здатного виконувати обидва зазначені перетворення.

У сучасній апаратурі для передавання мови використовують різні системи і типи перетворювачів, які можна класифікувати за принципом дії та призначенням. За принципом дії перетворювачі можуть бути: вугільними, електромагнітними, електродинамічними, конденсаторними (електростатичними) та п'єзоелектричними. Усі системи перетворювачів, на які їх поділяють за принципом дії, крім вугільних, можуть виконувати обидві функції; їх називають оборотними.

За призначенням перетворювачі поділяють на передавальні (мікрофони та ларингофони) і приймальні (телефони і гучномовці).

3.2. Загальний принцип побудови і дії обернених перетворювачів

Кожний перетворювач складається з акустичної, механічної та електричної частин. Акустична частина – це випромінювач або приймач акустичних коливань. Механічна частина перетворювача являє собою коливальну систему, зміщення якої і зумовлює процес перетворення енергії.

Електричне коло (ланцюг) перетворювача є генераторним виходом або приймачем електричного сигналу.

Основний принцип дії перетворювача визначають за взаємодією його механічної коливальної системи з електричним колом. Тому для встановлення закономірностей, властивих більшості основних типів перетворювачів, на першому етапі доцільно розглядати перетворювач як електромеханічну систему.

Із цього погляду будь-який лінійний обернений перетворювач можна подати у вигляді чотириполосника з механічним входом і електричним виходом (передавальні пристрої) або з електричним входом і механічним виходом (приймальні перетворювачі).

Узагальнену еквівалентну схему оберненого перетворювача показано на рис. 3.1.

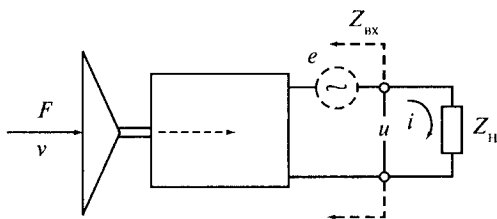


Рис. 3.1. Еквівалентна схема оберненого перетворювача

Якщо на механічному вході перетворювача діє змінна сила F , що створює коливальну швидкість v , то змінюються параметри (характеристики) електричного кола, що, у свою чергу, змінює електричні показники кола у вигляді електрорушійної сили (ЕРС), а у разі замкненого електричного кола створює електричний струм i .

Якщо ж до електричного входу перетворювача прикласти зовнішню напругу u , яка викликає струм i , то на механічному виході виникає сила F і коливальна швидкість v .

Відношення ЕРС e до коливальної швидкості v , яка її зумовила, дорівнює відношенню механічної сили F до сили струму i , що її спричинив, і називається *коефіцієнтом електромеханічного зв'язку* K :

$$K = \frac{e}{v} = \frac{F}{i}.$$

Коефіцієнт K характеризує ефективність перетворення: чим більше значення цього коефіцієнта, тим більша частина енергії одного виду перетворюється в енергію іншого виду.

3.3. Основні системи перетворювачів

3.3.1. Перетворювачі електромагнітної системи

У перетворювачах електромагнітної системи використовують обернений процес перетворення енергії під час взаємодії магнітного поля з феромагнітним якорем. Принцип дії такого перетворювача показано на рис. 3.2.

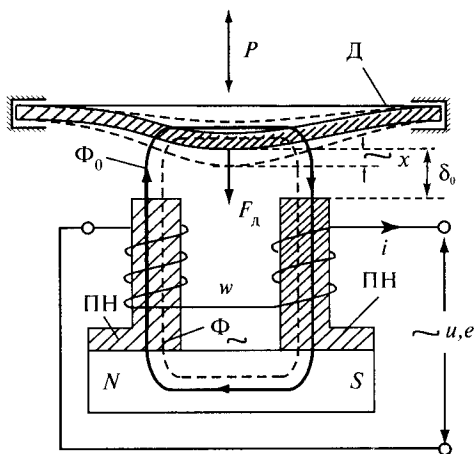


Рис. 3.2. Принцип дії електромагнітного перетворювача з простою електромагнітною системою

Магнітне коло перетворювача складають: постійний магніт NS з полюсними надставками ПН і пружною феромагнітною діафрагмою Д, яка виконує функцію якоря. Між діафрагмою і полюсними надставками розміщено повітряний зазор δ_0 , який визначається початковим прогином діафрагми. На полюсні надставки надіто котушки з обмоткою, яка має загальну кількість витків w .

Постійний магніт створює у повітряному зазорі магнітний потік $\Phi_0 = \frac{E_m}{r_m}$, де E_m – магніторушійна сила; r_m – опір магнітного кола.

Під час роботи перетворювача, виконуючого функції телефону або гучномовця, по обмотці проходить змінний струм i , який створює в повітряному зазорі додатковий змінний магнітний потік $\Phi \cong \frac{iw}{r_m}$. Залежно від напрямку струму змінний магнітний потік буде або збільшувати початкову силу притягання діафрагми F_a , або зменшувати її. У результаті цих дій діафрагма створюватиме коливальний рух зі змінною x в обидва боки

від початкового положення, який супроводжується випромінюванням звукових хвиль.

Під час роботи в режимі акустичні коливання – електричні коливання звукових хвиль – діють на діафрагму і надають їй коливального руху. При цьому змінюється повітряний зазор δ_0 , що змінює магнітний опір r_m , і відповідно магнітний потік Φ_0 . Змінний магнітний потік наводить у витках обмотки ЕРС e .

Силу притягання F_q обчислюють згідно з відомою формулою Максвелла. Сила притягання F_q , яку розвиває електромагніт, має вигляд

$$F_q = \frac{\Phi^2}{2\mu_0 S},$$

де Φ – магнітний потік; S – пронизана ним площа; μ_0 – магнітна проникність повітря (магнітна проникність вакууму $\mu_0 = 4\pi 10^{-7}$).

3.3.2. Перетворювачі електродинамічної системи

Розглянемо перетворення акустичних сигналів в електричні і навпаки за схемою, яку зображено на рис. 3.3.

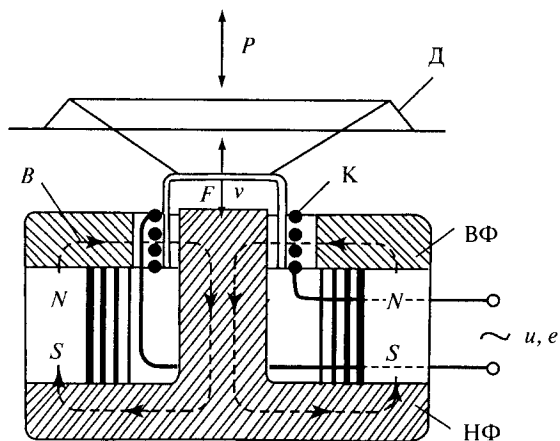


Рис. 3.3. Принцип побудови перетворювача електродинамічної системи

Магнітна система перетворювача складається з кільцевого магніту NS і полюсних надставок: верхнього фланця ВФ і нижнього фланця НФ з керном. Між верхнім фланцем ВФ і керном розміщено повітряний зазор, у якому знаходиться легка котушка K з обмоткою, яка може вертикально переміщуватися. З котушкою пов'язана діафрагма D – випромінювач або приймач звукових коливань.

Під час роботи перетворювача в режимі «електричний сигнал – акустичний звук» струм звукової частоти, який проходить витками котушки, взаємодіє з постійним магнітним полем так, що котушка або втягується в повітряний зазор, або виштовхується із нього залежно від напрямку струму. При цьому діафрагма випромінює звукові коливання.

Якщо величина магнітної індукції у повітряному зазорі, яка створюється постійним магнітом, дорівнює B , сила струму в обмотці i , а загальна довжина провідника в котушці l , тоді сила, яка діє на провідник (котушку),

$$F = Bil.$$

Під час роботи перетворювача в режимі «акустичний звук – електричні коливання» котушка набуває коливального руху зі швидкістю ϑ ; при цьому в її витках виникає ЕРС

$$e = Bl\vartheta.$$

Звідси коефіцієнт електромеханічного зв'язку для перетворювачів електродинамічного типу

$$K = \frac{F}{i} = \frac{e}{\vartheta} = Bl.$$

Позитивною якістю таких перетворювачів є відносно невеликі величини частотних і нелінійних спотворень. Це забезпечує можливість перетворення великих потужностей, наприклад, під час роботи гучномовця. Ці перетворювачі використовують як мікрофони і телефони для вимірювань та в апаратурі звукозапису.

3.3.3. Перетворювачі конденсаторної системи

У конденсаторній системі перетворювачів використовується взаємодія між обкладками зарядженого конденсатора, одна з яких рухома.

Принцип дії показано на рис. 3.4. Тут над нерухомою масивною пластинкою П розміщено пружну мембрану М, яка зазвичай є туго натягнутою плівкою з діуралюмінієвої фольги. Відстань між обкладками одержаного таким чином конденсатора дорівнює a . Вихід електричного кола замкнений на батарею і навантажувальний опір R .

Під час роботи в режимі перетворювача «електричний струм – акустичні коливання» до його електричного входу підводиться змінна напруга u , яка викликає зміну заряду на обкладках конденсатора. При цьому змінюється сила притягання між ними, внаслідок чого мембрана М набуває коливального руху. Батарея відіграє роль магніту в раніше розглянутих перетворювачах. Якби $E = 0$, то конденсатор заряджувався і розряджувався б з подвійною частотою відносно вхідного сигналу.

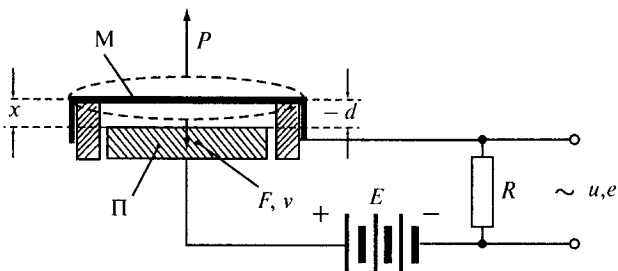


Рис. 3.4. Принцип дії перетворювача конденсаторної системи

Конденсаторні перетворювачі мають бути малогабаритними. Їх застосовують здебільшого для вимірювання звукових тисків.

3.3.4. Перетворювачі п'єзоелектричної системи

У п'єзоелектричних перетворювачах використовується здатність деяких кристалів виділяти на своїх гранях електричний заряд дією стиснення (прямий п'єзоэффект), або деформуватися під дією прикладеної напруги (обернений п'єзоэффект). П'єзоэффект властивий таким кристалам, як кварц, турмалін, сегнетова сіль, титанат барію і звичайний цукор. Для використання в електроакустичних перетворювачах із кристалів вирізаються квадратні або прямокутні пластинки, певним чином орієнтовані відносно його кристалографічних осей. Принцип побудови і дії п'єзоелектричного перетворювача показано на рис. 3.5. Його основною частиною є так званий біморфний п'єзоелемент П, який складається із двох пластинок, складених так, що під час механічної дії на їх зімкнутих гранях, оскільки на взаємно віддалених виділяються заряди однакового знака. Для вивідних електродів E п'єзоелемента використовують сталеві листи, один з яких розміщується між пластинками, а два інші – на їх зовнішніх гранях (рис. 3.5).

Три кути квадратного біморфного п'єзоелемента приклеюють до основи, на якій їх закріплюють, а четвертий кут залишають вільним.

Пластинки, які складають п'єзоелемент, складені так, що їх деформація виникає у взаємно перпендикулярних напрямках, тобто одна з пластинок прагне подовжитися по вертикальній осі і скоротитися по горизонтальній, тоді для другої цей процес буде оборотним.

Під час роботи в режимі «електричний сигнал – акустичний звук» до електричного входу перетворювача підводиться напруга мовного сигналу u , яка створює електричний струм i . Оскільки три кути п'єзоелемента закріплені, а напрям деформації пластин вільного кута взаємно протилежні, то останній буде здійснювати коливання у вертикальній площині, розвиваючи зусилля F і приводячи в дію діафрагму – випромінювач D .

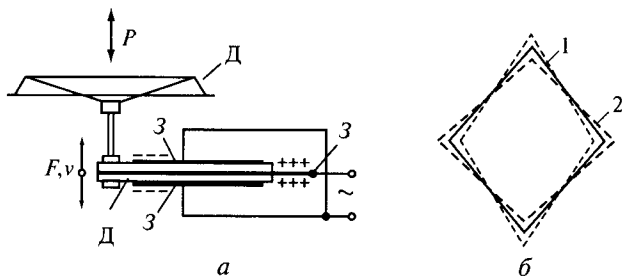


Рис. 3.5. Принцип дії п'єзоелектричного перетворювача (а) та біморфного п'єзоелемента (б)

Під час роботи перетворювача в режимі «акустичний звук – електричний сигнал» змінне зусилля F , яке діє на рухомий кут п'єзоелемента, надає йому швидкість ϑ і викликає появу на його гранях змінного електричного заряду, який, в свою чергу, створює різницю потенціалів e на електричному виході перетворювача.

Позитивні якості — простота побудови, відносно невелика вартість їх виготовлення, а негативні – труднощі відтворення низьких частот, температурна залежність, низька вологостійкість.

3.3.5. Вугільні перетворювачі

Вугільні перетворювачі – це необоротні активні перетворювачі, використовувані тільки в режимі «акустичний звук – електричний сигнал».

Дія вугільного перетворювача основана на властивості вугільного порошку змінювати свій опір електричному струму залежно від тиску на порошок, викликаючи його ущільнення або розрідження.

Принцип побудови вугільного перетворювача показано рис. 3.6.

Його основними частинами є камера К з вугільним порошком ВП, нерухомий НЕ та рухомий РЕ електроди і зв'язана з РЕ діафрагма Д. Рухомі та нерухомі електроди в сучасних перетворювачах виготовляють металевими.

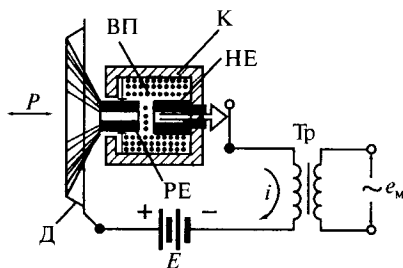


Рис. 3.6. Принцип побудови вугільного перетворювача

Звуковий тиск, який діє на діафрагму, надає їй коливального руху, причому ВП піддається перемінним змінам стиснення і розрідженням. При цьому змінюється його опір, а відповідно і електричний струм. У разі розімкнення зовнішнього кола на затискачах первинної обмотки виникає напруга звукової частоти, яка і розглядається як ЕРС.

На схемі показано, що джерелом живлення є місцева батарея. Якщо використовується джерело центральної батареї, то трансформатора немає. Таку схему застосовують, коли такі перетворювачі виконують функції мікрофонів у телефонних апаратах (ТА) з центральним живленням АТС телефонної мережі загального користування.

3.4. Експлуатаційні характеристики основних пристроїв термінального устаткування для передавання та приймання мовних звуків

До основних вузлів термінального устаткування для перетворення, передавання та приймання мовних звуків належать: мікрофони, телефони, ларингофони та гучномовці.

3.4.1. Мікрофони

В експлуатації знаходяться мікрофони різних систем: вугільні, електромагнітні, електродинамічні, конденсаторні та п'єзоелектричні. Їх якість оцінюється рядом параметрів, частина з яких регламентується згідно з міжнародними та державними стандартами.

Основні параметри та характеристики мікрофонів:

- чутливість;
- частотна характеристика чутливості (ЧХЧ);
- потужність, що віддається;
- опір;
- амплітудна характеристика;
- коефіцієнт нелінійних спотворень;
- характеристика направленості;
- шумостійкість.

Чутливість мікрофона. Чутливістю мікрофона називають відношення ЕРС, яку розвиває мікрофон, до звукового тиску, що діє на нього

$$\left(\frac{\text{мВ}}{\text{Нп/м}^2}\right):$$

$$K_m = \frac{e}{p},$$

де e – діюче значення ЕРС мікрофона; p – звуковий тиск.

Частотна характеристика чутливості. Приблизний вигляд ЧХЧ показано на рис. 3.7. Її нерівномірність пояснюється резонансними явищами в механоакустичній системі перетворювача.

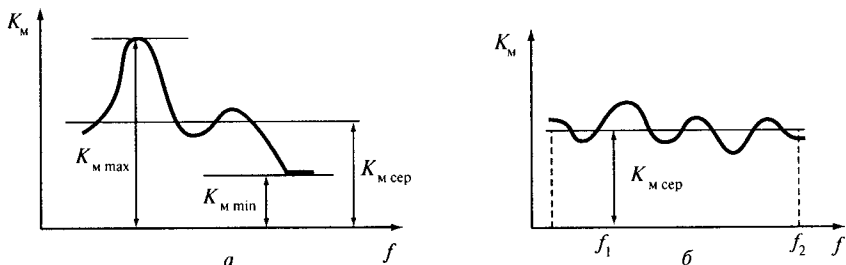


Рис. 3.7. Частотна характеристика чутливості мікрофона:
a – резонансного, *б* – широкосмугового

Нерівномірність ЧХЧ мікрофона – це показник частотних спотворень, який визначають за формулою

$$\beta = 20 \lg \frac{K_{m \max}}{K_{m \min}},$$

де $K_{m \max}$ – максимальне значення ЧХЧ; $K_{m \min}$ – мінімальне значення в цій смузі частот.

Для порівняння мікрофонів один з одним користуються поняттям «середня чутливість», під якою розуміють висоту прямокутника, площа якого рівновелика підінтегральній площі кривої ЧХЧ мікрофона в межах заданої смуги частот:

$$K_{m \text{ сep}} = \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} K_m(f) df.$$

За виглядом ЧХЧ мікрофони поділяють на резонансні та широкосмугові. Резонансні мікрофони мають у вузьких смугах високу чутливість, але велику нерівномірність. Широкосмуговим мікрофонам властива рівномірна ЧХЧ у широкому діапазоні спектра частот, але невелика середня чутливість.

Для повного порівняння мікрофонів не достатньо знати параметри чутливості. Для цього треба знати енергетичні параметри, до яких належать потужність, що видає мікрофон. Так, наприклад, з двох мікрофонів з однаковими середніми чутливостями більшу потужність видає на узгоджене навантаження той мікрофон, внутрішній опір Z_m якого буде менший. Тому для порівняння мікрофонів за цим показником їх зводять до стандартного значення опору, що дорівнює 600 Ом, і обчислюють зведену чутливість.

Зведеною чутливістю мікрофона називають чутливість, яку він мав би, якби модуль його внутрішнього опору дорівнював 600 Ом у разі віддачі тієї ж потужності:

$$K'_m = K_m \sqrt{\frac{600}{|Z_m|}}.$$

Середню зведену чутливість обчислюють аналогічно $K_{m \text{ сер}}$ за частотною характеристикою $K_m = \varphi(f)$.

Для спрощення обчислення $K_{m \text{ сер}}$ користуються формулою

$$K'_{m \text{ сер}} = K_{m \text{ сер}} \sqrt{\frac{600}{|Z_{m \text{ 1000 Гц}}|}}.$$

Потужність, яку видає мікрофон на навантаження. Можна використовувати як самостійний показник, який визначають за формулою

$$P_m = \frac{U^2}{|Z_m|},$$

де U – напруга, яку розвиває мікрофон на узгодженому з ним навантаженні $Z_n = Z_m$.

Напругу U ще називають віддачею мікрофона. Опір мікрофона Z_m визначає умови його узгодженості з навантаженням, на яке він має працювати.

Амплітудна характеристика мікрофона. Являє собою залежність ЕРС, яку розвиває мікрофон, від звукового тиску P , що збуджує мікрофон.

Приблизний вигляд амплітудної характеристики вугільного мікрофона показано на рис. 3.8.

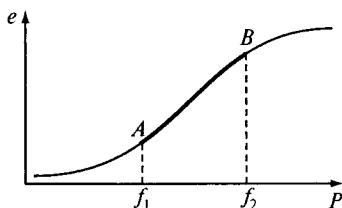


Рис. 3.8. Амплітудна характеристика вугільного мікрофона: AB – прямолінійна ділянка, на якій нелінійні спотворення мінімальні

Коефіцієнт нелінійних спотворень знаходять за формулою

$$K_n = \frac{\sqrt{e_2^2 + e_3^2 + e_4^2 + \dots}}{e_1} \cdot 100,$$

де e_i – дійсне значення амплітуди i -ї гармоніки ЕРС мікрофона.

Характеристика напрямленості мікрофона. Це залежність відносної чутливості $\frac{K_M}{K_{M \max}}$ від кута α між напрямом звукової хвилі та перпендикуляром до площини діаграми мікрофона. Приклад напрямленості показано на рис. 3.9.

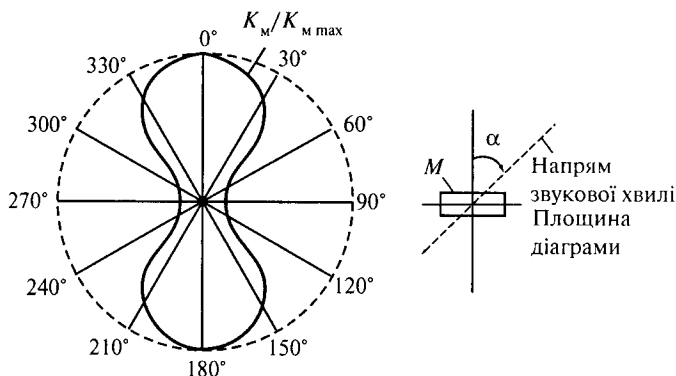


Рис. 3.9. Характеристика напрямленості мікрофона

Шумостійкість мікрофона визначається як 20 логарифмів відношення напруги U_M до напруги $U_{\text{ш}}$ на навантаженні:

$$\gamma_M = 20 \lg \frac{U_M}{U_{\text{ш}}},$$

де U_M – напруга, яку розвиває мікрофон на навантаженні від збудження мовою; $U_{\text{ш}}$ – напруга, яку розвиває мікрофон на навантаженні від збудження шумом.

3.4.2. Телефони

Найбільшого використання для телефонів набула електромагнітна система. Телефони електродинамічної системи застосовують головним чином у вимірвальній техніці.

Основні параметри і характеристики телефонів:

- чутливість;
- частотна характеристика чутливості;
- опір;
- амплітудна характеристика;
- коефіцієнт нелінійних спотворень.

Чутливість телефону. Чутливістю телефону називають відношення звукового тиску, який розвиває телефон, до напруги на його електричному вході:

$$K_T = \frac{P}{U},$$

де P – звуковий тиск, який розвиває телефон, Нп/м^2 ; U – напруга на його затискачах, В.

Частотна характеристика чутливості. Це функція чутливості від частоти $K_T = \varphi(f)$. Приблизний вигляд ЧХЧ телефону показано на рис. 3.10.

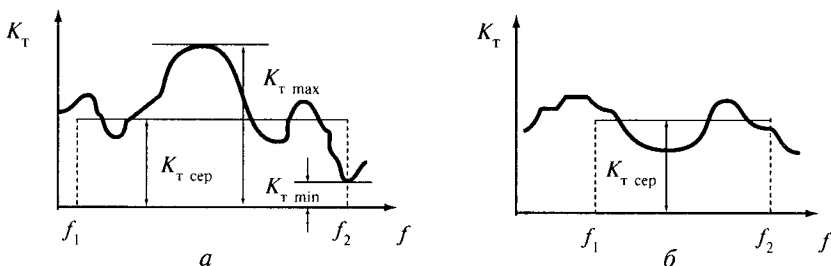


Рис. 3.10. Частотна характеристика чутливості телефону:
а – резонансної; б – широкосмугової

Так само, як і для мікрофона, нерівномірність ЧХЧ телефону визначають логарифмічним відношенням максимального ($K_{T \max}$) та мінімального ($K_{T \min}$) значень чутливості телефону в заданій смузі частот:

$$\beta = 20 \lg \frac{K_{T \max}}{K_{T \min}}.$$

Для порівняння телефонів користуються поняттям середньої чутливості, під якою розуміють висоту прямокутника, площа якого рівновелика підінтегральній площі кривої ЧХЧ телефону в межах заданої смуги частот:

$$K_{T \text{ сеп}} = \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} K_T(f) df.$$

Телефони, так само, як і мікрофони, за видом своєї ЧХЧ чутливості можна поділити на резонансні та широкосмугові (див. рис. 3.10).

Середня чутливість телефону без урахування його вхідного опору не дозволяє порівняти його з іншими телефонами за енергетичною ефектив-

ністю, тобто за споживанням ним потужності з електричного кола. Так, наприклад, з двох телефонів з однаковою середньою чутливістю, але з різним вхідним опором менше потужності з електричного кола споживає той телефон, у якого більший вхідний опір Z_T , оскільки потужність, яку споживає телефон, визначають за відомою формулою $P_T = \frac{U^2}{|Z_T|}$. Тому для

більш об'єктивного порівняння телефонів за цими показниками їх зводять до стандартного значення опору, що, як і для мікрофонів, становить $Z_T = 600$ Ом, і обчислюють зведену чутливість K_T .

Зведеною чутливістю телефону називають чутливість, яку він мав би, якби модуль його вхідного опору дорівнював 600 Ом за такої самої потужності, що споживається:

$$K_T = K_{T\text{сеп}} \sqrt{\frac{|Z_T|}{600}}$$

За частотною характеристикою $K_T = \varphi(f)$ знаходять середню зведену чутливість телефону $K_{T\text{сеп}}$. Для приблизних розрахунків як середнє значення зведеної чутливості можна брати:

$$K_{T\text{сеп}} = K_T \sqrt{\frac{|Z_{T\ 1000\ \text{Гц}}|}{600}}$$

де $Z_{T\ 1000\ \text{Гц}}$ – опір телефону на частоті 1000 Гц.

Звуковий тиск, який розвиває телефон для визначеної напруги на його затискачах, називають віддачею телефону.

Суттєвим експлуатаційним параметром телефону є вхідний опір Z_T , за яким визначають його узгодженість із вхідним опором кола або схеми, у яку його вмикають.

Амплітудна характеристика телефона. Це функціональна залежність звукового тиску, який розвиває телефон, від прикладеної на його вхідні затискачі напруги.

Коефіцієнт нелінійних спотворень обчислюють за тією самою формулою, яку використовують для розрахунків параметрів мікрофона.

3.4.3. Ларингофони

Ларингофон – це механоелектричний перетворювач, у якому процес мовлення на звукоприймальну частину здійснюється завдяки м'язам гортані.

Ларингофони будують з використанням вугільних та електромагнітних перетворювачів. Ларингофони не чутливі до акустичних шумів (рис. 3.11).

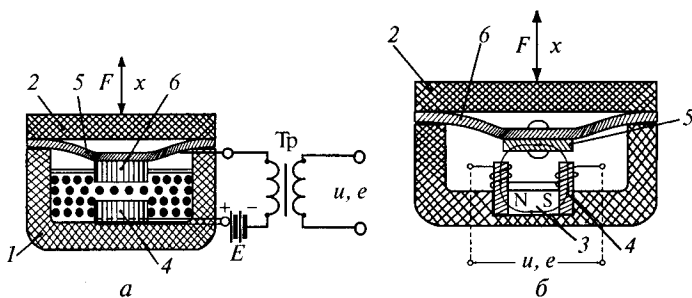


Рис. 3.11. Принцип дії ларингофона: *а* – вугільного, *б* – електромагнітного

Характеристики ларингофона. Чутливістю ларингофона називають відношення ЕРС ε , яку він розвиває, до коливальної швидкості його механічного збудження \dot{x} :

$$K_{\text{л}} = \frac{\varepsilon}{\dot{x}}.$$

Шумостійкість ларингофона визначають як логарифмічне відношення чутливості ларингофона під час механічного збудження $K_{\text{л}}$ до його чутливості під час акустичного збудження $K_{\text{л.а}}$:

$$\beta_{\text{ш}} = 20 \lg \frac{K_{\text{л}}}{K_{\text{л.а}}},$$

$$\text{де } K_{\text{л}} = \frac{\varepsilon}{\dot{x}}; K_{\text{л.а}} = \frac{\varepsilon}{P}.$$

Принцип дії ларингофона показано на рис. 3.11 (вугільного – рис. 3.11, *а*, електромагнітного – рис. 3.11, *б*). Вугільний ларингофон складається із корпуса 1, кришки 2, магніта 3, котушки 4, діафрагми 5.

Механічні коливання м'язів гортані зумовлюють коливання діафрагми, що, у свою чергу, зводиться до зміни відстані між нею та надставками котушки, при цьому змінюється опір магнітному потоку, що зумовлює появу змінного електричного струму (рис. 3.11, *б*). Як і для вугільних перетворювачів, роботу вугільних ларингофонів показано на рис. 3.11, *а*.

3.4.4. Гучномовець

Гучномовець застосовують у засобах гучномовного телефонного зв'язку, системах звукопідсилення для озвучення приміщень, майданів і в кінцевих перетворювачах у радіоприймачах та магнітофонах.

Принцип перетворення енергії аналогічний діям телефонів.

Основні експлуатаційні параметри і характеристики гучномовців:

- чутливість;
- частотна характеристика чутливості;
- номінальна електрична потужність;
- коефіцієнт нелінійності;
- коефіцієнт корисної дії (ККД);
- характеристика направленості.

Чутливість гучномовця визначають як відношення звукового тиску, що розвивається гучномовцем в точці, яка знаходиться на відстані 1 м від нього по осі випромінювання, до напруги на його електричному вході:

$$K_r = \frac{P_0}{U}$$

Користуються ще абсолютною осьовою чутливістю гучномовця:

$$K_{r.a} = \frac{P_0}{\sqrt{UI}}$$

де P_0 – звуковий тиск на відстані 1 м по осі випромінювання; U – діюче значення напруги; I – діюче значення електричного струму.

Частотна характеристика чутливості гучномовця визначається частотною характеристикою перетворювача, який використовують у гучномовці. Характеристику направленості гучномовця зображено на рис. 3.12.

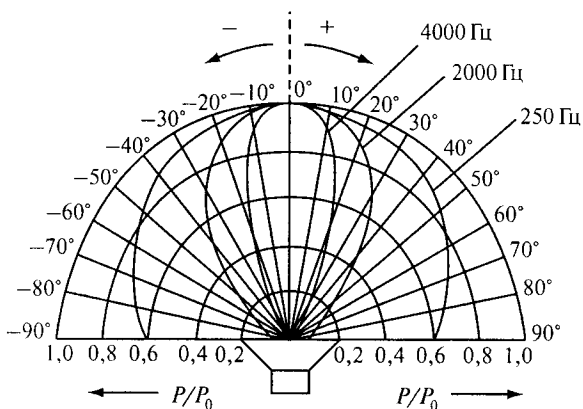


Рис. 3.12. Характеристика направленості гучномовця

Отже, перетворюючи звукові мовні сигнали в електричні коливання і навпаки, використовують фізичні явища змінювання опору як електричного, так і акустичного.

4. Телефонні апарати

4.1. Загальні положення.

Класифікація телефонних апаратів та їх схем

Термінали телефонної мережі – це звичайні телефонні апарати. У світі існують різноманітні конструкції ТА, хоч електрично всі вони подібні. На сьогодні це найбільш масовий термінал.

Телефонні апарати класифікують за такими показниками:

- призначенням;
- способом живлення;
- способом обслуговування з'єднань;
- конструкцією;
- характеристиками схем.

За призначенням ТА поділяють на апарати загального застосування: абонентські апарати загального користування та монетні (карткові) автомати), апарати спеціального призначення (військово-польові, корабельні, шахтні (вибухонебезпечні)), ТА для групового вмикання та ін.

За способом живлення розрізняють ТА місцевої батареї (МБ) і центральної батареї (ЦБ).

За способом обслуговування з'єднань, які визначаються типом телефонної станції, розрізняють апарати ручної телефонної станції (РТС) та апарати АТС.

За конструкцією ТА поділяють на настільні, стінні та переносні.

За характеристикою схеми існує розгалужена класифікація ТА, а саме: за способом розділення розмовного та викличного ланцюгів, за способом вмикання розмовних приладів та за наявністю місцевого ефекту.

Спосіб розділення розмовного та викличного ланцюга. Схема кожного ТА складається з розмовних і викличних пристроїв. Вона має бути побудована так, щоб під час виклику не було втрати енергії викличного струму в розмовній частині, а під час розмови – втрати енергії розмовного струму у викличній частині схеми. Це досягається застосуванням схеми з механічним і електричним розділенням викличних та розмовних ланцюгів.

Схемою з механічним розділенням ланцюгів називають таку схему ТА, в якій викличні пристрої відділяються від розмовних за допомогою механічного перемикача. Як такий перемикач можна використати важільний перемикач (ВП), який керується масою мікротелефонної трубки (слухавки), або натискний контакт, розміщений у слухавці (мікротелефонній трубці), і називається розмовним клапаном (РК).

Розглянемо дію механічного розділення ланцюгів схеми ТА (рис. 4.1). У спокійному стані схеми, в якому мікротелефон лежить на важелі, у лінію ввімкнено тільки викличні пристрої (дзвінок та індуктор). Розмовні прилади вмикаються, коли знімають мікротелефонну трубку.

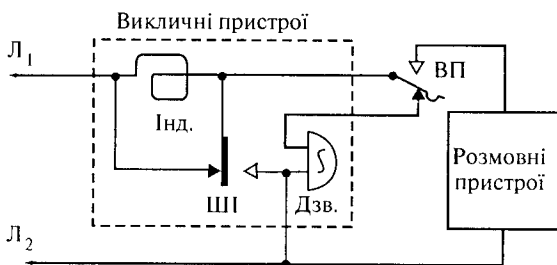


Рис. 4.1. Механічне розділення викличного та розмовного ланцюгів ТА

У схемі викличних пристроїв, у свою чергу, розрізняють викличні ланцюги вхідного та вихідного з'єднань.

У спокійному стані в лінію ввімкнено тільки дзвінок Дзв., а обмотка індуктора Інд. закорочена пружиною шунта індуктора (ШІ). Під час обертання рукоятки індуктора ці контакти перемикаються. Цією дією відмикається викличний прилад Дзв., а вмикається тільки індуктор.

Схемою з електричним розділенням ланцюгів називають таку схему ТА, у якій викличні пристрої відділяються від розмовних за допомогою конденсатора, який пропускає струм розмовних частот, але не пропускає струм викличної частоти.

Принцип електричного розділення ланцюгів показано на рис. 4.2. Ємність розподільного конденсатора C_p беруть зазвичай в границях 0,2...0,5 мкф.

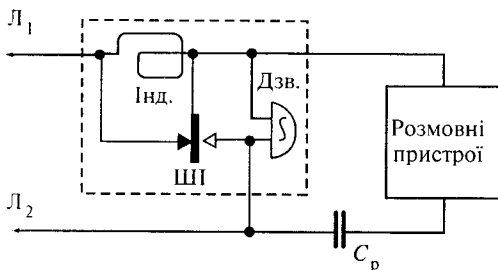


Рис. 4.2. Електрична схема розділення викличних та розмовних ланцюгів у схемі ТА МБ

Для ТА системи МБ нині переважно використовують електричне розділення викличних і розмовних ланцюгів, оскільки апарати з такою схемою завжди готові як до приймання виклику, так і до розмови; вони більш надійні, бо мають меншу кількість перемикальних контактів. Їх недолік – неминуча втрата енергії в схемі апарата.

Механічне розділення викличного і розмовного ланцюгів використовують для ТА системи ЦБ (ЦБ – АТС), для якої ємнісне розділення принципово непридатне.

Спосіб вмикання розмовних приладів. Залежно від способу вмикання розмовних приладів розрізняють постійні та змінні схеми ТА.

Постійною схемою називають таку схему ТА, в якій постійно вимкнуті під час розмови як телефон, так і мікрофон (рис. 4.3).

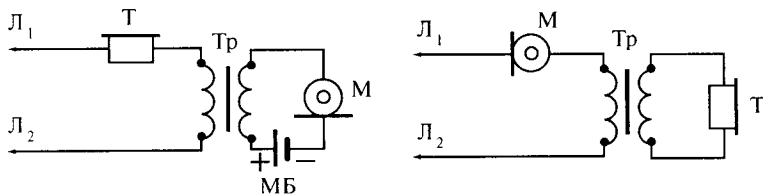


Рис. 4.3. Постійна схема вмикання розмовних приладів: а – МБ, б – ЦБ

Змінна схема – це така схема ТА, в якій вмикається мікрофон тільки під час передавання, а телефон – тільки під час приймання (рис. 4.4).

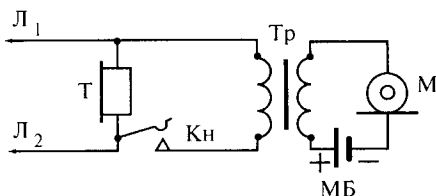


Рис. 4.4. Змінна схема вмикання розмовних приладів у ТА МБ

Змінну схему використовують тільки для схеми ТА МБ.

Нормально змінна схема знаходиться в стані приймання – у лінію ввімкнено телефон. З натисканням кнопки Кн телефон відмикається від лінії, а в лінію через трансформатор вмикається мікрофон. Розмовляючи, абонент діє за принципом «розмовляю – натискую, слухаю – відпускаю».

Постійна схема найпростіша. Вона містить найменшу кількість елементів, але її ККД не більше 50 %, оскільки під час передавання частина потужності, яку розвиває мікрофон, втрачається на своєму телефоні, а під

час приймання потужність корисного розмовного сигналу (струму) втрачається на своєму мікрофоні.

Змінна схема усуває цей недолік і забезпечує як під час передавання, так і під час приймання 100 % ККД, але вона незручна в експлуатації, тому що доводиться виконувати перемикання з приймання на передавання і навпаки, що не дозволяє перервати розмову співрозмовника. Тому у всіх абонентських телефонах загального користування натепер використовують постійну схему в ускладненому вигляді порівняно зі схемою, показаною на рис. 4.3.

Місцевий ефект. Місцевим ефектом називають явище прослуховування своєї розмови в телефоні свого апарата. Постійні схеми ТА, в яких розмовний струм, генерований мікрофоном, у неослабленому вигляді проходить через телефон свого апарата і викликає такий ефект, називають відповідно схемами з місцевим ефектом. Прикладами таких схем можуть бути схеми, показані на рис. 4.5.

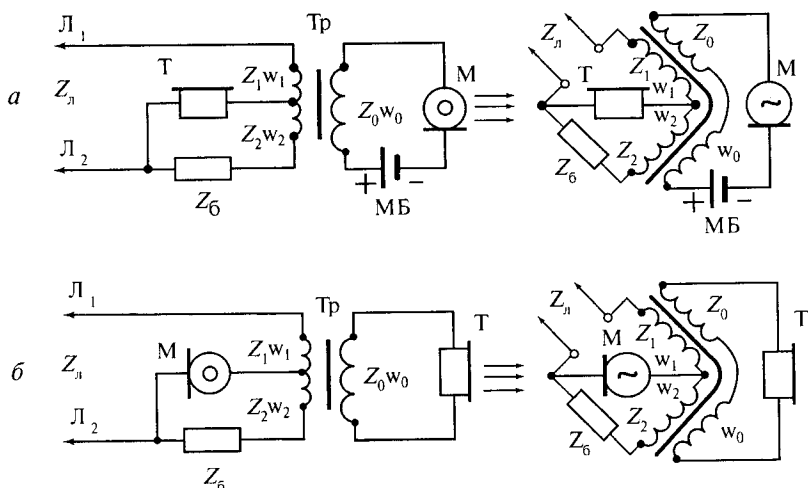


Рис. 4.5. Протимісцева мостова трансформаторна схема ТА: а – МБ; б – ЦБ

Схемам з місцевим ефектом властиві такі два основні недоліки.

1. Через властивість «адаптації звуку» під час слухання власної гучної розмови знижується чутливість слуху. Якщо слідом за цим телефон апарата збуджується слабким лінійним сигналом, тоді перші слова, а то й фрази, можуть бути загублені для сприйняття, оскільки вухо до слухання тихого сигналу пристосовується із затримкою.

2. Якщо розмова ведеться із шумного приміщення, то акустичний шум, який діє на свій мікрофон, потрапляє у вигляді електричного сигналу до свого телефону і викликає в усі маскування приймальної мови. Це означає, що схема з місцевим ефектом має низьку шумозахищеність.

Обидва недоліки схем з місцевим ефектом знижують практичну довжину телефонування.

Щоб уникнути цих недоліків, схеми сучасних ТА ускладнюють, що дозволяє звести місцевий ефект до мінімально можливого.

Протимісцевою схемою ТА називають таку схему, в якій у процесі передавання мови досягається значне послаблення струму, який проходить через свій телефон, порівняльно зі струмом, який поступає в лінію.

Використовують два різновиди протимісцевих схем:

- мостову,
- компенсаційну.

Мостова протимісцева схема – це така схема розмовної частини ТА, елементи якої разом з лінією утворюють схему містка змінного струму, в одну з діагоналей якого вмикається мікрофон (або напруга, яку розвиває мікрофон на виході мікрофонного трансформатора чи підсилювача), а в другу – телефон.

Канонічний вигляд мостових протимісцевих схем ТА, систем МБ показано на рис. 4.6.

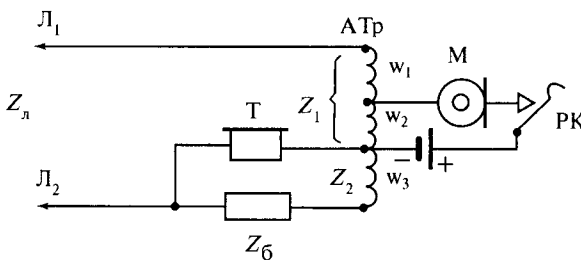


Рис. 4.6. Протимісцева мостова автотрансформаторна схема ТА МБ: Z_1, Z_2, Z_n, Z_6 – комплексний опір відповідно першої і другої півобмоток диференційного трансформатора, лінії зв'язку та балансового контуру

Відомо, що в збалансованій мостовій схемі добутки опорів протилежних плеч однакові. Звідси випливає, що опір балансного контуру має дорівнювати:

$$Z_6 = \frac{Z_2}{Z_1} Z_n. \quad (4.1)$$

Аналізуючи роботу мостової схеми в умовах ідеального балансу місцевого ефекту детальніше, не важко показати, що формула 4.1 набуває вигляду

$$Z_6 = \frac{w_2}{w_1} Z_n,$$

де w_1 і w_2 – кількість витків відповідно лінійної та балансової півобмоток.

Якщо диференціальний трансформатор симетричний $w_1 = w_2$, тоді опір балансового контуру має бути однаковий з вхідним опором лінії: $Z_n = Z_6$. Оскільки вхідний опір лінії має активну та ємнісну складові, то балансний контур має складатись з активного опору та ємності, причому:

$$R_6 - jX_6 = R_n - jX_n.$$

Компенсаційна протимісцева схема – це така схема розмовної частини ТА, за якої послаблення чутності своєї розмови досягається тим, що під час передавання мови в контурі, в який ввімкнено телефон, виникає взаємна компенсація напруги мікрофонного струму та наведеної ним ЕРС, що виявляються в протифазі.

Варіантів компенсаційних протимісцевих схем багато. Найбільшого поширення набула схема, яку використовують у системі ЦБ (АТС) (рис. 4.7).

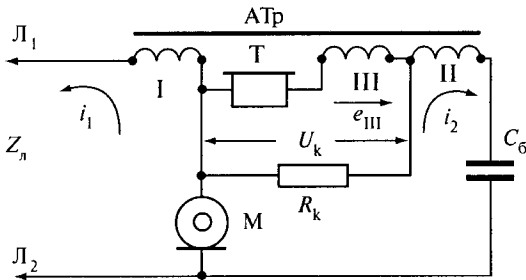


Рис. 4.7. Варіант протимісцевої компенсаційної автотрансформаторної схеми ТА ЦБ

4.2. Основні елементи схем телефонних апаратів

Основні дві частини схем ТА, як відомо, – розмовний та викличний вузли. До складу розмовного вузла входять мікрофони і телефони, а також схемні рішення з розгалуження розмовних ланцюгів під час передавання за допомогою мікрофона мовних сигналів і під час приймання за допомогою телефону.

У ТА АТС до складу викличного вузла входять номеронабирач (НН) – давач імпульсів для керування з'єднанням на АТС, і дзвінок змінного струму – приймач вхідного виклику. У ТА, які вмикаються в телефонні станції системи ТА ЦБ, до викличної частини належить приймач вхідного виклику – дзвінок. У ТА МБ до викличного вузла входять індуктор для формування вихідного виклику та дзвінок змінного струму для сприймання вхідного виклику.

Розглянемо стисло основні технічні характеристики цих приладів.

4.2.1. Номеронабирач

Номеронабирач являє собою завідний контактний механізм, за допомогою якого під час заведення виконується попереднє перемикання в схемі ТА, а під час зворотного обертання в початкове положення в лінію передається набірна комбінація, яка складається з періодичних замикань та розмикань лінійного шлейфа.

Електричну схему найпростішого НН зображено на рис. 4.8.

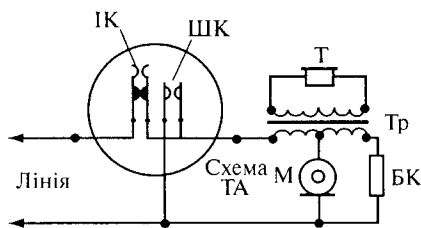


Рис. 4.8. Електрична схема найпростішого НН

Контактна система НН має чотири пружини, які утворюють два контакти: імпульсний – ІК і шунтувальний ШК. Ці контакти у разі набирання номера працюють так.

Під час заведення диском ШК замикається і накоротко замикає схему розмовної частини (майже відмикає) ТА. Це робиться для того, щоб збільшити значення струму в набірній частині відключенням від неї приладів розмовної схеми, а також, щоб у наступних періодичних розмиканнях і замиканнях шлейфа в телефоні, який під час набирання номера притиснутий до вуха абонента, не було клацання внаслідок збільшення чи зменшення струму в обмотці трансформатора.

Під час зворотного ходу диска ШК залишається замкнутий, а ІК розмикається і замикається стільки разів, скільки цифр набрав абонент. Кількість розмикань визначає цифру, яку приймають прилади АТС. Коли диск повертається в початкове положення, ІК залишається замкнутим, а ШК розмикається і заново вмикає схему ТА в лінію.

Такій схемі НН властиві два недоліки.

Перший недолік полягає в тому, що ШК запобігає клацанню в ТА тільки у разі зворотного повертання диска, тобто під час роботи ІК, але два клацання все таки не виключаються. Під час заведення диска замикання ШК зумовлює зниження струму в первинній обмотці трансформатора Тр, внаслідок цього виникає імпульс струму у вторинній його обмотці, який діє на телефон. Коли диск повертається в стан спокою, розмикання ШК викликає імпульс струму, який проходить через телефон. Отже, клацання в телефоні відбувається на початку та в кінці набирання абонентом кожної цифри.

Другий недолік полягає в тому, що через невелику відстань від першого отвору на набірному диску, який відповідає цифрі «1», може статися, що у разі швидкого набирання номера, який має цифру «1», проміжок часу між двома цифрами виявиться дуже малим для надійної роботи приладів АТС.

Ці недоліки усуваються ускладненням контактної системи НН. Схему семипружинного НН показано на рис. 4.9.

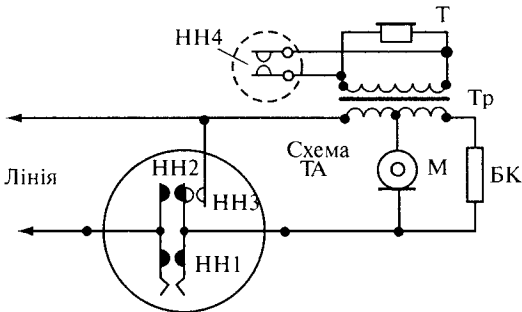


Рис. 4.9. Електрична схема типового НН

Контактна система в цьому НН працює в такій послідовності.

Під час заведення диска першим замикається контакт НН4, який шунтує телефон. При цьому ніякі зміни електричного режиму лінійного ланцюга вже не зможуть справляти на нього які-небудь впливи (діяння). Слідом за контактом НН4 перемикається контактна група НН2–НН3. Розімкнутий контакт НН2 підготовлює роботу імпульсного контакту НН1, а замкнутий контакт НН3 забезпечує коротке замикання лінійного шлейфа, який шунтує схема ТА.

Під час зворотного руху диска пульсує імпульсний контакт НН1.

Механізм, який керує пульсацією контакту НН1, побудований таким чином, що кількість його розмикань завжди на одне розмикання більша,

ніж набирає цифра, але наприкінці обертання диску, після того, як виконається передостаннє розімкнення імпульсного контакту, відбувається зворотнє перемикання групи НН2–НН3, контакт НН2 знову замикає на коротко імпульсний контакт НН1, що забезпечує гарантоване подовження міжсерійного часу на величину тривалості одного імпульсу, тобто на 10 мс.

Перехід у початкове положення контакту НН3 вмикає в лінію розмовну частину схеми ТА.

В останній момент, коли диск НН і контакт НН4, який знімає коротке замикання з телефону, але струм в лінійній обмотці трансформатора Тр до цього часу вже установився, то ніякого імпульсу струму в ланцюзі телефону не буде.

Контактною системою керує система із шести зубчастих колес і відповідних пружин.

Основні експлуатаційно-технічні характеристики НН – частота викличних імпульсів, імпульсний коефіцієнт і строк служби. Перші дві з них пояснює рис. 4.10.

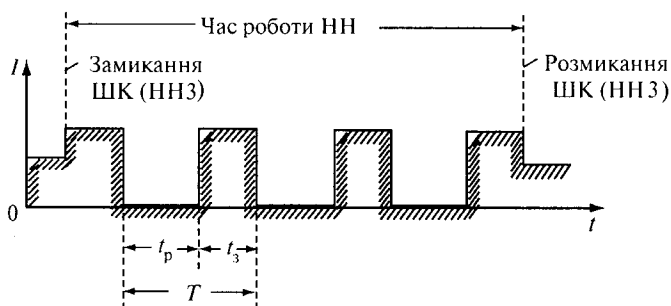


Рис. 4.10. Діаграма струму в лінії під час набірвання номера

Тут зображено діаграму струму в лінії під час набірвання абонентом цифри «3», що відповідає трьом розмиканням лінійного шлейфа. Тривалість одного періоду набірної комбінації T , що складається з часу розмикання t_p і часу замикання t_3 , беруть рівною 100 мс відповідно; за секунду НН може виконати 10 розмикань. Звідси частина викличних імпульсів НН $f_i = 10 \pm 1$ імпл./с. У границях цих допусків забезпечується нормальна робота АТС.

Тривалості замикання та розмикання не однакові. Для нормального НН вітчизняного виробництва $t_p = 62,5$ мс; $t_3 = 37,5$ мс.

Імпульсний коефіцієнт — відношення часу розмикання до часу замикавання:

$$K_i = \frac{t_p}{t_3}$$

Для зведених величин $K_i = 1,66$. Допустимі границі зміни $K_i = 1,4 \dots 1,8$. Перевищення часу розмикання над часом замикавання шлейфа відбувається тому, що основна робота комутаційних приладів АТС виконується під час розмикання контактів НН.

Строк експлуатації сучасних НН вимірюють кількістю робочих циклів без додаткового регулювання і становить 500 000 наборів.

4.2.2. Індуктор

Індуктор використовують для посилення викликів у ТА системи МБ; він являє собою малогабаритний генератор змінного струму, який приводиться в дію від руки. Принцип улаштування індуктора, застосований в польових ТА, показано на рис. 4.11.

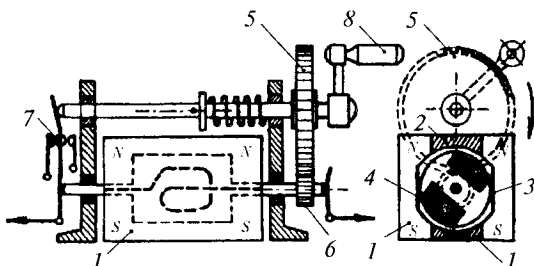


Рис. 4.11. Принцип улаштування індуктора:

1 – два прямокутні магніти; 2 – полюсні надставки; 3 – якорь; 4 – обмотки якоря;
5 і 6 – зубчасті колеса з передатним числом 1 : 6

Оскільки кількість обертів якоря в 6 разів більша від кількості обертів рукоятки, у витках обмотки, які перетинають силові лінії магнітного поля, виникає ЕРС із частотою 15...25 Гц залежно від швидкості обертання. Напруга, яку розвиває індуктор, може досягати 100 В.

4.2.3. Поляризований дзвінок змінного струму

Приймачем викликів у ТА всіх систем служить дзвінок змінного струму. Розглянемо найбільш використовувану спрощену конструкцію поляризованого дзвінка (рис. 4.12).

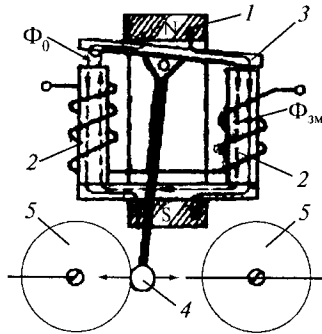


Рис. 4.12. Принцип устаткування поляризованого дзвінка змінного струму:
 1 – постійний магніт; 2 – осердя; 3 – якір; 4 – бойок; 5 – дзвінкові чашки;
 Φ_0 – постійний магнітний потік; $\Phi_{3м}$ – змінний магнітний потік,
 що періодично змінює свій напрям

На одній половині (рис. 4.12, праве плече) збігаються Φ_0 і $\Phi_{3м}$; загальний потік буде дорівнювати сумі $\Phi_n = \Phi_0 + \Phi_{3м}$, а в лівому плечі $\Phi_l = \Phi_0 - \Phi_{3м}$.

Якщо змінюється фаза викличного струму, то змінюється напрям $\Phi_{3м}$, тоді в лівому плечі $\Phi_l > \Phi_n$; притягання магніту в лівому плечі більше, ніж у правому і бойок стукає праву чашечку.

4.3. Підвищення зручності користування телефонним апаратом

У деяких типах ТА абонентам забезпечується можливість гучномовного приймання.

Тепер запроваджують ТА, в яких замість обертальних дискових НН використовують тастатурні НН, які бувають кнопковими або клавіатурними.

Зовнішній вигляд тастатури – кнопкове (клавішне) поле, на якому розміщено 10 кнопок (рис. 4.13).

Під час короткого часу натискання на будь-яку кнопку на АТС передається сигнал, який присвоюється певній цифрі (кодування).

Перевага тастатурного набирання полягає в тому, що цей спосіб значно скорочує витрати часу на процедуру набирання номера (шестизначний номер за допомогою дискового НН передається в середньому за 10 с, тоді як для тастатурного набирання для цього достатньо 4...5 с) та зменшує зусилля, які витрачає абонент на процес набирання. Остання обставина дуже важлива особливо для операторів, які за родом своєї діяльності виконують багатократні виклики. Разом з тим тастатурний апарат, безперечно, значно складніший за технічними схемними рішеннями, ніж дисковий.

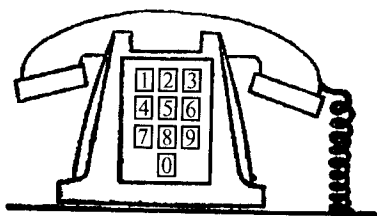


Рис. 4.13. Схема розміщення кнопок тастатурного поля

За час (40 років) впровадження ТА з тастатурним набором у світі розроблено дуже багато патентів на різноманітні варіанти тастатурних НН, у яких застосовують різні принципи кодування набраних номерів. Кожна фірма чи підприємство, які виготовляють ТА з тастатурним набором, видають свою продукцію як «ноу-хау».

Розглянемо в загальному вигляді принцип дії тастатурного НН з частотною системою кодування, яку вважають найбільш перспективною. Функціональну схему тастатурного НН цього типу показано на рис. 4.14.

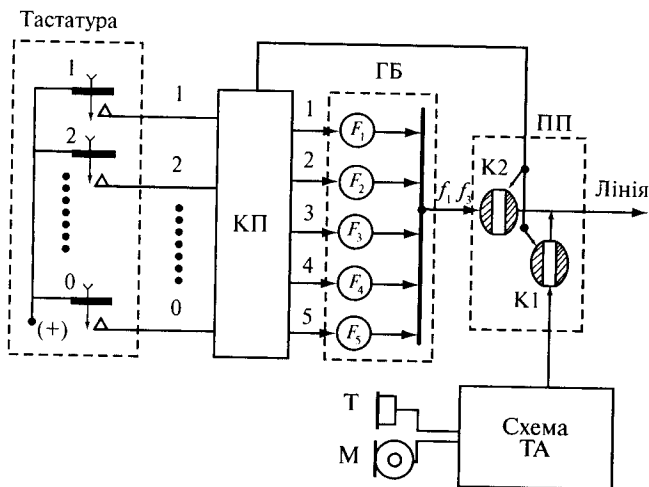


Рис. 4.14. Функціональна схема тастатурного НН з частотним кодуванням номерів

Кожній із 10 цифр тут присвоєно певну комбінацію з тональних частот, які одночасно посилаються в лінію. Дуже часто знаходить використання міжнародний частотний код «2 із 5», де будь-яка цифра передається поєднанням двох частот із п'яти можливих. Використовуються частоти: $f_1 = 700$ Гц, $f_2 = 900$ Гц, $f_3 = 1100$ Гц, $f_4 = 1300$ Гц, $f_5 = 1500$ Гц. Кодування виконується згідно з частотами, які наведено в табл. 4.1.

Таблиця 4.1

| Цифра | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| Поєднання частот | f_1f_2 | f_1f_3 | f_2f_3 | f_1f_4 | f_2f_4 | f_3f_4 | f_1f_5 | f_2f_5 | f_3f_5 | f_4f_5 |

На АТС кодова комбінація розшифровується та реєструється як відповідна цифра. Схема тастатурного НН містить такі складові:

- тастатурний кнопковий пристрій (тастатура);
- кодувальний пристрій (КП);
- блок генераторів (БГ), який містить 5 тональних генераторів частот від f_1 до f_5 ;
- перемикальний пристрій (ПП), який за допомогою двох клапанів К1 і К2 вмикає в лінію схему ТА або НН.

У спокійному стані схеми клапан К1 відкритий і з'єднує схему ТА з лінією, а клапан К2 закритий, унаслідок чого НН відімкнений від лінії.

Під час набирання тієї чи тієї цифри номера короткочасно натискується відповідна кнопка тастатури. Нехай це буде кнопка «2». Тоді по проводу «2» в КП подається сигнал, у результаті дії якого КП згідно з кодовою таблицею збуджує в генераторному блоці ГБ генератори F_1 і F_3 . На вихідних шинах БГ з'явиться напруга частот f_1 і f_3 . Одночасно КП видає в ПП сигнал тривалістю 40...50 мс, який закриває клапан К1, тим самим відмикаючи ТА від лінії і відкриваючи клапан К2. Це робиться для того, щоб тривалість посилок була однаковою і не залежала від того, скільки часу абонент буде тримати натиснутою тастатурну кнопку. За час відкритого положення К2 в лінію посилається кодовий імпульс частот f_1f_3 , який відповідає цифрі «2». Після закінчення терміну часу на посилення частотного імпульсу схема набуває початкового стану.

У разі дискових НН для передавання однієї цифри (від 1 до 10 імпульсів) витрачається близько 550 мс, тобто в 11 разів більше.

На сьогодні тастатурні апарати поступово витісняють дискові.

Подальший розвиток у напрямі автоматичного набирання номера – використання автовідповідача.

Ще більшу економію часу, який витрачається на набирання номера, забезпечують ТА з автоматичним або скороченим набиранням.

Принцип дії такого апарата полягає в тому, що за допомогою спеціального програмувального пристрою завчасно підготовлюються набірні комбінації для виклику найбільш частих необхідних абонентів, а функції виклику потрібного запрограмованого абонента зводяться лише до натискання відповідної кнопки. При цьому в схемі автоматичного набирання формується заздалегідь передбачена комбінація імпульсів, які автоматично поступають у лінію.

Автоматичне набирання не тільки економить час набирання, але й запобігає можливості помилок, які майже неминучі у разі набирання багатозначних номерів.

Подальший крок у напрямі розширення послуг абонента – це використання автовідповідача, який працює сумісно з ТА для записування повідомлень, які поступають від абонентів.

Залишаючи своє робоче місце, абонент переводить лінію зі звичайного ТА на автовідповідач. Якщо тепер по цій лінії поступає виклик, тоді він буде прийнятий автовідповідачем, причому абонент, який викликає, почує задалегідь записану фразу: «Абонента немає на місці. Для передачі йому вашого повідомлення вам надається така можливість».

Слідом за цим в лінію вмикається магнітофон у режимі «Запис». Повернувшись на місце, абонент прослуховує записане для нього повідомлення. Перемикання з ТА на автовідповідач може виконуватись і автоматично після визначеного числа без відповідних посилань виклику змінним струмом.

Таким чином, закінчено розгляд терміналу – телефонного апарата, його складових елементів та вузлів, їх роботу в загальному процесі виклику і обміну повідомленнями між абонентами, що спілкуються по телефонних мережах зв'язку.

4.4. Параметри і характеристики телефонних апаратів

Результуючу оцінку якості ТА визначають здебільшого за розбірливістю, яка забезпечується ТА для заданого тракту за певних акустичних умов.

Для того, щоб забезпечити виконання вимог до розбірливості, електроакустичні й електричні параметри ТА мають задовольняти задані норми.

4.4.1. Електроакустичні параметри

Основні електроакустичні параметри ТА:

- коефіцієнт приймання-передавання для розмовної частини апарата;
- рівень гучності викличного акустичного сигналу для викличної частини апарата.

Для апаратів загального користування основними контрольованими електроакустичними параметрами будуть еквіваленти затухання за гучністю:

$E_{з,пер}$ – еквівалент затухання передачі;

$E_{з,пр}$ – еквівалент затухання прийому;

$E_{з,м,е}$ – еквівалент затухання місцевого ефекту.

Для розуміння поняття про коефіцієнти приймання-передавання розглянемо схему рис. 4.15.

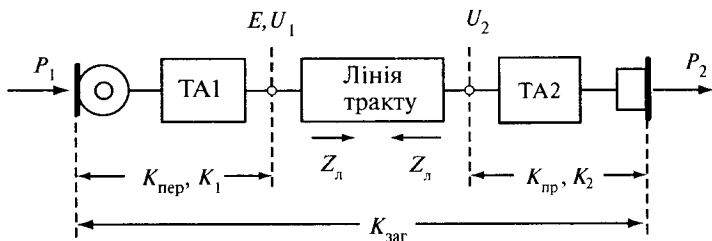


Рис. 4.15. Визначення коефіцієнтів приймання-передавання ТА

Передавальний телефонний апарат ТА1 при звуковому тиску P_1 визначеної частоти утворює на розімкнених затискачах ЕРС E , а при вмиканні на вхідний опір Z_L – напругу U_1 . У таких умовах вмикання приймального апарата ТА2 за напруги на його затискачах U_2 розвиває звуковий тиск P_2 .

Коефіцієнтом передачі (або чутливістю) ТА називають відношення ЕРС на розімкнутих лінійних затискачах апарата до звукового тиску заданої частоти на мікрофоні ($\text{мВ}/(\text{Нп}/\text{м}^2)$):

$$K_{пер} = \frac{E}{P_1}.$$

Коефіцієнтом передачі ТА за напругою називають відношення напруги, яку розвиває ТА на заданий вхідний опір, до звукового тиску заданої частоти на мікрофоні ($\text{мВ}/(\text{Нп}/\text{м}^2)$):

$$K_1 = \frac{U_1}{P_1}.$$

Коефіцієнтом приймання ТА називають відношення звукового тиску, який розвиває телефон апарата, до напруги частоти на лінійних затискачах ($\text{мВ}/(\text{Нп}/\text{м}^2)$):

$$K_{пр} = \frac{P_2}{U_2}.$$

Оскільки відношення P_2/U_2 не залежить від вхідного опору лінії, до якої підключено ТА, то коефіцієнт приймання в будь-яких дійсних умовах вмикання дорівнює K_2 .

Загальним коефіцієнтом приймання-передавання називають відношення звукового тиску, розвинутого на виході ТА (приймача), до звукового тиску частоти на мікрофоні ТА (передавача):

$$K_0 = \frac{P_2}{P_1}.$$

Рівень гучності акустичного сигналу, створюваного дзвінком апарата, вимірюється шумоміром на заданій відстані від апарата (зазвичай 50 см), напруга якого на його вході 50 В при частоті викличного сигналу 25 Гц.

4.4.2. Електричні параметри

Основні електричні параметри ТА визначаються робочими загасаннями:

- апарата на передавання;
- апарата на приймання;
- місцевого ефекту.

Робоче загасання апарата на передавання визначають як 10 десятичних логарифмів відношення потужності P_M , віддаваної мікрофоном на узгоджене навантаження, до потужності $P_{л1}$, віддаваної мікрофоном вхідному опору лінії $Z_{л}$ через схему ТА (дБ):

$$a_{\text{пер}} = 10 \cdot \lg \frac{P_M}{P_{л1}}.$$

Робоче загасання схеми ТА на приймання визначають як 10 десятичних логарифмів відношення потужності $P_{л2}$, віддаваної лінією на узгоджене з нею навантаження $Z_{л}$, до потужності P_T , віддаваної тією ж лінією телефона через схему ТА (дБ):

$$a_{\text{пр}} = 10 \cdot \lg \frac{P_{л2}}{P_T}.$$

Очевидно, що чим менші значення $a_{\text{пер}}$ і $a_{\text{пр}}$, тим менші будуть втрати в електричній частині схеми ТА.

Робоче загасання місцевого ефекту ТА визначають як 10 десятичних логарифмів відношення потужності P_M , віддаваної мікрофоном на узгоджене з ним навантаження, до потужності P_T , яку він віддає телефону свого ТА через його схему при навантаженні лінійних затискувачів апарата на опір $Z_{л}$ (дБ):

$$a_{\text{м.е}} = 10 \cdot \lg \frac{P_M}{P_T}.$$

Придушення місцевого ефекту буде тим більше, чим більше значення $a_{\text{м.е}}$.

Для виконання розрахунків інколи зручно користуватись другою характеристикою протимісцевих властивостей схеми ТА – загасанням протимісцевості $a_{м.е}$, яке визначають за формулою (дБ):

$$a_{м.е} = 20 \cdot \lg \frac{U_{м}}{U_{т}}$$

де $U_{м}$ і $U_{т}$ – напруга відповідно на вихідних затискачах мікрофона і вхідних затискачах телефону свого ТА, який навантажено на лінію з вхідним опором $Z_{л}$.

Типовий вигляд частотних залежностей величин $a_{пер}$, $a_{пр}$, $a_{ме}$ для сучасних ТА з вугільними мікрофонами показано на рис. 4.16.

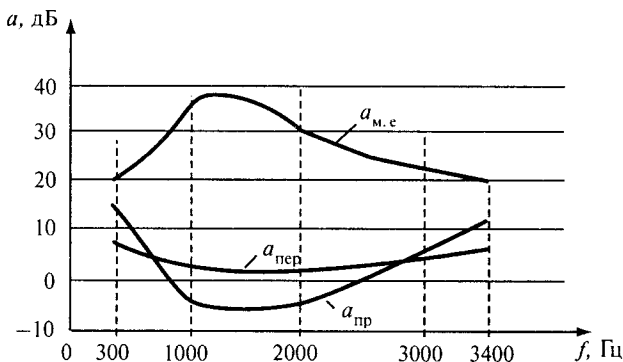


Рис. 4.16. Частотні залежності електричних параметрів ТА

Визначення значень $a_{пер}$, $a_{пр}$ та $a_{м.е}$ експериментальним шляхом дозволяє оцінювати якість роботи випробуваних ТА. Методи оцінювання якості передавання мовних звукових сигналів розглянуто у п'ятому і шостому розділах.

5. Оцінювання якості передавання мови

5.1. Загальні положення. Вимоги до якості передавання

Якість передавання мови – одна з найважливіших експлуатаційних характеристик будь-якої телефонної та радіотелефонної апаратури. Тому для виробництва і експлуатації цієї апаратури потрібно мати методи кількісної оцінки її якості.

Одна з можливих підходів до вирішення цього завдання – оцінювання якості переговорної апаратури аналізуванням вимірювальних електроакустичних і електричних характеристик (чутливість мікрофонів і телефонів, частотні характеристики загасання та підсилення). Якщо всі ці характеристики відповідають установленим нормам, то можна вважати, що і якість зв'язку за допомогою цієї апаратури буде задовільною. Для деяких випадків такого типу оцінка цілком достатня. Але вона не пряма і не дає можливості відповісти на запитання: «А яка буде якість за тієї чи тієї довжини лінії або за того чи того рівня шумової завади?»

Очевидно, що якщо всі характеристики одного апарата кращі за характеристики другого, то і якість передавання мови першим буде вищою (але невідомо, на скільки).

Якщо деякі характеристики одного апарата кращі за такі самі характеристики другого, але інші характеристики другого кращі ніж першого, в такому разі взагалі нічого певного не можна сказати у разі їх порівняння. Тому треба якість передавання мови оцінювати в числових одиницях. Спочатку отримали поширення суб'єктивні методи. Так, наприклад, якість ТА можна характеризувати величиною загасання між передавальним і приймальним апаратами, за якого оператор, що приймає, оцінює чутність переданого йому тексту як прийнятій. Це затухання отримало назву «перекриваним затуханням». Дотепер ще діють норми «перекриваного затухання» в технічних завданнях на технічні пристрої.

Така оцінка не задовольняє запити виробництва та експлуатації, оскільки суб'єктивна оцінка значною мірою залежить від думки особи, яка виконує випробування; не дозволяє отримати безперечну цифру, яка могла б бути відтворена другим випробувачем. Тому бажано було б мати показники якості передавання мови, які були б числовими і носили об'єктивний характер не залежно від думки та стану випробувача.

Складність установаження таких показників для оцінювання якості передавання мови зумовлено різними чинниками, які впливають на якість, а також тим, що кінцевими ланками мовного тракту є не технічні, а біологічні елементи: голосовий апарат мовця (розмовляючого) і слуховий апарат абонента, що приймає мову.

Відомо, що характеристики мови і слуху різних людей не тільки відрізняються, але й різні для однієї і тієї самої особи; вони значно змінюються залежно від зовнішніх умов.

Тому для отримання об'єктивних оцінок якості передавання мови широко застосовують статистичні методи, основані на усередненні великої кількості одиночних оцінок.

Різноманітність причин, від яких залежить якість передавання мови, а також різні вимоги до неї в різних умовах експлуатації зумовили появу різних методів її оцінки.

В основі визначення якості передавання мови лежить оцінка ступеня виконання вимоги до сприйняття прийманої мови.

Узагальнення вимог до якості передавання мови у всіх можливих випадках практики можна звести до таких трьох:

- 1) розбірливість (зрозумілість, ясність);
- 2) гучність (голосність);
- 3) природність (натуральність).

Розбірливість – першу вимогу, безперечно, можна назвати основною, оскільки вона відображає виконання системою приймання-передавання мови свого головного призначення – забезпечити, щоб той, хто слухає, правильно зрозумів зміст того, що йому передається. Для цього мова, яка приймається, повинна мати достатній ступінь розбірливості.

Гучність – визначає бажаний рівень прийнятих сигналів, який для оптимальних умов має бути таким, щоб не викликати стомленості напруги слухового апарата.

Природність (натуральність) – оцінює здатність системи відтворювати не тільки зміст мови, яка передається, але і її індивідуальне «забарвлення» (тембр), притаманне різним мовцям.

Для телефонної та радіотелефонної апаратури загального користування ця вимога другорядна. Але в спеціальних системах зв'язку, наприклад у системах, у яких потрібне впізнавання того, хто розмовляє, по голосу, або для художнього відтворення мови та музики.

Розглянемо об'єктивні методи оцінювання якості передавання мови, які мають в основі оцінку ступеня виконання розглянутих вимог.

5.2. Метод порівняння за гучністю

В основу методу покладено порівняння гучності передавання випробуваним трактом з гучністю такого самого передавання по еталонному тракту, характеристики якого стандартизовані.

Різниця між гучністю у разі передавання через еталонний тракт і гучністю у разі передавання через випробуваний тракт виражається в децибелах і називається *еквівалентом затухання (ЕЗ) за гучністю* випробуваної системи і служить показником її якості.

Вимірювання ЕЗ за гучністю засновано на здатності слухового апарата людини з великою точністю оцінювати однаковість двох гучностей (похибка такого порівняння не перевищує 2...3 дБ).

Схему вимірювання показано на рис. 5.1.

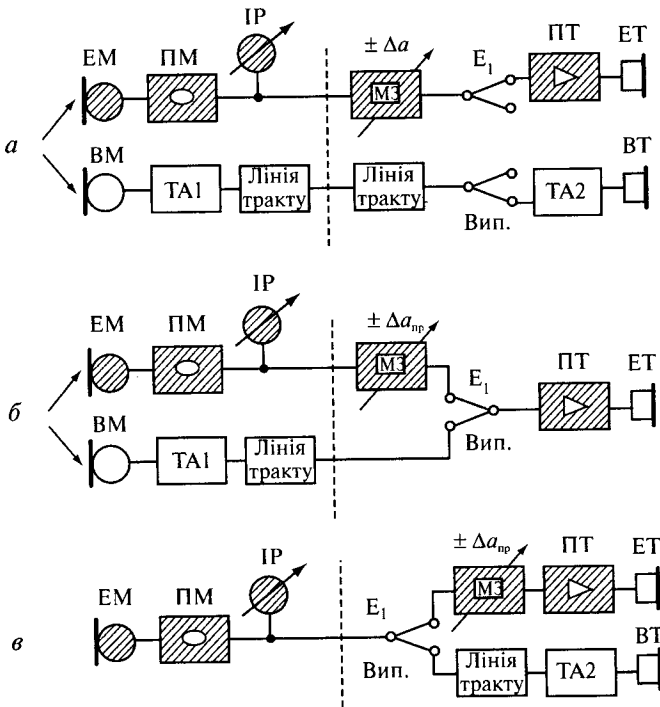


Рис. 5.1. Вимірювання еквівалента затухання гучності:

ЕМ, ЕТ, ПМ, ЕТ – еталонні мікрофон, телефон та їх підсилювачі;
 ІР – індикатор рівня; МЗ – магазин затухань; ВМ, ВТ, ТА1, ТА2 – випробувані мікрофон, телефон та їх телефонні апарати ТА1 і ТА2;

Вимір виконують у такій послідовності. Перед двома мікрофонами знаходиться оператор-розмовник, який багатократно читає умовну фразу, яку заздалегідь готують для певної мови для забезпечення середньої потужності мови (для російської мови фраза «не видали мы такого невода»).

Другий оператор знаходиться в приміщенні, звукоізолюваному від приміщення передачі, слухає поперемінно через еталонний і випробуваний тракту, вимикає та вмикає перемикач ЕТ – Випр.

Якщо в еталонному тракту чутно гучніше, тоді в магазин затухань (МЗ) додатково вводиться затухання ($+\Delta a$) доти, доки гучності не будуть вирівняні.

Якщо в еталонному тракту гучність гірша, тоді з МЗ виводиться затухання ($-\Delta a$) до забезпечення такої самої гучності, як у випробуваній системі. При цьому виявляється, що ЕЗ за гучністю буде дорівнювати цьому додатковому затуханню:

$$EЗ = L_e - L_v = \pm \Delta a,$$

де L_e, L_v – рівень гучності, який забезпечується еталонним і випробуваними трактами. Знак (+) свідчить про те, що гучність випробуваного тракту нижча ніж еталонного, а знак (-) – навпаки, еталонного нижча, ніж випробуваного.

Випробуваний мікрофон, ТА і лінійний тракт – це місцева лінія, до складу якої входять випробуваний ТА, лінія зв'язку з джерелом живлення АТС; такий самий склад має і приймальна частина випробуваної системи.

Передавальну та приймальну частини загального тракту вимірюють окремо за тією самою схемою.

Вимірювання виконує бригада з двох-трьох операторів з обчисленням середнього із шести (або більше) вимірів.

Згідно з нормами МККГТ значення загального еквівалента загасання за гучністю приймання-передавання ЕЗ для 97 % має не перевищувати 40 дБ, що приблизно дорівнює гучності мовців, які знаходяться на відстані 10 м один від одного.

Цей метод забезпечує отримання об'єктивних оцінок, які не залежать від думки оператора та його слухових чи мовних даних.

Недоліком цього методу є те, що він не враховує розбірливість, яку забезпечує випробуваний тракт. Не важко уявити випадки, коли гучність висока, але мова зовсім нерозбірлива.

Цим методом можна користуватись для порівняння ТА, у яких інші характеристики однакові.

Останніми роками значно удосконалено вимірювальну техніку в напрямі прискорення вимірювання та підвищення їх точності.

5.3. Метод артикуляції

В основі методу лежить оцінка ступеня виконання головної вимоги до якості передавання мови – її розбірливості.

Термін «артикуляція» в лінгвістиці означає роботу мовних органів людини під час вимовляння звуків мови. У зв'язку з цим метод оцінювання якості передавання мови, в основі якого лежить аналіз «артикуляції» спеціально складених випробуваних таблиць, який дозволяє визначати розбірливість елементів, що входять у випробуваний текст, називають «методом артикуляції».

Мірою якості передавання мови в цьому методі є розбірливість елементів мови, яку можна виміряти, або обчислити. У цьому розумінні розбірливість мови – це відношення кількості правильно прийнятих по випробувальному тракту елементів мови (звуків, складів, слів або фраз) до достатньо великої загальної кількості переданих елементів, виражених у відсотках.

Якщо, наприклад, передано 10 складів, з них 7 правильно, тоді відношення буде $\frac{7}{10}$, а якщо передати два десятки складів, то буде більша кількість правильно прийнятих, оскільки цифру не можна визнати за розбірливість. Тому, якщо збільшувати кількість передаваних елементів, то можна помітити, що це відношення прямує до деякої постійної величини S , а n_0 – загальна кількість передаваних елементів, n_a – кількість правильно прийнятих елементів.

Такий процес вимірювання розбірливості означає, що властивість розбірливості можна визначати як імовірність правильного приймання елементів мови по цьому тракту.

Залежно від того, які елементи мови передаються, розрізняють такі види розбірливості:

- D – розбірливість звуків;
- S – розбірливість складів;
- W – розбірливість слів;
- J – розбірливість фраз.

Процес вимірювання розбірливості будь-якого виду (рис. 5.2) полягає в тому, що передавальний та приймальний кінці випробуваного тракту розміщуються у звукоізолюваних приміщеннях, у яких створено нормальні для цієї апаратури акустичні умови (рівні шумів), після чого оператор (диктор), що передає, читає спеціально складену таблицю складів, слів або фраз (звукову розбірливість визначають за тими самими складовими таблицями), а оператор, що приймає, записує їх на призначеному для цього бланку.

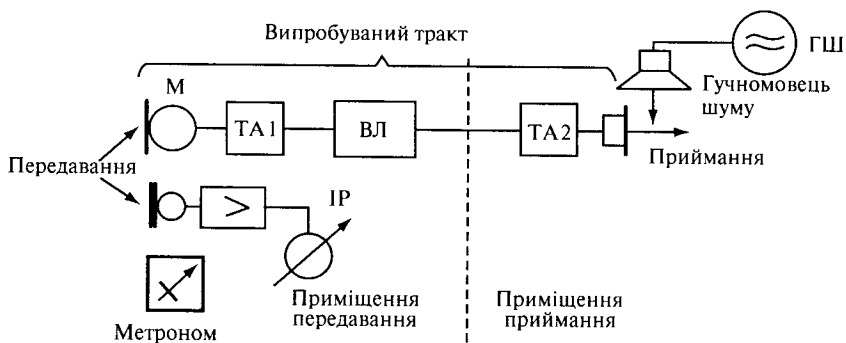


Рис. 5.2. Схема вимірювання розбірливості методом артикуляції

Після закінчення циклу приймання-передавання, який забезпечує достатній обсяг вимірів, прийняті таблиці зв'язуються з переданими і обчислюються відсотки правильно прийнятих елементів.

Основний вид розбірливості, за яким мають установлювати норми якості для більшості видів апаратури, – складова розбірливість S .

Під час вимірювань складової розбірливості передаються звукосполучення, які не мають змістового значення, тому освіта і розвиток операторів не будуть мати значення.

Виміряти розбірливість слів і фраз складніше, тим більше, що тут можлива догадка.

Крім зазначених видів розбірливості (D , S , W , J), які безпосередньо можна виміряти, у техніці передавання мови використовується поняття розбірливості формант A , яку не можна безпосередньо виміряти, але є можливість обчислити, оскільки звукам мови властивий певний формантний склад.

Усі види розбірливості (A , D , S , W , J) для кожної мови, що пов'язані між собою однозначною взаємозалежністю, з одного боку, відображають статистичні пропорції, що властиві фонетичній будові певної мови, з другого, – визначаються закономірностями діяльності свідомості для розпізнавання звукових комплексів.

Приблизний вигляд взаємозалежності між різними видами розбірливості показано на рис. 5.3–5.6 та відповідних таблицях (табл. 5.1–5.4).

Таблиця 5.1

| <i>A</i> | <i>S</i> , % |
|----------|--------------|
| 0,05 | 5,0 |
| 0,10 | 15,0 |
| 0,15 | 26,0 |
| 0,20 | 36,0 |
| 0,25 | 46,2 |
| 0,30 | 55,0 |
| 0,35 | 62,5 |
| 0,40 | 69,0 |
| 0,45 | 75,0 |
| 0,50 | 80,0 |
| 0,55 | 83,0 |
| 0,60 | 87,2 |
| 0,65 | 90,0 |
| 0,70 | 92,5 |
| 0,75 | 95,2 |
| 0,80 | 96,5 |
| 0,85 | 98,0 |
| 0,90 | 99,0 |
| 0,95 | 99,5 |
| 1,00 | 100,0 |

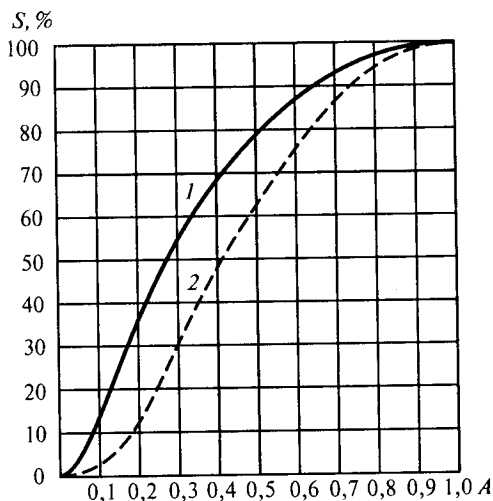


Рис. 5.3. Залежність складової розбірливості (*S*) від розбірливості формант (*A*):
1 – російська мова; 2 – англійська мова

Таблиця 5.2

| <i>D</i> , % | <i>S</i> , % |
|--------------|--------------|
| 10 | 0,08 |
| 20 | 0,70 |
| 30 | 2,30 |
| 40 | 5,00 |
| 50 | 11,20 |
| 60 | 20,00 |
| 70 | 32,20 |
| 80 | 49,10 |
| 90 | 71,70 |
| 100 | 100,00 |

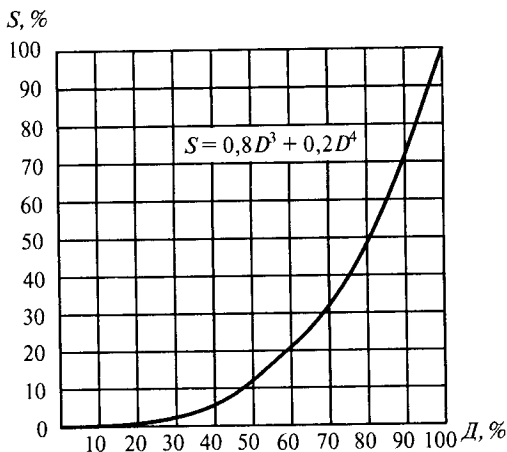
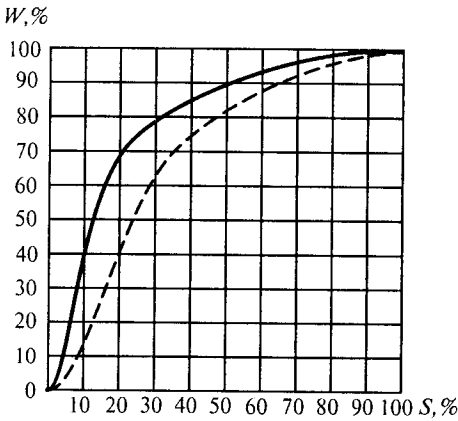


Рис. 5.4. Залежність складової розбірливості (*S*) від розбірливості звуків (*D*)

Таблиця 5.3

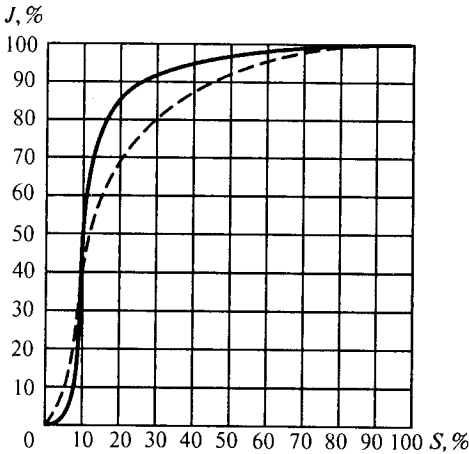
| $S, \%$ | $W, \%$ |
|---------|---------|
| 10 | 50,0 |
| 20 | 70,0 |
| 30 | 80,0 |
| 40 | 86,0 |
| 50 | 90,5 |
| 60 | 94,0 |
| 70 | 97,0 |
| 80 | 99,0 |
| 90 | 99,5 |
| 100 | 100,0 |

Рис. 5.5. Залежність словесної розбірливості (W) від розбірливості складів (S):

1 – російська мова; 2 – англійська мова

Таблиця 5.4

| $S, \%$ | $J, \%$ |
|---------|---------|
| 10 | 41,0 |
| 20 | 86,0 |
| 30 | 92,0 |
| 40 | 95,0 |
| 50 | 96,5 |
| 60 | 97,5 |
| 70 | 98,5 |
| 80 | 99,5 |
| 90 | 99,8 |
| 100 | 100,0 |

Рис. 5.6. Залежність фразової розбірливості (J) від розбірливості складів (S):

1 – російська мова; 2 – англійська мова

Для отримання цих залежностей потрібно створити складові артикуляційні таблиці для вимірювання розбірливості мови. Вони складаються за певними правилами. Ці правила враховують лінгвістичні (мовні) і технічні вимоги до таблиць. Лінгвістичні вимоги полягають у тому, щоб таблиці достатньою мірою відображали структуру мови. Технічні вимоги передбачають забезпечення максимальної економності під час виконання вимірів мінімальної надмірності (щоб не було звичних для слуху сполучень, які б сприяли відгадуванню і запам'ятовуванню випробуваних складів) і максимально можливою однорідністю для того, щоб зменшити розкид результатів одиночних вимірів.

Поєднання в таблицях лінгвістичних і технічних вимог можливе тільки у разі розумного компромісу, оскільки вони взаємно суперечливі.

Основний принцип створення складових таблиць, знайдений сумісними зусиллями зв'язківців і мовознавців, що забезпечує їх високі вимірювальні властивості, полягає в тому, що випробувані склади будуються за законами сполучення сусідніх звуків, але вони не повинні бути складами, по можливості, поширеними в реальних словах.

Для прикладу наведемо елементи запозиченої складової таблиці, як у підручнику М. Б. Покровського.

| | | | | | | |
|-------------|-------------|------------|------------|-------------|------------|------------|
| Ж | | | Ь | З | В | В |
| ДЛЕТ | СУРЬ | ПАЦ | ЛУШ | ЧТИС | НИФ | ВАФ |

Вимірювальний комплект має сто різних таблиць. Кожну таблицю побудовано згідно зі структурною формулою $0,8 \text{ ПГП} + 0,2 \text{ КГП}$, де велика літера П – приголосна, Г – голосна, К – консонанс (сполучення приголосних звуків).

Усі таблиці містять 50 складів і 160 звуків. Малими літерами показано додатковий варіант написання деяких складів.

Словесні таблиці зведено в комплект із 400 таблиць, у кожній – 50 слів.

Для вимірювання розбірливості комплектуються артикуляційні бригади з чотирьох – восьми операторів, що дозволяє усереднити вимірювальні параметри з урахуванням суб'єктивних властивостей операторів.

До складу бригад включають уважних осіб віком від 18 до 30 років, які не мають дефектів мови та слуху. Зазвичай бригади складаються з 50 % чоловіків і 50 % жінок.

Бригади перед початком роботи проходять двотижневий термін тренування. Досвід, набутий у процесі виконання цих робіт, показує, що нетреновані бригади допускають суб'єктивні помилки.

5.4. Метод вимірювання еквівалента загасання розбірливості мови AEN

Цей метод, скорочено названий AEN (від франц. Affaiblissement Equivalent pour la Nettete), оснований на порівнянні розбірливості мови на випробувальному тракті з розбірливістю на спеціальному еталонному тракті. Суть методу пояснює рис. 5.7.

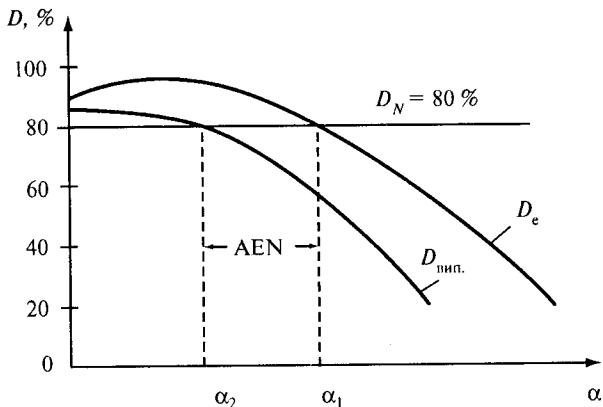


Рис. 5.7. Визначення еквівалента загасання за розбірливістю

Маємо еталонний вимірювальний тракт у складі передавальної та приймальної частин і магазину затухань, увімкненим між ними. Характеристики цієї еталонної системи, яка має назву ARAEN, стандартизовані Міжнародним консультативним комітетом з телефонії та телеграфії (МККГТ). Залежність звукової розбірливості D_e , отриманої під час передавання через систему, від загасання a , яке вводиться в тракт, буде виражатися кривою D_e .

Якщо тепер замінити приймальну та передавальну частини еталонної системи відповідними частинами випробуваної системи (наприклад, передавальним та приймальним ТА) і знову виміряти звукову розбірливість $D_{\text{вип}}$ за тих самих значень затухання a , що і в першому випадку, тоді отримаємо залежність $D_{\text{вип}} = f_2(a)$, яку зображено на тому ж графіку, що й крива $D_e = f_1(a)$.

Далі задають деяке контрольне значення розбірливості, наприклад $D_N = 80\%$, і визначають загасання a_1 і a_2 , які відповідають цьому значенню D_N для еталонного та випробувального трактів.

Різницю між цими загасаннями (в децибелах), що являє собою те загасання, яке потрібно ввести (вивести) в еталонну систему, щоб отримати таку саму розбірливість, як і на випробуваному тракті, називають еквівалентом загасання за розбірливістю:

$$AEN = a_1 - a_2.$$

Якщо випробувальний тракт гірший ніж еталонний, тоді AEN позитивний, а якщо кращий – негативний.

Для користування цим методом потрібно мати спеціальний еталонний вимірювальний тракт, що зменшує можливість широкого впровадження.

5.5. Оцінювання натуральності передавання мови

Оцінці натуральності мови, передаваної по трактах телефонного зв'язку, стали приділяти увагу тільки останніми роками. Це пояснюється тим, що у «звичних» телефонних трактах і трактах радіомовлення натуральність звучання мови корельовано з її якістю передавання за розбірливістю. Для таких трактів можна стверджувати, що чим вища розбірливість, тим вища натуральність. Це й зрозуміло, оскільки в цих умовах обидві ці розбірливості втрачаються одними і тими самими видами спотворень, головним чином частотними і нелінійними.

Однак у другій половині ХХ ст. у практику зв'язку почали впроваджувати різні спеціальні системи зв'язку, у яких використовуються різноманітні системи передавання даних. Тепер вважається, що (як підтверджується практикою) найбільш перспективні – системи передавання дискретних сигналів, тим більше, що тенденція розвитку систем та мереж зв'язку спрямована на створення єдиної інтегрованої технології телекомунікаційних мереж. Це зумовило потребу природний мовний сигнал перетворювати в дискретну (цифрову) форму. Прикладом можуть бути системи комп'ютерного та вокодерного зв'язку. У цих системах кореляція між розбірливістю та натуральністю значно зменшується, і за дуже високої розбірливості натуральність може бути низькою.

Неприродність звучання і неможливість розпізнавати голоси негативно оцінюють абоненти телефонних мереж. Тому на сьогодні оцінка якості передавання мови за натуральністю актуальна.

Загальноприйнятого кількісного критерію оцінки натуральності до цього часу ще не встановлено через складність цього завдання. Запропоновані для цього завдання методи не повністю задовольняють техніку передавання мови.

Один із таких методів – метод визначення відсотка узнавання. Суть методу така. Випробувальна бригада складається з тих, що розмовляти-

муть і тих, що слухатимуть. Ті, що слухатимуть, повинні знати голоси тих, що будуть розмовляти в оригіналі. У процесі випробування виконують достатню для статистичної обробки кількість заслуховувань мови тих, що розмовляють через випробувальний тракт. Відсоток випадків, у яких голос того, хто розмовляє впізнає той, що слухатиме відносно загальної кількості дослідів і буде оцінкою натуральності.

Недолік методу полягає в тому, що впізнавання не тотожне натуральності, а тільки його частина, тим більше він дуже громіздкий.

Відомий також метод оцінювання натуральності, оснований на визначенні коефіцієнта кореляції між вхідним і вихідним процесами. У загальному вигляді коефіцієнт кореляції R характеризує ступінь взаємного зв'язку між двома будь-якими процесами x і y і визначається виразом

$$R = \frac{(x - \bar{x}) * (y - \bar{y})}{\sqrt{(x - \bar{x})^2 * (y - \bar{y})^2}}$$

Стосовно до трактів зв'язку x – процес на вході, y – процес на виході, \bar{x} , \bar{y} – середньостатистичні значення x та y . Для отримання величини, яка відображала б слухове відчуття, запропоновано виміряти коефіцієнт кореляції в критичній смузі сприйняття з подальшим зваженням усередненням для тракту в цілому. Але допустимість у тих чи тих випадках значень коефіцієнтів кореляції залишається поки що не визначеною. Для використання цього коефіцієнта потрібна спеціальна досить складна апаратура.

Практичного впровадження у виробництво та експлуатацію апаратури зв'язку методи оцінювання натуральності поки що не отримали. Їх використовують переважно в науково-дослідних організаціях.

5.6. Електроакустичний розрахунок мовного тракту

5.6.1. Постановка завдання і вихідні дані

Як тільки були визначені фізичні характеристики мови та слуху і створено теорію розбірливості, яка ґрунтується на формантних поданнях, проектувальники телефонної та радіотелефонної апаратури отримали розрахункові інструментальні засоби, які дозволили їм розробляти нові пристрої на основі добре вивчених закономірностей.

Розрахунковий метод використовують для проектування ТА, спеціальних видів переговорних пристроїв і радіостанцій.

Можна виконувати такі розрахунки:

- перевірні – визначають якість передавання мови, яка досягається на тракті із заданими характеристиками;

- на заданій якості передавання – змінюючи значення деяких параметрів тракту, домагаються заданого значення показника якості;
- максимально можливої якості передавання – пошук оптимальних частотних і динамічних характеристик тракту.

Узагальнюючи ці три частинні розрахунки, сформулюємо постановку завдання: розрахувати якість передавання мови через певний тракт як функцію її електроакустичних і електричних параметрів та характеристик шумової завади.

Критерій якості передавання мови – це забезпечене на певному тракті значення розбірливості. Загалом розрахунки зводять до визначення складової розбірливості S . Поки що не знайдено способів аналітичного оцінювання якості передавання мови за показником натуральності, але для телефонних трактів зв'язку, у яких мова передається неперервними (аналоговими) сигналами, натуральність пропорційна розбірливості. Ці показники (розбірливість і натуральність) не корельовані у разі передавання мови цифровими методами.

Теоретична основа розрахунків – це встановлення зв'язків між розбірливістю і тими чинниками, від яких вона залежить.

До таких чинників належать:

- характеристики мови;
- характеристики слуху;
- електроакустичні параметри тракту;
- електричні параметри тракту;
- характеристики шумової завади.

Усі ці чинники розглядалися окремо. Перед тим як перейти до встановлення основних розрахункових співвідношень, доцільно визначити узагальнену характеристику тракту, до якої зводять сукупність загасань і підсилювань усіх його елементів.

Розглянемо схему (рис. 5.8), узагальнену для розрахунку розбірливості на тракті будь-якого типу, та встановимо для неї деякі визначення, якими потім користуватимемося.

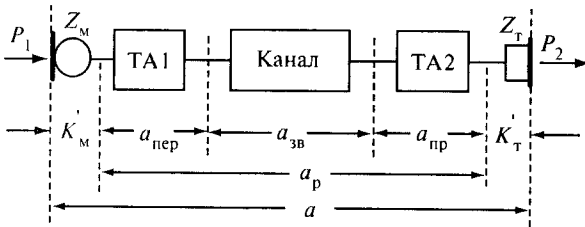


Рис. 5.8. Узагальнена розрахункова схема мовного тракту за заданих параметрів передавального і приймального апаратів

Повним загасанням мовного тракту a від звуку до звуку будемо називати логарифмічне відношення звукових тисків на мікрофоні (P_1) і в усі того, хто слухає (P_2), або, що те саме, різницю між рівнями цих звукових тисків (дБ):

$$a = 20 \cdot \lg \frac{P_1}{P_2}.$$

Обернене відношення цих звукових тисків – це повне підсилення тракту від звуку до звуку (дБ):

$$S_{\tau} = 20 \cdot \lg \frac{P_1}{P_2} = -a.$$

Якщо відомо такі параметри окремих елементів тракту, як чутливість перетворювачів K_M і K_T , модулі їх опору $|Z_M|$ і $|Z_T|$, робочі загасання передавальної та приймальної схеми $a_{\text{пер}}$ і $a_{\text{пр}}$ та власне загасання каналу зв'язку $a_{\text{зв}}$, то повне загасання тракту від звуку до звуку a можна виразити через параметри його елементів:

$$a = a_p - S'_{M-T},$$

де a_p – робоче загасання електричної ділянки тракту, для якої мікрофон розглядають як генератор, а телефон – як приймач; S'_{M-T} – результуюче зведене підсилення системи мікрофон–телефон.

Ці величини, в свою чергу, виражають так:

$$a_p = a_{\text{пер}} - a_{\text{зв}} - a_{\text{пр}},$$

де $a_{\text{пер}}$ – робоче загасання схеми ТА на передавання; $a_{\text{пр}}$ – робоче загасання схеми ТА на приймання; $a_{\text{зв}}$ – власне загасання лінії зв'язку:

$$S'_{M-T} = 20 \cdot \lg \left(K_M \sqrt{\frac{600}{|Z_M|}} K_T \sqrt{\frac{|Z_T|}{600}} \right) = 20 \cdot \lg K'_M K'_T.$$

Тут $K'_M = K_M \sqrt{\frac{600}{|Z_M|}}$ – зведена чутливість мікрофона; $K'_T = K_T \sqrt{\frac{|Z_T|}{600}}$ –

зведена чутливість телефону.

При цьому повне загасання тракту від звуку до звуку

$$a = a_{\text{пер}} - a_{\text{зв}} - a_{\text{пр}} - S'_{M-T}.$$

Інколи можна задавати коефіцієнти передачі K_1 і приймання K_2 ТА або інших переговорних пристроїв. Тоді узагальнена розрахункова схема тракту для передавання мови набуває такого вигляду, як на рис. 5.9.

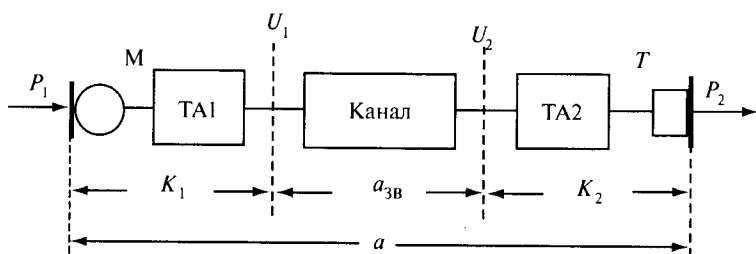


Рис. 5.9. Узагальнена розрахункова схема мовного тракту

При цьому одержимо вираз для загасання тракту від звуку до звуку:

$$a = 20 \cdot \lg \frac{P_1}{P_2} = 20 \cdot \lg \left(\frac{P_1 u_1 u_2}{u_2 u_2 P_2} \right).$$

Ураховуючи те, що $\frac{u_1}{P_1} = K_1$ – коефіцієнт передачі ТА за напругою, $\frac{P_2}{u_2} = K_2$ – коефіцієнт приймання ТА, а логарифмічне відношення

$20 \cdot \lg \frac{u_1}{u_2} = a_{зв}$ – загасання передачі каналом зв'язку, отримуємо

$$a = 20 \cdot \lg \frac{1}{K_1} + 20 \cdot \lg \frac{u_1}{u_2} + 20 \cdot \lg \frac{1}{K_2} = a_{зв} - 20 \cdot \lg K_1 K_2.$$

Величина $20 \cdot \lg K_1 K_2$ – результуюче підсилення приймання-передавання ТА. Тоді остаточною формула набуває вигляду

$$a = a_{зв} - S_{12}.$$

Тепер маємо всі потрібні вихідні дані для вирішення поставленого завдання – аналітичного розрахунку розбірливості за заданими характеристиками тракту.

5.6.2. Загальний метод розрахунку розбірливості за параметрами тракту

Загальну ідею розрахунку розбірливості мови можна розглядати в такій послідовності.

Оператор, який розмовляє, розглядається як деякий еквівалентний генератор, який передає на вхід тракту сигнал зі статистичними характеристиками. Зокрема, відомо середньостатистичний спектр і динамічний діапазон цього сигналу.

Приймальний оператор уподібнюється при цьому деякому еквівалентному приймачу, поріг спрацювання якого визначається порогом чутності, який змінюється під дією маскувальної завади.

У процесі передавання мови через електроакустичний тракт, який знаходиться між генератором та приймачем, частково втрачаються складові мови.

Знаючи кількісні характеристики первинного сигналу, приймача і тракту передавання, можна оцінити, яка частка із загальної кількості складових мовного сигналу, що містять змістову інформацію, дійде до приймача з рівнем, який перевищує поріг сприйняття. Ця частка (у відсотках) до загальної кількості складових, що містяться в первинному сигналі, і визначить шукану розбірливість на заданому тракті.

Ця загальна ідея набуде конкретної математичної форми після того, як будуть розглянуті деякі уточнені її обставини. Перша з них полягає в тому, що всі без винятку величини, від яких залежить розбірливість мови, є функціями частоти. Звідси випливає, що поставлене завдання можна вирішити розбиттям спектра, який займає мовний сигнал, на елементарні розрахункові смуги, достатньо вузькі для того, щоб можна було не зважати на частотну залежність потрібних параметрів у межах смуги і відносити їх до її середньої частоти. Тоді можна обчислити частку розбірливості, яка вноситься кожною розрахунковою смугою, а результируючий ефект отримати як суму ефектів у всіх розрахункових смугах.

Друга уточнювальна обставина пов'язана з першою і полягає в конкретизації виду розбірливості, який підлягає безпосередньому аналітичному розрахунку. Було розглянуто п'ять видів розбірливості, але таким видом розбірливості може бути і така, яка має властивість адитивності (підсумовування) за частотою.

Виявляється, що з усіх видів розбірливості властивість адитивності за частотою властива тільки формантній розбірливості A . Це випливає з самого поняття розбірливості формант. Дійсно, якщо в смузі $f_1 \dots f_2$ A_1 формант, а в смузі $f_3 \dots f_4$ A_2 формант, то у разі сумісного пропускання цих смуг по тракту буде прийнято $A_{12} = A_1 + A_2$, таким чином виконується властивість адитивності.

Такі співвідношення не зберігаються у разі приймання складніших елементів мови: звуків, складів, слів і фраз. Так, наприклад, звуки, форманти яких розміщені в різних частотних смугах, не можуть бути розпізнані ні під час приймання смуги $f_1 \dots f_2$, ні під час приймання смуги $f_3 \dots f_4$, вони розпізнаються тільки за сумісної дії на слуховий апарат.

І так, ще раз уточнимо ідею розрахунку.

Частотний діапазон передавання розподіляється на елементарні, досить вузькі смуги. У кожній смузі обчислюється форманта розбірливості ΔA_k .

Результуючу формантну розбірливість A для всього пропускового спектра визначають на підставі властивості адитивності формантної розбірливості за частотою як суму $A = \sum_{k \neq 1}^n \Delta A_k$, де n – кількість елементарних розрахункових смуг.

Перехід від величини розрахункової формантної розбірливості до будь-якого другого виду розбірливості (D, S, W, I) виконують за залежностями між цими видами розбірливості і формантною розбірливістю, встановленими для певної національної мови.

Методику розрахунку розбірливості створимо, узявши за основу визначення розбірливості як імовірність правильного приймання елементів мови на заданому тракті.

Отже, якщо розбірливість формант по всьому спектру A – імовірність правильного приймання формант під час пропускання всіх смуг, тоді ΔA_k можна трактувати як імовірність правильного приймання формант під час пропускання k -ї смуги.

Розглянемо тепер найпростіший випадок, коли мова передається в оптимальних умовах і той, що слухає свідомо (явно), приймає все, що передає той, що розмовляє.

Очевидно, що в цьому разі, якщо кількість випробувань велика, середнє значення величини ΔA_k буде постійним і дорівнюватиме максимальній імовірності приймання формант $\Delta A_{\max k}$ у заданій смузі. Величина $\Delta A_{\max k}$ – це та частка від загального складу формант у мовному сигналі, яка припадає на розрахункову смугу. Її можна отримати з фізичної характеристики мови (з формантного розподілу в інтегральній формі $A = \varphi(f)$).

Оскільки максимально можлива розбірливість формант під час пропускання всіх n розрахункових смуг для випадку передавання мови по тракту без втрат дорівнює одиниці, то для нього буде справедливе співвідношення

$$\sum_{k=1}^n \Delta A_{\max k} = 1.$$

Якщо тепер перейти від розгляду ідеальної системи до реальних умов передачі, коли частина інформації, яка є у вихідному мовному сигналі, втрачається під час передавання через тракт, тоді розбірливість формант у розрахунковій смузі вже не буде дорівнювати $\Delta A_{\max k}$, а буде меншою від неї. За цих умов

$$\Delta A_k = P_k \Delta A_{\max k},$$

де P_k – коефіцієнт сприйняття, який показує, яка частка від загальної кількості формант у k -й смузі дійде до вуха того, що слухає, з рівнем, вищим від порогового.

Це співвідношення можна отримати і на основі ймовірнісного підходу. Розбірливість у смузі ΔA_k можна розглядати як ймовірність складної події, яка виникає за одночасного початку двох незалежних подій, ймовірності яких відомі. Ймовірність того, що той, хто розмовляє, передасть ту чи ту кількість формант, дорівнює $\Delta A_{\max k}$; а ймовірність того, що слухає, сприйме ці форманти без втрат, дорівнює P_k . Тоді результуюча ймовірність буде дорівнювати їх співвідношенню.

Очевидно, що коефіцієнт сприйняття може набувати значення від 0 до 1. Це означає, якщо $P_k = 0$, то той, що слухає, зовсім не чує того, що розмовляє, а якщо $P_k = 1$, то той, що слухає, повністю сприймає інформацію без втрат (оптимальний випадок).

З урахуванням цих положень ймовірність правильного сприймання формант у будь-якій смузі вихідного спектра являє собою результуючу формантну розбірливість, яку визначають так:

$$A = \sum_{K=1}^n P_k \Delta A_{\max k}.$$

Вибір ширини елементарних смуг. До деякого ступеня цей вибір довільний; важливо, щоб ці смуги були досить вузькими. Якщо вибрати всі смуги однакової ширини, то величина $\Delta A_{\max k}$ для них буде різною. Зручно припустити, що межі смуг такі, що кожна наступна смуга забезпечує однаковий приріст формантної розбірливості $\Delta A_{\max k} = \text{const}$. Для цього випадку множник $\Delta A_{\max k} = \Delta A_{\max}$ можна винести за знак суми:

$$A = \Delta A_{\max} \sum_{K=1}^n P_k.$$

Зробивши останнє уточнення – кількість елементарних розрахункових смуг $n = 20$, так що $\Delta A_{\max} = \frac{A_{\max}}{20} = \frac{1}{20} = 0,05$, отримаємо кінцеву розрахункову формулу

$$A = 0,05 \sum_{K=1}^{20} P_k. \quad (5.1)$$

Спосіб отримання границь і середніх частот згаданих 20 елементарних розрахункових смуг показано на рис. 5.10. Значення середніх частот однакової розбірливості наведено в табл. 5.5. Ці границі отримано з формантного розподілу в інтегральній формі шляхом відліку приросту $\Delta A_{\max} = 0,05$.

Таблиця 5.5

| Номер смуг | $f_{\text{ср}}, \text{Гц}$ |
|------------|----------------------------|
| 1 | 250 |
| 2 | 500 |
| 3 | 650 |
| 4 | 800 |
| 5 | 950 |
| 6 | 1125 |
| 7 | 1300 |
| 8 | 1500 |
| 9 | 1700 |
| 10 | 1875 |
| 11 | 2050 |
| 12 | 2225 |
| 13 | 2425 |
| 14 | 2725 |
| 15 | 3100 |
| 16 | 3500 |
| 17 | 3850 |
| 18 | 4550 |
| 19 | 6150 |
| 20 | 8600 |

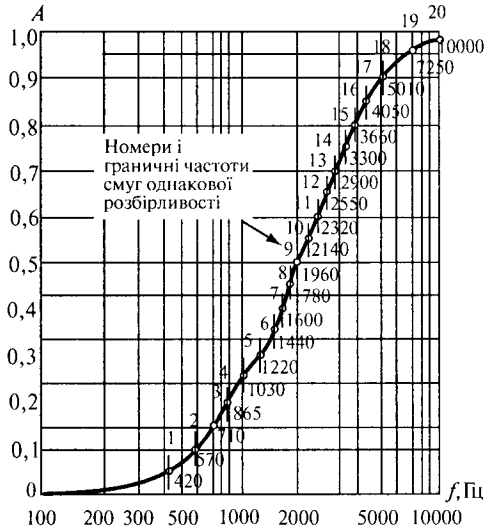


Рис. 5.10. Визначення границь і середніх частот смуг однакої розбірливості

Із формули (5.1) видно, що, по суті, розрахунок розбірливості зводиться до визначення коефіцієнта сприйняття P_k у розрахункових смугах. Виявляється, що ця величина є цілком визначеною функцією ефективного рівня відчуття формант у смузі E' . У цьому не важко переконатись, якщо виходити з таких міркувань.

У разі малих рівнів відчуття ні одна формантна складова не перевищує рівня порога чутності і тому не сприймається. Очевидно, що при цьому $P_k = 0$. Якщо тепер поступово збільшувати ефективний рівень відчуття прийманого сигналу, то дедалі більша і більша кількість формантних складових буде з'являтися над порогом і сприйматися тим, хто слухає. Значення коефіцієнта P_k буде збільшуватись згідно зі статистичним законом розподілу формант за рівнями. За деякого, достатньо великого рівня відчуття E' всі, навіть найслабші, формантні складові виявляються над порогом і сприймаються. Коефіцієнт P_k при цьому досягає свого максимального значення.

Подальше підвищення рівня прийманої мови вже не підвищує розбірливості, а, починаючи з деякого значення, навіть може зменшити її через первантаження слухового апарата зайвою гучністю.

Оскільки статистичний закон розподілу енергії формант за рівнями можна вивести з відомих характеристик амплітудного складу мови $Z = f(B)$, то з урахуванням природного розкиду рівнів голосів і порогів чутності можна отримати потрібну для розрахунків розбірливості залежність коефіцієнта сприйняття від ефективного рівня чутності формант $P = f(E')$, яку наведено в табл. 5.6 і на рис. 5.11.

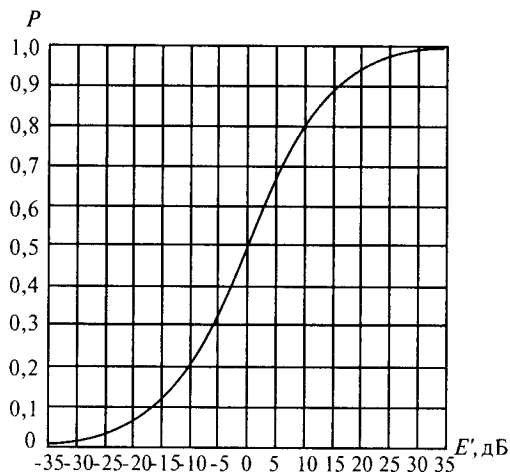


Рис. 5.11. Залежність коефіцієнта сприйняття типової випробувальної бригади від ефективного рівня чутності формант

Таблиця 5.6

| E' , дБ | P |
|-----------|------|
| -35 | 0,00 |
| -30 | 0,01 |
| -25 | 0,02 |
| -20 | 0,05 |
| -15 | 0,11 |
| -10 | 0,21 |
| -5 | 0,35 |
| 0 | 0,50 |
| 5 | 0,65 |
| 10 | 0,79 |
| 15 | 0,89 |
| 20 | 0,95 |
| 25 | 0,98 |
| 30 | 0,99 |
| 35 | 1,00 |

Аргумент цієї функції – ефективний рівень формант E' – для кожної розрахункової смуги можна визначити за загальною формулою, у якій враховуються параметри мови, слуху, тракту і шумової завади в цій смугі:

$$E' = B'_p + k - (a + M) - \beta_0,$$

де B'_p – середній спектральний рівень формант на передавальному кінці, який збуджує мікрофон; k – логарифмічна ширина критичної смуги слуху; a – повне загасання тракту від звуку до звуку; M – результуюче маскування від усіх видів шумів; β_0 – пороговий рівень для чистого тону (всі величини належать до середньої частоти цієї елементарної смуги).

За деякого досить великого рівня шуму в усі того, що слухає, результат приймання перестас залежати від слухових даних оператора.

Критерієм застосування цієї умови є значення ефективного рівня відчуття шуму $Z \geq 10$ дБ, де $Z = B_{ш} + \dots - \beta_0$.

Для таких випадків ефективний рівень чутності формант у смузі можна розраховувати за спрощеною формулою

$$E' = B'_p - a - B_{ш},$$

де $B_{ш}$ – спектральний рівень результуючого шуму в усі того, хто слухає.

Тепер можна сформулювати алгоритм виконання розрахунку розбірливості за параметрами тракту:

1) уточнити задачу розрахунку, спеціальні особливості заданого тракту і режиму його роботи, наявність усіх потрібних для розрахунку вихідних даних;

2) визначити фактичну кількість елементарних смуг, яку потрібно враховувати під час обчислення;

3) для кожної розрахункової смуги обчислити ефективний рівень чутності формант E'_k за формулами (E' – повну і скорочену);

4) за обчисленими величинами E'_k визначити коефіцієнти P_k за залежністю $P = f(E')$;

5) обчислити формантну розбірливість за формулою $A = 0,05 \sum_{k=1}^{n_1} P_k$.

6) визначити складову розбірливість S за залежністю $S = f(A)$ або інший будь-який вид розбірливості за взаємозалежностями між ними.

5.6.3. Особливість розрахунку розбірливості мови в умовах сильного шуму

Під час приймання-передавання мови в умовах сильного шуму виникають нові чинники, які впливають на результат приймання:

- підвищення рівня та деформація обвідної спектра мови під час форсування голосу на передавальному кінці;
- явище перевантаження слуху гучною мовою і шумом на приймальному кінці тракту.

Згадаємо статистичну залежність рівня інтенсивності мовного сигналу.

Якщо той, що розмовляє, знаходиться в умовах сильного шуму, то спостерігається явище мимовільного підвищення ним рівня свого голосу, яке називається форсуванням мови. Суть форсування: $B_{p \max}$ зміщується в бік більших частот, збільшується рівень B_p , енергія мови зростає приблизно в 25 разів (рис. 5.12).

Другий чинник ураховують за допомогою експериментально отриманої сім'ї кривих залежності коефіцієнта сприйняття від ефективного рівня чутності $P = f(E')$ за різних сумарних рівнів шуму в усі того, що слухає (рис. 5.13).

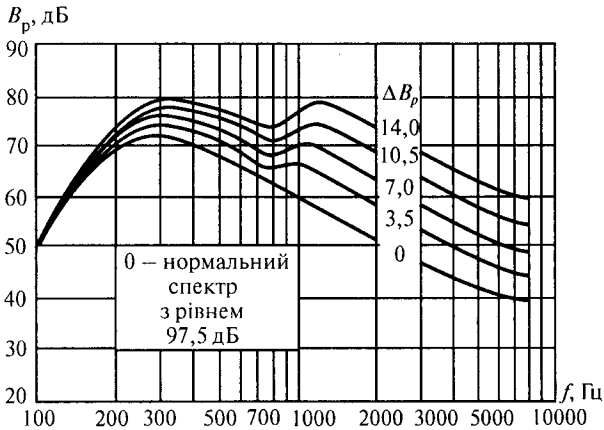


Рис. 5.12. Частотний спектр форсованої мови за різного ступеня форсування (V_p)

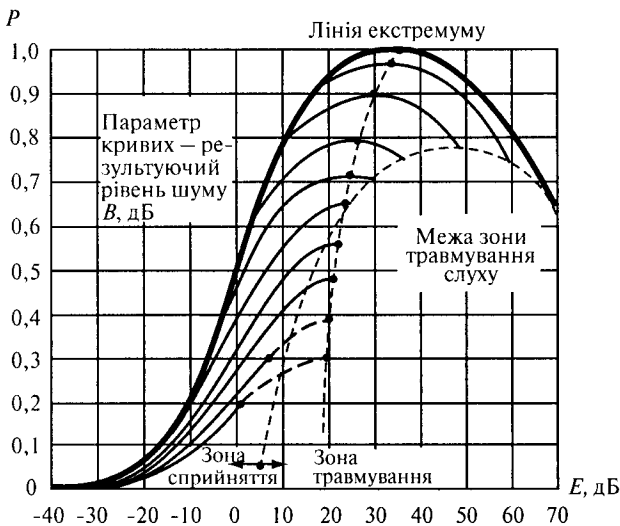


Рис. 5.13. Залежність коефіцієнта сприйняття P від ефективного рівня відчуття формант E' для різних рівнів шуму в ўсі того, що слухає

Слід зазначити, що за будь-якого рівня шуму спостерігається максимум (оптимум) найкращого відчуття. Якщо E' перевищить своє оптимальне значення, то спостерігається зменшення коефіцієнта доти, доки буде досягнуто поріг більового відчуття.

5.7. Методика розрахунку розбірливості за параметрами тракту (послідовність виконання розрахунків)

5.7.1. Постановка завдання

Розрахувати якість передавання мови за складовою розбірливості під час телефонування за допомогою заданого типу апаратів через дві автоматичні міжміські телефонні станції (АМТС), з'єднаних одна з одною каналом, створеним апаратурою ущільнення зі стандартними електричними параметрами. Телефонні апарати підключаються до своїх АМТС заданим типом кабеля визначеної довжини. Сумарний рівень шуму в приміщеннях приймання та передавання $B_{ш0} = 75$ дБ.

5.7.2. Вихідні дані

1. Розрахункова схема тракту (рис. 5.14).

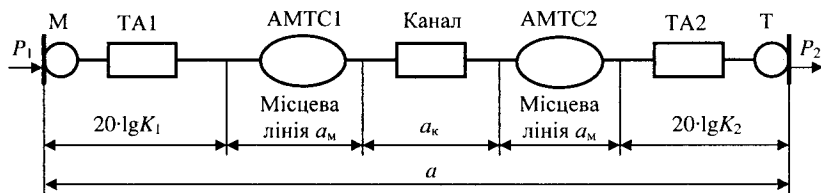


Рис. 5.14. Розрахункова схема тракту

2. Частотні характеристики параметрів ТА ($K_1, K_2, \gamma_m, a_3, a_{m,e}$):

$$K_1 = \frac{U_1}{P_1};$$

$$K_2 = \frac{P_2}{U_2};$$

K_3 – загальний ($U_1 = U_2$):

$$K_3 = \frac{P_2 U_1}{P_1 U_2} = K_1 K_2 = K_{12} \cdot 10^{-3}.$$

γ_m – шумозахищенність мікрофона:

$$\gamma_m = 20 \cdot \lg \frac{U_m}{U_{ш}};$$

Примітка: множник 10^{-3} – для узгодження однакових одиниць виміру.

U_m – напруга на виході мікрофона без шумів;

$U_{ш}$ – напруга на виході мікрофона під дією шуму в приміщенні;

a_3 – загасання, яке вносить заглушка;

$a_{м.е}$ – загасання протиміщевості, яке визначають за формулою

$$a_{м.е} = 20 \cdot \log \frac{U_m}{U_T},$$

де U_m, U_T – напруга відповідно на вихідних затискачах мікрофона та вхідних затискачах телефону; $a_{м.е}$ – у робочому діапазоні не менше 20 дБ.

3. Частотна характеристика $a_{м.е}$ – загасання, яке вносить місцева лінія тракту, що знаходиться між вихідними затискачами ТА та входом (виходом) каналу апаратури ущільнення $a_{м.е} = \varphi(f)$. Ця характеристика містить загасання, які вносить кабельна лінія визначеної довжини та міст АМТС, що постачає електричну енергію.

4. Залишкове загасання каналу в діапазоні 300...3400 Гц, що дорівнює $a_k = 6$ дБ (0,8 Нп).

5. Спектральна характеристика акустичного шуму із сумарним рівнем $B_{ш0} = 75$ дБ.

6. Звуковий тиск мови на мікрофон передавального апарата вважається відомим і дорівнює $p_1 = 1,5$ Нп/м² (сумарний рівень звуків мови 97,5 дБ).

5.7.3. Послідовність розрахунку та розрахункові вирази

1. Відповідно до заданої ширини смуги частот (300...3400 Гц), яка пропускається каналом, для розрахунку взяти 15 смуг однакової формантної розбірливості, починаючи з першої.

2. Розрахунок ефективного рівня відчуття формант виконують за скороченою формулою

$$E' = B'_p - a - B_{ш}.$$

3. Сумарне загасання тракту від звуку до звуку відповідно до розрахункової схеми (див. рис. 5.14) визначають за виразом

$$a = 2a_m + a_k - s_{12},$$

де $s_{12} = 20 \cdot \lg(k_{12})$.

4. Спектральний рівень результуючого шуму в усі того, що слухає, $B_{ш}$ визначають шляхом логарифмічного складання трьох його складових:

$$B_{ш} = B_{ш1} (+) B_{ш2} (+) B_{ш3},$$

де $B_{ш1} = (B_{ш0} - a_3)$ – шум, який проникає в ухо того, що слухає, з приміщення приймання; $B_{ш2} = (B_{ш0} - \gamma_m - a)$ – шум, який проникає з приміщення

передавання через електричний тракт і телефон; $B_{ш3} = (B_{ш0} - \gamma_m - a_{м.е} - S_{м-т})$ – шум, який проникає через ланцюг місцевого ефекту; $S_{м-т}$ – результуюче підсилення пари перетворювачів мікрофон – телефон:

$$S_{м-т} = 20 \cdot \lg K_m K_T \cdot 10^{-3},$$

де K_m – чутливість мікрофона; K_T – чутливість телефону.

Оцінюють порядок величини $B_{ш3}$. Оцінювання виконують на частоті 1000 Гц, де підсилення $S_{м-т}$ достатньо велике. Як приклад можна взяти значення K_m і K_T для відомих перетворювачів: для мікрофона типу ДЕМШ-1 $K_m = 1,0 \text{ мВ}/(\text{Н}/\text{м}^2)$, для телефону ТА-4 $K_T = 3,0 (\text{Н}/\text{м}^2)/\text{В}$. Загасання протимісцевості беруть таке, що дорівнює мінімальному значенню $a_{м.е}$. Тоді

$$S_{м-т} = 20 \cdot \lg K_m K_T \cdot 10^{-3} = 20 \cdot \lg 1 \cdot 3 \cdot 10^{-3} = -50,4 \text{ дБ}.$$

Рівень шуму, який проникає через ланцюг місцевого ефекту, буде дорівнювати:

$$B_{ш3} = (B_{ш0} - \gamma_m - a_{м.е} - S_{м-т}) = 41 - 2 - 20 - 50,4 = 31,4 \text{ дБ}.$$

Величину $B_{ш0}$ взято з графіка спектра шуму в приміщеннях передавання та приймання при частоті 1000 Гц, вона дорівнює 41 дБ.

Якщо порівняти значення $B_{ш3}$, спектральний рівень результуючого шуму розраховують з урахуванням лише двох його складових:

$$B_{ш} = B_{ш1} (+) B_{ш2}.$$

5. Оскільки спектр шуму $B_{ш} = \varphi(f)$ досить рівномірний, тоді значення коефіцієнта сприйняття P_k (після розрахунку величин E'_k для всіх розрахункових смуг) знаходимо завдяки кривій $P = f(E')$, яку вибирають з графіка (рис. 5.13) залежно від сумарного рівня шуму $B_{ш0\Sigma}$ у вусі того, що слухає.

6. Обчислюють фортматну розбірливість A за формулою

$$A = 0,05 \sum_{k=1}^{15} P_k.$$

7. Шукану складову розбірливість S визначають із графіка залежності $S = f(A)$.

6. Методи перетворення мовного сигналу

6.1. Загальні відомості та види перетворення сигналу

Досі ми розглядали такі системи телефонної передачі мови, у яких мовний сигнал незмінно зберігав у всіх ланках тракту свої частотні, амплітудні та часові характеристики, що аналогічні для натуральної мови, яку формує той, що розмовляє. Системи передавання мовних сигналів належать до аналогових систем передавання даних.

Але в ряді випадків можна побудувати значно ефективніші системи передавання мовних сигналів, якщо ця інформація буде передаватися по тракту зв'язку не у формі, яка визначається фізичною природою голосового апарату людини, а в суттєво зміненому вигляді. З погляду загальної теорії зв'язку такі зміни природної структури сигналу рівносильні його перекодуванню.

Основа створення нових систем передавання мовних сигналів у разі використання операції перекодування та досягнення ефективності – заздалегідь відомі загальні закономірності формування та сприйняття мовних звуків. Тому не передається по тракту повністю інформація мовного сигналу, а тільки його ознаки. А ознаки, завдяки знанням цих закономірностей, вкладаються в систему зв'язку. Це дає можливість передавати по тракту тільки ознаки (обмежену кількість сигналів), оскільки складові частини (вузли, пристрої та прилади) мають знання, як за ознаками можна відновити інформацію мовного сигналу, що надійшла від джерела інформації – оператора, що розмовляє.

Такі перетворення мовного сигналу відкривають простір для пошуку шляхів підвищення пропускну здатності напряму зв'язку за рахунок збільшення кількості каналів у заданому спектрі частот, підвищення завадостійкості мереж телекомунікацій, одночасної передачі по телефонному каналу мови та дискретно-цифрової інформації.

Розглянемо конкретні характеристики телефонного сигналу та виявимо можливості їх перетворення.

Із загальної теорії зв'язку відомо, що об'єм сигналу V є добутком трьох його умовних «вимірів»: частотного діапазону F , динамічного діапазону H та часу передавання T , тобто $V = FHT$.

Порівняємо об'єм телефонного сигналу $V_{\text{тел}}$ з об'ємом телеграфного сигналу $V_{\text{тлг}}$, які переносять ту саму текстову інформацію.

Нехай для спрощення розрахунків час передавання буде однаковий.

Припустимо, що мова передається в смузї завширшки $F_{\text{тлф}} = 3000$ Гц, динамічний діапазон якої $M = 40$ дБ, а телеграфна передача ведеться зі

швидкістю 200 дБ, для чого смуга частот $F_{\text{тлг}} = 100$ Гц, якщо $M = 60$ дБ (передаються два значення амплітуди сигналу $\pm E$).

Тоді об'єм телефонного сигналу $V_{\text{тлф}} = F_{\text{тлф}} H_{\text{тлф}} T_{\text{тлф}} = 3000 \cdot 40 \cdot 1 = 120\,000$ ум. од., а об'єм телеграфного сигналу $V_{\text{тлг}} = F_{\text{тлг}} H_{\text{тлг}} T_{\text{тлг}} = 100 \cdot 6 \cdot 1 = 600$ ум. од.

Отже, об'єм телефонного сигналу в цих умовах більший від телеграфного сигналу в $\frac{120\,000}{600} = 200$ разів.

Значне збільшення об'єму телефонного сигналу відносно телеграфного пояснюється тим, що незалежно один від одного передається та сама інформація, але в телефонному сигналі до суто текстової добавляється інформація про тембр (голос того, що розмовляє), тональність і динаміка. Основна причина цього явища полягає в тому, що телефонний сигнал має високий ступінь надмірності.

Надмірністю повідомлення або сигналу називають наявність у повідомленні елементів, які не потрібні для його правильного приймання.

Телефонний сигнал має нещільну структуру і його елементи, які несуть змістовну інформацію, далеко розосереджені один від одного, тоді як в телеграфному сигналі вони нібито щільно спресовані один з одним.

Але якби вся змістовна інформація мови була зосереджена у вузькій смузі 50...100 Гц замість декількох тисяч герців, то мова легко маскувалась би вузькосмуговими шумами зі збіжним спектром і потребувала б для спілкування людей трактів з дуже жорстко визначеними характеристиками. Широкосмуговий мовний сигнал може бути зруйнований завадою в тих чи тих ділянках, але зберігатиме достатню інформативність за рахунок інших.

Узагальнюючи ці міркування, можна стверджувати, що більша надмірність в об'ємі телефонного сигналу допускає значні його руйнування, при цьому все ж зберігається достатня інформаційна здатність. Отже, підвищується достовірність приймання мовного сигналу. Цю властивість використовують для передавання дискретної інформації, впроваджуючи коди зі штучно введеною в них надмірністю.

Скорочення об'єму передавального телефонного сигналу через зменшення штучно одного або декількох вимірів (при цьому враховується можливість цілеспрямованого руйнування та збереження інформаційної здатності) з одночасним зниженням вимог до відповідних характеристик каналу зв'язку – ширини пропущеної ним смуги частот і динамічного діапазону. Добуток цих параметрів каналу F_k, H_k, T_k складає і його об'єм V_k .

Такі перетворення можна з'ясувати з докладного розгляду тривимірної структури телефонного сигналу.

Зображення сигналу у вигляді паралелепіпеда отримують у тому разі, якщо розглядати граничні значення параметрів сигналу за всіма f , h , t (рис. 6.1).

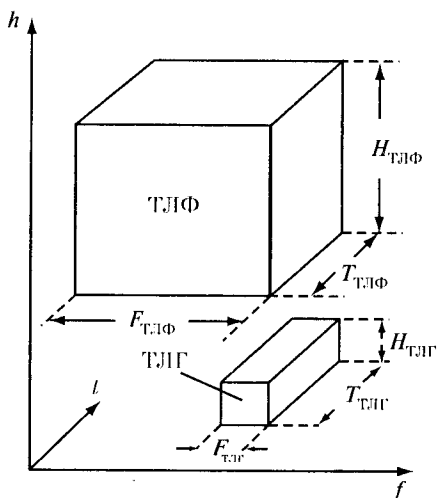


Рис. 6.1. Графічне порівняння об'єму телефонного і телеграфного сигналів

Якщо врахувати динамічні зміни миттєвого спектра та рівнів амплітуд його складових протягом деякого інтервалу часу T , то геометричне зображення телефонного сигналу може бути таким, як показано на рис. 6.2.

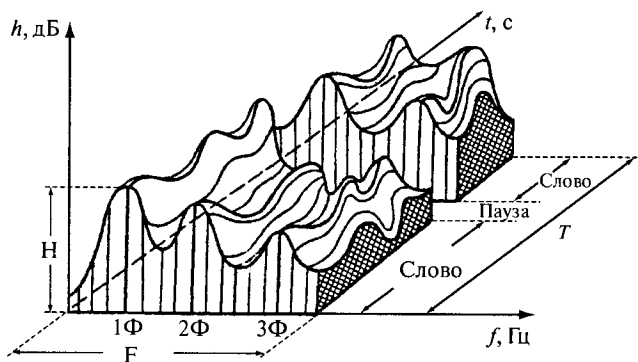


Рис. 6.2. Уточнення геометричного зображення телефонного сигналу:
 F – частотний діапазон; H – динамічний діапазон; T – деякий часовий інтервал, протягом якого розглядається мовний процес

Переріз зображення на рисунку геометричної фігури площиною, яка паралельна осям hf , являє собою змінний миттєвий амплітудно-частотний спектр процесу, а переріз площиною, яка паралельна осям ht , – змінне миттєве значення рівня амплітуди сигналу на частоті, яка відповідає взятому перерізу. На цьому рисунку показано під час передавання два слова, які розділено паузою.

Фактичний об'єм показаного тут сигналу визначають з виразу

$$\vartheta_{\text{тлф}} = \iint_{FT} h(f, t) df dt.$$

Аналізуючи рис. 6.2, можна відзначити такі чинники, які характеризують надмірність сигналу за трьома його вимірами:

1) за частотою: формантні ділянки розділено проміжками, у яких частотні складові майже не містять ніякої змістовної інформації, ширина формантних ділянок деяких звуків дуже велика і перевищує величину, потрібну для розпізнавання цих звуків;

2) за рівнями велика різниця між рівнями найгучніших і найтихіших звуків; зменшення цієї різниці не призведе до зниження розбірливості мови, а в деяких умовах може навіть підвищувати її;

3) за часом є паузи мови частково потрібні, а частково є втратою часу; як у першому, так і в другому випадку корисна інформація в канал не надходить.

Знаючи структуру мовного сигналу, інколи впроваджують такі його види перетворення:

- обмеження;
- трансформацію;
- компандування;
- перетворення неперервного (аналогового) мовного сигналу в дискретну форму.

Розглянемо основні закономірності, які пов'язані з цими видами перетворювань.

Обмеження – вилучення з передачі однієї або декількох частин сигналу без збереження інформації, яка міститься в цих вилучених частинах. Обмеження може бути частотним, амплітудним і часовим. Принцип частотного обмеження показано на рис. 6.3.

Частотне обмеження (I_0, I_1 – кількість інформації відповідно у вихідному і обмеженому сигналах), коли замість первинного сигналу, ширина смуги якої F_0 , у канал пропускається лише низькочастотна частина цього сигналу шириною F_1 . Інформація, яка міститься у високочастотній частині (незаштрихована частина первинного об'єму), затримується фільтром і безповоротно губиться. Частотне обмеження властиве майже будь-якій звичайній апаратурі телефонного зв'язку.

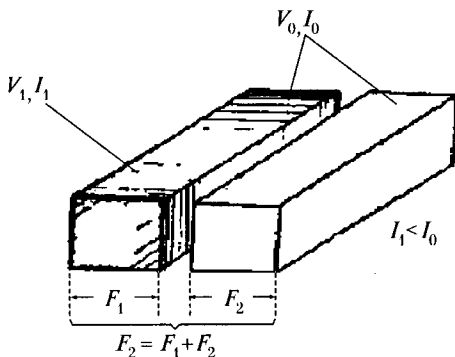


Рис. 6.3. Частотне обмеження: $F_0 = F_1 + F_2$; $I_1 < I_0$

Стандартний телефонний канал зі смугою пропускання 300...3400 Гц частотно-обмежений, який зберігає, однак, досить високу якість передавання. Це і буде впровадженням частотного обмеження.

6.1.1. Трансформація

Трансформація – це такий вид перетворення сигналу, за якого зменшуються одні його виміри за рахунок збільшення інших зі збереженням постійного (незмінного) об'єму сигналу. Принцип трансформації показано на рис. 6.4.

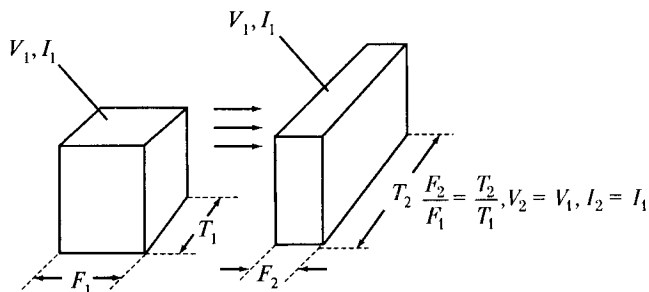


Рис. 6.4. Трансформація

Тут зменшення передаваного частотного спектра компенсується збільшенням тривалості передавання (часу).

Ефект трансформації з розтягненням мовного сигналу по осі часу можна спостерігати під час розмови по низькоякісному каналу з великим частотним або амплітудним обмеженням, якщо багатократними перетво-

рюваннями і повтореннями той, що розмовляє, домагається нарешті приймання всієї потрібної інформації.

6.1.2. Компандування

Компандування – це такий спосіб перетворення сигналу, за якого на передавальному кінці тракту виконується стиснення мовного сигналу за одним, двома або трьома вимірами, а на приймальному кінці – відновлення нормального об'єму сигналу шляхом розширення тих його вимірів, які були стиснені на передавальному кінці.

Компандування включає два взаємно оборотні процеси – компресію (стиснення) та експандування (розширення).

Принцип компандування за вимірами F і H за незмінного часу показано на рис. 6.5.

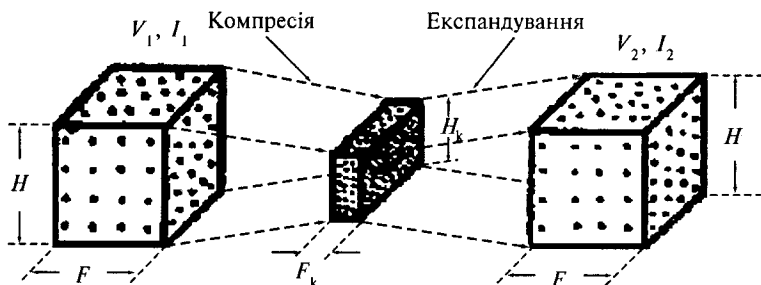


Рис. 6.5. Компандування

Найбільше практичне значення має частотне компандування мовного сигналу. Амплітудне компандування виконує допоміжну роль. Часове компандування рідко використовують через складність системи.

Перетворення неперервного мовного сигналу в дискретну форму полягає в перетворенні натуральної мови (телефонного сигналу) в послідовність імпульсів, яка передається і обробляється проміжними ланками тракту дискретними приладами, а на приймальному кінці знову перетворюється в натуральний мовний сигнал.

Розглянувши повне подання мовного сигналу в тривимірному просторі та аналізуючи зміни сигналу по трьох осях, доходимо висновку, що можна штучно змінювати будь-які параметри без суттєвого руйнування головного показника сигналу – розбірливості мови.

Зосередимо увагу на деяких видах перетворення, широко використовуваних у середовищі телекомунікацій, та принципах їх реалізації.

6.2. Частотне компандування мови.

Основні відомості та загальний принцип вокодера

Компандування як вид перетворення передбачає можливість зміни за всіма вимірами мови. Значного поширення отримав спосіб зміни частоти (основний параметр) і амплітуди (додатковий параметр). Цей спосіб уперше було використано для створення пристроїв, які отримали назву вокодер (від англ. voice coder – кодувальник голосу).

Розглянемо основні відомості та загальний принцип дії вокодера.

Можливі два способи компандування мови – безпосередній та параметричний. Стосовно частотного компандування безпосередній спосіб полягає в тому, що частотна компресія відбувається прямим перетворенням спектра мови без його аналізу та розкладання на складові, а експандування – без застосування місцевих джерел сигналів.

Відомо просту ідею безпосереднього компандування мови: на передавальному кінці тракту мова пропускається через подільник частоти, який ділить всі частоти складових спектра на n . При цьому, якщо вихідний початковий спектр має ширину ΔF , тоді в канал надходить спектр шириною $\Delta F/n$. На приймальному кінці цей спектр має проходити через помножувач частоти, який множить усі частоти на одну й ту саму частоту n і знову відновлює вихідну ширину спектра ΔF .

Ця приваблива ідея виявляється майже не реалізовною, тому що всі відомі подільники та помножувачі частоти виконують дії (операції) ділення та множення із заданим коефіцієнтом n тільки у разі надходження на їх вхід однієї фіксованої частоти, а під час збудження комплексом частот цей закон порушується. Тому принцип безпосереднього компандування не знайшов застосування.

Параметричним компандуванням називають такий спосіб перетворення сигналу, за якого його компресію виконують виділенням з нього обмеженої кількості параметрів, що повільно змінюються, завдяки яким сигнал може бути відновлений, а експандування виконується за рахунок місцевих джерел, які керуються цим комплексом параметрів.

Таким чином, за частотного параметричного компандування в канал поступає не сам спектр мови, а лише відомості про його характерні особливості (параметри). Ці відомості передаються за допомогою простих сигналів у вузькій смузі частот, що й забезпечує потрібну компресію.

Вихідний спектр відновлюється дією цих сигналів на рівномірний спектр, який створюється місцевим генератором. Місцевий генератор моделює ті особливості мовного сигналу, відомості про яких не потрібно передавати через канал.

Залежно від прийнятої системи параметрів, завдяки яким виконується відновлення первісного мовного сигналу, розрізняють основні типи вокодерів: смугові, формантні, гармонічні та фонемні.

Розглянемо загальний принцип дії вокодера будь-якої системи, який пояснюється узагальненою блок-схемою (рис. 6.6).

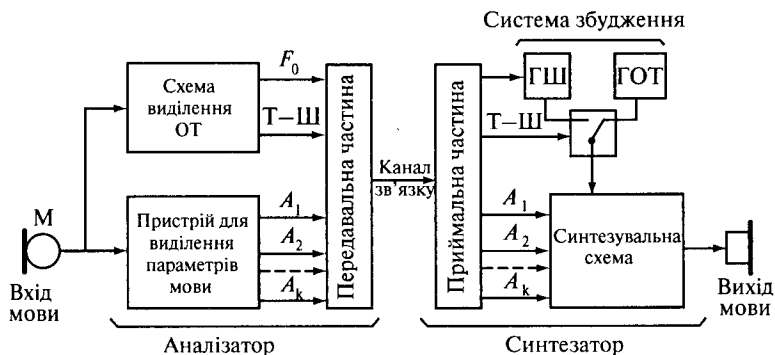


Рис. 6.6. Узагальнена блок-схема вокодер

Головними частинами вокодерного тракту є аналізатор, який виявляє параметри мовного сигналу, система передавання, яка забезпечує проходження інформації про ці параметри через канал зв'язку у вузькій смузі частот, і синтезатор, який відновлює первісний мовний сигнал.

Аналізатор вокодера складається з пристрою для виділення параметрів мовного сигналу $A_1, \dots, A_i, \dots, A_k$ та схеми виділення ОТ.

Для того щоб на приймальному кінці мова могла бути відновлена з достатньо високою розбірливістю і хоча б із задовільною якістю звучання, потрібно передати в синтезатор такі сигнали:

1) сигнал про те, чи передається у певний момент часу звук з тональним (Т) або шумовим (Ш) збудженням; ця інформація передається параметром Т – Ш (тон–шум);

2) якщо передається тональний звук, тоді має бути переданий сигнал про частоту ОТ – параметр F_0 ;

3) сигнали, за якими можна відновити миттєвий спектр звуку і його зміни з часом, тобто параметри $A_1, \dots, A_i, \dots, A_k$.

Останній формує пристрій для виділення параметрів мови, конкретний принцип дії якого залежить від типу вокодера.

Сигнали F_0 і Т – Ш «втягаються» з мови схемою виділення ОТ, яка будується незалежно від конкретного типу вокодера. Варіант структури такої схеми показано на рис. 6.7.

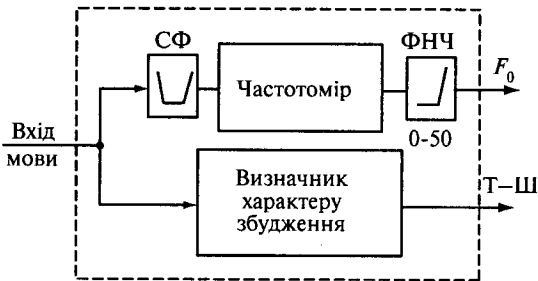


Рис. 6.7. Схема виділення ОТ

Мовний сигнал, який поступає на вхід системи, одночасно діє на два блоки – визначник характеру збудження Т – Ш і виділювач ОТ. Перший з них аналізує характер спектра і видає на вихід сигнал «Тон» або «Шум». Другий блок, який складається зі смугового фільтра СФ, що виділяє діапазон частот, у яких може знаходитись ОТ, частотоміра та вихідного фільтра низьких частот ФНЧ, визначає частоту ОТ і видає на вихід сигнал F_0 у вигляді постійної напруги, величина якої пропорційна частоті ОТ f_0 .

У синтезаторі вокодера також можна виділити дві основні частини, які функціонально зв'язані з відповідними двома частинами аналізатора. Це – місцеві джерела збудження у складі генератора шуму суцільного спектра ГШ, генератора основного тону ГОТ та перемикача Т – Ш і синтезувальної схеми. Остання є не що інше, як модель мовного тракту змінної форми спектра, створеного збуджувачем (ГОТ або ГШ), відповідно до сигналів, які надходять від аналізатора.

Розглянемо принцип дії синтезатора. Від приймальної частини системи передавання синтезатор отримує сигнали про параметри мовного процесу Т – Ш, F_0 , A_1 , ..., A_k , ..., A_k . Кожний з них діє на свій приймальний пристрій.

Якщо в певний момент часу передається глухий звук, тоді сигналом Т – Ш перемикач Т – Ш устанавлюється в положення «Ш» і в синтезувальну схему надходить рівномірний суцільний спектр від ГШ. Якщо вимовляється дзвінкий звук, то перемикач Т – Ш переводиться в положення «Т» і в синтезувальну схему буде видаватися дискретний спектр гармонік ОТ частоти f_0 , значення якої беруть відповідно до значення рівня сигналу F_0 .

Одночасно сигнали A_1 , ..., A_k поступають у схему, що синтезується, і змінюють обвідну суцільного або дискретного спектра, формуючи тим самим спектральні максимуми та мінімуми відповідно до змінних спектра первісного звуку. Синтезований таким чином сигнал надходить на мовний вихід синтезатора.

Послідовність процесів, які утворюються у вокодері під час аналізу та синтезу мови, показано на рис. 6.8, де на графіках *a*, *б*, *в*, *г* зображено перетворення спектра, які використовуються під час компресії та експандування у разі вимовлення дзвінкого звуку. У разі синтезу глухого звуку відмінність полягатиме в тому, що спектри $B_P(f)$ і $B_0(f)$ будуть суцільними.

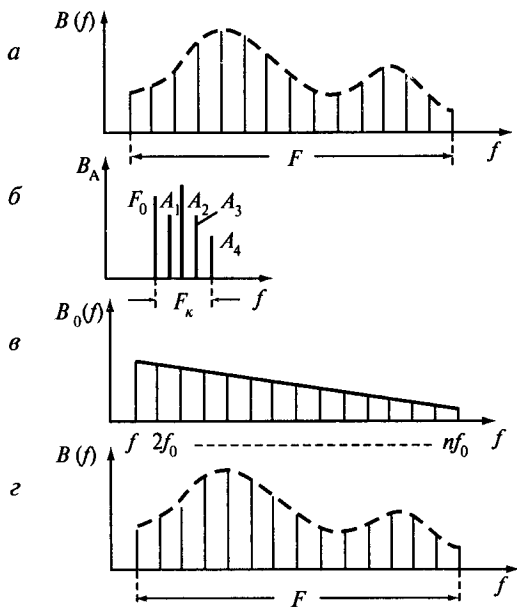


Рис. 6.8. Процес перетворення мови під час параметричного компандування: *a* – миттєвий спектр мови шириною F ; *б* – інформація про параметри спектра, який надходить в канал вузької смуги F_k ; *в* – спектр частот, який виробляє ГОТ; *г* – миттєвий спектр мови, відновлений під час взаємодії параметрів F_0, A_1, A_2 зі спектром, який виробляє ГОТ

Відношення ширини спектра мови до компресії F до ширини спектра компресованої мови F_k називають *коефіцієнтом компресії* вокодера:

$$\eta = \frac{F}{F_k}.$$

Коефіцієнт компресії вокодера будь-якого типу можна орієнтовно визначити, якщо відомо кількість параметрів m , які передаються по каналу зв'язку, і ширину смуги частот Δf , потрібну для передавання кожного параметра.

Але для підвищення надійності передавання беруть смугу частот шириною $2\Delta f$. Тоді ширина смуги частот компресованого сигналу

$$F_k = 2m\Delta f.$$

Розглянуті загальні принципи будови та дії вокодерів дозволяють легко розібратися і в конкретних особливостях вокодерів різних систем. По суті, вони відрізняються один від одного принципом дії пристрою для виділення параметрів мови в аналізаторі та синтезувальної схеми в синтезаторі.

Розглянемо використання вокодера на конкретному типі – смуговому вокодері (рис. 6.9).

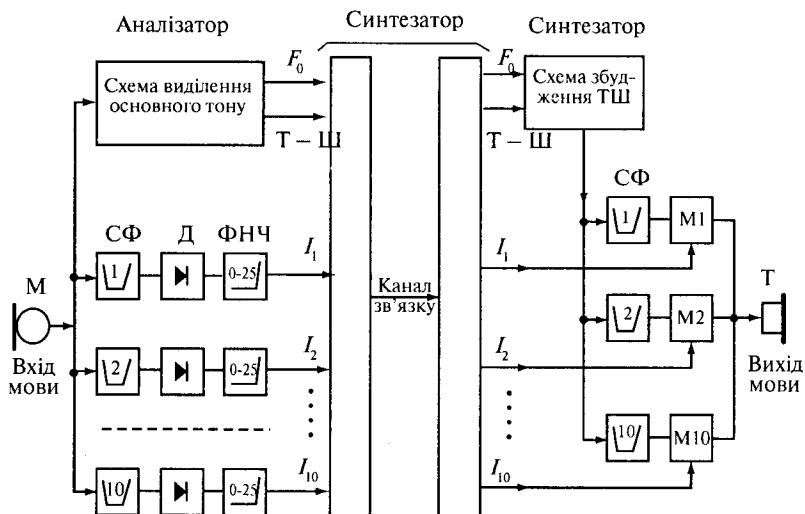


Рис. 6.9. Принцип дії смугового вокодера

У смуговому вокодері параметрами, які описують поточний спектр, є середні рівні енергії мови в смугах, на які ділиться частотний діапазон. Такі смуги називають спектральним каналом. Чим більше спектральних каналів, тим точнішим буде перетворення, але зменшиться коефіцієнт компресії.

Аналізатор смугового вокодера складається зі схеми виділення ОТ (принцип дії якого розглянуто раніше) та пристрою виділення параметрів обвідної спектра мови. Цей пристрій являє собою сукупність деякої кількості (від 1 до 20) спектральних каналів, у яких виконується визначення середнього рівня мови. Очевидно, чим більше буде утворено спектральних каналів, тим більша точність апроксимації спектра, але тим менше значення коефіцієнта компресії.

Кожний смуговий спектральний канал пропускає ширину смуги через СФ, потім випрямляє детектором Д з наступним пропусканням його через ФНЧ смугою пропускання 0...25 Гц. На виході кожного ФНЧ виділяється випрямлений струм, значення якого пропорційне енергії спектра в смузі. Ці сигнали, що позначають як I_1, \dots, I_{10} , являють собою параметри, які визначають форму обвідної поточного спектра.

Ці сигнали разом з двома іншими Т – Ш, F_0 , отриманими на виході схеми виділення ОТ, у вузькому частотному спектрі передаються на приймальну частину вокодера.

Синтезатор вокодера складається зі схеми – системи збудження (ГШ і ГОТ зі своїм перемикачем Т – Ш) і синтезувальної схеми, яка містить k спектральних каналів (у цьому разі – 10). Кожний з них має смуговий фільтр (СФ1 – СФ10) та модулятор (М1 – М10).

Система збудження Т – Ш видає на смугові фільтри широкий спектр частот – суцільний або дискретний залежно від сигналу Т – Ш. Кожна смуга, на яку ділить гребінка СФ1 – СФ10, надходить на модулятор. Підсилення або послаблення, які отримують у модуляторах частотні складові спектра тієї чи тієї смуги, залежить від значення параметрів I_1, \dots, I_{10} , які надходять із прийманої частини системи передавання.

Оскільки ці параметри відображають рівні мови в спектральних каналах аналізатора, то на мовному виході синтезатора вокодера відтворюється картина миттєвого спектра мови на передавальному кінці.

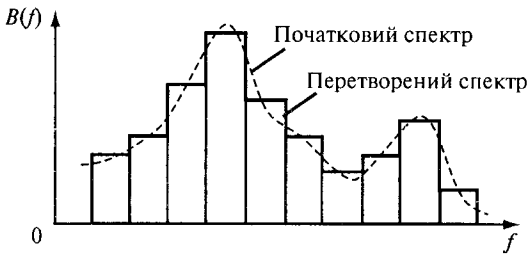


Рис. 6.10. Вхідний і перетворений спектр у смуговому вокодері

Похибка перетворення, як видно із рис. 6.10, пов'язана з тим, що реальна обвідна мовного спектра замінена тут ступеневою функцією. Мірою приближення останньої до реальної кривої є збільшення кількості ступенів, тобто збільшення кількості спектральних каналів.

Ширина смуги частот компресованої мови для вокодера такого типу з 12 передаваними параметрами:

$$F_k = 2m\Delta f = 2 \cdot 12 \cdot 25 = 600 \text{ Гц.}$$

У разі вихідного спектра в границях 300...3400 Гц коефіцієнт компресії

$$\eta = F/F_k = 3100/600 \approx 5.$$

Виділення основного тону. Найважче завдання у техніці частотного компанування – виділення параметрів ОТ.

Виникаючі тут труднощі визначають такими обставинами:

- широкі границі змін ОТ (70...350 Гц);
- діапазон ОТ не передається, або слабо передається звичайними мікрофонами;
- інколи спостерігається швидка зміна частоти ОТ, яка потребує розширення смуги частот, яка відводиться для передачі параметра F_0 в системі передавання.

Це завдання до кінця не вирішено.

У форматних вокодерах відновлення мовного сигналу виконують за інформацією про формантні максимуми. Параметрами, які передаються каналом зв'язку, є сигнали про середні частоти і рівні формант.

6.3. Принципи амплітудного та часового перетворення мови

У техніці передавання мови трапляються види перетворення мовних сигналів і амплітудне обмеження, амплітудне компандування та часове перетворення.

Амплітудне обмеження має місце під час передавання мови трактами з нелінійними характеристиками. Спеціальні амплітудні обмежувачі використовують для зменшення впливу на другі ланцюги та полегшення режиму підсилювальних та інших елементів тракту передавання мови.

Амплітудне компандування застосовують для зменшення динамічного діапазону мовного сигналу під час його проходження ділянками тракту з нелінійними амплітудними характеристиками. Відмінність між цими способами перетворення полягає в тому, що амплітудне компандування зберігає майже повністю інформацію, яка є в натуральному мовному сигналі.

Часове перетворення мови, на перший погляд, здається, не має практичного значення, оскільки розмова між двома співрозмовниками по телефону із заданою кількістю інформації не може бути скорочена або розтягнута порівняно з тривалістю, яка відповідає нормальній швидкості мови. Для окремої розмови воно так і є. Але коли група співрозмовників користується загальним пучком каналів, тривалість розмови однієї пари та кількість одночасно розмовляючих пар, то виявляється, що діючі при цьому статистичні закономірності дозволяють використати часове «компандування». Компандування розуміють як досягнення скорочення сума-

рного часу займання пучка каналів для проведення заданої кількості розмов.

Системи, які реалізують ідею статистичного аналізу використання пучків каналів та навантаження телефонними розмовами, називають системами статистичного часового ущільнення.

Принцип дії таких систем полягає в тому, що абоненту, який починає розмову, одночасно надається канал зв'язку до потрібного йому абонента, скануванням визначається його номер, який записується в запам'ятовувач, і пробний пристрій шукає вільний канал. На основі інформації номера абонента, який викликає, і номера вільного каналу електронний комутатор підключає його до знайденого каналу. Водночас ця інформація передається через спеціальний канал у запам'ятовувальний пристрій приймального кінця (адреси абонента та номера каналу) і вибраний абонент електронним комутатором підключається до вказаного каналу. Система симетрична.

6.4. Перетворення мовного сигналу в дискретну форму

Один з напрямів розвитку техніки передавання мови – створення систем, у яких мова передається каналами зв'язку не в натуральній формі, аналогічній реальному неперервному акустичному процесу, а у вигляді послідовності дискретних імпульсів. Процес перетворення неперервних сигналів в дискретну форму називають дискретизацією.

Дискретизація мовного сигналу забезпечує переваги таких систем:

1) системи інтегрованого обслуговування (ISDN), у яких інформація телефонних розмов, телеграфного, фототелеграфного, комп'ютерного та факсимільного зв'язку передається одним типом, використовуються ті самі канали та комутаційне устаткування;

2) підвищується завадостійкість;

3) зменшуються вимоги до комутаційного обладнання, використовуваного в мовному тракті;

4) з'являється можливість безпосереднього введення мовної інформації в комп'ютер.

Ці якості імпульсних систем не даються «даремно» – складність устаткування таких систем значно збільшується порівняно із загальними аналоговими системами телефонного зв'язку і значно розширюється смуга частот, яку використовують для передавання неперервної (аналогової) мови.

Розглянемо стисло основні способи, які застосовують для перетворення мови в послідовність імпульсів.

Для передавання сигналів по лініях, або трактах зв'язку їх належить перетворювати у форму, яка враховує властивості середовища переда-

вання. Цей процес отримав назву «модуляція». У пункті приймання виконується обернене перетворення – демодуляція.

Розглянемо деякі види модуляції, пов'язані з перетворенням аналогових сигналів у послідовність імпульсів.

6.4.1. Амплітудно-імпульсна модуляція

Дискретизація мовного сигналу шляхом використання амплітудно-імпульсної модуляції (АІМ) оснований на теоремі Котельникова.

Реальні сигнали завжди мають кінцеву тривалість і обмежену смугу частот. Граничні частоти спектра визначаються властивостями системи передавання та самим отримувачем. Так, для передачі дискретних повідомлень смугу частот визначають за швидкістю передавання.

Для функцій з обмеженим спектром В. О. Котельников довів чудову теорему, яка лежить в основі дискретизації неперервних сигналів. Згідно з цією теоремою функція $f(t)$, яка не містить частот, вищих за F , повністю визначається послідовністю своїх значень у моменти, які віддалені один від одного на $\Delta t = 1/2F$ (у секундах). Це доведення відкрило можливість замість того, щоб передавати повністю неперервний мовний сигнал, передавати лише послідовність імпульсів, які віддалені один від одного на відстань Δt і мають величину, що відповідає значенням мовного сигналу в моменти відліку. Процес утворення амплітудно-модульованих імпульсів показано на рис. 6.11.

Для пояснення процесу застосування АІМ у телефонному зв'язку на рис. 6.12 показано схему розмовного тракту з АІМ.

На мовному вході розмовного тракту є фільтр низьких частот (ФНЧ), який обмежує спектр сигналу з метою зведення його до умов теореми Котельникова. Через цей фільтр проходить вхідний мовний сигнал $U_1(t)$. Далі цей сигнал поступає на електронний контакт ЕК1, який керується частотою $f = 1/\Delta t = 2f_{\max}$. Якщо мовного сигналу на вході немає, а на ЕК1 ввімкнено сигнал керування, то на виході його буде вихідна імпульсна послідовність $U_1(t)$. Якщо спектр вхідний, обмежений частотою 400 Гц, то отримаємо:

$$\Delta t = \frac{10^6}{2 \cdot 4000} = 125 \text{ мкс}; \quad f = \frac{1}{125} \cdot 10^6 = 8000 \text{ Гц.}$$

За одночасної дії на ЕК1 імпульсів і мови виконується модуляція амплітуди імпульсів і на виході ЕК1 з'являється амплітудно-модульована імпульсна послідовність $U_3(t)$. Отже, ЕК1 виконує одночасно функції клапана, який пропускає мову в фіксовані моменти, розділені проміжками Δt , і модулятора імпульсів, що змінює амплітуду імпульсів відповідно до величини мовного сигналу.

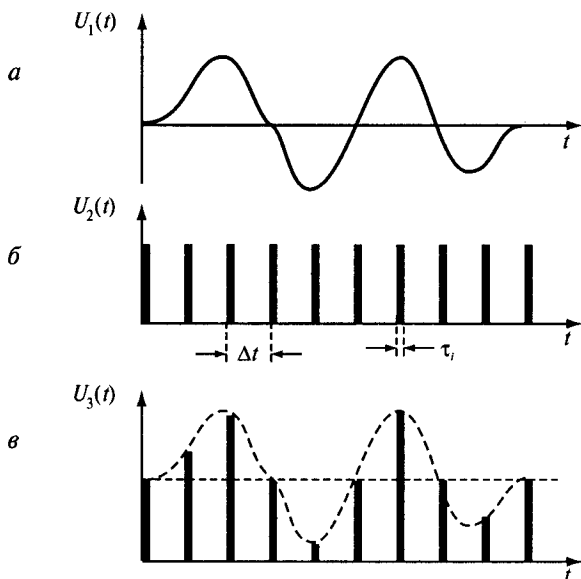


Рис. 6.11. Перетворення мовного сигналу в дискретну форму АІМ:
a – початковий мовний пристрій; *б* – початкова імпульсна послідовність;
в – амплітудно-модульована імпульсна послідовність

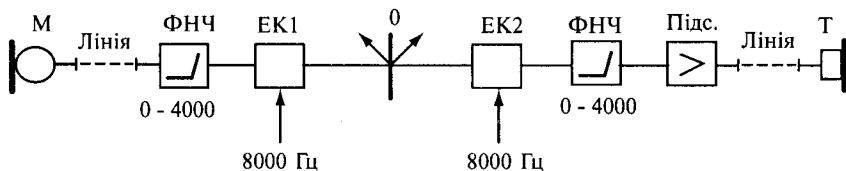


Рис. 6.12. Схема розмовного тракту з АІМ

Тракт приймання має також електронний контакт ЕК2, який працює синхронно з ЕК1 і забезпечує періодичне з'єднання приймального апарата з передавальним, і ФНЧ, що виконує функції демодулятора. На виході цього фільтра знову виділяється низькочастотна складова модульованого сигналу. Рівень цього сигналу буде значно нижчим відносно вихідного мовного сигналу, оскільки енергія мовного сигналу в проміжках імпульсів не використовується, а тривалість імпульсів τ_i становить усього 0,5...10 мкс. Для компенсації цих витрат застосовують підсилювачі.

Амплітудно-імпульсна модуляція малостійка відносно амплітудних і фазових спотворень, а також шумових завад у каналах зв'язку.

Для імпульсної передачі мови каналами зв'язку застосовують імпульсно-кодову модуляцію (ІКМ).

Імпульсно-кодова модуляція від АІМ відрізняється тим, що вводиться ще одна операція над імпульсом під час його передавання по каналу зв'язку – операція квантування амплітуди електричного сигналу. Квантування – це ділення миттєвої амплітуди на деяку кількість рівнів (рівнів квантування). Кожний рівень кодується двійковим числом. Двійкове число присвоюється рівню заздалегідь відповідно до шкали квантування.

Шкала квантування рівнів може бути лінійною, логарифмічною або побудованою за будь-яким іншим законом.

Кількість рівнів шкали квантування вибирають згідно з вимогами точності відтворення реальної часової характеристики сигналу.

Для реалізації ІКМ схема тракту передавання мовного сигналу розширюється: до виходу електричного контакту ЕК1 приєднується кодувальний пристрій (див. рис. 6.12), який має набір кодових комбінацій відповідно до рівня сигналу (рис. 6.13).

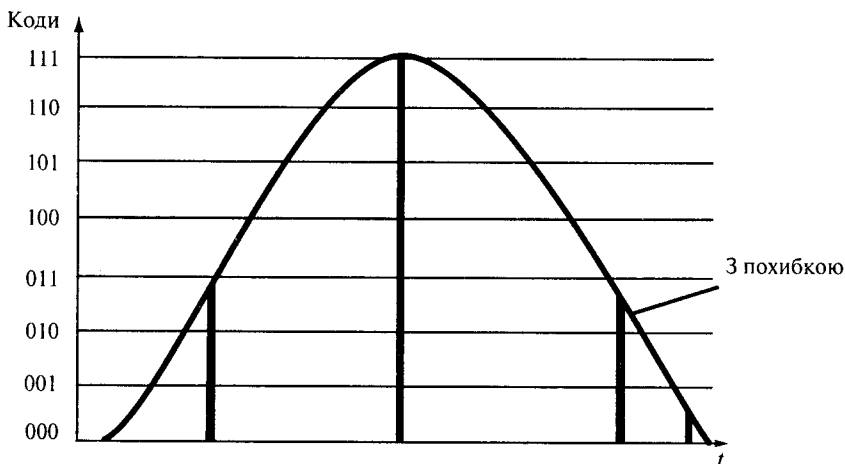


Рис. 6.13. Квантування амплітуди електричного сигналу

Задовільним відтворенням характеристик мовного сигналу та його перетворення кодувальним пристроєм вважається використання 64 фіксованих рівнів ($N = 64$):

$$n = \log_2 N = \log_2 64 = 6,$$

тобто шестирозрядне двійкове число.

Загальна швидкість передавання імпульсів для одного мовного повідомлення у вигляді ІКМ-сигналу

$$R = f \log_2 N = 8000 \cdot 6 = 48\,000,$$

де f – частота квантування за часом.

Ширину смуги частот, потрібну для передавання послідовності імпульсів визначають за теоремою Котельникова:

$$\Delta f \geq \frac{1}{2f_{\max}}.$$

З попереднього матеріалу очевидно, що якість передавання мовних звукових сигналів у цифровому вигляді залежить від фіксованої величини рівня квантування: чим більший рівень квантування, тим якісніше за показником розбірливості передавання мовних повідомлень каналами телефонного зв'язку, що, в свою чергу, потребує розширення смуги частот каналу.

Отже, у другій частині розглянуто принципові етапи створення системи телефонного зв'язку: теорія формування мовних звукових сигналів, що утворюють телефонне повідомлення, будується на адекватних фізико-математичних моделях фізичної оптики; процеси перетворення мовних звукових сигналів у електричні і навпаки – електричних сигналів у звукові – обґрунтовуються фізичними явищами, у яких пропорційно змінюються параметри, зумовлені механічними діями (зміна положення координат стану елементів) та законами електродинаміки. Принципи перетворення однієї форми енергії (звукової, електричної) у другу реалізуються у різноманітних термінальних пристроях телефонної системи зв'язку. Методи перетворення аналогових (неперервних) природних мовних звукових сигналів в електричній формі їх подання в дискретну (цифрову) форму дозволяють створювати сучасні технології передавання дискретної інформації каналами телефонного зв'язку (їх розглянуто у другій частині підручника).

Контрольні питання

1. Системи передавання та обміну повідомленнями. Функції і задачі.
2. Термінальне устаткування: призначення, функції і класифікація.
3. Аналогові системи передавання мови людини.
4. Хвильова теорія подання мовних звуків.
5. Звук і його основні властивості.
6. Енергетична характеристика звуку, тиск та інтенсивність звуку.
7. Зв'язок між інтенсивністю і тиском звуку.
8. Класифікація звуків. Часові і спектральні характеристики звуків.
9. Спектральна характеристика неперіодичних звуків.
10. Рівні інтенсивності та звукового тиску простих і складних періодичних звуків.
11. Рівні спектрів складних неперіодичних звуків.
12. Складання рівнів звуків від декількох джерел.
13. Шуми та їх характеристики.
14. Механізм утворення мовних звуків. Форманти і фонемні звуки.
15. Інтегральні характеристики мови. Спектр мови і формантний спектр.
16. Динамічний діапазон мови. Формантний розподіл в інтегральній формі.
17. Механізм слухового сприйняття мови.
18. Характеристики слуху. Поріг чутності та больовий поріг.
19. Рівень відчуття. Критична смуга слуху. Ефективний рівень відчуження.
20. Адаптація слуху. Маскування звуків.
21. Перетворення мовних звуків у електричні сигнали. Загальні принципи побудови перетворювачів.
22. Електромагнітна система перетворювачів.
23. Електродинамічна система перетворювачів.
24. Перетворювачі конденсаторної системи.
25. Перетворювачі п'єзоелектричної системи.
26. Вугільні перетворювачі.
27. Мікрофон. Системи перетворення. Експлуатаційні характеристики.
28. Ларингофон. Принцип дії. Система перетворення. Експлуатаційні характеристики.
29. Телефон. Системи перетворення. Експлуатаційні характеристики.
30. Телефонні апарати. Класифікація. Викличні та розмовні пристрої.
31. Розділення викличного та розмовного ланцюгів.
32. Місцевий ефект. Протимісцеві схеми.
33. Основні елементи телефонних апаратів.
34. Типова схема дискового номеронабирача. Діаграма струму в лінії.
35. Принцип дії викличних пристроїв.
36. Електроакустичні та електричні параметри телефонних апаратів.
37. Тастатурний номеронабирач із частотним кодуванням номерів.
38. Оцінка якості передавання мови. Вимоги до якості передавання мови.
39. Метод оцінки якості передавання мови – порівняння за гучністю.
40. Метод артикуляції. Види розбірливості.
41. Метод вимірювання еквівалента затухання за розбірливістю (AEN).
42. Електроакустичний розрахунок мовного тракту. Постановка завдання. Вихідні дані.
43. Метод розрахунку розбірливості за параметрами тракту.

44. Визначення коефіцієнта сприйняття як функції ефективного рівня відчуття формант.
45. Загальний алгоритм розрахунку розбірливості мови.
46. Методи перетворення мовних звукових сигналів.
47. Тривимірний модель телефонного сигналу. Використання методів перетворення звукових сигналів на цій моделі.
48. Методи частотного компандування.
49. Принцип дії частотного вокодера.
50. Частотні діаграми процесу перетворення мови під час параметричного компандування.
51. Смуговий вокодер частотного параметричного компандування.
52. Амплітудне перетворення мови.
53. Перетворення мовного сигналу в дискретну форму.
54. Амплітудно-імпульсна модуляція.
55. Імпульсно-кодова модуляція.
56. Визначення потрібної ширини смуги частот для $N = 64, 128, 256$ рівнів квантування.

Термінальне устаткування дискретної дії та передавання дискретної інформації каналами телекомунікаційних мереж

7. Передавання дискретної інформації каналами телекомунікаційних мереж

У цій частині розглядаються термінальні устаткування дискретної дії. Організація зв'язку між ними утворюється пристроями передавання дискретної інформації і каналами тональної частоти телефонних мереж електров'язку. До такого термінального устаткування належать: термінали-комп'ютери, факсимільні апарати, електронні апарати телеграфного та фототелеграфного зв'язку. Розглянемо передавання дискретної інформації трактами обміну між кінцевим устаткуванням телекомунікаційних мереж.

7.1. Основні поняття та визначення

Вивчаючи телекомунікаційні системи і мережі, доволі часто вживають поняття «інформація» і «повідомлення». Ці близькі за змістом поняття складні і дати їх точне визначення через більш прості нелегко. Слово «інформація» походить від латинського *informatio* – роз'яснення, ознайомлення, поінформування. Зазвичай під інформацією розуміють сукупність відомостей, даних про будь-які події, явища або предмети. Сукупність відомостей, даних стає знанням лише після їх інтерпретації з урахуванням цінності та змісту цих відомостей. Отже, інформацію в широкому розумінні можна визначити як сукупність знань про оточуючий нас світ.

Отже, під інформацією розуміють певні властивості матерії, які сприймаються спостерігачем із навколишнього матеріального світу. При цьому спостерігачем може бути людина, живий організм або технічний пристрій.

Якщо інформації надати певну форму, тоді її називають повідомленням. Форми подання інформації дуже різноманітні: письмовий текст, мова, музика, графічне зображення, електромагнітне поле, взаємне розміщення предметів у просторі і т. ін. Загальне завдання системи зв'язку полягає у передаванні повідомлень від людини до технічного пристрою, яка спостерігає деяку матеріальну систему, другій людині або технічному пристрою, яка не може безпосередньо спостерігати за матеріальною системою, що їх цікавить.

Спостережувану систему разом із спостерігачем називають джерелом інформації, а людину або технічний пристрій, яким передають результати спостережень, – одержувачем інформації.

Джерело інформації вибирає повідомлення із деякої множини повідомлень, яке характеризує властивості спостережуваної системи відносно її стану. Вибране повідомлення передається одержувачу інформації.

Ці повідомлення можна подавати у вигляді функції часу $x(t)$. Функція $x(t)$ може бути неперервною у часі і описувати мову людини. Функція $x(t)$ може бути дискретною у часі і описувати дискретні відомості. Це може бути друкований текст, який складається зі знаків алфавіту, що має кінцеву кількість символів (букв, цифр, розділових знаків), а також подання мови людини в цифровому вигляді.

Для передавання інформації від джерела одержувачу потрібно мати будь-який переносник, який мав би достатню стійкість своїх параметрів у просторі та часі, щоб була можливість виявити його на приймальному.

Фізична природа і характеристики переносника пов'язані з певним фізичним середовищем між джерелом і одержувачем інформації – каналом зв'язку. Для утворення каналів електричного зв'язку використовують природне фізичне середовище (нарколоземний і космічний простори) або штучно створені середовища (провідникові, хвилевідні або світловодні лінії зв'язку).

Поняття «канал» містить не тільки фізичне середовище, в якому поширюється сигнал, але й сукупність технічних засобів, призначених для передавання сигналів.

У каналах електричного зв'язку переносником є електромагнітне поле або електричний струм. Для того, щоб переносник міг виконувати свою функцію – передавання інформації, його характеристики мають змінюватись. Щоб переносник став моделлю повідомлення, між його характеристиками і повідомленням має існувати однозначна, заздалегідь встановлена закономірність – інформаційна відповідність.

Переносник, параметри якого знаходяться в однозначній інформаційній відповідності з повідомленням, називають сигналом.

Незалежно від фізичної суті сигналу можна розглядати множину сигналів як множину деяких функцій $z(t)$. Для дискретних повідомлень із множини A множина сигналів Z також дискретна. Для відновлення за прийнятим сигналом переданого повідомлення має бути однозначна відповідність між кожним із можливих сигналів множини Z і відповідним повідомленням із множини A .

Процес передавання інформації полягає в тому, що повідомлення від джерела інформації перетворюється в сигнал, який передається каналом зв'язку одержувачу (споживачу). Споживач, який отримав цей сигнал, на основі знань закону відповідності сигналів та повідомлень «витагує» вміщену в повідомленні інформацію. Канал зв'язку разом з джерелом інформації і її одержувачем за заданих методів перетворення повідомлень у сигнал і відновлення повідомлення прийнятим сигналом називають системою зв'язку. Схему системи зв'язку в загальному вигляді показано на рис. 7.1.

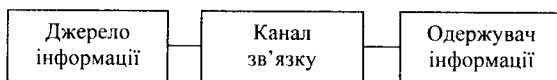


Рис. 7.1. Схема системи зв'язку в загальному вигляді

7.2. Перетворення повідомлень у сигнал

Систему перетворення повідомлення в сигнал на передавальному кінці та обернене перетворення на приймальному можна подати у вигляді таблиці (кодової), у якій кожному повідомленню із кінцевої множини A зіставлено певні сигнали з кінцевої множини Z . Перетворення повідомлення в сигнал у теорії інформації називають кодуванням.

У разі дуже великих розмірів множини A , яка вміщує так багато різних повідомлень, то скласти і зберегти кодову таблицю майже неможливо. Тому такий спосіб складання кодових таблиць використовують тільки для обмеженої кількості елементів множини A .

Для більшості випадків замість безпосереднього складання кодових таблиць виконують розбиття всіх повідомлень із множини A на послідовність елементарних повідомлень або символів (букв), які утворюють кінцеву множину X , що вміщує невелику кількість елементів. Оскільки всі можливі повідомлення дискретного джерела утворюють злічену множину і можуть бути пронумеровані, то ці числа можна виразити в зручній системі числення і брати кожну цифру номера як елементарне повідомлення.

Отже, кожне повідомлення із множини A можна подати у вигляді послідовності елементів x множини X : $a = (x^{(1)}, x^{(2)}, \dots, x^{(L)})$. Тут L позначає

довжину повідомлення в абетці (алфавіті) X . При цьому складання таблиці значно спрощується і зводиться до зіставлення дуже обмеженої кількості елементарних повідомлень множини X з відповідними їм сигналами із множини Z . Що являє собою елементарне повідомлення? У більшості випадків замість безпосереднього складання кодових таблиць виконують розчленування всіх повідомлень з множини A на послідовність елементарних повідомлень або символів (букв), які утворюють кінцеву множину X , що містить обмежену кількість елементів.

У разі великого обсягу алфавіту символів X зазвичай удаються до додаткового перетворення: переходу від алфавіту символів X об'ємом L до нового алфавіту символів Y , який має об'єм $k < L$.

Правило перетворення елементарних повідомлень алфавіту X (чисел, букв, розділових знаків) у символи алфавіту Y і навпаки називають кодом.

Кодування у вузькому розумінні слова (будемо виключно користуватись ним надалі) визначають як відображення дискретних повідомлень послідовністю вибраних символів. Метод кодування обумовлюється не тільки особливостями джерела інформації, але інколи і властивостями каналу зв'язку.

У разі використання двійкового коду з основою 2 модульований параметр може набути лише двох фіксованих дискретних значень – значущих позицій (значень постійного струму або напруги).

Дискретну модуляцію інколи називають маніпуляцією.

Границі між одиничними елементами, на яких виконуються зміни значущих позицій, називають значущими моментами (на рис. 7.2 показано буквою a).

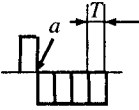

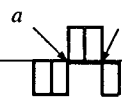
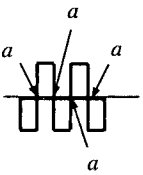
| Символ первинного алфавіту | Е | К | Н | Р |
|----------------------------|---|--|---|---|
| Двійкове число | 10000 | 11110 | 00110 | 01010 |
| Кодова комбінація |  | Одиничні елементи  |  |  |

Рис. 7.2. Приклади комбінацій п'ятиелементного коду (переносник – постійний струм)

Кількість одиничних елементів, які передаються за 1 с, називають швидкістю модуляції і визначають як

$$B = \frac{1}{T}.$$

Одиниця швидкості модуляції – бод (Бод) – швидкість, яка відповідає передачі одного одиничного елемента за секунду.

Передавання інформації характеризують також швидкістю передавання інформації, під якою розуміють кількість інформації, яка поступає через систему передавання від джерела інформації до одержувача за одну секунду. Її вимірюють кількістю бітів за секунду (біт/с). Швидкість передавання залежить від швидкості модуляції, статичних властивостей джерела, типу каналу, системи сигналів, завад у каналі, методів приймання. Швидкість модуляції не можна плутати зі швидкістю передавання інформації, а бод ототожнювати з біт/с. Кількісно ці характеристики збігаються тільки для симетричних двійкових каналів без завад.

У симетричному каналі без завад у випадку $k_m > 1$, $V = Bk_m$, де k_m – кратність модуляції, тобто кількість бітів інформації, яку переносить одиничний елемент сигналу за одну секунду. Наприклад, за чотирирівневого передавання $k_m = 2$, $V = 2B$. У загальному випадку швидкість передавання інформації може бути більшою або меншою від швидкості модуляції.

Один із головних параметрів коду – його основа, яка дорівнює основі вибраної системи числення і відповідає об'єму кодового алфавіту, тобто кількості різних символів у коді k . Кожному елементарному повідомленню (символу) відповідає деяка послідовність кодових символів, яку називають кодовою комбінацією.

Отже, результатом кодування дискретної інформації є подання її у вигляді послідовності чисел – у цифровій формі.

Найбільшого застосування в техніці передавання дискретної інформації набули коди з основою 2, які називають двійковими або бінарними. Широке використання двійкових кодів зумовлено: простотою реалізації, надійністю, швидкою дією елементів двійкової логіки, їх малою чутливістю до дії на них зовнішніх завад, зміною напруги живлення, а також простотою двійкової арифметики. Тому надалі будемо користуватись і двійковими кодами.

Елементами двійкового коду є символи: одиниця (1) і нуль (0). Символи, які застосовують для подання даних у двійковій системі числення (наприклад, 0 або 1), називають бітами.

Кодування повідомлень виконує спеціальний пристрій, який входить до складу джерела інформації; його називають кодером. Відповідно в одержувача інформації є другий пристрій – декодер. Призначення деко-

дера – ототожнення кожної прийнятої кодової комбінації сигналу з відповідним її елементом повідомлення (символом).

Процес формування сигналу (перетворення повідомлення в сигнал) вклучає, окрім кодування, операцію модуляції.

Під дискретною модуляцією розуміють процес перетворення послідовності кодових символів у послідовність електричних сигналів. Обернений процес відповідно називають демодуляцією.

Модуляція полягає в зміні того чи того або одночасно декількох параметрів переносника. У каналах зв'язку переносник використовує найчастіше синусоїдні коливання або періодичну послідовність імпульсів.

Зазначимо, що в кодерах джерела переносник дискретного повідомлення використовує лише постійний струм.

Кожний елемент коду в процесі модуляції перетворюється у відрізок переносника тривалістю T , який називають одиничним елементом.

7.3. Кодування джерелом інформації

Одне з призначень кодера джерела – подання повідомлення у вигляді послідовності кодових комбінацій. Код, у якому повідомлення видаються із кодера джерела, називають первинним на відміну від вторинного коду, який використовують для передавання повідомлення каналами зв'язку.

Розрізняють два способи виведення кодових комбінацій із джерела повідомлень: послідовний і паралельний. У кодерах першого типу символи кодових комбінацій виводяться одним ланцюгом послідовно один за одним. У кодерах другого типу всі символи однієї кодової комбінації виводяться одночасно; для цього використовується стільки електричних ланцюгів, скільки розрядів має код.

Залежно від способу обмінювання інформацією джерела з каналом зв'язку розрізняють некеровані і керовані джерела повідомлень.

Джерела першого типу видають повідомлення у канал в деякі моменти часу, які визначаються самим джерелом і не залежать від каналу зв'язку. Так працює телефонний зв'язок: якщо встановити зв'язок, то джерело інформації (розмовляюча людина і мікрофон) видає повідомлення залежно тільки від моменту його появи.

Джерела другого типу повідомлення запам'ятовують і зберігають, а в канал зв'язку видаються визначеними порціями на вимогу (запит) передавального пристрою каналу (передавача). Прикладом керованого джерела повідомлень є термінал – комп'ютер, який накопичує повідомлення за вказівкою оператора-передавача або на запит оператора-одержувача, а потім передає їх у канал згідно із вказівкою передавача.

У цій частині курсу будемо вивчати теоретичні положення організації дискретних сигналів, які використовуються джерелами другого типу (джерела першого типу розглянуто у першій частині курсу).

Перетворення дискретного повідомлення в сигнал полягає у виконанні двох операцій: кодування та модуляції.

Джерелами інформації є термінали: комп'ютер, факсимільний апарат, електронний ТА.

Кодування визначає закон побудови сигналу.

Найпростіший приклад дискретного повідомлення – текст. Будь-який текст складається з обмеженої кількості елементів: букв, цифр і розділових знаків. Оскільки їх кількість для будь-якої мови обмежена, то їх доцільно перенумерувати і звести передавання повідомлення до передавання послідовності чисел.

Будь-яке число з основою m можна подати у вигляді

$$N = \dots + a_2 m^2 + a_1 m^1 + a_0 m^0,$$

де $a_0, a_1, a_2 \dots$ – коефіцієнти, які набувають значень від 0 до $m - 1$.

Так, для десяткової системи числення, наприклад число 265, можна подати у такому вигляді:

$$265 = 2 \cdot 10^2 + 6 \cdot 10^1 + 5 \cdot 10^0.$$

Якщо задати значення m (будь-яке), то отримаємо будь-яку систему числення. Так, якщо $m = 2$, матимемо двійкову систему, у якій числа записуються за допомогою лише двох знаків: 0 і 1. Наприклад, число 8 у двійковій системі записується як 1000, що відповідає виразу $1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0 = 8$.

Розрізняють коди рівномірні та нерівномірні. Прикладом рівномірного коду є код Бодо (п'ятизначний, табл. 7.1), а нерівномірного – код Морзе (табл. 7.2).

Пояснення утворення кодів наведено в табл. 7.1 і 7.2.

У кодї Морзе сигнал, який відповідає одній одиниці, називають точкою, трьом одиницям – тире. Нуль використовують для розділення точки і тире, а три нулі – для розділення знаків.

Таблиця 7.1

| Елементи повідомлення | Код | Сигнал |
|-----------------------|-------|--------|
| А | 10000 | |
| Б | 00110 | |
| В | 01101 | |
| Г | 01010 | |

Таблиця 7.2

| Елементи повідомлення | Код | Сигнал |
|-----------------------|-------|--------|
| А | • — | |
| Б | — ••• | |
| Е | • | |
| Т | — | |

7.4. Канали передавання дискретної інформації

Система зв'язку для передавання дискретної інформації – це сукупність каналу передавання дискретної інформації, джерела дискретної інформації і її одержувача із заданими методами перетворення дискретного повідомлення в сигнал і відновлення дискретного повідомлення згідно з прийнятим сигналом. Схему системи зв'язку для передавання дискретної інформації в загальному вигляді показано на рис. 7.3.

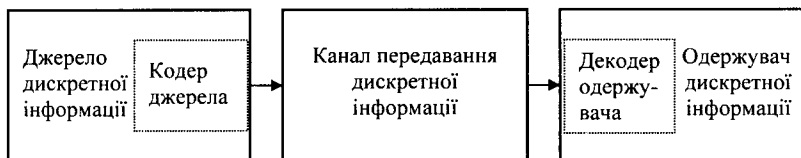


Рис. 7.3. Узагальнена схема зв'язку для передавання дискретної інформації

Тут під каналом передавання дискретної інформації розуміють сукупність технічних засобів і фізичного середовища, призначених для відповідного перетворення та передавання сигналів.

Кодер джерела повідомлень перетворює послідовність елементів повідомлення в послідовність кодових елементів і кожному кодовому символу ставить у відповідність певний елемент сигналу. Отже, первинний сигнал на виході кодера джерела являє собою деяку послідовність одиничних елементів постійного струму. Такий первинний сигнал здебільшого не може бути безпосередньо переданий каналом зв'язку. Так, наприклад, у дискретно-неперервних каналах виникає потреба спектрального узгодження характеристик дискретного джерела інформації і неперервного переносника в каналі, оскільки основна частина енергії послідовності двійкових імпульсів припадає на діапазон низьких частот, що не пропускається каналом.

Спектральне узгодження, а також узгодження за амплітудою сигналів джерела і каналу виконує пристрій перетворення сигналів (ППС) – один з основних складових передавача і приймача, тобто пристрій, що перетворює сигнали джерела в лінійні сигнали і лінійні сигнали в сигнали одержувача інформації.

Первинний сигнал, що виробляється джерелом, перетворюється ППС передавача у лінійний сигнал (сигнал у каналі). Цей процес називають модуляцією переносника в каналі зв'язку. Він полягає у зміні тих чи тих параметрів переносника (амплітуди, частоти, фази і т. ін.) відповідно до зміни первинного сигналу. Відповідно обернена операція – перетворення лінійного сигналу в первинний, який сприймається декодером одержувача інформації – здійснюється в ППС приймача і називається демодуляцією.

Проміж ППС на боці джерела інформації і ППС на боці одержувача інформації знаходиться фізичне середовище, яке має забезпечити доставляння сигналу каналом зв'язку. Процес передавання сигналу каналом зв'язку здійснюється завдяки фізичному явищу – поширенню сигналу на довгих лініях.

Під час передавання дискретних сигналів каналами зв'язку виникають перехідні процеси, зумовлені обмеженням частотних спектрів пере-

давальних сигналів, нелінійністю фазочастотної характеристики (ФЧХ) і несталістю амплітудно-частотної характеристики (АЧХ) каналу. Перехідні процеси призводять до спотворення форми сигналу, а також збільшують його тривалість. У результаті виникає часткове взаємне перекриття сусідніх передавальних сигналів. Їх називають міжсимвольними спотвореннями. Зовнішні завади утворюють електромагнітне поле, яке разом з електромагнітним полем передавального (корисного) сигналу створюють нове загальне електромагнітне поле, яке за формою безперечно відрізняється від поля корисного сигналу, що формує сигнал. Цей сигнал буде відрізнятися від сигналу, який надсилається в канал зв'язку ППС передавача.

Сукупність усіх причин, які викликають невизначеність приймального сигналу взагалі, називають *шумами*, або *завадами*. Завади можуть бути різними: завади, які спотворюють сигнал, але методи боротьби з ними дозволяють відновити первинний сигнал, і завади, які так спотворюють сигнал, що його неможливо відновити методами оброблення сигналів і їх відносять до помилок. Наявність помилок у каналі потребує вжиття заходів щодо їх зменшення.

Для цього існують два основні способи запобігання цим явищам:

- використання таких систем сигналів у каналах і таких методів приймання, які б забезпечили найменшу ймовірність помилок;
- виправлення помилок використанням для передавання сигналів каналами зв'язку завадостійких кодів.

Завадостійкі коди відрізняються від простих безнадлишкових кодів, мають надлишковість, використовуювану для виявлення і виправлення помилок. Підвищення правильності передавання сигналів використанням завадостійких кодів потребує введення в передавач спеціальних пристроїв, які б перекодували первинні сигнали (кодові комбінації), що надходять від кодера джерела. Такі пристрої мають загальну назву «кодери каналу», що являють собою основну складову пристрою захисту від помилок передавача (ПЗПпер).

Відповідно у приймачі декодер каналу виконує операцію виявлення і виправлення помилок за допомогою надлишковості, яка вміщується у завадостійкому коді. Декодер каналу є складовою частиною ПЗПпр. Кодові комбінації, звільнені від надлишковості, являють собою уже прості кодові комбінації, які надходять у кодер одержувача інформації.

Основні операції перетворення інформації для обміну між користувачами (абонентами) розглянуто в підрозд. 7.3. Тепер для деталізації процесів перетворення сигналів у системі передавання дискретної інформації та функцій, які виконують складові частини і пристрої, що реалізують їх, розглянемо її структурну схему (рис. 7.4).



Рис. 7.4. Структурна схема системи передавання дискретної інформації

Функції, які виконують складові СПДІ, а також пристрої (вузли), що реалізують ці функції, полягають в такому:

1. Джерело інформації своїм внутрішнім кодером перетворює інформацію в двійкову (бінарну) форму (для найбільш поширеного часткового випадку дискретного каналу).

2. Пристрій захисту від помилок під час передавання інформації каналом зв'язку виконує функцію зменшення помилок, спричинених дією завад на корисні сигнали. Для цього є два способи:

1) застосування таких сигналів у каналах і таких методів приймання, які забезпечують якнайменшу ймовірність помилки;

2) виправлення помилок використанням для передавання сигналів каналами зв'язку завадостійких кодів.

3. Пристрій перетворення сигналів передавальною частиною системи (ППСпер) виконує функцію перетворення сигналів, які виникають на виході ПЗПпер (це двопозиційні сигнали вихідного тригерного регістра у вигляді постійних напруг), у керованому послідовність імпульсів, які модулюються одним з видів модуляції та забезпечують енергетичний рівень цих сигналів.

4. Канал зв'язку, на який діють завади, виконує функцію поширення елементарних сигналів, які спотворюються характеристиками каналу та параметрами діючих завад.

5. Пристрій перетворення сигналів приймальною частиною системи (ППСпр) виконує функції ресстрації та синхронізації. Ресстрація приймальних одиночних сигналів полягає у визначенні самого сигналу значущими позиціями та значущими моментами і видача кодових комбінацій ПЗПпр.

6. Пристрій захисту від помилок приймальної частини (ПЗПпр) виконує операцію виявлення та виправлення помилок за допомогою надмірності, яка міститься в завадостійкому коді.

7. Декодер-одержувач отримує від ПЗПпр послідовність кодових комбінацій (символів) первинного (ненадмірного) коду, де перетворюється в послідовність символів повідомлення, яка видається одержувачу (користувачу).

Зазначимо таке:

- об'єднання ПЗПпер і ППСпер називають передавачем СПДІ, ППСпр і ПЗПпр–приймачем;
- частину схеми СПДІ, яка охоплює ППСпер, канал зв'язку і ППСпр, називають дискретним каналом;
- дискретний канал, доповнений пристроями ПЗПпер і ПЗПпр, називають каналом передавання дискретної інформації.

Отже, поняття «канал» далеко не однозначне. Тому каналом зазвичай називають ту частину системи передавання дискретної інформації, яку за умовами розв'язуваного завдання неможливо або небажано змінювати.

7.5. Вимоги до системи передавання дискретної інформації

Системою передавання дискретної інформації називають сукупність джерел і одержувачів повідомлень, передавачів, приймачів і каналів зв'язку, що забезпечують передавання дискретної інформації з певними властивостями.

Найважливіші зовнішні характеристики системи такі:

- правильність передавання інформації;
- швидкість передавання інформації;
- надійність системи передавання інформації.

Важливість цих характеристик полягає в тому, що недостатня правильність робить прийняту інформацію безкорисною, а інколи й шкідливою. Не менше значення має і швидкість передавання інформації, яка впливає на своєчасність доставки повідомлення одержувачу. Показник надійності системи означає здатність її працювати протягом тривалого часу із задоволенням вимог до перших двох її характеристик.

Не слід забувати, що отримання високих показників правильності, швидкості та надійності СДПІ завжди пов'язано з її вартістю.

До внутрішніх характеристик СПДІ належать ефективна пропускна здатність і завадостійкість. Ефективна пропускна здатність – це відношення пропускної здатності системи до пропускної здатності каналу, яке характеризує ступінь використання пропускної здатності каналу.

Завадостійкістю називають здатність системи протистояти шкідливій дії завади. Ця величина не має абсолютної оцінки, тому зазвичай користуються порівняльною оцінкою завадостійкості різних систем. Пропускна

здатність системи суперечить завадостійкості системи. Тому в практичній діяльності доводиться відшукувати компромісні рішення. Це і є основне завдання теорії передавання дискретної інформації.

Для пошуку шляхів отримання цих розв'язків зведемо всі послідовності операції перетворення кодових комбінацій і сигналів, які виконують окремі пристрої – складові СПДІ, у математичний опис загального алгоритму роботи СДПІ, використовуючи при цьому структурну схему (див. рис. 1.3).

7.6. Математичний опис загального алгоритму роботи системи передавання дискретної інформації

Для кращого уявлення про систему перетворення сигналів у СПДІ і виявлення як теоретичних, так і інженерних завдань є відомий шлях – створення математичної моделі, у якій відображено основні закономірності процесів перетворення. Відомо, що створення математичної моделі, яка описує адекватно поведінку об'єкта, що вивчається, ґрунтується на детальному поданні окремих складових загального процесу у вигляді функціональних залежностей, які потім підлягають системному аналізу з виявленням несуттєвих параметрів і врахуванням їх впливу на загальний процес. Процес побудови формалізованої математичної моделі – багаторівневий процес. Перший ступінь цього процесу – формування загального алгоритму функціонування у часі на вербальному рівні та введення (в цьому разі операторів) процедур перетворення сигналів.

Джерело дискретної інформації створює на виході свого кодера випадкову дискретну послідовність $\{y_i\}$. Унаслідок перетворень у ПЗПпер ця послідовність на його виході перетворюється в дискретну випадкову послідовність $\{y_{ki}\}$, яка має більший розмір завдяки введенню надмірності. У результаті модуляції у ППСпер утворюється послідовність сигналів $\{z_{ki}\}$, яка однозначно відповідає випадковій дискретній послідовності $\{y_{ki}\}$.

Сукупність операцій перетворення повідомлення в сигнал називають *способом передавання*. Спосіб передавання можна подати співвідношенням

$$\{z_{ki}(t)\} = U_{\text{пер}}\{y_i\} = U_{\text{мод}}U_{\text{код}}\{y_i\},$$

де $U_{\text{пер}}$ – оператор способу передавання; $U_{\text{мод}}$ – оператор модуляції; $U_{\text{код}}$ – оператор кодування.

Сигнал передається каналом зв'язку, у якому зменшуються його енергетичні параметри внаслідок поглинання енергії і спотворення через несталість параметрів передавання у часі, неоднорідностей нелінійності частотних характеристик і т. ін. Тому сигнали, які надходять на вхід

приймача $\{z_{ni}(t)\}$, можуть значно відрізнятись від переданих сигналів $\{z_{ki}(t)\}$ як своєю енергією, так і частотними і спектральними характеристиками:

$$\{z_{ni}(t)\} = U_{\text{кан}}\{z_{ki}(t)\} = U_{\text{пер}}U_{\text{кан}}\{y_i\},$$

де $U_{\text{кан}}$ – оператор каналу.

Оскільки в каналі діє завада, яка є випадковим процесом $n(t)$, то на виході приймача діє неперервний випадковий процес

$$\xi(t) = \{z_{ni}(t)\} + n(t).$$

У приймачі після попереднього підсилення з метою компенсації загасання сигналу в каналі виконується демодуляція сигналу в ППСпр і на виході створюється дискретна послідовність $\{z'_{ni}\}$, яка має відповідати послідовності $\{z_{ki}\}$ на виході ППСпер. Але через вплив завад і спотворень у каналі та похибок перетворення пристроями системи повної відповідності між $\{z'_{ni}\}$ і $\{z_{ki}\}$ можливо і не буде. У ПЗПпр і декодері одержувача виконується перетворення послідовності кодових імпульсів у послідовність кодових комбінацій $\{y'_i\}$, яка має відповідати переданій послідовності $\{y_i\}$. Повнота цієї відповідності залежить від властивості каналу, способів приймання, типу ПЗП та використовуваного завадостійкого коду. Дискретна послідовність $\{y'_i\}$ видається одержувачу.

Сукупність операцій перетворення сигналу в повідомлення називають *способом приймання*. Спосіб приймання характеризується співвідношенням

$$\{y_i\} = U_{\text{пр}}[\{z_{ni}(t)\} + n(t)],$$

де $U_{\text{пр}} = U_{\text{дем}}U_{\text{дек}}$ – оператор способу приймання; $U_{\text{дем}}$ – оператор демодуляції; $U_{\text{дек}}$ – оператор декодування.

Тоді весь процес передавання дискретної інформації можна подати так:

$$\{y'_i\} = U_{\text{дек}}U_{\text{дем}}[U_{\text{кан}}U'_{\text{мод}}U'_{\text{код}}\{y_i\} + n(t)].$$

Основне завдання побудови системи передавання дискретної інформації полягає в тому, щоб одержати послідовність $\{y'_i\}$, яка б якнайменше відхилялась від переданої послідовності $\{y_i\}$, забезпечивши при цьому високу швидкість передавання інформації, високу надійність і допустиму вартість.

У загальному вигляді вирішення завдання нашоветується на великі труднощі.

Будуючи систему передавання, завжди враховують заданий ансамбль повідомлень джерела, завади і параметри передавання каналу.

Загальне завдання поділяють на такі окремі два.

1. Оптимізація підсистеми «дискретний канал». У результаті вирішення цього завдання вибирають найкращий тип сигналу (вид перенос-

ника, спосіб модуляції) для передавання і відповідний йому найкращий спосіб приймання.

2. Оптимізація підсистеми «кодер – декодер» у припущенні, що дискретний канал задається матрицею трансформації множини $\{y_{ki}\}$ у множини $\{y_{ni}\}$.

Вирішення цих двох часткових завдань і є певною мірою вирішенням загального завдання. Але тільки до деякої міри, оскільки функції окремих ділянок системи передавання тісно переплітаються і оптимальна в цілому система передавання, можливо, не збігається із системою, яка складається з окремих отриманих оптимальних її ділянок. Тому компромісне вирішення загального завдання шукають на шляху оптимізації окремих ланок з урахуванням характеристик інших ланок.

7.7. Основні характеристики каналів зв'язку та їх вплив на передавання даних

7.7.1. Основні характеристики каналів зв'язку

Канали зв'язку утворюються за допомогою різних типів ліній зв'язку та різних систем передавання. У загальному випадку їх характеристики можуть значно відрізнитися одна від одної. Будемо розглядати некомутовані канали тональної частоти (ТЧ) (а для комутованих – тракти).

Характеристики каналів зв'язку значною мірою визначають основні показники якості передавання дискретної інформації: швидкість, правильність, час доставляння, надійність, ефективність. Тому знання цих характеристик потрібні для розробників і створювачів СПДІ.

Канал зв'язку розглянемо як перетворювач сигналів, на вхід якого поступає сигнал $S_{вх}(t) \rightarrow S_{вих}(t)$. Оскільки в реальних каналах ідеальна відповідність не витримується $S_{вх}(t) = S_{вих}(t)$, сигнал на виході каналу відрізняється від сигналу на його вході; при цьому вирізняють детерміновані та випадкові зміни.

Детерміновані зміни сигналу визначаються структурою каналів і полягають у деяких змінах масштабу сигналу (підсилення або загасання), зміщені у часі (затримка) і зміні форми (спотворення). Випадкові зміни сигналу викликаються завадою, яка діє в каналі, і полягають у випадковій зміні тих самих показників – масштабу, затримки, спотворення форми сигналу.

Із детермінованих змін сигналу найбільший інтерес для вивчення становлять спотворення, оскільки затримка у часі принципово не може бути зменшена, а зміни масштабу компенсуються підсиленням або послабленням сигналу за допомогою пристроїв автоматичного регулювання.

Спотворення форми сигналу можуть бути лінійними і нелінійними. До лінійних належать спотворення, спричинені АЧХ і ФЧХ, що відрізняються від ідеальних характеристик.

Амплітудно-частотну характеристику каналу ТЧ задають зазвичай частотною характеристикою залишкового загасання. Залишковим загасанням називають рівень між рівнями сигналів на вході та виході каналів зв'язку.

Фазочастотна характеристика каналу ТЧ задається нерівномірністю групового часу сповільнення і для зручності вимірювання оцінюється не за значенням фази, а за її похідною – груповим часом проходження.

Нерівномірність групового часу проходження призводить до того, що різні частотні складові спектра передавальних сигналів одержують в каналі різні часові затримки і поступають на вхід приймача в різні моменти часу.

Нелінійні спотворення виникають у каналі ТЧ і внаслідок залежності їх залишкового загасання від рівня вхідного сигналу і за рахунок нелінійності амплітудної характеристики групового тракту, до якого входить канал ТЧ. Для кожного каналу ТЧ такі продукти не є спотвореннями, а є завадами, тому що залежать від випадкової величини – загальної завантаженості групового тракту.

Під завадою будемо розуміти будь-яке випадкове діяння (вплив) на сигнал у каналі зв'язку, яке перешкоджає правильному прийманню сигналів. При цьому відокремимо випадковий характер діяння, оскільки боротьба з регулярними завадами не викликає труднощів.

У каналах зв'язку діють як адитивні завади, тобто випадкові процеси, які накладаються на передавальні сигнали, так і мультиплікативні завади, виражені у випадкових змінах характеристик каналу. Ці завади мають імпульсивний характер. Один з основних параметрів імпульсної завади – її амплітуда, тобто величина максимального викиду напруги. Ця величина випадкова, бо вона залежить від багатьох чинників: типу і потужності джерела завад, місця проникнення завади в канал зв'язку, частотних характеристик каналу тощо.

7.7.2. Спотворення дискретних сигналів у каналах зв'язку

Сигнали, що надходять у канал зв'язку (рис. 7.5, а), та сигнали в каналах зв'язку, спотворені неідеальними характеристиками каналу та наявністю різних завад, зазнають змін і в приймач надходять у спотвореному вигляді (рис. 7.5, б).

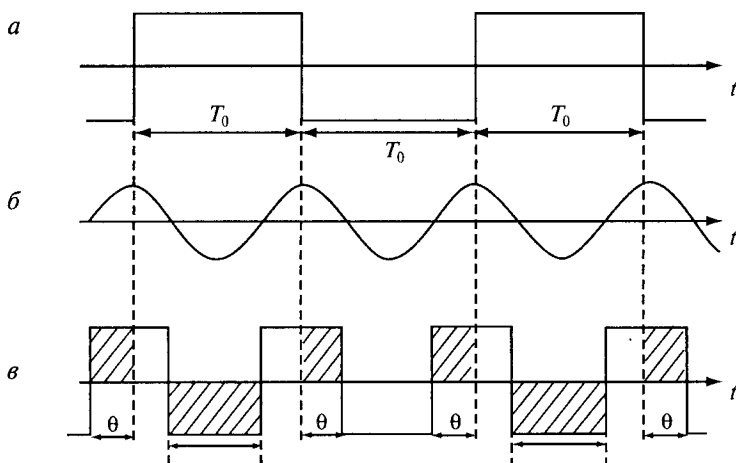


Рис. 7.5. Спотворення дискретних сигналів у каналах зв'язку

Після перетворення і демодуляції відновлений дискретний сигнал відрізняється від переданого (рис.7.5, в). Ці відмінності, які називають спотвореннями, полягають у тому, що значущі моменти прийнятого сигналу не завжди збігаються з переданими, а зміщуються на деяку величину θ . Окрім того, можливе появлення додаткових пар значущих моментів, які знаходяться на відстані τ один від одного. У зв'язку з цим домовились розрізняти два види спотворень дискретних сигналів:

- крайові спотворення (КС);
- спотворення дроблення (СД).

Крайові спотворення – це спотворення дискретних сигналів, що зумовлені зміщенням значущих моментів відносно ідеальних.

Крайові спотворення розділяють на регулярні та випадкові. До регулярних КС належать характеристичні та з переважанням.

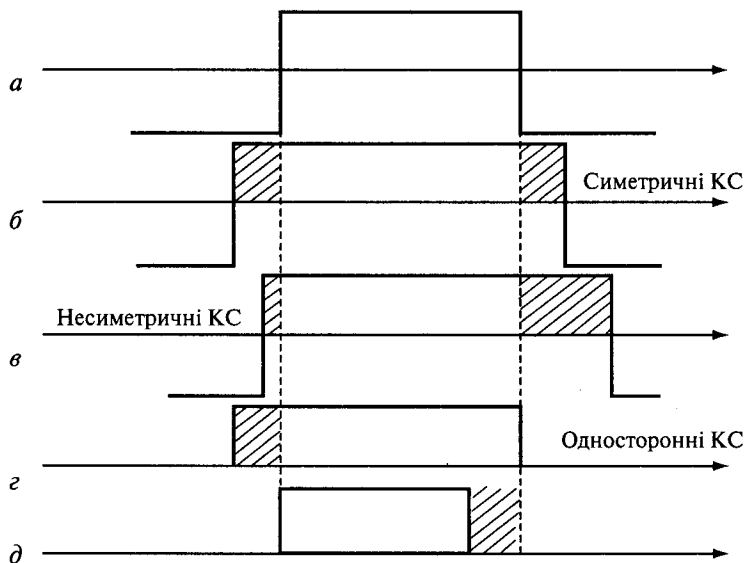


Рис. 7.6. Часові діаграми спотворення сигналів

7.7.3. Характеристичні крайові спотворення

Зумовлюються перехідними процесами в каналах зв'язку і виникають, коли перехідний процес не встигає закінчитися за час одиночного значущого інтервалу T_0 . Їх величина залежить від швидкості модуляції, ширини пропускання каналу, форми АЧХ і ФЧХ каналу.

Принцип виникнення характеристик спотворень ілюструє рис. 7.7.

Суцільними лініями показано передані (рис. 7.7, а) і прийняті (рис. 7.7, г) одиничні елементи під час передавання послідовності з переміжних одиниць і нулів («точок»), а пунктирними (рис. 7.7, б, д) – під час передавання трьох одиниць і одного нуля.

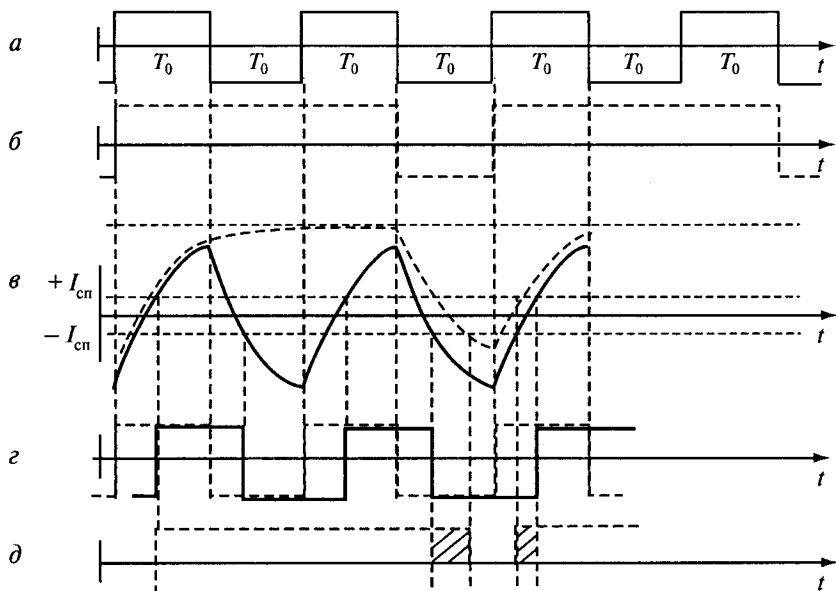


Рис. 7.7. Часова діаграма характеристикних спотворень

Як видно з рис. 7.7, в, під час передавання «точок» параметр сигналу не встигає досягнути сталого значення за час одиночного значущого інтервалу T_0 , але завдяки тому, що криві наростання і спадання параметра однакові, то КС немає. Водночас приймання послідовності з трьох одиниць і одного нуля виконується із спотворенням: четвертий одиничний елемент скоротився на початку і в кінці сигналу.

8. Методи перетворення дискретних сигналів для передавання каналами зв'язку – методи модуляції

8.1. Загальний підхід до опису сигналів

У системах зв'язку використовують різні перетворення сигналів. Одним із найважливіших перетворень є модуляція – змінювання параметрів деякого переносника («несучої») згідно із законом первинного сигналу $b(t)$. Так утворюється сигнал на виході модулятора $u(t)$, який здатний передаватись певною лінією зв'язку. Відомо, що канал телефонного зв'язку пропускає смугу частот 300...3400 Гц, тоді як раніше процес перетворення повідомлень у сигнали був зведений до формування інформаційних сигналів постійного електричного струму, які є носіями інформації виду «1» або «0». Отже, імпульсні одинокі струмові посилки є носіями інформації. Відомо також, що будь-який сигнал можна подати у вигляді множини гармонічних коливань різних частот.

Згадаймо, що інформація є джерелом, яке має свій власний кодер, що перетворює змістовий алфавіт у кодові посилки двійкових символів «0» і «1» і надає їм фізичну форму у вигляді електричних напруг на виходах відповідного регістра. Такі регістри утворюються тригерними комірками. Кожному розряду кодоутворення відповідає певний тригер, який має два виходи. Їх називають відповідно: лівий вихід – одиничним, а правий – нульовим. Якщо записується число, яке відображає відповідний алфавітний символ, то на лівих виходах тригерів виникає напруга високого рівня, що означає відповідність одиниці. Тепер, уявімо, що регістр складається з k -тригерів і одночасно на всіх лівих виходах тригерів з'являються сигнали високого рівня, що відповідає кодовій комбінації, яка відображає якийсь алфавітний символ.

Завдання дискретного каналу СПДІ – передавання цієї кодової комбінації по одному каналу зв'язку (наприклад, по двопровідній лінії зв'язку). Для цього потрібно багатолінійність виходів регістра звести до однолінійного передавання, що досягається застосуванням часового розподілу дії однієї лінії. Таку операцію виконують зсувні регістри, або окремі перетворювачі. Це обумовлює загальний підхід до опису сигналів. Найчастіше сигнали надходять у часовій або частотній формі.

Часове подання сигналу – це його опис за допомогою певної функції часу. Форма запису такої функції може бути різною. При цьому намагаються подати сигнал у вигляді ряду (полінома). Таке подання дозволяє

спростити математичну модель складного сигналу і виконати всі потрібні розрахунки.

Часове подання сигналу може бути замінено частотним. У цьому разі сигнал подається у вигляді спектра, який показує, як розподілена вся множина амплітуд і фаз сигналу частотами, що складають сигнал.

Систему функцій для подання сигналів вибирають, виходячи з практичної і математичної зручності, і передусім за типом використовуваного аналізатора сигналів. Під аналізатором сигналів розуміють технічний пристрій, призначений для визначення всіх потрібних характеристик сигналу.

У загальному вигляді будь-який сигнал $S(t)$ можна подати сукупністю елементарних сигналів $\eta_k(t)$, помножених на коефіцієнт C_k , і записати так:

$$S(t) = \sum_{k=0}^{\infty} C_k \eta_k(t).$$

Систему функцій $\{\eta(t)\}$, до складу якої включається набір функцій $\eta_0(t)$, $\eta_1(t)$, $\eta_2(t)$, ..., називають базисною системою, а подання сигналу у вигляді наведеної суми функцій – розкладом сигналу системою базисних функцій.

Найчастіше як функції $\eta_k(t)$ використовують гармонічні функції, функції Лежандра, Чебишова, Уолша, Ерміта, Мат'є та ін.

Для опису сигналів використовують такі найпростіші (типові) функції, як одиночний стрибок, функція знака й одиночний імпульс (дельта-функція).

Типові функції, використовувані для подання сигналів, мають задовольняти такі основні вимоги: бути досить простими в математичному описі; за їх допомогою мають подаватися будь-які сигнали (функції); бажано, щоб типова функція відображала ті дії (функції), які трапляються на практиці.

Функція знака. Функцію знака, позначену як $\text{sign}(t)$, називають ще сигнатурою, оскільки її назва походить від латинського слова *signum* (знак). Функція має постійну величину (амплітуду), що дорівнює одиниці, знак якої стрибком змінюється під час переходу змінної t через нуль (рис. 8.1.а).

Аналітичний вираз для функції знака має вигляд

$$\text{sign}(t) = \begin{cases} 1, & \text{якщо } t > 0, \\ 0, & \text{якщо } t = 0, \\ -1, & \text{якщо } t < 0. \end{cases}$$

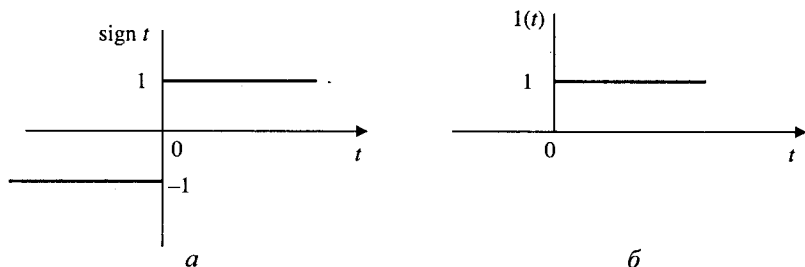


Рис. 8.1. Зображення функції стрибка (а) й одиничної функції (б)

Одинична функція (одиничний стрибок). Одиничну функцію позначають через $1(t)$; вона має постійне значення, що дорівнює одиниці, яке досягається миттєво (стрибком) в якийсь момент часу (в частковості, якщо $t = 0$). Аналітичний запис одиничної функції, яку зображено на рис. 8.1 б:

$$1(t) = \begin{cases} 0, & \text{якщо } t < 0, \\ 1/2, & \text{якщо } t = 0, \\ 1, & \text{якщо } t > 0, \end{cases}$$

Слід звернути увагу на те, що функція в момент часу $t = 0$ набуває значення $1/2$, що пояснюється взаємозв'язком одиничної функції із функцією знака, описаним виразом

$$1(t) = 1/2[1 + \text{sign}(t)].$$

Одиничну функцію називають ще функцією включення, оскільки множення $1(t)$ на який-небудь сигнал рівносильне його включенню в момент $t = 0$. Реально функція знака і одинична функція не існують, вони лише відображають дуже швидке зростання сигналу (струму, напруги). У загальному випадку функцію типу «стрибок» записують так:

$$f(t) = A \cdot 1(t - t_0).$$

Цей вираз описує сигнал типу «стрибок», який виникає в момент часу $t = t_0$ з амплітудою A .

За допомогою функцій «стрибок» можна будь-яку функцію подати як суму стрибків. Вивчаючи процеси формування сигналів, зосередимо увагу на прямокутні сигнали обмеженої тривалості τ . Тому розглянемо процес подання функції $f(t)$, що описує прямокутний імпульс, амплітудою A і тривалістю τ (рис. 8.2, а). Цю функцію можна подати у вигляді суми стрибків (рис. 8.2) у моменти часу $-t/2$ і $+t/2$. Складаючи обидва стрибки (рис. 8.2, б, в), отримаємо початковий прямокутний імпульс

$$f(t) = \begin{cases} A, & \text{якщо } -\frac{\tau}{2} < t < +\frac{\tau}{2}, \\ 0 & \text{для всіх інших } t, \end{cases}$$

який можна записати в такому вигляді:

$$f(t) = A \left\{ 1\left(t + \frac{\tau}{2}\right) - 1\left(t - \frac{\tau}{2}\right) \right\}.$$

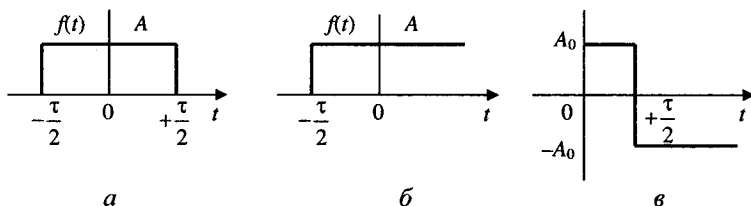


Рис. 8.2. Зображення $f(t)$ як прямокутний імпульс (а) і його утворення сукупністю стрибків (б і в)

Розглянуті типові функції зручні для опису сигналів під час дослідження реакції різних систем на вхідні діяння, до яких у цих випадках належать діяння, спрямовані на створення послідовності періодичних сигналів (імпульсів) з одиничних функцій, що утворюються на виходах регістрів джерела інформації. Однак вони не дають уявлення про спектр сигналу, знати який потрібно для оцінювання спотворення сигналів та вибору смуги пропускання технічного устаткування, до яких значною мірою належать канали зв'язку. Для того, щоб з'ясувати вимоги періодичних імпульсних (кодових) послідовностей до пропускнув здатності каналів (трактів) передавання дискретної інформації розглянемо подання періодичних сигналів рядом Фур'є.

Подання періодичних сигналів рядом Фур'є. Послідовність кодових комбінацій, що знімається з виходів власного кодера джерела інформації або з виходів кодера ПЗПер, формується в послідовність імпульсів однакової (тактової) тривалості у вигляді однополярних (рис. 8.3, а) або двополярних (рис. 8.3, б) сигналів.

Для подання таких сигналів рядом Фур'є нагадаємо, що функція $f(t)$ на інтервалі T_0 називається парною, якщо $f(t)_{T_0} = f(-t)_{T_0}$, і непарною, якщо $f(t)_{T_0} = -f(-t)_{T_0}$.

Очевидно, що прямокутні сигнали (струм, напруга) – парні функції.

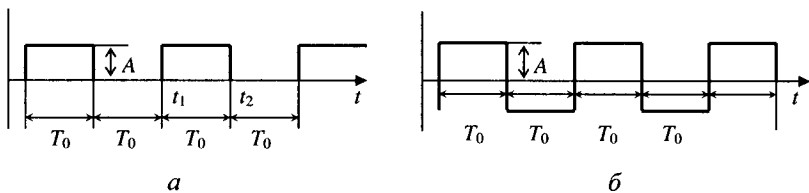


Рис. 8.3. Часові діаграми утворення послідовностей:
 а – однополярні; б – двополярні

Загальний вигляд тригонометричної форми ряду Фур'є для функції періодичного сигналу $f(t) = f(t + T_0)$ в інтервалі $T_0 = \frac{2\pi}{\omega_0} = t_2 - t_1$ (де $t_1 \leq t \leq t_2$ – інтервал задання функції $f(t)$) задається так:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\omega_0 t + b_n \sin n\omega_0 t).$$

Амплітуди косинусоїдних, синусоїдних членів і постійної складової в цьому розкладі визначають за виразами:

$$a_n = \frac{2}{T_0} \int_0^{T_0} f(t) \cos(n\omega_0 t) dt;$$

$$b_n = \frac{2}{T_0} \int_0^{T_0} f(t) \sin(n\omega_0 t) dt, \quad n = 1, 2, \dots$$

$$\frac{a_0}{2} = \frac{A}{2} = \frac{1}{T_0} \int_0^{T_0} f(t) dt.$$

Амплітуда A_n і фаза φ_n n -ї гармоніки зв'язані з величинами a_n і b_n співвідношеннями $A_n = \sqrt{a_n^2 + b_n^2}$, $a_n = A_n \cos \varphi_n$, $b_n = A_n \sin \varphi_n$, $\varphi_n = \arctg \frac{b_n}{a_n}$.

Для першої функції $f(t)$ відомо, що коефіцієнти b_n у загальній формі розкладу виявляються нулями. Тому парна функція має ряд Фур'є, складений з одних косинусів:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos n\omega_0 t.$$

Отже, сигнал прямокутної форми можна розкласти на деяку множину гармонічних коливань. Частоти будь-якої гармоніки кратні частоті першої (основної) гармоніки і знаходяться в простих кратних співвідношеннях. Тому періодичний сигнал довільної форми можна розглядати як ре-

зультат накладання одну на одну нескінченної кількості гармонік, а також постійної складової. Кількість гармонік, на яку розкладається сигнал, називають спектром частот цього сигналу. Для реальних сигналів, які передаються мережею електрозв'язку, спектр частот для їх неспотвореного передавання має бути нескінченно широким. Однак на практиці використовують обмежений спектр частот, за якого існує нехай не ідеальна, але допустима якість передавання, яка не спричиняє суттєвого спотворення сигналу, а отже, і повідомлення. Більше за все цей висновок придатний для широкосмугових каналів.

Якщо для передавання даних використовують канали тональної частоти телефонного зв'язку, які мають обмежений пропускний частотний спектр 300...3400 Гц, то звідси випливає, що дискретні сигнали постійного струму не можна передавати цими каналами. Це означає, що треба враховувати середовище передавання – канал тональної частоти (його пропускну частотну смугу).

Зазначені обставини враховують під час уведення процедури перетворення сигналів, за якої відшукується такий сигнал, який пристосовується до вимог середовища передавання і зберігає основні властивості первинного сигналу. Таке перетворення сигналу називають модуляцією. У методах модуляції використовують деякий переносник сигналу у вигляді гармонічного коливання високої частоти – несучого коливання (часто кажуть «несуча частота»). У цьому разі модуляція являє собою процес змінування одного або декількох параметрів (частоти, амплітуди, фази) згідно із законом первинного сигналу, тобто несуче коливання набуває ознаки первинного сигналу. На приймальному кінці системи здійснюється обернене перетворення (демодуляція), за якого з модульованого сигналу виділяється первинний сигнал.

Отже, у разі амплітудної модуляції (АМ) у несучому коливанні змінюється амплітуда. Якщо змінюється частота несучого коливання, маємо частотну модуляцію (ЧМ), а якщо змінюється фаза несучого коливання, – фазову модуляцію.

Для передавання дискретного сигналу через аналогове середовище (канали тональної частоти телефонного зв'язку) також використовують модуляцію первинним дискретним сигналом гармонічного несучого коливання. При цьому, як і раніше, можливі три види модуляції: амплітудна, частотна і фазова. Перейдемо до розгляду цих видів модуляції.

8.2. Ампліудна модуляція. Спектр модульованого сигналу за ампліудної модуляції

За ампліудної модуляції переносником інформації є амплітуда несучого коливання.

Нехай рівняння переносника має вигляд

$$U = U_m \cos(\omega_0 t + \varphi_0),$$

де U_m – амплітуда напруги переносника; ω_0 – його кутова частота; φ_0 – його початкова фаза.

Для загального пояснення процесів перетворення сигналів у СПДІ, зважаючи на особливості, які доводиться враховувати під час передавання реальними каналами зв'язку (фазочастотні спотворення, мале співвідношення несучої і модульованої частот, частотна і фазова модуляції, багатократна модуляція, передавання повідомлень періодичною послідовністю прямокутних імпульсів), потрібно розглянути передусім сам принцип ампліудної модуляції і отримані спектральні характеристики заново утвореного сигналу.

Найкраще це буде зробити, якщо передати сигнал, який має одну частоту, у вигляді простого гармонічного коливання, що змінюється за законом косинуса,

$$u_k = U_\Omega \cos \Omega t,$$

де u_k – миттєве значення напруги сигналу в момент часу t ; Ω – колова частота коливання.

Це гармонічне коливання має бути перетворено в гармонічне коливання частотою, набагато більшою ніж Ω , зі збереженням інформації, яку несе це коливання. Для цього потрібно мати переносник цієї інформації. Нехай функції переносника виконує також гармонічне коливання частотою $\omega \gg \Omega$ вигляду $u_m = U_m \cos \omega t$ (для кращого розуміння вважаймо, що початкові фази дорівнюють нулю).

Припустімо, що амплітуда модульованого сигналу відповідно до керувального сигналу u_m набуде вигляду

$$U_m + U_\Omega \cos \Omega t = U_m \left(1 + \frac{U_\Omega}{U_m} \cos \Omega t \right).$$

Починаючи з моменту здійснення модуляції, амплітуда високочастотного коливання, тобто коливання, буде мати таку форму запису:

$$u = U_m \left(1 + \frac{U_\Omega}{U_m} \cos \Omega t \right) \cos \omega t.$$

Відношення $m = \frac{U_\Omega}{U_m}$ називають *коефіцієнтом модуляції* за амплітудної модуляції. Процес модуляції зображено на рис. 8.4.

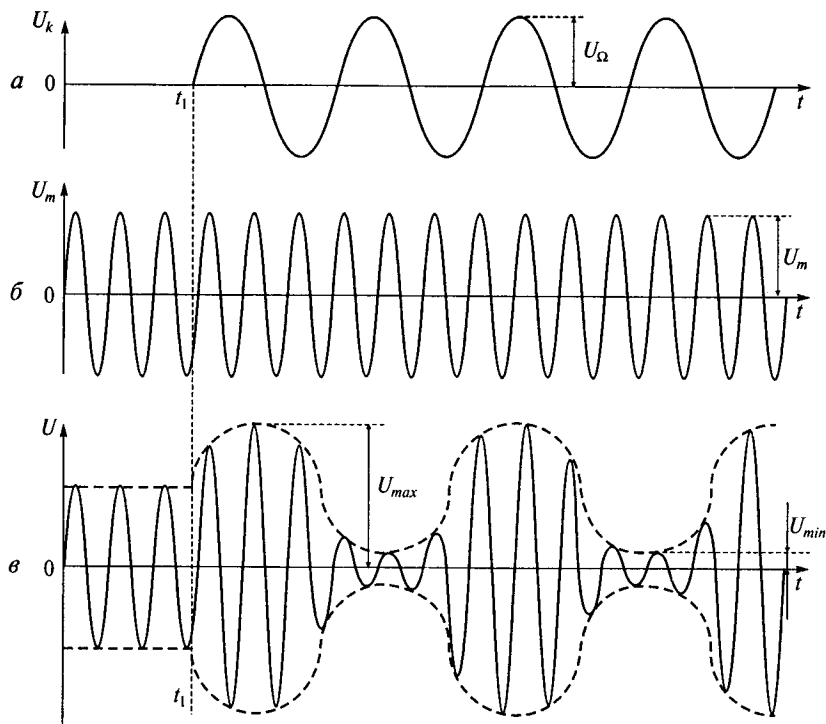


Рис. 8.4. Амплітудна модуляція: a – модульований первинний сигнал $u_k = U_\Omega \cos \Omega t$;

b – модулювальний сигнал $u_m = U_m \cos \omega t$;

c – модульований сигнал $u = U_m \left(1 + \frac{U_\Omega}{U_m} \cos \Omega t \right) \cos \omega t$

Визначимо амплітуду модульованого сигналу:

$$U_{\max} = U_m + U_\Omega;$$

$$U_{\min} = U_m - U_\Omega.$$

Звідси

$$U_m = \frac{1}{2}(U_{\max} + U_{\min}), U_{\Omega} = \frac{1}{2}(U_{\max} - U_{\min}),$$

тоді

$$m = \frac{U_{\Omega}}{U_m} = \frac{U_{\max} - U_{\min}}{U_{\max} + U_{\min}}.$$

Очевидно, що m змінюється від 0 до 1, або $0 \leq m \leq 100\%$.

Модульований сигнал – складний сигнал (див. рис. 8.4, в). Визначимо спектр частот отриманого модульованого сигналу, модуляція якого здійснювалася однією частотою $\Omega \leq \omega$ (Ω – низька частота, ω – висока частота):

$$\begin{aligned} u &= U_m \left(1 + \frac{U_{\Omega}}{U_m} \cos \Omega t\right) \cos \omega t = U_m (1 + m \cos \Omega t) \cos \omega t; \\ u &= U_m \cos \omega t + U_m m \cos \Omega t \cos \omega t; \end{aligned} \quad (8.1)$$

$$u = U_m \cos \omega t + \frac{1}{2} U_m m \cos(\Omega + \omega)t + \frac{1}{2} U_m m \cos(\omega - \Omega)t.$$

З рівняння (8.1) видно, що модульоване коливання складне, тобто складається з коливання несучої частоти ω та двох коливань з боковими частотами – $(\omega + \Omega)$ і $(\omega - \Omega)$. Спектр такого модульованого коливання показано на рис. 8.5, а.

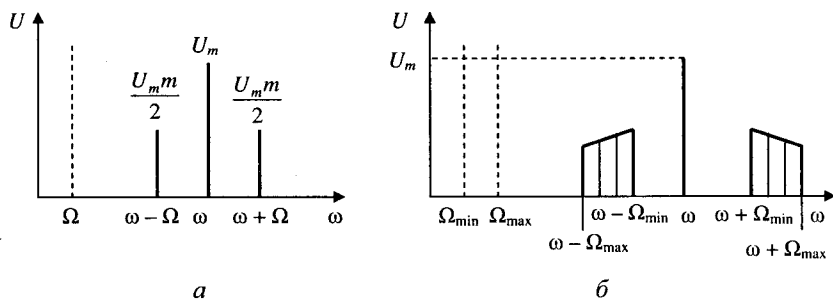


Рис. 8.5. Спектри сигналів за амплітудної модуляції:

- а – первинний сигнал (одногоармонічний);
- б – первинний сигнал (багатоармонічний)

Якщо первинний сигнал (модульований) складний і має не одну частоту, а спектр частот, розміщений на інтервалі $\Omega_{\min} \dots \Omega_{\max}$, то спектр модульованого коливання буде мати дві бокові смуги, як показано на рис. 8.5, б.

У системах передавання дискретної інформації форми послідовності сигналів (інформаційних) мають вигляд відеоімпульсів (див. підрозд. 7.2) – імпульсів постійного струму прямокутної форми. Тут і надалі будемо розглядати спектри модульованих коливань у випадках, коли $u_m(t)$ – періодична послідовність прямокутних імпульсів, а переносник – гармонічне коливання.

Графічне зображення амплітудно-модульованих сигналів за дискретної модуляції для випадку, коли амплітуда первинного сигналу дорівнює амплітуді моделювального сигналу $U_k = U_m$ (за цієї умови $m = \frac{U_k}{U_m} = 1$), показано на рис. 8.6.

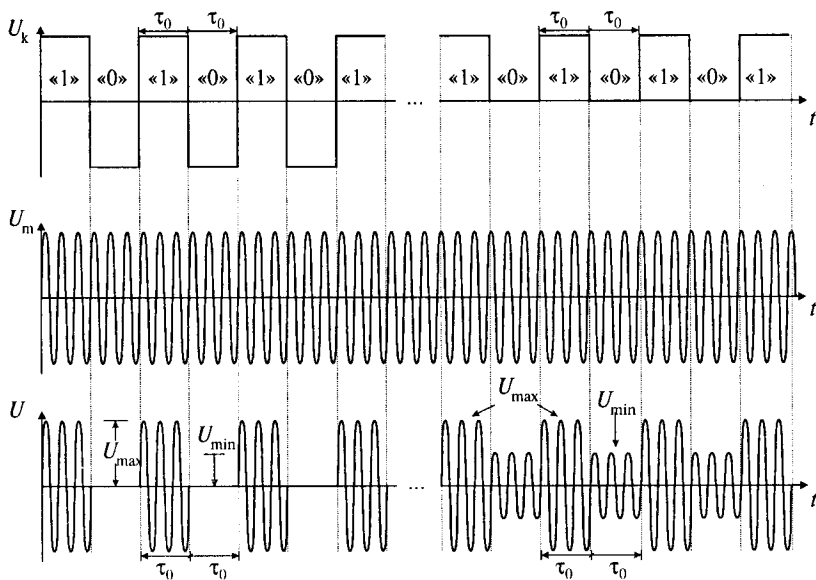


Рис. 8.6. Амплітудно-модульовані сигнали за дискретної модуляції точок ($m = 1$), якщо $U_k = U_m$: а – первинні сигнали (різнополярні); б – первинні сигнали (однополярні)

Для передавання дискретної інформації часто доводиться визначати ширину смуги частот, яку має канал зв'язку для того, щоб пропустити імпульсні сигнали із заданою швидкістю, або швидкістю передавання, яка може бути отримана у каналі, якщо задано ширину смуги пропускання. Для цього потрібно порівняти ширину каналу і ширину частотного діапазону, який займають передавальні сигнали, тобто розрахувати спектри

за різними методами модуляції з різним характером передавальних імпульсних послідовностей. Отже, вивчення імпульсних сигналів дозволяє правильно вибирати швидкості передавання і ширину каналу для розроблення та експлуатації СПДІ.

Для приблизного вирішення поставлених завдань достатньо знати тільки спектри, а для більш точного – визначити значення тих спотворень, які отримують імпульси в процесі проходження каналом зв'язку. Прямокутні імпульси, які подаються на вхід каналу, перетворюються на його виході в «згладжені», скруглені сигнали. Зменшується або збільшується тривалість імпульсу. З'являються «провісники» і «хвости» імпульсу, які заважають передаванню сусідніх імпульсів. Ці спотворення імпульсів треба уміти розраховувати і для того, щоб на практиці вибирати найбільш раціональні способи їх зменшення.

Знання перехідних процесів, які виникають під час передавання імпульсних сигналів каналами зв'язку, допоможуть оцінити якість передавання дискретної інформації.

Загальну теорію спектрів застосовують для пояснення процесів у СПДІ. Вивчивши спектри і перехідні процеси, можна визначити ширину фільтрів і допустиму швидкість у системах передавання даних.

У СПДІ широко використовують імпульси постійного струму різної форми. Найчастіше застосовують імпульси прямокутної форми. Таку форму зазвичай мають інформаційні та синхронізувальні імпульси, якщо немає помітних спотворень.

Як відомо із теоретичних курсів, періодичний сигнал будь-якої складної форми можна подати у вигляді суми гармонічних коливань з частотами, кратними частоті повторення цього сигналу. Ці коливання називають гармоніками спектра сигналу. Як відомо, таке розкладання рядом Фур'є має вигляд

$$u(t) = C_0 + \sum_{k=1}^{\infty} C_k \cos(k\Omega t - \varphi_k),$$

де $u(t)$ – розкладений періодичний сигнал; k – номер гармонік ($k = 1, 2, \dots$); C_0 – постійна складова; C_k – амплітуда гармонік; φ_k – початкові фази гармонік; Ω – колова частота повторення:

$$\Omega = \frac{2\pi}{T},$$

де T – період повторення функції $u(t)$.

Сукупність амплітуд C_k і C_0 називають спектром амплітуд (або просто спектром), а сукупність початкових фаз φ_k – спектром фаз.

Амплітуди гармонік можна розрахувати за формулою

$$C_k = \sqrt{a_k^2 + b_k^2}, \quad (8.2)$$

де

$$a_k = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t) \cos k\Omega t dt; \quad (8.3)$$

$$b_k = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t) \sin k\Omega t dt. \quad (8.4)$$

Постійну складову обчислюють за формулою

$$C_0 = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t) dt. \quad (8.5)$$

Інтеграли у формулах (8.2)–(8.4) легко розв'язуються. Обчислимо як приклад спектр сигналу під час передавання прямокутних імпульсів. Для цього спочатку розглянемо передавання «точок», що відповідає сигналам виду 1 : 1 (рис. 8.7, а). Спектр цього сигналу можна записати за формулами (8.2) і (8.3) так:

$$C_k = |a_k| = \left| \frac{2}{T} \int_{-\frac{\tau_0}{2}}^{\frac{\tau_0}{2}} U_0 \cos k\Omega t dt \right| = U_0 \left| \frac{\sin \frac{\pi k}{2}}{\frac{\pi k}{2}} \right|; \quad (8.6)$$

$$C_0 = \frac{1}{T} \int_{-\frac{\tau_0}{2}}^{\frac{\tau_0}{2}} U_0 dt = \frac{U_0}{2}. \quad (8.7)$$

Спектр «точок» показано на рис. 8.7, б. Стала складова дорівнює $\frac{U_0}{2}$. Основну енергію вміщує перша гармоніка, усі парні гармоніки пропадають.

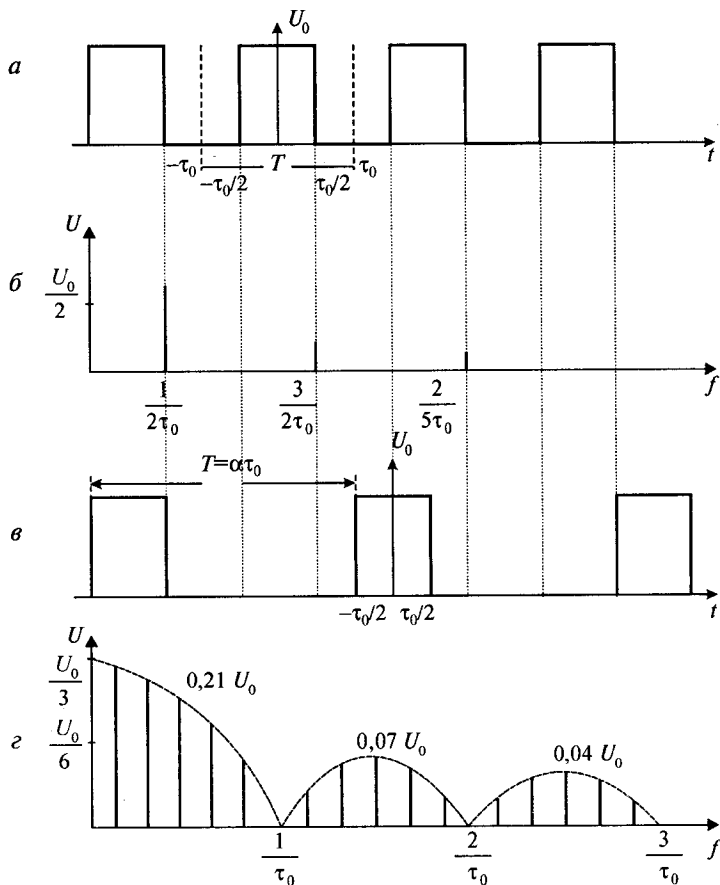


Рис. 8.7. Сигнали виду 1:1 («точки») і виду 1: ($\alpha - 1$):
 а – форма імпульсів; б – спектр; в – форма імпульсів; г – спектр

Розглянемо тепер спектр імпульсної послідовності виду 1 : 5. Нехай в загальному випадку

$$T = \alpha\tau_0,$$

тобто період повторення довший від елементарного імпульсу в α разів.

Це буде сигнал виду 1: ($\alpha - 1$), у якому є імпульси струму тривалістю τ_0 і паузи тривалістю $(\alpha - 1)\tau_0$. Такий сигнал показано на рис. 8.7, в. Коефіцієнт α часто називають відстанню між імпульсами або щільною імпульсної послідовності.

Оскільки сигнал є парною функцією у часі, то всі b_k обертаються в нуль. Амплітуди гармонік будуть дорівнювати:

$$C_k = |a_k| = \left| \frac{2}{T} \int_{-\frac{\tau_0}{2}}^{\frac{\tau_0}{2}} U_0 \cos k\Omega t dt \right| = \frac{2U_0}{\alpha} \left| \frac{\sin \frac{\pi k}{\alpha}}{\frac{\pi k}{\alpha}} \right|. \quad (8.8)$$

Постійна складова згідно з виразом (8.3)

$$C_0 = \frac{U_0}{\alpha}. \quad (8.9)$$

Наприклад, для сигналу виду 1 : 5 (якщо $\alpha = 6$) отримуємо

$$C_k = \frac{U_0}{3} \left| \frac{\sin \frac{\pi k}{6}}{\frac{\pi k}{6}} \right|; \quad C_0 = \frac{U_0}{6}. \quad (8.10)$$

Спектр цього сигналу показано на рис. 8.7, з. Порівняно зі спектром сигналу 1 : 1 спектр сигналу 1 : 5 набагато складніший. Від нуля до частоти $1/\tau_0$ тепер розміщується 5 гармонік (а з «точками» розміщувалась лише одна). Унаслідок збільшення кількості гармонік амплітуда кожної з них зменшилась.

Розгляд цього окремого випадку і загальної формули (8.8) дозволяє замінити такі закономірності спектра послідовності прямокутних імпульсів виду 1 : (α):

1. Амплітуда спектральних складових спадає за законом функції

$$\left| \frac{\sin x}{x} \right|.$$

2. Амплітуда спектральних складових обертається в нуль на частотах $1/\tau_0, 2/\tau_0, 3/\tau_0$ і т. д.

3. В основному діапазоні частот від нуля до $1/\tau_0$ розміщується $\alpha - 1$ гармонік через однакові проміжки.

4. Стала складова (компонента з нульовою частотою)

$$C_0 = \frac{U_0}{\alpha}.$$

5. Амплітуда гармонік зі зменшенням частоти прямує до величини $2C_0$.

Якщо припустити, що період повторення сигналу стає дедалі більшим і більшим, то, розглядаючи рис. 8.7, з, можна зрозуміти, яким буде спектр одиночного прямокутного імпульсу. Для обчислення цього спектра у наведених формулах (8.8)–(8.10) потрібно покласти $\alpha \rightarrow \infty$. Спектр дедалі

матиме таку форму, яку показано на рис. 8.7, *з*, але в діапазоні від нуля до частоти $1/\tau_0$ буде розміщуватись нескінченно велика кількість гармонік. Спектр стане суцільним. На відміну від нього спектр періодичної функції називають лінійчастим (або дискретним).

Математично спектр одиночного імпульсу, як будь-якої неперіодичної функції, потрібно визначати уже не рядом Фур'є, а інтегралом Фур'є:

$$u(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega t} d\omega,$$

де $u(t)$ – розглядуваний неперіодичний сигнал; $\Omega = 2\pi f$ – поточна колова частота; $S(\Omega)$ – комплексний спектр, який можна обчислити з виразу

$$S(\Omega) = \int_{-\infty}^{\infty} u(t) e^{-j\Omega t} dt. \quad (8.11)$$

Зазвичай цікавляться лише спектром амплітуд $|S(\Omega)|$ (точніше спектральною щільністю амплітуд).

Спектр одиночного прямокутного імпульсу (рис. 8.8, *а*) за формулою (8.10) буде дорівнювати:

$$S(\Omega) = u_0 \int_{-\frac{\tau_0}{2}}^{\frac{\tau_0}{2}} e^{-j\Omega t} dt = u_0 \tau_0 \frac{\sin \pi \tau_0 f}{\pi \tau_0 f}.$$

Форму спектра амплітуд $|S(\Omega)|$ показано на рис. 8.8.

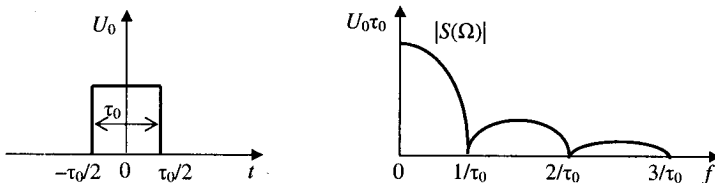


Рис. 8.8. Прямокутний одиночний імпульс та його спектр за аналізом рис. 8.7

Тут і надалі будемо розглядати спектри модульованих коливань у випадках, коли $f(t)$ – періодична послідовність прямокутних імпульсів і переносник – гармонічне коливання.

До цього ми розглядали побудову спектрів модульованих коливань, які задаються послідовністю прямокутних імпульсів з різною відстанню (щільністю) між імпульсами, і встановили: якщо відстань α буде весь час збільшуватись і прямувати $\alpha \rightarrow \infty$, тоді будемо мати одиночний імпульс. Порівнюючи спектри, якщо $\alpha = 2$, $\alpha = 6$ та $\alpha \rightarrow \infty$, бачимо, що форма

спектрів залишається однаковою і змінюється тільки кількість амплітуд (гармонік) в інтервалі частот від 0 до $1/\tau_0$ та інтервалах, кратних йому, а якщо $\alpha \rightarrow \infty$, дискретний спектр перетворюється в суцільний, як це показано на рис. 8.7.

Якщо відомо спектр модульованого первинного сигналу, тоді немає потреби розкладати модульований імпульс рядом Фур'є. Достатньо знати спектральний розклад модульованих імпульсів і підставити його у вираз (8.11).

Розглянемо, наприклад, періодичні прямокутні імпульси (див. рис. 8.6, б), тобто амплітудну модуляцію. Тривалість елементарної послідовності позначимо через τ_0 , а період проходження імпульсів цієї послідовності – T . Модульовані імпульси згідно з формулами (8.6) і (8.7) можна записати так:

$$f(t) = \frac{u_0}{\alpha} + \frac{2u_0}{\alpha} \sum_{k=1}^{\infty} \frac{\sin \pi \frac{k}{\alpha}}{\pi \frac{k}{\alpha}} \cos k\Omega t, \quad (8.12)$$

де Ω – колова частота повторення; $\Omega = \frac{2\pi}{T}$; α – щільність імпульсної по-

слідовності; $\alpha = \frac{T}{\tau_0}$.

Для послідовності 1 : 1 (див. рис. 8.7, а) $\alpha = 2$, для послідовності 1 : ($\alpha - 1$) (рис. 8.7, в) – α . Підставимо формулу (8.12) у вираз для U , отримаємо

$$U_{AM} = \frac{U_M}{\alpha} \cos(\omega_0 t + \varphi_0) + \frac{U_M}{\alpha} \sum_{k=1}^{\infty} \frac{\sin \pi \frac{k}{\alpha}}{\pi \frac{k}{\alpha}} \{ \cos[(\omega_0 + k\Omega)t] + \cos[(\omega_0 - k\Omega)t] \}. \quad (8.13)$$

Формула (8.12) показує, що які всі модульовані коливання імпульсної сигнали з АМ мають, окрім несучої частоти, ще верхню і нижню бокові смуги. Складові верхньої бокової мають частоти $\omega_0 + k\Omega$, а складові нижньої бокової: $\omega_0 - k\Omega$. Амплітуди складових залежать від коефіцієнта модуляції m і щільності імпульсної послідовності α . У формулах (8.12) і (8.13) k – номер бокових частот ($k = 1, 2, \dots$).

Приклади спектрів АМ сигналів видів 1 : 1 ($\alpha = 2$) і 1 : 5 ($\alpha = 6$), обчислені за цими двома формулами, показано на рис. 8.9, в, г. Для розрахунку було взято два коефіцієнти: $m = 0,5$ і $m = 1$.

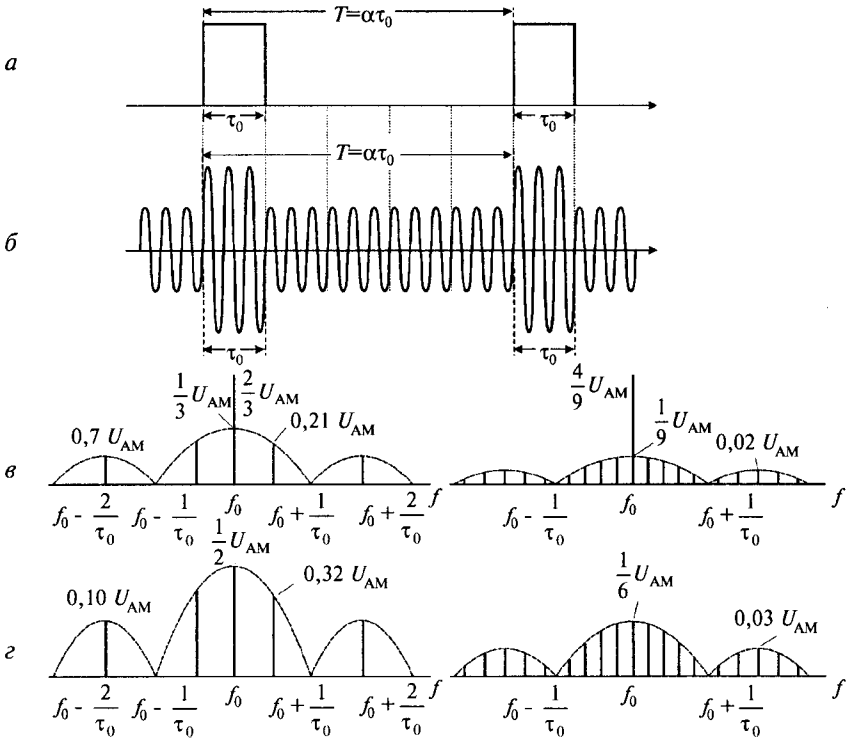


Рис. 8.9. Спектри АМ сигналів під час передавання періодично повторюваних прямокутних імпульсів:

a – форма модульованих імпульсів; *б* – форма змодульованих імпульсів; *в* – спектри модульованих імпульсів, якщо $m = 0,5$; *г* – спектри, якщо $m = 1$

Розглядаючи рис. 8.9, можна помітити ряд закономірностей у спектрах імпульсних сигналів з АМ:

- 1) форма бокових смуг спектра модульованого сигналу аналогічна формі спектра модульованих імпульсів;
- 2) спектр модульованого сигналу в два рази ширший від спектра модульованих імпульсів;
- 3) форма спектра завжди симетрична відносно несучої частоти, тобто нижня і верхня бокові смуги мають однакову форму;
- 4) амплітуда несучої частоти вписується в обвідну спектра за стовідсоткової модуляції. Зі зменшенням коефіцієнта модуляції енергія несучої зростає, а енергія бокових смуг спадає.

8.3. Частотна модуляція. Спектр модульованого сигналу за частотної модуляції

У разі частотної модуляції передавана (первинна) інформація відображається на коливаннях несучого сигналу – змінюється його частота відповідно до закону зміни первинного сигналу, загальний вираз якої має вигляд

$$\omega_0(t) = \omega_0[1 + m'f(t)],$$

де $f(t)$ – первинний (модульований) сигнал, який подає інформацію; ω_0 – несуча частота; m' – коефіцієнт пропорційності, який встановлює зв'язок між модульовальним сигналом і зміною частоти високочастотного (модульовального) сигналу, використовуваного для передавання інформації.

Частотно-модульований сигнал у загальному вигляді описується виразом

$$u(t) = U_m \cos[\varphi(t)],$$

де $\varphi(t)$ – миттєва фаза сигналу, оскільки зміна частоти супроводжується зміною фази сигналу;

$$\varphi(t) = \int_0^t \omega(t) dt + \varphi_0.$$

Тоді

$$\varphi(t) = \omega_0 t + \Delta\omega \int_0^t f(t) dt + \varphi_0. \quad (8.14)$$

Отже, напругу, модульовану за частотою, можна записати у вигляді

$$u(t) = U_m \cos(\omega_0 t + \Delta\omega \int_0^t f(t) dt + \varphi_0).$$

Позначимо у формулі (8.14) величину $\Delta\omega \int_0^t f(t) dt$ через $\theta(t)$; тоді її можна переписати так:

$$u(t) = U_m \{ \cos[\omega_0 t + \varphi_0] \cos \theta(t) - \sin[\omega_0 t + \varphi_0] \sin \theta(t) \}. \quad (8.15)$$

У разі дискретної частотної модуляції модульованим сигналом є періодична послідовність прямокутних імпульсів (рис. 8.10, а), причому позитивному модульованому сигналу відповідає частота «натиснення» ω_n , а негативному модульованому сигналу – частота «відтиснення» ω_b .

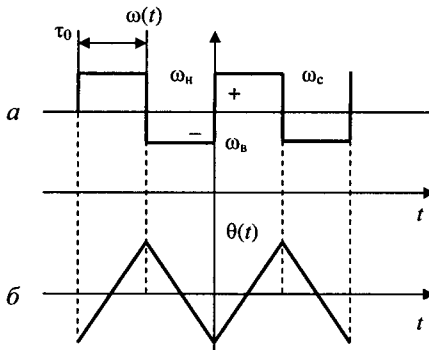


Рис. 8.10. Модульований сигнал (а) і перехідна фаза ЧМ (б)

У виразі (8.15) ω_0 відповідає середній частоті $\omega_c = (\omega_n + \omega_b)/2$, а $\Delta\omega = (\omega_n - \omega_b)/2$ – девіація частоти. Перехідна фаза сигналу при цьому змінюється за таким законом (рис. 8.10, б):

$$\theta(t) = \begin{cases} -\Delta\omega(t + \frac{\tau_0}{2}), & \text{якщо } -\tau_0 < t < 0; \\ \Delta\omega(t - \frac{\tau_0}{2}), & \text{якщо } 0 < t < \tau_0. \end{cases}$$

Подамо $\cos \theta(t)$ і $\sin \theta(t)$ у вигляді їх розвинення у ряди Фур'є:

$$\left. \begin{aligned} \cos \theta(t) &= A_0 + \sum_{k=2,4,6\dots}^{\infty} A_k \cos k\Omega_1 t; \\ \sin \theta(t) &= \sum_{k=1,3,5\dots} B_k \sin k\Omega_1 t, \end{aligned} \right\} \quad (8.16)$$

де Ω_1 – колова частота повторення імпульсів;

$$\begin{aligned} A_0 &= \frac{1}{2\tau_0} \int_{-\tau_0}^{\tau_0} \cos \theta(t) dt = \\ &= \frac{1}{2\tau_0} \left\{ \int_{-\tau_0}^0 \cos \left[-\Delta\omega(t + \frac{\tau_0}{2}) \right] dt + \int_0^{\tau_0} \cos \left[\Delta\omega(t - \frac{\tau_0}{2}) \right] dt \right\} = \frac{\sin \frac{\Delta\omega\pi}{\Omega_1}}{2\Omega_1}; \end{aligned}$$

$$A_k = \frac{2}{\tau_0} \int_{\frac{-\tau_0}{2}}^{\frac{\tau_0}{2}} \cos \theta(t) \cos k\Omega_1 t dt = \frac{4}{\pi} \frac{m \sin(\pi m / 2)}{m^2 - k^2};$$

$$B_k = \frac{4}{\pi} \frac{m \cos(\pi m / 2)}{m^2 - k^2}; \quad m = \frac{\Delta \omega}{\Omega_1} - \text{індекс модуляції.}$$

Підставивши вираз (8.16) у формулу (8.15), отримаємо

$$u(t) = U_m \frac{\sin(\pi m / 2)}{\pi m / 2} \cos(\omega_c t + \varphi_0) + U_m \frac{2}{\pi} \sum_{k=2,4,6\dots}^{\infty} \frac{m \sin(\pi m / 2)}{m^2 - k^2} \times$$

$$\times \left\{ \cos \left[(\omega_c - k\Omega_1)t + \varphi_0 \right] + \cos \left[(\omega_c + k\Omega_1)t + \varphi_0 \right] \right\} +$$

$$+ U_m \frac{2}{\pi} \sum_{k=1,3,5\dots}^{\infty} \frac{m \cos(\pi m / 2)}{m^2 - k^2} \times$$

$$\times \left\{ \cos \left[(\omega_c - k\Omega_1)t + \varphi_0 \right] - \cos \left[(\omega_c + k\Omega_1)t + \varphi_0 \right] \right\}. \quad (8.17)$$

Із виразу (8.17) видно:

- спектр частотно-модульованого коливання складатиметься із несучої частоти, верхньої та нижньої бокових смуг;
- парні і непарні бокові частоти підпорядковуються різним законам і відрізняються за фазою на 90° ;
- спектри бокових частот відрізняються від спектра модульованого сигналу; форма спектра залежить від індексу модуляції.

Частотну модуляцію можна реалізувати двома способами. Перший спосіб модуляції виконується перемиканням незалежних генераторів, один з яких генерує частоту $f_1 = f_n$, що відповідає сигналу $+h$, а другий – частоту $f_2 = f_n$, що відповідає сигналу $-h$. При цьому перехід від сигналу $+h$ до сигналу $-h$ обернено супроводжується стрибком фази, бо фази частот f_1 і f_2 змінюються незалежно. Перший спосіб частотної модуляції називають «зі стрибком фази». Для модуляції другим способом використовують один генератор, частота якого у значущі моменти змінюється від f_1 до f_2 або обернено без стрибка фази (модуляція «з неперервною фазою»).

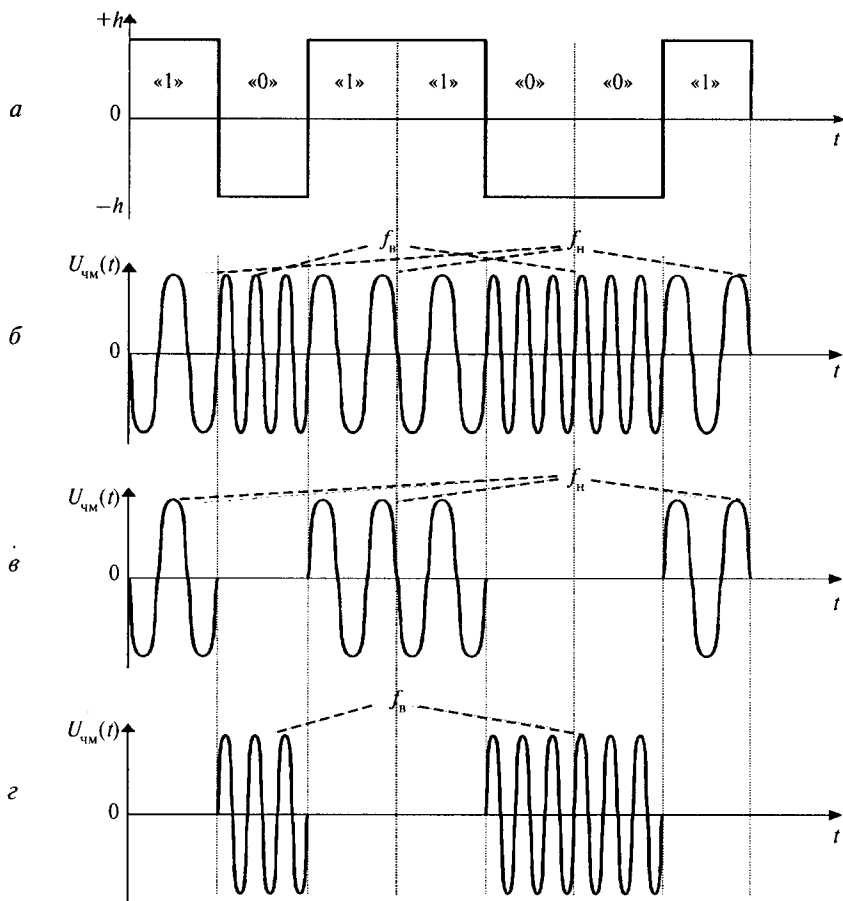


Рис. 8.11. Дискретна частотна модуляція: a – модульований (первинний) сигнал; b – змодульований сигнал ЧМ; c, d – складові сигналу ЧМ

Частотну модуляцію «зі стрибком фази» можна розглядати як суму двох АМ сигналів з різними частотами несучих коливань f_1 і f_2 , модульованих у різні проміжки часу, як показано на рис. 8.11. На цьому рисунку зображено форму сигналів за частотної модуляції «з неперервною фазою».

8.4. Фазова модуляція.

Спектр модульованого сигналу за фазової модуляції

У разі фазової модуляції передавана (первинна) інформація відображається на зміні коливання несучого сигналу, а саме: змінює його фазу відповідно до закону первинного сигналу $f(t)$, загальний вираз якої можна записати так:

$$\theta = \omega_0 t + \varphi_0 + \Delta\varphi f(t),$$

де $f(t)$ – первинний (модульований) сигнал, який передає інформацію; ω_0 – несуча частота; φ_0 – початкова фаза несучого коливання; $\Delta\varphi$ – максимальна зміна початкової фази.

Миттєве значення фазомодульованої напруги має такий вигляд:

$$u(t) = U_m \cos \theta = U_m \cos[\omega_0 t + \varphi_0 + \Delta\varphi f(t)], \quad (8.18)$$

де $\Delta\varphi$ – індекс ФМ (девіація фази). Чим більше змінюється модульований параметр, тим, очевидно, легше відрізнити одне від одного значення передаваних сигналів під час їх приймання. Тому значення девіації фази треба вибирати якомога більшим, тобто $\Delta\varphi = 90^\circ$.

У разі модуляції серії прямокутних імпульсів (рис. 8.12, а) фазомодульований сигнал при $\Delta\varphi = 90^\circ$ має вигляд, як показано на рис. 8.12, б. Модульований фазою сигнал можна подати як суму двох модульованих за амплітудою сигналів, які мають однакову частоту ω_0 , але відрізняються значенням початкової фази. Зокрема, для випадку $\Delta\varphi = 90^\circ$ (рис. 8.12, б) ці сигнали (U_{Am1} і U_{Am2}) зображено на рис. 8.12, в і г. Їх несучі частоти відрізняються за фазою на $2\Delta\varphi = 180^\circ$.

Перепишемо вираз (8.18) у такому вигляді:

$$u(t) = U_m \{ \cos(\omega_0 t + \varphi_0) \cos[\Delta\varphi f(t)] - \sin(\omega_0 t + \varphi_0) \sin[\Delta\varphi f(t)] \}.$$

Побудуємо спектр для випадку дискретної ФМ:

$$f(t) = \begin{cases} +1, & \text{якщо } (n-1)\tau_0 \leq t \leq n\tau_0; \\ -1, & \text{якщо } n\tau_0 \leq t \leq (n+1)\tau_0. \end{cases}$$

При цьому

$$u(t) = U_m [\cos(\omega_0 t + \varphi_0) \cos \Delta\varphi - f(t) \sin(\omega_0 t + \varphi_0) \sin \Delta\varphi]. \quad (8.19)$$

У разі біполярного модульованого сигналу (див. рис. 8.6, а)

$$f(t) = \sum_{k=1}^{\infty} \frac{\sin(\pi k / 2)}{\pi k / 2} \cos k\Omega t. \quad (8.20)$$

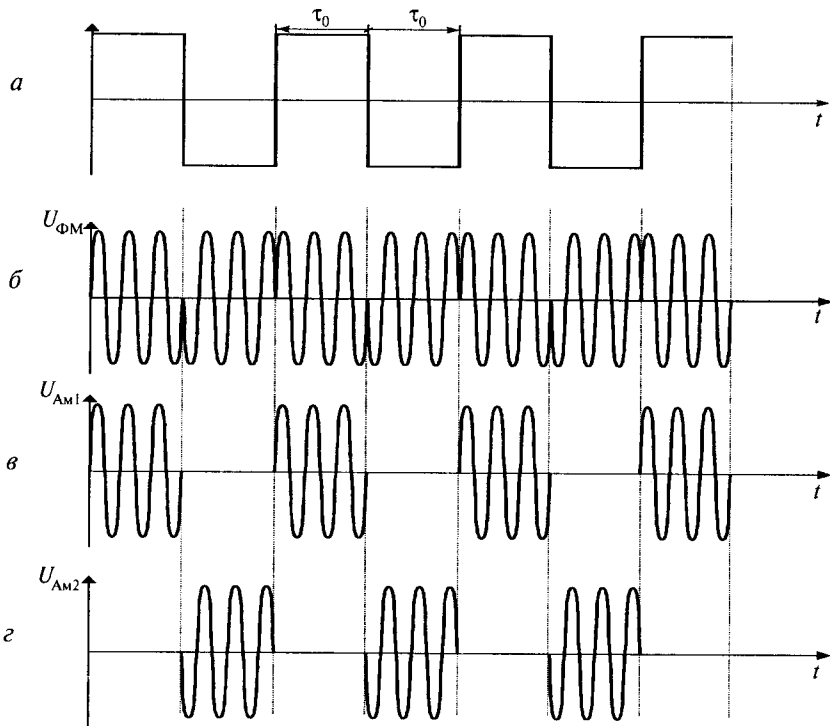


Рис. 8.12. Дискретна ФМ: *a* – модульований сигнал;
б – фазомодульований сигнал ФМ; *в*, *г* – складові сигнали ФМ

Підставивши вираз (8.20) у формулу (8.19), отримаємо

$$\begin{aligned}
 u(t) &= U_M \left\{ \cos(\omega_0 t + \varphi_0) \cos \Delta\varphi - \sin \Delta\varphi \sum_{k=1}^{\infty} \frac{\sin(\pi k / 2)}{\pi k / 2} \cos k\Omega t \sin(\omega_0 t + \varphi_0) \right\} = \\
 &= U_M \cos \Delta\varphi \cos(\omega_0 t + \varphi_0) + \frac{U_M}{2} \sum_{k=1}^{\infty} \frac{\sin \Delta\varphi \sin(\pi k / 2)}{\pi k / 2} \sin(\omega_0 t + k\Omega t + \varphi_0) + \\
 &\quad + \frac{U_M}{2} \sum_{k=1}^{\infty} \frac{\sin \Delta\varphi \sin(\pi k / 2)}{\pi k / 2} \sin(\omega_0 t - k\Omega t + \varphi_0).
 \end{aligned}$$

Отже, у загальному випадку спектр фазомодульованого колювання містить несучу частоту, симетрично від якої розміщуються бокові складові, які віддалені частотними інтервалами, кратними частоті модуляції.

У розглядуваному випадку ($\Delta\varphi = 90^\circ$) спектр ФМ дорівнює спектру АМ у разі придушення несучого коливання. Спектри фазомодульованих сигналів з різними індексами модуляції та спектр амплітудно-модульованого сигналу показано на рис. 8.13.

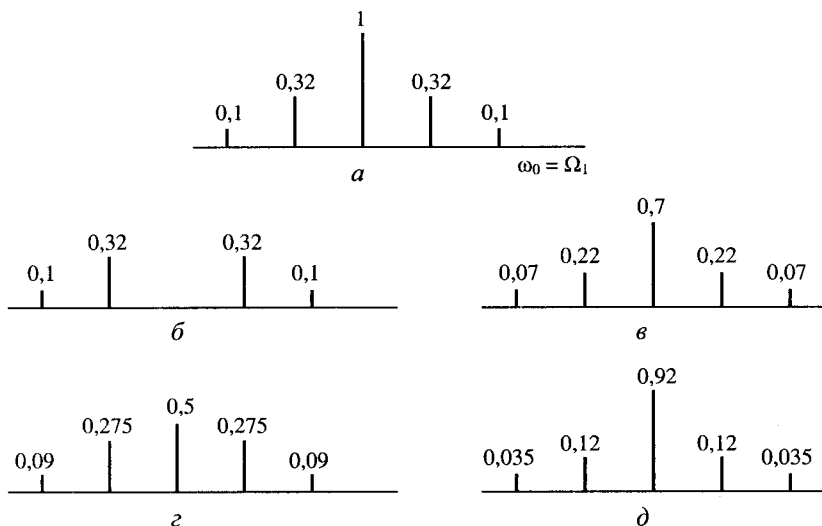


Рис. 8.13. Спектри амплітудно-модульованого (а) і фазомодульованих сигналів для $2\Delta\varphi = 180^\circ$ (б), 120° (в), 90° (г) і 45° (д)

Отже, можна зробити висновок про те, що у разі ФМ періодичної послідовності прямокутних імпульсів:

- ширина спектра ФМ (вірніше її частини, яка несе основну енергію сигналу) дорівнює ширині спектра АМ і не залежить від індексу модуляції;
- амплітуди бокових частот фазомодульованого сигналу відрізняються від таких у разі АМ на величину $\sin \Delta\varphi$.

Розглянемо способи утворення фазомодульованих послідовностей дискретних сигналів прямокутної форми напруги. Ці способи називають відповідно до виду ФМ: абсолютної та відносної.

Сигнали у разі двійкової ФМ можна записати в такому вигляді:

$$u_i(t) = \begin{cases} u_1(t) = u_0 \sin(\omega_0 t + \varphi_1), & \text{якщо } x = x_1; \\ u_2(t) = u_0 \sin(\omega_0 t + \varphi_2), & \text{якщо } x = x_2, \end{cases} \quad (8.21)$$

де x – послідовність передаваних інформаційних символів.

Абсолютна фазова модуляція полягає в тому, що одиничному елементу кожної полярності (позитивної або негативної) відповідає передавання в канал сигналу, зсунутого за фазою на кут φ відносно несучої частоти генератора передавача. Приклад передаваної послідовності одиничних елементів показано на рис. 8.14, *a*, а несучої частоти генератора – на рис. 8.14, *б* і *г*. За час передавання позитивних одиничних елементів напруга сигналу знаходиться у фазі з напругою генератора ($\varphi = 0^\circ$), а за час передавання негативних одиничних елементів – у протифазі з напругою генератора ($\varphi = 180^\circ$), як це показано на рис. 8.14, *в* і *д*. Вираз для цього випадку можна переписати в такому вигляді:

$$u_i(t) = U_0 \sin(\omega_0 t + \varphi_x); \quad \varphi_x = \begin{cases} \varphi_1 = 0, & \text{якщо } x = x_1; \\ \varphi_2 = \pi, & \text{якщо } x = x_2. \end{cases}$$

Отже, у разі передавання двійкової послідовності фаза є випадковою величиною, однозначно зв'язаною з випадковою послідовністю передаваних інформаційних символів x .

Демодуляцію оснований на можливості їх фазового розрізнення за допомогою детектора, чутливого до фази сигналу. Для цього у приймачі має бути опорний сигнал у вигляді копії передаваного сигналу з точністю до фази (синхронного і синхронного з напругою моделювального сигналу (коливання) несучої на передавальному кінці. Створити таку опорну (когерентну) напругу, фаза якої була б постійною, дуже важко.

Можливі три способи отримання опорної напруги:

1) від місцевого генератора, висока стабільність частоти коливань якого забезпечує підтримку синфазності опорної напруги;

2) за допомогою допоміжних пілот-сигналів, які передаються неперервно по тому самому частотному каналу або по спеціальному частотному каналу;

3) безпосередньо із приймального сигналу інформації.

Перший спосіб не забезпечує потрібної синфазності та синхронності коливань, оскільки фаза і частота будь-якого високостабільного генератора змінюється під впливом ряду чинників (зміни температури навколишнього середовища, зміни напруги живлення, спрацювання деталей і т. ін.).

Другий спосіб отримання опорної напруги також не знайшов практичного застосування, оскільки його реалізація призводить до втрат спектра у частотному каналі і потужності на передавання пілот-сигналів.

Найбільшого поширення набули схемні рішення виділення когерентної напруги безпосередньо із приймального інформаційного сигналу. Розглянемо принцип дії цього способу.

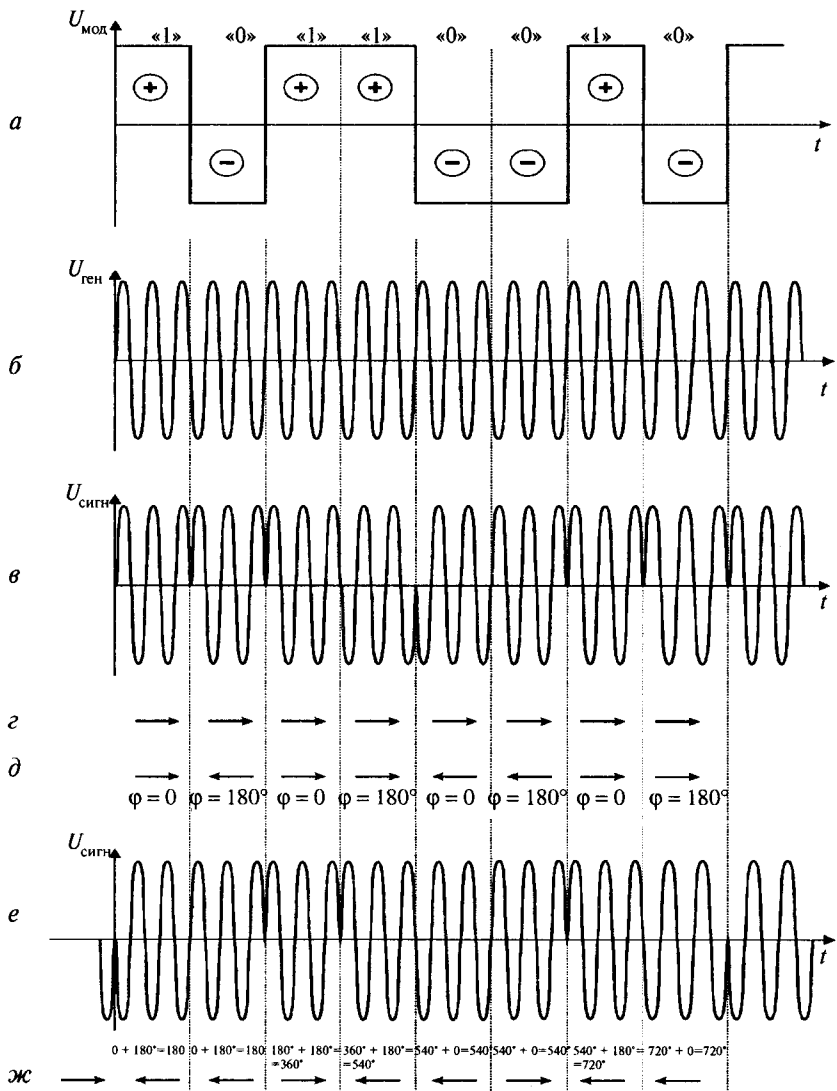


Рис. 8.14. Часові діаграми у разі абсолютної та відносної ФМ

Якщо сигнал, який надходить з каналу зв'язку (рис. 8.14, в), випрямляється двопівперіодним випрямлячем, то змінна складова випрямленої напруги буде мати подвійну частоту без стрибків фази. Отриману подвійну частоту використовують для формування опорного коливання за допомогою подільника частоти (ділить на два), що і є заміною місцевого генератора.

На жаль, ділильні пристрої у разі короткочасного припинення приймання або сильної завади можуть довільно повернути фазу опорної напруги на 180° . Це спричинить зміну полярності посилок на обернену, таким чином призведе до «оберненої роботи». Приймання стане неможливим, оскільки всі знаки будуть прийняті неправильно. Це і є причиною того, що протягом 20 років цей принцип не застосовувався.

Слід зазначити, що внаслідок завад на лінії, флукутацій параметрів каналів зв'язку у часі та інших причин параметри, що надійшли на приймальний кінець сигналу, відрізнялись від тих характеристик приймача, з якими порівнювався сигнал і за допомогою яких виділялась закладена у ньому інформація.

Теоретичні й експериментальні дослідження характеру флукутацій сигналу в каналах зв'язку та аналіз причин їх виникнення в телекомунікаційній техніці зумовлюють розширення впливу флукутації на сигнали. Отже, на відрізок в декілька одиниць або десятків елементарних посилок (залежно від швидкості змінювання параметрів каналу зв'язку і швидкості передавання) можна визначити з достатньою для практики точністю, що канал зв'язку має постійні параметри, котрі в наступний такий же відрізок часу трохи змінюються. Усе це зумовило формулювання нового методу передавання сигналів, за яким інформація, яка переноситься кожною елементарною посилкою, буде виявлятися на приймальному кінці не шляхом порівняння її з характеристиками приймача, а порівнянням з однією із попередніх посилок, або із сусідньою попередньою посилкою. Отже, метод порівняння ґрунтується на порівнянні параметрів сигналу, які відповідають різним моментам часу. На передавальному кінці кожна посилка також маніпулюється відносно попередньої посилки.

Оскільки кожна посилка порівнюється з попередньою, то в літературі установилась назва цього методу як відносного методу передавання сигналів.

Цілком очевидно, що відносний метод передавання можна застосовувати до різних видів модуляції. Використання цього методу до ФМ зумовило появу відносної фазової модуляції (ВФМ). Для реалізації ВФМ у передавальних та приймальних пристроях потрібний запам'ятовувальний пристрій (ланцюг затримки), за допомогою якого можна було б сумістити у часі дві послідовні посилки, що дозволятиме порівнювати їх параметри.

Відносна фазова модуляція має відмітну особливість, яка визначається способом формування передавального фазомодульованого сигналу. Якщо за абсолютної ФМ фаза несучої змінюється з кожною зміною полярності передаваних посилок (див. рис. 8.14, *в*), то за ВФМ вона змінюється під час передавання кожної елементарної послідовності тільки однієї полярності, наприклад позитивної (рис. 8.14, *е*), або, що те саме, з надходженням кожної одиниці. Це означає, що у разі передавання підряд декількох позитивних посилок (одиниць) фаза несучої буде змінюватись з початку кожної елементарної послідовності.

Нехай під час передавання позитивного одиничного елемента фаза сигналу зсувається відносно фази попереднього сигналу на $\varphi = 0^\circ$. Тоді (рис. 8.14) протягом часу передавання першого (позитивного) одиничного елемента фаза сигналу буде протилежна деякій початковій фазі допоміжного сигналу (рис. 8.14, *е*). Напочатку передавання другого (негативного) одиничного елемента фаза сигналу не змінюється, бо негативному одиничному елементу відповідає фазовий зсув $\varphi = 0^\circ$ відносно попереднього одиничного елемента (рис. 8.14, *е, ж*). Для визначення інформації, яка міститься у першому одиничному елементі повідомлення; перед ним має бути попередній допоміжний одиничний елемент з довільною фазою несучої частоти.

Реалізація ВФМ потребує однакової для передавача і приймача відповідності між значеннями послідовності a_i і значеннями різниці фаз $\varphi_i - \varphi_{i-1}$. Будемо вважати, що послідовностям $a_i = 0$ і $a_i = 1$ відповідають:

$$\left. \begin{aligned} \varphi_i - \varphi_{i-1} &= \Delta\varphi_1 = 0, \\ \varphi_i - \varphi_{i-1} &= \Delta\varphi_2 = 180^\circ. \end{aligned} \right\}$$

Умовимося також, що передаванню $a_i = 0$ відповідає передавання послідовності негативної полярності, а $a_i = 1$ – передавання послідовності позитивної полярності, тобто

$$\left. \begin{aligned} a_i = 0 &\rightarrow -A, \\ a_i = 1 &\rightarrow +A. \end{aligned} \right\}$$

У цьому разі закон маніпуляції фази несучого коливання у разі ВФМ буде таким. Під час передавання послідовності $a_i = 1$ фаза несучого коливання стрибком змінюється на 180° відносно фази $(i - 1)$ -ї послідовності, а під час передавання послідовності $a_i = 0$ вона остається такою самою, що й у $(i - 1)$ -й послідовності.

8.5. Багатопозиційні методи модуляції

Розглянуті методи дискретної модуляції являли собою двійкові методи, оскільки модульовальні параметри несучої частоти (амплітуда, частота, фаза) набували двох можливих значень. Поряд з двійковими методами модуляції існують методи, за яких модульовальний параметр може набувати $m > 2$ значень. Модуляцію, за якої модульовальний параметр може мати більше двох розрізняваних станів, називають *багаторівневою*. Так, наприклад, відомі багаторівневі методи АМ, багаточастотні методи ЧМ, багатократні методи ВФМ, сумісне використання різних методів модуляції.

Багаторівневі методи АМ не отримали широкого застосування через низьку завадостійкість. У радіозв'язку по короткохвильових каналах широко використовують двократну частотну модуляцію. Метод багатократної модуляції можна розглядати і як своєрідне кодування, за якого кожному значенню модульовального параметра ставиться у відповідність не двійкова одиниця інформації, яка дорівнює одному біту, як це має місце в однократних методах модуляції, а декілька бітів. Так, наприклад, за двократної частотної модуляції кожній частоті відповідає 2 біт (1 дібіт) інформації.

Можливість відносно простими засобами створити в одному частотному каналі декілька каналів передавання двійкової інформації – одна з важливих переваг методу ФМ, що сприяло його поширенню (у модифікації відносної ФМ) у практиці передавання двійкових сигналів.

Розглянемо використання ВФМ для утворення багаторівневої ФМ. Принцип багатократної ВФМ передавання полягає в тому, що кожній послідовності посилок, які передаються окремим каналом, присвоюється певне значення фазового кута несучого коливання. Якщо по кожному каналу передаються двійкові послілки і кількість каналів дорівнює N , то загальна кількість можливих комбінацій посилок буде дорівнювати 2^N . Отже, для N -кратної ФМ передавання треба використати 2^N різних значень фазових зсувів несучого коливання.

Зазвичай значення різниці фаз сусідніх посилок вибирають кратними деякому мінімальному куту маніпуляції, який відрізняється від нуля і дорівнює

$$\Delta\varphi_{\min} = \frac{2\pi}{2^N}.$$

Для наочності в табл. 8.1 наведено значення $\Delta\varphi_{\min}$ з різною кількістю каналів N .

Таблиця 8.1

| N | 1 | 2 | 3 | 4 | 5 |
|------------------------|-------------|------------|------------|--------------|--------------|
| $\Delta\varphi_{\min}$ | 180° | 90° | 45° | $22,5^\circ$ | $11,2^\circ$ |

Розглянемо метод двократної системи відносної фазової модуляції (ДВФМ). Суть методу ДВФМ полягає в тому, що для передавання інформації першим двійковим каналом використовують, наприклад, фазові положення несучої 0 і 180° , а для передавання інформації другим двійковим каналом – фазові положення 90° і 270° . Положення векторів коливань показано на рис. 8.15.

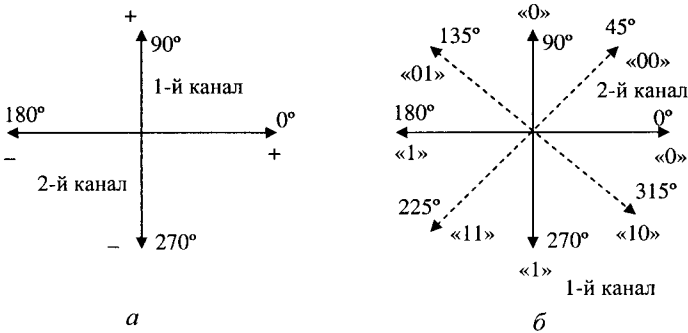


Рис. 8.15. Векторні діаграми, які пояснюють фазові положення векторів двох підканалів у разі ДВФМ: *a* – фазові положення векторів каналів; *б* – фазові положення під час сумісної роботи каналів

Отже, за найбільш поширених дво- і трикратних ВФМ $\Delta\varphi_2 = \frac{\pi}{2}$, $\Delta\varphi_3 = \frac{\pi}{4}$. Правило кодування у разі ДВФМ за першим варіантом ілюструє табл. 8.2.

Таблиця 8.2

| | | | | |
|---|---|---------|-------|----------|
| Символ 1-го каналу | 0 | 0 | 1 | 1 |
| Символ 2-го каналу | 0 | 1 | 0 | 1 |
| Різниця фаз $\Delta\varphi = \varphi_n - \varphi_{n-1}$ | 0 | $\pi/2$ | π | $3\pi/2$ |

Правило кодування за другим варіантом, рекомендованим міжнародною організацією зв'язку (рекомендація V26') ілюструє табл. 8.3.

Таблиця 8.3

| | | | | |
|---|---------|----------|----------|----------|
| Символ 1-го каналу | 0 | 0 | 1 | 1 |
| Символ 2-го каналу | 0 | 1 | 0 | 1 |
| Різниця фаз $\Delta\varphi = \varphi_n - \varphi_{n-1}$ | $\pi/4$ | $3\pi/4$ | $7\pi/4$ | $5\pi/4$ |

Перший варіант дещо простіший щодо реалізації, але другий варіант дозволяє виділяти із приймального сигналу синхронного опорного коливання у разі тривалого передавання каналом комбінації символів, які складаються з одних нулів. Дійсно, за першим варіантом поворотів фази не відбувається, результатом чого виявляється неможливість виділити тактову частоту, що призводить до порушення синхронізації.

У разі ДВФМ для передавання інформації першим двійковим каналом використовують, наприклад, фазові зсуви $\pi/2$ і $3\pi/2$, а другим двійковим каналом – 0 і 2π , що зображено векторними діаграмами на рис. 8.15. На цьому рисунку суцільними лініями показано фазові положення векторів окремих каналів, а пунктиром – фазові положення векторів у разі сумісної роботи двох каналів, які відповідають другому варіанту. Таким чином, будь-якій кодовій комбінації одиничних елементів у кожному з двійкових каналів відповідає певний зсув за фазою (табл. 8.4).

Таблиця 8.4

| | | | | | | |
|---|----|-------------------------|--------------------------|-------------------------|--------------------------|--------------------------|
| Символ 1-го каналу | 0 | 1 | 1 | 0 | 1 | 0 |
| Символ 2-го каналу | 0 | 0 | 1 | 0 | 1 | 1 |
| φ, \dots° | 45 | $45 + 315 =$ $= 360$ | $360 + 225 =$ $= 225$ | $225 + 45 =$ $= 270$ | $270 + 225 =$ $= 135$ | $135 + 135 =$ $= 270$ |
| $\Delta\varphi = \varphi_n - \varphi_{n-1}$ | 45 | 315 | 225 | 45 | 225 | 135 |

Демодулятор ДВФМ побудовано так, що у разі зсуву фаз між попереднім і наступним одиничними елементами на 45° на виходах обох каналів отримуються нулі, якщо $\Delta\varphi = 135^\circ$, на виході першого каналу – нуль, другого – одиниця і, якщо $\Delta\varphi = 315^\circ$, на виході першого каналу – одиниця, а другого – нуль.

У разі ДВФМ на такій самій швидкості модуляції, що і за ВФМ, забезпечується у два рази більша ефективна швидкість передавання, оскільки кожний фазовий стан відповідає не одному біту інформації (як у разі ВФМ), а двом бітам (по одному в кожному каналі).

Аналогічно ДВФМ можна побудувати ТВФМ – трикратну відносну фазову модуляцію. При цьому знижується завадостійкість порівняно з ДВФМ і сильно ускладнюється обладнання. Тому питомі швидкості передавання збільшуються за сумісного використання амплітудної і фазової

модуляції, причому для скорочення спектра сигналу застосовують методи передавання однієї бокової смуги частот. Приклад шестипозиційного сигналу, який створено комбінацією трирівневої амплітудної та однократної фазової модуляції АМ-ФМ показано на рис. 8.16.

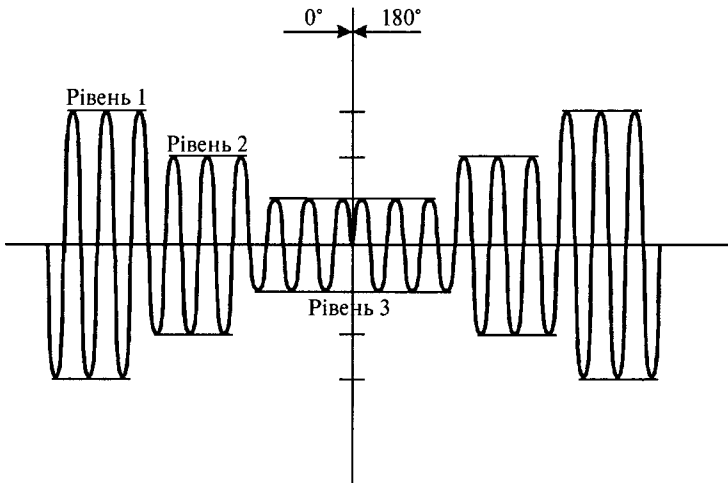


Рис. 8.16. Шестипозиційний сигнал (три рівні амплітуд і однократна фазова модуляція)

У модемах для телефонних каналів здебільшого використовують три види модуляції: частотну, відносну фазову і квадратурну амплітудну модуляції, яку часто називають багатопозиційною амплітудно-фазовою. До цього ми детально розглянули частотну та відносну фазову модуляції з визначенням спектрів модульованих послідовностей імпульсів. Частотна модуляція дуже завадостійка, оскільки завади телефонного каналу спотворюють переважно амплітуду, а не частоту сигналу. Однак за частотної модуляції неекономно витрачається ресурс смуги частот телефонного каналу. Тому цей вид модуляції застосовується у низькошвидкісних протоколах, які дозволяють здійснювати зв'язок каналами з низьким відношенням сигнал/шум.

Частіше застосовують чотирифазову ВФМ (ВФМ-4) або ДВФМ з двома наборами фаз: залежно від значення дібіта (00, 01, 10, 11) фаза змінюється на 0° , 90° , 180° , 270° або 45° , 135° , 225° , 315° відповідно. При цьому, якщо кількість кодованих бітів більша від трьох (8 позицій повороту фази), різко знижується завадостійкість ВФМ. Тому для високошвидкісного передавання даних ВФМ не використовують.

8.6. Квадратурна амплітудна модуляція

У разі квадратурної амплітудної модуляції (КАМ, QAM – Quadrature Amplitude Modulation) змінюється як фаза, так і амплітуда сигналу, що дозволяє збільшити кількість кодованих бітів і при цьому суттєво підвищити завадостійкість. Тепер використовують способи модуляції, у яких кількість кодованих на одному бодовому інтервалі інформаційних бітів може досягнути 8–9, а кількість позицій сигналу в сигнальному просторі – 256–512.

Квадратурне подання сигналів – зручний і досить універсальний засіб їх опису. Квадратурне подання можна записати як коливання лінійної комбінації двох ортогональних складових – синусоїдної і косинусоїдної:

$$S(t) = x(t)\sin(\omega t + \varphi) + y(t)\cos(\omega t + \varphi), \quad (8.22)$$

де $x(t)$ і $y(t)$ – біполярні дискретні величини.

Така дискретна модуляція (маніпуляція) здійснюється по двох каналах на несучих, зсунутих на 90° один відносно одного, тобто знаходяться в квадратурі (звідси і назва подання і методу формування сигналів).

Пояснимо процес формування сигналів чотирифазової модуляції ФМ (ФМ-4) на прикладі роботи квадратурної схеми (рис. 8.17).

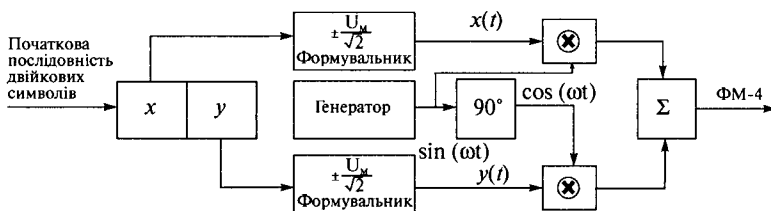


Рис. 8.17. Схема квадратурного модулятора

Початкова послідовність двійкових символів тривалістю T за допомогою регістра зсуву розділюється на непарні імпульси y , які подаються в квадратурний канал ($\cos \omega t$), і парні – x , які надходять у синфазний канал ($\sin \omega t$). Обидві послідовності імпульсів надходять на входи відповідних формувальників маніпулювальних імпульсів, на виходах яких утворюються послідовності біполярних імпульсів $x(t)$ і $y(t)$. Маніпулювальні імпульси мають амплітуду $U_m/\sqrt{2}$ і тривалість $2T$. Імпульси $x(t)$ і $y(t)$ надходять на входи каналних перемножувачів, на виходах яких формуються двофазові $(0, \pi)$ ФМ коливання. Після підсумовування вони утворюють сигнал ФМ-4. Згідно з методом формування сигнал ФМ-4 також називають *квадратурним фазомодульованим сигналом* (QPSK – Quadrature PSK).

Якщо одночасно змінюються символи в обох каналах модулятора (з 10 на 01, або з 00 на 11), у сигналі ДВФМ виникає стрибок фази на 180° (π). Такі стрибки фази також відбуваються і за звичайної двофазової модуляції (ФМ-2), викликаючи паразитну амплітудну модуляцію обвідної сигналу. У результаті цього під час проходження сигналу через вузько-смуговий фільтр виникають «провали» обвідної до нуля. Такі зміни не бажані, оскільки призводять до збільшення енергії бокових смуг і завад у каналі зв'язку.

Чотирифазова ФМ зі зсувом (OQPSK – Offset QPSK) (рис. 8.18) дозволяє уникнути стрибків фази на 180° і, отже, глибокої модуляції обвідної. Сигнал у квадратурній схемі формується так само, як і в модуляторі ФМ-4, за винятком того, що маніпуляційні елементи інформаційної послідовності $x(t)$ і $y(t)$, зсунені за часом на тривалість одного елемента T (рис. 8.18, б, в). Зміна фази у разі такого зсування модульованих потоків визначається лише одним елементом послідовності, а не двома, як за ФМ-4. У результаті стрибків фази на 180° не буває, оскільки кожний елемент послідовності, яка надходить на вхід модулятора синфазного або квадратурного каналу, може викликати зміну фази на 0° , $+90^\circ$ або -90° .

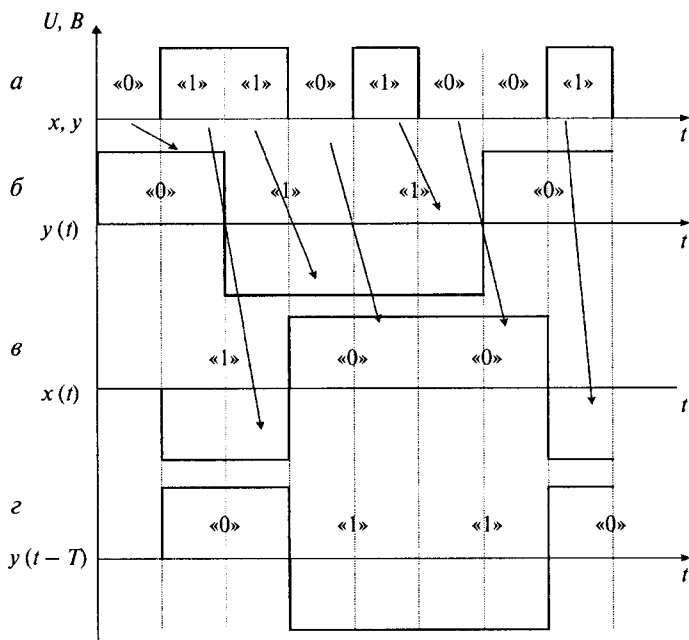


Рис. 8.18. Формування сигналів OQPSK

Для виразу для опису сигналу (8.22) характерна взаємна незалежність багаторівневих маніпулювальних імпульсів $x(t)$, $y(t)$ у каналах, тобто одиничному рівню в одному каналі може відповідати одиничний або нульовий рівень у другому каналі. У результаті вихідний сигнал квадратурної схеми змінюється не тільки за фазою, але й за амплітудою. Оскільки в кожному каналі здійснюється амплітудна маніпуляція, цей вид модуляції називають *квадратурною маніпуляцією* зі зміною амплітуди (QASK – Quadrature Amplitude Shift Keying) або просто КАМ.

Використовуючи геометричне зображення, кожний сигнал КАМ можна зобразити вектором у сигнальному просторі. Якщо позначати тільки кінці векторів, для сигналів КАМ отримаємо зображення у вигляді сигнальної точки, координати якої визначаються значеннями $x(t)$ і $y(t)$. Сукупність сигнальних точок утворює *сигнальне сузір'я* (signal constellation). Структурну схему модулятора і сигнальне сузір'я для випадку, коли $x(t)$ і $y(t)$ набувають значень ± 1 , ± 3 (чотирирівнева КАМ), показано на рис. 8.19. Значення ± 1 , ± 3 визначають рівні модуляції і мають відносний характер. Сузір'я містить 16 сигнальних точок, кожна з яких відповідає чотирьом передаваним інформаційним бітам.

Комбінація рівнів ± 1 , ± 3 , ± 5 може сформувати сузір'я із 36 сигнальних точок. Проте з них у протоколах ITU-T використовують тільки 16 рівномірно розподілених у сигнальному просторі точок.

Існує декілька способів практичної реалізації чотирирівневої квадратурної амплітудної модуляції.

Застосування багатопозиційної КАМ у чистому вигляді пов'язано з проблемою недостатньої завадостійкості. Тому у всіх сучасних високошвидкісних протоколах (процедурах) КАМ використовують сумісно з решітчастим кодуванням – спеціальним видом згорткового кодування. У результаті з'явився новий спосіб модуляції – *треліс-модуляція* (TSM – Trellis Coded Modulation). Вибрану певним чином комбінацію конкретної КАМ завадостійкого коду в технічній літературі ще називають *сигнально-ковою конструкцією* (СКК). Сигнально-кова конструкція дозволяє підвищити завадостійкість передавання інформації поряд зі зниженням вимог до відношення сигнал/шум у каналі на 3...6 дБ. При цьому кількість сигнальних точок збільшиться вдвоє за рахунок додавання до інформаційних бітів одного надлишкового біта, утвореного згортковим кодуванням. Розширений таким чином блок бітів підлягає все тій же КАМ. У процесі демодуляції виконується декодування прийнятого сигналу за алгоритмом Вітербі. Цей алгоритм з використанням уведеної надлишковості та знань передісторії процесу приймання дозволяє за критерієм максимальної правдоподібності вибрати із сигнального простору найбільш достовірну еталонну точку.

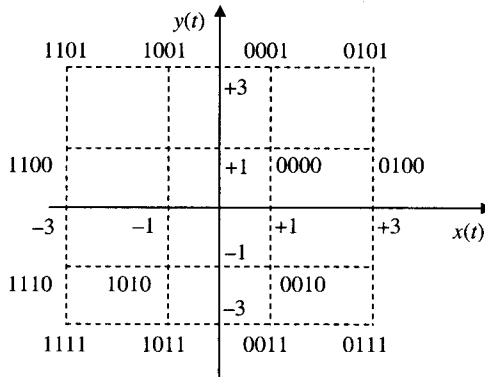
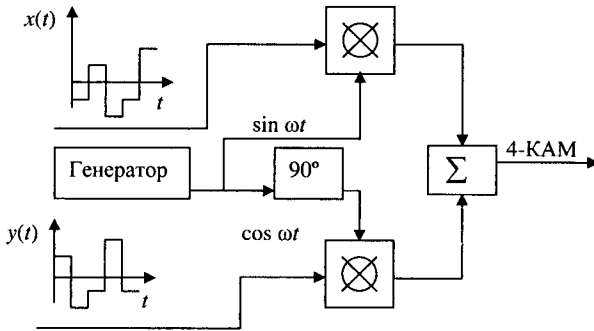


Рис. 8.19. Схема модулятора і сигнальна діаграма КАМ-4

Вибір способів модуляції і кодування зводиться до пошуку такого заповнення сигнального простору, за якого забезпечується висока швидкість та висока завадостійкість. Комбінування різних ансамблів багатопозиційних сигналів і завадостійких кодів породжує багато варіантів сигнальних конструкцій. Узгоджені варіанти, які забезпечують поліпшення енергетичної і частотної ефективності, і є сигнально-кодovими конструкціями. Пошук найкращої СКК – одне з найскладніших завдань теорії зв'язку.

9. Методи реєстрації дискретних сигналів

Одна з основних функцій приймальної частини СПДІ – реєстрація значущої позиції сигналу, який надходить із лінії зв'язку на приймальні пристрої системи. Оптимальний вибір сигналу, який зазнає впливу характеристик каналу і завад, що діють на цей канал, виконується передавальною частиною СПДІ.

Спотворення сигналів каналами зв'язку як перетворювачем сигналів порушує бажану рівність $S_{\text{вх}} = S_{\text{вих}}$, де $S_{\text{вх}}$ – сигнал на вході каналу зв'язку; $S_{\text{вих}}$ – сигнал на виході каналу зв'язку.

Отже, спотворені сигнали надходять у приймальну частину СПДІ, яка представлена пристроєм перетворення сигналів ППСпр. В основі більшості таких приймачів лежить фільтрація із заданими частотними характеристиками та частотною смугою пропускання, а процес реєстрації зводиться до аналізу обвідної приймального сигналу.

Способи приймання сигналу залежно від методу виділення моделювальної функції можна поділити на дві групи: когерентне і некогерентне детектування. За когерентного детектування опорні сигнали являють собою точну копію передавальних сигналів (з точністю до початкової фази). У разі некогерентного детектування апріорні відомості про початкову фазу частоти носія не враховують.

Приймач приймає рішення на підставі аналізу прийнятого сигналу (імпульсу постійного струму, обвідна якого являє собою зображення спотвореного сигналу), причому час цього аналізу не перевершує τ_0 – тривалості імпульсної послідовності передавального сигналу, оскільки після закінчення цього часу здійснюється приймання наступного сигналу зі всієї послідовності кодової комбінації. За видами прийняття рішень (реєстрації) способи приймання поділяють на три групи: стробування, інтегральні та кореляційні.

Інформація передається у вигляді послідовності одиночних сигналів, які під час проходження каналом зв'язку можуть бути спотворені.

Установлення значущої позиції (0 або 1) кожного прийнятого одиночного сигналу (елемента) називають *реєстрацією*, а пристрій, який виконує цю процедуру, – *реєструвальним пристроєм* (РП).

Перейдемо до розгляду методів реєстрації, які використовують для створення ППСпр.

9.1. Реєстрація сигналів методом стробування

Метод стробування найпростіший. Він полягає в тому, що приймальний сигнал аналізується протягом інтервалу τ_a в одній точці, тобто виконується одновимірна вибірка випадкового процесу $\tau(t)$, яка потім порівнюється з порогом, і за результатом цього порівняння приймається рішення. У цьому разі сигнал підлягає перетворенню типу згортки з δ -функцією, тобто

$$r'(t) = \int_0^{\tau_0} r(t)\delta(t - t_a)dt,$$

де $r'(t)$ – сигнал на виході РП; $r(t)$ – сигнал на його вході; $0 \leq t_a \leq \tau_0$ – момент реєстрації.

Суть реєстрації дискретних сигналів методом стробування полягає в тому, що одиночний сигнал аналізується не на всьому інтервалі T_0 , а на деякій частині $t_a < T_0$. При цьому значення одиночного сигналу до початку аналізу і після прийняття рішення не враховуються.

Зазвичай інтервал часу аналізу t_a («строба») вибирається значно менший від T і розміщується в середній частині значущого інтервалу T_0 . Це пояснюється тим, що середня частина одиночного сигналу менше спотворена, ніж крайня.

Поріг реєстрації сигналу вибирають таким, щоб він дорівнював половині номінального значення амплітуди сигналу.

Принцип дії РП методом стробування розглянемо за рис. 9.1.

Схема складається із двох двовходових комірок I (один вхід $I2$ інверсний) і вихідного тригера T . На перші входи комірок $I1$ і $I2$ подається вхідний сигнал (рис. 4.1, б), а на другі входи – імпульси реєстрації (PI), тривалість яких $t_a < T_0$. Причому момент надходження PI збігається з прийманням середини T_0 . Із рис. 9.1 видно, що з надходженням одиничних елементів із спотвореннями вони будуть зареєстровані правильно (тригер сформує сигнал без помилки), якщо спотворення не збігаються з імпульсами реєстрації.

Якщо спотворення виникають у момент дії імпульсу реєстрації (елемент 5), то на виході РП буде зареєстровано неправильно п'ятий елемент.

Для виявлення характерних спотворень у процесі вимірювання каналів використовують послідовності контрольних сигналів 5 : 1 (п'ять одиниць і один нуль) або 1 : 5 (одна одиниця і п'ять нулів).

Регулярні крайові спотворення типу переважання – це результат порушення симетрії живильних напруг тракту передавання. Ці порушення призводять до того, що одиничні елементи однієї полярності виявляються всі, наприклад, подовженими, а другої – укороченими.

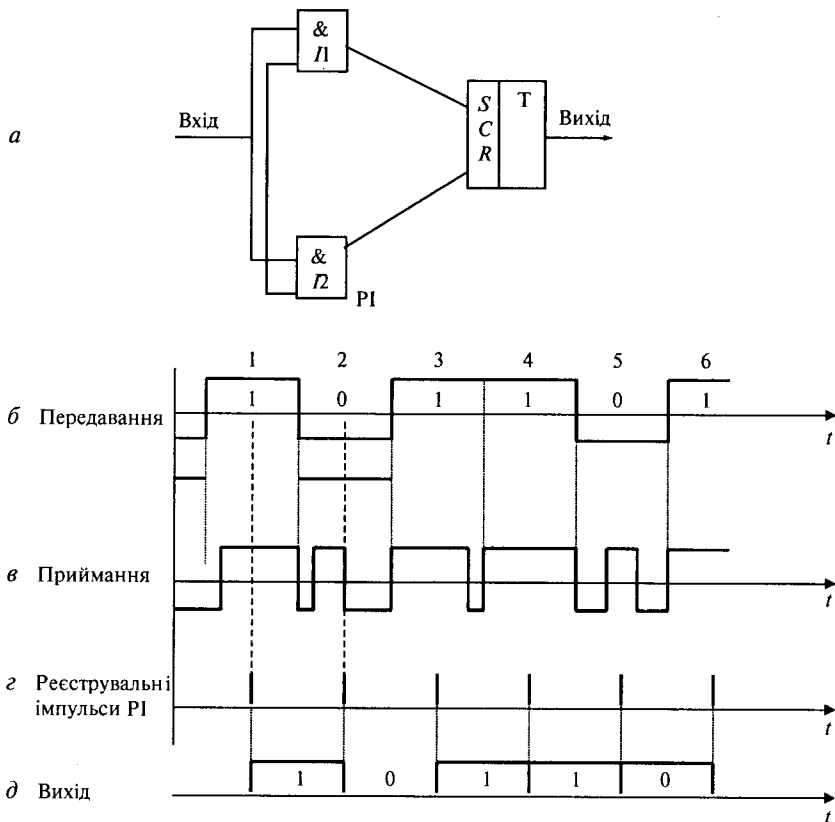


Рис. 9.1. Структурна схема реєструвального пристрою методом стробування (*a*) і часова діаграма (*б, в, г, д*)

Такі спотворення називають спотвореннями переважання. Регулярні спотворення самі по собі не призводять до помилок, оскільки їх можна врахувати. Але коли вони збігаються з випадковими спотвореннями, збільшується імовірність виявлення помилок.

Випадкові крайові спотворення спричиняють зовнішні впливи на електричний сигнал під час його проходження каналом зв'язку. Джерелом зовнішніх впливів можуть бути будь-які атмосферні явища (грозові розряди, магнітні бурі і т. ін.), імпульсні завади, стрибкоподібні зміни рівня приймального сигналу тощо.

Спотворення дробленням – це спотворення дискретних сигналів, спричинені появою додаткових пар значущих моментів (одного або декількох) усередині значущого інтервалу. Причини дроблення – завади, які перевищують рівень корисного сигналу (переважно імпульсного), і короткочасові перерви.

Кількісно спотворення дроблень оцінюють за двома параметрами:

- тривалістю дроблень;
- моментами часу початку дроблень.

Тривалість дроблень вимірюється абсолютним значенням τ , яке дорівнює інтервалу часу між значущим моментом сигналу, що утворюється під дією завади.

Дроблення, які мають однакову тривалість τ , неоднаково впливають на результати приймання одиночних сигналів. Одні з них можуть призвести до помилки (замість одиниці нуль і навпаки нуль замість одиниці), а інші ні. Тому, крім тривалості дроблення, треба знати момент виникнення дроблення відносно значущого моменту – це відстань x (рис. 9.2). Різниця крайових спотворень та спотворень дробленням виникають випадково.

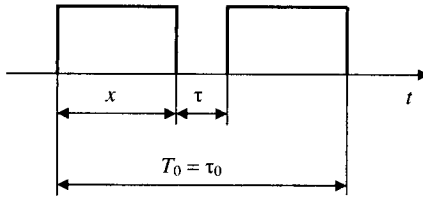


Рис. 9.2. Часова діаграма виникнення дроблення

У зв'язку зі слабкою захищеністю методу стробування від дії дроблень завжди вживають додаткових заходів, які зменшують вплив дроблень (тривалість строби беруть $t_a = T_0/4$).

Тривалість прийнятих методом стробування рішень визначають одновимірними статистичними характеристиками випадкового процесу $r(t)$. І хоча це рішення виконується в середній частині елементарного сигналу, що найменше зазнає спотворень, існує певна ймовірність прийняття помилкового рішення.

9.2. Реєстрація сигналів інтегральним методом

Прагнення зменшити ймовірність виникнення помилки природно приводить до потреби прийняття рішення на підставі не однократного, а багатократного аналізу сигналу або його неперервного аналізу на інтервалі τ_0 . Такий метод реєстрації називають інтегральним, оскільки перетворення, якому підлягає процес $r(t)$, має вигляд

$$r'(t) = \int_0^{\tau_0} r(t) dt,$$

а рішення приймається в момент часу t_a на підставі порівняння отриманої суми миттєвих значень $r(t)$ від 0 до τ_0 .

Для підсумовування декількох відліків складаються як корисні сигнали, так і миттєві значення завад. Величина корисного сигналу у всіх доданках (складових) частин однакова, тому їх сума збільшується пропорційно кількості відліків на інтервалі інтегрування n , а її квадрат – пропорційно n^2 . Середній же квадрат суми незалежних випадкових величин зростає пропорційно n . У результаті вигравш інтегрального методу у відношенні сигнал/завада дорівнює n .

Інтегральний метод можна реалізувати неперервним і дискретним підсумовуванням відліків.

Суть інтегрального методу реєстрації дискретних сигналів полягає в тому, що одиничний елемент аналізується на інтервалі $t_a < T_0$ і рішення про його значення приймається в кінці інтервалу t_a .

Аналіз сигналу, що надходить на вхід РП, можна виконувати двома методами:

1) неперервним у часі спостереженням за вхідним сигналом (цей спосіб ще називають *аналоговим інтегрувальним методом* (АІМ));

2) протягом великого інтервалу тривалості сигналу T_0 у дискретні моменти виконується багаторазове стробування з накопиченням значень для кожного моменту і кінцевим підрахуванням кількості правильно прийнятих елементарних стробових імпульсів та дотриманням правила більшості під час прийняття рішення про його значення.

Цей метод одержав назву «*дискретний інтегральний метод реєстрації*» (ДІМ). Розглянемо основні принципи положення цих методів.

9.3. Аналоговий інтегральний метод реєстрації

За цим методом визначення значень значущої позиції виконується неперервно протягом інтервалу t_a – інтервалу аналізу (спостереження).

Основна ідея методу полягає в тому, що з надходженням із лінії зв'язку кожного спотвореного сигналу до приймача елемент РП накопичує енергію, пропорційну інтервалу аналізу t_a одиночного значущого інтервалу. Таку залежність можна отримати у разі використання інтегрувального ланцюга, який має постійну інтегрування $\tau = RC$. Керування роботою інтегрувального ланцюга комірки та визначення результату інтегрування за кожний інтервал аналізу t_a здійснюється за допомогою порогового пристрою.

Розглянемо цей процес під час роботи функціональної схеми (рис. 9.3).

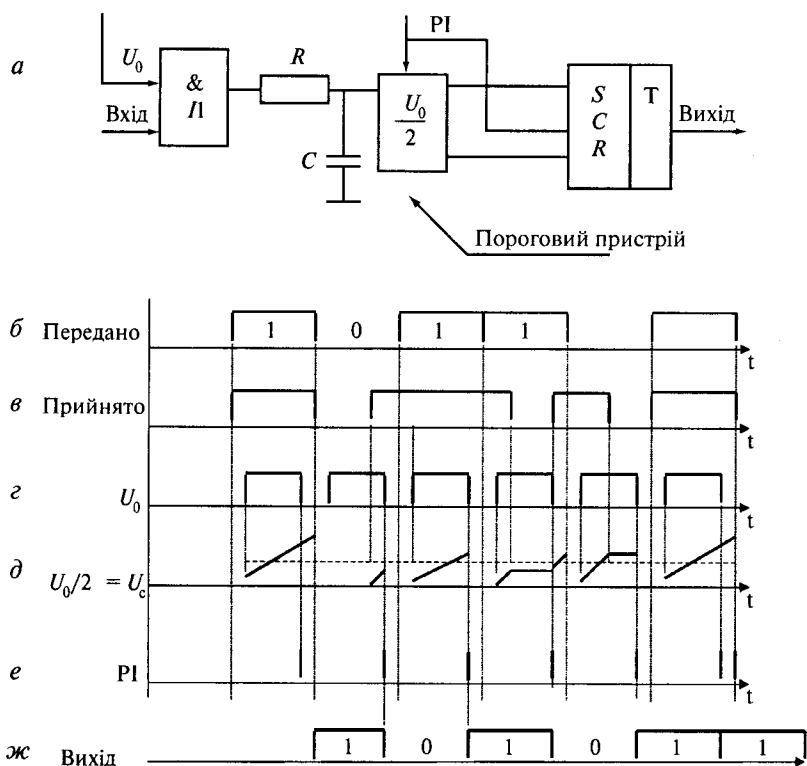


Рис. 9.3. Структурна схема РП (а) методом АІМ та часова діаграма його роботи (б, в, г, д, е, ж)

У цій схемі за час дії одиничного елемента на смності C інтегрувальної комірки накопичується енергія, пропорційна інтервалу аналізу t_a одиничного значущого інтервалу. Така залежність забезпечується тим, що постійна інтегрувального ланцюга в декілька разів (3–5) більша, ніж T_0 . Завдяки цьому напруга на C буде лінійно збільшуватись залежно від часу заряду.

Реєстрацію приймальних одиночних сигналів розглянемо на часовій діаграмі (рис. 9.3), на якій: b – послідовність сигналів на виході ППС передавальної частини СПДІ; v – послідовність сигналів на вході порогового пристрою (обвідні спотворених корисних сигналів); z – послідовність опорних тактових сигналів; d – послідовність сигналів, які виникають на виході інтегрувального ланцюга (RC); e – послідовність реєструвальних імпульсів; $ж$ – послідовність сигналів на виході реєструвального тригера T .

У разі одностороннього надходження на вході $П$ сигналів з лінії зв'язку і опорного U_0 на виході інтегрувального ланцюга виникає напруга і надходить на вхід порогового пристрою з порогом спрацювання $U_0/2$.

Під час приймання неспотвореного одиночного сигналу або зі спотворенням, у якому неспотворена частина інтервалу аналізу більша за $t_a/2$, конденсатор C заряджається до величини, більшої від $U_0/2$ (U_0 – напруга на конденсаторі під час приймання неспотвореного сигналу). У момент появи на вході порогового пристрою імпульсу реєстрації (PI) і рівня сигналу $\geq U_0/2$ на його виході виникає сигнал, який діє на C , вхід тригера T , який переводить його в стан, що відповідає переданому сигналу.

Якщо одиничний елемент спотворюється так, що неспотворена частина менша ніж $t_a/2$, тоді напруга на C буде менша за $U_0/2$ і з надходженням імпульсу реєстрації пороговий пристрій переведе тригер у стан, який відповідає сигналу виникнення помилки.

З часової діаграми видно, що цей метод привносить затримку сигналу на величину t_a , оскільки реєстрація здійснюється наприкінці кожного інтервалу аналізу прийнятого одиничного елемента, тоді отримана послідовність елементів буде зсунута на величину t_a відносно переданої.

9.4. Дискретний інтегральний метод реєстрації

За дискретним інтегральним методом реєстрації кожний одиночний сигнал, який надходить з лінії зв'язку на інтервалі аналізу t_a оцінюється декількома імпульсами в ξ точках, підраховуються однакові значення і за критерієм більшості приймається рішення про значення цього елемента.

Для уникнення невизначеності зазвичай кількість точок беруть не парним.

Дієздатність методу розглянемо за структурною схемою (рис. 9.4). Реєструвальний пристрій складається із схеми збігу Π , на яку надходять приймальні одиночні сигнали з лінії зв'язку, стробувальні імпульси (СІ) та сигнали з нульового виходу лічильника стробів (ЛС). Виходи ЛС підключаються до входів реєструвального тригера (синхронізованого однотактного RS-тригера). На рис. 9.4 показано часову діаграму, на якій: b – послідовність сигналів, які надходять у канал зв'язку; v – послідовність сигналів, які надходять на вхід пристрою реєстрації у спотвореному вигляді; z – послідовність стробувальних імпульсів СІ; d – послідовність стробувальних імпульсів, які надходять на вхід ЛС; e – сигнали на виході ЛС; $ж$ – послідовність реєструвальних імпульсів, які надходять на «нульовий вхід» ЛС та на вхід C тригера; i – послідовність імпульсів на виході реєструвального тригера.

Правило більшості реалізується ЛС згідно з виразом $(\epsilon + 1)/2$, де ϵ – визначена непарна кількість стробувальних імпульсів.

У початковому стані на перший вхід системи Π надходить періодична послідовність груп стробувальних імпульсів, в кожній групі однакова кількість коротких за довжиною стробувальних імпульсів і дорівнює априорі визначеному числу ϵ , а на третій вхід схеми Π надходить сигнал з нульового виходу ЛС (нульового виходу тригера ЛС). Розглянемо процес реєстрації першого сигналу (рис. 9.4) за умови, що він надійшов без спотворень. У цьому разі перші три стробувальні імпульси проходять Π і надходять на лічильний вхід (+1) ЛС. Згідно з алгоритмом роботи ЛС на виході (рис. 9.4, a) виникає сигнал, відповідно на виході (рис. 9.4, v) зникає сигнал; цією дією забороняється надходження стробувальних імпульсів на ЛС (рис. 9.4, e). У разі надходження реєструвального імпульсу в кінці інтервалу аналізу t_a тригер під дією сигналів на входах S і C переходить в одиничний стан (рис. 9.4, i). Отже, коли ЛС зареєструє три стробувальні імпульси, тоді не виникає помилки.

Помилка виникає під час реєстрації шостого сигналу для випадку, коли t_d – тривалість інтервалу дроблення, яке виникло під дією завади і заборонило ЛС зареєструвати число $(5 + 1)/2 = 3$.

Формування групи стробувальних імпульсів показано на рис. 9.5.

На діаграмі показано f_b – велику частоту задавального генератора (ЗГ), $1/T_0$ – тактову частоту. Очевидно, що для збільшення інтервалу аналізу потрібно збільшувати кількість стробувальних імпульсів так, щоб t_a вкладався в межі тактового інтервалу T_0 ; при цьому кількість імпульсів непарна. Однак під час визначення ϵ слід враховувати можливі крайові спотворення, які вносяться каналом зв'язку.

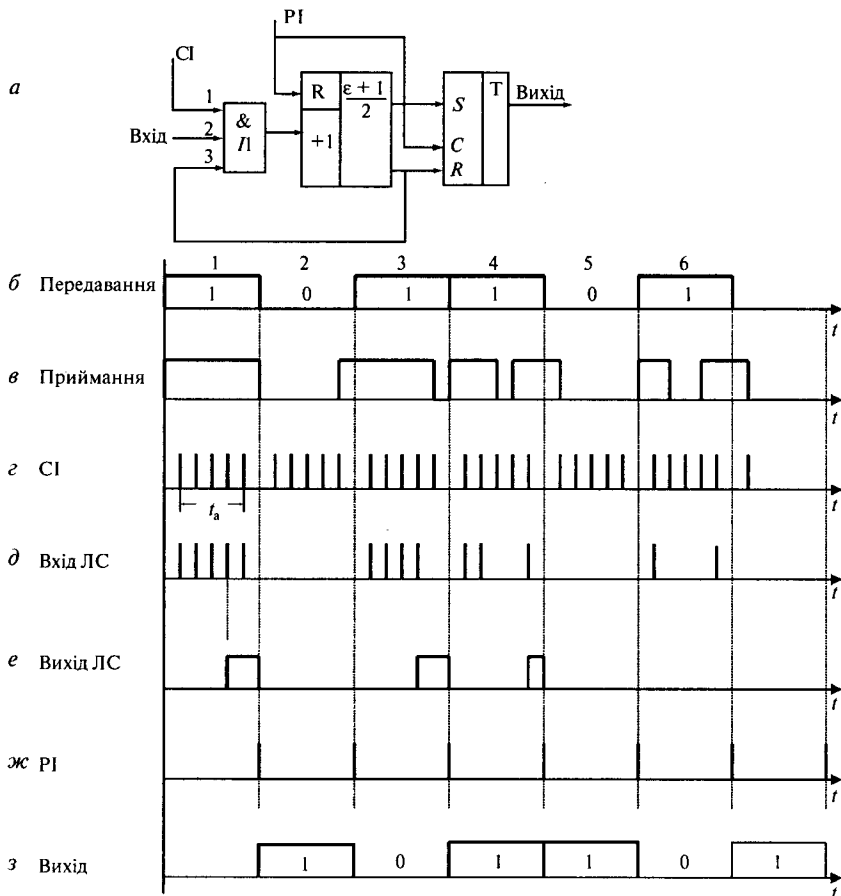


Рис. 9.4. Структурна схема РП цифровим інтегральним методом (*a*) і часова діаграма його роботи (*б, в, г, д, е, ж, з*)

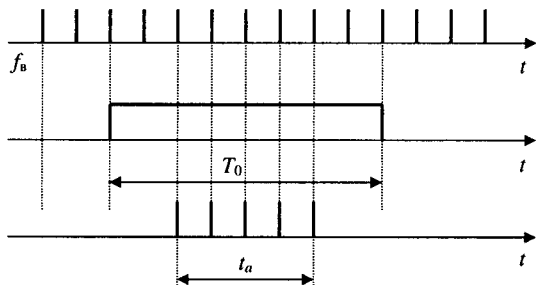


Рис. 9.5. Часова діаграма формування групи стробувальних імпульсів на інтервалі аналізу t_a

Але під дією завад, особливо, коли спотворення дроблення в середині T_0 і тривалість його більші, ніж Δt , тоді реєстрація виконується майже в кінці інтервалу T_0 .

Отже, РП знаходиться в режимі очікування. Цей режим підтримується послідовністю стробувальних імпульсів та сигналом заборони лічення стробувальних імпульсів. Реєстрація завершується помилкою реєструвального імпульсу і водночас ЛС установлюється у нульове положення.

10. Методи синхронізації та фазування в системі передавання дискретної інформації

10.1. Потреба в синхронізації та фазуванні

Дві послідовності подій синхронні, якщо відповідні події в них виникають одночасно.

Синхронізація – це процес установлення та підтримання синхронного стану.

У системах зв'язку одна послідовність подій утворюється в передавачі, а друга – у приймачі.

Дискретні повідомлення передаються сигналами послідовністю елементів певної довжини. Це потребує синхронізації відліків часу на передавачі та приймачі. Дуже часто відліки часу називають «тактами», а синхронізацію відліків часу – «тактовою синхронізацією». На приймальному кінці передане повідомлення відновлюється обробленням кожного елемента повідомлення пристроєм реєстрації елементів. Синхронна послідовність відліків часу, з якої формуються імпульси опитування (імпульси реєстрації), має знаходитись у певному фазовому відношенні до прийманих одиночних елементів дискретного повідомлення.

Відстань між імпульсами опитування залежить від способу обробки дискретних повідомлень.

Відомі два способи оброблення повідомлень.

1. Оброблення повідомлень в цілому. Це, якщо кодова комбінація подає повідомлення як єдиний складний сигнал і до кодової комбінації входить n одиночних сигналів тривалістю τ_0 , тоді опитувальні імпульси мають йти слідом з інтервалом $n\tau_0$ і збігатися з початком (кінцем) кожної прийманої кодової комбінації.

2. Оброблення повідомлення окремо кожного складового елемента, або поелементне оброблення кодової комбінації. Ідея цього способу оброблення – потрібно мати синхросигнали. Тому розрізняють два підвиди синхронізму: синхронізм елементами і синхронізм циклами.

Синхронізм елементами, або синхронізація – це процес установлення відповідності між значущими моментами одиночних елементів під час передавання та приймання.

Фазування циклів (або синхронізація циклами) – це процес установлення відповідності між фіксованими місцеположеннями в кодовій комбінації значущих моментів на передавальному та приймальному кінцях.

Потребу у синхронізації та фазуванні можна показати на прикладі синхронної системи передавання дискретних повідомлень (рис. 10.1).

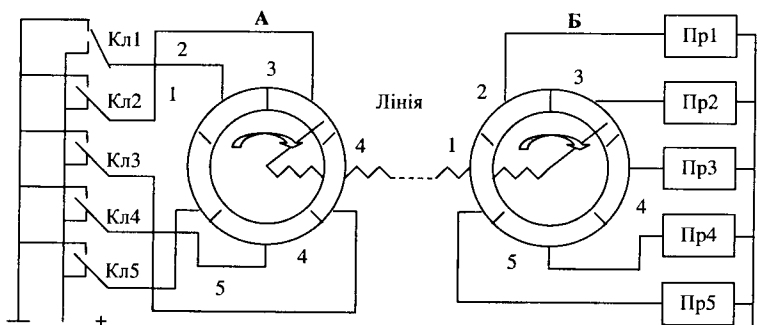


Рис. 10.1. Схема принципу синхронного зв'язку

Кодова комбінація, яка складається з п'яти елементів, довжина кожного з яких дорівнює τ_0 , утворюється неперервно працюючим передавальним розподільником станції за один оберт. Якщо щітки розподільників станцій А і Б мають однакову кутову швидкість (розподільники синхронізовані) і однакову фазу (виконано циклове фазування), тоді до лінії в кожний момент часу підключено однойменні контакти розподільників.

Інформація від клавіші Кл1 надходить до приймача Пр1, від Кл2 – до Пр2 і т. д. Порушення фазування циклами призводить до неправильного приймання всієї інформації, бо до лінії будуть підключені неоднойменні контакти.

10.2. Поелементна синхронізація

У попередньому розділі ми розглянули методи реєстрації дискретних сигналів, які передбачають наявність спеціальних реєструвальних сигналів, за допомогою яких приймаються рішення в двійкових дискретних каналах позитивної або негативної послідовності, що відповідають символам даних «1» або «0».

Нагадаємо, що реєструвальні сигнали повинні мати строго визначене фазове співвідношення з одиночним значущим інтервалом T_0 СПДІ, а саме, у процесі реєстрації методом стробування вони включаються в дію (момент дії) всередині інтервалу T_0 .

Виникає завдання – яким чином установити і підтримати ці співвідношення.

Процес установлення та підтримки потрібних фазових співвідношень між значущими моментами передавальних і приймальних одиничних елементів СПДІ називають *поелементною синхронізацією*.

Синхронний стан приймальної частини відносно передавальної в СПДІ можна підтримувати, застосовуючи для приймання автономне джерело частоти. Така система буде найідеальнішою щодо завадостійкості синхронізації. Для створення автономного джерела частоти використовують будь-який генератор. Генерована кожним із цих генераторів частота завжди буде мати відхилення від свого номінального значення, спричинена впливом різних чинників: навколишньою температурою, тиском, зміною напруги джерел живлення тощо, а також неточністю установлення номінальної частоти. Відхилення генерованої частоти від номінального значення характеризується значенням нестабільності. Під абсолютною нестабільністю розуміють різницю Δf (без урахування знака) між номінальним f_0 і дійсним або результируючими значеннями частоти f_0' , тобто

$$\Delta f = f_0 - f_0'.$$

Відносною нестабільністю генератора називають відношення абсолютної нестабільності до номінального значення частоти:

$$K_r = \frac{\Delta f}{f_0}.$$

Розглянемо процес поелементної синхронізації, використовуючи часову діаграму порушення співвідношень між значущими моментами передавальних і приймальних двійкових сигналів СПДІ (рис. 10.2).

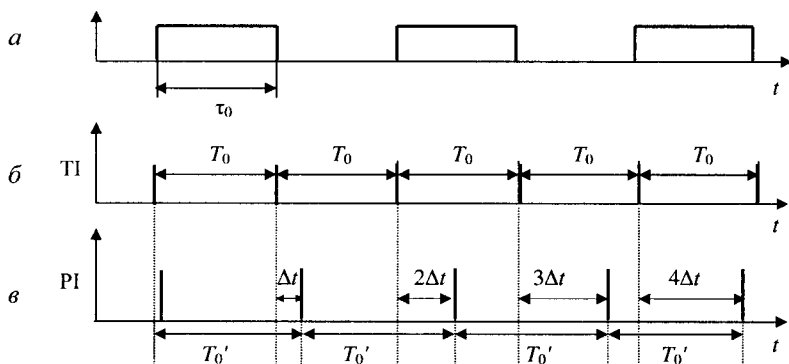


Рис. 10.2. Часова діаграма розбіжності між тактовими частотами послідовностей передавача та приймача: *a* – передавальна послідовність; *б* – її тактові інтервали; *в* – тактова послідовність приймача

Згідно із загальною схемою структури пристрою перетворення сигналів передавальна та приймальна частини містять ЗГ. Задавальний генератор на передавальному кінці використовують для формування тактових імпульсів (ТІ), які і визначають швидкість модуляції. Задавальний генератор на прийальному кінці використовують для формування імпульсів реєстрації та імпульсів стробування. Отже, ЗГ ППС і ЗГ ППСпр і утворюють дві послідовності подій, які в СПДІ мають бути синхронними.

У ППС застосовують високостабільні задавальні генератори з коефіцієнтом нестабільності частоти $K_r = \frac{\Delta f}{f_0}$ ($10^{-4} \dots 10^{-8}$).

Для відомого значення нестабільності генераторів визначимо час допустимої перерви, за якої буде зберігатись синфазний стан між приймальною та передавальною частинами СПДІ із заданою точністю. Зазвичай для генераторів указується частотна відносна нестабільність. Але на практиці зручніше користуватися не частотою, а оберненою їй величиною – часовою характеристикою, тобто періодом повторення, в одиницях якого виражаються тривалість елементарних посилок і період повторення реєструвальних імпульсів.

Якщо відносна частотна нестабільність $K_r = \frac{\Delta f}{f_0}$, то часова нестабільність

$$K'_r = \frac{\Delta T}{T_0}, \quad (10.1)$$

де $\Delta T = T_0' - T_0$, T_0 і T_0' – номінальний і результуючий періоди коливань або повторення імпульсів.

У свою чергу,

$$f_0' = \frac{1}{T_0'}, \quad f_0 = \frac{1}{T_0}.$$

Підставивши в K_r значення Δf і f_0 , отримаємо

$$K_r = \frac{\Delta f}{f_0} = \frac{\frac{1}{T_0} - \frac{1}{T_0'}}{\frac{1}{T_0}} = \frac{T_0' - T_0}{T_0'} = \frac{\Delta T}{T_0 + \Delta T},$$

або
$$K'_r = \frac{\Delta T}{T_0} = \frac{K_r(T_0 + \Delta T)}{T_0} = K_r \left(1 + \frac{\Delta T}{T_0}\right) = K_r(1 + K'_r) = \frac{K_r}{1 - K_r}.$$

Ураховуючи те, що K_r змінюється в межах $10^{-4} \dots 10^{-8}$, тобто $K_r \ll 1$, покладемо

$$K'_r = K_r. \quad (10.2)$$

Тоді коливання приймальної частини СПДІ на одну елементарну послілку відносно передавальної частини зміститься через

$$L = \frac{T_0}{\Delta T}, \quad (10.3)$$

де $T_0 = \tau_0$ – тривалість елементарної послілки.

Залежність (10.3) отримано за умови, що частота генератора передавальної частини системи не змінюється (не відхиляється від номінальної f_0), тобто $f_0 = \text{const}$.

Якщо врахувати нестабільність другого генератора передавальної частини системи, частота якого у найнесприятливіх випадках відхиляється від номінального значення в бік, протилежний приймальній частині, тоді обидва генератори розійдуться фазою на одну елементарну послілку за

$$L' = \frac{T_0}{2\Delta T} \text{ (посілок),}$$

або, ураховуючи рівняння (10.1) і (10.2), отримаємо

$$L' = \frac{1}{2K_r}.$$

Насправді потрібно визначити час зміщення синхронізувального імпульсу не на цілий елементарний імпульс, а на деяку його частину ε ($1 > \varepsilon > 0$), за якої система ще могла б правильно реєструвати приймані послілки. Відповідно до цього кількість послілок, через які змістяться розподільники фазою на задану частину ε елементарної послілки, буде становити

$$L'' = \frac{\varepsilon}{2K_r}. \quad (10.4)$$

Із формули (10.4) випливає, що обсяг повідомлення, за якого відбувається розсинхронізація розподільників у допустимих границях, залежить тільки від коефіцієнта нестабільності і не залежить від швидкості передавання даних.

Тепер можна за допустимого обсягу повідомлення знайти час t , за який буде передано це повідомлення, або, що те саме, допустиму перерву зв'язку, після якої треба знову установити фазові співвідношення:

$$T = L''t.$$

Ураховуючи, що

$$B = \frac{1}{\tau_0},$$

де B – швидкість передавання, бод, отримаємо залежність допустимого часу перерви зв'язку від швидкості передавання даних:

$$t_{\text{пер}} = \frac{\varepsilon}{2K_r B}. \quad (10.5)$$

Для різних значень ε , K_r і B виконаємо розрахунки за формулами (10.4) і (10.5), результати зведемо в табл. 10.1.

Таблиця 10.1

| Коефіцієнт нестабільності K_r | Допустимий обсяг повідомлення L за різних зсувів фаз | | | Час, через який зсув фаз $\varepsilon = 0,2$ за швидкості передавання B , кбіт/с | | |
|---------------------------------|--|---------------------|---------------------|--|------------|----------|
| | $\varepsilon = 0,4$ | $\varepsilon = 0,2$ | $\varepsilon = 0,3$ | $B = 0,6$ | $B = 19,2$ | $B = 48$ |
| 10^{-3} | | | | | | |
| 10^{-4} | | | | | | |
| 10^{-5} | | | | | | |
| 10^{-6} | | | | | | |
| 10^{-7} | | | | | | |
| 10^{-8} | | | | | | |
| 10^{-9} | | | | | | |

Із таблиці видно, який обсяг повідомлення можна передавати без виникнення помилок через нестабільність генератора, але збільшення переданого обсягу призведе до виникнення помилок. Для того, щоб уникнути помилок через зсув фаз, треба після передавання допустимого обсягу даних установити генератори в початковий стан, а потім розпочати передавання нової порції повідомлень, або впроваджувати додаткові заходи для підтримки синхронізму.

У формулі (10.5) у знаменнику поставлено двійку. Це означає, що нестабільність генераторів передавальної та приймальної частин СПДІ однакова, але коли нестабільності різні, тоді в знаменнику потрібно замість $2K_r$ записати суму нестабільностей.

10.3. Способи підтримки синхронізму

Відхилення робочих частот від номінальних – причина порушення синхронізму. Інформацію про ці порушення називають синхронізувальною інформацією, яка потрібна для роботи пристроїв синхронізації. Таку інформацію можна отримати трьома способами, які і визначають способи підтримки синхронізму.

1. Автономний спосіб отримання синхронізувальної інформації на приймальній частині використовують у тих випадках, коли швидкість передавання інформації в каналі зв'язку підтримується з високою точністю, а передавальна частина СПДІ має дуже високу стабільність параметрів. Такими еталонами часу можуть бути місцеві генератори з високою стабільністю.

На перший погляд здається, що це найпростіший спосіб отримання синхронізувальної інформації, який і забезпечує найвищу завадостійкість системи синхронізації. Але його реалізація пов'язана з деякими труднощами. Основні з них такі:

- проблема першопочаткового входу в зв'язок: значно ускладнюється сам пристрій синхронізації;
- залежність точності синхронізації від постійності поширення електромагнітної енергії каналом зв'язку та постійності фазових співвідношень між переданими і прийнятими сигналами;
- місцеві генератори – еталони часу – мають кінцеве значення стабільності.

2. Спосіб синхронізації шляхом передавання синхронізувальної інформації окремим каналом дозволяє побудувати досить ефективну систему синхронізації. Але він має суттєві недоліки. Знижується пропускна здатність системи зв'язку, оскільки частина смуги каналу займається під передавання синхронізувальної інформації. Організація додаткового каналу призводить до значного ускладнення каналоутворювальної апаратури, тому цей спосіб знайшов обмежене застосування. Якщо інформація з одного пункту в другий передається одночасно великою кількістю каналів, то тільки тоді доцільно синхронізувальну інформацію передавати окремим каналом.

3. Спосіб передавання синхронізувальної інформації каналом передавання повідомлень дозволяє передавати синхронізувальну інформацію. Такий спосіб називають синхронізацією робочими посилками. Цей спосіб порівняно зі способом передавання сигналів окремим каналом має ряд суттєвих переваг:

- найбільш висока ефективність використання каналу зв'язку;

– незалежність установлення фази синхронізувального сигналу відносно інформаційного від групової швидкості поширення сигналів у каналі зв'язку.

Однак ці переваги легко долаються незначним ускладненням ПС. Тому тепер майже у всіх системах передавання синхронізація здійснюється робочими сигналами.

Незалежно від того, передається чи не передається синхронізувальна інформація окремим каналом або сумісно з каналом передавання інформації, основою синхронізатора є генератор, належно синхронізований. Вимоги до синхронізатора в обох випадках однакові: забезпечити для приймання фазові співвідношення між синхронізувальними сигналами і сигналами передавання інформації в заданих границях.

Пристрої синхронізації, виконані на основі генератора з регульованою фазою, називають пристроями синхронізації з автоматичним підстроюванням фази місцевого генератора, або ЗГ.

Отже, маємо три способи підтримки синхронізму.

Перший спосіб підтримки синхронізму забезпечується завдяки високій стабільності генераторів, але сучасні високостабільні генератори не можуть забезпечити достатнього рівня синхронізації.

Другий спосіб досить ефективний, але потребує створення окремого каналу тільки для підтримки синхронізму. Спеціальні сигнали для підтримки синхронізму одержали назву «пілот-сигнали».

Третій спосіб найбільш ефективний тим, що за допомогою додаткових елементів у ППСпер виконується спостереження за зміною положення прийманих значущих моментів та приймаються рішення щодо зміни частоти в бік значення оптимальної (номінальної) частоти.

Розглянемо детально принцип підтримки синхронізму на підставі виявлення відхилень від номінального значення параметрів.

Одержану інформацію про розбіжність параметрів сигналів процесу синхронної підтримки роботи передавальних та приймальних пристроїв СПДІ використовують для коригування тактових сигналів підстроювання параметрів пристроїв, які задають тактовість роботи передавача та приймача СПДІ.

Головним елементом формування тактової послідовності імпульсів є ЗГ.

Пристрої синхронізації можна реалізувати різними способами: без посередника і з посередником.

Безпосередній спосіб формування тактових імпульсів ПС називають спосіб без зворотного зв'язку (розімкнених ПС). У розімкнених ПС тактові імпульси формуються на підставі виявлення інформації з інформаційних (робочих) імпульсів за допомогою високовибірних резонансних пристроїв, тому їх називають резонансними.

Такі пристрої використовують у регенераторах цифрових систем передавання. Недолік – малий час підтримки синхронізму.

У замкнених ПС тактові імпульси формуються ЗГ ППСпр, а фазові – співвідношенням їх зі значущими моментами приймальних робочих імпульсів, що аналізуються і коригуються. Це і є принцип фазового підстроювання частоти генератора ППСпр. У СПДІ знайшли широке впровадження замкнені ПС.

У замкнених ПС виявлені сигнали розбіжності використовують для діяння на зміну частоти ЗГ.

Діяння на частоту ЗГ можна реалізувати двома способами:

- 1) без посередника;
- 2) з посередником.

За безпосередньої дії на частоту ЗГ приблизно на порядок знижується стабільність ЗГ, що є суттєвим недоліком такого ПС. За діяння з посередником на частоту генератора ПС цього недоліку немає, тим більше, що можна мати один високостабільний генератор з подільниками частоти для декількох комплектів апаратури, або забезпечити від одного генератора роботу з різними швидкостями модуляції.

10.4. Основні вимоги до пристроїв синхронізації

До цього ми розглядали деякі методи підтримки синхронізму та принципи побудови пристроїв, які реалізують ці методи. Тепер розглянемо основні вимоги до ПС з дискретним керуванням і постійним корекційним ефектом. Знання цих вимог дозволять грамотно вибирати і конструювати пристрої системи синхронізації.

Система синхронізації має забезпечити такі вимоги.

1. Задану точність синхронізації з урахуванням діяння всіх чинників, які впливають як на канал зв'язку, так і на саму апаратуру передавання інформації. Під точністю (похибкою) синхронізації будемо розуміти відносну величину відхилення (без урахування знаків) моментів реєстрації посилок від моментів, які відповідають точній синхронізації. Похибка синхронізації δ_c складається із статичної похибки $\delta_{ст}$, яка привноситься самим ПС приймальної частини, і динамічної похибки $\delta_{дин}$, яка привноситься ПС у разі впливу сигналів синхронізації (тактових сигналів), спотворених завадами в каналі зв'язку:

$$\delta_c = \delta_{ст} + \delta_{дин}.$$

Статична похибка, по суті, – це апаратна похибка, яка визначається фазовим зсувом ресструвальних імпульсів від дійсних моментів під час синхронізації неспотвореними сигналами синхронізації. Інакше кажучи,

статична похибка визначається розбіжностями частот ЗГ передавача і приймача і не враховує спотворення вхідного сигналу.

Динамічна похибка $\delta_{\text{дин}}$ характеризує ступінь збільшення відхилення ресетрувальних імпульсів у разі переведення з роботи «на себе» на роботу через канал зв'язку, тобто фазові зсуви тактових імпульсів, викликані спотвореннями елементів інформаційного сигналу (зсування значущих моментів). Як було показано, спотворення тривалості приймальних елементів можуть бути викликані появою помилкових сигналів керування на виході фазового дискримінатора, а отже, і на виході інерційного елемента. Ці сигнали можуть здійснювати помилкове коригування ПС у бік зсуву фаз.

2. Заданий час входження в синхронізм як під час початкового входження у зв'язок, так і після перерви в роботі. Цей параметр не менш важливий, ніж точність синхронізації, оскільки він характеризує швидкодію системи синхронізації. Заданий час входження в синхронізм інколи називають часом синхронізації, який залежить від початкового фазового збігу та кроку коригування. Для ПС з постійним коригувальним ефектом у результаті кожного підстроювання фазовий зсув зменшується на величину Δt , тому для входження в режим синхронізму потрібно буде декілька кроків залежно від значення початкового збігу фаз.

3. Зберігання синфазності під час короткочасних перерв зв'язку. Ця вимога порівняно легко виконується в системах синхронізації з автоматичним підстроюванням фази місцевого ЗГ, який має задану стабільність. Зберігання синфазності – це час підтримки синфазного стану, або час, протягом якого фаза синхросигналів не вийде за допустимі границі зсуву фаз у разі припинення роботи ПС з підстроювання фази. Для уникнення випадків втрати синхронізму під час приймання інформаційних сигналів одного й того ж знака в апаратурі СПДІ установлюють спеціальні пристрої (скремблери). Останні формують кодові послідовності на передавальному кінці такими, щоб унеможливити в них подібні поєднання послідовності одиничних елементів. Обернене перетворення на приймальній кінці виконують дескремблери.

10.5. Синхронізація безпосереднім діямням на частоту задавального генератора

Принцип роботи системи синхронізації у разі безпосереднього діямня на частоту ЗГ розглянемо за структурною схемою ПС за посилками з автоматичним підстроюванням фази місцевого ЗГ (рис. 10.3). Основу пристрою складають фазовий дискримінатор і керувальний елемент.

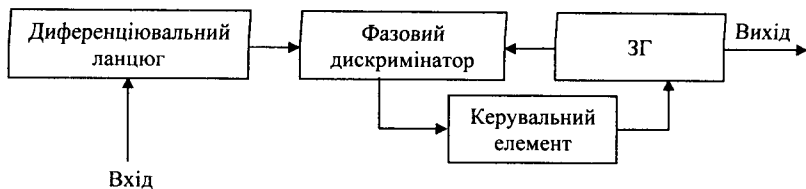


Рис. 10.3. Структурна схема ПС за безпосереднього керування частотою ЗГ

Синхронізувальна інформація подається на фазовий дискримінатор у вигляді прямокутних імпульсів постійного струму. Ці електричні посилки використовуються для синхронізації. Вони диференціюються ланцюгом, а потім короткі імпульси надходять до фазового дискримінатора.

Фазовим дискримінатором називають пристрій, який через порівняння із сигналами, що надходять від ЗГ, виділяє сигнал неузгодженості, який вказує значення і знак фазової неузгодженості між сигналами синхронізації, що надходять із каналу зв'язку, і коливаннями ЗГ.

Керувальним елементом будемо назвати пристрій, який під дією сигналу неузгодженості, отриманого із дискримінатора, змінює частоту ЗГ. Частоти генератора змінюються внаслідок зміни параметрів L і C резонансного контура генератора під дією визначеного сигналу збігу фаз.

Згадаймо, що інформаційні (робочі, синхронізувальні) посилки формуються передавачем СПДІ з урахуванням дозволеної смуги частот каналу зв'язку. Тривалість імпульсних посилок називають тактовою, а отже, відповідно і частоту – тактовою. Тому і вихідний елемент ЗГ буде не що інше як формувач тактових сигналів, узгоджених за параметрами з тактовими сигналами, які надходять з лінії зв'язку.

Розглянемо роботу системи синхронізації. Припустімо, що в момент входження в зв'язок коливання місцевого генератора виявились зсунутими відносно диференційованого імпульсу на кут φ в бік випередження (рис. 10.4). Тоді на виході фазового дискримінатора з'являється керувальний сигнал, значення і полярність якого будуть залежати від значення і знака неузгодження (розбіжності) φ двох порівнюваних коливань. Цей сигнал – сигнал розбіжності (помилки) діє (або безпосередньо, або після виділення із нього постійної складової) на керувальний елемент, який, у свою чергу, зменшує частоту ЗГ і відповідно зменшує фазовий зсув.

Якщо коливання місцевого генератора відстають від коливань генератора передавальної частини СПДІ на виході фазового дискримінатора (рис. 10.4, б), тоді виробляється сигнал помилки від'ємного знака, який своєю дією на керувальний елемент збільшує частоту ЗГ.

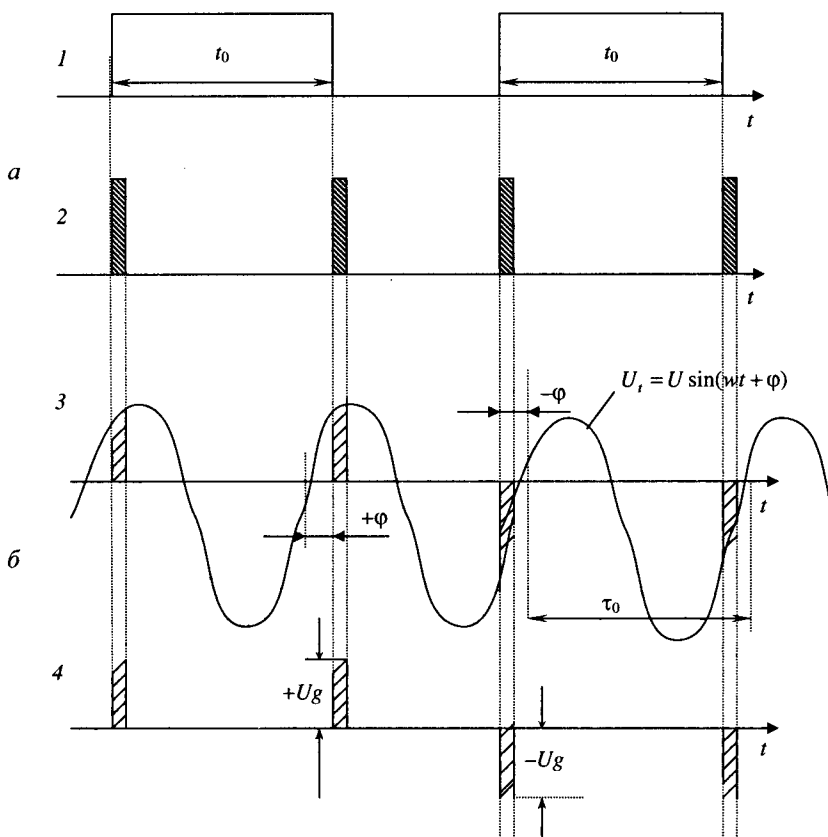


Рис. 10.4. До принципу роботи ПС для випадків випередження (а) і відставання (б):
 1 – приймальні посилки; 2 – сигнали на виході диференційованого ланцюга;
 3 – синусоїдні коливання на виході ЗГ; 4 – напруга на виході фазового дискримінатора

Загальні недоліки усіх пристроїв синхронізації з безпосереднім діянням на генератор (порівняно з пристроями з посереднім діянням) такі: зниження стабільності генератора приблизно в 10 разів унаслідок діяння на один із його параметрів; неможливість використання одного генератора для роботи декількох ПС; ПС з безпосереднім діянням на частоту ЗГ застосовують тоді, коли потрібне спрощення схеми або коли пристрої з посереднім діянням на генератор важко реалізувати, наприклад, у разі високих швидкостей модуляції. Але цей недолік можна усунути, використовуючи визначення відповідних параметрів складових частин ПС з посереднім діянням на частоту ЗГ. Перейдемо до розгляду ПС такого типу.

10.6. Синхронізація з посереднім діянням на частоту задавального генератора

У системах синхронізації з посереднім діянням на частоту ЗГ (рис. 10.5) керувальний елемент (керувальний пристрій) підключається до проміжного перетворювача, який змінює фази вихідних імпульсів. Завдяки цьому установлюється синхронізм між сигналами, які надходять із каналу зв'язку, і сигналами, які надходять з проміжного перетворювача.

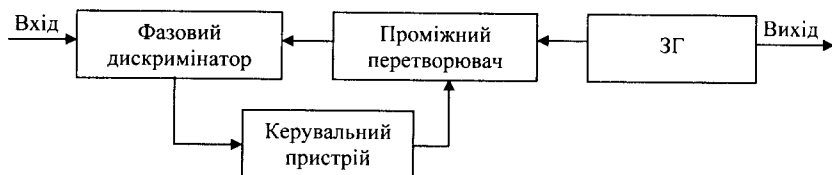


Рис. 10.5. Структурна схема ПК з посереднім діянням на частоту генератора

У системах з посереднім діянням проміжний перетворювач створюється здебільшого з використанням подільника частоти, або додатковим генератором.

У першому випадку фазування здійснюється діянням на подільник частоти, у результаті чого або змінюється коефіцієнт ділення подільника, або додаються чи віднімаються імпульси, які надходять на вхід подільника із ЗГ.

Для синхронізації цим способом найбільше застосування знайшли керувальні пристрої, які виконують додавання або віднімання імпульсів, що надходять на вхід подільника частоти. Пристрої, які виконують синхронізацію таким способом, називають пристроями з дискретним або цифровим керуванням.

Система синхронізації з посереднім керуванням частотою додаткового генератора за принципом здійснення регулювання така сама, як і звичайна система з безпосереднім діянням на частоту генератора. Тому вона дозволяє реалізувати принцип повільного регулювання, що в кінцевому підсумку усуває спотворення вихідної інформації через коливання параметрів і в той же час запобігає дестабілізації ЗГ. Вона подібна до систем з дискретним керуванням, перерва зв'язку визначається стабільністю ЗГ. Від дискретної системи вона відрізняється виключно простотою побудови ПК.

Розглянемо принцип побудови ПК, у якому реалізується метод підтримки синхронізму за дискретного керування з посереднім діянням на частоту ЗГ, схему якого зображено на рис. 10.6.

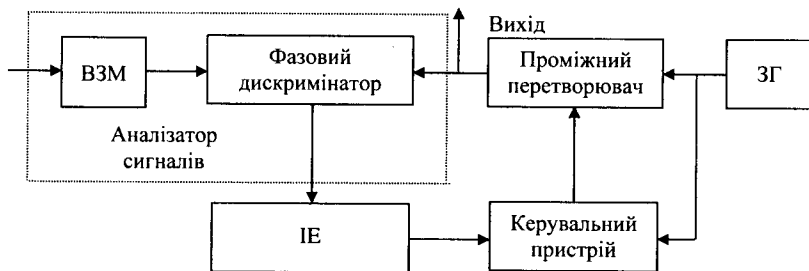


Рис. 10.6. Пристрій синхронізації з посереднім дискретним керуванням частотою ЗГ

Пристрій синхронізації складається з аналізатора сигналів, який виконує функції виділення значущих моментів (ВЗМ) приймального сигналу (функцію ВЗМ звично виконує диференціальний ланцюг) і фазовий дискримінаціатор, який вимірює фазові неузгодження значущих моментів приймального сигналу та сигналу, сформованого місцевим ЗГ разом з проміжним перетворювачем.

Вивчаючи процес зсуву фаз ЗГ ППСпер і ЗГ ППСпр під час передавання та приймання послідовності сигналів, визначають допустиме значення зсуву фаз. Цим самим установлюється потреба зовнішнього втручання для зміни частоти ЗГ ППСпр.

Сигнали зсуву фаз на кожному такті Δt можна використовувати для підстроювання частоти ЗГ ППСпр, але раніше був розглянутий процес передавання послідовності тактових інформаційних сигналів і встановлено, що за визначеної довжини в n імпульсів (або $t_{\text{доп}}$) зсув фаз не призводить до виникнення помилки.

Для формування сигналу змінювання частоти місцевого ЗГ потрібно часткові відхилення Δt , які виникають на кожному тактовому інформаційному (робочому) сигналі, складати доти, доки не буде витримано рівність $\Delta t k = t_{\text{доп}}$. Процедуру формування цього сигналу виконують за допомогою лічильника (реверсивного); кінцевий сигнал лічильника за визначеного k і буде сигналом для змінювання частоти місцевого ЗГ. Цей лічильник зі схемою формування сигналу називають інерційним елементом. Функціонально ІЕ – це пристрій, який затримує на $t_{\text{доп}}$ початковий зсув фаз.

Роботу вимірника збігу фаз – фазового дискримінаціатора – розглянемо на схемі та часових діаграмах (рис. 10.7).

Схема фазового дискримінаціатора складається із двох комірок збігу П1 і П2, на входи яких надходять імпульси від ВЗМ прийнятих сигналів, а також дві послідовності тактових імпульсів Т1, Т2 у протифазі одна одній. Частоти Т1, Т2 у два рази більші від тактових інформаційних Т0.

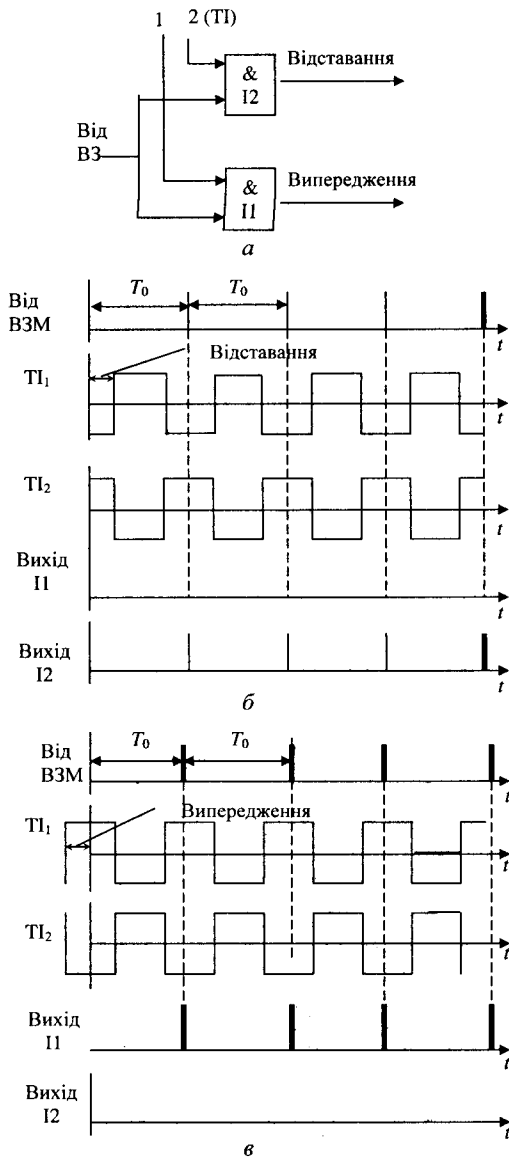


Рис. 10.7. Схема фазового дискримінатора (а), часових діаграм відповідно появи сигналів випередження (б) та відставання частоти місцевого ЗГ(в)

Роботу фазового дискримінатора пояснюють часові діаграми. На рис. 10.7, б зображено випадок, коли частота місцевого ЗГ відстає від частоти ЗГ ППСпер і на виході I2 з'являється послідовність імпульсів, які інформують про відставання. Якщо буде синфазне явище, тоді ні на виході I1, ні на виході I2 не буде формуватись послідовність імпульсів. Таким чином, на виході фазового дискримінатора з'являється послідовність імпульсів, які несуть інформацію про відставання і не інформують про часовий момент цих відставань.

Для випадку, коли місцевий ЗГ випереджає ЗГ передавальної частини СПДІ, часову діаграму показано на рис. 10.7, в. Аналогічно, як і у разі відставання частоти, так і випередження, послідовність імпульсів виникає на виході схеми I1.

Отже, між приймальними сигналами і місцевими тактовими сигналами імпульсів на виходах фазового дискримінатора немає. У разі виникнення розфазування короткі імпульси з'являються на виході схеми I1 або схеми I2 залежно від знака збігу фази.

Під час розгляду схеми фазового дискримінатора здавалося, що одиничні елементи вхідної послідовності не спотворені і вимірювання збігу фаз між границями (значащими моментами) і синхронізувальними імпульсами досить точно характеризує дійсний збіг фаз. У реальних умовах спостерігається зміщення значущих моментів від їх ідеального положення. Ці зміщення спричинено впливом завад та інших спотворень (стрибками рівня, фазовими спотвореннями і т. ін.). Зміщення значущих моментів навіть у режимі синфазності потребує коригування фази, оскільки частина зміщення значущих моментів буде збігатися з режимом відставання місцевого генератора, а частина – з режимом випередження. Ураховуючи, що зміщення значущих моментів у бік відставання і випередження рівноймовірно, вплив цих спотворень можна зменшити, увімкнувши між дискримінатором і керувальним пристроєм ІЕ.

Отже, ІЕ забезпечує можливість враховувати як регулярні відставання або випередження частот ЗГ, так і відставання або випередження, спричинені впливом завад, особливо які привносять такі спотворення, як перевантаження. Очевидно, що ІЕ формує сигнали або відставання, або випередження в моменти часу, які визначаються часовою характеристикою $t_{\text{доп}}$.

У системах синхронізації з дискретним керуванням змінювати частоту найпростіше за допомогою керувального пристрою посереднім діям на проміжний перетворювач, створений на базі подільника частоти або додаткового генератора.

Якщо фазування виконує керувальний пристрій, діючи на подільник частоти, то результатом може бути або зміна коефіцієнта ділення подільника, або додавання чи віднімання імпульсів, які надходять на вхід подільника із ЗГ.

Для синхронізації цим способом найбільш застосовний керувальний пристрій, який додає або віднімає імпульси, які надходять на вхід подільника частоти. Пристрої, які виконують синхронізацію таким способом, називають пристроями з дискретним, або цифровим керуванням.

Принцип роботи керувального пристрою такого типу розглянемо за схемою, зображеною на рис. 10.8. Схему побудовано на логічних елементах: АБО1 – схема збирання, НІ-І2 – схема збігу з одним інверсним входом та І3 – схема збігу.

На перші входи схем НІ-І2 і І3 надходять сигнали випередження та відставання відповідно. На другі входи подаються імпульси високої частоти ЗГ: ТІ3 – тактова послідовність роботи ЗГ, та ТІ2 – тактова послідовність ЗГ, зсунута на півперіод.

У процесі роботи керувального пристрою можливі три випадки.

1. Сигналів відставання та випередження немає. Це означає, що синхронізм не порушено. Зсунута на півперіод послідовність ТІ2 без перешкод проходить через НІ-І2 і АБО і надходить на вхід проміжного перетворювача (рис. 10.8, *а*).

2. На вході І3 виникає імпульс відставання, який за часом збігається з одним імпульсом послідовності ТІ3. Цією дією формується імпульс, що використовується як імпульс додавання до послідовності ТІ2, тим більше що періоди ТІ3 і ТІ2 зсунуті на півперіод (рис. 10.8, *б*).

3. На вході НІ-І2 виникає імпульс випередження, який забороняє проходження одного імпульсу з послідовності ТІ2. Це означає прорідження цієї послідовності на один імпульс (рис. 10.8, *в*). Ця дія в послідовності ТІ2 зумовлює появу ділянки з вилученням імпульсів. Таким чином, на вході проміжного перетворювача діє послідовність з прорідженими ділянками.

Очевидно, що робота керувального пристрою зводиться до змінювання частоти ЗГ унаслідок добавлення імпульсу в послідовність у разі відставання ЗГ приймальної частини відносно ЗГ передавальної частини СПДІ, або віднімання імпульсу з цієї послідовності у разі випередження частоти ЗГ пр. відносно частоти ЗГ пер.

У пристроях синхронізації з посереднім діянням на ЗГ фаза підстроюється проміжним перетворювачем, через який проходить місцева тактова послідовність (послідовність ЗГ приймальної частини СПДІ). Для функціонування проміжного перетворювача найчастіше використовують подільник частоти послідовності імпульсів, які надходять із керувального пристрою послідовністю ЗГ пр.

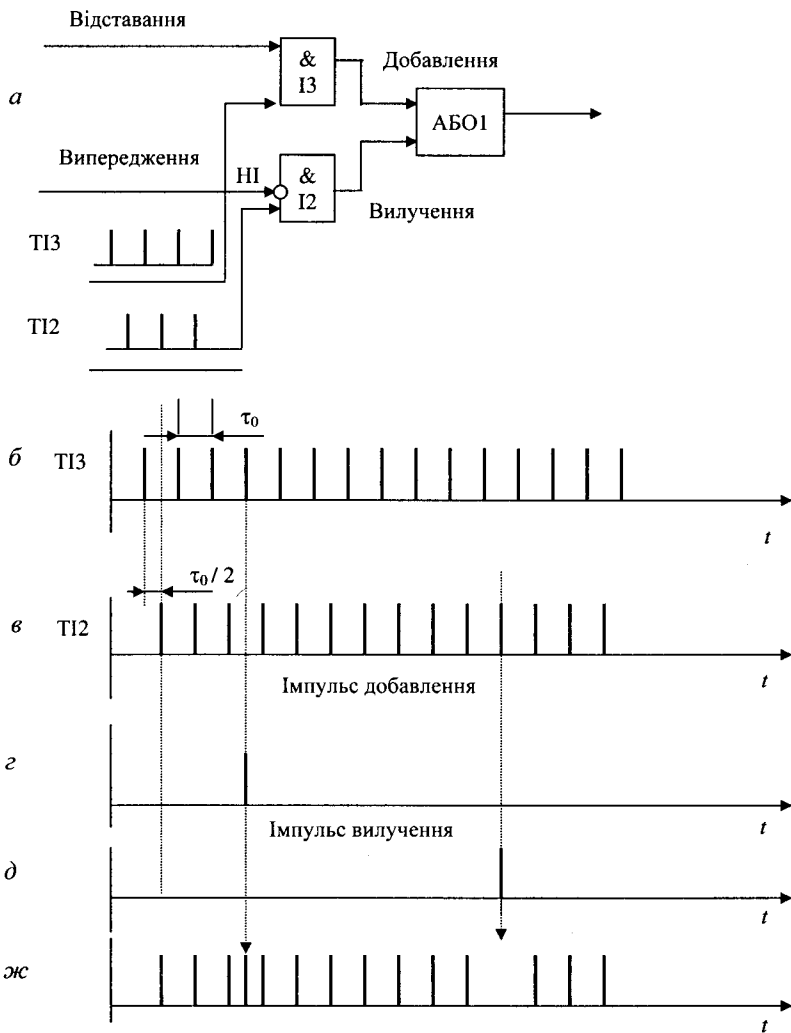


Рис. 10.8. Принцип роботи керувального пристрою послідовністю ЗГ:
 а – структурна схема; б, в, г, ж – часова діаграма пристрою

Принцип змінювання фази в процесі поділу частоти можна пояснити за допомогою рис. 10.8. Генератор виробляє послідовність імпульсів високої частоти в m разів більшу від тактової частоти: $f_{\text{Вч}} = mf_{\text{т}} = \frac{1}{T_0}$.

Величину m називають *коефіцієнтом ділення подільника*. Процес ділення ілюструється часовою діаграмою, показаною на рис. 10.8, *а*. На рисунку зображено початкову послідовність імпульсів частотою $f_{\text{Вч}}$ і отриману на виході тактову послідовність імпульсів частотою $f_{\text{т}}$. Коефіцієнт ділення $m = 8$. При цьому на вихід видається кожний восьмий імпульс вхідної послідовності імпульсів.

Якщо на вхід подільника подати додатковий імпульс, то фаза вихідних імпульсів (рис. 10.8, *б*) зсунеться в бік випередження на значення, яке дорівнює періоду послідовності імпульсів генератора: $\Delta t = 1/f_{\text{Вч}}$. Додаткові імпульси не повинні збігатися у часі з імпульсами високочастотного генератора. Ця вимога виконується за допомогою процедури двох послідовностей місцевого генератора, зсунутих на півперіод.

Якщо один з імпульсів, які подаються на подільник, вилучити, то фаза вихідних імпульсів зсунеться на таке саме значення Δt у бік відставання (рис. 10.8, *в*). Отже, додаванням або відніманням імпульсів легко змінювати фазу тактової послідовності в потрібному напрямі.

Розглянутий принцип дії проміжного перетворювача, одну із функцій якого виконував подільник, характеризується тим, що на основі вимірювання зсуву фаз ЗГ виробляються імпульси додавання або віднімання. Цією дією визначається знак неузгодженості фаз і не визначається величина неузгодженості. Такі ПС називають пристроями з постійним корекційним ефектом.

У телекомукаційних засобах застосовують і ПС зі змінним корекційним ефектом, у яких кількість доданих або віднятих імпульсів пропорційна зсуву фаз.

Для ПС з постійним коригувальним ефектом сигнал про необхідність коригування частоти формується кінцевим станом ІЕ, який відповідає граничному значенню $t_{\text{доп}}$. Але можна використовувати не граничне значення $t_{\text{доп}}$, а проміжні значення t , що легко отримати, оскільки ІЕ – багатовихідний лічильник і залежно від кількості зареєстрованих відхилень частот знімається той чи той сигнал.

10.7. Системний аналіз окремих процесів перетворення сигналів та їх взаємодія в єдиній системі передавання дискретної інформації

Раніше ми розглядали загальний процес передавання інформації від джерела її виникнення до одержувача у вигляді окремих взаємопогоджених процесів перетворення дискретних сигналів, послідовність яких зведена до логічної схеми, до якої входять процеси: кодування; перетворення сигналів у форму, зручну для передавання по лінії зв'язку (каналах передавання даних); обернене перетворення сигналів (демодуляція) та їх реєстрація за належної синхронізації тактовних послідовностей; декодування прийнятих послідовностей сигналів та їх запам'ятовування для одержувача.

Пристрої, які виконують відповідні функції перетворення сигналів, називають відповідними пристроями перетворення сигналів. Система взаємозв'язаних пристроїв перетворення сигналів визначає структурну схему СПДІ.

Для узагальнення процесів перетворення сигналів і взаємодії пристроїв, які їх реалізують, розглянемо динаміку передавання та приймання послідовності дискретних сигналів – переносників повідомлення.

Розпочнемо з процесу синхронізації. Нагадаємо, що дві послідовності подій синхронні, якщо відповідні події в них виникають одночасно. Одна послідовність подій створюється в передавальній частині системи (у передавачі), а друга – у приймальній частині системи (у приймачі). Тривалість виникнення події можна інтерпретувати як тривалість дії одного елементарного сигналу. Звідси для створення послідовності тривалостей для кожного сигналу (події) потрібно створити послідовність часових однакових відліків як на передавачі, так і на приймачі.

Для розглядуваної системи синхронізації відліки за часом називають тактовими, а синхронізацію – тактовою синхронізацією. У літературі із систем передавання дискретних сигналів трапляється термін «розподільник тактових сигналів». Це пристрій, який надає певну тривалість для відповідної події.

Основний елемент такого розподілення – ЗГ. На передавальній частині СПДІ ЗГ створює одну послідовність дискретних сигналів (відліків), а на приймальному – другу послідовність сигналів (відліків).

Утворення тактових послідовностей розглянемо на часовій діаграмі (рис. 10.9).

Кожний тактовий відрізок призначений для певного елемента кодової послідовності; його ще називають одиничним елементом. Тактова частота визначає швидкість модуляції. Умова – $f_{\text{вч}} \gg f_r$. Інтервал частот $f_{\text{вч}} \dots f_r$ і самі частоти використовуються для підтримки синхронізму і реєстрації сигналів.

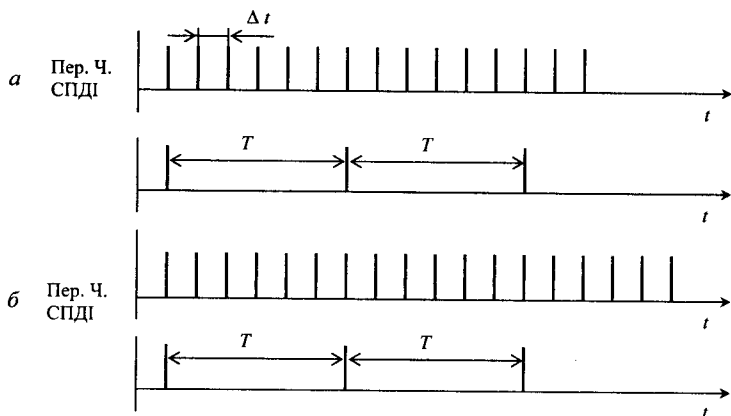


Рис. 10.9. Утворення тактових послідовностей СПДІ: *a* – передавальна частина (Пер.Ч.СПДІ); *б* – приймальна частина (Пр.Ч.СПДІ); $f_{\text{вч}}$ – власна частота ЗГ передавальної частини СПДІ; f_t – частота тактових сигналів

На передавачі частоту f_t утворює відповідний подільник: $f_{\text{вч}} = mf_t$, де m – коефіцієнт ділення, який входить до складу формувача тактових імпульсів. Тактові імпульси використовують для керування перетворювачем сигналів у форму, зручну для передавання каналом зв'язку згідно з прийнятим способом модуляції.

Модуляція полягає в змінах того чи того або одночасно декількох параметрів переносника залежно від закону змінювання інформаційного сигналу. Тоді значення дискретного сигналу модулюється на інтервалі T і такий процес одержав назву «маніпуляція».

Розглядаючи процес одночасного утворення тактових послідовностей на передавальній та приймальній частинах СПДІ, можна дійти висновку, що в межах відповідних тактів передавач формує сигнал, який спрямовує в лінію зв'язку інформаційний сигнал, а в цей час приймач сприймає його і реєструє таким, яким він надійшов.

Але відомо, що сигнал, який поширюється лінією зв'язку, підлягає впливу як власних характеристик, так і зовнішніх завад. Цей вплив виражається в зміні (спотворенні) його енергетичних та часових параметрів і він не надходить до приймача суворо у відповідному такті, а зі зсувом у часі і зміною його обвідної.

Розглянемо детально заходи боротьби з цими явищами. Не розриваючи логічну послідовність перетворення сигналів, розглянемо процес утворення тактових послідовностей. Поки що уявімо, що $f_{\text{вч}}$ передавача дорівнює $f_{\text{вч}}$ приймача і відповідно f_t їх однакові.

Приймальна частина СПДІ виконує обернену функцію шляхом перетворення, звільняючи цим самим модульований сигнал від переносника і виділяючи його обвідну, цим самим перетворює сигнал в імпульс постійного струму, виділяючи значущі моменти сигналу (демодуляція). Ці сигнали змінюються відносно передавальних, які можуть зумовити зміну самої значущої позиції.

Реєстрацію і встановлення значущої позиції виконують спеціальні пристрої реєстрації, які реалізують відповідні методи реєстрації (стробування, інтегральні аналогові та цифрові). Для створення цих пристроїв потрібно мати сигнали, розміщення яких відносно тактового імпульсу і частоти $f_{вч}$ генератора показано на рис. 10.10.

Для реалізації методу стробування потрібно мати імпульс (короткий), який має розміститися всередині інтервалу T , а для реалізації інтегрального аналогового методу – пороговий імпульс тривалістю інтервалу аналізу t_a і короткий імпульс реєстрації. Взагалі тривалість стробувальних імпульсів та імпульсів реєстрації однакова.

Із часової діаграми видно, що всі потрібні сигнали для будь-якого методу реєстрації можна отримати, використовуючи послідовність генератора імпульсів – ЗГ. Ці процеси ми розглядали відносно тактової послідовності сигналів, створених завдяки стабільності ЗГ. Але основна причина розбіжності в часі тактових послідовностей полягає безперечно в нестабільності генераторів, які використовуються для їх формування.

Правильне приймання інформації в таких системах можливе тільки тоді, коли моменти надходження кожного корисного сигналу заздалегідь відомі з достатньою точністю. Але це можливо тільки у разі, якщо відповідні тактові інтервали виникають водночас. Це потребує установалення та підтримки синхронізму. Зсув фаз зумовлений нестабільністю генератора і виражений або випередженням, або відставанням ЗГ передавача та приймача.

Для усунення цього явища є засоби підтримки синхронізму, які зводяться до впровадження високостабільних генераторів, використання окремих каналів (спеціальних) та пілот-сигналів, а також виявлення розбіжностей інформаційних (робочих) сигналів. Найбільшого поширення набув останній засіб. Суть його зводиться до того, що відхилення тактових частот від номінальних може спричинити розсинхронізацію через n переданих інформаційних імпульсів. Визначено, що через критичне значення часу $t_{\text{доп}} = \frac{\epsilon}{2k_r R}$ для найбільш несприятливих умов, коли ЗГ неста-

більні на обох частинах СПДІ, їх потрібно налагоджувати.

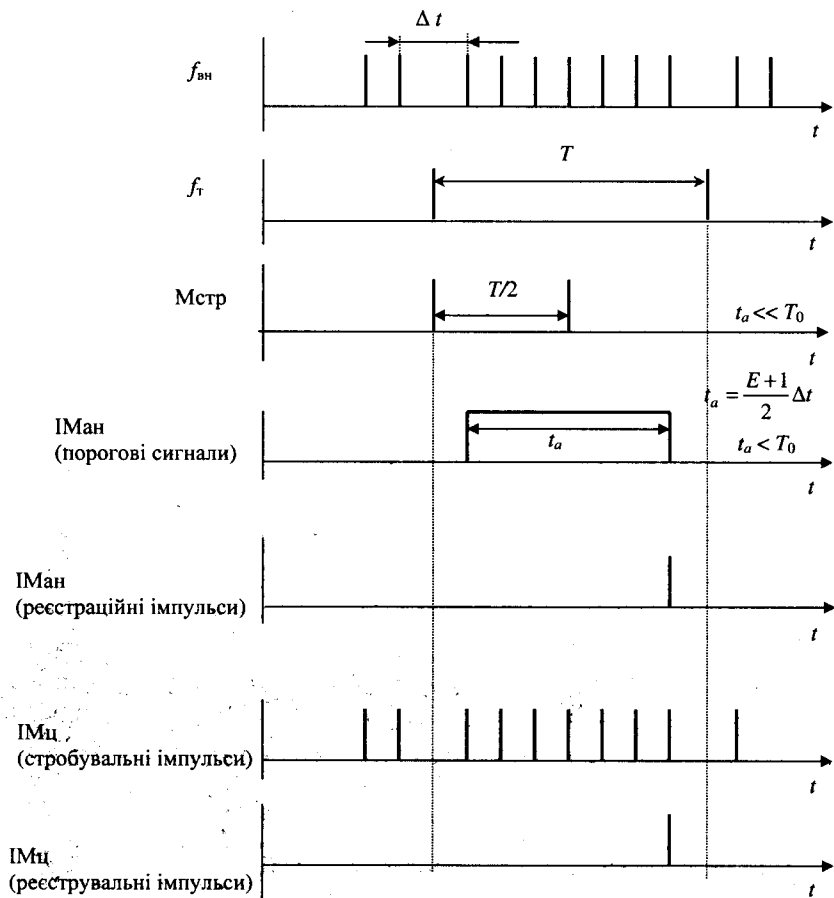


Рис. 10.10. Часові діаграми формування сигналів реєстрації: Мстр – метод стробування одиничним імпульсом (стробом), ІМан – інтегральний аналоговий метод, ІМц – інтегральний цифровий метод

Функції установлення та підтримки синхронізму виконує спеціальний пристрій. Принцип дії та складові цього пристрою розглянуто окремо (підрозд. 10.2).

Вимірювання відхилень виконує фазовий дискримінація, для дії якого потрібно мати дві послідовності імпульсів, частоти яких у два рази більші від тактових (рис. 10.11).

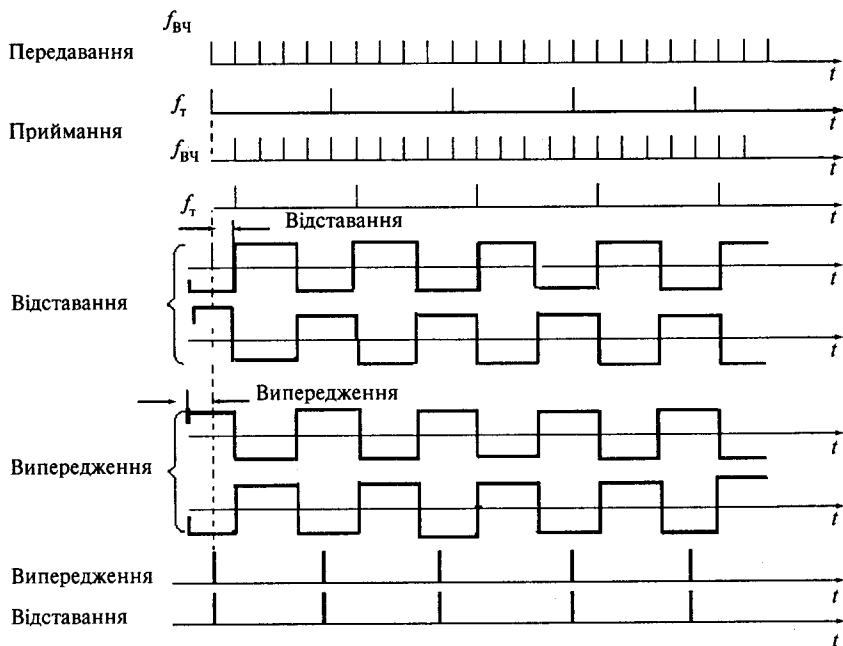


Рис. 10.11. Часові діаграми реєстрації сигналів відставання або випередження у разі зсуву фаз генераторів

Для вимірювання відхилень формуються дві тактові послідовності в протифазі. Одна з них збігається з послідовністю тактових імпульсів, яка і буде видавати імпульси відхилення або в бік відставання, або в бік випередження.

Такі послідовності легко одержати за допомогою відповідного подільника послідовності ЗГ приймача.

Сигнали на виході фазового дискримінатора можна використовувати для зміни частоти ЗГ приймача, але безпосередній принцип керування на порядок зменшує його стабільність. Тому поширення одержав метод керування частотою ЗГ з уведенням проміжного перетворювача.

Для цього імпульси відхилення за допомогою інерційного елемента накопичуються реверсивним лічильником до моменту, коли потрібно втручатись у зміну частоти ЗГ. Накопичування імпульсів визначають з

$$\text{виразу } t_{\text{доп}} = \frac{\varepsilon}{2k_r R}, \text{ або } k = \frac{\varepsilon}{2k_r R}.$$

Отже, сигнали про зміну частоти синхронізації керувальним пристроєм використовують для збільшення або зменшення кількості імпульсів, що надходять із ЗГ, як це показано на часовій діаграмі (рис. 10.12). Але слід зазначити, що для прийняття інформації керувальним пристроєм необхідно мати послідовність імпульсів ЗГпр і зсунуту на півперіод таку саму послідовність (рис. 10.12).

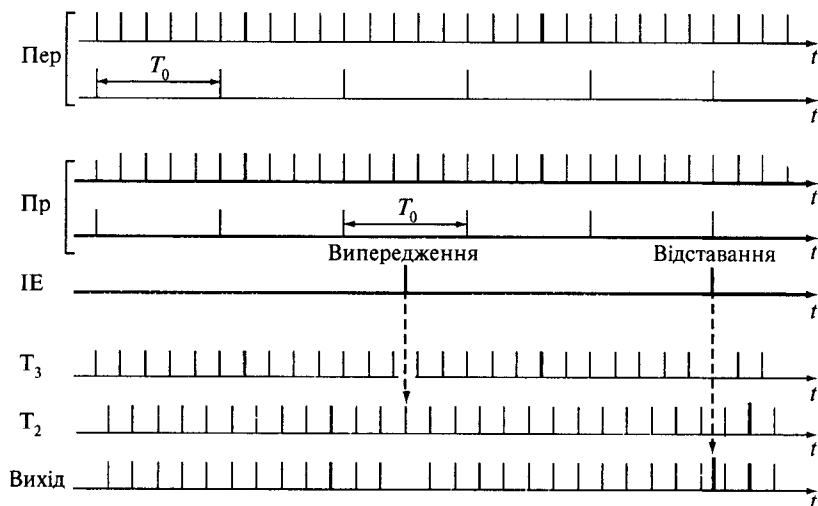


Рис. 10.12. Часова діаграма формування нової послідовності з урахуванням імпульсів відставання або випередження: Пер – передавач; Пр – приймач

На виході керувального пристрою утворюється зсунута послідовність відносно послідовності ЗГ, причому з уведенням імпульсу, якщо ЗГпр відстає, або з усуненням з послідовності одного імпульсу – якщо випереджає.

Завдяки введенню зсунутої послідовності одержимо послідовність ЗГпр з уведеними або усуненими імпульсами, що використовуються для зміни тактової частоти.

Для того щоб одержати змінні значення тактової частоти, тепер вихід ЗГпр буде виходом керувального пристрою. Але це ще не тактова послідовність. Для її створення потрібний спеціальний пристрій – проміжний перетворювач.

Нагадаємо, як можна з послідовності імпульсів утворити нову тактову частоту із часової діаграми (рис. 10.13). Із діаграми видно, що якщо додати імпульс у послідовність, тоді подільник перетворювача обчислює постійну кількість m імпульсів.

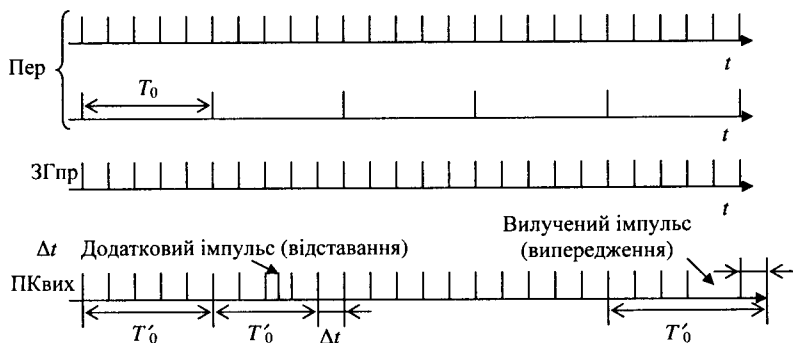


Рис. 10.13. Часова діаграма до процесу змінювання фази у разі відставання або випередження: ПКвих – вихід пристрою керування

Після такого детального аналізу процесів перетворення сигналів у системі та виявлення потрібних сигналів, які можна отримати від базової послідовності імпульсів ЗГпр за допомогою проміжного перетворювача, який виконує формування:

- імпульсів стробування для методу реєстрації стробування;
- опорних сигналів і реєструвальних імпульсів для інтегрального аналогового методу реєстрації;
- послідовності імпульсів для цифрового методу реєстрації і реєструвального імпульсу;
- протифазних послідовностей для виявлення відставання та випередження ЗГпр відносно ЗГпер;
- сигналів для зміни тактової частоти (прямої та зсунутої на півперіод послідовності ЗГпр);
- послідовності імпульсів з уведенням або усуненням із послідовності ЗГпр;
- тактової послідовності подільником проміжного перетворювача з керованої послідовності, яка з'являється на виході пристрою керування синхронізацією.

Досі ми розглядали процеси перетворення сигналів окремими пристроями, які є складовими частинами дискретного каналу СПДІ за однонаправленою роботи. Але сучасні пристрої, використовувані для організації комп'ютерного зв'язку каналами тональної частини телефонного зв'язку, складаються з передавальної та приймальної частин і називаються модемами.

Щоб краще з'ясувати процеси функціонування двостороннього півдуплексного обміну послідовностями сигналів, розглянемо структурну схему (рис. 10.14).

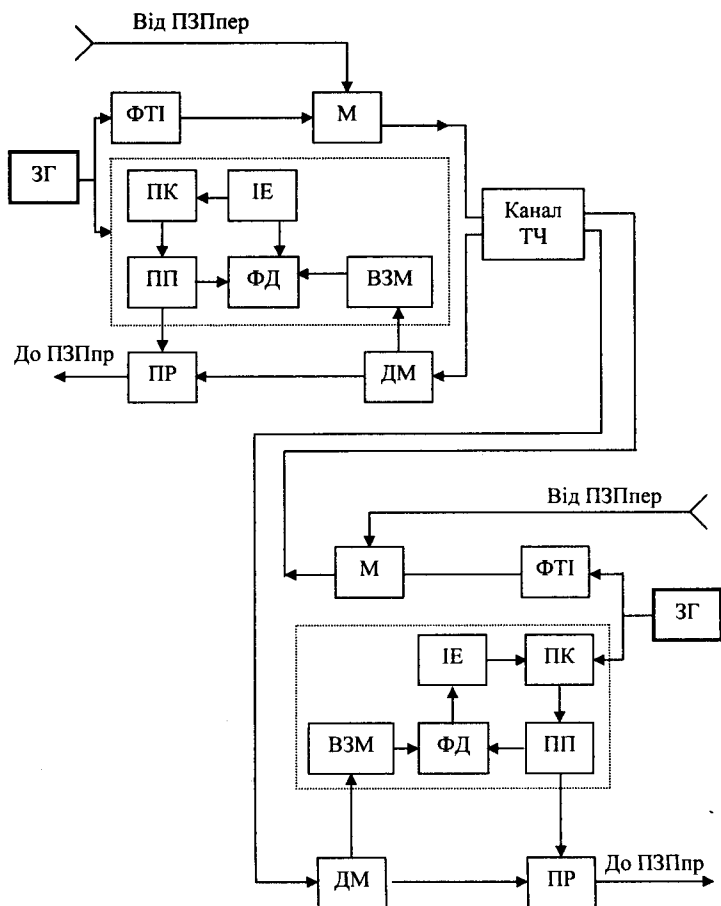


Рис. 10.14. Структурна схема дискретного каналу СПДІ для двостороннього передавання дискретної інформації:
 ФД – фазовий дискримінація; М – модулятор; ДМ – демодулятор;
 ПП – проміжний перетворювач; ФТІ – формувач тактових імпульсів;
 ПР – пристрій ресинхронізації

Таким чином, наведена структурна схема (рис. 10.14), часові діаграми окремих процесів перетворення сигналів (рис. 10.9–10.13) та принципові схемотехнічні пристрої, що описуються у попередніх підрозділах, дозволяють повністю прослідкувати процес перетворення сигналів під час організації двостороннього зв'язку між взаємодіючими абонентськими пунктами.

11. Моделювання передавання дискретної інформації дискретним каналом

11.1. Загальні положення

Методи формування сигналів – переносників передавальних первинних інформаційних сигналів мають такі властивості, що дозволяють розпізнавати інформацію, яка міститься в цих сигналах. Але не існує методів перетворення простих інформаційних сигналів, таких як прямокутні імпульси постійного струму, в значно складніші сигнали, які забезпечували б гарантовану достовірність передавання каналами зв'язку.

Безперечно, що будь-який сигнал зазнає впливу з боку середовища, через яке він поширюється. Це можуть бути канали електрозв'язку, волоконно-оптичні та канали радіозв'язку. Ці впливи визначаються характеристиками цих каналів і по-різному діють на сигнал, що поширюється через них. Крім цього впливу, на передавальний сигнал діють зовнішні завади, які також змінюють форму сигналів. Проблемою зменшення такого впливу на передавальний сигнал займається теорія сигналів. І не випадково у науково-технічній літературі та наукових спеціальних періодичних виданнях і дисертаційних роботах обґрунтовуються пропонувані методи, засоби та принципи створення сигналів чи систем передавання, які підвищують, наприклад, достовірність передавання інформації, але допускають можливі спотворення сигналів, що спричиняють виникнення помилок у передаваній (прийнятій) інформації.

Як відомо, на передавальні сигнали (їх форму) впливають характеристики, або параметри каналу, а тому спотворення інформаційних сигналів можна обчислити, але самі параметри каналу залежать від зміни їх значень під дією зовнішнього середовища (температури, енергії живлення і т. ін.), зокрема зовнішніх завад. Більшість зовнішніх завад виникають випадково, що ускладнює їх урахування у процесі формування передавальних сигналів. Тому і вискоєфективні сигнали не в змозі протистояти сукупності випадкових завад, особливо тим зовнішнім завадам, які мають імпульсний характер з амплітудами, близькими, або більшими від амплітуд інформаційних сигналів протилежної полярності, що призводить до виникнення помилок у приймальній інформації.

Умови передавання інформації дискретними сигналами передбачають потребу апріорно (ззадалегідь) передбачати можливості виникнення уражень такого ступеня, що це призведе до помилки в передаваній інформації.

Детально зупинимося на процесах передбачення явищ, які зв'язані з випадковостями виникнення помилок в передавальних кодових послідовностях або із встановленням закономірностей їх появи.

Очевидно, що строго відобразити навіть форми сигналів-завад, які можуть виникнути і діяти на інформаційні сигнали, майже неможливо, та у цьому і немає особливої потреби.

Один із потужних способів вивчення розглядуваних процесів – це їх моделювання. Моделювання трактується як філософська категорія – метод пізнання. Розрізняють фізичне моделювання і математичне моделювання.

Під час фізичного моделювання модель відтворює досліджуваний процес (оригінал) зі збереженням його фізичної природи. Цей вид моделювання передбачає передавання дискретної інформації між двома комп'ютерами (абонентськими машинами) через штучну лінію зв'язку і штучний генератор завад, що потребує побудови еквівалентної лінії зв'язку та генератора, який би формував очікувані послідовності та форми сигналів-завад. Переваги фізичного моделювання перед натурним експериментом полягають у тому, що в разі реалізації процесу моделі можуть значно відрізнятися від умов, які властиві процесу-оригіналу, і їх можна вибирати, виходячи зі зручності та простоти дослідження.

Фізичне моделювання має обмежену сферу застосування. Свідомо більш широкі можливості має математичне моделювання. Під математичним моделюванням розуміють спосіб дослідження різних процесів через вивчення явищ, які мають різний фізичний зміст, але описуються однако-вими математичними співвідношеннями. Для вивчення будь-якого процесу методом математичного моделювання потрібно побудувати його математичний опис, або математичну модель.

Математична модель реальної системи, до якої належить досліджувана система передавання дискретної інформації, – це абстрактний формально описаний об'єкт, вивчити який можна математичними методами, зокрема за допомогою математичного моделювання. Процесу формалізації будь-якого реального процесу передують попереднє вивчення структури складових його явищ, результатом чого є змістовий опис процесу – перша спроба чітко висловити закономірності, характерні для досліджуваного процесу, і постановку прикладної задачі. Змістовий опис – це початковий матеріал для наступних етапів формалізації: побудови формалізованої схеми процесу і математичної моделі для нього.

Формалізована схема процесу – це проміжна ланка між змістовним описом і математичною моделлю.

11.2. Узагальнена модель дискретного каналу

Розглянемо побудову формалізованої схеми процесів і їх взаємодії під час функціонування СПДІ у спрощеному вигляді. Уявімо собі спрощений варіант передавання даних СПДІ (рис. 11.1).



Рис. 11.1. Спрощений варіант СПДІ, що пояснює загальний процес передавання даних

Складові частини цього варіанта СПДІ: $A = (a_1, a_2, \dots, a_n)$ – послідовність імпульсних посилок, які є переносниками закодованої інформації, де a_i може набувати значення 1 або 0; модулятор – пристрій, що формує новий сигнал, який відображає значення передавального імпульсу і наділяє його тривалістю $T = \frac{1}{f_T}$; f_T – тактова частота; канал зв'язку – при-

стрій, через який передаються послідовно модульовані імпульсні посилки; джерело завад – джерело (генератор) послідовності імпульсних завад; пристрій, який перший у приймальній частині СПДІ приймає з каналу зв'язку інформаційні сигнали, що надходять.

Аналогом дії каналу зв'язку у разі сумісного надходження двох послідовностей інформаційних сигналів та сигналів-завад може бути подання дискретного каналу як суматора за модулем два.

Під час вивчення принципів побудови пристроїв перетворення сигналів та методів реєстрації і синхронізації сигналів, реалізованих цими пристроями як складовими дискретного каналу, є очевидним, що правильність реєстрації одиничних елементів цифрових сигналів даних залежить від багатьох чинників – завад зовнішнього і внутрішнього походження, різноманітних спотворень, спричинених реальними, а не ідеальними характеристиками тракту передавання сигналів. Тому в реальних умовах приймання одиничних елементів відбувається з помилками. Це означає, що замість символу даних «1» приймається символ «0» і навпаки. Оскільки завади з'являються і впливають на корисні інформативні послідовності випадково, то й помилки в дискретному каналі мають випадковий характер.

У загальному випадку, якщо на вхід дискретного каналу подати деяку сукупність «0» і «1», позначену умовно $A = (a_1, a_2, \dots, a_n)$, то в результаті дії завад у каналі на виході дискретного каналу з'явиться дещо видозмінена послідовність $A^* = (a^*_1, a^*_2, \dots, a^*_n)$, як це показано на рис. 11.2.

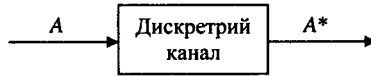


Рис. 11.2. Дискретний канал із завданнями

Аналізуючи залежності між A і A^* , неважко бачити, що рис. 11.3 можна зобразити і у вигляді, показаному на рис. 11.3.

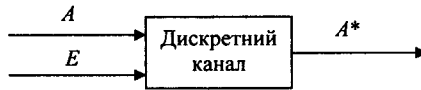


Рис. 11.3. Дискретний канал з вектором помилок E

Отже, цілком обґрунтованим є подання дискретного каналу як перетворювача двох послідовностей:

A – послідовності корисних інформаційних сигналів (кодових комбінацій);

E – послідовності символічних сигналів, що зумовлюють появу помилок в A .

У разі такого формального подання дискретного каналу послідовності A та E в часі абсолютно синхронні. Тоді процесам перетворення, передавання і відображення (реєстрації) послідовностей, очевидно, можна надати формального вигляду як суми за модулем два двох синхронних послідовностей:

$$A^* = A \oplus E = (a_1 \oplus e_1, a_2 \oplus e_2, \dots, a_n \oplus e_n) = (a^*_1, a^*_2, \dots, a^*_n). \quad (11.1)$$

Із виразу (11.1) видно, що основне завдання побудови математичної моделі дискретного каналу передавання інформації зводиться до пошуку математичного виразу, аналітичного чи імітаційного алгоритму, що описує закономірності та числові значення параметрів вектора $E = (e_1, e_2, \dots, e_n)$.

Одержання даних, потрібних для встановлення закономірностей виникнення вектора E , пов'язано з вимірюванням, реєстрацією й обробкою експериментальних вимірів дискретного каналу. Кожний елемент вектора помилок E – це результат складання за модулем два відповідних елементів векторів A та A^* , тобто

$$e_n = a_n \oplus a^*_n.$$

Нагадаємо правило складання за модулем два (табл. 11.1).

Очевидно, що якщо всі елементи вектора E дорівнюють одиниці, то вихідна послідовність тотожно дорівнює вхідній. З огляду на це модель дискретного каналу можна подати як суматор за модулем два (рис. 11.4).

Таблиця 11.1

| A | E | A^* |
|-----|-----|-------|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 1 | 0 |
| 1 | 0 | 1 |

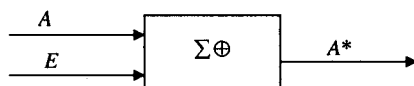
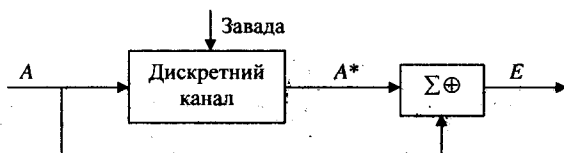


Рис. 11.4. Модель дискретного каналу

Отже, функціонування дискретного каналу можна зобразити формальною збільшеною схемою суматора за модулем два, на входи якого надходять синхронно два потоки: A – потік інформаційних сигналів; E – потік сигналів-завад. Під дією процедури $\Sigma \oplus$ потоки A і E утворюють вихідний потік A^* . Звідси теорія передавання дискретних сигналів займається методами та засобами, які забезпечують мінімуми впливу завад на корисні сигнали.

Але для того, щоб зменшити кількість одиниць у послідовності E , потрібно знати закономірність їх виникнення. Для встановлення закономірності виникнення помилок (одиниць у послідовності E) треба мати багаторазову реєстрацію (експерименти) спостережуваних явищ – виникнення помилок. Для цього розглянемо варіанти спостереження та реєстрації помилок.

Вимірювання помилок. Узагальнену схему знаходження вектора помилок E показано на рис. 11.5.

Рис. 11.5. Вимірювальна система вектора E

Методи вимірювання вектора помилок E можна розділити на дві групи: 1) за цифровим інформаційним сигналом; 2) за випробувальним сигналом. У першій групі методів вимірювання виконують з використанням певних ознак структурної регулярності цифрового сигналу. Такими ознаками є:

- повторюваність комбінацій циклової і надциклової синхронізації;
- чергування полярності імпульсів лінійного сигналу;
- обмеження кількості нулів, що надходять підряд (пауз, тривалість яких більша від нормованої);
- обмеження цифрової суми коду;
- обмеження кількості імпульсів однієї полярності в модифікованих квазітрійкових кодах, що надходять підряд.

Усі зазначені ознаки можна використовувати для виявлення (детектування) помилок у цифровому сигналі, що дозволяє зробити підрахунок помилок. Перша група методів знайшла застосування переважно в системах телеконтролю лінійних і групових трактів.

У другій групі методів вирізняють методи вимірювання з використанням сигналів регулярного типу і сигналів типу псевдовипадкових послідовностей. Перевага першої групи методів – можливість організації контролю і вимірювання без припинення зв'язку. Водночас ці методи забезпечують невисоку точність вимірів.

Набагато більшу точність можна одержати використанням псевдовипадкової послідовності (ПВП) як вимірювального сигналу. У цьому разі відомо структуру сигналу на передавальному кінці. Відповідно на приймальному кінці можна генерувати аналогічну послідовність і виконувати побітове порівняння двох ПВП для виділення помилок. Недолік цього методу – складність синхронізації двох послідовностей. З теорії завадостійких кодів відомо методи, що дозволяють обробляти прийняту ПВП із виокремленням помилок, що не потребують складніших від тактової видів синхронізації. Цей метод і використовуватимуть для вимірювань.

Дискретний канал вивчають на експериментальній установці – типовій системі передавання даних, схему якої зображено на рис. 11.6.

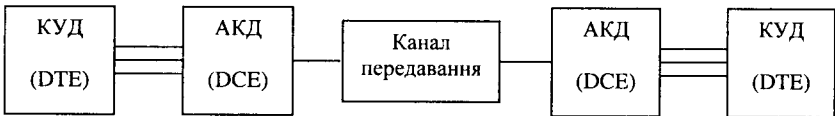


Рис. 11.6. Структурна схема дискретного каналу передавання даних:
 КУД (DTE) – узагальнена назва кінцевого устаткування даних (Data Terminal Equipment); АКД (DCE) – узагальнена назва апаратури каналу даних (Data Communication Equipment)

Часто як DTE може виступати персональний комп'ютер.

Функція DCE полягає в забезпеченні можливості передавання інформації між двома і більшою кількістю DTE по каналу певного типу, наприклад по телефонному. DCE може бути аналоговим модемом, якщо використовується аналоговий канал, чи, наприклад, пристроєм обслугову-

вання каналу (CSU/DSU – Channel Service Unit / Data Service Unit), якщо використовується цифровий канал типу E1/T1 чи ISDN.

Вимірювання виконують з використанням тестових рекурентних послідовностей з наступним їх порівнянням з аналогічним еталонним тестом. Завдання одержання експериментальних даних полягає в спостереженні за вектором E і визначенні кількості помилок за одиницю часу (наприклад, за хвилину), чи кількості помилок на деякий обсяг переданої інформації. На підставі цих даних можна встановити закономірності розподілу помилок.

Установлення закономірностей виникнення випадкових подій вивчається в курсі «Теорія імовірностей і математична статистика» (розділ «Статистичний аналіз випадкових величин»).

11.3. Аналітичний статистичний аналіз експериментальних даних

11.3.1. Поняття та завдання математичної статистики

Багаторазове спостереження за деякими випадковими явищами дозволяє помітити їх певні закономірності. Вивчивши ці закономірності, можна деякою мірою керувати випадковими явищами, передбачати і враховувати їх у своїй діяльності.

Очевидно, що вивченню підлягають тільки такі випадкові явища, які можна, принаймні, спостерігати необмежену кількість разів. Такі випадкові явища називають масовими. При цьому найбільший інтерес становлять випадкові масові явища, що мають певні закономірності.

Математична статистика – це розділ математики, що вивчає методи збирання (реєстрації), систематизації (опису) й оброблення (аналізу) результатів спостережень випадкових масових явищ з метою виявлення існуючих закономірностей. Висновки про закономірності, яким підпорядковуються явища, що досліджуються методами математичної статистики, завжди ґрунтуються на обмеженій вибірковій кількості спостережень. Тому природно припустити, що ці висновки за більшої кількості спостережень можуть виявитися іншими. Щоб висловити більш обґрунтоване судження про досліджуване явище, математична статистика спирається на теорію імовірностей.

На відміну від математичної статистики, яка займається результатами спостережень над випадковими величинами, теорія імовірностей формально-логічно вивчає закономірності випадкових явищ і використовує абстрактні описи дійсності з математичними моделями випадкових явищ. Розглянемо основні завдання математичної статистики.

1. Визначення закону розподілу випадкової величини. Закономірності спостережуваних випадкових явищ виявляються тим точніше й чіткіше, чим більший обсяг статистичного матеріалу.

На практиці доводиться опрацьовувати обмежений обсяг експериментальних даних. Тому результати спостережень і їх обробка завжди містять більший або менший елемент випадковості. Виникає питання щодо того, які ознаки спостережуваного явища постійні, стійкі і дійсно властиві явищу, а які випадкові і виявляються лише через обмежену кількість спостережень. Тому методика оброблення статистичного матеріалу має виявляти все типове і відкидати нетипове. У зв'язку з цим виникає характерна для математичної статистики задача знаходження закону розподілу випадкової величини.

2. Перевірка правдоподібності гіпотез. Отриманий з дослідів статистичний матеріал може лише з більшою чи меншою правдоподібністю підтверджувати або спростовувати справедливість висунутої гіпотези про закон розподілу розглянутого випадкового явища $F(x)$. Крім того, якщо розглядаються дві випадкові величини, то належить установити, залежні вони, чи ні. Для цього математична статистика виробила критерії згоди і спеціальні прийоми.

3. Знаходження невідомих параметрів розподілу. У результаті оброблення статистичного матеріалу закон розподілу можна і не визначати, тому що він або відомий, або для його визначення дуже мало даних. Тоді виникає більш вузька задача – визначити основні параметри, тобто числові характеристики випадкової величини.

За невеликої кількості статистичних даних можливий елемент випадковості, а тому про точне визначення числових характеристик говорити не доводиться. Отже, задача зводиться до знаходження «оцінок» чи «придатних значень» для визначуваних параметрів, тобто таких наближених значень, які у разі масового застосування приводили б до менших помилок, ніж усі інші. Відшукування «оцінок» тісно пов'язано із визначенням їх точності та надійності. Слід пам'ятати, що який би метод оцінювання не застосовувався, отриманий результат потрібно надалі перевіряти. Інакше можна відхилити правильну гіпотезу, або ж прийняти неправильну.

Вивчаючи якісну чи кількісну ознаку, що характеризує сукупність однорідних об'єктів, не завжди можна обстежити кожний об'єкт досліджуваної сукупності. Наведемо такий приклад. Електричну лампочку умовимося вважати стандартною, якщо тривалість її горіння не менше ніж 1200 год. За якістю продукції зобов'язаний стежити завод-виробник. Досліджувати кожну лампочку на тривалість горіння майже неможливо, крім того, це суперечить здоровому глузду. Як же одержати уявлення про якість виготовленої продукції? Нехай завод має поставити споживачу

партію готових виробів. Замість даних про якість усіх електричних лампочок партії досить мати точні знання про якість невеликої її частини, відібраних випадково. За тривалістю горіння відібраних лампочок можна судити про якість усіх лампочок партії. Практика підтверджує, що зроблені висновки бувають досить надійними.

Вихідна інформація для дослідження – множина значень випадкової величини, отриманих у результаті спостережень над нею; їх називають *випадковою вибіркою*, або просто вибіркою. Наприклад, вибірка (x_1, x_2, \dots, x_n) значень деякої випадкової величини x розглядається як результат n незалежних повторних вимірів, причому всі вимірювання мають виконуватись в однакових умовах. А сукупність усіх можливих (іноді їх називають мислимыми) значень досліджуваної випадкової величини називають генеральною сукупністю. Кількість об'єктів у генеральній сукупності й у вибірці називають їх об'ємами. Генеральна сукупність може мати як кінцевий, так і нескінченний обсяг.

Слід зауважити, що коли зростає обсяг вибірки (n), багато вибірових статистик збігаються за ймовірністю з відповідними параметрами теоретичного розподілу величини x . Тому кожен вибірку розглядають як вибірку з теоретично нескінченної генеральної сукупності, розподіл ознаки в якій збігається з теоретичним розподілом імовірностей величини x . Його називають розподілом генеральної сукупності, а його параметри – параметрами генеральної сукупності. У багатьох застосуваннях теоретична генеральна сукупність є ідеалізацією дійсної сукупності, з якої отримано вибірку.

11.3.2. Побудова варіаційного ряду

Найчастіше вибірки являють собою ряди експериментальних значень, які не зовсім зручні для безпосереднього аналізу. Для вивчення даних їх передусім треба згрупувати. Розмістимо значення досліджуваної величини (ознаки) у міру зростання. Цю операцію називають групуванням даних по інтервалах, чи *ранжуванням* статистичних даних. Наприклад, для вибірки 0, 1, 1, 0, 0, 2, 0, 1, 0, 3 ранжований ряд має вигляд: 0, 0, 0, 0, 0, 1, 1, 1, 2, 3.

З ранжованого ряду випливає, що ознака набула чотирьох різних значень. Надалі різні значення ознаки умовимось називати *варіантами*, а під *варіюванням* розуміти зміну значень ознаки. Якщо ознака за своєю сутністю така, що різні її значення не можуть відрізнитися одне від одного менше, ніж на деяку скінченну величину, то ця ознака *дискретно варіюється*. Наприклад, у результаті ранжування ряду 0, 1, 1, 0, 0, 2, 0, 1, 0, 3 одержують варіанти 0, 1, 2 і 3.

Надалі умовимося позначати варіанти так: M_i ($i = 1..k$), де k – кількість варіантів, або інакше, груп. Очевидно, що різні варіанти трапляються неоднакову кількість раз. Число, що показує, скільки разів трапився варіант M_i у ряді спостережень, називають частотою варіанта F_i . Крім того, замість частоти варіанта іноді доцільно розглядати її відношення до загальної кількості спостережень n , що називають відносною частотою, і позначають через P_i . Справедлива така рівність:

$$P_i = \frac{F_i}{n}.$$

Усі зазначені величини зводять у таблицю, яку називають частотною таблицею, чи варіаційним рядом. Складемо подібну таблицю для ряду значень, отриманих реальними вимірюваннями вектора E (кількість помилок за часовий інтервал) у дискретному каналі. Вимірювання проводять протягом 509 год. Кожне значення наведеного ряду відповідає кількості зареєстрованих помилок за годину (табл. 11.2).

Таблиця 10.1

| Варіант M_i | Частота варіанта F_i | Відносна частота варіанта P_i |
|---------------|------------------------|---------------------------------|
| 0 | 315 | 0,619 |
| 1 | 142 | 0,279 |
| 2 | 40 | 0,078 |
| 3 | 9 | 0,018 |
| 4 | 2 | 0,004 |
| 5 | 1 | 0,002 |
| – | $\Sigma = 509$ | $\Sigma = 1$ |

Побудована таблиця дає більш наочне уявлення про дані. Якщо перегляд первинних даних не дозволяє скласти уявлення про варіювання значень ознаки (особливо за значних обсягів вибірки), то, розглядаючи варіаційний ряд, можна зробити такі висновки: ознака коливається від 0 до 5; найчастіше трапляється 0; частота варіанта зменшується від 0 до 5.

Очевидно, що значення, набуті ознакою, можуть відрізнитися одне від одного на як завгодно малу величину, тобто ознака може набути будь-якого значення в деякому числовому інтервалі. Ознаку називають такою, що неперервно варіюється. Побудова дискретного варіаційного ряду в цьому разі не дасть бажаних результатів (надто багато різних значень ознаки, яку спостерігали). Для більшої переконливості об'єднують у групи значення ознаки, які коливаються в межах деякого інтервалу. Тоді за

аналогією частота F_i^i буде показувати, у скількох спостереженнях ознака набула значень, що належать тому чи тому інтервалу. Таку частоту називають інтервальною, а відношення її до загальної кількості спостережень – інтервальною відносною частотою P_i^i . Таблицю, що дозволяє зробити висновок про розподіл частот (чи відносних частот) між інтервалами варіювання значень ознаки, називають інтервальним варіаційним рядом.

Інтервальний варіаційний ряд будують на підставі даних спостережень за ознакою, що неперервно варіюється, та за ознакою, що дискретно варіюється, але кількість спостережуваних варіантів дуже велика. Дискретний варіаційний ряд будують тільки для ознаки, що дискретно варіюється.

Іноді інтервальний варіаційний ряд умовно замінюють дискретним. Тоді середнє значення інтервалу беруть за варіант M_i , а відповідну інтервальну частоту – за F_i .

Для побудови інтервального варіаційного ряду треба передусім визначити величину інтервалу. Для визначення оптимального інтервалу h , тобто такого, за яким побудований інтервальний ряд не був би занадто громіздким і водночас дозволив виявити характерні ознаки розглянутого явища, можна використовувати формулу Стерджеса:

$$h = \frac{X_{\max} - X_{\min}}{1 + 3,322 \cdot \lg n},$$

де X_{\max} і X_{\min} – відповідно максимальний і мінімальний варіанти. Якщо h – дробове число, то за значення інтервалу варто взяти або найближче ціле число, або найближчий нескладний дріб. Крім того, використання табличних даних значно прискорює подальші процеси оброблення інформації.

За допомогою даних таблиці можна побудувати гістограму. Для цього в прямокутній системі координат по осі абсцис відкладають відрізки, що зображують інтервали варіювання, і на цих відрізках як на основі будують прямокутники з висотами, що дорівнюють відносним частотам відповідного інтервалу. У результаті одержують східчасту фігуру – гістограму.

11.3.3. Знаходження статистичних характеристик

Наступний етап оброблення результатів вимірювань – знаходження показників, які деякою мірою характеризують вибірку: значення ознаки, навколо якого концентруються спостереження (міри цієї якісної особливості називають середніми величинами), і розсіювання спостережень навколо середніх величин (міри цієї особливості називають показниками варіації). Ці показники називаються *статистичними характеристиками*. Середні величини є немовби «представниками» усього ряду спостере-

жень, оскільки навколо них концентруються значення спостережуваної ознаки. Зазначимо, що тільки для якісно однорідних спостережень має сенс обчислювати середні величини. Найбільш поширена середня величина – *середня арифметична*. Найпростіша формула для обчислення:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n},$$

де n – обсяг вибірки.

Обчислимо середнє арифметичне для варіаційного ряду:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{509} = 0,5147.$$

Якщо за спостереженнями побудовано варіаційний ряд, то

$$\bar{x} = \frac{\sum_{i=1}^k M_i F_i}{n},$$

де k – кількість варіантів (груп); n – обсяг вибірки; M_i – i -й варіант, якщо ряд дискретний, або центр інтервалу, якщо ряд інтервальний; F_i – частота i -го варіанта.

Середні величини, характеризуючи варіаційний ряд числом, не відображають мінливості значень ознаки, що спостерігалася, тобто варіацію. Для цього використовують декілька різних показників варіації, найбільш відомим з яких є емпірична, чи експериментальна дисперсія (s^2). Під варіацією розуміють середню арифметичну квадратів відхилень результатів спостережень від їх середньої арифметичної:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n},$$

де n – обсяг вибірки.

Розрахуємо експериментальну дисперсію для варіаційного ряду:

$$s^2 = \frac{\sum_{i=1}^n (x_i - 0,5147)^2}{509} = 0,5995.$$

Якщо за результатами спостережень побудовано варіаційний ряд, то

$$s^2 = \frac{\sum_{i=1}^k M_i^2 F_i - n\bar{x}^2}{n}.$$

Середнє арифметичне й експериментальна дисперсія – найважливіші статистичні характеристики як вибірки, так і генеральної сукупності, а тому повинні бути обов'язково обчислені в ході статистичного аналізу.

11.3.4. Поняття про оцінювання параметрів

Оцінювання параметрів. У загальному значенні зміст оцінювання методів можна сформулювати як сукупність методів, що дозволяють робити науково обґрунтовані висновки про числові параметри розподілу генеральної сукупності за випадковою вибіркою з неї. Щодо математичного сподівання генеральної сукупності, то статистичне оцінювання параметрів полягає в тому, щоб знайти таку вибіркову характеристику, що дозволила б одержати, по можливості, найбільш точне та надійне уявлення про потрібний параметр. Склад вибірки випадковий, тому висновки про параметри генеральної сукупності, зроблені за вибірковими даними, можуть бути помилковими. Зі зростанням кількості елементів вибірки ймовірність правильного висновку збільшується. Тому будь-якому рішення, прийнятому за статистичною оцінкою параметрів, намагаються поставити у відповідність імовірність, що характеризує ступінь імовірності прийнятого рішення.

Сформулюємо задачу оцінювання параметрів у загальному вигляді. Нехай X – випадкова величина, що підлягає закону розподілу $F(X, \Theta)$, де Θ – параметр розподілу, числове значення якого невідоме. Досліджувати всі елементи генеральної сукупності для обчислення параметра Θ не можливо, тому про цей параметр намагаються судити за вибірками з генеральної сукупності.

Будь-яку однозначно визначену функцію результатів спостережень, за допомогою якої роблять висновок про значення параметра Θ , називають оцінкою (або статистикою) параметра Θ . Розглянемо деяку множину вибірок обсягом n кожна. Вибіркову оцінку параметра Θ , обчислену за i -ю вибіркою, позначимо через Θ_{n_i} . Оскільки склад вибірки випадковий, то можна вважати, що Θ_{n_i} набуде невідомого наперед числового значення, тобто є випадковою величиною.

Відомо, що випадкова величина визначається законом розподілу і числовими характеристиками a , отже, і вибіркову оцінку можна також описувати законом розподілу і числовими характеристиками.

Для того щоб відобразити випадковий характер вибірки обсягу n з генеральної сукупності, позначимо її через (X_1, X_2, \dots, X_n) , а вибірккову оцінку параметра Θ – через $\tilde{\Theta}_n$. Отже, можна записати $\tilde{\Theta}_n = f(X_1, X_2, \dots, X_n)$. Вибір оцінки, що дозволяє одержати добре наближення оцінюваного параметра, – основне завдання теорії оцінювання.

Основні властивості оцінок – це властивості незміщеності, ефективності та здатності до використання.

Оцінку $\tilde{\Theta}_n$ параметра Θ називають незміщеною, якщо її математичне сподівання дорівнює оцінюваному параметру Θ , тобто $M(\tilde{\Theta}_n) = \Theta$. Вимога незміщеності гарантує, що під час оцінювання параметрів систематичні помилки не виникатимуть.

Незміщену оцінку $\tilde{\Theta}_n$, що має найменшу дисперсію серед усіх можливих незміщених оцінок параметра Θ , обчислених за вибірками такого самого обсягу, називають ефективною оцінкою.

Оцінку $\tilde{\Theta}_n$ параметра Θ називають спроможною, якщо вона підкоряється закону великих чисел, тобто за досить великої кількості незалежних спостережень n з імовірністю, близькою до одиниці, можна стверджувати, що різниця між $\tilde{\Theta}_n$ і Θ за абсолютною величиною виявиться меншою за як завгодно мале невід’ємне число τ , або

$$P\{|\tilde{\Theta}_n - \Theta| < \tau\} > 1 - \eta,$$

де η – невід’ємне число, близьке до нуля. Отже, у разі використання спроможних оцінок справджується збільшення кількості членів вибірки, оскільки дедалі менш імовірною стає можливість появи істотної помилки в оцінці невідомого параметра.

Розглянемо способи одержання основних властивостей оцінок. Одним з перших методів оцінювання параметрів був метод моментів, розроблений К. Пірсоном.

Нехай відомий закон розподілу випадкової величини X визначається декількома параметрами: $\Theta_1, \Theta_2, \dots, \Theta_q$, їх числове значення невідоме. Щоб знайти вибірккові оцінки, виконують n незалежних спостережень x_i , ($i = \overline{1, n}$) над випадковою величиною X . Згідно з методом моментів потрібно q перших моментів випадкової величини X порівняти з вибіркковим моментом, отриманим з експериментальних даних. Формули для розрахунку першого початкового (математичного сподівання) і другого центрального (дисперсію) моментів за вибіркою наведено в підрозд 2.5.

Теоретичним обґрунтуванням методу моментів служить закон великих чисел, відповідно до якого у разі великого обсягу вибірки вибіркові моменти близькі до моментів генеральної сукупності.

За методом моментів можна одержувати усереднені оцінки. Отже, надійність висновків, зроблених з його використанням, залежить від кількості спостережень.

Як приклад застосування методу моментів розглянемо випадкову величину X , що має нормальний закон розподілу та визначається двома параметрами: математичним сподіванням $M(X)$ і дисперсією $D(X)$. За цим методом невідоме математичне сподівання оцінюється середньою арифметичною, а дисперсія — вибірковою дисперсією s^2 .

Метод моментів рекомендують застосовувати на практиці, оскільки він приводить до порівняно простих обчислень. Крім того, існують методи максимальної правдоподібності та найменших квадратів, які хоча й приводять до більш складних обчислень, ніж метод моментів, але всі оцінки, одержані за їх допомогою, зазвичай виявляються більш надійними і кращими, особливо за малої кількості спостережень.

Поняття про довірчий інтервал. Оцінку невідомого параметра генеральної сукупності одним числом називають точковою оцінкою. Поряд із точковим оцінюванням статистична теорія оцінювання параметрів займається питаннями інтервального оцінювання. Завдання інтервального оцінювання в загальному вигляді можна сформулювати так: за даними вибірки побудувати числовий інтервал, щодо якого q із наперед обраною імовірністю можна стверджувати, що всередині цього інтервалу знаходиться оцінюваний параметр. Інтервальне оцінювання особливо потрібне за малої кількості спостережень, коли точкова оцінка малонадійна.

Довірчим інтервалом $[\tilde{\Theta}_n^{(1)}; \tilde{\Theta}_n^{(2)}]$ для параметра Θ називають такий інтервал, щодо якого можна із заздалегідь обраною імовірністю $p = 1 - \alpha$, близькою до одиниці, стверджувати, що він містить невідоме значення параметра Θ , тобто $P[\tilde{\Theta}_n^{(1)} < \Theta < \tilde{\Theta}_n^{(2)}] = 1 - \alpha$. Чим менший для обраної імовірності $[\tilde{\Theta}_n^{(2)} - \tilde{\Theta}_n^{(1)}]$, тим точніша оцінка невідомого параметра Θ , і, навпаки, якщо цей інтервал великий, то оцінка, отримана за його допомогою, малопридатна для практики. Оскільки кінці довірчого інтервалу $\tilde{\Theta}_n^{(1)}$ і $\tilde{\Theta}_n^{(2)}$ залежать від елементів вибірки (ними визначаються), то значення $\tilde{\Theta}_n^{(1)}$ і $\tilde{\Theta}_n^{(2)}$ можуть змінюватись від вибірки до вибірки. Імовірність $p = 1 - \alpha$ називають довірчою імовірністю, а число α — рівнем значущості.

Якими принципами варто керуватися для вибору довірчої імовірності? Вибір довірчої імовірності не є математичною задачею, а визначається конкретною розв'язуваною проблемою. Наведемо такий приклад. Нехай на двох підприємствах імовірність випуску придатних виробів $p = 1 - \alpha = 0,99$, тобто ймовірність випуску бракованих виробів $\alpha = 0,01$. Чи можна в рамках математичної теорії, тобто не цікавлячись характером виробів, що випускаються, з'ясувати, мала чи велика ймовірність α ?

Нехай одне підприємство випускає електролампи, а друге – парашути. Якщо на 100 ламп виявиться одна бракована, то з цим можна миритися за умови, що викинути 1 % ламп дешевше, ніж перебудувувати технологічний процес. Якщо ж на 100 парашутів трапляється один бракований, це може мати серйозні наслідки, що недопустимо. Отже, у першому випадку ймовірність браку α прийнятна, а у другому – ні, тому довірчу ймовірність $p = 1 - \alpha$ варто вибирати, виходячи з конкретних умов.

Статистична гіпотеза і закони розподілу ймовірностей. Відомо, що часто для з'ясування справжності того чи того факту вдаються до висловлення гіпотез, які можна перевірити статистично, тобто, спираючись на результати спостережень у випадковій вибірці. Під статистичною гіпотезою розуміють усяке висловлення про генеральну сукупність, яку перевіряють за вибіркою. Статистичні гіпотези класифікують на гіпотези про закони розподілу і гіпотези про параметри розподілу. Розглянемо тільки гіпотези про закони розподілу.

Якщо припустити, що значення експериментального ряду розподілено за деяким відомим розподілом, то це рівносильно гіпотезі про те, що статистичні характеристики вибірки, такі як середнє арифметичне й експериментальна дисперсія, тотожно дорівнюють числовим характеристикам випадкової величини, таким як математичне сподівання і дисперсія. З огляду на це нескладно знайти параметри розподілу через відомі статистичні характеристики.

Розподіли ймовірностей, найбільш використовувані для перевірки гіпотез, наведено в табл. 11.3.

Якщо вважати, що значення варіаційного ряду розподілено за законом Пуассона, то, як видно з табл. 11.3 розподілів, математичне сподівання має збігатися за величиною з дисперсією. У цьому випадку це не так ($0,5147 \neq 0,5995$), але ці значення можна усереднити:

$$\bar{x} = s^2 = \frac{0,5147 + 0,5995}{2} = 0,5571.$$

Таблиця 11.3

| Розподіл | Функція розподілу | Математичне сподівання | Дисперсія |
|---------------------------------|---|--------------------------------------|---|
| Біноміальний | $f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$ | $\bar{x} = np$ | $s^2 = np(1-p)$ |
| Геометричний | $f(x) = p(1-p)^{x-1}$ | $\bar{x} = \frac{1}{p}$ | $s^2 = \frac{(1-p)}{p^2}$ |
| Від'ємний біноміальний | $f(x) = \frac{(n+x-1)!}{x!(n-1)!} p^n (1-p)^x$ | $\bar{x} = \frac{n(1-p)}{p}$ | $s^2 = \frac{n \cdot (1-p)}{p^2}$ |
| Пуассонівський | $f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$ | $\bar{x} = \lambda$ | $s^2 = \lambda$ |
| Експоненціальний (показниковий) | $f(x) = \frac{1}{b} e^{-\frac{x}{b}}$ | $\bar{x} = b$ | $s^2 = b^2$ |
| Нормальний (гауссів) | $f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}}$ | $\bar{x} = \mu$ | $s^2 = \sigma^2$ |
| Логнормальний | $f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\log \frac{x}{m})^2}{2\sigma^2}}$ | $\bar{x} = m e^{\frac{\sigma^2}{2}}$ | $s^2 = m^2 e^{\frac{\sigma^2}{2}} (e^{\frac{\sigma^2}{2}} - 1)$ |
| Ерланга | $f(x) = \frac{\left(\frac{x}{b}\right)^{c-1}}{b(c-1)!} e^{-\frac{x}{b}}$ | $\bar{x} = bc$ | $s^2 = b^2 c$ |
| Гіперболічний | $f(x) = \frac{nh^n}{(x+h)^{n+1}}$ | $\bar{x} = \frac{h}{n-1}$ | $s^2 = \frac{2h^2}{(n-1)(n-2)}$ |
| Логістичний | $f(x) = \frac{e^{-\frac{(x-a)}{k}}}{k(1+e^{-\frac{(x-a)}{k}})^2}$ | $\bar{x} = a$ | $s^2 = \frac{k^2 \pi^2}{3}$ |

Отже, $\lambda = 0,5571$, а передбачуваний закон щільності розподілу імовірності має вигляд

$$f(x) = \frac{0,5571^x \cdot e^{-0,5571}}{x!}.$$

11.3.5. Перевірка статистичної гіпотези

Критерії узгодженості дозволяють визначити причини розбіжностей між теоретичною кривою і статистичним розподілом – чи це чисто випадковий обставини через обмежену кількість спостережень, чи це помилковий вибір теоретичного розподілу. Для цього слід на підставі наявних статистичних матеріалів перевірити гіпотезу H про те, що випадкова величина X підкоряється деякому визначеному закону розподілу. Цей закон може бути заданий найчастіше у вигляді функції $F(x)$ або щільності $f(x)$ розподілу, або ж у вигляді сукупності ймовірностей того, що випадкова величина X потрапить у межі i -го розряду.

Важливо зазначити, що за допомогою перевірки статистичних гіпотез можна лише відкинути гіпотезу, що перевіряється, але ніколи не можна довести її справедливості.

Для прийняття чи спростування гіпотези H потрібно вибрати якусь характеристику V , яка б показувала ступінь розбіжності теоретичного і статистичного розподілів. Як показник V можна брати: суму квадратів відхилень теоретичних ймовірностей від відповідних частот; максимальне відхилення статистичної функції розподілу від теоретичної і т. ін.

За такого вибору показник V являє собою випадкову величину, закон розподілу якої залежить від закону розподілу X і від кількості вимірів n . Розглянемо застосування показника V .

Нехай у результаті серії випробувань встановлено, що обраний показник V набув значення U , і, крім того, закон розподілу величини V відомий. Потрібно встановити ймовірність обраної теоретичної кривої. Якщо гіпотеза H правильна, то величина U пояснюється випадковими причинами. З урахуванням цього припущення необхідно розглядати ймовірність того, що через випадкові причини показник відхилення V виявиться не менше, ніж спостережуване значення U . Якщо ця ймовірність $P(V > U)$ мала, то гіпотезу H можна вважати малоправдоподібною. Велике значення ймовірності події $V > U$ свідчить про правдоподібність висунутої гіпотези H щодо закону розподілу досліджуваної випадкової величини.

Показник відхилення V треба вибирати таким, щоб його закон розподілу майже не залежав від $F(x)$ за досить великого n . Цій вимозі відповідають критерії згоди. Розглянемо його.

Критерій χ^2 Пірсона. Критерій згоди χ^2 Пірсона – найчастіше використовуваний критерій для перевірки гіпотези про закон розподілу.

Розіб'ємо весь діапазон зміни X на ℓ інтервалів $\Delta_1, \Delta_2, \dots, \Delta_\ell$ і підрахуємо кількість елементів F_i , що потрапили в кожний з інтервалів Δ_i . Припускаючи відомим теоретичний закон розподілу $F(x)$, завжди можна визначити p_i (ймовірність попадання випадкової величини X в інтервал Δ_i), тоді теоретичне число значень випадкової величини X , що потрапили в

інтервал Δ_i можна розрахувати за виразом np_i . Результати об'єднуємо і заносимо в табл. 11.4.

Таблиця 11.4

| Інтервали Δ_i | Δ_1 | Δ_2 | ... | Δ_i | ... | Δ_l |
|---------------------------|------------|------------|-----|------------|-----|------------|
| Емпіричні частоти F_i | F_1 | F_2 | ... | F_i | ... | F_l |
| Теоретичні частоти np_i | np_1 | np_2 | ... | np_i | ... | np_l |

Перевіряючи узгодженість теоретичного і статистичного розподілів, часто розглядають розбіжність між теоретичними ймовірностями p_i і спостережуваними частотами F_i/n . Як міру такої розбіжності вибирають суму квадратів відхилень

$\left(\frac{F_i}{n} - p_i\right)^2$, узятих з деякою «вагою» C_i , тобто

виражають V як

$$V_i = \sum_{i=1}^l C_i \left(\frac{F_i}{n} - p_i\right)^2.$$

Вагові коефіцієнти C_i потрібні для того, щоб відхилення різних інтервалів не можна було вважати рівноправними.

Англійський математик Карл Пірсон показав, що якщо

$$C_i = \frac{n}{p_i},$$

то за великих n закон розподілу величини V не залежить від функції розподілу $F(x)$ і кількості дослідів n , а залежить тільки від кількості розрядів. Цей закон зі збільшенням n наближається до розподілу χ^2 .

У результаті одержимо

$$V = \chi^2 = \sum_{i=1}^l \frac{(F_i - np_i)^2}{np_i}.$$

Розподіл χ^2 залежить від параметра r – числа степенів вільності цього розподілу. Число степенів вільності r дорівнює числу розрядів l мінус число незалежних умов s (зв'язків), накладених на частоти F_i/n , тобто $r = l - s$.

Такими зв'язками можуть бути:

– обов'язкова умова $\sum_{i=1}^l \frac{F_i}{n} = 1$;

– $\sum_{i=1}^l x_i \frac{F_i}{n} = M[x]$ (якщо треба підібрати теоретичний розподіл за умови збіжності теоретичних і статистичних середніх значень);

$$- \sum_{i=1}^l (x_i - \bar{x})^2 \frac{F_i}{n} = D_x$$
 (якщо потрібна збіжність теоретичної і статистичної дисперсії) і т. ін.

Існують спеціальні таблиці розподілу χ^2 , за допомогою яких для кожного значення χ^2 і числа степенів вільності r можна знайти ймовірності того, що величина V перевершить розраховане значення χ^2 .

Наскільки малою має бути ймовірність $P(V > \chi^2)$ для того, щоб спростувати гіпотезу, невідомо. Практично встановлено, що якщо $P \leq 0,01$, доцільно перевірити експеримент, і в разі повторної появи істотних відхилень вибрати більш придатний закон розподілу.

Якщо ймовірність $P(V > \chi^2)$ порівняно велика, то можна визнати розбіжності між теоретичним і статистичним розподілами несуттєвими і вважати їх випадковими.

Отже, правило застосування критерію χ^2 таке: розрахувавши значення χ^2 і вибравши рівень значущості критерію α , за таблицею χ^2 -розподілу визначати $\chi_{r;\alpha}^2$. Якщо $\chi^2 > \chi_{r;\alpha}^2$, то гіпотезу відхиляють, якщо $\chi^2 \leq \chi_{r;\alpha}^2$, то гіпотезу приймають.

Зазначимо, що неодмінною умовою застосування критерію Пірсона є наявність у кожному з інтервалів щонайменше 5 – 10 спостережень. Якщо кількість спостережень в окремих інтервалах дуже мала (1 чи 2), то треба об'єднати відповідні інтервали.

Перевіримо відповідність даних віріаційного ряду, наведеного у табл. 11.2, розподілу Пуассона за допомогою критерію χ^2 , якщо довірчий інтервал 0,95. Для цього використовуємо формулу $f(x) = \frac{0,5571^x \cdot e^{-0,5571}}{x!}$,

а результат зводимо в табл. 11.5.

Таблиця 11.5

| Інтервали Δ_i | 0 | 1 | 2 | 3 | 4 | 5 | Σ |
|-------------------------------|------|------|------|---|------|---|----------|
| Емпіричні частоти F_i | 315 | 142 | 40 | 9 | 2 | 1 | 509 |
| Теоретичні частоти np_i | 291 | 162 | 45 | 9 | 1 | 1 | 509 |
| $\frac{(F_i - np_i)^2}{np_i}$ | 1,98 | 2,47 | 0,56 | | 0,09 | | 5,1 |

Далі знаходимо за таблицею χ^2 -розподілу величину $\chi_{r;\alpha}^2(2; 0,95) = 0,95$, де 2 – число степенів вільності. Тоді $\chi^2 \leq \chi_{r;\alpha}^2$, гіпотезу не відхиляємо, вона може бути правильною.

Критерій Колмогорова – Смирнова. Інший поширений критерій згоди – критерій А. М. Колмогорова, що для практичного розрахунку значно простіший від критерію χ^2 . При цьому показником відхилення служить максимальне значення модуля різниці між статистичною функцією розподілу $F^*(x)$ й апроксимуючою теоретичною функцією $F(x)$:

$$V = D = \max |F^*(x) - F(x)|.$$

А. М. Колмогоров показав, що незалежно від $F(x)$ за необмеженого зростання кількості дослідів n імовірність події

$$P(D\sqrt{n} \geq \lambda)$$

прямує до межі

$$P(\lambda) = 1 - \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2\lambda^2}. \quad (11.1)$$

Значення $P(\lambda)$, розраховані за формулою (11.1), зведено в таблиці Колмогорова – Смирнова.

Отже, послідовність застосування критерію А. М. Колмогорова така: виконують порівняння на тому інтервалі, де експериментальний розподіл має найбільше (за абсолютною величиною) відхилення від теоретичного розподілу, потім цю абсолютну різницю R порівнюють з критичним значенням, обчисленим за формулою

$$D_{\text{кр}} = \frac{z}{\sqrt{n}},$$

де n – обсяг вибірки; z – коефіцієнт, значення якого визначають за таблицею Колмогорова – Смирнова і яке залежить від значення рівня значущості α .

Якщо максимальна різниця $R > D_{\text{кр}}$, гіпотезу відхиляють, а якщо $R \leq D_{\text{кр}}$, то гіпотезу приймають.

Критерій Колмогорова застосовують лише тоді, якщо $F(x)$ заздалегідь відома і відомі входні в неї параметри, що трапляється рідко, тому що з теоретичних міркувань може бути відомий лише закон розподілу, а параметри визначають для конкретного числового прикладу.

Умови використання критеріїв. Цілком природно виникає питання: коли варто використовувати критерій χ^2 -Пірсона, а коли критерій Колмогорова – Смирнова? Очевидно, що кожен критерій має свої сильні та слабкі сторони і дає тільки вказівки у разі перевірки тієї чи тієї гіпотези. Утім відомо, що визначальний чинник для вибору критерію – обсяг вибірки n . Так, якщо $n \geq 100$, варто використовувати критерій χ^2 -Пірсона,

у діапазоні $10 \leq n \leq 99$ – критерій Колмогорова – Смирнова, а якщо $n \leq 10$, – критерій Крамера – фон Мізеса.

Якщо всі гіпотези, що перевіряються, були відхилені, то для установлення закономірностей розподілу використовують метод імітаційного моделювання на основі генератора псевдовипадкових чисел.

У цьому розділі наведено основні теоретичні положення математичної статистики, які можна використовувати для побудови математичної моделі, яка описує вектор E -помилку, що виникають у дискретному каналі передавання інформації.

Обґрунтовані теоретичні положення зведемо в упорядковану систему процедур і подамо її у вигляді окремої методики як інструментальних засобів ідентифікації законів розподілу помилок, що виникають у дискретних каналах СПДІ.

11.3.6. Принципи побудови моделі дискретного каналу

До цього ми розглядали у формальному вигляді дискретний канал передавання інформації, який зводився до дії суматора за модулем два на два вхідні дискретні синхронні потоки – A і E , детально ознайомилися зі статистичним аналізом обробки результатів зареєстрованих помилок і встановленням закономірності їх виникнення. При цьому уявляли, що потік кодових комбінацій без уведення надлишковості або, інакше, без використання завадостійких кодів до кодових комбінацій, утворюється внутрішнім кодером джерела інформації.

Формалізована модель дискретного каналу буде дієздатна тоді, коли можна на ній відобразити будь-який вхідний потік дискретних повідомлень і ідентифікований потік помилок, а також вихідний потік, який утворюється в результаті операції «модуль два». До повної функції моделі каналу належить порівняльний аналіз вхідного A та вихідного A^* потоків. Останню функцію моделі введено для розширення її можливості щодо дослідження впровадження методів підвищення достовірності передавання дискретної інформації: завадостійкого кодування та розв'язувального зворотного зв'язку.

Очевидно, що побудова такої моделі – це трудомісткий процес, та і сама модель громіздка.

Зазначимо, що джерело дискретної інформації (термінальні устаткування дискретної дії) на виході свого кодера утворює випадкову дискретну послідовність $\{a_i\}$, яку відповідні пристрої перетворюють у послідовність випадкових сигналів, і являє собою не що інше, як вхідний потік дискретних сигналів.

Уявімо, що дискретний канал – перетворювач двох послідовностей $A = (a_1, a_2, \dots, a_n)$ і $E = (e_1, e_2, \dots, e_n)$ в одну $A^* = (a_1^*, a_2^*, \dots, a_n^*)$, причому елементи a_i і e_i в часі виникають синхронно. Тоді в будь-який момент часу дискретний канал можливо буде знаходитись в одному зі станів: «1» або «0» відповідно до передавального сигналу, але під дією відповідного у часі елемента e_i передавані «1» або «0» можуть змінити свій стан. Процес зміни стану каналу або на сигнальному рівні зміни значущої позиції передавального сигналу можна подати у вигляді ймовірнісних переходів у можливі стани системи (каналу). Ці зміни зображено на діаграмах (рис. 11.7).

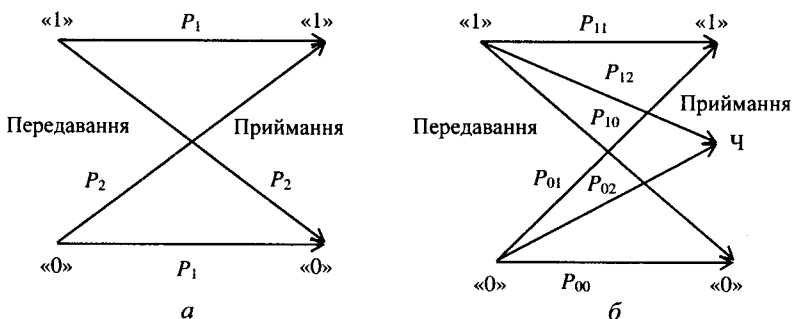


Рис. 11.7. Діаграми ймовірнісних переходів: *a* – у двійковому каналі; *б* – у симетричному каналі зі стиранням

На рис. 11.7, *б* показано і відповідні ймовірності наслідків p_{11} , p_{00} , p_{10} , p_{01} , $p_{0ч}$, $p_{1ч}$, де перший індекс позначає символ на вході, а другий – на виході дискретного каналу.

Якщо $p_{0ч} \neq 0$ і $p_{1ч} \neq 0$, тоді дискретний канал називають каналом зі стиранням, якщо ж $p_{0ч} = p_{1ч} = 0$, – то каналом без стирання.

Якщо $p_{11} = p_{00}$, $p_{10} = p_{01}$, тоді такий канал називають симетричним.

Вхідний потік A – це послідовність символів a_i , які передаються каналом протягом тактового інтервалу часу T_b , і послідовність E символів e_i (вхідний потік послідовності помилок).

Ці міркування відносно дії двох потоків випадкових послідовностей у дискретному каналі мають ідентичну збіжність з поведінкою в часі математичних випадкових дискретних процесів, у яких стани системи, які описують процес, можуть змінюватися у наперед фіксовані моменти часу. Такі випадкові дискретні процеси зручно описувати за допомогою ланцюгів Маркова. Ланцюги Маркова – це потужний математичний апарат, використовуваний для створення аналітичних моделей дискретних каналів передавання інформації, тому наведемо їх основні означення і властивості.

Основними для марковських процесів є поняття «стани системи» і «переходи з одного стану в другий». Система знаходиться в довільному стані, якщо вона повністю описується значеннями змінних, які задають цей стан. Система здійснює перехід з одного стану у другий, якщо описані її змінні змінюються від значень, які задають один стан, на значення, які визначають другий. Нехай ці стани можуть змінювати один одного (або зберігаються) тільки у наперед фіксовані моменти часу $t_0, t_1, t_2, \dots, t_i$. Надалі ці моменти часу будемо називати позиціями (кроками) і позначати їх номерами $0, 1, 2, \dots, i$. Зручно вважати, що час між переходами постійний, що повністю відповідає тривалості передавання сигналу або бодового інтервалу, який визначається тактовою частотою.

Припустімо для загального випадку, що існує N станів, перенумерованих числами від 1 до N . Якщо робота системи описується простим марковським процесом, тоді ймовірність її переходу зі стану i в стан j протягом наступного інтервалу часу є функцією тільки i і j і не залежить від поведінки системи до моменту її попадання у стан i . Інакше кажучи, можна ввести набір умовних імовірностей $[p_{ij}]$ того, що система, яка знаходиться у стані i , після чергового переходу опиниться у стані j . Оскільки вона обов'язково має попасти у деякий стан після цього переходу, то

$$\sum_{j=1}^N p_{ij} = 1.$$

Тут враховується і ймовірність p_{ij} того, що система залишиться у стані i . Оскільки p_{ij} є імовірностями, то

$$0 \leq p_{ij} \leq 1.$$

Подібні матриці називають стохастичними. Для повного опису простого ланцюга Маркова, окрім переходу, потрібно задати початкові ймовірності p_i ($1, 2, \dots, N$) станів системи на початковому кроці. Одне з основних питань для вивчення ланцюгів Маркова полягає в тому, щоб, знаючи матрицю – рядок початкових імовірностей $[p_i]$, і матрицю переходів $[p_{ij}]$, визначити ймовірності p_{ik} i -го стану системи через k -кроків і дослідити поведінку p_{ik} , якщо $k \rightarrow \infty$.

Як уже зазначалось, особливістю більшої частини існуючих моделей дискретних каналів є те, що канал стаціонарний, симетричний і в ньому не виникають помилки синхронізації (вставки та випадіння) і стирання. За цих обмежень модель дискретного каналу зводиться до моделі задової послідовності \vec{E}_i (у нескінченних межах). Цю послідовність потрібно визначати так, щоб послідовність задових символів \vec{e}_i (на кінцевому

відрізку від 1 до N) можна було вважати її реалізацією з достатнім ступенем точності. Існує велика кількість моделей з указаними обмеженнями.

Протягом тривалого часу, коли немає даних для реальних каналів, вважали, що помилки у каналах зв'язку виникають незалежно. У разі незалежних помилок достатньо знати єдиний параметр p_n – імовірність помилки – для визначення розподілу випадкової величини. На підставі теореми Бернуллі імовірність появи в n -елементній комбінації, що дорівнює t помилок $p(t, n)$, визначається біноміальним розподілом

$$p(t, n) = C_n^t p_n^t (1 - p_n)^{n-t}, \text{ якщо } 0 \leq t \leq n. \quad (11.2)$$

Із формули (11.2) видно, що ймовірність приймання неспотвореної комбінації ($t = 0$)

$$p(0, n) = (1 - p_n)^n,$$

а ймовірність приймання помилкової комбінації, тобто комбінації, у якій виникає хоч би одна помилка,

$$p(\geq 1, n) = 1 - p(0, n) = 1 - (1 - p_n)^n;$$

імовірність появи m і більше помилок у кодовій комбінації довжиною n

$$p(\geq m, n) = \sum_{t=m}^{t=n} C_n^t p_n^t (1 - p_n)^{n-t}.$$

Найпростіша модель помилок дає грубе наближення до реальних каналів. Подальше дослідження потоку помилок у реальних каналах показали, що незалежні помилки порівняно рідкісні, особливо для передавання цифрової інформації телефонними каналами зв'язку. Було виявлено, що помилки є залежними і мають тенденцію до групування (пакування). Далі було помічено, що у ряді випадків і самі пакети помилок залежні й утворюють «пакети пакетів».

Завадостійкі коди, оптимальні для каналів з незалежними помилками, виявились неефективними під час їх використання у реальних каналах. Для вибору ефективних кодів потрібно якомога більше знати про статистику помилок та математичні моделі каналів, які якнайточніше відображали б реальний характер помилок. Одна з перших запропонованих моделей дискретного каналу, у якій використовується математична схема ланцюгів Маркова, – це модель Гілберта. У ній дискретний канал розглядають як систему, що може знаходитись у будь-який момент часу в одному із двох станів – задовільному, коли помилки неможливі, і незадо-

вільному, коли виникають незалежні помилки з імовірністю ε . Послідовність станів утворюють простий ланцюг Маркова. У цій моделі послідовність E повністю визначається матрицею перехідних імовірностей:

$$P = \begin{vmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{vmatrix}.$$

Для відображення процесу пакетування помилок імовірності зміни станів каналу p_{01} і p_{10} мають бути значно меншими від імовірностей збереження стану каналу p_{00} і p_{11} .

Імовірність помилки у каналі

$$p_n = \varepsilon p_1 = \frac{\varepsilon p_{01}}{(p_{01} - p_{10})}.$$

Дискретний канал моделі Сміта – Боуена – Джойса має три стани. Один стан – незадовільний, за якого виникають незалежні помилки з імовірністю ε_0 , близькою до 0,5. Такий стан відповідає пакетам помилок. Два інші стани – задовільні (один з них відображає проміжок між пакетами помилок, а другий – проміжок між групами пакетів). Імовірності незалежних помилок у разі задовільних станів ε_1 і ε_2 значно менші за 0,5, але і не обов'язково однакові. Послідовність трьох станів – простий ланцюг Маркова. Матриця перехідних імовірностей

$$P = \begin{vmatrix} P_{00} & 0 & P_{02} \\ 0 & P_{11} & P_{12} \\ P_{20} & P_{21} & P_{22} \end{vmatrix}.$$

Імовірність помилки у дискретному каналі

$$p_n = \varepsilon_0 P_0 + \varepsilon_1 P_1 + \varepsilon_2 P_2.$$

Отже, модель Сміта – Боуена – Джойса дозволяє відобразити не тільки наявність пакетів, а й найпростіші їх групування. Одна з перших моделей каналу із залежними помилками – це модель Мертца, призначена для опису часових і амплітудних характеристик завад. На базі моделі імпульсних шумів вивчають характеристики послідовності помилок, які виникають у каналі під час передавання цифрової інформації.

У цій моделі за експериментально отриманими даними вимірювань було виявлено велику розбіжність між розподілом амплітуд імпульсних завад та гауссівим розподілом, а також подібність з поведінкою гіперболічної функції. На цій підставі вважають, що розподіл амплітуд імпульс-

них завад підпорядкований гіперболічному закону високого порядку зі щільністю амплітуди

$$p(x) = \frac{k}{(x+h)^m},$$

де x – амплітуда; k, h, m – константи.

Інтегральна ймовірність $p(x)$ амплітуд, що дорівнює або більша за x ,

$$p(x) = k \int_x^{\infty} \frac{dv}{(v+h)^m} = \frac{k}{(m-1)(x+h)^{m-1}}.$$

Нормувальна константа має забезпечити отримання $p(x) = 1$ у разі накопичення від 0 до ∞ відповідно:

$$1 = \frac{k}{(m-1)h^{m-1}} \text{ або } k = (m-1)h^{m-1}.$$

Часовий розподіл імпульсних завад виражено в термінах помилок. При цьому припускається, що помилки виникають пакетами, розподіл – пуассонівський. Самі помилки в межах одного пакета також передбачаються підпорядкованими розподілу Пуассона.

Для відображення пакетів помилок у моделях дискретних каналів можна використати моделі процесів, які вивчаються в теорії відновлення. Теорія відновлення отримала численні застосування у таких дослідженнях, як теорія надійності, теорія масового обслуговування, теорія запасів і багато інших.

Для розглядуваних процесів передавання дискретних сигналів каналом, у якому виникають за випадковими законами послідовності помилок з тенденціями угруповання (пакетування), привабливі альтернативні процеси відновлення. Для таких процесів придатні безпосередні процеси відновлення з дискретним часом.

Процес відновлення з дискретним часом – двійкова послідовність \vec{D}_i (рис. 11.8, а), яка однозначно визначається заданням розподілів $P(\Lambda_j)$ – довжин серій нулів Λ_j і $P(L_j)$ – довжин серій одиниць L_j .

Випадкова величина Λ_j – це довжина серії нулів λ_j , де $j = 0, 1, 2, \dots$ – порядковий номер серії (рис. 11.8, б), випадкова величина L_j – довжина серії одиниць l_j , де $j = 0, 1, 2, \dots$ – порядковий номер серії (рис. 11.8, в). Двійковій послідовності \vec{D}_i відповідають свої порядкові номери $i = 0, 1, 2, \dots$

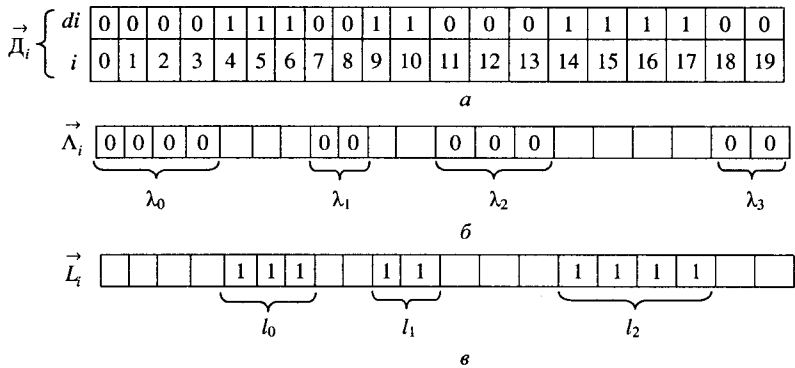


Рис. 11.8. Двійкові послідовності: a – стани \vec{D}_i ; $б$ – серії нулів $\vec{\Lambda}_j$;
 $в$ – серії одиниць \vec{L}_j

Якщо випадкові величини Λ_j статистично незалежні у сукупності, тоді послідовність \vec{D}_i називають потоком одиниць з обмеженою післядією, заданим одновимірним розподілом $P(\lambda_j)$. Якщо у сукупності незалежні випадкові величини L_j , тоді послідовність \vec{D}_i називають потоком нулів з обмеженою післядією. Його визначають заданням одновимірних розподілів $P(l_j)$.

Послідовність \vec{D}_i розглядуваної моделі можна розбити на відрізки з одиничних елементів двох видів – пакети помилок та інтервали між пакетами.

Для загального випадку послідовність \vec{D}_i – це не що інше, як вектор \vec{E} послідовності сигналів-завад, які спричиняють помилки у приймальних інформаційних послідовностях пристроями СПДІ.

Отже, знання характеристик складових вектора \vec{E} дозволяє вирішувати нові завдання пошуку методів і заходів щодо запобігання помилкам.

11.3.7. Методика аналізу експериментальних даних – зареєстрованих помилок та устанавлення їх закономірності виникнення

Збір і аналіз даних. Важлива функція для моделювання помилок, які виникають у каналах передавання даних, – збір та аналіз інформації про них. Відомо, що всі можливості боротьби із завадами на рівні сигналів повністю використані апаратними методами СПДІ в частині дискретного каналу. Завади, які спотворюють повністю сигнал, або його інверсують чи стирають, потрібно навчитись виявляти на етапі збору даних, а потім їх проаналізувати, устанавивши закономірності їх виникнення.

Щоб зібрати такі дані, треба виконати експериментальні виміри. Експериментальні виміри виникнення помилок зводяться до побудови вектора помилок $E = (e_1, e_2, \dots, e_n)$. Довжина цієї кодової комбінації має бути не меншою за певну величину, яку визначають методами математичної статистики.

Опис статистичних (вимірних) даних. Під час збору реальних вимірних даних для встановлення закономірності їх виникнення стикаються з проблемою зведення «сирих» даних до зручного для аналізу вигляду. Для вирішення цієї проблеми використовують способи пошуку та опису найважливіших властивостей набору даних. Ці способи дозволяють організувати дані за рахунок втрат деякої змістовної інформації.

Один із методів перетворення даних до зручного вигляду – групування даних за класами (інтервалами, варіантами). Дані потім зводять у таблицю, яка містить кількість попадань елементів даних у кожний клас. Таку таблицю називають частотною. Вона зазвичай дозволяє мати загальне уявлення про дані.

Зібрані дані підсумовують здебільшого у вигляді розподілу відносних частот (гістограми). Для прикладу розглянемо дані спостережень появи запитів на встановлення телефонного зв'язку. Нехай було проведено спостереження за появою запитів за час тривалістю 509 год. У кожному годинному інтервалі зареєстровано від 0 до 5 запитів. Результати групового зведення наведено в табл. 11.6.

За даними цієї таблиці побудовано гістограму (рис. 11.9).

Гістограма для відносних частот появи запитів. Чи відповідає частота спостережуваних вибірових значень тій частоті, з якою вони мають з'являтися за деякого ймовірного розподілу, який відповідає певному теоретичному закону.

Закінчивши побудову гістограми, аналітик починає підбирати відповідний теоретичний закон розподілу ймовірностей.

Таблиця 11.6

| Кількість запитів N | Кількість одногодинних інтервалів з N запитами | Відносна частота |
|-----------------------|--|------------------|
| 0 | 315 | 0,619 |
| 1 | 142 | 0,279 |
| 2 | 40 | 0,078 |
| 3 | 9 | 0,018 |
| 4 | 2 | 0,004 |
| 5 | 1 | 0,002 |
| – | $\Sigma = 509$ | $\Sigma = 1$ |

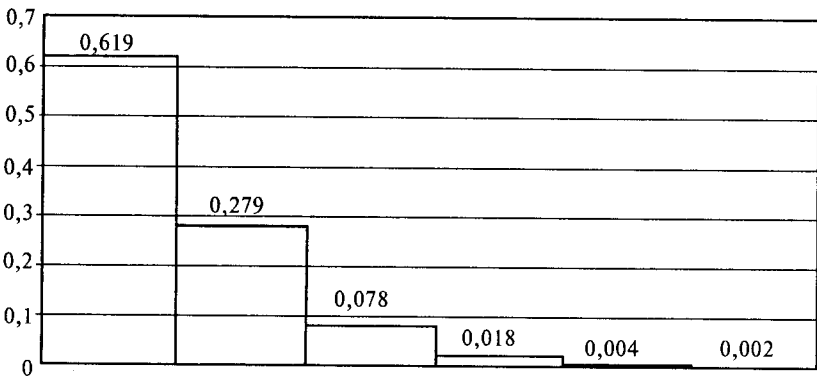


Рис. 11.9. Графічне зображення розподілу відносних частот

Візуальне порівняння отриманої гістограми з деякими кривими теоретичних розподілів дозволив виявити подібність з розподілом Пуассона. Але таке візуальне порівняння дозволяє лише припустити, до якого теоретичного розподілу потрібно спробувати «підігнати» експериментальне значення, і ніколи не дає достатніх підстав, щоб остаточно прийняти деяку гіпотезу (теоретичний розподіл).

Після того, як аналітик підібрав один, або декілька теоретичних розподілів (наприклад, пуассонівський), за яким він припускає можливість з'ясувати експериментальні дані, він визначає параметри, які треба перевірити за статистичними критеріями.

Якщо припустимий розподіл є функцією двох параметрів, решту звичай оцінюють на основі вибіркової середньої і вибіркової дисперсії.

Коли експериментальні дані розбиті на групи, тоді середнє і дисперсію можна обчислити за формулами:

$$\text{– середнє} \quad \bar{x} = \frac{\sum_{i=1}^k M_i F_i}{n}; \quad (11.3)$$

$$\text{– дисперсія} \quad S^2 = \frac{\sum_{i=1}^k M_i^2 F_i - n\bar{x}^2}{n-1}, \quad (11.4)$$

де n – повний обсяг вибірки, $n = \sum_{i=1}^k F_i$; k – кількість груп (інтервалів) вибірки; M_i – значення i -ї групи; F_i – частота появи i -го інтервалу.

Обчислення статистичних параметрів для дискретних даних наведено в табл. 11.7.

Таблиця 11.7

| M_i | F_i | $M_i F_i$ | $M_i^2 F_i$ |
|-------|-------|-----------|-------------|
| 0 | 315 | 0 | 0 |
| 1 | 142 | 142 | 142 |
| 2 | 40 | 80 | 160 |
| 3 | 9 | 27 | 81 |
| 4 | 2 | 8 | 32 |
| 5 | 1 | 5 | 25 |
| – | 509 | 262 | 440 |

Підставивши числові дані із табл. 11.7 у формули (11.3) і (11.4), одержимо

$$\bar{x} = \frac{262}{509} = 0,5147; \quad S^2 = \frac{440 - 509 \cdot 0,5147^2}{509} = 0,5995.$$

Статистичний аналіз. Візуально було зроблено припущення, що дані табл. 11.6 можуть відповідати розподілу Пуассона. Із статистики відомо, що середнє цього розподілу дорівнює дисперсії (позначеної через λ), але з табл. 11.7 видно, що для експериментальних даних середнє не дорівнює дисперсії: $0,5147 < 0,5995$. Це вже причина відмовитися від гіпотези про те, що експериментальний розподіл – пуассонівський.

Але в цьому конкретному випадку маємо практичні і теоретичні підстави не відмовлятися від цієї гіпотези. Коли ймовірність деякої події для одного часового інтервалу така сама, як і для будь-якого іншого, а здійс-

нення будь-якої події не впливає на ймовірність її повторної появи, тоді є підстава очікувати розподіл Пуассона.

Додаткові підстави для цього отримуємо, якщо в будь-якому часовому інтервалі має місце висока ймовірність появи нульових інтервалів і якщо середня кількість подій в кожному наступному інтервалі мала.

Дані табл. 11.6 задовольняють ці критерії. Якщо й надалі вважатимемо можливий розподіл Пуассона, тоді можна припустити, що λ дорівнює середньому значенню вибіркового середнього і середній дисперсії, тобто

$$\lambda = \frac{0,5147 + 0,6007}{2} = 0,5577.$$

Оцінка за критерієм згоди χ -квадрат. Для статистичної оцінки гіпотези про те, що сукупність експериментальних або вибірових даних незначно відрізняється від тієї, яку можна очікувати за деякого теоретичного розподілу, треба оцінити визначений параметр.

Один з параметрів, що дозволяє оцінювати розбіжність між спостережуваними й очікуваними частотами, – величина χ^2 (χ -квадрат). Цей критерій запропонував у 1903 р. Пірсон.

Статистику χ^2 визначають за виразом

$$\chi^2 = \sum_{i=1}^k \frac{(f_{0i} - f_{ei})^2}{f_{ei}},$$

де f_0 – спостережувана частота для кожного інтервалу; Σ – передбачувана теоретичним розподілом сума розбіжностей за всіма інтервалами; f_e – очікувана частота для кожного інтервалу.

Якщо $\chi^2 = 0$, тоді спостережувані і теоретично передбачені значення частот точно збігаються, якщо $\chi^2 > 0$, тоді повного збігу немає. Чим більша величина χ^2 , тим більша розбіжність між спостережуваними й очікуваними значеннями. Якщо $\chi^2 > 0$, тоді слід порівнювати розрахункові значення з табличними значеннями χ^2 , щоб оцінити, наскільки спостережувані значення визначаються тільки випадковими причинами.

Значення статистики χ^2 табульовані для різних чисел степенів вільності та різних рівнів довірчої імовірності $(1 - \alpha)$.

Для практичного використання цієї статистики висунуто нульову гіпотезу H_0 про те, що між спостережуваним і очікуваним теоретичними розподілами з однаковими параметрами немає значних розбіжностей.

Якщо під час перевірки цієї гіпотези розрахункова величина χ^2 буде мати значення більше від критичного табличного значення (для заданого рівня довірчої імовірності і відповідного числа степенів вільності), тоді можна зробити висновок, що за заданого рівня довірчої імовірності спо-

стерезувані частоти значно відрізняються від очікуваних, і тоді належить відмовитись від гіпотези H_0 .

Застосовуючи метод перевірки гіпотез за критерієм згоди χ^2 , варто пам'ятати таке:

- відносні значення частот (у відсотках) не можна брати, треба користуватись даними прямих спостережень або абсолютними значеннями частот;
- значення спостережуваних частот для кожного інтервалу має дорівнювати 5 і більше, якщо ж ні, тоді суміжні інтервали слід об'єднати;
- число степенів вільності задається виразом $\nu = k - 1 - m$, де k – число інтервалів; m – число параметрів, визначуваних дослідним шляхом або на підставі вибірових даних.

Якщо треба розглянути тільки дискретний розподіл, то варто продовжити ідентифікацію експериментальних даних, наведених у табл. 11.6.

Нехай потрібно перевірити за даними табл. 11.6 їх відповідність розподілу Пуассона з довірчим інтервалом 0,95. Розподіл Пуассона виражається так:

$$P(x = n) = P_x(n) = \frac{\lambda^n e^{-\lambda}}{n!},$$

де $P\{x = n\}$ – імовірність виникнення n подій; $e = 2,71828$; λ – позитивна константа, яка одночасно є і середнім, і дисперсією.

Раніше було обчислено, що для цього ряду $\lambda = 0,5577$, тому гіпотезу H_0 можна сформулювати так: не існує суттєвої різниці між спостережуваними даними і даними, отриманими з розподілу Пуассона за довірчого інтервалу 0,95 та з математичним сподіванням, або середнім $\lambda = 0,5577$. Узявши в формулі розподілу Пуассона це значення за λ і підставивши послідовно $n = 0, 1, 2, 3, 4, 5$, отримаємо дані, які наведено в табл. 11.8.

Таблиця 11.8

| n | $P(n)$ | f_e | f_0 | $\frac{(f_0 - f_e)^2}{f_e}$ |
|-----|--------|-------|-------|-----------------------------|
| 0 | 0,571 | 291 | 315 | 1,98 |
| 1 | 0,319 | 162 | 142 | 2,47 |
| 2 | 0,089 | 45 | 40 | 0,56 |
| 3 | 0,017 | 9 | 9 | |
| 4 | 0,003 | 1 | 2 | 0,09 |
| 5 | 0,001 | 1 | 1 | |
| – | 1,000 | 509 | 509 | 5,10 |

Щоб одержати f_e треба відповідно величину $P(n)$ помножити на 509. Масою розрахункову величину $\chi^2 = 5,10$. Знаходячи критичне значення, величини χ^2 із додатка для довірчого інтервалу 0,95 та числа степенів вільності $4 - 1 - 1 = 2$, дістаємо $\chi^2 = 5,99$. Отже, оскільки розрахункова величина χ^2 менша від критичного значення χ^2 , гіпотези не відхиляємо. Решту груп значень у розрахунку об'єднуємо з тим, щоб отримати значення частоти, в крайньому разі не менше 5 у кожній групі. Тоді замість 6 груп отримуємо 4. Так само, визначаючи число степенів вільності, зменшуємо його на одиницю, оскільки для розрахунку очікуваної частоти використовувалась величина λ , отримана із спостережень.

Оцінка критерієм Колмогорова – Смирнова. Ще один широко використовуваний критерій для статистичної перевірки гіпотез запропонував у 1939 р. професор Смирнов, а потім теоретично довів А. М. Колмогоров. Його використовують у разі, коли перевірні розподіли неперервні та відомі середнє і дисперсія сукупності. Подібно до критерію χ^2 критерій Колмогорова – Смирнова можна використовувати для перевірки відповідності між розподілом сукупності емпіричних даних та деяким конкретним теоретичним розподілом.

Перевірку виконують заданням інтегральної функції, отриманої з теоретичного розподілу, та її порівнянням з інтегральною функцією розподілу емпіричних даних.

Порівняння ґрунтується на вибірковому інтервалі, у якому експериментальний розподіл має найбільше абсолютне відхилення від теоретичного. Далі цю абсолютну різницю порівнюють з критичними значеннями з метою визначення, чи може таке відхилення бути випадковим за законом розподілу.

Для ілюстрації візьмемо дані, які наведено в табл. 11.6 і які використовувались для перевірки критерієм χ^2 на відповідність розподілу Пуассона із середнім значенням 0,5577 і $n = 0,509$. Для цього потрібно отримати два інтегральні розподіли із спостережуваних даних та з теоретичного розподілу і знайти абсолютні різниці для всіх груп значень випадкової величини. Результати обчислення за критерієм Колмогорова – Смирнова зведено в табл. 11.9.

Як бачимо з таблиці, найбільшу абсолютну різницю 0,048 отримано в інтервалі, який відповідає нульовій кількості запитів. Тому цю різницю потрібно порівняти з критичним значенням, наведеним у таблиці критичних чисел Колмогорова – Смирнова. Із цієї таблиці видно, що якщо $n = 509$ і $\alpha = 0,05$, критичне значення

$$D_{кр} = \frac{1,36}{\sqrt{n}} = \frac{1,36}{\sqrt{509}} = \frac{1,36}{22,56} = 0,0603.$$

Таблиця 11.9

| Кількість запитів | Спостережувана частота | Спостережувана ймовірність | Теоретична ймовірність | Інтегральна ймовірність | Інтегральна ймовірність | Абсолютна різниця |
|-------------------|------------------------|----------------------------|------------------------|-------------------------|-------------------------|-------------------|
| 0 | 315 | 0,619 | 0,571 | 0,619 | 0,571 | 0,048 |
| 1 | 142 | 0,279 | 0,319 | 0,898 | 0,890 | 0,008 |
| 2 | 40 | 0,078 | 0,089 | 0,979 | 0,979 | 0,003 |
| 3 | 9 | 0,018 | 0,017 | 0,996 | 0,996 | 0,002 |
| 4 | 2 | 0,004 | 0,003 | 0,999 | 0,999 | 0,001 |
| 5 | 1 | 0,002 | 0,001 | 1,000 | 1,000 | 0,000 |

Оскільки найбільша різниця 0,048 менша від критичного значення, то гіпотезу про те, що експериментальний розподіл пуассонівський, не відхиляємо.

Природно, постає питання, коли слід користуватися критерієм χ^2 , а коли критерієм Колмогорова – Смирнова. За відносно малих обсягів вибірок критерій χ^2 взагалі не застосовують, тоді користуються критерієм Колмогорова – Смирнова. До того ж, коли обсяг вибірки настільки малий, що для користування критерієм χ^2 доводиться об'єднувати сусідні групи (інтервали), потужність цього критерію деякою мірою знижується. Однак, якщо обсяг вибірки великий, найкраще використовувати критерій χ^2 .

Кожний критерій має свої сильні і слабкі сторони, тому щодо вибору можна дати загальні рекомендації.

Критерій χ^2 дуже потужний для великих вибірок ($n \geq 100$).

Деякі автори показують, що отримані за критерієм Колмогорова – Смирнова кращі результати для $n \geq 30$ і що немає достатньої підстави проти застосування цього критерію і для $99 \geq n \geq 10$.

Використовуючи як критерій χ^2 , так і критерій Колмогорова – Смирнова, аналітик має можливість задати кількість груп або інтервалів вибірки.

Правильний вибір цієї кількості має велике значення, оскільки вона визначає ступінь вільності для користування критерієм, і чим їх більше, тим надійніше критерій розрізняє характер розподілу.

Для критерію χ^2 кількість груп частіше визначають за умови, що у кожному групі попало не менше п'яти експериментальних точок. Водночас для користування критерієм Колмогорова – Смирнова дані можна як групувати, так і зараховувати кожне спостереження до окремої групи; остання умова дає можливість ефективного аналізу у разі малих вибірок.

Отже, розглянуто приклад, коли випадкова величина мала дискретний характер.

Якщо вивчати характер впливу завади на корисний сигнал, то передусім треба уміти знаходити розподіл зміни амплітуди та тривалості дії завади; ці параметри мають неперервний характер.

Імовірнісні закони розподілу цих параметрів встановлюють також використанням критеріїв згоди. Механізм ідентифікації такий, як і в дискретному випадку.

Розглянемо приклад, якщо амплітуда сигналу-завади змінюється в межах 40...46 В. Під час спостереження вимірювальна техніка зареєструвала 374 завади; їх значення попадають в інтервали, наведені в табл. 11.10.

Таблиця 11.10

| Номер інтервалу | $x_i - x_{i+1}$ | m_i |
|-----------------|-----------------|----------------|
| 1 | 40...41 | 20 |
| 2 | 41...42 | 112 |
| 3 | 42...43 | 154 |
| 4 | 43...44 | 73 |
| 5 | 44...45 | 13 |
| 6 | 45...46 | 2 |
| | - | $\Sigma = 374$ |

Вихідні дані використаємо для обчислення: $\bar{x} = 42,37$; $S^2 = 0,94$. Рівень значущості $\alpha - 0,01$.

Перевірити гіпотезу про те, що сигнал-завада має, наприклад, нормальний закон розподілу. Потрібно виконати аналогічні дії, як із розподілом Пуассона для вимірних помилок.

12. Телекомунікаційні системи передавання зображень

12.1. Основні принципи побудови систем факсимільного зв'язку

Системи факсимільного зв'язку призначено для передавання заздалегідь підготовлених нерухомих оригіналів. Для передавання зображень оригіналів електричними засобами потрібно реалізувати три фізичні процеси:

- 1) перетворення променевої енергії, відбитої від об'єкта, в електричний сигнал (аналіз зображення);
- 2) передавання електричного сигналу каналом зв'язку;
- 3) перетворення прийнятого електричного сигналу в зображення (синтез зображення).

Надалі електричні сигнали, які відображають властивості оригіналу, будемо називати факсимільними сигналами.

Оригінали відрізняються загальною структурою, кольором, контрастом, мінімальними розмірами тонких ліній і штрихів, точок і проміжків між ними. Характерними прикладами різних структур зображень є рукописний текст і рисунки, машинописний текст, рисунки і графіки, синоптичні карти, фотографії, газетні полоси. Зображення можуть бути чорнобілими та кольоровими.

Розглянемо етапи перетворення факсимільних сигналів:

1. Формування елементарних площинок (растр-елементів на поверхні оригіналу розгортальним елементом передавального апарата). Розмір елементарних площинок рядка розгортки визначає роздільну здатність апарата, а розмір у напрямі, перпендикулярному до рядка розгортки, визначає щільність факсимільного зображення.

2. Розгортка оригіналу розгортальним пристроєм-передавачем виконується переміщенням розгортального елемента по поверхні носія запису. Розгортка виконується по рядках і по кадрах. У плоскісних розгортках розгортка рядками виконується переміщенням розгортального елемента, а розгортка по кадру – поступальним рухом розгортальної поверхні. У барабанних розгортках рух і по рядку, і по кадру здійснюється одночасним обертанням і поступальним переміщенням уздовж осі обертання розгортального барабана із зображенням. Нормовані параметри розгортки – швидкість, напрям, відстань, щільність, модуль і коефіцієнт взаємодії, формат документа. Пристрій розгортки – розгортальний елемент і фотоелектричний перетворювач складають аналізуючий пристрій передавача.

3. Перетворення оптичних щільностей елементарних площинок оригіналу в часову послідовність електричних сигналів за допомогою фотоелектричних перетворювачів. Фотоелектричні перетворювачі факсимільної апаратури мають забезпечувати пропорційну залежність амплітуди вихідного сигналу від яскравості зображення; оброблення та перетворення факсимільного сигналу до вигляду, зручного для передавання по каналах зв'язку.

4. Передавання факсимільного сигналу по каналу (мережі) зв'язку.

5. Приймання факсимільного сигналу.

6. Перетворення сигналу до вигляду, зручного для керування записувальним пристроєм. Перетворення полягає в підсиленні прийнятого сигналу, перетворення його у відеосигнал і в разі потреби – у формуванні визначеним способом фронтів імпульсів для поліпшення спрацювання записувального пристрою.

7. Розгортка факсимільного бланка на приймальному кінці, яка здійснюється синхронно і синфазно розгортці оригіналу на передавальному апараті.

8. Запис факсимільних сигналів на носій запису (на світлочутливі матеріали, барвниками на звичайний папір, на спеціальний папір, який має здатність змінювати колір під дією прийнятого сигналу або утворювати потенційний рельєф під дією світла чи електричного діяння з послідовним проявленням цього рельєфу спеціальними барвниками). Пристрої розгортки і запису являють собою синтезувальний пристрій приймача.

9. Фазування факсимільних апаратів полягає у встановленні однакового положення розгортальних елементів передавального та приймального апаратів відносно початку рядка.

10. Синхронізація факсимільних апаратів полягає у встановленні однакових швидкостей розгортки в передавальному та приймальному апаратах.

11. Керування циклом приймання – передавання полягає в забезпеченні виконання операцій передавання – приймання і процедур взаємодії факсимільних апаратів у їх логічній послідовності.

Кожну з наведених технічних операцій виконує відповідний технічний пристрій. Конструктивно об'єднана сукупність цих пристроїв являє собою факсимільний апарат.

12.2. Принципи передавання зображень

Факсимільний зв'язок – вид документального зв'язку, призначений для передавання не тільки змісту, але й зовнішнього вигляду самого документа. Суть факсимільного методу передавання полягає в тому, що передаване зображення (оригінал) розбивається на окремі елементарні площинки, які скануються зі швидкістю розгортки 60, 90, 120, 180 або 240 рядків/хв. Сигнал яскравості, пропорційний коефіцієнту відбиття таких елементарних площинок, перетворюється у цифрову форму і передається каналом зв'язку з використанням того чи того способу модуляції. На приймальному кінці каналу ці сигнали перетворюються в елементи зображення і відтворюються (записуються) на приймальному бланку.

Факсимільний метод передавання зображення ґрунтується на властивостях відбиття світлового потоку від поверхні, на яку він падає. Відомо, що це явище характеризується відношенням величини відбитого світлового потоку до падаючого світлового потоку і називається коефіцієнтом дифузного відбиття:

$$\rho = \frac{\Phi_{\text{в}}}{\Phi_{\text{п}}},$$

де $\Phi_{\text{п}}$ – падаючий світловий потік; $\Phi_{\text{в}}$ – відбитий світловий потік.

Коефіцієнт відбиття залежить від виду поверхні, на яку падає світло. Для дзеркальних поверхонь він становить 0,9...0,95, для білого паперу 0,7...0,8, а для чорних полів – 0,03. Тобто значення коефіцієнта залежить від ступеня почорніння. Отже, у разі переміщення падаючого світлового потоку вздовж площини, яка має різний ступінь почорніння, відбитий світловий потік буде змінним. Для перетворення світлового потоку в електричний струм використовують фотоелементи, які мають визначену світлову характеристику. Світловою характеристикою фотоелемента називають залежність значення фотоструму від значення світлового потоку, яка змінюється лінійно. Лінійна залежність фотоструму від світлового потоку забезпечує правильне передавання різних відтінків зображення. Відношення фотоструму, вираженого в мікроамперах, до світлового потоку, вираженого в люменах, називають інтегральною чутливістю фотоелемента.

Зображення на вході системи характеризується розподілом яскравості $L(x, y)$ або коефіцієнтом відбиття $\rho(x, y)$. В аналізуючому пристрої виконується електронно-оптичне перетворення і кодування. Робота перетворювача ґрунтується на використанні зовнішнього або внутрішнього фотоелемента: кожному елементу зображення однозначно відповідає деяке значення фотоструму $L(x, y) \rightarrow i_{\text{ф}}(x, y)$ або $\rho(x, y) \rightarrow i_{\text{ф}}(x, y)$. У цьому разі

перетворення і кодування взаємозв'язані: у величині фотоструму від елементів зображення закладено інформацію про розподіл яскравості або коефіцієнта відбиття в оптичному зображенні.

Відомості про стан кожного елемента зображення можна передавати одночасно або послідовно протягом часу дискретизації ΔT . Для послідовного передавання потрібно N каналів зв'язку, що у разі $N \gg 1$ нереально. Тому у факсимільному зв'язку використовується послідовне у часі передавання сигналів. При цьому маємо такі самі N каналів, але з розділенням їх за часом.

Отримання сигналу зображення в разі послідовного передавання ілюструє рис. 12.1. Нехай первинне зображення має вигляд сукупності елементів, що утворюють чотири рядки по п'ять елементів у кожному (рис. 12.1, а). Тут перший стовпчик (а) складається із білих елементів, чорним кольором позначено стилізовану літеру «г», усі інші елементи мають сірий колір.

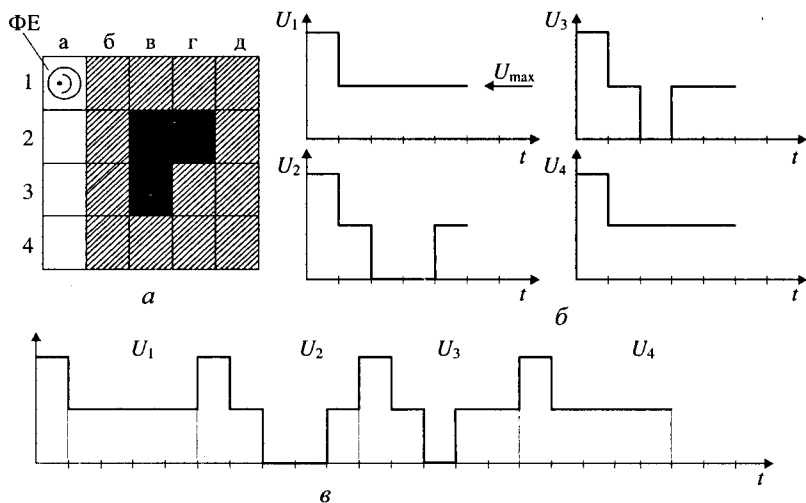


Рис. 12.1. Послідовне у часі отримання зображення

Електронно-оптичне перетворення здійснюється за допомогою фотоелемента (ФЕ), на робочу поверхню якого проєціюється один елемент зображення. Фотострум i_{ϕ} , що виникає у ФЕ, проходить по навантажувальному резистору, утворюючи напругу U . Фотоелемент послідовно (зліва направо) проходить елементи зображення першого рядка. Від білого елемента (1а) отримаємо U_{max} , від сірих – $0,5U_{max}$ (сигнал U_1 першого рядка показано на рис. 12.1, б). Потім ФЕ переміщується вздовж другого рядка.

Якщо ФЕ знаходиться на чорних елементах зображення (рис. 12.1, *в* і *г*), струм $i_{\phi} = 0$ і відповідно $U_{\min} = 0$ (сигнал U_2 на рис. 12.1, *б* і т. д.). Графік сумарної напруги $U(t)$ сигналів рядків усього зображення показано на рис. 12.1, *в*.

Отже, відбувається процес послідовного в часі перетворення яскравості всіх елементів зображення в електричний сигнал. У цьому полягає суть факсимільного аналізу зображення: двовимірну функцію розподілу яскравості (коефіцієнта відбиття) перетворено в одновимірну функцію часу:

$$L(x, y) \rightarrow U(t); \rho(x, y) \rightarrow U(t).$$

Це виявилось можливим за сумісного використання фото ефекту і розгортки. Розгортка – послідовне в часі переміщення розгортального елемента в площині оптичного зображення за визначеним періодичним законом. Розгортальний елемент – це робоча поверхня (вхідний отвір) ФЕ. Під час кодування забезпечується однозначний зв'язок між світловим потоком, який діє на ФЕ, і фотострумом на його виході.

12.3. Структура факсимільних апаратів і системи факсимільного зв'язку

Розглянемо узагальнену схему передавання зображень системою, яка складається з передавальної та приймальної частин, що з'єднуються між собою каналом зв'язку і являють собою узагальнену структурну схему типового факсимільного апарата (ФА) (рис. 12.2).

Розглянемо складові пристрої ФА та системи факсимільного зв'язку (див. рис. 12.1). Розгортальні пристрої (сканери) узгоджують переміщення у вертикальному та горизонтальному напрямках (з різними швидкостями) аналізуючого (синтезуючого) растра-елемента за оригіналом (копією), завдяки чому виконується рядкове зчитування (відновлення) зображення.

Світлооптична система (СОС) складається із джерела світла та системи лінз і дзеркал, що забезпечують засвічення невеликої елементарної ділянки зображення (елемента дозволу, растра-елемента) сфокусованою світловою плямою (зчитувальною апертурою) і попадання відбитого світлового променя на фотоелектронний перетворювач (ФЕП), де й утворюється відеосигнал.

Аналізуючі пристрої. Основну частину будь-якого аналізуючого і синтезуючого пристроїв складають сканери (механічні, електронні та комбіновані). Механічні сканери реалізують традиційні види розгортки: барабанну, дугову, площинну.

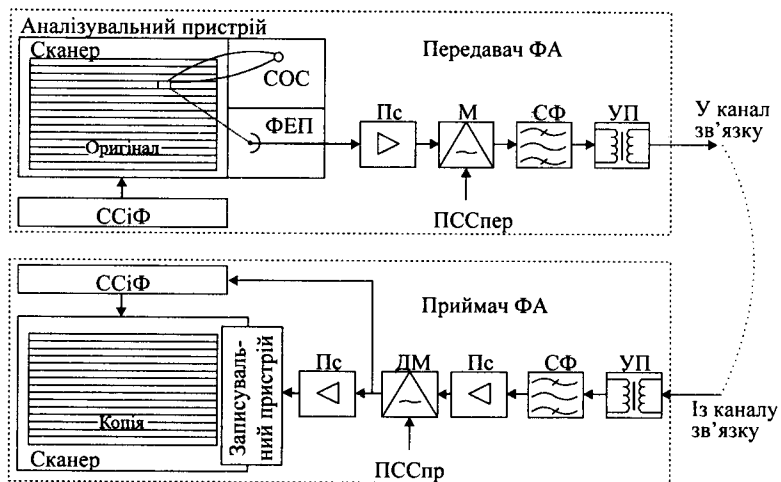


Рис. 12.2. Узагальнювальна структурна схема типового ФА

У разі використання барабанної розгортки зображення закріплюється (зшивкою, пружиною) на зовнішній поверхні швидкообертового барабана, над яким уздовж його осі на каретці переміщується світлооптична голівка. Недостаток такої розгортки – неможливість автоматизувати подачу бланків на барабан та їх знімання.

Під час дугового (колового) розгортання оригінал (копія) розміщується на внутрішній поверхні порожньої циліндричної камери, а сама камера (або голівка) повільно переміщується уздовж циліндра.

Площинне розгортання здійснюється повільним поздовжнім переміщенням (подачею) оригіналу або копії, які лежать на стрічці <транспортера>, з одночасним швидким поперечним перетином (розгортанням) світлового променя зазвичай за допомогою хитного (обертового) дзеркала (призми) СОС.

Розширення функціональних можливостей ФЕП, отримання інформації з потрібною точністю та в потрібному обсязі за мінімальних апаратних і програмних витрат можна досягнути використанням багатоелементних фотоприймачів (БЕФП).

Найбільш технологічно освоєний фотоприймач – фоторезисторний БЕФП (ФР БЕФП), принцип роботи якого полягає у змінненні провідності фоторезисторів (первинного перетворювача) під дією оптичного випромінювання.

Найбільш перспективні фотодіоди БЕФП (ФР БЕФП), які мають високу швидкодію, добре розроблену технологію і працюють на низьких рівнях освітленості.

Останніми роками застосовують прилади із зарядовим зв'язком (ПЗЗ), принцип роботи яких полягає у перетворенні розподілення освітленості зображення, у розподіленні неосновних носіїв заряду, зберіганні їх у потенціальних ямах і передаванні зарядів до вихідного підсилювача. Залежно від фізичної реалізації розрізняють два типи ПЗЗ: з поверхневим перенесенням заряду та із захованим (об'ємним) каналом. Переваги ПЗЗ із захованим каналом полягають у збільшенні швидкодії та ефективності перенесення заряду, у зменшенні рівня шумів.

Застосування БЕФП у фотоелектричних перетворювачах визначає ряд основних вимог: роботу в широкому спектральному діапазоні, високі інтегральну чутливість і роздільну здатність, велику кількість елементів, високу швидкість, відтвореність параметрів від елемента до елемента. Цьому переліку вимог найбільше відповідають ПЗЗ.

Синтезувальні пристрої. Синтезувальний пристрій ФА (див. рис. 12.2) складається із сканера, записувального пристрою та системи синхронізації і фазування (ССіФ).

Основна функція синтезувального пристрою полягає в об'єднанні процесів декодування та електронно-оптичного перетворення. У факсимільному зв'язку зображення синтезують різними методами: на фотоматеріал – модульованим за яскравістю джерелом світла, на термочутливий електрохімічний матеріал – дією змінного за часом електричного струму і т. ін. Нанесення зображення на копію здійснюється локальним змінням у невеликому околі синтезованого растра-елемента фізичних або хімічних характеристик навколишнього середовища або самого носія зображення: освітленості носія, його температури, тиску на носій і т. ін. При цьому для створення записувального елемента використовують світлооптичну голівку, голку, штифт, кульку, спіраль, форсунку та інші пристрої.

Фотографічний запис здійснюється світловим променем на фотоплівку або фотопapір у темній камері, застосовується рідко і тільки у ФА, які забезпечують дуже високу якість копії. Електростатичний запис перспективний, оскільки забезпечує досить високу роздільну здатність (до 8 ліній/мм), високу швидкість запису (до 6000 рядків/хв) і дозволяє проявляти зображення на світлі.

Перетворювач спектра сигналу (ПСС) передавача ФА складається із підсилювача (Пс), модулятора (М), смугового фільтра (СФ), узгоджувального пристрою (УП) – симетризованого трансформатора – формують модульований сигнал з потрібними амплітудою та смугою частот і узгоджують вихід ФА із входом каналу зв'язку (рис. 12.2).

Перетворювач спектра сигналу (ПССпр) приймача ФА (УП, СФ, Пс і демодулятора (ДМ)) відновлюють факсимільний відеосигнал, який керує яскравістю растрів-елементів, що наносяться один за одним на копію за-

писувальним пристроєм. Перетворювач спектра сигналу передавача і приймача разом утворюють модем.

Узгоджена робота аналізуючого і синтезуючого пристроїв на передавальному та приймальному ФА забезпечується ССіФ.

12.4. Модеми сучасних факсимільних апаратів

Сучасні ФА виконують цифрове оброблення відеосигналу і цифрове передавання каналом зв'язку.

Переваги цифрових ФА:

- можливість передавання зображень універсальними каналами єдиної інтегральної цифрової мережі без розширення займаної смуги частот;
- більш ефективне використання каналів зв'язку завдяки скороченню надлишковості відеосигналу добре розвиненими цифровими методами;
- більш висока завадостійкість;
- поліпшення технічних характеристик ФА (габаритів, мас, енергоємності, надійності);
- можливість широкої автоматизації усіх технологічних операцій шляхом використання мікропроцесорної техніки.

Схему модема цифрового ФА показано на рис. 12.3. У пристрої квантування за часом дворівневий сигнал, який надходить з виходу електронного сканера, дискретизується за часом і перетворюється у цифрову форму. У кодері-компресорі надлишковий цифровий відеосигнал скорочується згідно з одним із рекомендованих МККТТ алгоритмом (наприклад, модифікований код Хаффмена); рідше реалізується адаптивне кодування з передбаченням, планарне і адресно-позиційне кодування. У декодер-декомпресор входять блок логічної обробки на інтегральних логічних елементах і буферний запам'ятовувальний пристрій великої ємності. Блоки скремблер і дескремблер – стандартні елементи цифрових систем зв'язку. Двійковий сигнальний код на вході модема може мати довільну статистичну структуру, яка не завжди задовольняє вимоги синхронного способу передавання. Основними такими вимогами є:

- частота змінювання символів (1, 0) має забезпечити надійне виділення тактової частоти безпосередньо із приймального сигналу;
- спектральна щільність потужності передавального сигналу має бути, по можливості, постійною і зосередженою в заданому діапазоні частот з метою зниження взаємного впливу каналів.

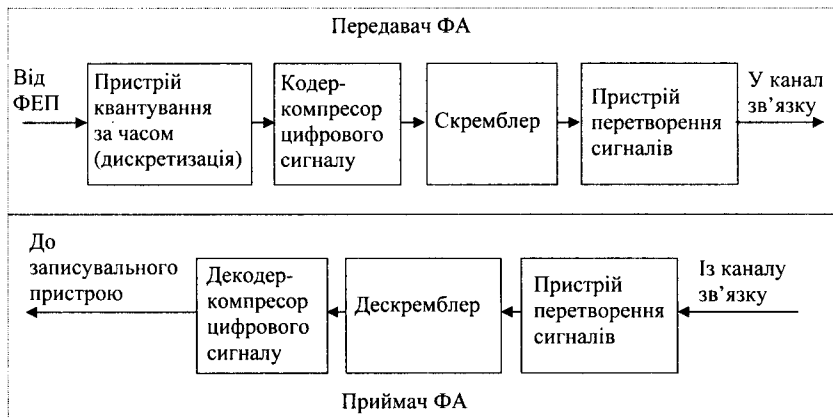


Рис. 12.3. Структурна схема модема цифрового ФА

Зведені вимоги потрібно виконувати незалежно від структури переданого повідомлення. Тому в синхронних модемах початкова послідовність двійкових послілок часто підлягає обробленню. Зміст такого оброблення полягає в отриманні послідовності, у якій статистика виникнення нулів і одиниць наближається до випадкової, що дозволяє задовольнити дві зазначені вимоги.

Одним із способів такої обробки є скремблювання (scramble – перемішування). Скремблювання – це обернене перетворення структури цифрового потоку без зміни швидкості передавання з метою отримання властивості випадкової послідовності. Скремблювання виконується на передавальному кінці системи за допомогою скремблера, який реалізує логічну операцію підсумовування за модулем два початкового і псевдовипадкового двійкових послідовностей сигналів. На приймальному кінці здійснюється обернене перетворення – дескремблювання, виконуване дескремблером.

Основна частина скремблера – генератор псевдовипадкової послідовності у вигляді лінійного v -каскадного регістра із зворотними зв'язками, який формує послідовність максимальної довжини $2^v - 1$. Розрізняють два основні типи скремблерів-дескремблерів – самосинхронізувальні і з початковим установленням (адитивні).

Схему самосинхронізувальної пари «скремблер – дескремблер» показано на рис. 12.4. Особливістю самосинхронізувального скремблера є те, що він керується самою скремблювальною послідовністю, тією, що надходить у канал. Тому в цьому разі не потрібне спеціальне установлення станів скремблера і дескремблера, оскільки вони виявляються ідентичними в результаті запису в їхні регістри зсуву скремблювальної послідовності.

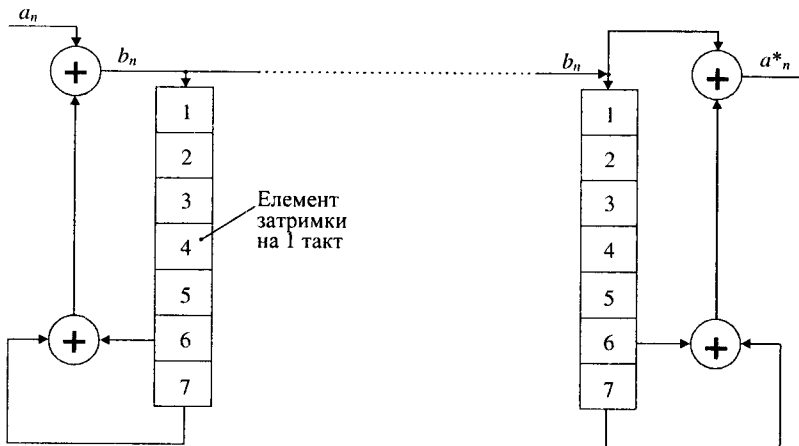


Рис. 12.4. Схема скремблювання із самосинхронізацією

У разі втрати синхронізму між скремблером і дескремблером час його відновлення не перевищує кількості тактів, що дорівнює кількості комірок регістра-скремблера. На приймальному кінці виділення інформаційної послідовності здійснюється складанням за модулем два прийнятої скремблювальної послідовності з псевдовипадковою послідовністю регістра. Наприклад, для схеми, зображеної на рис. 12.4, вхідна послідовність a_n за допомогою скремблера згідно з виразом $b_n = a_n \oplus (b_{n-6} \oplus b_{n-7})$ перетворюється у двійкову послідовність b_n , відправлювану в канал. У приймачі з цієї послідовності таким самим регістром зсуву, як і в передачі, формується послідовність $a_n = b_n \oplus (b_{n-6} \oplus b_{n-7})$, ідентична послідовності a_n . Це легко перевіряється перетворенням першого виразу до вигляду $a_n = b_n \oplus (b_{n-6} \oplus b_{n-7})$ і порівнянням отриманого виразу з попереднім.

Один із недоліків самосинхронізувальних скремблерів – дескремблерів – притаманна їм властивість розмноження помилок. Так, для схеми на рис. 12.4 у разі однієї помилки у послідовності b_n помилковими виявляються також символи 6 і 7. У загальному випадку помилково прийнятий біт буде виявлятися α разів, де α – кількість зворотних зв'язків. Цей недолік обмежує кількість зворотних зв'язків у регістрі зсуву, яке майже не перевищує $\alpha = 2$, тобто поліном регістра є триномом вигляду $x^y + x^u + 1$. Другий недолік самосинхронізувальних скремблерів – можливість появи на його вході «критичних ситуацій», коли вихідна послідовність набуває періодичного характеру з періодом, меншим від довжини ПВП. Для запобігання таким ситуаціям у скремблері і дескремблері згідно з рекоменда-

ціями ITU-T передбачаються спеціальні додаткові схеми контролю, які виявляють періодичність елементів на вході і порушують її.

Недоліків, властивих самосинхронізувальному скремблеру – дескремблеру, майже не буває у разі адитивного скремблювання. Однак для цього потрібно попереднє ідентичне установлення станів реєстрів скремблера і дескремблера.

Система синхронізації і фазування (див. рис. 12.2). Такі системи забезпечують узгоджену роботу розгортальних пристроїв передавача і приймача (обов'язково за однакових модулів або значень коефіцієнтів взаємодії) за рахунок однакових швидкостей розгортання та однакового початкового положення зчитувальної апертури на початку роботи ФА. Перше досягається автономною, мережевою або примусовою синхронізацією; друге – за допомогою автоматичного, напівавтоматичного або ручного фазування. У разі автоматичної синхронізації розгортувальні пристрої передавача і приймача живляться незалежно – кожний від свого високостабільного джерела змінного струму. У разі мережевої синхронізації як джерело живлення використовують мережу змінного струму. Фазування провадять один раз перед сеансом. Для узгодження моментів початку сканування оригіналу і копії на приймальному і передавальному кінцях системи примусово встановлюють різні швидкості розгортання. Обидва ФА на початку (кінці) кожного рядка виробляють спеціальні сигнали фазування (імпульс або кодову комбінацію), і в разі точного збігу цих сигналів приймальний ФА автоматично переводиться в синхронний режим роботи.

У новітніх моделях ФА роботою автоматичного пристрою керує мікропроцесор.

12.5. Кодування передаваної інформації і скорочення надмірності

12.5.1. Факсимільний сигнал

Факсимільний сигнал визначають за властивостями оригіналу, зображення якого передається, і параметрами аналізуючого пристрою. Оригінали розрізняються між собою загальною структурою, кольором, значенням оптичних щільностей у найбільш світлих і темних місцях, контрастом, мінімальними розмірами тонких ліній і штрихів, крапок і проміжків між ними.

Характерними прикладами різних структур зображень виявляються рукописний текст і рисунки, машинописний текст, креслення і графіки, синоптичні карти, фотографії, газетні полоси. Зображення можуть бути чорно-білими та кольоровими.

Для факсимільного передавання зображень використовують оригінали, ширина яких 220 мм і довжина 50, 75, 100 і 300 мм, а стандартний формат – 210 × 297 мм.

Елементи оригіналу мають різні значення коефіцієнта відбиття падаючого світлового потоку: від $\rho_{\max} = 0,9 \dots 0,8$ для білого паперу до $\rho_{\max} = 0,03 \dots 0,05$ для чорної фарби. Зв'язок між коефіцієнтом відбиття і оптичною щільністю D , яка характеризує ступінь почорніння окремих ділянок зображення, визначають за співвідношенням $D = \lg \frac{1}{\rho}$. Зазвичай для ори-

гіналів максимальна оптична щільність $D_{\max} = 1,6 \dots 1,7$, якщо коефіцієнт відбиття падаючого світлового потоку ρ_{\min} , мінімальна $D_{\min} = 0,03 \dots 0,04$, якщо ρ_{\max} . Інтервал оптичних щільностей $\Delta D = D_{\max} \dots D_{\min}$ становить 1,6...1,7. Оскільки яскравість освітлюваних ділянок пропорційна коефіцієнту відбиття, тоді контраст зображення оригіналу

$$K_{\text{зоб}} = \frac{L_{\max}}{L_{\min}} = \frac{\rho_{\max}}{\rho_{\min}}$$

може досягати 30. У реальних умовах величина контрасту значно нижча.

Рукописний і машинописний тексти, рисунки, креслення мають мінімальну товщину ліній (штрихів) 0,1 мм, у зображенні газетних полос мінімальні розміри деталей становлять 0,07 мм і проміжків між ними 0,15 мм.

Перед передачею півтонових ілюстрацій виконують їх растрівання, тобто їх відтворення у вигляді сукупності різних за розміром чорних крапок на світлому фоні або білих точок на темному фоні (рис. 12.5) з різними оптичними щільностями. При цьому інтегральна оптична щільність зображення залежить від співвідношення площинок пропусків і друкованих елементів. Крапки знаходяться на лініях растра, які розміщуються під кутом 45° . Щільність цих ліній (лініатура) зазвичай становить 25 або 36...40 ліній/см, відповідно відстань між лініями растра 0,25...0,4 мм. Приклади растрованих зображень показано на рис. 12.5.

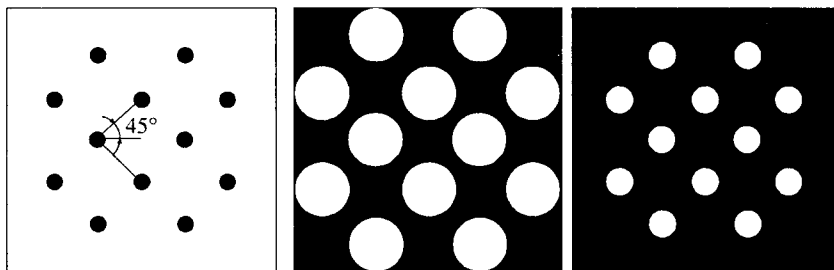


Рис. 12.5. Растрівання зображень різної оптичної щільності (у збільшеному масштабі)

Розглянемо більш детально структуру зображення газетних полос. Газетні полоси вміщують текстовий матеріал і заголовки, які набирають різними шрифтами, ілюстрації і оформлювальні елементи. Шрифти розрізняються конфігурацією (гарнітурою) і розміром (кеглем). Розмір шрифту вимірюють у пунктах (1 пункт дорівнює 0,353 мм). Номер кегля (кг) відповідає висоті літери (знака, букви). Широко використовуються кеглі розміром 6, 8, 9 і 10 пунктів (відповідно 2,2; 2,93; 3,3; 3,67 мм).

Зображення газетної полоси анізотропне з чітко вираженою періодичністю в напрямі, перпендикулярному до рядків тексту. Періодичність визначається тільки розміром кегля і не залежить від особливостей шрифту. Фотографії також мають періодичну структуру; заголовки у зображенні газетної полоси не мають вираженої періодичності.

На виході аналізуючого пристрою в загальному випадку отримується аналоговий багаторівневий початковий сигнал, у якому закладено інформацію про розподіл коефіцієнта відбиття $\rho(x, y)$ у площині оригіналу (рис. 12.6). Поточне значення сигналу в процесі розгортання пропорційне середньому значенню ρ у межах апертури. Після попереднього підсилення початковий сигнал u_1 надходить на формувач. Досягнувши визначеного (установленого або автоматично підтримуваного) рівня напруги $u_1 = U_{\text{сер}}$, формувальний пристрій спрацьовує і на його виході підтримується постійне одиничне значення напруги протягом проміжку часу, поки $u_1 > U_{\text{сер}}$.

Коли u_1 стає меншим від $U_{\text{сер}}$, схема повертається у початковий стан і вихідна напруга формувача набуває нульового значення. У результаті утворюється послідовність імпульсів, яку показано на рис. 12.6, б. Це дозволяє звільнитися від нерівномірності сигналу фону, не задрукованого фарбою (неоднорідність паперової підкладки) на світлих ділянках оригіналу, а також від нерівномірності продрукування тексту, растрованих елементів, ліній, штрихів. Білим (світлим) ділянкам на зображенні відповідає нульове значення сигналу «0», чорним (темним) – значення «1».

Пояснимо характеристику «передавання півтонів». Квантування сигналу за рівнем для передачі двоградаційних (графікових) зображень здійснюється просто, оскільки сигнал принципово дворівневий («чорне» і «біле»). Для цього використовують простий пороговий формувальний пристрій (типу тригера Шмітта). Якщо треба передати півтонове зображення цифровим способом, амплітуду сигналу квантують на 16, 32 або 64 фіксовані рівні яскравості.

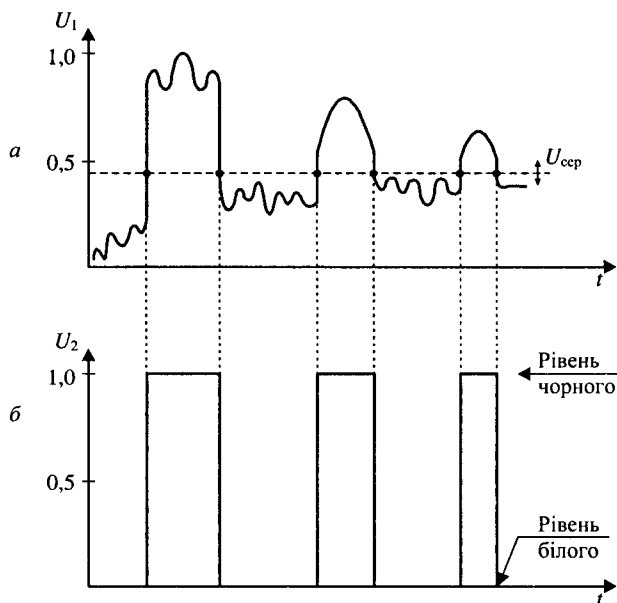


Рис. 12.6. Початковий аналоговий (а) і дискретизований (б) факсимільні сигнали

Швидкість передавання в сучасних ФА має декілька конкретних градацій. Перед передаванням повідомлення ФА за командами мікропроцесора відправляє в лінію (у приймач переприймального пристрою на вузлі комутації) спеціальний еталонний сигнал з максимальною швидкістю передавання 9600 біт/с. На приймальному кінці отриманий сигнал порівнюється з еталонном. Не отримавши підтвердження про достатню якість приймання, факс-відправник знижує швидкість до 7200 біт/с і повторює спробу. Може виявитися, що внаслідок впливу на канал завад швидкість автоматично буде знижена до 4800 і навіть до 2400 біт/с, і лише після підтвердження необхідної якості приймання еталонного сигналу починається передавання повідомлення.

Спосіб кодування у разі стиснення зображень. Уявімо, наприклад, сторінку машинописного тексту. Під час зчитування машинописного тексту, навіть при дозволі ліній на міліметр (лін/мм), багато рядків факса будуть зчитувати міжряддя тексту і не будуть нести ніякої інформації. Тому факсимільне повідомлення має велику надмірність, яку не важко оцінити так. Припустімо, що на сторінці 1800 знаків машинописного тексту (30 рядків і 60 знаків у рядку). Після її кодування, наприклад восьмизначним кодом, дискретне повідомлення буде становити приблизно 14000 біт. Для відправлення такого повідомлення зі швидкістю передавання 2400 біт/с

потрібно близько 5 с. У разі дискретизації для факсимільного передавання з роздільною здатністю по рядках 8 ліній/мм і по вертикалі 7,7 ліній/мм дискретне двоградаційне повідомлення буде вмещувати $1728 \text{ (елементів)} \times 2287 \text{ рядків} = 951936 \text{ біт}$, тобто в 280–300 разів більше, ніж кодування кодом телетекса. Час його передавання тим самим каналом (2400 біт/с) становить близько 27 хв. Якщо роздільна здатність по вертикалі 3,85 ліній/мм, час передавання дорівнює 13 хв.

Для тексту надмірність факсимільного повідомлення завжди більша за 100. З метою скорочення надмірності дискретний факсимільний сигнал за допомогою кодування стискають, у результаті чого надмірність дискретного (двоградаційного) факсимільного повідомлення зменшується в 5–15 разів і відповідно у стільки ж разів скорочується час його передавання каналом зв'язку.

Для стиснення зображень використовують модифікований код Хаффмена (МКХ) або метод порядкового кодування довжини сегментів (КДС). Є і більш складні коди (READ).

12.5.2. Передавання інформації і скорочення надмірності

Для передавання зображення достатньо укласти оригінал у плоский отвір на корпусі апарата, установити режим «Передача», набрати телефонний номер адресата-абонента і натиснути кнопку «Старт». Для приймання достатньо ввімкнути автомат-апарат, який в будь-який час самостійно прийме повідомлення. Для копіювання у плоский отвір укладають оригінал, натискають клавішу «Копія» і через декілька секунд отримують копію оригіналу.

Термінал забезпечує виконання таких функцій:

- напівавтоматичне і автоматичне передавання повідомлень;
- автоматичне приймання і реєстрація повідомлень;
- автоматичне установлення з'єднань, ідентифікація правильності установлення з'єднань;
- автоматичне проставляння відтиску штампа на оригіналі і копії документа;
- автоматична реєстрація службової інформації на контрольній стрічці (операційний журнал);
- накопичення добового обсягу повідомлень;
- автоматичне сортування повідомлень за напрямками (для систем з накопиченням повідомлень).

Термінал, що реалізує зазначені функції, складається з таких блоків:

- первинного кодера (аналізувального пристрою) і декодера (синтезувального пристрою);
- блоків стиснення і накопичення інформації;

- блоку сполучення з мережею зв'язку (повідомлення правильності) і перетворення сигналів;
- блоків напівавтоматичного і автоматичного установаження з'єднань, керування та індикації та живлення.

Сучасні термінали будують у вигляді уніфікованих блоків, об'єднання яких на основі програмного керування дозволяє формувати апаратуру різних груп. Модульний принцип побудови не тільки розширює можливості впровадження терміналів, але й підвищує їх надійність, створює передумову зниження вартості за рахунок збільшення серійності виробництва.

Факсимільний сигнал на виході кодера джерела являє собою послідовність випадкових за довжиною серій імпульсів, що відповідають послідовностям растрових елементів чорного та білого кольорів. Для зменшення наявної у конкретній послідовності інформаційної надмірності використовують два методи: збільшення і передбачення. Метод збільшення полягає в тому, що повідомлення розбивають не на елементи, а на відрізки, кожний з яких вміщує по декілька елементів, а практична його реалізація – у передаванні не самих серій імпульсів, а тільки інформації про їх довжину або про положення елементів, на яких змінюється колір.

Метод передбачення ґрунтується на тому, що характер імовірнісних взаємозв'язків між частинами повідомлення в ряді випадків дозволяє, прийнявши деяку частину повідомлення передбачити його відсутню частину. Відповідно до рекомендацій МККТТ.4, урахувавши неоднакову ймовірність різних довжин серій, для кодування факсимільних сигналів використовують модифікований код Хаффмена. Принцип побудови коду Хаффмена полягає в тому, що частим повідомленням відповідають короткі кодові комбінації, тоді як рідкісним – довгі. Такий нерівномірний код виявляється більш економним, ніж рівномірні коди завдяки меншій кількості елементарних сигналів, що припадають на один символ.

Крім одновимірного кодування, допускається використання двовимірного – модифікованого коду вибору відносної адреси елемента. У цьому коді позиція кожного змінного елемента рядка кодується з урахуванням або наперед зміненого елемента, або позиції відповідного змінюваного елемента опорного рядка, який розміщується безпосередньо над кодованим. Кожний рядок після кодування використовується як опорний для наступного кодованого рядка. Завдяки частковому урахуванню кореляції растрів-елементів і у вертикальному напрямку двовимірний код дозволяє отримати великі значення коефіцієнтів стиснення порівняно з одновимірним.

Методи ефективного кодування дозволяють значно стискати факсимільну інформацію. Але одночасно зі зменшенням надмірності знижується завадостійкість факсимільного методу передавання, і це зниження тим сильніше, чим більше значення коефіцієнта стиснення коду.

Контрольні питання

1. Основні поняття та визначення передавання дискретної інформації.
2. Перетворення повідомлень у сигнал. Поняття кодування – декодування, модуляції – демодуляції.
3. Кодування джерелом інформації.
4. Канал передавання дискретної інформації.
5. Функції та призначення складових пристроїв дискретної інформації.
6. Вимоги до СПДІ.
7. Математичний опис алгоритму роботи СПДІ.
8. Основні відомості про сигнали. Параметри сигналів.
9. Модуляція сигналів. Загальна ідея керування параметрами сигналів.
10. Амплітудно-змодульований сигнал.
11. Частотно змодульований сигнал.
12. Фазомодульований сигнал, абсолютна та відносна модуляція.
13. Принцип багатократною модуляції.
14. Кодування повідомлень. Рівномірні та нерівномірні коди. Коди Бодо та Морзе.
15. Принцип побудови пристроїв перетворення сигналів у СПДІ.
16. Методи реєстрації дискретних сигналів.
17. Реєстрація сигналів методом стробування.
18. Аналоговий інтегральний метод реєстрації сигналів.
19. Дискретний інтегральний метод реєстрації сигналів.
20. Завади та спотворення. Джерела виникнення завад.
21. Основні характеристики каналів зв'язку та їх вплив на передавання даних.
22. Характеристичні крайові спотворення. Часова діаграма сигналів типу «точка – точка» та групи сигналів однієї полярності.
23. Поняття та визначення синхронізації та фазування. Обґрунтування потреби в синхронізації та фазуванні.
24. Визначення поелементної синхронізації. На часовій діаграмі показати початок порушення синхронізму і сформулювати завдання визначення параметрів підтримки синхронізму.
25. Знайти смісно-часові параметри розсинхронізації.
26. Визначити допустимий час синхронного стану та обсяг переданої за цей час інформації.
27. Способи підтримки синхронізму.
28. Синхронізація сигналів і повідомлень. Основні визначення. Фазові співвідношення.
29. Умови порушення синхронізації. Показати на діаграмі роботи дискретного каналу зсув за фазою на одному періоді реєстрації.
30. Вимоги до ПС.
31. Замкнені пристрої синхронізації. Структурна схема.
32. Фазовий дискримінатор. Схема і часова діаграма.
33. Формування сигналів випередження та відставання.
34. Побудувати реверсивний лічильник – інерційний елемент ПС.
35. Пристрій керування синхронізатором.
36. Призначення та принцип дії проміжного перетворювача.
37. Системний аналіз окремих процесів перетворення сигналів і їх взаємодія в єдиній СПДІ.

38. Мета та завдання статистичного дослідження помилок, що виникають у трактах передавання даних.
39. Мета ідентифікації емпіричних даних вимірних помилок.
40. Призначення критеріїв згоди для статистичного оброблення результатів вимірювання.
41. Правила користування критерієм χ^2 -Пірсона. Степені вільності. Довірчий інтервал.
42. Правила користування критерієм Колмогорова – Смирнова. Критичні числа.
43. Рекомендації до впровадження критерію χ^2 -Пірсона і критерію Колмогорова – Смирнова.
44. Загальний алгоритм (механізм) статистичного оброблення помилок, які виникають у результаті впливу завад на тракти обмінювання даними.
45. Системний аналіз спотворень і помилок.
46. Використання результатів статистичного оброблення помилок.
47. Принципи побудови системи факсимільного зв'язку.
48. Методи передавання нерухомих зображень.
49. Узагальнювальна структурна схема ФА.
50. Аналізуювальні пристрої ФА.
51. Синтезуювальні пристрої ФА.
52. Структурна схема модема факсимільного апарата. Призначення та функції складових частин.
53. Факсимільний сигнал.
54. Передавання факсимільної інформації і методи скорочення надмірності.

Системний аналіз, надійність, завадостійкість телекомунікаційної системи під час керування обміном інформацією

13. Загальні принципи керування дискретним (інформаційним) каналом

13.1. Багаторівневий підхід. Протокол. Інтерфейс. Стек протоколів

Універсальна теза – користь стандартизації, справедлива для всіх галузей, а для комп'ютерних мереж має особливе значення. Суть мережі – це об'єднання різного обладнання. Це означає, що проблема сумісності – одна з найбільш гострих. Без прийняття усіма виробниками загальноприйнятих правил побудови обладнання прогрес у створенні мереж був би неможливий. Тому весь розвиток комп'ютерної галузі відображено в стандартах. Будь-яка технологія тільки тоді має «законний» статус, коли її зміст закріплюється у відповідному стандарті.

Ідеологічна основа стандартизації комп'ютерних мереж – це багаторівневий підхід до розроблення засобів мережевої взаємодії. Саме на основі цього підходу було розроблено стандартну семирівневу модель взаємодії відкритих систем, яка стала своєрідною універсальною мовою мережевих фахівців.

Розглянемо взаємодію двох процесів A і C , кожний з яких реалізується відповідною комп'ютерною програмою. Нехай програма користувача A – програма статистичної обробки даних вимірювання помилок, а програма C – програма керування базою даних (для цього прикладу подано масиви даних критичних чисел Пірсона та Колмогорова – Смирнова).

Схему взаємодії двох безпосередньо зв'язаних процесів показано на рис. 13.1.



Рис. 13.1. Взаємодія двох процесів

У разі потреби в передаванні інформації програма користувача А видає масив інформації. Програми керування поданням і сеансом додають до цього масиву заголовок і кінцевик, утворюючи таким чином повідомлення. У заголовок останнього записується інформація, яка дозволяє виокремити повідомлення та зазначити потрібні адреси. У кінцевіку розміщуються символи, які дозволяють виявити помилки в повідомленні після його передавання другому процесу.

Повідомлення має довільну довжину. Тому його перед передаванням ділять на частини – блоки, максимальний розмір кожного з яких обмежений визначеним значенням. Ці блоки передаються (див. рис. 13.1) процесу Є. Тут програми керування поданням і сеансом збирають із цих блоків повідомлення, яке аналізують, а потім відкидають його заголовок і кінцевик. Отриманий масив інформації передається програмі користувача Є.

Така безпосередня взаємодія процесів можлива тільки в тих рідкісних випадках, коли ці процеси розміщені в одному і тому ж комп'ютері. Зазвичай у комп'ютерній мережі процеси знаходяться в різних комп'ютерах і взаємодіють через апаратуру передавання даних, фізичні канали, комунікаційні, а часто й інтерфейсні комплекси. У цьому разі до кожного процесу (рис. 13.2) має бути добавлений новий елемент програмної структури комп'ютерної мережі – програма керування передаванням інформації від одного процесу до іншого. Точку, розміщену між процесом і програмою керування передаванням, називають портом.

У розглядуваному випадку структура передавального масиву інформації ускладнюється; його називають фрагментом, який має два заголовки (рис. 13.2): процесу і передавання. Заголовок передавання вміщує керувальні дані, які описують процес передавання: тип масиву інформації, який передається від одного процесу до другого, адресу початкового і кінцевого процесів (точніше адресу портів), ідентифікатор (ім'я) цього масиву.

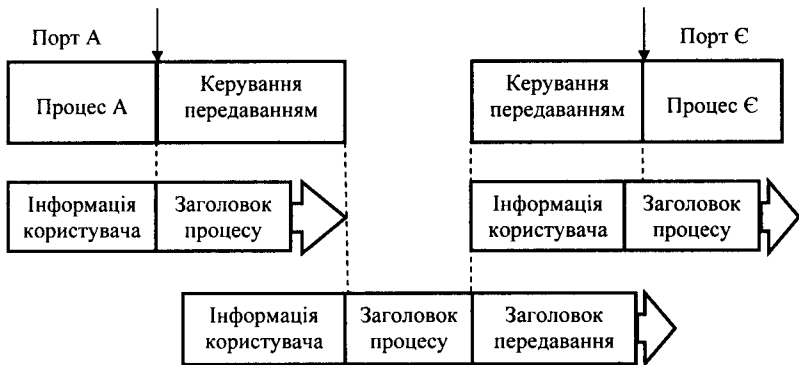


Рис. 13.2. Зв'язок процесів, розміщених у двох комп'ютерах

У разі потреби в передаванні процес А (рис. 13.2) видає блок інформації, який передається програмі керування передаванням. Остання додає до нього заголовок передавання і отриманий фрагмент інформації спрямовує другій програмі керування передаванням. Ця програма відкидає у фрагментів заголовок передавання, читає його, а частину масиву інформації, який залишився, передає процесу Є. Тут із блоків збирається повідомлення з інформацією користувача А і передається програмі користувача Є. Аналогічно передається інформація від процесу Є до процесу А.

У комп'ютерній мережі один і той самий процес може бути пов'язаний з кількома процесами, виконуваними різними комп'ютерами, розміщеними в різних місцях. Тому поряд з програмою керування передаванням вводиться (рис. 13.3) ще один елемент – програма керування мережею. Лінії, які з'єднують у комп'ютерних мережах програми керування мережею, називають транспортними каналами. Основне завдання програми керування мережею – організація маршрутизації пакетів у комп'ютерній мережі.

Згідно з новою програмою до фрагменту додається заголовок пакета. У ньому розміщується вся керувальна інформація, потрібна для маршрутизації пакетів.

Створений таким чином пакет передається програмі (рис. 13.4), зв'язаній з процесом Є.

Тут програми керування мережею і передаванням, використовуючи відповідні заголовки пакета і фрагмента, передають процесу Є прийнятий блок.

Зазначимо, що в деяких комп'ютерних мережах програми керування передаванням і мережею об'єднуються в одну комплексну програму.

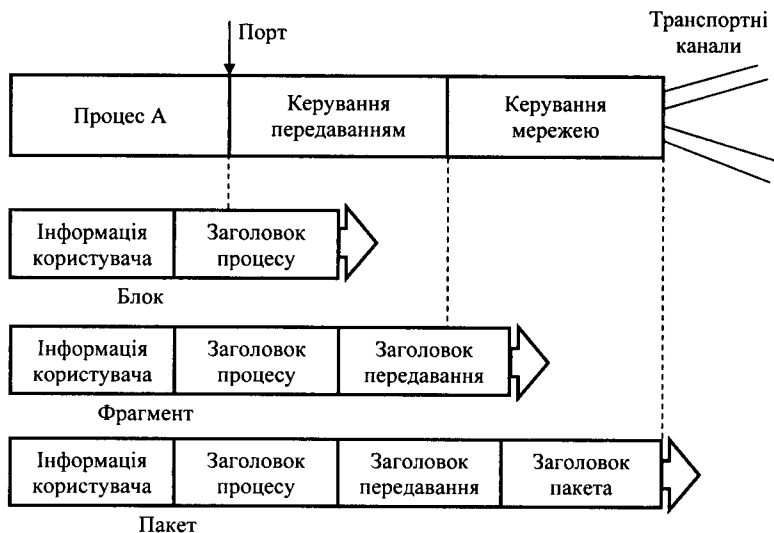


Рис. 13.3. Процедура утворення пакета

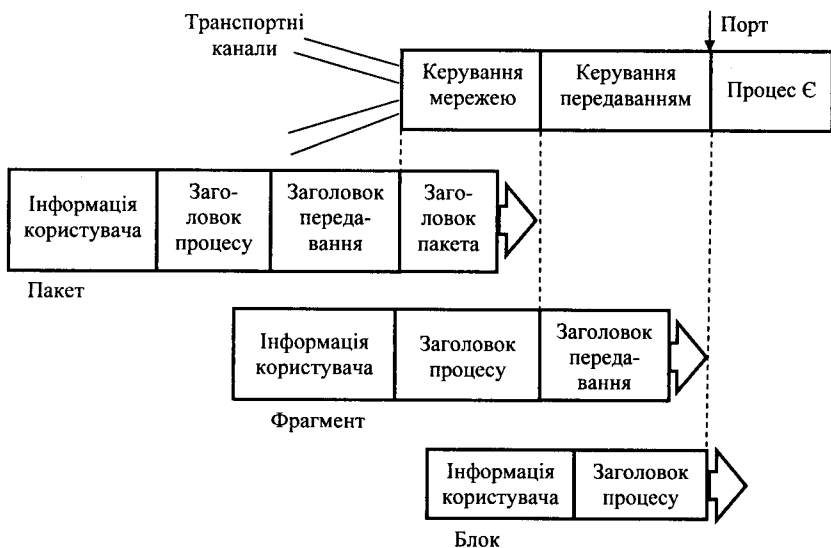


Рис. 13.4. Процедура приймання пакета

Між будь-якою парою суміжних комп'ютерів розміщено канал, керування роботою якого виконує ще один елемент – програма керування інформаційним каналом. Відповідно до цього такий канал називають інформаційним.

Програма керування інформаційним каналом перед передаванням пакета по інформаційному каналу оброблює його (рис. 13.5) спереду заголовком каналу і ззаду кінцевиком каналу. Заголовок каналу містить інформацію, потрібну для керування інформаційним каналом. У кінцевик каналу включається інформація, завдяки якій забезпечується перевірка пакета після передавання його інформаційним каналом. У результаті додавання до пакета заголовка і кінцевика каналу створюється масив інформації, який називають кадром. Кадри в інформаційному каналі йдуть неперервним потоком один за одним. Вони відділяються один від одного керувальними символами, які називають прапорцями.

Отже, виникає шарувата програмна структура комп'ютерної мережі, яка складається з елементів, розміщених на семи рівнях.

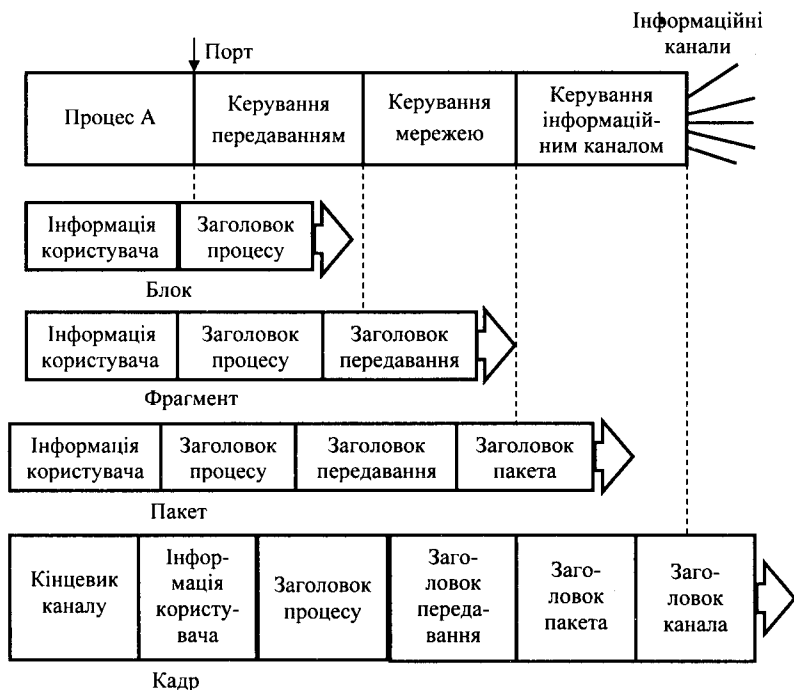


Рис. 13.5. Процедура утворення кадру

Багаторівневе подання засобів мережевої взаємодії має свою специфіку, пов'язану з тим, що в процесі обміну повідомленнями беруть участь два комп'ютери і в одному випадку потрібно організувати узгоджену роботу двох «ієрархій». Передаючи повідомлення, обидва учасники мережевого обміну повинні прийняти множину погоджень. Наприклад, вони мають узгодити рівні та форму електричних сигналів, спосіб визначення довжини повідомлення, домовитись про методи контролю достовірності і т. ін. Інакше кажучи, узгодження приймають для всіх рівнів, починаючи із найнижчого – рівня передавання бітів – до найвищого, який реалізує сервіс користувачу мережі.

Процедура взаємодії комп'ютерів у мережі може бути описана у вигляді набору правил взаємодії кожної пари взаємодіючих сторін.

Формалізовані правила, які визначають послідовність і формат повідомлень, якими обмінюються мережеві компоненти, розміщені на одному рівні, але в різних вузлах, називаються протоколом.

Модулі, які реалізують протоколи сусідніх рівнів і знаходяться в одному комп'ютері користувача також взаємодіють один з одним відповідно до чітко визначених правил і за допомогою стандартизованих форматів повідомлень. Ці правила називають інтерфейсами.

Засоби кожного рівня мають відпрацьовувати, по-перше, свій власний протокол, а по-друге, інтерфейси із сусідніми рівнями.

Ієрархічний організований набір протоколів, достатній для організації взаємодії вузлів у мережі, називають стеком комунікаційних протоколів.

Отже, незалежність протоколів за рівнями один від одного і робить привабливим рівневий підхід.

13.2. Модель відкритих систем OSI

Протокол – це узгодження, які приймають два взаємодіючі об'єкти (у цьому разі два працюючі в мережі комп'ютери). Але це не означає, що вони мають обов'язково бути стандартними, хоча на практиці під час реалізації мереж прагнуть використати стандартні протоколи. Це можуть бути фірмові, національні або міжнародні стандарти.

На початку 80-х років ряд міжнародних організацій зі стандартизацією – Міжнародна організація стандартів (International Standards Organization – ISO) і сектор стандартизації електрозв'язку Міжнародного союзу електрозв'язку (ITU-T) (раніше Міжнародний консультативний комітет з телефонії і телеграфії – МККТТ) – розробили модель, яка відіграла значну роль у розвитку мереж. Цю модель називають моделлю взаємодії відкритих системи (Open System Interconnection – OSI) або моделлю OSI. Модель OSI визначає різні рівні взаємодії систем, дає їм стандартні імена

і вказує, які функції має виконувати кожний рівень. Модель OSI розроблено на основі великого досвіду, набутого під час створення комп'ютерних глобальних мереж, переважно в 70-ті роки.

У моделі OSI (рис. 13.6) засоби взаємодії мають сім рівнів: прикладний, представницький, сеансовий, транспортний, мережевий, каналний і фізичний. Кожний рівень відповідає окремим визначеним функціям мережевих пристроїв.

Модель OSI описує тільки системні засоби взаємодії, які реалізуються операційною системою, системними утилітами, системними апаратними засобами. Модель не має засобів взаємодії додатків кінцевих користувачів. Свої особисті протоколи взаємодії додатків реалізують, звертаючись до системних засобів.

Підіб'ємо підсумок, узагальнюючи зверненням додатка (програма користувача) до прикладного системного рівня, наприклад, зверненням до файлової служби (у цьому прикладі масиви критичних чисел Пірсона та Колмогорова – Смирнова). На підставі цього запиту програмне забезпечення прикладного рівня формує повідомлення стандартного формату. Повідомлення складається із заголовка і поля даних (інформація користувача). У цьому випадку заголовок, очевидно, має передати через мережу прикладного рівня комп'ютеру-адресату повідомлення про те, яку роботу йому треба виконати.

Після формування повідомлення прикладний рівень спрямовує його вниз стеком представницького рівня. Протокол представницького рівня на основі інформації, яку він отримав від заголовка прикладного рівня, виконує потрібні дії і додає до повідомлення особисту службову інформацію – заголовок представницького рівня комп'ютера-адресата. Отримане повідомлення передається вниз сеансовому рівню, який, у свою чергу, додає свій заголовок і т. д. Зрештою повідомлення досягає нижнього, фізичного рівня, який передає його каналом передавання комп'ютеру-адресату. До цього часу повідомлення «обростає» заголовками всіх рівнів (рис. 13.6).

Коли повідомлення через мережу надходить до комп'ютера-адресата, воно приймається його фізичним рівнем і послідовно пересувається вгору з рівня на рівень. Кожний рівень аналізує і обробляє заголовки свого рівня, виконуючи для цього рівня функції, а потім відкидає цей заголовок і передає повідомлення вищому рівню.

У моделі OSI розрізняють два основні типи протоколів.

Комп'ютер 1

Комп'ютер 2

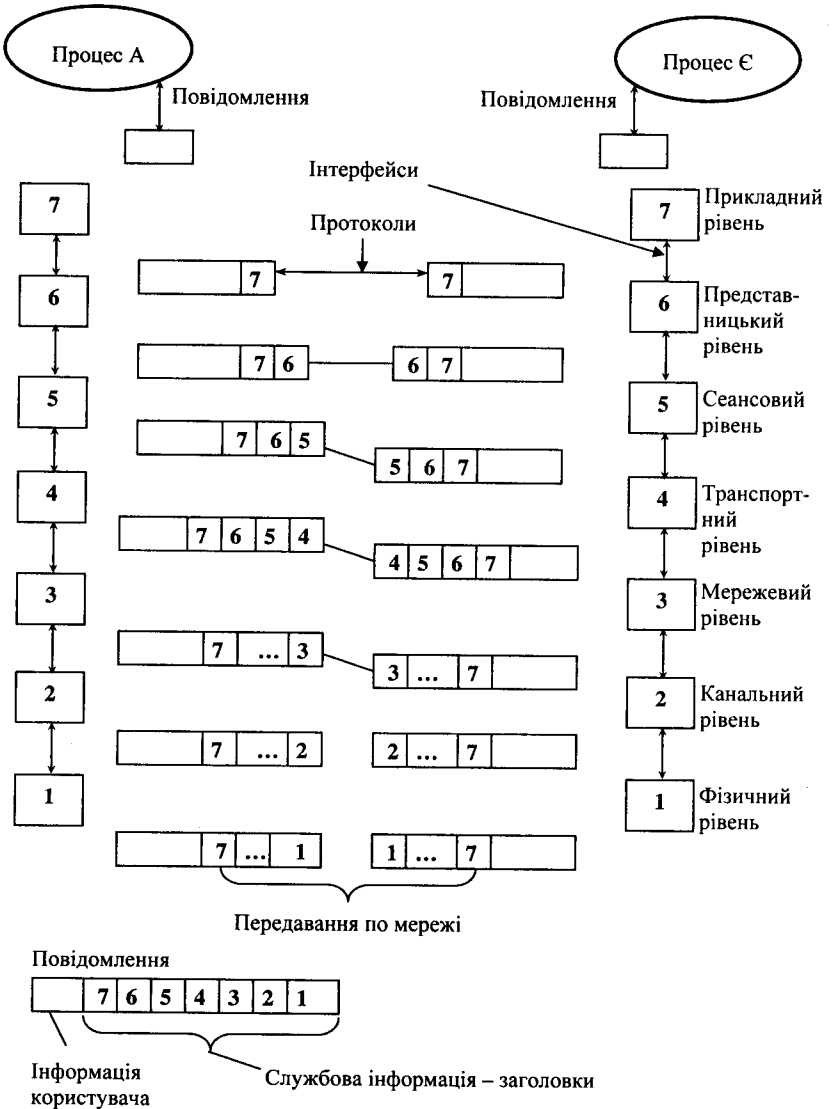


Рис. 13.6. Модель взаємодії відкритих систем ISO/OSI

Перший тип протоколів – протоколи з установленням з'єднання перед обміном даними. Відправник і одержувач повинні спочатку встановити з'єднання і, можливо, вибрати деякі параметри протоколу, який вони будуть використовувати для обміну даними. Після закінчення діалогу вони мають розірвати це з'єднання. Телефон – це приклад взаємодії, оснований на установленні з'єднання.

Другий тип протоколів – протоколи без попереднього установлення з'єднання. Такі протоколи називають також дейтаграмними протоколами. Відправник просто передає повідомлення, коли воно готове.

Для взаємодії комп'ютерів використовують протоколи обох типів.

13.3. Керування інформаційним каналом

Будь-яка мережева технологія має забезпечувати надійне і швидке передавання дискретної інформації лініями зв'язку (каналами передавання). І хоч технології дуже розрізняються, вони ґрунтуються на загальних принципах передавання даних. Ці принципи знаходять своє втілення у методах подання двійкових одиниць і нулів за допомогою імпульсних або синусоїдних сигналів у лініях зв'язку різної фізичної природи, методах виявлення і коригування помилок, методах компресії та методах комутації.

Для передавання дискретних даних каналами зв'язку застосовують два основні типи фізичного кодування – на основі синусоїдного несучого сигналу та послідовності прямокутних імпульсів. Перший спосіб називають модуляцією або аналоговою модуляцією, другий спосіб – цифровим кодуванням. Це і буде фізичний рівень, який детально вивчався в розділі «Системи передавання дискретної інформації».

Сформоване повідомлення з «обрамленням» заголовками потрібно передати фізичному каналу. Цю функцію виконує протокол керування інформаційним каналом, розміщеним на другому рівні OSI. Цей протокол визначає значну кількість функцій керування, до яких передусім належать:

- керування процедурами ініціалізації та деініціалізації каналу (установлення та припинення зв'язку);
- упакування переданої інформації в кадри перед її передаванням і розпакування інформації після її передавання фізичним каналом;
- генерування та читання керувальних кадрів;
- забезпечення прозорості інформаційного каналу;
- передавання та приймання підтверджень про прийом кадрів;
- утворення перевірних символів і перевірка вмісту кадрів після їх передавання фізичним каналом;

- повторне передавання тих кадрів, які виявилися загубленими або в яких виявлено помилки;
- керування потоками кадрів в обидва боки фізичного каналу.

У фізичних каналах використовують два методи передавання. Перший метод полягає в тому, що передавання виконується одночасно декількома каналами (паралельними електричними ланцюгами). Найчастіше передаються біти одного байта (вісім бітів) у тому разі, коли потрібні дуже високі швидкості передавання. Другий метод – це послідовне (біт за бітом) передавання інформації одним електричним ланцюгом фізичного каналу. Для передавання ним одного байта потрібно визначитися щодо передавання бітів, які його утворюють. Тоді передавальна частина направляє у канал по одному біту кожного байта, а приймальна – збирає байт із прийраних бітів.

У разі обміну даними на фізичному рівні одиницею інформації є біт, тому засоби фізичного рівня завжди підтримують побітову синхронізацію між передавачем і приймачем.

Канальний рівень оперує кадрами і забезпечує синхронізацію між приймальним і передавальним комп'ютерами користувачів на рівні кадрів. Приймач розпізнає початок першого байта кадру, межі полів кадру і закінчення кадру. Зазвичай достатньо забезпечити синхронізацію на зазначених двох рівнях (бітовому та кадровому), щоб передатчик і приймач могли забезпечити стійкий обмін інформацією. Однак за поганої якості лінії зв'язку для підвищення надійності передавання вводять додаткові засоби синхронізації на рівні байтів. Такий режим роботи називають асинхронним або старт-стопним. Другою причиною використання такого режиму роботи є наявність пристроїв, які генерують байти даних у випадкові моменти часу. Так працює клавіатура дисплея або термінального устаткування, з якого користувач вводить дані для обробки комп'ютером.

В асинхронному режимі кожний байт даних супроводжується спеціальними сигналами «Старт» і «Стоп» (рис. 13.7, а). Призначення цих сигналів полягає в тому, щоб повідомити приймач про надходження даних і щоб дати приймачу достатньо часу для виконання деяких функцій, пов'язаних із синхронізацією, до надходження наступного байта. Сигнал «Старт» має тривалість в один тактовий інтервал, а сигнал «Стоп» може тривати один, півтора або два такти, тому кажуть, що використовується один, півтора або два біти як стопові сигнали.

Асинхронним описаний режим називають тому, що кожний байт може бути зсунутий у часі відносно побітових тактів попереднього байта. Така асинхронність передавання байтів не впливає на коректність прийнятих даних, оскільки на початку кожного байта здійснюється додаткова синхронізація приймача з джерелом за рахунок бітів «Старт». Більш «вільніші» часові допуски визначають низьку вартість обладнання асинхронної системи.

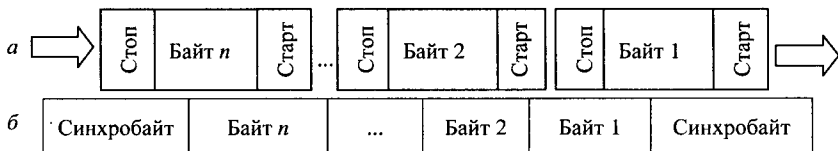


Рис. 13.7. Асинхронне (а) і синхронне (б) передавання на рівні байтів

За синхронного режиму передавання старт-стопних бітів між кожною парою байтів немає. Дані користувача збираються в кадр, який «обрамляється» байтами синхронізації (рис. 13.7, б). Байт синхронізації – це байт, який містить заздалегідь відомий код (наприклад, 01111110), надходження якого повідомляє приймач про надходження кадру. У разі його отримання приймач має вийти в байтовий синхронізм з передавачем.

Інколи використовують декілька синхробайтів для забезпечення більш надійної синхронізації приймача і передавача. Оскільки під час передавання довгого кадру у приймача може виникнути розсинхронізація під час роботи з передавачем, використовують самосинхронізувальні коди.

13.4. Методи керування передаванням даних каналного рівня (інформаційного каналу). Протоколи керування

Канальний рівень забезпечує передавання пакетів даних, які надходять від протоколів верхніх рівнів, вузла призначення, адресу якого також указує протокол верхнього рівня. Протоколи каналного рівня оформляють передані їм пакети в кадри визначеного формату, розміщуючи указану адресу призначення в одне з полів такого кадру, а також супроводжують кадр інформацією про застосований спосіб завадостійкості.

Найсуттєвіші характеристики методу керування передавання, а отже, і протоколу, який працює на каналному рівні, такі:

- асинхронний (синхронний);
- байтзорієнтований/бітзорієнтований;
- з попереднім устанавленням з'єднання (дейтаграмний);
- з виявленням спотворення даних (без виявлення);
- з виявленням утрачених даних (без виявлення);
- з відновленням спотворення і утрачених даних (без виявлення);
- з підтримкою динамічної компресії даних (без підтримки).

Багато з цих властивостей характерні не тільки для протоколів каналного рівня, але й для протоколів вищих рівнів.

13.4.1. Асинхронні протоколи

Найстаріший вид зв'язку – телеграфія, підтримувалась асинхронним протоколом. Асинхронні протоколи оперують не з кадрами, а з окремими символами, які представлені кодовими образами – байтами зі старт-стоповими сигналами. Одиницею передаваних даних був не кадр даних, а окремий символ. Деякі символи були керувальні, наприклад символ <CR> приписував телетайпу або дисплею виконувати повернення каретки на початок рядка.

В асинхронних протоколах використовують стандартні набори символів, найчастіше ASCII або МТК-5. Оскільки перші 32 або 27 кодів у цих наборах – спеціальні коди, які не відображаються на дисплеї або принтері, то вони використовуються асинхронними протоколами для керування режимом обміну даними.

Поступово асинхронні протоколи ускладнювалися і стали поряд з окремими символами використовувати цілі блоки даних, які називаються кадрами.

13.4.2. Синхронні байтзорієнтовані та бітзорієнтовані протоколи

У синхронних протоколах між пересилальними байтами (символами) немає старткових і стопових сигналів, тому окремі символи в цих протоколах пересилати не можна. Усі обміни кадрами виконуються кадрами, які мають у загальному випадку заголовок, поле даних і кінцевик (рис. 13.8).

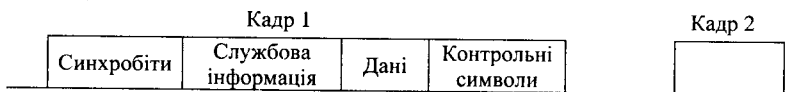


Рис. 13.8. Кадри синхронних потоків

Усі біти кадру передаються неперервним синхронним потоком, що значно прискорює передавання даних. Оскільки байти в цих протоколах не відокремлюються один від одного службовими сигналами, то однією з перших функцій є розпізнавання меж байтів. Потім приймач має знайти початок і кінець кадру, а також визначити межі кожного поля кадру – адреси призначення, джерела, інших службових полів заголовка, поля даних і контрольних кодових комбінацій (контрольних символів), якщо вони є.

Більшість протоколів припускають використання в кадрі поля даних змінної довжини. Інколи і заголовок може мати змінну довжину.

Синхронні протоколи каналного рівня бувають двох типів: байтзорієнтовані (символьнозорієнтовані) і бітзорієнтовані. Для обох типів харак-

терні одні і ті самі методи синхронізації бітів. Головна відмінність між ними полягає в методи синхронізації символів і кадрів.

Зупинимось на розгляді широковідомого одного з перших протоколів байтзорієнтованого типу, у характеристиках якого є попереднє установлення з'єднання, виявлення спотворень даних, виявлення загублених даних та їх відновлення.

13.4.3. Байтзорієнтований протокол двійкового синхронного зв'язку BSC

Цей протокол розробила фірма IBM для реалізації в системах телеоброблення і має скорочене позначення BSC (Binary Synchronous Communications). Протокол BCS – це набір правил керування обміном даними між користувачами двоточкового або багатоточкового з'єднання. Ці правила визначають синхронне напівдуплексне передавання даних по цьому з'єднанню на основі використання знаків первинного коду ASCII.

Формат кадру в протоколі BSC. Кадр – це масив, який складається з декількох частин, в яких записується певна інформація. У кадрі можуть бути поле адрес, інформаційне поле, поле керування тощо. Ці поля, об'єднуючись, утворюють три основні частини кадру: заголовок, основу і кінцевик. Структура цих частин визначає розміри та місцезнаходження різних полів і називається форматом кадру.

Протокол передбачає декілька способів передавання пакета, який приймається з третього рівня програмної структури OSI.

Повністю пакет може бути розміщений в основі кадру (рис. 13.9) і «обрамлений» символами, які утворюють заголовок і кінцевик кадру. У результаті отримується інформаційний кадр, який передає пакет інформаційним каналом.

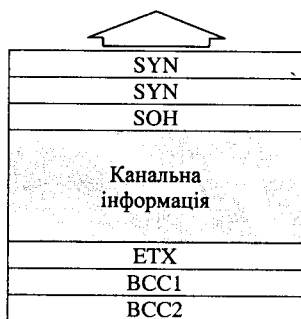


Рис. 13.9. Формат інформаційного кадру

Якщо пакет не розміщується в одному кадрі, тоді його ділять на частини. Для цього в протоколі передбачено, щоб кожна частина розміщувалася в окремому кадрі і щоб пакет у кадрі був розділений керувальними символами (рис. 13.10) на декілька частин. Така потреба виникає тоді, коли довгий пакет треба перевіряти частинами або передавати великий масив інформації без реверса інформаційного каналу. Таким способом можна також передавати і декілька взаємозв'язаних пакетів, які є частинами одного повідомлення.

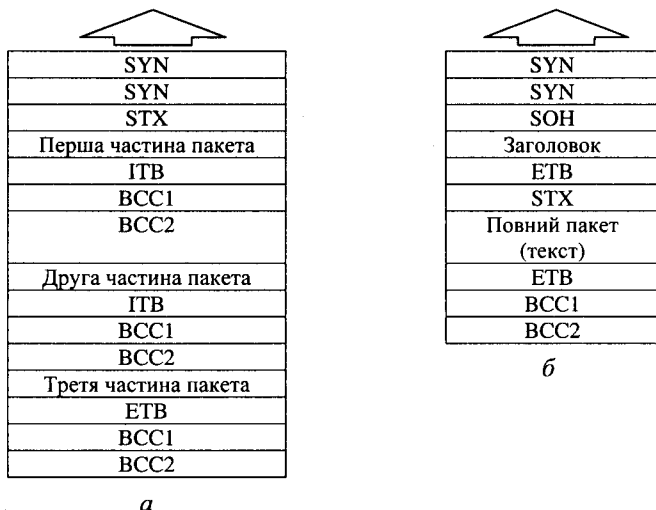


Рис. 13.10. Інформаційний кадр: *a* – пакет, розділений на частини; *б* – повний пакет у кадрі

Крім інформаційних кадрів, протоколом BSC передбачається використання керувальних кадрів, формат (структуру) яких показано на рис. 13.11.



Рис. 13.11. Формат керувального кадру

Керувальні кадри інформують про початок і кінець сеансу передавання, повідомляють про те, є чи немає помилок у передаваних пакетах, потребують очікування передавання чергового інформаційного кадру.

Побудова інформаційних і керувальних кадрів ґрунтується на введенні керувальних восьмибітових символів (команд), які розділяються на відповідні кадри. В обох типах кадрів використовується синхронізувальний символ SYN для встановлення і підтримки символної (знакової) синхронізації між приймальною і передавальною станціями. Крім того, його часто використовують для підтримки дискретного каналу в робочому стані у разі, коли бракує інформації для передавання цим каналом. Процедурою протоколу встановлено, що має бути не менше двох SYN. До керувальних символів інформаційних кадрів належать:

SOH (Start of Heading) – ПЗ (початок заголовка);

STX (Start of Text) – ПТ (початок тексту);

ETB (End of Transmission Block) – КБ (кінець блоку або пакета);

ETX (End of Text) – КТ (кінець тексту).

Символи керувального кадру такі:

ENQ (Enquiry) – запит на сеанс зв'язку та запит повторного зв'язку, якщо останній запізнився або надійшов спотворений;

ACK 0/1 (Acknowledgment) – Так (підтвердження приймання);

NAK (Negative Acknowledgment) – Ні (символ, який сповіщає передавача приймачем про прийняття інформаційного кадру з помилками і потребує повторного передавання);

WACK – Так (символ підтвердження правильного приймання, але тимчасово не має можливості прийняти наступний кадр);

RV1 – Так/Ні (символ, який повідомляє про правильне приймання, але вимагає припинити передачу наступних кадрів);

EOT (End of Transmission) – КП (кінець передавання – кінець сеансу зв'язку).

Кожний інформаційний кадр закінчується двома символами (BCC1, BCC2), які забезпечують перевірку пакета після його передавання фізичним каналом.

Найчастіше використовують три способи перевірки: перевірення кожного символу пакета на парність (непарність); позовжний контроль бітів інформації (уздовж кадру); циклічне перевірення інформації – чисел, які кодують передавану інформацію. Вони додаються, отримана сума ділиться на константу, частка відкидається, а остача служить двобайтовим контрольним символом (BCC1, BCC2).

Процедури керування передаванням даних. Сеанс зв'язку між парою комп'ютерів користувачів, які з'єднані інформаційним каналом, здійснюється в п'ять етапів (їх ще називають фазами):

1. Підключення каналу. Ця фаза потрібна тоді, коли з'єднання між комп'ютерами здійснюється через комутовану телефонну мережу. У разі використання орендованого каналу зв'язку потреби в цій фазі немає.

2. Запит на передавання. Для цього комп'ютер, який має передати один або декілька кадрів, надсилає запит на це другому комп'ютеру.

3. Передавання кадрів. У цій фазі інформаційний канал передає кадри. Один із комп'ютерів стає передавальним, а другий – приймальним.

4. Закінчення передавання. Передавши потрібну кількість кадрів, передавальний комп'ютер повідомляє про кінець сеансу зв'язку.

5. Вимкнення каналу. Після використання комутованого каналу зв'язку останній відключається. У разі використання орендованого каналу зв'язку потреби в цій фазі немає.

Приклади взаємодії комп'ютерів під час передавання двох пакетів, упакованих у кадри, показано на рис. 13.12.

Спочатку комп'ютер I направляє комп'ютеру II керувальний кадр, який виконує запит на початок сеансу зв'язку. У відповідь на нього комп'ютер II передає керувальний кадр підтвердження. Після цього комп'ютер I відправляє комп'ютеру II інформаційний кадр, у якому розміщено пакет А. Отримавши від нього керувальний кадр підтвердження про те, що пакет надійшов без помилок, комп'ютер I відправляє інформаційний кадр з пакетом В. Отримавши підтвердження про те, що в цьому пакеті після передавання каналом немає помилок, комп'ютер I відправляє керувальний кадр, який повідомляє про закінчення сеансу зв'язку.

У разі виявлення помилок приймальним комп'ютером в інформаційному кадрі він передає (рис. 13.13) керувальний кадр, у якому розміщено символи SYN, SYN, NAK, і вимагає повторного передавання інформаційного кадру. Після цього інформаційний кадр знову передається інформаційним каналом до комп'ютера II.

Протоколом BSC передбачено два способи подання інформації в пакеті. У першому з них (нормальному) вся інформація в пакеті кодується визначеним зразком, у другому – вміст пакета може кодуватися будь-яким способом, а її прозорість забезпечується застосуванням спеціальних засобів. Другий спосіб використовують для передавання інформації з плаваючою комою, блоків десяткової інформації, програм, написаних мовами комп'ютерів, даних, отриманих під час експерименту. Що стосується керувальних символів, які включаються в заголовки та кінцевики різних кадрів, то вони завжди кодуються визначеним способом.

Прозорість. Для керування інформаційним каналом відповідна апаратура (програмні засоби) має відрізнити керувальні символи (STX, ETB, SYN та ін.) від вмісту кадру, який передається в його основі. Проте за довільного способу кодування символів основи кадру в ньому можуть з'явитися такі, які мають ті самі коди, що й керувальні символи. Тобто в кадрі можуть бути символи, які не відрізняються від керувальних.

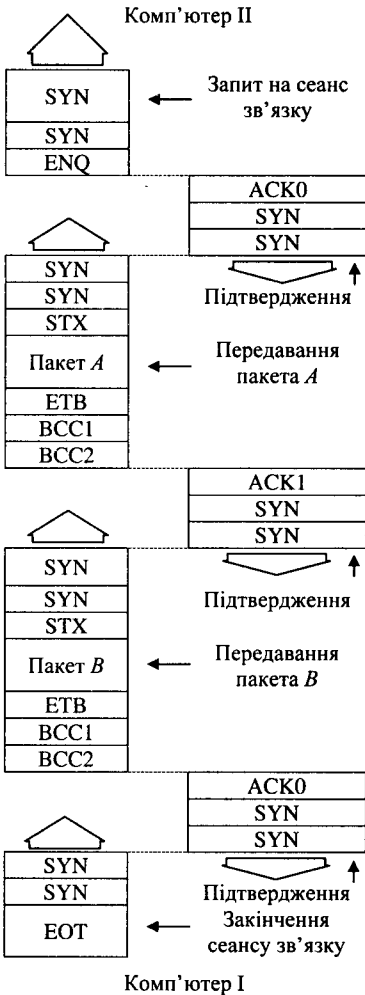


Рис. 13.12. Процедура передавання фізичним каналом

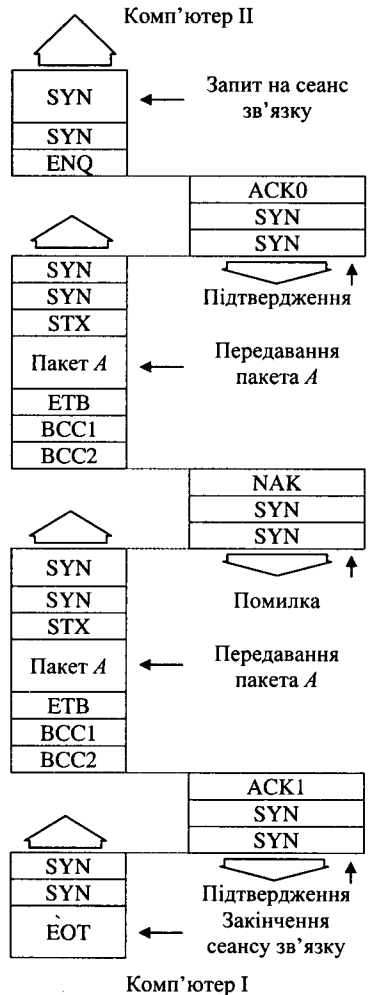


Рис. 13.13. Процедура повторного передавання кадру

Для того щоб відрізнити символи основи кадру від керувальних символів, використовують спеціальний символ DLE і виконують таку процедуру (рис. 13.14):

- перед кожним керувальним символом заголовка і кінцевика дописують символ DLE;

– перед кожним символом DLE, який з'являється в основі кадру, додають ще один символ DLE. Це операція – байтстафінг (наповнення кадру байтами DLE). Вона виконується перед передаванням інформаційного кадру в інформаційний канал. Після приймання цього кадру з інформаційного каналу виконується обернена операція – вилучення символів DLE, які були вставлені перед передаванням його в канал.

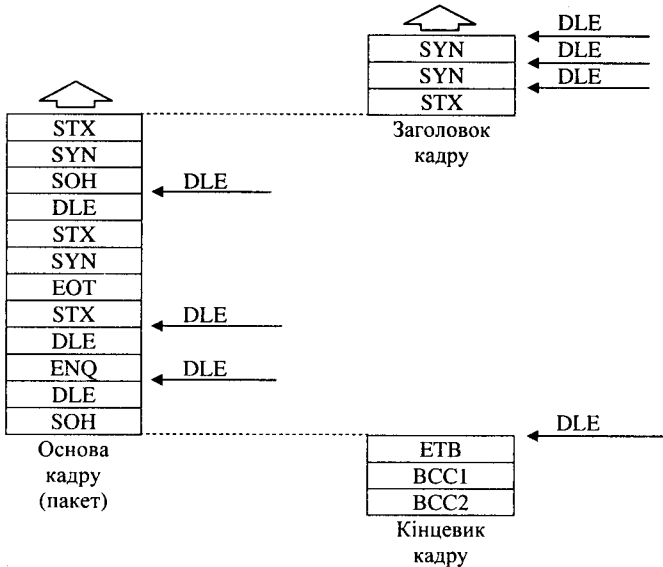


Рис. 13.14. Частини інформаційного кадру перед збиранням

Оскільки символ DLE більше ніде не використовується, то після виконання операції байтстафінгу всі керувальні символи стануть двобайтовими. Це дозволяє отримати алгоритм аналізу приймання з каналу кадру. Поява в кадрі символа DLE (рис. 13.15) є попередженням про те, що наступний символ буде керувальним, якщо він відрізняється від DLE. Якщо цей символ збігається з DLE, тоді він не є керувальним.

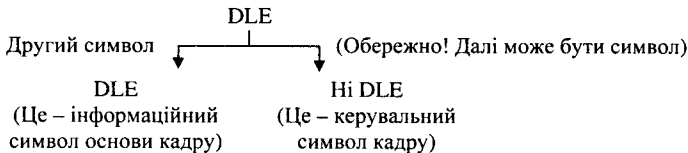


Рис. 13.15. Алгоритм виявлення символів керування каналом

Завдяки цьому за будь-яких кодів, використовуваних у пакеті, який вкладено в основу кадру, керувальні символи виділяються із загального потоку символів.

13.4.4. Бітзорієнтований протокол двійкового синхронного зв'язку HDLC

Основна особливість наведеного протоколу синхронного передавання полягає у його високій ефективності відносно протоколу BSC з використання надлишкової службової інформації та забезпечення «прозорості» передавання, що дозволяє передавати довільні тексти.

Уперше бітзорієнтований протокол розробила фірма IBM. Його скорочена назва SDLC (Synchronous Data Link Control – керування синхронним каналом передавання даних). На базі цього протоколу з деякими модифікаціями в ISO розроблено міжнародний протокол HDLC (High-Level Data Link Control – керування каналом передавання даних високого рівня).

Протокол HDLC визначає множину процедур керування каналом, із яких під час створення конкретної телекомунікаційної мережі вибираються потрібні. Тексти протоколу HDLC ще називають версіями протоколу, які відрізняються одна від одної.

Режими роботи протоколу HDLC. Режим роботи протоколу (версій) визначається правилами ініціалізації. У телекомунікаційних комп'ютерних мережах завдяки вкладеним мережевим програмам у версію протоколу можна заздалегідь (апріорі) визначати порядок надання права першому ініціювати зв'язок між абонентськими комп'ютерами. Це правило реалізується складовими програмними компонентами загального системного програмного мережевого комплексу.

Уведемо поняття «первинна та вторинна функції» і визначимо їх права та можливості. Первинна функція виконується тим абонентським комп'ютером, за яким закріплено в мережі право відповідати за встановлення та підтримку зв'язку. Надалі будемо розглядати ці функції у разі потреби встановити комп'ютерний зв'язок між двома абонентськими комп'ютерами телекомунікаційної мережі.

Розглянемо можливості протоколу HDLC у разі встановлення режимів роботи з обміну інформацією між двома абонентськими комп'ютерами (АК). Відповідно до протоколу BSC кожний АК може в будь-який момент часу звернутись до іншого АК з проханням щодо організації передавання даних. У протоколі HDLC ці функції регламентуються введенням керувальної інформації, яка надається у вигляді команд. Тоді згідно з концепцією цього протоколу АК, який виконує первинні функції керування інформаційним каналом, подає команди для вторинного АК, який формує інформацію у вигляді відповіді. Залежно від того, як в АК (їх ще називають

станціями) визначено і розміщено первинні та вторинні функції, створюються два типи інформаційного каналу: асиметричний та симетричний.

З метою організації керування одну станцію фізичної ланки вважають первинною, а другу – вторинною. Передбачається, що первинна станція несе відповідальність за організацію потоку даних під час операції відновлення після збоїв для своєї ланки. За нормальної роботи у вигляді реакції на команду очікується відповідь або ланцюжок відповідей.

Деякі станції можуть поєднувати в собі функції первинних і вторинних станцій. У цьому разі їх називають комбінованими. Протокол передавання між комбінованими станціями може бути симетричним, тобто обидві станції ланцюга виконують однаковий набір керувальних функцій. Такий спосіб дуже важливий для комп'ютерних мереж; його називають збалансованим.

Перейдемо до розгляду режимів роботи комп'ютерної взаємодії абонентських станцій, використовуючи введені функції ініціалізації.

У звичних умовах вторинна абонентська станція працює в одному з двох режимів: у режимі нормальної відповіді, у якому станція розпочинає передавання тільки в тому разі, якщо первинна станція дає відповідну команду; у режимі синхронної відповіді, у якому станція може розпочати передавання без команди від первинної станції.

Режим нормальної відповіді. У цьому режимі вторинна станція може починати передавання тільки у відповідь на командний кадр від первинної станції. Її відповідь складається з одного чи декількох кадрів, які вказують, який кадр відповіді останній. Після цього вторинна станція позбавляється можливості передавання доти, доки не отримає нову команду. Відповідальність за цей ланцюг несе первинна станція, яка має забезпечити опитування, надсилаючи команди вторинним станціям. Первинна станція відповідає за тайм-аути, повторні передавання і всі форми дій відновлення.

Режим синхронної відповіді. У режимі синхронної відповіді вторинна станція може за своєї ініціативи починати передавання кадру або групи кадрів. У цих кадрах може розміщуватись інформація користувача або вони можуть слугувати тільки цілям керування. Розпочавши передавання, вторинна станція відповідає за тайм-аути і повторне передавання, якщо на надіслані нею кадри не будуть отримані підтвердження правильності приймання.

У режимі синхронної відповіді вторинна станція може на деякий час відключатись і не опитуватись первинною станцією. Переключившись знову в робочий стан, вона має ініціювати передавання первинної станції.

Для деяких спеціальних типів фізичних ланцюгів процедури керування передбачають наявність первинних і вторинних станцій, але потребу-

ють ініціювання передавання вторинними станціями. Так, наприклад, працює сім'я протоколів ALONA, впроваджених до пакетного радіозв'язку.

Збалансований режим. У мережах із структурою довільного графу бажано забезпечити двосторонні інтенсивні інформаційні потоки між вузлами. Важливо, щоб ефективність передавань була максимальною, а час передавання найкоротшим. Для таких випадків і розроблено режим асинхронної відповіді, що дозволяє будь-якому вузлу ініціювати передавання. Однак для підвищення ефективності HDLC вводить збалансований клас процедур.

У разі збалансованої роботи передбачається, що кожна станція виконує функції і первинної, і вторинної, тобто є комбінованою станцією. Усі станції працюють за однаковим набором протоколів. Будь-яка станція може надсилати і приймати як команди, так і відповіді. Станція несе однакову відповідальність за нейтралізацію помилок. Збалансовані процедури працюють в одному режимі. Його називають асинхронним збалансованим режимом – АВМ (Asynchronous Balanced Mode), оскільки будь-яка станція може починати передавання в будь-який час. Іноді його називають просто збалансованим режимом.

У багатьох мережах з комутацією пакетів вузли – комутатори (маршрутизатори) пакетів – працюють у збалансованому режимі.

Незбалансовані процедури раніше були розроблені у передбаченні, що функції первинних станцій більш економно розмішувати у великих вузлах. Вторинні станції, кількість яких набагато більша, бажано зробити дешевшими. Масовий випуск великих інтегральних схем і значне зниження цін на комп'ютери персонального користування дозволило вкласти однакові керувальні функції в усе термінальне устаткування (персональні комп'ютери, комунікаційні вузли та мережеве устаткування). Надання переваги класу збалансованих процедур зумовлено саме цими обставинами.

Схематичне зображення взаємодії первинної і вторинної станцій показано на рис. 13.16.

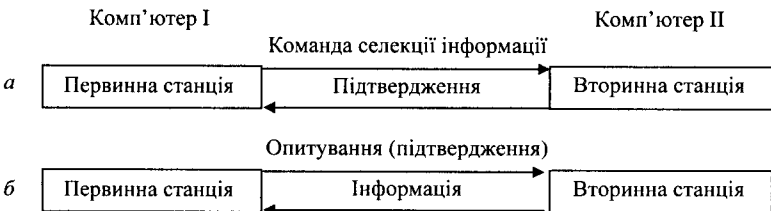


Рис. 13.16. Взаємодія первинної і вторинної станцій

Взаємодію первинної і вторинної станцій на випадок, коли первинна станція (головна) має інформацію для передавання вторинній станції, показано на рис. 13.16, а. Для цього первинна станція надсилає команду селекції (вибір відповідної вторинної станції для загального багатоточкового з'єднання) і передає інформацію вторинній. Взаємодію первинної і вторинної станцій на випадок, коли вторинна станція має інформацію для передавання первинній станції, показано на рис. 13.16, б. Для такої взаємодії використовують команду опитування. Команда опитування – це дозвіл первинної станції взяти на себе функції керування передаванням і всю відповідальність первинної з підтвердженням первинної. У такому сеансі взаємодії вторинна станція може відключитись від опитування, але в разі перемикання знову в робочий стан вона має ініціювати передавання первинної станції.

Для створення симетричного інформаційного каналу кожна станція виконує як первинні, так і вторинні функції, завдяки чому з двох асиметричних інформаційних каналів утворюється один симетричний (рис. 13.17).

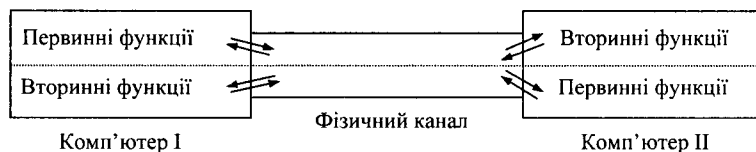


Рис. 13.17. Первинні та вторинні функції комп'ютера – термінального устаткування

Формат кадру HDLC. Обмін інформацією між взаємодіючими комп'ютерами в мережі здійснюється через передавання кадрів. Як і в протоколі BSC, тут використовують два види кадрів: інформаційні та керувальні. Кожний інформаційний кадр зберігається в передавальному комп'ютері доти, доки від приймального комп'ютера не надійде підтвердження про те, що кадр передано без помилок.

Структуру кадру, який застосовується в протоколі HDLC, або формат кадру, показано на рис. 13.18.

Кадр складається з трьох частин: трибайтового (у виняткових випадках – чотирибайтового) заголовка, основи кадру (інформаційного поля) розміром від 0 до N байтів і трибайтового кінцевика.

Кожний кадр починається і закінчується прапорцем (восьмибітовим керувальним символом), який відокремлює один кадр від другого в безперервному потоці інформації, яка проходить інформаційним каналом.

Прапорець, розміщений на початку кадру, називають відкриваючим, а прапорець, який знаходиться в кінці кадру, – закриваючим.

Кожний прапорець представлений у кадрі восьмибітовим кодом 01111110.



Рис. 13.18. Формат (структура) кадру HDLC

Протокол HDLC передбачає використання трьох типів кадрів, які визначають три фази (етапи) роботи двох комп'ютерів, які взаємодіють через інформаційний канал:

1. Ненумерований (U-кадр) призначається для виконання функцій фази установлення та припинення зв'язку пари комп'ютерів, з'єднаних інформаційним каналом (запит і згода на встановлення зв'язку, повідомлення про помилки у форматі кадру).

2. Супервізорний кадр (S-кадр) забезпечує виконання функцій фази керування передаванням інформаційних кадрів (повідомлення про готовність або неготовність до приймання, про те, є помилки місткості кадру чи немає їх, запит на повторне передавання кадрів).

3. Інформаційний (J-кадр), використовуваний у фазі передавання інформації інформаційним каналом від одного комп'ютера до другого.

Через інформаційний канал, який з'єднує пару комп'ютерів, передаються два зустрічні змішані потоки інформаційних кадрів. Для чіткого керування цими потоками та безпомилкової ідентифікації кадрів у кожному комп'ютері утворюється два лічильники (рис. 13.19). Один з них підраховує інформаційні кадри, які передає комп'ютер I в інформаційний канал, а другий рахує кадри, які отримує комп'ютер I із цього каналу. Значення, які видаються цими лічильниками, позначають відповідно $V(S)$ і $V(R)$. Якщо інформаційні кадри мають нормальне поле керування (1 байт), тоді лічильники видають циклічні значення, що дорівнюють 0, 1, 2, 3, 4, 5, 6, 7, 0, 1, 2..., якщо ж 2 байт, тоді значення лічильників дорівнюють 0, 1, 2, ..., 126, 127, 0, 1, 2...

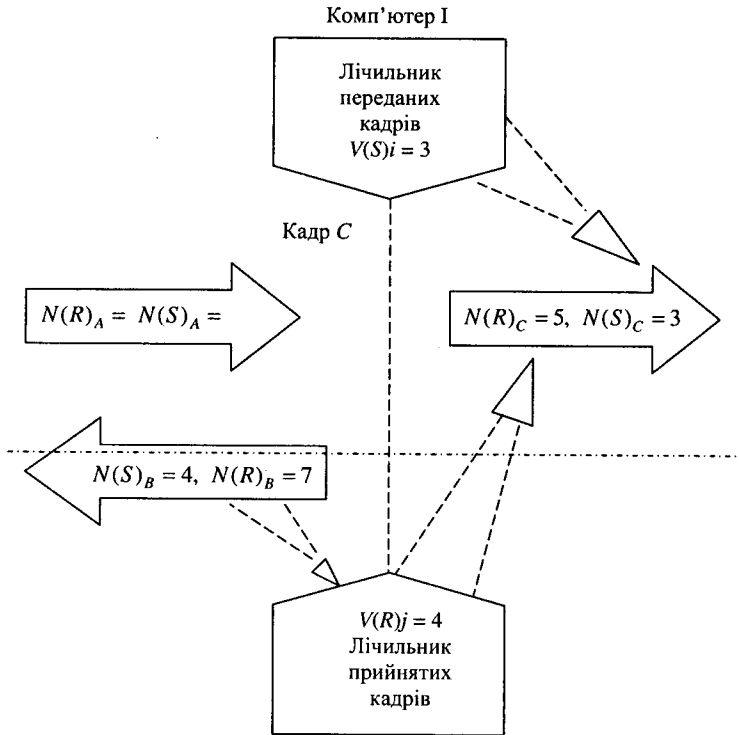


Рис. 13.19. Схема нумерації інформаційних кадрів

Безпосередньо перед відправленням інформаційного кадру через інформаційний канал у нього записуються:

$N(S)$ – номер передавального кадру;

$N(R)$ – номер очікуваного (під час приймання) кадру.

Ці номери визначаються (наприклад, кадр C) показанням лічильника комп'ютера I (рис. 13.19):

$$N(S)_C = V(S)_1, \quad N(R)_C = V(R)_1.$$

Якщо процедуру передавання кадру виконано без помилок, тоді номер переданого комп'ютером II кадру B має збігатися з показанням лічильника прийнятих кадрів у комп'ютері I (рис. 13.19):

$$N(S)_B = V(R)_1.$$

Якщо цю умову порушено, тоді прийнятий кадр вилучається, а комп'ютер II надсилає вимогу щодо повторного передавання кадру. Значення лічильника $V(R)$ у цьому разі не змінюється.

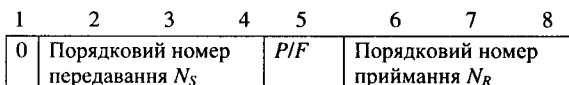
Якщо цю умову виконано і кадр передано без помилок, тоді він приймається комп'ютером I і значення лічильника $V(R)$ змінюється на одиницю.

Відповідно до виконаних умов у передавальному кадрі (кадрі C , див. рис. 13.19) $N(R)_C = N(C)_B + 1$. Тому кадр $N(R)$, що надійшов у комп'ютер II, повідомляє, що інформаційні кадри з номерами до $N(R) - 1$ передано правильно і прийнято комп'ютером I.

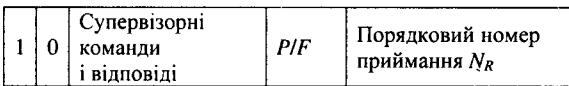
У супервізорних кадрах також розміщується інформація номера очікуваного інформаційного кадру (детальніше розглянемо це питання, вивчаючи супервізорні кадри та їх використання для підтримування інформаційного обміну).

Керувальний байт кадрів HDLC. Керувальний байт трьох типів кадрів розміщено третім у заголовку формату кадру (рис. 13.20).

J-кадр: кадр, у якому переноситься інформація користувача



S-кадр: супервізорний кадр



U-кадр: нумерований кадр

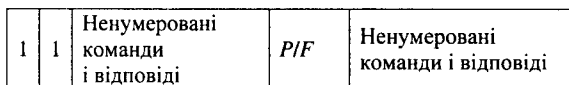


Рис. 13.20. Формат керувального байта в трьох типах кадрів

Керувальний байт показано на рис. 13.20. Він виконує декілька функцій керування обміном інформації.

П'ятий біт байта називають бітом опитування, коли кадр видається первинною станцією. Біт опитування встановлюється в одиницю у командах, які вимагають відповіді. Відповідь може містити декілька кадрів.

Останній кадр у відповіді на команду вміщує біт кінця передавання, що дорівнює одиниці, і на нього потрібне підтвердження.

Обмін між взаємодіючими комп'ютерами (станціями) контролюється за допомогою порядкових номерів. Зазвичай це трибітові номери, які показано на рис. 13.20 (за бажанням номери можуть бути розширені до 7 біт).

Відомо, що кожний інформаційний кадр ідентифікується порядковим номером $N(S)$ у розрядах керувального байта з 2 до 4. Приймання інформаційного кадру засвідчується посилкою підтвердження, у якій вміщується порядковий номер наступного інформаційного кадру, який приймальна станція очікує отримати, $N(R)$. При цьому передбачається, що всі попередні кадри з номерами $N(R)$ були прийняті правильно. Супервізорний кадр, у якому може передаватися підтвердження, вміщує $N(R)$ в останніх трьох бітах керувального байта. Підтвердження може бути також передано в інформаційному кадрі. Тому інформаційний кадр вміщує два порядкові номери: свій особистий номер $N(S)$ і номер, яким підтверджується прийнятий кадр, $N(R)$.

Два біти у S-кадрі визначають його тип. Отже, S-кадри можуть бути чотирьох типів.

В U-кадрах немає порядкових номерів, тому для ідентифікації типу U-кадрів використовують 5 біт. Отже, може бути 32 типи U-кадрів, однак не всі ці типи задіяно.

Супервізорні S-кадри. За допомогою S-кадрів виконується керування передаванням інформації, яку несуть J-кадри. У S-кадрів немає бітів між заголовком і кінцевиком. Інакше кажучи, це шестибайтові (48-бітові) кадри. Їх усього чотири, і кожний із них може бути або командою, або відповіддю. Тип задається в бітах 3 і 4 (див. рис. 13.20). У S-кадрі є порядковий номер приймання $N(R)$.

Значення бітів 3 і 4 пов'язані з типами таких S-кадрів:

00: RR (Recive Ready) – до приймання готовий;

01: REJ (Reject) – перепитування;

10: RNR (Recive Not Ready) – до приймання не готовий;

11: SREJ (Selective Reject) – селективне перепитування.

Кадр RR (до приймання готовий). Кадр посилається первинною або вторинною станцією для того, щоб:

- дати підтвердження, що вона правильно прийняла кадри до $N-1$ включно;
- повідомити, що вона готова до приймання кадру $N(R)$ (це нормальний механізм керування потоком інформаційних кадрів, які передаються без помилок).

Кадр RR можна використати для опитування вторинної станції, повідомляючи фактично таке: «Очікується J-кадр з номером $N(R)$; якщо такий кадр готовий, надсилайте його». Коли вторинна станція не готова надіслати дані, вона відповідає кадром RR, який означає: «Поки що я не маю J-кадру для передачі; наступний J-кадр, який очікую від Вас, – кадр з номером $N(R)$ ».

Кадр RNR (до приймання не готовий). Кадр надсилається первинною або вторинною станцією, щоб повідомити, що вона тимчасово знаходиться в стані зайнятості і не може приймати інформаційні кадри. Порядковий номер $N(R)$ повідомляє про те, що по закінченні стану зайнятості станція очікує такий кадр і служить також підтвердженням, що попередні кадри прийнято правильно. Про закінчення стану зайнятості можна сигналізувати за допомогою інших допустимих S-кадрів і деяких типів U-кадрів.

Кадр REJ (перепитування). Кадр надсилається первинною або вторинною станцією для запиту передавання або повторного передавання інформаційного кадру $N(R)$ і наступних за ним кадрів. Передбачається, що попередні $N(R)$ кадри прийнято правильно.

У кожний конкретний момент часу для кожного напрямку передавання фізичним каналом може бути тільки один стан перепитування. Стан перепитування знімається після того, як буде правильно прийнято J-кадр з номером $N(S)$, що дорівнює $N(R)$ у кадрі REJ.

Кадр SREJ (селективне перепитування). Кадр надсилається первинною або вторинною станцією для запиту передавання одного J-кадру з номером $N(R)$. Він підтверджує правильність приймання кадрів до $N(R)$ та інформує, що кадр $N(R)$ не був прийнятий правильно. Після передавання селективного перепитування станція прийме тільки J-кадр з номером $N(R)$. Особливий стан SREJ буде знятий тільки після отримання запитуваного J-кадру.

Селективне перепитування призначено для способу оброблення помилок і відоме як селективний запит повторного передавання.

Підтвердження. Станція, яка передає кадри, має отримувати підтвердження правильності приймання. Приймальна станція, однак, не зобов'язана надсилати окремі підтвердження на кожний кадр. Підтвердження вміщує в полі $N(R)$ порядковий номер наступного очікуваного кадру. Оскільки це свідчить про те, що кадри з меншими номерами прийнято правильно, підтвердження може стосуватися декількох кадрів.

S-кадри будь-якого із чотирьох типів слугують підтвердженням. Але для підвищення ефективності підтвердженням може використовуватись і інформаційний кадр (J-кадр). Підтвердження будь-якої форми вміщує номер $N(R)$ інформаційного кадру, який станція, що видала підтвердження, очікує отримати наступним.

U-кадри. Розглянуті J-кадри та S-кадри пов'язані з передаваннями, які контролюються порядковими номерами.

Кадри третього типу – U-кадри (ненумеровані кадри) не вміщують порядкових номерів.

U-кадри призначені для виконання таких функцій. Завдяки їм реалізується механізм передавання інформації, що не контролюється порядковими номерами. Кількість типів U-кадрів може досягати 32 (тип кодується п'ятьма бітами). Це дозволяє мати широкий набір керувальних команд і відповідей, а у разі потреби можна розширити керувальні можливості HDLC. U-кадри можуть містити поле інформації; такої можливості не мають S-кадри.

U-кадри використовують як команди, за допомогою яких виконується запуск станцій, зміна режиму передавання та відключення станцій. Це особливо зручно для ліній з номеронабирачами та непостійно підключених станцій.

U-кадри дозволяють відмовлятися від неправильних команд або обробляти особливі ситуації, які виникають, наприклад, у разі занадто довгих інформаційних кадрів, що перевищують розміри буферів на станціях.

Команди і відповіді U-кадрів. Зміст деяких типів U-кадрів фіксований. Одні з них є командами, інші – відповідями.

Основні команди задають режими передавання та відключають вторинну станцію:

- SNRM (Set Normal Response Mode) – установити режим нормальної відповіді (команда 1100P001);
- SARM (Set Asynchronous Response Mode) – установити режим асинхронної відповіді (команда 1111P000);
- SABM (Set Asynchronous Balanced Mode) – установити асинхронний збалансований режим (команда 1111P100);
- SNRME (Set Normal Response Mode Extended) – установити розширений режим нормальної відповіді (команда 1111P011);
- SARME (Set Asynchronous Response Mode Extended) – установити розширений режим асинхронної відповіді (команда 1111P010);
- SABME (Set Asynchronous Balanced Mode Extended) – установити розширений асинхронний збалансований режим (команда 1111P110);
- UA (Unnumbered Acknowledge) – ненумероване підтвердження про прийняття до виконання команди (відповідь 1100F110);
- DISK (Disconnect) – розірвати зв'язок, команда припинення роботи інформаційного каналу (команда 1100P010);
- DM (Disconnected Mode) – режим роз'єднання (команда 1111F000);

- TEST (Test) – тест команд і відповідей, використовуваний для тестування інформаційного каналу (команда/відповідь 1100 P/F 111);
- CMDR (Command Reject) – відмова від команди (відповідь 1110F001), нумерований кадр, який передає цю команду (також має трибайтову основу). У трибайтовому полі цієї основи записується інформація про причини відмови від виконання цієї команди. Вони такі:
 - 1) таких команд не буває;
 - 2) запрошена командою операція не реалізована на цій станції;
 - 3) поле J занадто велике і перебільшує розміри приймальних буферів станції;
 - 4) значення $N(R)$ не узгоджується з останнім надісланим значенням $N(S)$.

Кінцевик і прозорість кадру. Кінцевик кадру – замикальна частина всього кадру. Він складається із двобайтового поля перевірки кадру та однобайтового поля, в якому розміщується закриваючий прапорець.

Поле перевірки кадру, яке позначається в протоколі літерами FCS, вміщує інформацію, яка дозволяє виконати перевірку кадру після його передавання інформаційним каналом. Версія протоколу HDLC не стандартизує метод перевірки згідно з алгоритмом завадостійкого кодування. МККТТ пропонує такий метод кодування і перевірки, в якому двобайтовий символ перевірки є додатком до одиниці суми (за модулем два) таких двох доданків: остача від ділення за модулем два виразу $X^k(X^{15} + X^{14} + \dots + X^2 + X + 1)$ на поліном $X^{16} + X^{12} + X^5 + 1$, де k – підсумована кількість бітів адресного, керувального та інформаційного полів; остача від множення на X^{16} та наступного ділення за модулем два ємності адресного, керувального й інформаційного полів на поліном $X^{16} + X^{12} + X^5 + 1$.

Цей алгоритм, на перший погляд, складний. Проте він реалізується досить просто за допомогою спеціального лічильника і забезпечує високий ступінь точності перевірки.

Після поля перевірки кадру слідує (див. рис. 13.18) закриваючий прапорець, яким закінчується будь-який (інформаційний і керувальний) кадр.

Якщо за одним кадром одразу не передається другий кадр, тоді простір між ними може заповнюватись множиною прапорців, що надходять один за одним.

Прозорість. Щоб спрямувати кадр в інформаційний канал, потрібно, щоб у послідовності бітів, що передаються, не було жодного символа, який би збігався з прапорцем. Тому передавальний комп'ютер перед послідовною послідовністю бітів, яка визначається кадром, переглядає цю послідовність і виконує операцію бітстафінгу, як це показано на рис. 3.21.

| Початкові дані – будь-яка комбінація | Передавальні дані з нулем, вставленим після кожних п'яти суміжних одиниць | Приймальний комп'ютер «відкидає» кожний нуль після п'яти суміжних одиниць |
|--|---|---|
| 10101010101010 | 10101010101010 | 10101010101010 |
| 11111111111110 | 11111011111011 | 111111111111 |
| 01111110 | 01111100 | 01111110 |
| 01111110 | 011111010 | 01111110 |

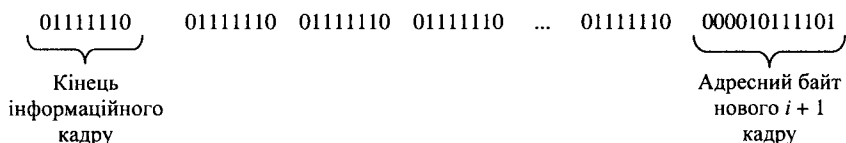
Ці комбінації тепер не вміщують шість суміжних одиниць і відповідно не будуть сприйматись як прапорець (01111110)

Рис. 13.21. Уставляння бітів у протоколі HDLC

Процедури керування обміном інформацією. Розглянуті кадри протоколу HDLC дозволяють організувати обмін між двома комп'ютерами – термінальним устаткуванням телекомунікаційних (комп'ютерних) мереж. Тим більше, що контроль за обміном інформаційними кадрами виконується завдяки впровадженню циклічного рахунку порядкових номерів $N(S)$ – передавальних інформаційних кадрів, та номерів $N(R)$ – очікуваних від передавальної станції. Для запису цих номерів у нормальному полі керування заголовка кадру виділено три, а в розширеному – сім двійкових разрядів, що забезпечують відповідно рахунок від 0 до 7 або від 0 до 127. Тому передавальний комп'ютер може, не дочекавшись підтвердження, відправити в інформаційний канал відповідно до 7 або 127 інформаційних кадрів.

Для передавання кадрів між ними можуть виникати проміжки, зумовлені підготовкою чергового кадру для передавання, або нездатністю передавального комп'ютера відразу почати передавання після отримання відповіді на прийнятий ним попередній кадр.

Стан інформаційного каналу, коли обмін відбувається з виникненням проміжків або без них, називають активним станом. У цьому стані неперервно надсилаються прапорцю байти. Приймальна станція, виявивши прапорцевий байт, який не супроводжується наступним прапорцевим байтом, визнає, що розпочалось передавання кадру. Отже:



Пасивний стан інформаційного каналу означає, що передавальна станція змушена закінчити сеанс зв'язку під час передавання інформаційного кадру, не чекаючи повного передавання цього кадру. Це називається вимушеним обривом кадру. Існують два типи обривання кадру.

Перехід до пасивного стану потрібний тоді, коли передавальна станція не передає повністю кадр, але відновлює передачу після передавання керувального кадру. Такий вид обривання потребує надсилання кодової комбінації з мінімальною кількістю 15 одиниць – 11111111 1111111.

Перехід до пасивного стану потрібний ще й тоді, коли передавальна станція прагне обіграти передавання інформаційного кадру (відміна кадру) і відразу передати новий J-кадр. Такий вид обривання потребує надсилання кодової комбінації з восьми одиниць і прапорця – 11111111 01111110.

У протоколі HDLC передбачається часовий контроль за процесом кадрового обміну.

Первинна станція, що надіслала кадр, у керувальному полі якого встановлено біт опитування, очікує відповідь протягом заданого часу. Не отримавши її, станція знову надсилає кадр, що запитує відповідь. Час витримки називається тайм-аутом. У первинній станції є годинник, за допомогою якого вона може вимірювати інтервал часу, коли лінія не працює, і за достатньої його тривалості з'ясувати, чи трапилося відхилення від норми. Часто відповідь не надходить через помилки, що виникають під час передавання. Повторне передавання дозволяє усунути ці помилки. Якщо ж вони не усуваються, то первинна станція повторює запит певну кількість разів. Якщо і це не приводить до успіху, то вважається, що на рівні протоколу фізичного рівня відновлення неможливе і потрібно звернутись до засобів вищих рівнів.

Використовують два тайм-аути. Перший – якщо немає відповіді від вторинної станції (тайм-аут на відсутність відповіді), другий – якщо отримано неправильний сигнал (тайм-аут на некоректну або неправильну відповідь).

Ні тривалість тайм-аутів, ні кількість спроб у разі повторення послілок не регламентовані в HDLC. Ці кількості змінюються від однієї реалізації до другої і залежать від характеристик каналу зв'язку та часу, потрібного для формування відповіді.

Керування інформаційним обміном. Взаємодія двох комп'ютерів характеризується трьома фазами (етапами) роботи: встановлення, підтримка та припинення зв'язку. До встановлення зв'язку інформаційний канал знаходиться у пасивному стані, а лічильники кадрів – у нульовому положенні.

Організація взаємодії можлива у двох випадках:

- 1) первинна станція прагне встановити зв'язок із вторинною;
- 2) вторинна станція прагне встановити зв'язок з первинною.

У першому випадку первинна станція надсилає U-кадр з однією із команд SNRM, SARM, SABM, SNRME, SARME, SABME з автоматичним увімкненням таймера T1.

Якщо вторинна станція цю команду прийняла без помилок і в неї є можливість установити зв'язок та увійти до робочого режиму обміну інформацією, тоді вона передає відповідь U-кадром з командою-підтвердженням UA. При цьому на вторинній станції лічильники $N(S)$ і $N(R)$ установлюються в нуль, а зв'язок визнається встановленим.

Якщо вторинна станція не може увійти до робочого режиму обміну інформацією, тоді вона передає U-кадр з відповіддю-командою DM.

Для встановлення зв'язку станція передає в канал неперервну послідовність прапорців – канал переводиться в активний стан. Нагадаймо, що канал визнається активним, якщо через нього передається кадр, послідовність передчасного переривання кадру або послідовність прапорців, які тимчасово заповнюють проміжок між кадрами.

Процес функціонування інформаційного каналу в разі взаємодії двох комп'ютерних станцій зі встановленням, підтримкою та припиненням зв'язку, де первинною станцією надіслано пропозицію SNRM установити роботу у нормальному режимі і від вторинної станції отримано команду-підтвердження UA на цю роботу, показано на рис. 13.22.

У фазі встановлення зв'язку одночасно з передаванням керувального U-кадру з командою SNRM вмикається таймер T1. Приймальна станція, якщо вона готова до взаємодії, відповідає підтвердженням U-кадру з командою UA. Передавальна станція, отримавши згоду, починає передавати інформаційні J-кадри.

Одночасно з передаванням підтвердження UA на приймальній станції лічильники $N(S)$ і $N(R)$ установлюються в нульовий стан, а з'єднання вважається установленим.

Якщо приймальна (вторинна) станція не може увійти в робочий режим обміну інформацією, тоді вона передає відповідь DM, яка є відповідним відмовленням від установлення зв'язку, а з'єднання вважається не встановленим.

Після приймання коректної відповіді UA передавальна (первинна) станція вмикає таймер T1 і переводить свої змінні $V(S)$ і $V(R)$ у нульовий стан, а з'єднання вважається установленим.

За нормального установлення зв'язку первинна станція в U-кадрі з командою, наприклад SNRM у керувальному байті, вставляє $P = 1$, а приймальна станція у відповіді U-кадром з DM вставляє $F = 1$.

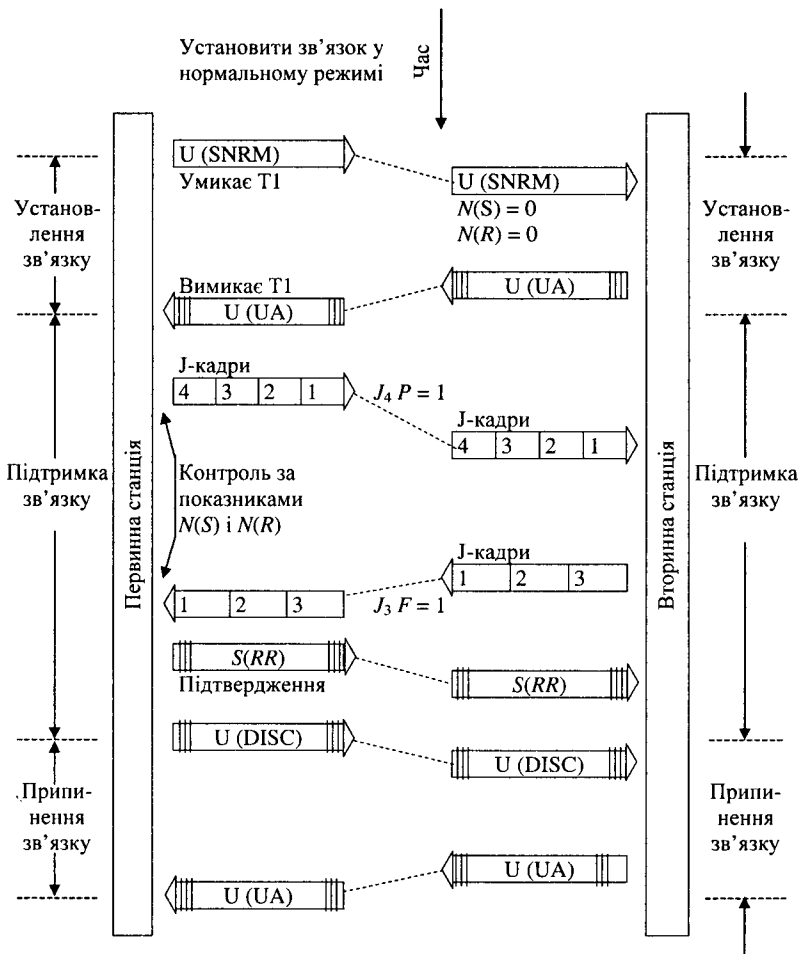


Рис. 13.22. Фази (етапи) взаємодії двох комп'ютерних станцій

Якщо команда SNRM або відповіді UA і DM були прийняті з помилками, тоді вони не враховуються станціями і на передавальній станції закінчиться регламентований час таймера T1. Останнє є для цієї станції вказівкою на повторне передавання U-кадру з командою SNRM. Зазначена послідовність дій може продовжуватись доти, доки не буде правильно прийнято відповідь UA чи DM або витрачено кількість регламентованих повторів, визначену параметром N2.

Таким чином, встановлюється з'єднання і розпочинається фаза інформаційного обміну – фаза підтримки зв'язку. Кожному передаваному J-кадру присвоюється порядковий номер $N(S)$ за показником лічильника $V(S)$, прийнятий J-кадр з номером $N(S)$ порівнюється з показником лічильника $N(R)$ приймальної станції; якщо вони однакові, тоді лічильник $V(R)$ збільшується на одиницю. Так формується для відповіді передавальної (первинної) $N(R)$ очікуваний черговий J-кадр, якщо такий є.

Ця відповідь з очікуванням $N(R)$ має бути надіслана своєчасно до закінчення відведеного часу тайм-аутом T_2 . Проміжок часу визначається з моменту надходження J-кадру до моменту обов'язкового передавання відповіді, яка має надійти на протилежну станцію до часу T_1 , встановленого для цього кадру, якого стосується ця відповідь ($T_1 > T_2$).

Фаза підтримки зв'язку закінчується із закінченням обміну інформаційними кадрами. Передавальна (первинна) станція отримала останній J-кадр від приймальної станції, де встановлює $F = 1$ – останній кадр, і якщо вона не має інформації для передавання, то повідомляє вторинну станцію S-кадром командою RR з $N(R)$, номером кадру, який очікується від неї. Це підтвердження того, що всі кадри, які вона надсилала, прийнято правильно.

Після закінчення інформаційного обміну взаємодія припиняється фазою припинення зв'язку. Для припинення зв'язку передавальна (первинна) станція передає U-кадр з командою DISK $P = 1$ з одночасним увімкненням таймера T1. Приймальна (вторинна) станція, прийнявши правильно команду DISK, передає відповідь U-кадром з командою UA і $F = 1$. Первинна станція, отримавши UA, вимикає таймер T1 і переходить в режим роз'єднання.

Якщо у фазі підтримки зв'язку або у фазі припинення зв'язку будуть виявлені помилки під час передавання кадрів чи збійні ситуації, тоді використовуються часові контрольні параметри T_1 , T_2 і $N2$.

Розглянемо деякі ситуації внутрішнього механізму обміну інформаційними кадрами у фазі підтримки зв'язку. Фрагмент обміну між первинною і вторинною станціями, коли первинна станція надсилає вторинній один J-кадр і вторинна станція прагне надіслати первинній також один J-кадр, показано на рис. 13.23.

Ситуацію, в якій одна (первинна) станція надсилає пакет послідовних J-кадрів, а у відповідь приймальна (вторинна) станція надсилає свій пакет кадрів з більшою кількістю J-кадрів, показано на рис. 13.24.

Ситуацію, за якою вторинна станція першою має потребу в передаванні первинній одного J-кадру, а другою – потребу в передаванні вторинній одного J-кадру, показано на рис. 13.25.

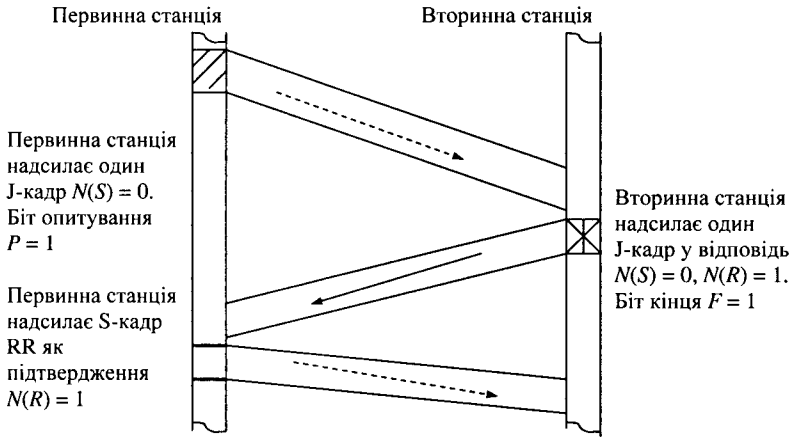


Рис. 13.23. Фрагмент обміну одиночними J-кадрами

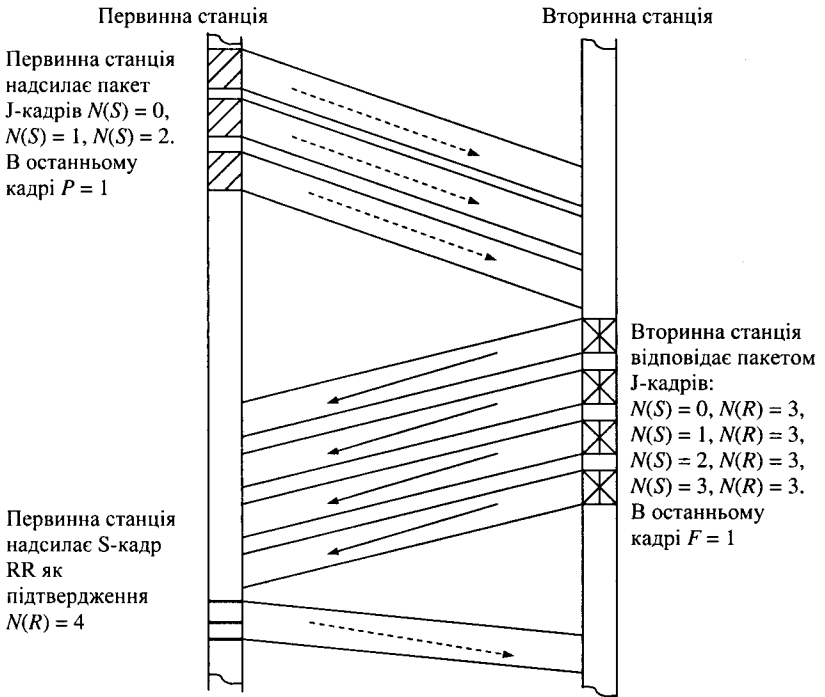


Рис. 13.24. Фрагмент обміну пакетами J-кадрів

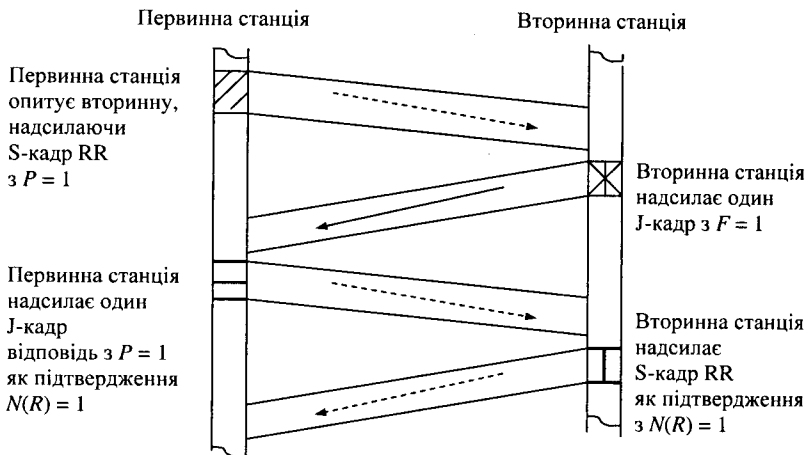


Рис. 13.25. Фрагмент обміну одиночними потоками за ініціативи вторинної станції

Ситуацію, коли ініціативу першою розпочати інформаційну взаємодію бере на себе вторинна станція під час обміну пакетами J-кадрів, зображено на рис. 13.26.

У режимі нормальної відповіді вторинна станція не може ініціювати обмін і тому первинна станція опитує вторинну, щоб з'ясувати, чи є в неї дані для передавання. Якщо їх немає, тоді вторинна станція відповідає на опитування S-кадром, зазвичай типу RR (до приймання готовий). Якщо дані є, тоді вони надсилаються у J-кадрах.

Порівняльний аналіз протоколу HDLC з розвиненими технологіями Frame Relay (FR) та ATM. У цьому розділі не ставиться мета провести глибокий детальний аналіз протоколу HDLC порівняно із сучасними технологіями передавання даних FR та ATM, а показати його як базовий і розвинений. Але докладне вивчення HDLC як бітзорієнтованого синхронного дає змогу швидкого засвоєння сучасних методів передавання даних.

На базі HDLC створено і широко використовується технологія X.25, у якій формат пакета вкладається у конверт HDLC. Основний недолік технології X.25 – ряд принципів обмежень щодо швидкості. Через це технологію X.25 вважають повільною і застарілою.

Для швидких ліній зв'язку з використанням основи протоколу HDLC створено технологію – протокол ретрансляції кадрів Frame Relay – бітзорієнтований синхронний протокол, формат кадру якого дуже схожий на формат кадру HDLC. Головний недолік FR – повільна комутація пакетів у вузлах мережі.

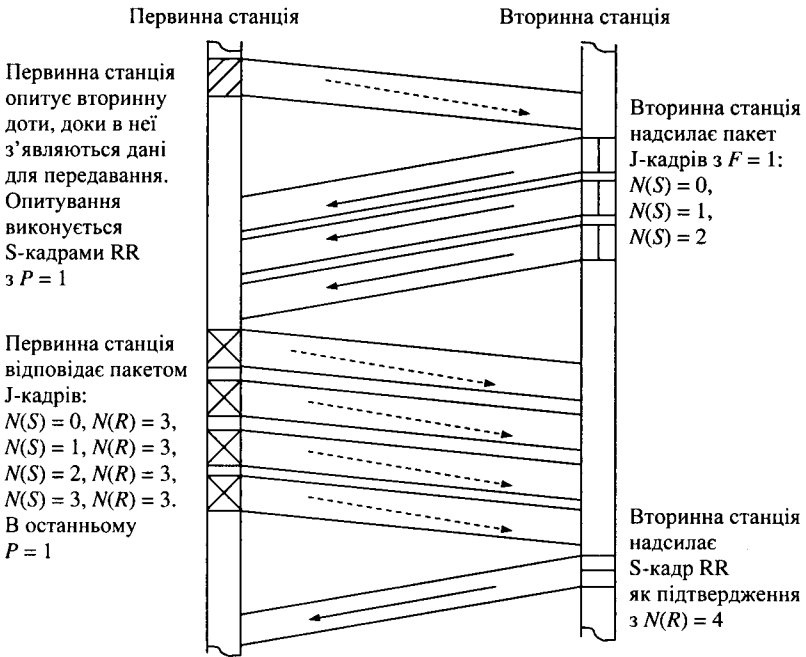


Рис. 13.26. Фрагмент обміну пакетами J-кадрів за ініціативи обміну вторинної станції

Для підвищення швидкості комутації пакетів з'явилася концепція асинхронного режиму перенесення (Asynchronous Transfer Mode) ATM, основна ідея якої – пакетна комутація з мінімумом функцій, виконуваних вузлами комутації на рівні ланки з метою підвищення рівня часової прозорості мережі.

Докладно особливості FR, ATM та X.25 вивчаються в курсі «Глобальні інформаційно-комп'ютерні мережі».

Щоб виконати аналіз і вибір транспортного протоколу для проектованої мережі, розробник повинен насамперед визначитись, з яким основним видом трафіку буде працювати ця мережа магістральними каналами. Якщо це здебільшого передавання трафіку локальної мережі, тоді основними вимогами до характеристик такої мережі будуть вимоги до семантичної прозорості, тобто ймовірність приймання інформаційного елемента з помилками, а для інтерактивних застосувань (таких, як розподілені бази даних) – невеликі значення затримки доставляння повідомлень користувачу.

У разі використання магістральних каналів корпоративної мережі для передавання прикладень, чутливих до затримок, наприклад голосу, факсу, відеозображень, крім вимог до семантичної прозорості і малого значення коефіцієнта помилок, додається ще й вимога до часової стабільності та синхронізації моментів надходження, передавання й оброблення на приймальному кінці даних користувача. Важливим стає той факт, що більшість існуючих на цей момент цифрових систем передавання плезіохронної ієрархії мають показник P_{ber} , що характеризує якість їх роботи, у той час, як передаваний таким каналом трафік HDLC більш чутливий до показника P_{per} .

Для того щоб розробити рекомендації щодо застосування протоколів залежно від типів передаваної інформації, потрібно спочатку виділити ці типи, охарактеризувавши їх основні вимоги до мереж передавання даних. Найбільш основні типи інформації разом зі сферою впровадження й основними вимогами наведено у табл. 13.1.

Таблиця 13.1

| Тип інформації | Застосування |
|--------------------------|---|
| Рухомі зображення і звук | Передавання на відстань голосу, ТВ-сигналів, рухомих зображень, документів і т. ін. |
| Звук | Передавання на відстань голосу, синхронний переклад |
| Дані | Високошвидкісне передавання файлів і даних, система тривоги і т. ін. |
| Документи | Обмін документами |

Основні вимоги під час передавання інформації перших двох типів ставляться до часової прозорості мережі. Це робить застосування протоколу HDLC у мережах з такою інформацією просто неприйнятним, оскільки наявність засобів коригування та відновлення інформації, а також часта їх подвійність у разі об'єднання з протоколами верхніх рівнів спричиняє великі затримки.

Як варіант протоколу HDLC в експлуатації знаходяться мережі під керуванням протоколу Frame Relay. Вони оптимізовані для передавання даних, однак характеризуються великими значеннями варіації часу затримки і допускають втрату кадрів.

Мережі ATM, які створюються на високоякісних лініях зв'язку (звичай волоконно-оптичних), якнайкраще підходять для такої інформації, оскільки особливих засобів коригування вони не містять, а за фіксованої довжини пакета відносно HDLC і Framy Relay досягають малих затримок.

Дані другого і третього типів також будуть швидше передаватися в мережах ATM, які спочатку і розроблялись для передавання даних.

Проте використання пакета фіксованої довжини швидше недолік, ніж перевага для даних цього типу.

Потрібно визначити, що з використанням каналів низької якості в Україні актуальним стає використання мереж HDLC, які характеризуються наявністю потужних засобів корекції помилок.

Для якісної роботи мереж АТМ належить забезпечити досить низьку ймовірність виникнення помилок. Тобто технологія АТМ, досягнувши високих швидкостей передавання, втратила в засобах корекції і стала непридатною на неякісних мережах.

14. Методи підвищення достовірності передавання дискретної інформації

У попередніх розділах було розглянуто вплив завад на інформаційний (корисний) сигнал, який передається каналом передавання даних. Для боротьби із завадами на сигнальному рівні (фізичному) використовують перетворення сигналів передавальною частиною СПДІ (ресстрацію і синхронізацію приймальної частини цієї системи). Якщо ці методи не справляються з цим завданням, тоді вступають у дію методи структурної боротьби із завадами, які спричиняють виникнення помилок. Структура незахищеного коду (без уведення примусової надлишковості) – простого коду – може призвести до виникнення помилки. Це викликано перехідними електричними процесами. Усунення цього недоліку частково досягається введенням додаткової процедури – скремблюванням (формуванням суміші корисних кодових комбінацій з псевдовипадковими кодовими комбінаціями).

Якщо зазначені методи не спроможні боротися з впливами завад, використовують структурні методи виявлення та коригування кодових послідовностей. Але для того, щоб впроваджувати ці методи і засоби, потрібно навчитись заздалегідь передбачати виникнення цих «небажаних» випадкових подій. Методи, які дозволяють на підставі попереднього дослідження каналу передавання дискретної інформації встановити закономірність виникнення помилок, яка зводиться до визначення складових вектора E помилок, докладно розглянуто в розд. 11.

14.1. Класифікація методів

Методи підвищення правильності передавання інформації численні і різноманітні. Для зручності аналізу розподілимо ці методи на три групи. Більшість причин зниження правильності передавання пов'язано з властивостями каналів зв'язку, через які передається дискретна інформація. Тому до першої групи належать методи експлуатаційного та профілактичного характеру. Вони покликані, якщо не усунути, то скоротити кількість джерел завад і зменшити інтенсивність їх дії і спотворень, які спричиняють помилки під час передавання дискретної інформації: нерівномірність АЧХ і ФЧХ, імпульсні шуми і короткочасні перерви, нестабільність рівня передавання, зсуви несучих частот і т. ін. Це досягається технічними і організаційними заходами.

До другої групи включено методи, спрямовані на підвищення завадостійкості передавання одиничних елементів дискретної інформації (перетворення сигналів, які менше зазнають впливу завад). Такими заходами є підвищення відношення сигнал/завада внаслідок збільшення амплітуди, тривалості або спектра частот корисного сигналу, застосування найбільш завадостійких методів модуляції та удосконалених методів приймання сигналів і т. ін. Але можливості цього напряму підвищення завадостійкості сигналів досить обмежені. Запобігання впливу завад на сигнали розглянуто в розд. 14. Усі способи підвищення якості передавання пов'язані з використанням найбільш завадостійких методів передавання.

Третя група методів підвищення правильного передавання дискретної інформації – заходи, спрямовані на виявлення і виправлення помилок у прийнятій інформації. Ці методи можна реалізувати системами без зворотного зв'язку і системами зі зворотним зв'язком. Для реалізації цих заходів потрібно передаваній інформації надати нові властивості, які б використовувала приймальна частина системи передавання для виконання функцій виявлення або виправлення помилок.

Оскільки в основі більшості методів підвищення правильності в системах як зі зворотним зв'язком, так і без зворотного зв'язку лежить використання надмірних кодів, перейдемо до вивчення самих кодів.

14.2. Надмірне кодування дискретної інформації

Вивчення процесу виникнення помилок у дискретному каналі сприяло побудові загальної моделі дискретного каналу. На вхід каналу подається послідовність двійкових символів a_1, a_2, a_3, \dots , які являють собою рівномірний ненадмірний код. Функція складових частин дискретного каналу полягає в доставлянні цієї послідовності символів одержувачу з мінімальними спотвореннями, незважаючи на наявність помилок $e_1, e_2, e_3, \dots, e_n$ в дискретному каналі. У процесі кодування символи a_i , подані рівномірним ненадмірним k -елементним кодом, перетворюються в n -елементні кодові слова b_i коригувального коду (n, k) .

Основне завдання надмірного кодування – виявлення та виправлення помилок на виході дискретного каналу. Воно полягає в тому, що в рівномірному блоковому коді використовуються не всі $N = 2^n$ можливих кодів комбінацій, а лише деяка їх частина – $N_0 < N$ (де N_0 – кількість кодів комбінацій, потрібних для кодування елементів представленого об'єкта). Кодові комбінації, які використовуються для передавання N_0 , називають дозволеними, а решту $(N - N_0)$, які не використовуються – забороненими. Якщо під час передавання дискретним каналом одна з дозволених перетворюється в одну із заборонених, то тим самим виявля-

ється наявність помилок. Очевидно, якщо сукупність помилок у каналі перетворює кодову комбінацію в іншу, також дозволена комбінацію, тоді помилки не можуть бути виявлені.

Отже, в основі кодування з виявленням помилок лежить розбиття всієї множини кодових комбінацій N на дві неперетинні підмножини: підмножина дозволених комбінацій $\{A_i\}$ і підмножина заборонених комбінацій $\{B_i\}$ (рис. 14.1).

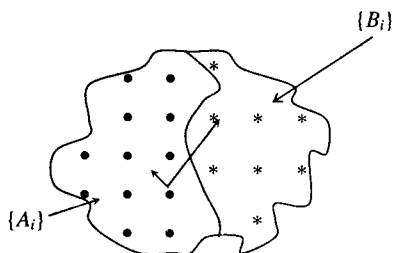


Рис. 14.1. Підмножини дозволених і заборонених комбінацій

Коди, які дозволяють виявляти помилки, називають кодами з виявленням помилок. Будь-який надмірний код здатний виявляти помилки в тих $N_0(N - N_0)$ випадках, коли будь-яка з N_0 кодових комбінацій переходить у будь-яку $N - N_0$ із заборонених. Загальна кількість можливих випадків дорівнює N_0N , оскільки будь-яка з N_0 кодових комбінацій може перейти під час передавання в будь-яку із наявних N . Отже, частина виявлених помилкових комбінацій становить $N_0(N - N_0)/N_0N = 1 - (N_0/N)$.

Для виправлення помилок підмножину заборонених комбінацій $\{B_i\}$ ($i = 0, 1, 2, \dots, N - N_0$) потрібно розбити на N_0 неперетинних підмножин M_k . Кожна із цих множин приписується одній із кодових комбінацій A_k . За довільного розбиття на підмножини M_k , кількість яких $(N - N_0)/N_0$, помилка завжди виправляється в $[(N - N_0)/N_0]N_0 = N - N_0$ випадках. Отже, відношення кількості виправлених помилкових комбінацій до кількості виявлених помилкових комбінацій $N_0(N - N_0)$ становить $1/N_0$. Отже, кількість виправлених помилок завжди менша, ніж виявлених.

Коди, які не тільки виявляють помилки, але й виправляють їх, називаються кодами з виправленням помилок. Такі коди підвищують правильність передавання з помилками і тому їх називають завадостійкими. Таким чином, завадостійке кодування полягає в накладенні на передавані повідомлення деяких додаткових умов через уведення надмірної інформації. У процесі декодування перевіряється виконання цих умов. Якщо помилок немає, умови виконуються і прийнята інформація вважається правильною.

Отже, виправлення або виявлення помилок завадостійким кодом досягається застосуванням у ньому більшої кількості елементів n , ніж це треба для передавання інформації k . Дійсно, тільки в цьому разі можна множини із 2^n можливих кодових комбінацій на виході дискретного каналу розкласти на 2^k непорожніх підмножин M_k і вибрати для кожної із них свого представника (A_i).

Передбачалось, що розбиття на підмножини M_k здійснюється випадково. Очевидно, що від способу цього розбиття залежать потенційні коригувальні здатності коду. Для вирішення питання про доцільність розбиття на підмножини M_k використаємо наочне геометричне подання.

Цифрові методи кодування та перетворення дискретних символів в електричні (фізичні) сигнали ґрунтуються переважно на теорії двійкового кодування.

У теорії двійкового кодування основну роль відіграє лінійний арифметичний простір $GF(2^n)$ над полем із двох елементів $GF(2) = \{0, 1\}$. Елементи цього поля нулів і одиниць складаються і перемножуються за такими правилами: $0 \oplus 0 = 0$, $0 \oplus 1 = 1$, $1 \oplus 1 = 0$ (складання за модулем два – mod 2), $0 \times 0 = 0$, $0 \times 1 = 0$, $1 \times 1 = 1$. Відповідно до визначення простір $GF(2^n)$ складається із 2^n наборів, які, у свою чергу, складаються із нулів та одиниць.

Таким чином, кожна n -розрядна двійкова кодова комбінація являє собою вектор у просторі $GF(2^n)$. Хемінг визначив відстань $d(X, Y)$ у просторі $GF(2^n)$ як кількість місць, у яких координати векторів не збігаються. Легко перевірити, що уведена так відстань задовольняє три аксіоми метрики. Цю метрику лінійного простору $GF(2^n)$ називають метрикою Хемінга. Відстань Хемінга $d(X, Y)$ – це природна міра розрізнення двох n -значних кодових комбінацій, зображених векторами X і Y . У теорії кодування важливу роль відіграє число $d(0, X) = w(X)$, яке отримало назву «вага кодової комбінації X ». Отже, вага $w(X)$ дорівнює кількості відмінних від нуля координат вектора X .

Відстань Хемінга узгоджена з операцією складання векторів у просторі $GF(2^n)$, тобто $d(X + Z, Y + Z) = d(X, Y)$ для будь-яких трьох векторів $X, Y, Z \in GF(2)^n$. На основі цієї властивості легко зробити висновок, що $d(X, Y) = w(X + Y)$ для будь-яких векторів.

Для наочності корисно подати 2^n векторів простору $GF(2^n)$ у вигляді вершин стандартного n -мірного куба з ребром, що дорівнює одиниці, кожна вершина якого задається одним з можливих наборів із нулів та одиниць. Будь-який конкретний код є тоді деякою підмножиною множини вершин куба. Модель тривимірного куба для трирозрядного двійкового коду, кодові комбінації якого – 000, 001, 010, 011, 100, 101, 110, 111 – показано на рис. 14.2. У цій геометричній моделі відстань Хемінга між векторами X і Y

інтерпретується як довжина найкоротшого шляху, який проходить по ребрах куба і з'єднує відповідну вершину з початком координат.

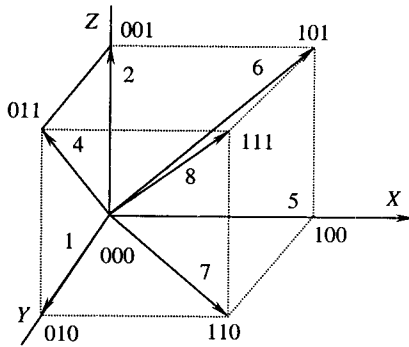


Рис. 14.2. Геометрична модель трирозрядного коду

14.3. Принципи виявлення і виправлення помилок надлишковими кодами – завадостійким кодуванням

Характеристикою помилок, які виникають у дискретному каналі, є вектор-помилка \vec{E} , відображений у вигляді двійкового набору тієї самої порядності, що й кодові інформаційні комбінації. Нуль в \vec{E} означає, що помилки у відповідних позиціях не виникають, а одиниці займають позиції, які спотворюються під час передавання, так що кількість одиниць в \vec{E} дорівнює кратності помилок q . Кодову комбінацію \vec{A}^* отримано складанням переданої комбінації \vec{A} з вектором помилок за модулем два: $\vec{A}^* = \vec{A} + \vec{E}$.

Під дією помилок кінці кодових векторів будуть зміщуватися в інші точки. Для характеристики положення цих «помилкових» точок використаємо відстань Хемінга між вершинами куба. Заміна в кодовій комбінації одиниці нулем або навпаки змінює цю відстань на одиницю, заміна одиниці нулем змінює відстань на дві одиниці і т. д.

Для того щоб у разі виявлення помилок одна з дозволених кодових комбінацій не перетворилась у другу дозволена комбінація, вона має бути тим більша, чим вища кратність виявлюваних помилок. Оскільки відстані між дозволеними кодовими комбінаціями неоднакові, найбільш критичною щодо коригувальних властивостей коду буде найменша для цього коду відстань. Таку відстань (d_0) називають кодовою відстанню за Хемінгом.

Аналогічно у випадку кодів з виправленням помилок кодова відстань визначає виправну здатність коду. Дійсно, чим більша d_0 , тим більше прийнята кодова комбінація буде подібною до переданої, ніж до інших дозволених комбінацій.

Таким чином, ми маємо справу з дозволеними і забороненими кодівими комбінаціями.

Розглянемо простий код Бодо – п’ятирозрядний код, який використовують для кодування у двійковій системі числення 32 символи (1-й символ – 00000, 2-й символ – 00001, ..., 32-й символ – 11111). Усі кодові комбінації – дозволені. Нехай послідовні комбінації 00010, 00011, 00100 відображають відповідно $(i - 1)$, i та $(i + 1)$ символи. Якщо в $(i - 1)$ комбінації останню цифру 0 замінити на цифру 1, тоді $(i - 1)$ кодова комбінація перетвориться в i -ту кодову комбінацію і навпаки, якщо останню цифру в i -й кодовій комбінації 1 замінити на 0, тоді i -та кодова комбінація перетвориться в $(i - 1)$ комбінацію. А тепер, якщо розглядати перетворення $(i + 1)$ або в i -ту, або в $(i - 1)$ комбінацію, то треба виконати декілька змін цифр, щоб це перетворення відбулося.

Отже, розглядуваний код має тільки дозволені кодові комбінації, тим більше, як уже з’ясовано, найменша кількість перетворень однієї комбінації в іншу – 1 (перехід i -ї в $(i - 1)$ комбінації і навпаки), тобто це буде мінімальна відстань між кодовими комбінаціями.

Розглянутий приклад демонструє можливі переходи однієї дозволеної комбінації в іншу, причому для перетворення i -ї в $(i - 1)$ і навпаки за однієї зміни цифри в кодовій комбінації, а для перетворення $(i + 1)$ в i -ту і навпаки, потрібно виконати три зміни цифр, тобто відстань між кодовими комбінаціями буде дорівнювати трьом.

За таких змін кодових комбінацій після декодування один закодований символ буде відображати другий.

Як відомо, цифри в кодових комбінаціях змінювалися через вплив завад на канали передавання даних.

У цьому прикладі відстань між $(i - 1)$ та i -ю кодовими комбінаціями дорівнює одиниці, оскільки перехід однієї комбінації у другу здійснюється зміною лише однієї позиції комбінації, а перехід i -ї у $(i + 1)$ і навпаки виконується зміною трьох позицій. Відстань A_i – кодову комбінацію від $A_{(i-1)}$ і $A_{(i+1)}$ знаходять за правилом складання за модулем два:

$$\begin{array}{r}
 A_{i-1} = 00010 \\
 \oplus \\
 A_i = 00011 \\
 \hline
 00001 \Rightarrow d = 1;
 \end{array}
 \qquad
 \begin{array}{r}
 A_i = 00011 \\
 \oplus \\
 A_{i+1} = 00101 \\
 \hline
 00111 \Rightarrow d = 3.
 \end{array}$$

Очевидно, що одна кодова комбінація на другу змінюється стільки разів, скільки зміниться позицій кодової комбінації.

У цьому прикладі кодування виконувалось для $M_0 = 32$ символів (букв), або $M_0 = 2^{n=5} = 32$, де кожній букві відповідала відповідна кодова комбінація, довжина якої п'ять розрядів, або всі можливі кодові комбінації призначаються для того чи того символу, тобто немає жодної вільної кодової комбінації.

Розширимо поле кодових комбінацій для 32 символів (букв). Для прикладу візьмемо довжину кодової комбінації $n = 7$, тоді поле кодових комбінацій $M = 2^{n=7} = 128$. Різниця $M - M_0$ – це кількість надлишкових комбінацій. Серед 128 комбінацій потрібно вибрати 32 комбінації і призначити кожній з них відповідний символ. Ці кодові комбінації називають дозволеними, а ті, що залишились, – забороненими. Серед 32 семирозрядних кодових комбінацій обов'язково може бути виявлена одна або декілька таких пар комбінацій, для яких відстань між ними буде мати мінімальне значення – d_0 .

Для пояснення використання цього розбиття множини M на підмножини M_0 та $M - M_0$ розглянемо загальний випадок (рис. 14.3).

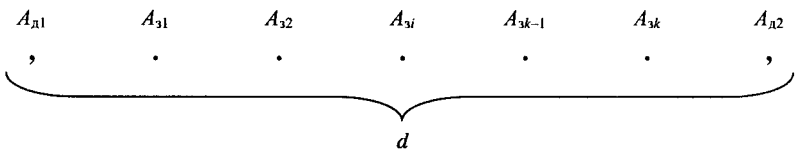


Рис. 14.3. Виявлення умов переходу однієї дозволеної комбінації A_{d1} у A_{d2} та навпаки: A_{d1} і A_{d2} – дозволені кодові комбінації; $A_{31}, A_{32}, \dots, A_{3i}, \dots, A_{3k-1}, A_{3k}$ – заборонені комбінації

Очевидно, що перехід A_{d1} (A_{d2}) у будь-яку A_{3i} із множини заборонених буде виявлюваний і зареєстрований як помилка. Якщо перехід A_{d1} у A_{d2} і навпаки буде здійснений з кодовою відстанню d_0 , то цей перехід не виявляється, звідси умова виявлення помилки буде

$$q \leq d_0 - 1,$$

де q – кратність помилок.

Для доведення виявлюваної та виправної здатності коду розглянемо співвідношення між d_0 і q . Нехай передавалась кодова комбінація A_i , а прийнятий вектор – Y , і при цьому відбулася t -кратна помилка, тобто $d(A_i, Y) = t$ (рис. 14.4).

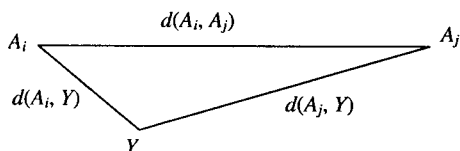


Рис. 14.4. До визначення коригувальної здатності коду

Позначимо через A_j будь-яку другу кодову комбінацію цього коду, тоді з нерівності трикутника випливає, що

$$d(A_j, Y) = d(A_i, A_j) - d(A_i, Y) = d(A_i, A_j) - t. \quad (14.1)$$

Якщо для будь-якої комбінації коду $d(A_j, Y) > 0$, тоді помилку (перехід A_i в Y) буде завжди виявлено. Отже, для виявлення всіх t -кратних помилок потрібно і достатньо, щоб $d(A_i, A_j) - t > 0$, тобто для будь-яких A_i і A_j $d(A_i, A_j) > t$ і, зокрема,

$$d_0 > t. \quad (14.2)$$

Для виправлення всіх t -кратних помилок потрібно і достатньо, щоб

$$d(A_j, Y) > d(A_i, Y) = t, \quad (i \neq j) \quad (14.3)$$

Із рівнянь (14.1) і (14.3) випливає $d(A_i, A_j) > 2t$ і, зокрема,

$$d_0 > 2t. \quad (14.4)$$

Якщо кратність виправних помилок q , а виявлюваних r , то, щоб виправити всі q -кратні помилки, має бути $d_0 > 2q$, а щоб виявити всі r -кратні помилки – $d_0 > r$. Отже, код, який може виправляти помилки кратністю q , може виявляти помилки кратністю $r = 2q$. Ураховуючи, що t і d_0 – цілі числа, нерівності (14.2) і (14.4) можна переписати у вигляді:

$$d_0 = r + 1; \quad (14.5)$$

$$d_0 = 2q + 1. \quad (14.6)$$

Аналогічно можна показати, що для одночасного виправлення помилок кратністю q і виявлення помилок кратністю $2r$ потрібно і достатньо, щоб

$$d_0 = q + r + 1, \text{ якщо } r \geq q.$$

У безнадлишкових кодах $d_0 = 1$, тому такі коди не мають здатності виявляти і виправляти помилки.

Формули, які установлюють залежність між кодовими відстанями і кратністю виправних і виявлюваних помилок, мають просту геометричну інтерпретацію. Візьмемо довільну точку множини, яка належить M_k , і розглянемо навколо неї кулю радіусом t , тобто множину точок із M_k , які знаходяться від розглядуваної точки на відстані не більше ніж t . Очевидно, що під дінням будь-якого вектора помилок ваги $r \leq t$ розглядувана

точка не вийде за межі цієї кулі. А оскільки за умови $d_0 \geq t + 1$, тобто всередині цієї кулі немає інших кодових точок (дозволених комбінацій), ця умова забезпечує виявлення помилок. Цю умову ілюструє модель (рис. 14.5) на площині, де кулі замінюються колами (на рисунку показано частини цих кіл).

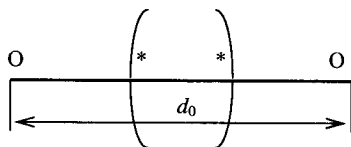


Рис. 14.5. До умови виявлення помилок кодом ($t = 2, d_0 = 3$):
 O – дозволена кодова комбінація; * – заборонена кодова комбінація

Якщо використати такий самий код для виправлення t і менше помилок, тоді декодування має виконуватись за таблицею, у якій під кожною кодовою комбінацією записано всі точки на кулі радіусом t із центром у цій комбінації. Щоб кожна з вписаних точок трапилась у таблиці один раз, відстань між двома кодовими комбінаціями має бути не менше $(2t + 1)$. Інакше кажучи, кожен кодову комбінацію немовби обведено захисною кулею, причому ці кулі не перетинаються. Будь-яка із прийнятих кодових комбінацій попадає в одну і тільки в одну із цих куль і декодується у центр тієї кулі, якій вона належить.

Як видно, із рівнянь (14.5) і (14.6) не випливає, що помилки більшої кратності, ніж r і q , не будуть виявлятися і виправлятися, оскільки d_0 – це мінімальна відстань між кодовими комбінаціями, а відстань між окремими комбінаціями більша, ніж d_0 . Однак відсоток виявлених і виправлених помилок кратністю вище r і q порівняно невеликий.

Отже, завадостійке кодування повідомлень має забезпечувати з достатньо високою імовірністю виявлення факту наявності помилок у прийнятих повідомленнях або виявлення і виправлення цих помилок.

14.4. Загальна характеристика завдання завадостійкого кодування

Розглянемо загальну постановку завдання завадостійкого кодування за обмеженої довжини кодованих повідомлень. Від джерела повідомлень надходять послідовності символів b_1, b_2, \dots, b_k , які називаються інформаційними послідовностями. Кодування кожної послідовності b_1, b_2, \dots, b_k приводить до взаємно-однозначної відповідності передаване повідомлення (кодова послідовність) $\beta_1, \beta_2, \dots, \beta_n$. Унаслідок впливу завад на виході каналу буде прийнято повідомлення $\beta'_1, \beta'_2, \dots, \beta'_n$, частина сим-

волів якого може відрізнитись від символів передаваного повідомлення: $\beta'_v \neq \beta_v$. Якщо під час приймання кожний символ обов'язково ідентифікується з одним із символів алфавіту каналу B (канал без стирання), тобто завжди $\beta'_v \in B$, то єдиною формою помилки може бути трансформація символів.

Завадостійке кодування повідомлень як методологія підвищення правильності передавання повідомлень має забезпечити з досить високою імовірністю або виявлення факту наявності помилок у прийнятих повідомленнях, або виявлення і виправлення цих помилок. Відповідно розрізняють коди з виявленням помилок і коди з виправленням помилок.

Коди, які виявляють і виправляють помилки, називаються завадостійкими або коригувальними. Коригувальна здатність коду ґрунтується на його надлишковості. Якщо немає надлишковості, коли будь-яка комбінація символів алфавіту каналу може являти собою передаване повідомлення, достовірне виявлення помилок неможливе: на приймальному кінці невідомо, яке саме повідомлення могло передаватися (інакше його передавання не несло б інформації і було б позбавлено змісту).

Для того щоб надлишковістю можна було скористатись для виявлення або виправлення помилок, її потрібно вносити за певними правилами.

Розглядаючи коригувальну здатність кодів, будемо враховувати лише надлишковість, унесену в повідомлення під час кодування, а повідомлення, що надходять від джерела, вважатимемо безнадлишковими.

За формою внесення надлишковості розрізняють систематичні та несистематичні коди.

У систематичних кодах символи інформаційної послідовності входять без змін у передаване повідомлення, займаючи у ньому відведені їм позиції, а кодування зводиться до внесення у передаване повідомлення додаткових (надлишкових) символів, які зв'язані певною залежністю із символами інформаційної послідовності.

У несистематичних кодах символи інформаційної послідовності у явному вигляді в передаване повідомлення не входять, а можуть бути установлені за відомими залежностями, які зв'язують їх із символами цього повідомлення.

За способом кодування розрізняють блокові та неперервні коди.

Під час блокового кодування інформаційна послідовність розбивається на блоки символів (b_1, b_2, \dots, b_k) фіксованої довжини k , кожному з яких ставиться у відповідність певна комбінація символів алфавіту каналу ($\beta_1, \beta_2, \dots, \beta_n$), яку називають кодовим словом. Код під час блокового кодування визначає закон формування кодового слова, який відповідає цьому блоку інформаційних символів. Для цифрової інформації такими блоками зручно реалізувати передавані числа. У випадку систематичного

блокового коду кодове слово, що відповідає блоку інформаційних символів (b_1, b_2, \dots, b_k) , можна подати у вигляді $b_1, b_2, \dots, b_k, \beta_{k+1}, \dots, \beta_n$, де $\beta_{k+1}, \dots, \beta_n$ – надлишкові символи. Блоковий код, який ставиться у відповідність блоку інформаційних символів довжиною K і має довжину n , будемо позначати через n, k . Надлишкові символи не несуть додаткової інформації про джерело повідомлення (вони однозначно визначаються інформаційними символами b_1, b_2, \dots, b_k). Тому кодове слово несе таку саму кількість корисної інформації, що і відповідний блок інформаційних символів. У разі безнадлишковості вхідного повідомлення блок із k інформаційних символів несе кількість інформації $H = k \log_2 m$, де m – обсяг алфавіту джерела.

Максимальна кількість інформації, яку може вміщувати слово із n символів каналу за такого самого обсягу алфавіту m , дорівнює $H_{\max} = n \log_2 m$.

Тому надлишковість коду (n, k) :

$$\rho_H = 1 - \frac{H}{H_{\max}} = 1 - \frac{k \log_2 m}{n \log_2 m} = \frac{n - k}{n} = \frac{r}{n},$$

де $r = n - k$ – кількість надлишкових символів, які вносяться під час кодування.

За неперервного кодування кожний символ передаваного повідомлення визначають за рекурентними співвідношеннями, які зв'язують його з відповідними символами інформаційної послідовності:

$$\beta_{ij} = \left[\sum_{v=-q}^{3-1} C_{vj} b_{i+v} \right] (\text{mod } m), \quad j = 0, 1, \dots, l-1. \quad (14.7)$$

Значення правої частини рівняння (14.7) визначають за модулем m , який позначає, що після виконання звичних обчислень утримується лише залишок від ділення отриманого значення на основу коду m (наприклад, 9 за модулем 4 дає 1). Виконання обчислювань у рівнянні (14.7) за модулем m потрібне для того, щоб отримані значення символів β_{ij} належали алфавіту каналу, тобто набули одне із значень $0, 1, \dots, m - 1$.

Рекурентне співвідношення (14.7) на i -му кроці кодування групі символів вхідної послідовності $b_{i-g}, \dots, b_i, \dots, b_{i+s-1}$ ставить у відповідність l символів кодової послідовності $\beta_{i0}, \dots, \beta_i, l-1$; на $(i+1)$ -му кроці групі символів $b_{i-g+1}, \dots, b_{i+1}, \dots, b_{i+s}$, зсунутої на одну позицію, становить у відповідність наступні l символів кодової послідовності $\beta_{i+1,0}, \dots, \beta_{i+1}, l-1$ і т. д. Отже, на кожний символ інформаційної послідовності припадає l символів передаваної кодової послідовності, що відповідає надлишковості коду $\rho_H = (l-1)l$. Праві частини рекурентних співвідношень (14.7) яв-

ляють собою ніби згортку відповідної ділянки інформаційної послідовності, тому коди і називають згортковими.

Згорткові (неперервні) коди можуть бути як систематичними, так і несистематичними. Систематичні згорткові коди отримуються в окремому випадку, коли для одного із значень індексу j (наприклад, $j = 0$) коефіцієнти у формулі (14.7) набувають значення

$$c_{v0} = \begin{cases} 0, & \text{якщо } v \neq 0; \\ 1, & \text{якщо } v = 0, \end{cases}$$

відповідно $\beta_{j0} = b_j$. У цьому разі передавана кодова послідовність має такий вигляд, як наведено в дужках $(\dots, b_i, \beta_{i1}, \dots, \beta_{il-1}, b_{i+1}, \beta_{i+1,1}, \dots, \beta_{i+1, l-1}, \dots)$.

14.5. Класифікація завадостійкості кодів

У попередньому розділі визначено способи використання надмірності для підвищення достовірності передавання дискретної інформації, які і визначають напрями створення завадостійких кодів. Теорія завадостійкого кодування визначає методи побудови алгоритмів, які описують математичні моделі, що являють собою відповідні коди.

Дотепер розроблено багато різних завадостійких кодів, які відрізняються один від одного основою q , відстанню d , надмірністю, структурою, функціональним призначенням, енергетичною ефективністю, кореляційними властивостями, алгоритмами кодування і декодування, формою частотного спектра. Щоб краще розуміти і використовувати на практиці завадостійкі коди, потрібно їх класифікувати, що дозволить розрізняти різні коди за характерними ознаками. Типи кодів, які розрізняються за особливостями структури, функціональним призначенням, фізичними властивостями коду як сигналу, показано на рис. 14.6.

Як видно, всі надмірні коди поділено на два класи: блокові і неперервні. У блокових кодах передавана інформаційна послідовність розбивається на окремі блоки – кодові комбінації, які кодуються і декодуються незалежно один від одного. У неперервних кодах, які ще називають рекурентними згортковими, або ланцюжковими, передавана інформаційна послідовність не розділюється на блоки, а перевіріні елементи розміщуються у визначеному порядку між інформаційними. Процеси кодування і декодування також мають неперервний характер.

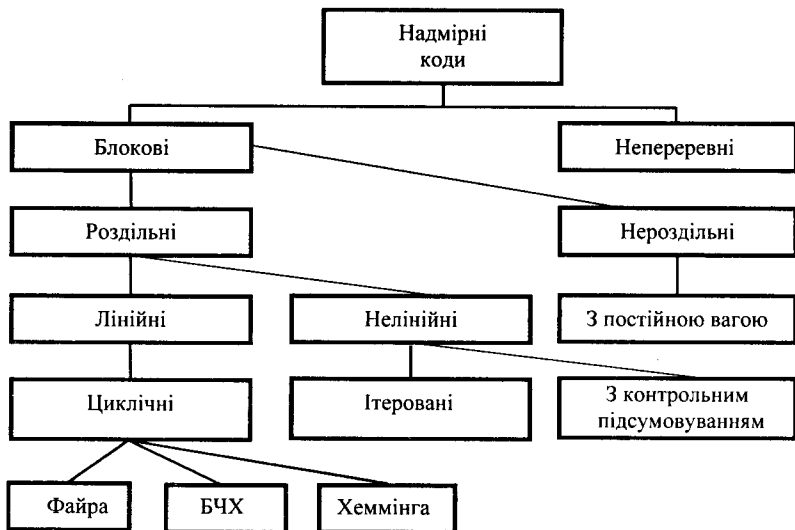


Рис. 14.6. Типи завадостійких кодів

Блокові коди ділять на роздільні та нероздільні. У роздільних кодах інформаційні елементи і перевірні елементи у всіх кодових комбінаціях займають одні й ті самі позиції. Тому роздільні зазвичай позначають як коди (n, k) , де n – розрядність коду; k – кількість інформаційних елементів. У нероздільних кодах немає ділення на інформаційні та перевірні елементи. До них належать коди з постійною вагою, усі комбінації яких мають одну вагу, і, зокрема, рекомендований МККТТ телеграфний код $N3$ – семирозрядний код «3 із 7» з постійною вагою, яка дорівнює 3. Ці коди можуть слугувати тільки для виявлення помилок змінуванням ваги.

Роздільні коди поділяють на лінійні (систематичні) і нелінійні (несистематичні). Лінійні коди так називають тому, що їх перевірні елементи являють собою лінійні комбінації інформаційних елементів. Вони утворюють велику групу кодів і широко застосовуються на практиці. Велику й особливу підгрупу лінійних кодів утворюють циклічні коди.

Нелінійні коди використовують значно рідше, ніж лінійні. Прикладом нелінійного коду є код з контрольним підсумовуванням, у якому перевірні розряди – сума одиниць у кодовій комбінації, зокрема ітерований код.

Більшість розглянутих кодів можна використовувати як для виявлення, так і для виправлення помилок. Однак треба зважити на те, що будь-який надмірний код завжди має більшу надмірність, ніж потрібно для виявлення і виправлення помилок, тому що надмірність мають усі кодові комбінації, а помилки виникають тільки в деяких з них. А оскільки надмірність, яка

потрібна для виявлення помилок, менша від надмірності, потрібної для їх виправлення, то для уникнення зайвої надмірності часто обмежуються використанням кодів з виявленням помилок, а підвищення правильності виконується використанням оберненого вирішального зв'язку.

Для підвищення правильності передавання дискретної інформації коди використовують у таких режимах:

- виправлення помилок;
- виявлення помилок;
- виправлення і виявлення помилок.

Режим виправлення помилок у СПДІ використовують дуже рідко. Його застосовують тоді, коли обернений зв'язок неможливий або недоцільний, наприклад через жорсткі вимоги до допустимої затримки чи великого часу поширення сигналів.

Інколи режим виправлення помилок використовують для передавання каналами з незалежними помилками або з короткими пакетами помилок. Найчастіше в цьому разі застосовують згорткові, ітеративні або каскадні коди. Оскільки в більшості реальних каналів спостерігаються пакети помилок довжиною у декілька десятків і навіть сотень символів, тоді для їх виправлення знадобились би коди, довжина яких відповідає кодовій комбінації, яка вимірюється тисячами розрядів, що важко реалізувати технічно. Тому в переважній більшості існуючих і проєктованих СПДІ підвищення правильності базується на виявленні помилок і повторному передаванні за допомогою оберненого зв'язку кодових комбінацій, у яких були виявлені помилки. Інколи виявляється ефективним виправлення помилок малої кратності з виявленням решти помилок.

Для виявлення помилок здебільшого використовують коди двох типів: циклічні й ітеративні (матричні), що зумовлено простотою апаратурної або програмної реалізації кодування і декодування.

14.6. Лінійні (систематичні) коди. Коди Хеммінга

Математична теорія завадостійкого кодування може, зокрема, ґрунтуватися на теорії лінійних векторних просторів. При цьому загальна теорія будується для випадків, коли основа коду m – просте число, тобто ділиться лише само на себе і на одиницю, а число розрядів n фіксоване (рівномірний код).

Використовуючи теорію лінійних векторних просторів, кодове слово b довжиною n розглядають як вектор у лінійному n -мірному просторі кодових слів R_n , а символи b_0, b_1, \dots, b_{n-1} – як його компоненти.

Лінійним називають блоковий рівномірний код, для якого будь-яка лінійна композиція кодових слів (двох або більше), які належать множині дозволених слів B_d , дає дозволене кодове слово, тобто для будь-яких векторів $b^{(i)}, b^{(j)} \in B_d$ і будь-яких значень символів λ_i, λ_j , які належать алфавіту каналу, справедлива умова

$$b^{(k)} = \lambda_i b^{(i)} + \lambda_j b^{(j)} \in B_d. \quad (14.8)$$

Уточнимо операції множення та складання символів за лінійних перетворень кодових слів. Для того щоб компоненти результуючого вектора $b^{(k)}$ склались тільки із символів каналу $0, 1, \dots, m-1$ (це потрібно для того, щоб $b^{(k)} \in B_d$), операції множення та складання компонент векторів $b^{(i)}$ і $b^{(j)}$ для лінійного перетворення (14.8) визначають за модулем m . Наприклад, використовуючи операції з трійковими символами $0, 1, 2$, отримаємо

$$(2+1)(\text{mod } 3) = 0; (2+2)(\text{mod } 3) = 1; (2 \cdot 2)(\text{mod } 3) = 1.$$

Виконання операцій множення і складання символів коду за модулем m надає їм властивість замкненості: у результаті цих операцій отримують символи, які належать початковій множині. Операції множення і складання за модулем m зберігають асоціативність, дистрибутивність і комутативність, властиві звичайним операціям множення і складання. У лінійній алгебрі множину, яка задовольняє умову замкненості відносно двох заданих операцій і яка має властивості асоціативності, дистрибутивності і комутативності, називають полем Галуа і позначають $GF(m)$. Той факт, що у використовуваних перетвореннях операціями (14.8) символи коду утворюють поле Галуа, дозволяє говорити (у термінах лінійної алгебри), що лінійний простір векторів кодових слів R_n будується над полем Галуа $GF(m)$.

У просторі R_n кодових слів вводиться також скалярний добуток векторів a і b :

$$(ab) = \left[\sum_{v=0}^{n-1} a_v b_v \right] (\text{mod } m).$$

Вектори (колові слова), для яких $(ab) = 0$, називаються ортогональними.

Подання передавальних і розпізнавальних кодових слів у лінійному векторному просторі R_n дозволяє математично описувати спотворення кодових слів вектором $e \in R_n$, що задовольняє умову

$$b' = b + e,$$

де b і b' – вектори передавального та приймального (розпізнаваного) кодових слів.

14.7. Властивості та завдання побудови лінійних (систематичних) кодів

Лінійним кодом (n, k) називають лінійний k -вимірний підпростір простору $GF(2^n)$. Ураховуючи, що поле $GF(2)$ складається із двох елементів – 0 і 1, приходимо до еквівалентного формулювання: двійковий код називається лінійним, якщо сума двох будь-яких кодових слів є знову кодове слово. Множина кодових векторів двійкового лінійного коду одночасно задовольняє і аксіоми групи зі складання векторів, тобто утворює кінцеву комутативну групу. Тому лінійні двійкові коди інколи називають груповими кодами.

Якщо джерело інформації виробляє двійкові інформаційні послідовності в деяку підмножину із 2^k кодових слів довжиною n , то розмірність підпростору коду (n, k) дорівнює k , тоді існують такі k лінійно незалежних векторів (слів) a_1, \dots, a_k , що 2^k лінійних комбінацій $\alpha_1 a_1 + \dots + \alpha_k a_k$, де $\alpha_i = 0, 1$, утворюють код. Вектори a_1, \dots, a_k називають твірними елементами коду (n, k) . Як твірні елементи можна взяти будь-яку множину із k векторів коду, сума будь-яких i з яких не дорівнює нулю ($i = 1, 2, \dots, k$).

Розглянемо властивості лінійних кодів.

1. Код, який вміщує $M = m^k$ кодових слів у дозволеному наборі, може бути заданий k лінійно-незалежними векторами $b^{(i)}$, ($i = 1, 2, \dots, k$), які складають базисну матрицю L розміром $k \times n$:

$$L = \begin{bmatrix} b^{(1)} \\ \dots \\ b^{(k)} \end{bmatrix} = \begin{bmatrix} \beta_0^{(1)} & \beta_1^{(1)} & \dots & \beta_{n-1}^{(1)} \\ \dots & \dots & \dots & \dots \\ \beta_0^{(k)} & \beta_1^{(k)} & \dots & \beta_{n-1}^{(k)} \end{bmatrix}.$$

Множина дозволених кодових слів, визначених базисною матрицею L , позначимо через B^L . Будь-яке кодове слово з дозволеного набору B^L , визначеного матрицею L , можна подати у вигляді лінійної композиції вхідних до неї векторів:

$$b^{(j)} = \sum_{i=1}^k \lambda_i b^{(i)}, \quad \lambda_i \in GF(m). \quad (14.9)$$

Легко бачити, що загальна кількість дозволених кодових слів, утворених за законом (14.9), дорівнює кількості дозволених комбінацій з k коефіцієнтів λ_i , які набувають одного значення з $0, 1, \dots, m-1$, тобто закон (14.9) забезпечує утворення заданого числа $M = m^k$ різних кодових слів довжиною n . Таким чином, базисна матриця розміром $k \times n$ задає лінійний код (n, k) .

Простий код є лінійним кодом (n, n) , базисна матриця якого включає повний набір n -вимірних лінійно-незалежних векторів (кількість лінійно-незалежних векторів у n -мірному просторі не може перевищувати n).

2. Множині дозволених кодів слів B^L , визначеній матрицею L , відповідає множина B^H кодів слів, ортогональних кодів слів дозволеного набору: якщо $b^{(i)} \in B^L$, а $b^{(j)} \in B^H$, тоді $(b^{(i)}b^{(j)}) = 0$. Множину B^H можна задати $n - k$ лінійно-незалежними векторами, які є складовими H розміром $(n - k) \times n$:

$$H = \begin{bmatrix} b^{(k+1)} \\ \dots \\ b^{(n)} \end{bmatrix} = \begin{bmatrix} \beta_0^{(k+1)} & \beta_1^{(k+1)} & \dots & \beta_{n-1}^{(k+1)} \\ \dots & \dots & \dots & \dots \\ \beta_0^{(n)} & \beta_1^{(n)} & \dots & \beta_{n-1}^{(n)} \end{bmatrix}.$$

Матриця H називається перевіркою. Множина B^H вміщує $m^{(n-k)}$ кодів слів, що відповідають усім можливим лінійним композиціям векторів, що складають матрицю H . Властивості множин B^L і B^H взаємні: множина B^H також являє собою лінійний код, для якого матриця H є базисною, а матриця L – перевіркою.

Перевірна матриця H так само, як і базисна матриця L повністю визначає лінійний код. Вона дозволяє подавати лінійний код (n, k) у систематичній формі:

$$b^{(i)} = (b_0^{(i)}, b_1^{(i)}, \dots, b_{k-1}^{(i)}, \beta_k^{(i)}, \dots, \beta_{n-1}^{(i)}),$$

де перші k символів $b_0^{(i)}, b_1^{(i)}, \dots, b_{k-1}^{(i)}$ – інформаційні, а решта $n - k$ символів $\beta_k^{(i)}, \dots, \beta_{n-1}^{(i)}$ – надмірні, які забезпечують можливість виявлення та виправлення помилок (їх ще називають перевірними символами).

Перевірні символи дозволеного кодового слова $b^{(i)}$ за заданих інформаційних символів $b_0^{(i)}, b_1^{(i)}, \dots, b_{k-1}^{(i)}$ можна визначити із $n - k$ лінійних рівнянь, визначених умовами ортогональності дозволеного кодового слова кодів слів до перевірки матриці H :

$$S_\nu = (b^{(i)}b^{(k+\nu)}) = 0, \quad \nu = 1, 2, \dots, n - k. \quad (14.10)$$

Рівняння (14.10) можна переписати у вигляді

$$\sum_{i=k}^{n-1} b_j^{(k+\nu)} \beta_j^{(i)} = - \sum_{j=0}^{k-1} b_j^{(i)} b_j^{(k+\nu)}, \quad \nu = 1, 2, \dots, n - k. \quad (14.11)$$

Розв'язавши систему $n - k$ рівнянь (14.11) відносно $n - k$ перевірних символів $\beta_j^{(i)}$, отримаємо для них лінійні вирази вигляду

$$\beta_j^{(i)} = \sum_{i=0}^{k-1} c_{ij} b_i^{(i)}, \quad j = k, k+1, \dots, n-1. \quad (14.12)$$

Формула (14.12) дозволяє визначити $n - k$ перевірних символів $\beta_k^{(i)}, \beta_{k+1}^{(i)}, \dots, \beta_{n-1}^{(i)}$ дозволеного кодового слова за відомим k інформаційним символом (коефіцієнти c_{ij} , які визначають закон формування перевірних символів, виражаються через елементи $b_j^{(k+v)}$ перевірної матриці H).

Формула (14.12) – дуже зручна форма подання регулярного лінійного коду (n, k) , який широко використовують у практиці зв'язку. При цьому для утворення дозволених кодових слів за відомим інформаційним символом достатньо зберігати в пам'яті $k \times (n - k)$ значень коефіцієнтів c_{ij} , що входять до рівняння (14.12). Оскільки у разі задання коду з основою m для кожного з цих коефіцієнтів можна вибрати одне з m значень, тоді всього можна вибрати $m^{k(n-k)}$ варіантів коефіцієнтів у виразі (14.12) і побудувати $m^{k(n-k)}$ різних лінійних кодів (n, k) . Завдання зводиться до того, щоб вибрати оптимальний варіант коду.

Критерій наявності помилок у прийнятому слові b' – порушення умов його ортогональності векторів $b^{(k+v)}$ перевірної матриці. Результати перевірки дотримання умов ортогональності дають $n - k$ значень скалярних добутоків

$$S_v = (b' b^{(k+v)}) = \sum_{j=0}^{n-1} b'_j b_j^{(k+v)}, \quad v = 1, 2, \dots, n - k, \quad (14.13)$$

які можна розглядати як компоненти деякого вектора S , що називається синдромом перевірки. Для неспотворених кодових слів дозволеного набору, які задовольняють умови ортогональності векторів $b^{(k+v)}$,

$$S_v = 0, \quad v = 1, 2, \dots, n - k.$$

Таким чином, синдром S являє собою нульовий вектор, усі компоненти якого дорівнюють нулю. За наявності спотворень синдром дозволяє встановити їх характер і виконати виправлення деяких видів помилок.

Для двійкових кодів скалярний добуток (14.13) вироджується в так звані контрольні суми – суми за модулем два символів двійкового коду b' , що відповідають ненулевим позиціям v -го рядка перевірної матриці. При цьому рівність нулю синдрому забезпечується у випадку парного числа одиниць у коді b' на позиціях, які входять у кожен із контрольних сум. Тому для двійкового коду перевірку ортогональності (14.13) називають перевіркою парності.

Для прикладу розглянемо дуже поширений клас лінійних кодів – двійкові коди Хеммінга. У двійковому коді Хеммінга для виявлення і виправлення помилок визначаються контрольні суми за модулем два, кожна з яких включає деяку кількість інформаційних символів і один із перевірних.

Виявлюваний код Хеммінга вміщує один перевірний символ (у молодшому розряді), який дорівнює сумі за модулем два всіх інформаційних символів:

$$\beta_{n-1} = \sum_{j=0}^{n-2} b_j \pmod{2}. \quad (14.14)$$

Формула (14.14) є окремим випадком формули (14.12), якщо $m = 2$, $k = n - 1$ і $c_{\mu} = 1$, $\mu = 0, 1, \dots, n - 2$. Єдина контрольна сума цього коду – це сума за модулем два всіх символів, до яких і включається перевірний:

$$\left[\sum_{j=0}^{n-2} b_j + \beta_{n-1} \right] \pmod{2} = 0. \quad (14.15)$$

Формула (14.15) є окремим випадком формули (14.11), якщо $k = n - 1$, а всі складові єдиного вектора, що входять до перевірної матриці, дорівнюють одиниці: $b_j^{(n)} = 1$, $j = 0, 1, \dots, n - 1$. Очевидно, що якщо спотворень у каналі немає, контрольна сума (14.15) для розпізнаного кодового слова дорівнює нулю, а в разі будь-якого одиночного спотворення (або в разі непарного числа спотворених символів) дорівнює одиниці. Відмінна від нуля контрольна сума виявляє факт (але не місце) одиночного або непарного числа спотворень сигналів прийнятого кодового слова. У разі парного числа спотворень останні не виявляються. Отже, цей код достовірно виявляє лише одиночні помилки. Кратні помилки виявляються за непарної їх кількості. Код Хеммінга, що виявляє помилки, має кодову відстань $d_{\min} = 2$, найближчу до передавального дозволеного кодового слова, яке вирізняється одним інформаційним символом, обов'язково вирізняється і значенням перевірного символу. Відстань $d_{\min} = 2$ до найближчого дозволеного кодового слова має місце для будь-якого передаваного слова, що відповідає максимально можливому за заданої кодової відстані набору дозволених кодових слів. Тому код Хеммінга, що виявляє одиночні помилки, оптимальний. Водночас він і досконалий код для виявлення одиночних помилок, оскільки всі можливі кодові слова або належать дозволеному набору, або відповідають випадку спотворення одного символу дозволеного слова.

У виправному коді Хеммінга двійкове число, складене з контрольних сум (синдром коду), має визначати порядковий номер спотвореного розряду кодового слова. Для цього контрольні суми складаються так, щоб

сума S_j містила символи кодового слова, які мають одиницю в j -му розряді в двійковому запису їх порядкового номера. Незавжно бачити, що при цьому символи, порядковий номер яких у двійковому запису виявляється одиницею j -го розряду, тобто дорівнює 2^{j-1} , $j = 1, 2, 3, \dots, n$, увійдуть тільки в одну контрольну суму S_j . Саме на цих позиціях (1-й, 2-й, 4-й, 8-й і т. д.) у коді Хеммінга і розміщуються перевірні символи $\beta_1, \beta_2, \beta_4, \beta_8$ і т. д. Відповідним їх вибором усі контрольні суми обертаються в нуль. Якщо немає спотворень, усі контрольні суми дорівнюють нулю. Одиночна помилка, наприклад, спотворення 1-го розряду кодового слова, призводить до обернення в одиницю контрольних сум, номери яких відповідають номерам розрядів двійкового запису числа 1, що вміщують одиниці. Тому, записуючи значення контрольних сум у порядку спадання їх номерів, тобто записуючи синдром $\dots S_4 S_3 S_2 S_1$, отримуємо у двійковому запису номер 1 спотвореного розряду.

Для наочності двійковий запис номерів розрядів кодового слова зведено у табл. 14.1 (таблицю наведено для 15-розрядних кодів).

Таблиця 14.1

| Порядковий номер розряду коду в десятковому запису | Значення розрядів двійкового запису порядкового номера розрядів коду | | | |
|--|--|---|---|---|
| | 1 (молодший) | 2 | 3 | 4 |
| 1 (молодший) | 1 | 0 | 0 | 0 |
| 2 | 0 | 1 | 0 | 0 |
| 3 | 1 | 1 | 0 | 0 |
| 4 | 0 | 0 | 1 | 0 |
| 5 | 1 | 0 | 1 | 0 |
| 6 | 0 | 1 | 1 | 0 |
| 7 | 1 | 1 | 1 | 0 |
| 8 | 0 | 0 | 0 | 1 |
| 9 | 1 | 0 | 0 | 1 |
| 10 | 0 | 1 | 0 | 1 |
| 11 | 1 | 1 | 0 | 1 |
| 12 | 0 | 0 | 1 | 1 |
| 13 | 1 | 0 | 1 | 1 |
| 14 | 0 | 1 | 1 | 1 |
| 15 | 1 | 1 | 1 | 1 |

Для 15-розрядного коду згідно з табл. 14.1 контрольні суми мають такий вигляд:

$$\begin{aligned}
 S_1 &= \beta_1 \oplus b_3 \oplus b_5 \oplus b_7 \oplus b_9 \oplus b_{11} \oplus b_{13} \oplus b_{15}; \\
 S_2 &= \beta_2 \oplus b_3 \oplus b_6 \oplus b_7 \oplus b_{10} \oplus b_{11} \oplus b_{14} \oplus b_{15}; \\
 S_3 &= \beta_4 \oplus b_5 \oplus b_6 \oplus b_7 \oplus b_{12} \oplus b_{13} \oplus b_{14} \oplus b_{15}; \\
 S_4 &= \beta_8 \oplus b_9 \oplus b_{10} \oplus b_{11} \oplus b_{12} \oplus b_{13} \oplus b_{14} \oplus b_{15},
 \end{aligned}
 \tag{14.16}$$

де b_i – значення інформаційних символів коду; $\beta_1, \beta_2, \beta_4, \beta_8$ – значення перевірних символів; \oplus – знак додавання за модулем два.

Значення перевірних символів β_i , які обертають у нуль контрольні суми (14.16), дорівнюють:

$$\begin{aligned}
 \beta_1 &= b_3 \oplus b_5 \oplus b_7 \oplus b_9 \oplus b_{11} \oplus b_{13} \oplus b_{15}; \\
 \beta_2 &= b_3 \oplus b_6 \oplus b_7 \oplus b_{10} \oplus b_{11} \oplus b_{14} \oplus b_{15}; \\
 \beta_4 &= b_5 \oplus b_6 \oplus b_7 \oplus b_{12} \oplus b_{13} \oplus b_{14} \oplus b_{15}; \\
 \beta_8 &= b_9 \oplus b_{10} \oplus b_{11} \oplus b_{12} \oplus b_{13} \oplus b_{14} \oplus b_{15}.
 \end{aligned}
 \tag{14.17}$$

Процедура формування виправного коду Хеммінга включає такі операції:

1) розміщення інформаційних символів у порядку зростання розрядів зі збереженням незайнятими розрядів з номерами 2^{i-1} , тобто 1-го, 2-го, 4-го, 8-го і т. д.;

2) обчислення перевірних символів, які займають розряди з номерами 2^{i-1} , за формулами (14.17).

Процедурою визначення передаваного інформаційного коду для використання виправного коду Хеммінга передбачено:

1) обчислення контрольних сум за формулами (14.16);

2) визначення порядкового номера спотвореного символу, якщо не всі контрольні суми дорівнюють нулю, і виправлення його на другий двійковий символ;

3) виділення інформаційного коду вилученням розрядів з номерами 2^{i-1} , $i = 1, 2, \dots$

Наприклад, якщо інформаційний код має вигляд 1001101 (будемо користуватись записом двійкових чисел за порядком зменшення номерів розрядів зліва направо), тоді його символи розміщуються у коді Хеммінга в такій послідовності: $100\beta_8110\beta_41\beta_2\beta_1$. Визначивши за формулами (14.17) значення перевірних символів $\beta_1, \beta_2, \beta_4, \beta_8$, отримаємо такий код Хеммінга: 10011100101. Якщо, наприклад, під час приймання 5-й розряд цього коду помилково буде прийнятий як одиниця, тоді, обчисливши контрольні суми, отримаємо $S_1 = 1, S_2 = 0, S_3 = 1, S_4 = 0$. Значення синдрому, який

визначає номер спотвореного розряду у двійковому запису, дорівнює $S_4S_3S_2S_1 = 0101$, тобто для виправлення помилки потрібно змінити символ у 5-му розряді.

Правильне виправлення кодового слова з використанням коду Хеммінга має місце для одиночних помилок, що знаходяться у повній відповідності з кодовою відстанню, яка в цьому разі дорівнює трьом. Дійсно, якщо у кодовому слові змінити інформаційний символ, двійковий запис порядкового номера, який вміщує мінімальну кількість одиниць, а саме: дві одиниці (одну одиницю у двійковому запису має лише порядковий номер контрольного символу), тоді одночасно змінюються два контрольні символи, що входять у ті самі контрольні суми. Отже, найближче до кожного передаваного слова дозволене кодове слово вирізняється трьома символами. Тому виправний код Хеммінга, який забезпечує максимальний набір дозволених кодових слів, якщо $d_{\min} = 3$, оптимальний. Зазначимо, що цей код, якщо його використати не як виправний, а як код, що виявляє помилки, як і належить при $d_{\min} = 3$, достовірно виявляє дві помилки. Дійсно, номери двох спотворених символів у двійковому запису відрізняються хоча б одним розрядом, а це означає, що якщо хоча б в одну контрольну суму ввійде тільки один зі спотворених символів, вона зміниться.

Отже, для утворення виправного коду Хеммінга навіть з одним інформаційним символом треба мати два перевірні символи (у двох молодших розрядах); за кількості інформаційних символів від 2 до 4 – три перевірні символи (у 1-, 2-, 4-му розрядах); за кількості інформаційних символів від 5 до 11 – чотири перевірні символи (у 1-, 2-, 4- і 8-му розрядах) і т. д.

Крім розглянутих виявлюваного і виправного кодів Хеммінга, можна використати комбінований код, який виправляє одиночні помилки і достовірно виявляє двійні помилки. Для цього виправний код Хеммінга доповнюється ще одним молодшим розрядом, який дорівнює сумі (за модулем два) всіх його розрядів. Тоді під час приймання складаються контрольні суми виправного коду без урахування молодшого розряду і, крім того, сума всіх розрядів прийнятого коду (за модулем два). Якщо частина контрольних сум дорівнює одиниці, тоді прийняте слово за наведеними правилами виправляється тільки за умови, що сума всіх символів також дорівнює одиниці. Якщо сума всіх символів дорівнює нулю, що свідчить про наявність двох (або парного числа) помилок, тоді кодове слово не виправляється (або правильно виправляються тільки одиночні помилки), а лише встановлюється факт наявності двох (або парного числа) помилок. Отже, якщо кількість помилок не більше двох, розглядуваний код достовірно виправляє одиночні та виявляє двократні помилки (у разі більшої кратності можливі пропуски або неправильне виправлення помилок). Неважко показати, що кодова відстань для цього коду дорівнює чотирьом.

Дійсно, як було показано, без урахування молодшого розряду дозволені кодові слова відрізняються не менше, ніж трьома символами. Але кодові слова, які відрізняються в останніх розрядах трьома символами, обов'язково (за умови парності) відрізняються і молодшими розрядами, тобто дозволені кодові слова в цілому відрізняються не менше ніж чотирма символами. Таким чином, і в цьому випадку має місце гранична корисувальна здатність (виправлення однієї та виявлення двох помилок), яка відповідає кодовій відстані $d_{\min} = 4$. Код забезпечує максимальний набір дозволених кодових слів, якщо $d_{\min} = 4$, і є оптимальним.

14.8. Побудова лінійного систематичного коду

Будь-який лінійний код еквівалентний деякому систематичному коду щодо збереження списку вагів, а отже, і відстаней Хеммінга. Перші k символів кодових слів, які збігаються із символами джерела, називаються інформаційними, а решта $n - k = r$ - перевірними. Швидкість коду R для систематичного коду - це відношення кількості інформаційних символів k до довжини кодового слова n . Не випадково, що систематичні коди, довжина яких дорівнює блоку N , а число інформаційних символів k , коротко називають (n, k) -кодами.

Побудуємо код на передавачі системи. Будь-яку кодову комбінацію можна подати в такому вигляді:

$$c_1, c_2, \dots, c_k; l_1, l_2, \dots, l_r,$$

де $\vec{c} = (c_1, c_2, \dots, c_k)$ - заданий вектор, що відображає інформаційну кодову комбінацію - інформація джерела; $\vec{l} = (l_1, l_2, \dots, l_r)$ - перевірна кодова комбінація.

Нехай на передавачі у запам'ятовувальному пристрої записано матрицю $[A]$ розміром $r \times k$, яка має вигляд:

$$[A] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rk} \end{bmatrix}; a_{ij} = \begin{cases} 0, \\ 1. \end{cases} \quad (14.18)$$

Знайдемо вектор \vec{l} як лінійне перетворення

$$\vec{l} = [A]\vec{c}_k$$

за відомим правилом

$$l_1 = c_1 a_{11} \oplus c_2 a_{12} \oplus \dots \oplus c_k a_{1k}, j = 1, 2, \dots, r;$$

$$l_j = c_1 a_{j1} \oplus c_2 a_{j2} \oplus \dots \oplus c_k a_{jk},$$

де l_j – символи вектора \bar{l}_r (перша контрольна група). Таким чином, утворено нову послідовність довжиною $n = k + r$ з уведеною надмірністю (перевірною), яка лінійно зв'язана з інформаційною послідовністю символів.

Для того щоб визначити здатність уведеної надмірності виявлення або виправлення помилок, доцільно розглянути загальний алгоритм роботи приймача (приймальної частини СПДІ).

У канал передавання дискретної інформації надходить послідовність сигналів, які відображають послідовність символів $c_1, c_2, \dots, c_k, l_{k+1}, l_{k+2}, \dots, l_{k+r}$. На вихід каналу або на вхід приймача надходить послідовність $c'_1, c'_2, \dots, c'_k, l'_{k+1}, l'_{k+2}, \dots, l'_{k+r}$.

У приймачі записано матрицю $[A]$ (14.18). Визначимо за тим самим правилом другу контрольну групу:

$$l''_j = a_{j1}c'_1 \oplus a_{j2}c'_2 \oplus \dots \oplus a_{jk}c'_k. \quad (14.19)$$

Порівняємо контрольну групу, що надійшла з каналу передавання, з новою контрольною групою (14.19):

$$x_1 = l'_{k+1} \oplus l''_{k+1}, x_2 = l'_{k+2} \oplus l''_{k+2}, \dots, x_i = l'_{k+i} \oplus l''_{k+i}, \dots, x_r = l'_{k+r} \oplus l''_{k+r};$$

$$X = [x_1, x_2, \dots, x_r];$$

де $X = [x_1, x_2, \dots, x_r]$; – контрольне число (синдром); якщо $X = [0, 0, \dots, 0]$, помилок немає, якщо хоча б одне $x_j = 1$, виявляється помилка.

14.9. Циклічні коди

Формування циклічних кодів основане на відображенні кодових слів елементами алгебри поліномів. Це дозволяє, крім операцій підсумовування кодових слів і множення їх на символ λ , які використовувались в алгебрі векторів, додатково ввести операції множення та ділення кодових слів і отримати за рахунок цього додаткові ознаки дозволених слів. Кодове слово довжиною n описується поліномом степеня $n + 1$, коефіцієнтами якого є символи кодового слова:

$$b = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_{n-1} x^{n-1}.$$

Для того щоб у результаті операцій, використовуваних для формування циклічних кодів, отримати поліноми, які описують ці коди, в усіх операціях над поліномами потрібно дотримувати дві умови:

1) дії над коефіцієнтами поліномів (додавання, віднімання, множення на символ λ) виконувати за модулем m ;

2) множення поліномів виконувати за модулем $x^n - 1$, тобто як результат множення брати залишок від ділення звичайного добутку поліномів на $x^n - 1$.

Перша умова потрібна для того, щоб коефіцієнти одержаних поліномів належали алфавіту каналу, а друга – щоб степінь цих поліномів відповідав довжині коду n .

Надалі в усіх випадках, крім спеціально обумовлених, ці умови братимемо до уваги. Добуток поліномів за модулем $x^n - 1$ позначатимемо символом \otimes , суму коефіцієнтів поліномів за модулем m – символом \oplus , а різницю – символом \ominus .

Циклічний код зазвичай задають породжувальним поліномом $F(x)$. Дозволені кодові слова циклічного коду відповідають поліномам, кратним породжувальному поліному $F(x)$. При цьому породжувальний поліном має задовольняти дві умови;

1) бути незвідним, тобто не ділитися ні на який інший поліном;

2) на нього має ділитися без остачі двочлен вигляду $x^n - 1$ (тут звичайна операція ділення).

Зведення операції ділення поліномів зумовлює дію віднімання коефіцієнтів за модулем m . Якщо звичайна різниця коефіцієнтів позитивна, ніяких затруднень щодо отримання її значення за модулем m не виникає. Для отримання n значень за модулем m від від'ємного числа до нього потрібно додавати (один або декілька разів) величину m до отримання першого позитивного значення або нуля, які і будуть результатом. Наприклад: $1 \ominus_3 3 = 1$; $1 \ominus_8 3 = 6$.

В окремому випадку двійкових кодів операцію віднімання коефіцієнтів не виконують, оскільки операції додавання і віднімання за модулем два еквівалентні: $1 \oplus 1 = 0$; $1 - 1 = 0$; $0 \oplus 1 = 1$; $0 - 1 = 1$. Це дозволяє у разі використання двійкових кодів за всіма операціями над поліномами (включаючи і ділення) виконувати тільки підсумовування коефіцієнтів за модулем два (двочлен $x^n - 1$ при цьому замінюють двочленом $x^n - 1$).

Циклічні коди є блоковими, рівномірними та лінійними. Лінійність циклічних кодів впливає з того, що будь-яка лінійна комбінація двох циклічних кодів, які діляться на породжувальний поліном $F(x)$, буде обов'язково ділитися на цей самий породжувальний поліном, тобто буде належати дозволеним словам цього циклічного коду. На дозволений набір кодових слів у цьому випадку накладаються додаткові обмеження порівняно зі звичайними лінійними кодами: ділення на породжувальний

поліном; ця ознака суттєво спрощує апаратурну реалізацію кодувальних і декодувальних пристроїв.

Базисну матрицю циклічного коду, який є лінійним, можна записати у такому вигляді:

$$L(x) = \begin{bmatrix} F(x) \\ xF(x) \\ \dots \\ x^{k-1}F(x) \end{bmatrix}. \quad (14.20)$$

Матрицю (14.20) у формі поліномів можна переписати у формі коефіцієнтів (символів), записуючи в одному стовпці коефіцієнти з однаковим степенем x . Найвищий степінь, який відповідає поліному в останньому рядку матриці (14.20), дорівнює $k - 1 + r$, де r – степінь породжувального полінома $F(x)$. Оскільки рядки цієї матриці, виражені поліномами за модулем $x^n - 1$, не можуть мати степінь вищий за $n - 1$, то має задовольнятися умова $k - 1 + r = n - 1$ або $k = n - r$. Неважко впевнитися, що всі поліноми у виразі (14.20) відповідають дозволеному набору слів певного циклічного коду (діляться на $F(x)$) і лінійно-незалежні, тобто матриця (14.20) дійсно є базисною матрицею лінійного коду (n, k) .

Перевірною матрицею цього коду утворюється аналогічно, тільки замість породженого полінома $F(x)$ у ній використовується так званий генераторний поліном

$$g(x) = (x^n - 1)/F(x), \quad (14.21)$$

ступінь якого дорівнює k . Вона має такий вигляд:

$$H(x) = \begin{bmatrix} g(x) \\ xg(x) \\ \dots \\ x^{r-1}g(x) \end{bmatrix}.$$

В ортогональності поліномів легко переконатися, оскільки на підставі (14.21)

$$x^i F(x) \otimes_n x^j g(x) = x^{i+j} \otimes_n (x^n - 1) = 0.$$

Можна показати, що якщо кодове слово $b = (\beta_0, \beta_1, \dots, \beta_{n-1})$ належить до дозволеного набору даного циклічного коду, то до нього належить і кодове слово $b^{(1)} = (\beta_{n-1}, \beta_0, \beta_1, \dots, \beta_{n-2})$, отримане із першої циклічної перестановки символів. Легко переконатися, що кодовому слову $b^{(1)}$ відпо-

відає поліном $b^{(1)}(x)$, який являє собою добуток $xb(x)$ за модулем $x^n - 1$. Тому, якщо $b(x)$ ділиться на породжувальний поліном $F(x)$, то на нього ділиться і поліном $b^{(1)}(x)$, а кодове слово $b^{(1)}$ належить тому самому циклічному коду, що й початкове слово b . Ця властивість і визначила назву «циклічний» для цього коду.

Формування дозволених слів циклічного коду згідно із заданим інформаційним словом, який має k символів, можна виконати таким чином. Утворимо поліном $G(x)$ степеня $k - 1$, що відповідає передаваному k -розрядному інформаційному слову, і помножимо його на x^r , де r – степінь породжувального полінома $F(x)$, $k + r \leq n$. Отриманий поліном степеня $k + r - 1$ розділимо на породжувальний поліном $F(x)$. Віднімемо від полінома $x^r G(x)$, молодший член якого має степінь r , а остачу $g(x)$ від його ділення – на $F(x)$ (порядок полінома $g(x)$, принаймні на одиницю, нижчий від порядку полінома $F(x)$, тобто не вищий за $r - 1$). Отримаємо поліном $Q(x) = x^r G(x) - g(x)$, який ділиться на $F(x)$ без остачі і визначає дозволене кодове слово. Молодші r розрядів цього кодового слова (перевірні) відповідають коефіцієнтам остачі $g(x)$, а старші розряди, починаючи з $(r - 1)$ -го, відповідають інформаційному слову, зміщеному на r розрядів у результаті операції множення полінома інформаційного слова $G(x)$ на x^r .

Циклічний код має систематичну форму, при цьому всі r перевірних символів ідуть підряд, займаючи молодші розряди слова. Число його перевірних символів r визначається степенем породжувального полінома, а загальне число розрядів n – степенем двочлена $x^n - 1$. Максимальне число інформаційних символів $k = n - r$.

Для прикладу визначимо дозволене слово семизначного двійкового циклічного коду, визначуване інформаційним словом 1011 і породжувальним поліномом $F(x) = x^3 + x^2 + 1$, що відповідає двочлену $x^7 + 1$. Поділимо поліном $x^r G(x) = x^6 + x^4 + x^3$ на поліном $F(x)$ (нагадаємо, що для двійкових кодів віднімання коефіцієнтів замінюється їх підсумовуванням за модулем два):

$$\begin{array}{r}
 \oplus x^6 \quad \quad \quad + x^4 + x^3 \\
 \hline
 x^6 + x^5 \quad \quad \quad + x^3 \\
 \hline
 \oplus x^5 + x^4 \\
 \hline
 x^5 + x^4 \quad \quad \quad + x^2 \\
 \hline
 \quad \quad \quad \quad \quad \quad x^2
 \end{array}
 \left| \begin{array}{r}
 x^3 + x^2 + 1 \\
 \hline
 x^3 + x^2
 \end{array} \right.$$

Остача $g(x) = x^2$. Отже, $Q(x) = x^r G(x) + g(x) = x^6 + x^4 + x^3 + x^2$, а шуканий циклічний код, отриманий за допомогою полінома $Q(x)$, має вигляд 1011100 (відсутнім у поліномі $Q(x)$ степеням відповідають коефіцієнти, які дорівнюють нулю).

Для виявлення помилок у циклічному коді замість операції ділення на породжувальний поліном $F(x)$ зручно користуватись операцією множення за модулем $x^n - 1$ на генераторний поліном $g(x)$ (для двійкових кодів, коли використовується тільки підсумовування коефіцієнтів за модулем два, множення поліномів виконують відповідно за модулем $x^n - 1$).

Поліном $\tilde{Q}(x)$, що відповідає прийнятому коду, можна подати у вигляді суми $\tilde{Q}(x) = Q(x) \oplus n(x)$, де $Q(x)$ – поліном, що відповідає переданому циклічному коду; $n(x)$ – поліном, що відповідає «коду помилки», який має одиниці на місці спотворених символів і нулі в решті позицій.

Помноживши $\tilde{Q}(x)$ на генераторний поліном $g(x)$ за модулем $x^n - 1$ (для двійкових кодів – за модулем $x^n + 1$), отримаємо

$$\tilde{Q}(x) \otimes g(x) = \frac{Q(x)}{F(x)} \otimes (x^n - 1) + n(x) \otimes g(x) = n(x) \otimes g(x),$$

оскільки $g(x) = \frac{(x^n - 1)}{F(x)}$, а $Q(x)$ ділиться без остачі на $F(x)$ за визначенням

циклічного коду. Добуток $n(x) \otimes g(x)$ не дорівнює нулю, якщо тільки «код помилки» не збігається ні з одним із дозволених слів циклічного коду цього типу. Отже, ознакою наявності помилок є відмінність від нуля добутку $\tilde{Q}(x) \otimes g(x)$; при цьому виявляється помилка будь-якої кратності, якщо тільки вона не збігається з одним із дозволених слів цього циклічного коду.

Множення за модулем $x^n + 1$ для двійкових кодів легко реалізується використанням кільцевого регістра з n комірок із суматором за модулем два на вході (рис. 14.7). Нехай на вхід кільцевого регістра послідовно надходять, розпочинаючи з молодших розрядів, коефіцієнти звичайного добутку поліномів $Q(x) \cdot g(x)$. Цей добуток, що має вигляд

$$Q(x) \cdot g(x) = p(x)(x^n + 1) + n(x) \otimes g(x), \quad (14.22)$$

розбивається на частину, яка кратна $x^n + 1$, і остачу $n(x) \otimes g(x)$. Перший добуток у правій частині рівняння (14.22) розбивається на пари членів вигляду $x^{n+i} + x^i$, які взаємно «гасяться» під час проходження через кільцевий регістр. Дійсно, у той момент, коли на вхідний суматор (рис. 14.7) надходить коефіцієнт 1 при члені x^{n+i} , по ланцюгу оберненого зв'язку із n -ї комірки регістра в нього надходить коефіцієнт 1 при x^i . Складаючись за модулем два, вони «гасяться». У результаті в момент надходження у регістр останнього (старшого) члена полінома добутку (14.22) у кільцевому регістрі будуть зафіксовані коефіцієнти полінома, який дорівнює $n(x) \otimes g(x)$.

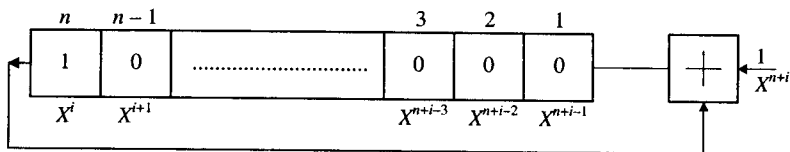


Рис 14.7. Схема отримання коду Хеммінга

Одиниць у кільцевому регістрі після закінчення цієї процедури немає. Це означає, що помилок у прийнятому кодовому слові або немає, або поліном помилки кратний породжувальному поліному $F(x)$; наявність же одиниць слугує критерієм спотворень цього кодового слова. Отже, розглянута найпростіша схема дозволяє виявити наявність помилок, які не збігаються з дозволеними кодовими словами.

Спотворення помилок, тобто визначення полінома помилки $n(x)$, можливе лише стосовно помилок заданої структури, оскільки в загальному випадку за відомим добутком $n(x) \otimes g(x)$ неможливо однозначно відновити значення відповідного звичайного добутку $n(x) \cdot g(x)$ і визначити $n(x)$. Структуру помилок можна задавати числом спотворених символів та інтервалами між ними. Помилка заданої структури повністю визначається положенням молодшого із спотворених розрядів кодового слова. Її поліном можна подати у такому вигляді:

$$n(x) = x^{\mu-1} n_0(x), \quad (14.23)$$

де μ – номер молодшого зі спотворених розрядів; $n_0(x)$ – помилка структури, що починається з першого розряду.

Остача в кільцевому регістрі, що відповідає помилці (14.23), дорівнює

$$x^{\mu-1} [n_0(x) \otimes g(x)],$$

тобто відрізняється від $n_0(x) \otimes g(x)$ зсувом вправо (у бік старших розрядів) на $\mu - 1$ розрядів регістра. Положення помилки визначається числом розрядів, на яке треба просунути залишок у кільцевому регістрі, щоб він збігся з «еталонним» поліномом $n_0(x) \otimes g(x)$, який зберігається у пам'яті. Зокрема, для одиночних помилок «еталоном» для порівняння служить генераторний поліном $g(x)$.

14.10. Неперервне згорткове кодування

До цього ми розглядали блокове кодування, за якого послідовність інформаційних символів розбивається на окремі блоки, які кодуються незалежно один від одного. Закодована таким чином послідовність стає послідовністю незалежних кодових комбінацій однакової довжини.

У разі неперервного кодування кожний символ інформаційної послідовності визначається за рекурентними співвідношеннями, які пов'язані з його відповідними символами інформаційної послідовності. Клас неперервних (рекурентних) кодів, які формуються неперервними перевірними елементами, перемежуються (чергуються) з інформаційними вздовж усієї довжини кодової послідовності і підпорядковуються одному й тому самому рекурентному співвідношенню. Праві частини рекурентних співвідношень являють собою нібито згортку відповідної ділянки інформаційної послідовності, тому їх і називають згортковими. В основу їх побудови покладено принцип формування перевірних розрядів підсумовуванням за модулем два кожного інформаційного розряду з деяким набором попередніх розрядів. Приклад найпростішого згорткового кодера показано на рис. 14.8.

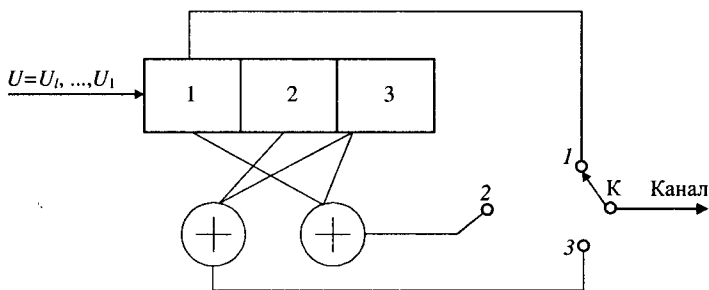


Рис 14.8. Приклад структурної схеми згорткового кодера

Вхідна інформаційна послідовність $U = U_1, U_2, U_3, \dots, U_l$ надходить у регістр багатотактного фільтра ($N = 3$), деякі з комірок якого з'єднані з двома суматорами за модулем два. На контакт 1 комутатора К подається інформаційний символ, на контакти 2 і 3 – перевірні символи. Після надходження кожного інформаційного символу комутатор (сканер) зчитує три символи (інформаційний і два перевірні) і передає їх у канал. Оскільки в розглядуваному випадку в канал подаються як інформаційний, так і перевірні символи, то вихідна кодова послідовність є систематичною. Однак виходу з першої комірки на контакт 1 може й не бути або на цей контакт подаватимуться і виходи інших комірок, результатом чого буде

несистематичний код. У загальному випадку згортковий кодер буде мати такий вигляд, як показано на рис. 14.9.

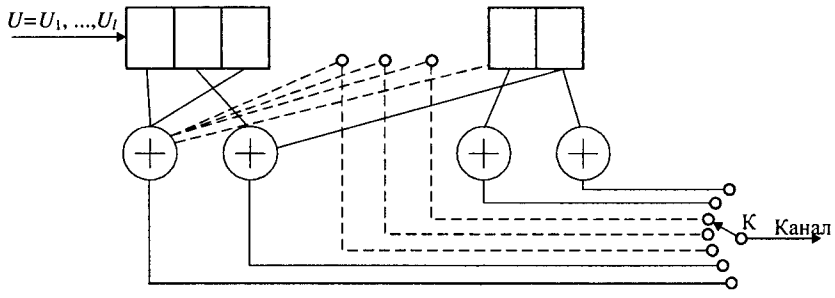


Рис. 14.9. Структурна схема двійкового згорткового кодера

Зчитування вихідних символів комутатор виконує після кожного надходження l інформаційних символів, де $l = 1, 2, 3, \dots$, причому зазвичай l кратно N , тобто $Nl = n$.

Згорткові коди задаються рекурентною формулою для визначення символів кодової послідовності, яка має для двійкових згорткових кодів такий вигляд:

$$\beta_{ij} = \left\{ \sum_{v=-g}^{s-1} c_{vj} b_{i+v} \right\} (\text{mod } 2); \quad j = 0, 1, \dots, l-1, \quad c_{vj} = 0, 1, \quad (14.24)$$

де b_i – символи вхідної інформаційної послідовності.

Функціональну схему кодувального пристрою, який забезпечує формування символів за законом (14.24), показано на рис. 14.10.

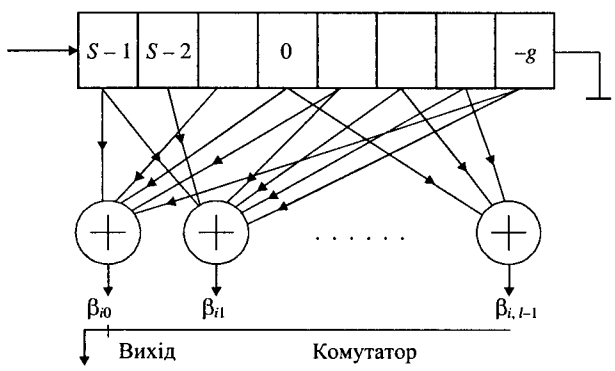


Рис. 14.10. Функціональна схема кодувального пристрою

Символи вхідної інформаційної послідовності надходять на регістр зсуву, який складається з $k = s + g$ комірок (відповідно до позначень рівняння (14.24) коміркам присвоєно номери від $-g$ до $s - 1$). Символи β_{ij} формуються за допомогою суматорів за модулем два, на які з комірок регістра зсуву, які відповідають ненулевим значенням c_{vi} у рівнянні (14.24), надходять інформаційні символи. У той такт роботи регістра зсуву, коли в нульовій його комірці знаходиться символ b_i (відповідно у вхідній комірці знаходиться символ b_{i+s-1} , а у вихідній $-b_{i-g}$), на суматорах формуються символи $\beta_{i0}, \beta_{i1}, \dots, \beta_{i,l-1}$. Комутатор в i -му такті роботи регістра зсуву, протягом якого залишаються незмінними символи b_i , що заповнюють його, по черзі подає символи β_{ij} у канал. У наступному такті всі символи в регістрі зсунуться вправо на одну комірку: у нульовій комірці буде символ b_{i+1} , у вхідну комірку надійде символ b_{i+s} , а у вихідну $-b_{i-g+1}$. Відповідно на суматорах формуються символи кодової послідовності $\beta_{i+1,0}, \beta_{i+1,1}, \dots, \beta_{i+1,l-1}$ і т. д. Отже, з надходженням на вхід кодувального пристрою кожного нового символа інформаційної послідовності у канал надходять l символів кодової послідовності.

Для випадку систематичного згорткового коду на рис. 14.11 зображено модифікацію кодувального пристрою. Тут на один із входів комутатора (наприклад, на перший) безпосередньо з нульової комірки регістра надходить інформаційний символ b_i , тобто $\beta_{i0} = b_i$.

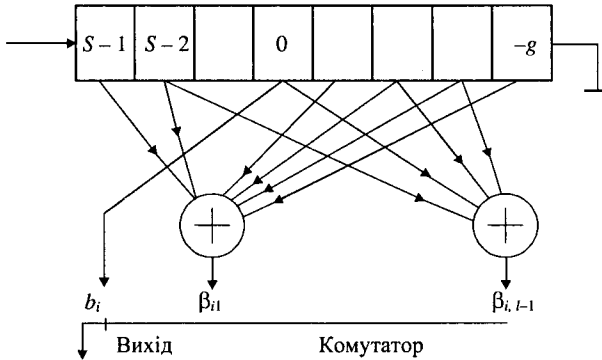


Рис. 14.11. Принцип формування згорткового коду

Основна відмінність від кодування блоків систематичних кодів полягає у неперервному оновленні групи інформаційних символів, які беруть участь у формуванні чергових l символів кодової послідовності (під час блокового кодування всі символи кодового слова формуються з використанням одного й того самого блоку інформаційних символів, який повністю оновлюється під час переходу до кодування наступного слова). Макс

симальну кількість інформаційних символів, які беруть участь у формуванні кожного вихідного символу згорткового коду, що визначається кількістю комірок регістра зсуву K , називають довжиною кодових обмежень. Ця характеристика близька за змістом до довжини блоку інформаційних символів для блокових кодів. Інколи довжину кодових обмежень визначають за каналними (вихідними) символами. Тоді вона виражається величиною $N = l \cdot k$. Кожний символ вхідної інформаційної послідовності може брати участь у формуванні до $lk - 1$ символів вихідної послідовності згорткового коду. Неважко бачити, що ускладнення кодувально-го пристрою зі збільшенням довжини кодових обмежень характеризується лінійним законом. Тому кодувальні пристрої не накладають суттєвих обмежень на реалізованість згорткових кодів.

За способом декодування згорткові коди можна поділити на два класи: коди із синдромним коригуванням і коди з використанням послідовного декодування.

Коди із синдромним коригуванням використовуються тільки систематичними згортковими кодами. Спосіб коригування ґрунтується на утворенні з прийнятої кодової послідовності контрольних сум вигляду

$$s_{ij} = \beta'_{ij} \oplus \sum_{v=g}^{s-1} c_{vj} b'_{i+v}, \quad j = 0, 1, \dots, l-1, \quad (14.25)$$

у кожному з яких, як і у випадку блокових лінійних кодів, входить один із перевірних символів β'_{ij} . Принцип коригування такий самий, як і для блокових лінійних кодів: закони формування контрольних символів і контрольних сум вибираються так, щоб синдром помилки, який складається зі значень контрольних сум, дозволяв визначити спотворені символи. Якщо немає спотворень, усі контрольні суми дорівнюють нулю, а синдромом помилки є нуль-вектор E .

Згорткові коди з використанням послідовного декодування характеризуються тим, що вони не мають будь-яких обмежень як на структуру зв'язків між символами інформаційної і кодової послідовностей, так і на структуру виправних помилок (у межах коригувальної здатності коду). Це створює принципову можливість, використовуючи несистематичні коди, отримувати вигоду в ефективності кодів більше ніж у $\frac{1}{(1-\rho_i)}$ разів.

При цьому вигоду у $\frac{1}{(1-\rho_i)}$ разів досягається завдяки тому, що у несистематичному коді всі символи кодової послідовності (а не тільки надмірні) беруть участь у коригуванні помилок, а додатковий вигоду – завдяки тому, що немає обмежень на вибір їх зв'язків з інформаційними симво-

лами. Однак практична реалізація цього виграшу можлива лише за умови створення алгоритмів декодування несистематичних згорткових кодів довільної структури, обчислювальна складність яких не виходила б за можливості діючих ЕОМ.

У загальному вигляді схему декодування згорткових кодів довільної структури можна було б подати таким чином. Для всіх можливих реалізацій вхідної інформаційної послідовності за формулами (14.24) визначаються відповідні дозволені кодові послідовності. Декодування полягає у відборі найбільш правдоподібних із них щодо прийнятої кодової послідовності (наприклад, найближчої до прийнятої послідовності за метрикою Хеммінга). Проте такий «лобовий» метод занадто громіздкий і практично непридатний (для інформаційної послідовності, яка вміщує n символів, треба було б розглянути 2^n дозволених кодових послідовностей). Тому для створення реалізованих алгоритмів декодування несистематичних згорткових кодів довільної структури потрібно знайти такий підхід до розв'язання задачі, який дозволив би суттєво скоротити обсяг перебирання дозволених кодових послідовностей для пошуку найбільш правдоподібних із них, тобто так, щоб складність декодувального пристрою зростала як малий ступінь довжини кодових обмежень.

Для відображення варіантів перебирання зручно користуватись «деревом інформаційних послідовностей». Дерево, яке відображає всі можливі реалізації вхідної інформаційної послідовності, показано на рис. 14.12. Від його основи йдуть дві гілки, що відповідають двом можливим значенням (0 або 1) першого символу інформаційної послідовності. Кожна з них дає дві гілки, що відповідають можливим значенням другого символу. Варіанти перебирання визначаються можливими «траєкторіями» переміщення по гілках дерева інформаційних послідовностей, починаючи від його основи.

Можливість суттєвого скорочення обсягу перебирання кодових послідовностей під час декодування несистематичних згорткових кодів залежить від таких двох обставин.

1. Кореляційні зв'язки між символами кодової послідовності обмежені ділянкою, яка визначається довжиною кодових обмежень K . Тому, якщо замість розпізнавання всього прийнятого повідомлення виконувати послідовне розпізнавання окремих його ділянок, що мають тривалість, порівнянну з довжиною кодових обмежень, тоді для перших символів виділеної ділянки послідовності будуть ураховані всі їх зв'язки з рештою символів і ймовірність їх розпізнавання майже не погіршиться. Це дозволяє реалізувати процедуру декодування у вигляді послідовного розпізнавання цих ділянок з прийняттям на кожній такій ділянці остаточного рішення лише щодо першого символу і зміщення кожного разу розпізна-

ваної ділянки на один символ. Оскільки кількість дозволених комбінацій визначається показниковою залежністю від тривалості розглядуваної ділянки інформаційної послідовності, загальний обсяг послідовностей, які підлягають перебиранню, при цьому різко скорочується.

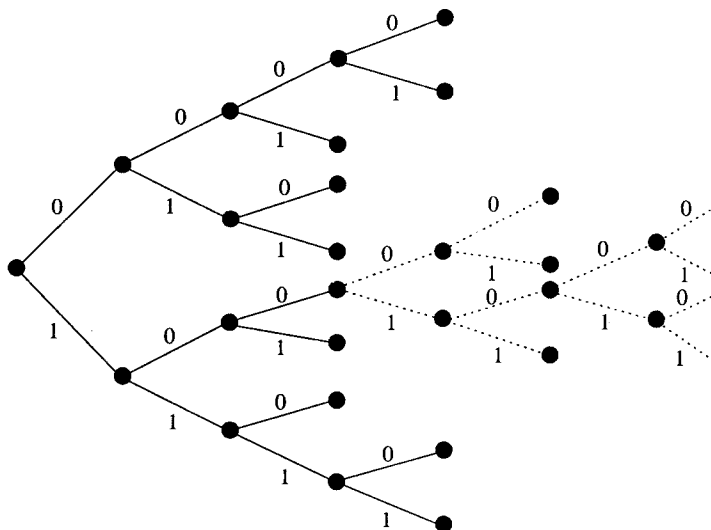


Рис. 14.12. Зображення кодового дерева

2. Більшість можливих комбінацій інформаційних символів мало-правдоподібні і тому кодові послідовності, що їм відповідають, у разі порівняння з прийнятою послідовністю можуть бути відкинуті після формування перших же символів. Це також приводить до різкого скорочення обсягу перебирання (для більшості варіантів розгляд припиняється вже на самому початку оцінюваної ділянки траєкторії переміщення по дереву інформаційних послідовностей).

Ідеї, які покладено в основу методу послідовного декодування, уперше запропонував Возенкрафт. Поява методу послідовного декодування викликала підвищений інтерес до згорткових кодів у теорії і практиці кодування: були запропоновані нові більш економічні алгоритми послідовного декодування, стала досліджуватися їх ефективність.

14.10.1. Згорткове кодування із синдромним коригуванням

Суттєвою особливістю згорткових кодів із синдромним коригуванням є більша порівняно з блоковими кодами можливість вибору інформаційних символів, які беруть участь у формуванні перевірних символів. Якщо для блокових кодів перевірні символи і контрольовані ними інформаційні символи мають бути зосереджені у межах кодового слова, то у згорткових кодах є можливість їх «розосередити» (за такою ознакою блокові коди належать до зосереджених, а згорткові – до розосереджених кодів). У разі впливу некорельованих завад, коли спотворення окремих символів можна розглядати як випадкові незалежні події, розрізнення кодів за цією ознакою не є суттєвим. Але якщо на канал діє корельована завада, що призводить до появи пачки помилок, що зачіпають цілий ланцюг символів (зазвичай це сильні короткочасні завади типу розряду блискавки, короткочасних загасань сигналу і т. ін.), розосереджені коди мають суттєві переваги. Якщо забезпечити достатнє рознесення між інформаційними і відповідними їм перевірними символами, тобто таке, щоб вони не підлягали одночасному впливу пачки помилок, тоді стає можливим виправити всі помилки у межах цієї пачки незалежно від їх кратності. Тому в тих випадках, коли за умов зв'язку доводиться враховувати можливість дії пачок помилок, згорткові коди із синдромним коригуванням дуже ефективні.

Для прикладу розглянемо найпростіший двійковий згортковий код Фінка – Хагельбергера, який широко використовують у практиці зв'язку. Це систематичний згортковий код з надмірністю 0,5 такої структури:

$$\dots b_{i-h} \beta_{i-h} \dots b_{i-1} \beta_{i-1} b_i \beta_i \dots b_{i+h+1} \beta_{i+h+1} \dots$$

Формула для обчислення перевірних символів цього коду має вигляд

$$\beta_i = b_{i-h} \otimes b_{i+h+1} \quad (14.26)$$

і є окремим випадком формули (14.24), коли коефіцієнти c_{v_i} відмінні від нуля, тільки якщо $v = -h$ і $v = -h + 1$. Відповідні синдрому помилки для цього коду згідно з формулою (14.25) дорівнюють:

$$S_i = \beta'_i \oplus b'_{i-h} \oplus b'_{i+h+1}. \quad (14.27)$$

Із рівнянь (14.26) і (14.27) видно, що перевірні символи β_i отримуються підсумовуванням симетрично розміщених відносно їх інформаційних символів кодової послідовності, віддалених на $2h + 1$ позицій у кожний бік:

$$\dots b'_{i-2h-1} \dots \beta'_{i-h-1} \dots b'_i \dots \beta'_{i+h} \dots b'_{i+2h+1} \dots$$

Виберемо параметр h так, щоб виконувались такі умови: якщо інформаційний символ b_i підлягає діяння пачки помилок, тоді ні ця, ні сусідні

пачки помилок не повинні захоплювати ні перевірні символи β_{i-h} і β_{i-h-1} , у формувани яких бере участь символ b_i , ні інформаційні символи b_{i+2h+1} і b_{i-2h-1} , які беруть участь у формуванні тих самих перевірних символів. Як видно з рис. 14.13, для того, щоб одна і та сама пачка помилок одночасно із символом b'_i не могла захопити і символ β'_{i+h} або β'_{i-h-1} , її максимальна довжина l_{\max} має задовольняти умову

$$l_{\max} \leq 2h. \quad (14.28)$$

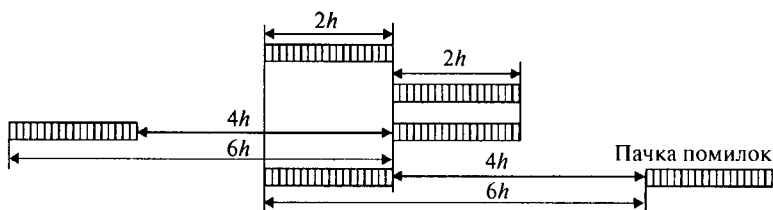


Рис. 14.13. Діаграма побудови згорткового коду

Для того щоб кінець попередньої і початок наступної пачки помилок не могли захопити символи b'_{i-2h-1} і b'_i або b'_i та b'_{i+2h+1} , мінімальний інтервал між початками сусідніх пачок L_{\min} має задовольняти умову

$$L_{\min} \geq 6h + 1. \quad (14.29)$$

У разі виконання умов (14.28) і (14.29) спотворення інформаційного символу L буде супроводжуватися одночасним оберненням в одиницю пари складових синдрому помилок:

$$S_{i-h-1} = \beta'_{i-h-1} \oplus b'_{i-2h-1} \oplus b'_i; \quad (14.30)$$

$$S_{i+h} = \beta'_{i+h} \oplus b'_i \oplus b'_{i+2h+1}, \quad (14.31)$$

індекси яких відрізняються на $2h + 1$ (решта символів у правих частинах рівнянь (14.30) і (14.31) не будуть підлягати дії пачок помилок). Легко переконатися, що в разі дотримування умов (14.28) і (14.29) спотворення перевірного символу зумовлює обернення в одиницю лише складової синдрому з тим самим індексом, складові ж синдрому з індексом, що відрізняється у той чи той бік на $2h + 1$, зберігає нульове значення.

Отже, правила коригування коду Фінка – Хагельбергера зводиться до такого: якщо складова синдрому S_j відмінна від нуля, а складова S_{j-2h-1} дорівнює нулю, тоді виправлення раніше прийнятих інформаційних символів не виконується; якщо поруч зі складовою S_j не дорівнює нулю і складова S_{j-2h-1} , тоді виправляється інформаційний символ b'_{j-h} .

Код Фінка – Хагельбергера дозволяє виправляти пачки помилок, якщо тільки мінімальний інтервал між ними більше ніж у три рази перевищує їх максимальну довжину (інакше неможливо вибрати параметр H коду, який задовольняє одночасно умови (14.28) і (14.29).

14.10.2. Згорткові коди з використанням послідовного декодування

Якщо за синдромного коригування сфера ефективного застосування згорткових кодів обмежувалась лише каналами, які підлягають впливу пачок помилок, то з появою алгоритмів послідовного декодування їх сфера використання суттєво розширилась. Запропоновано алгоритми послідовного декодування, обчислювальна складність яких не виходить за межі існуючих технічних можливостей навіть для згорткових кодів з досить великою довжиною кодового обмеження.

Суть усіх алгоритмів послідовного декодування полягає в цілеспрямованому пошуку «траєкторії» переміщення по дереву інформаційних послідовностей (див. рис. 14.12), яка відповідала б найбільш правдоподібній реалізації інформаційної послідовності. Під цілеспрямованістю тут розуміють такий спосіб пошуку цієї траєкторії, який потребує перебирання лише невеликої частини гілок цього дерева. Мірою правдоподібності траєкторії для посимвольного приймання може служити відстань за Хеммінгом між реалізацією кодової послідовності, що відповідає цій траєкторії, та фактично прийнятою послідовністю символів. У міру просування по розглядуваній траєкторії і нарощування кількості символів в оцінюваній кодовій послідовності відстань між нею і відповідною ділянкою прийнятої послідовності у разі неправильного вибору траєкторії буде швидко збільшуватись. Це дозволяє припинити просування по цій траєкторії, як тільки відстань досягає встановленого порогового значення d_1 (порогове значення d_1 у міру просування по траєкторії збільшується).

Існують різні підходи до процесу перебирання реалізацій за послідовного декодування.

Ми обмежились розглядом основної ідеї алгоритму послідовного декодування, запропонованого Возенкрафтом. Ключова позиція прийнятої Возенкрафтом стратегії послідовного декодування полягає у заміні завдання пошуку найбільш імовірною з усіх дозволених кодових послідовностей завданням вилучення всіх малоймовірних кодових послідовностей лише за великої кількості спотворень. Кожний крок запропонованої Возенкрафтом процедури послідовного декодування полягає у перебиранні відрізків траєкторій, які можуть служити продовженням вибраної на попередніх кроках початкової її ділянки і закріпленні першої гілки відібраного продовження. Установлюється максимальна довжина n відріз-

ків траєкторії, розглядуваних на кожному кроці декодування (глибина перебирання), і поріг d_1 , який визначає допустиму відстань між відповідними ділянками розглядуваної і прийнятої кодової послідовності. Окремі варіанти «відсіюються», як тільки ця відстань перебільшується. Перебирання припиняється з отриманням першого ж варіанта відрізка траєкторії довжиною n , для якого ця відстань не перевищує пороговий рівень d_1 , після чого перша її гілка (символ) включається у декодовану початкову ділянку й уся процедура вибирання варіантів продовження траєкторії повторюється від нової вузлової точки. Якщо на якому-небудь кроці декодування з порогом d_1 виявляються «відсіяними» усі варіанти продовження траєкторії заданої довжини n , процедура на цьому кроці повторюється з новим значенням порога $d_2 > d_1$. Обчислювальна складність цього алгоритму послідовного декодування визначається вибором глибини перебирання n , значеннями порогових рівнів d_1, d_2, \dots, d_l і кількістю градацій порогових рівнів (зазвичай обмежуються двома градаціями).

У спеціальній літературі, у якій публікуються праці із завадостійкого кодування, зокрема з послідовного декодування, можна знайти множину різних алгоритмів, які відрізняються як окремими особливостями процедури декодування, так і підходом до розв'язання задачі. Цікавим виявився, зокрема, запропонований А. Вітербі алгоритм декодування згорткових кодів, оснований на використанні зрізаної функції правдоподібності. Основною перевагою цього алгоритму, який використовується лише за малої довжини кодових обмежень ($K \leq 8$), є фіксація обчислювального навантаження на символ.

Останнім часом велику увагу приділяють каскадним кодам, які ґрунтуються на послідовному кодуванні двома різними кодами (код, який використовується для першого кодування, називають зовнішнім, а код, який використовується для другого кодування і формує символи, що надходять у лінію, – внутрішнім). Зокрема, дуже ефективними виявляються каскадні коди, використовувані у вигляді зовнішнього – відповідні класи блокового коду, а у вигляді внутрішнього – згорткові коди.

15. Надійність функціонування системи передавання дискретної інформації

15.1. Надійність систем передавання дискретної інформації

Одна з основних вимог, які ставляться до СПДІ, – вимога забезпечення високої надійності. Надійністю називають властивість системи забезпечувати нормальне виконання всіх заданих функцій протягом регламентованого проміжку часу зі збереженням у заданих границях усіх якісних показників. У розглядуваному випадку будемо розуміти СПДІ як тракт передавання дискретної інформації, тобто сукупність технічних засобів і каналів зв'язку, які забезпечують передавання дискретної інформації між двома безпосередньо зв'язаними джерелом і отримувачем інформації. Відповідно до функціонального призначення СПДІ надійністю тракту передавання дискретної інформації називають здатність тракту забезпечувати передавання певного обсягу інформації із заданою правильністю за встановлений проміжок часу. У загальному випадку тракт передавання даних (ТПД) (рис. 15.1) складається із декількох паралельно ввімкнених каналів зв'язку, апаратури передавання даних (АПД) та апаратури обслуговування (АО).

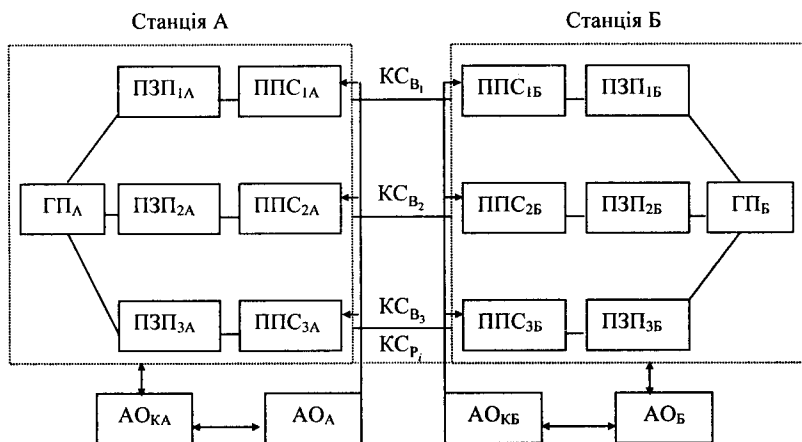


Рис. 15.1. Структурна схема ТПД за паралельної роботи каналів передавання даних

Канали зв'язку, які входять до складу ТПД, можуть бути різних типів, якості та довжини. Крім основних робочих каналів ($КС_{в}$), організується деяка кількість резервних ($КС_{р}$). Контроль за станом основних та резервних каналів здійснюється за допомогою апаратури обслуговування каналів зв'язку ($АО_{к}$), а заміна несправних каналів зв'язку резервними – за допомогою спеціальних пристроїв резервування. Основною складовою частиною ТПД є канал передавання даних (симплексний або дуплексний). Схему симплексного каналу передавання даних показано на рис. 15.2, у дуплексному каналі міняються місцями передавальна та приймальна частини.

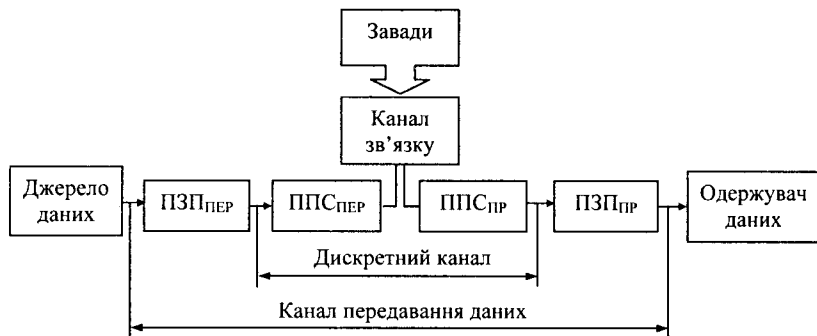


Рис. 15.2. Схема симплексного каналу передавання даних

Відповідно до вищенаведеного визначення надійності ТПД остання може знизитись у таких випадках:

- потрібна правильність передавання даних забезпечується, але час для передавання заданого обсягу інформації перевищує нормативний;
- час передавання в межах нормативного, але точність передавання нижча від потрібної;
- правильність нижча від потрібної і час передавання менший від нормативного.

Час передавання повідомлення t визначається з моменту введення в передавач СПДІ першого розряду блоку (знака) до моменту виведення із приймача СПДІ останнього розряду цього самого блоку (знака). Він складається з апаратурних затримок (часу перетворення інформації в передавачі і приймачі СПДІ) і часу поширення сигналу каналом зв'язку: $t = t_a + t_k$, де t_a – апаратурні затримки; t_k – час поширення каналом зв'язку. У системах без зворотного зв'язку $t_a = \text{const}$, а t_k може бути як постійним (наприклад, у каналах кабельних ліній зв'язку), так і змінним (наприклад, у каналах зв'язку через штучні супутники Землі).

У системах зі зворотним зв'язком t_a і t_k – величини змінні і визначаються виразами:

$$t_a = t_{a0} + n\Delta t_a; t_k = t_{k0} + n\Delta t_k,$$

де t_{a0} – апаратурні затримки за відсутності помилок у каналі; t_{k0} – час поширення сигналу за відсутності помилок у каналі; Δt_a – додаткова апаратурна затримка внаслідок погіршення якості каналу; Δt_k – збільшення часу поширення сигналу за рахунок повторення передавання інформації про погіршення якості сигналу; $n = 0, 1, 2, \dots$ – кількість повторів передавання інформації.

Отже, $t = t_{a0} + t_{k0} + n\Delta t$, де $\Delta t = \Delta t_a + \Delta t_k$ – змінна затримка, яка визначається якістю каналу.

Якщо немає помилок у каналі ($n = 0$), $t = t_{a0} + t_{k0} = \text{const}$. У цьому разі передача інформації по тракту передавання каналу характеризується номінальною швидкістю v_n , тобто середньою кількістю блоків (знаків), які видаються ПЗПпер абоненту (в апаратурі обробки даних) за одиницю часу. За наявності у каналі помилок реальна швидкість передавання інформації v_p , тобто середня кількість блоків (знаків), які видає ПЗПпер за одиницю часу, величина змінна; у системах з оберненим зв'язком v_p залежить від кількості повторень передавання інформації. Відношення реальної швидкості до номінальної називається ефективністю тракту передавання даних:

$$\alpha = \frac{v_p}{v_n} = \frac{t_{a0} + t_{k0}}{t_{a0} + t_{k0} + n\Delta t} \leq 1.$$

Зазвичай реальну швидкість у трактах передавання даних вибирають таким чином, щоб виконувалось співвідношення $t \leq t_n$ (де t_n – нормативний час передавання) за умови, що в каналах немає завад і відмов.

У трактах передавання даних (див. розд. 13) задана правильність забезпечується повторним передаванням інформації. Однак велика кількість цих повторень може привнести додатковий час до нормативного часу передавання заданого масиву інформації. Тому причини зниження надійності у системах зі зворотним зв'язком фактично зумовлюються перевищенням нормативного часу передавання.

Події, які призводять до повного або часткового припинення виконання системою заданих функцій або до перевищення значень робочих показників установлених границь, називають відмовою. У разі несправностей тракту не завжди виникає відмова. Пояснимо це прикладом.

Нехай потрібно передавати за кожний часовий інтервал тривалістю t дискретну інформацію обсягом W .

Передавання виконується зі швидкістю v (біт/с) у разі надмірності використуваного коду α , α – частка надмірних розрядів у кодовій комбінації. Час, який витрачається безпосередньо для передавання інформації обсягом W зі швидкістю v , становить $W(\alpha + 1)/v$. Зазвичай часовий інтервал тривалістю t (який називають інколи інтервалом обміну або передавання) з метою створення деякого резерву часу вибирають більшим ніж $W(\alpha + 1)/v$. Резерв часу на інтервалі тривалістю t

$$\theta = t - [W(\alpha + 1)/v].$$

Ця величина – максимально допустима сумарна тривалість усіх несправностей за час t . Отже, до відмов тракту будуть призводити всі ті несправності, сумарна тривалість яких перевищує θ . Тому цю величину можна називати критерієм ТПД. Із цього випливає, що залежно від величини θ одні і ті самі несправності тракту будуть або спричиняти, або не спричиняти відмови. Критерій відмови визначається значною мірою характером переданої інформації. Для повільно старіючої інформації (наприклад, систематичних даних) критерій відмови може становити десятки хвилин і навіть годин. Для передавання швидкостаріючої інформації (наприклад, у системах керування повітряним рухом, затримки пакетів мовних повідомлень і т. ін.) критерій відмови вимірюється секундами і навіть частками секунди.

15.2. Класифікація відмов

Відмови поділяють на поступові та раптові. Оскільки поступові відмови зазвичай виявляються під час випробувань елементів і пристроїв у процесі виробництва та регламентованих робіт у процесі експлуатації, то на надійність ТПД переважно впливають раптові відмови. Розрізняють також стійкі відмови та ті, що самоусуваються.

Системи передавання даних належать до класу відновлюваних систем. Тривалість стійкої відмови обумовлюється часом її усунення, який залежить від системи контролю, ліквідації обслуговуючого персоналу, наявності резервного та запасного обладнання, систем ремонту.

Тракт передавання даних складається з двох основних частин: каналів зв'язку й АПД. Тому надійність тракту визначається надійністю каналів зв'язку й АПД. Надійність каналів зв'язку має свої особливості щодо надійності електронної техніки, яку використовують для створення пристроїв АПД. Ці особливості полягають в утворенні різних систем передавання, довжина яких від декількох до багатьох тисяч кілометрів. Вони містять як лінійні споруди, так і станційні пристрої. Розглянемо питання надійності каналів зв'язку для найбільш складних і водночас найбільш

цікавих для передавання даних – каналів тональної частоти, створених на міжміських і кабельних магістралях.

Надійність каналів тональної частоти зазвичай визначають не тільки надійністю лінійних споруд і станційного обладнання, але й експлуатаційною надійністю, обумовленою системою організації технічної експлуатації і кваліфікацією технічного персоналу. Очевидно, доцільно визначати надійність каналів зв'язку на підставі статистичної обробки даних про їх відмови.

15.3. Характеристики надійності трактів передавання даних

Для оцінювання надійності ТПД телекомунікаційних систем розглянемо терміни і поняття теорії надійності:

- відмова – подія, яка полягає у повній або частковій втраті працездатності виробу (елементом, вузлом, пристроєм, системою);
- час роботи до відмови – випадковий інтервал часу від початку роботи виробу до першої відмови;
- час роботи між відмовами – випадковий інтервал часу від початку роботи після закінчення чергового відновлення до моменту наступної відмови;
- середній час між відмовами – математичне сподівання випадкового часу між відмовами;
- резервування – метод підвищення надійності пристроїв за допомогою додаткових (резервних) пристроїв;
- ковзний резерв – спосіб резервування групи основних пристроїв групою резервних, за якого будь-який резервний пристрій може бути підключений замість будь-якого з тих, що відмовили;
- навантажений резерв – спосіб резервування, за якого резервний пристрій знаходиться у робочому режимі;
- ненавантажений резерв – спосіб резервування, за якого резервний пристрій не може відмовити до моменту ввімкнення.

Імовірність безвідмовної роботи пристрою $P(t)$ – імовірність події, яка полягає в тому, що виріб працює безвідмовно протягом потрібного інтервалу часу t , починаючи з моменту часу $t = 0$. Відповідно до цього означення

$$P(t) = P\{\xi > t\} = 1 - F(t),$$

де ξ – випадкова тривалість часу безвідмовної роботи виробу до появи відмови; $F(t)$ – функція розподілу випадкової величини ξ .

Статистичне визначення імовірності безвідмовної роботи

$$\tilde{P}(t) = \frac{N(t)}{N(0)} = 1 - \frac{n(t)}{N(0)},$$

де $N(t)$ – це кількість справних виробів у момент часу, тобто кількість виробів, що не відмовили протягом потрібного інтервалу часу t ; $N(0)$ – кількість справних виробів у початковий момент часу $t = 0$; $n(t)$ – кількість виробів, що відмовили до моменту часу t .

Інтенсивність відмов виробів $\lambda(t)$ у момент часу t – умовна щільність імовірності відмови устаткування для моменту часу t за умови, що до моменту часу t відмови устаткування не сталось. Згідно з цим визначенням

$$\lambda(t) = \frac{1}{1 - F(t)} F'(t) = \frac{f(t)}{P(t)},$$

де $f(t)$ – щільність розподілу часу безвідмовної роботи. Статистичне визначення інтенсивності відмов виробу

$$\tilde{\lambda}(t) = \frac{n(t + \Delta t) - n(t)}{N(t)\Delta t} = -\frac{N(t + \Delta t) - N(t)}{N(t)\Delta t},$$

де $\tilde{\lambda}(t)$ – кількість виробів, справних до моменту часу t ; $n(t)$ – кількість виробів, які відмовили до моменту часу t .

Якщо за розрахунковими частковими значеннями інтенсивностей відмов для кожного проміжку часу побудувати гістограму і з'єднати ці значення плавною кривою, отримаємо функцію $\lambda(t)$ (рис. 15.3), яку називають λ -характеристикою. Ця характеристика, як видно з рис. 15.3, має три характерні ділянки: I – відповідна початковому періоду (періоду прироблення приладу); II – період нормальної експлуатації; III – зумовлена зносом і старінням масових елементів (резисторів, конденсаторів і т. ін.) виробів (устаткувань). Наявність горизонтальної ділянки кривої інтенсивності відмов устаткування значно спрощує розрахунок надійності систем.

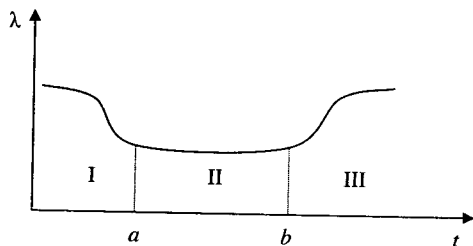


Рис. 15.3. λ -характеристика

Середній час роботи устаткування до першої відмови T_1 – математичне сподівання часу безвідмовної роботи устаткування до першої відмови:

$$T_1 = \int_0^{\infty} x \alpha F(x),$$

де $F(x)$ – функція розподілу часу безвідмовної роботи.

Статистичне визначення

$$T_1 = \frac{1}{N(0)} [\xi_1 + \xi_2 + \dots + \xi_{n(0)}] = \frac{1}{N(0)} \sum_{i=1}^{N(0)} \xi_i,$$

де $N(0)$ – загальна кількість устаткування; ξ_i – тривалість роботи до першої відмови устаткування.

Середній час між відмовами T – математичне сподівання часу безвідмовної роботи устаткування між двома послідовними відмовами;

$$T^{(k)} = \int_0^{\infty} x \alpha F_k(x); \quad T = \frac{1}{N} \sum_{k=1}^N T^k,$$

де $F_k(x)$ – функція розподілу часу безвідмовної роботи між k і $k+1$ відмовами; N – кількість спостережуваних відмов. Якщо $F_k(x) = F(x)$ для всіх k , тобто устаткування після ремонту повністю відновлює свої властивості, тоді $T = T_1$.

Статистичне визначення середнього часу між відмовами

$$T = \frac{1}{N} \sum_{k=1}^N \frac{1}{M} (\xi_k^{(1)} + \xi_k^{(2)} + \dots + \xi_k^{(M)}) = \frac{1}{N} \sum_{k=1}^N \frac{1}{M} \sum_{i=1}^M \xi_k^{(i)},$$

де M – загальна кількість устаткування; $\xi_k^{(i)}$ – випадкова величина часу між k і $k+1$ відмовами для i -го устаткування.

Інтенсивність відновлення устаткування t – умовна щільність імовірності відновлення для моменту часу t (відрахованого від моменту початку відновлення) за умови, що до моменту часу t відновлення не виконано. Якщо через $\Phi(t)$ позначити функцію розподілу часу відновлення, тоді

$$\mu(t) = \frac{\Phi'(t)}{1 - \Phi(t)}.$$

Статистичне визначення інтенсивності відновлення устаткування:

$$\tilde{\mu}(t) = \frac{n(t + \Delta t) - n(t)}{n(t)\Delta t},$$

де $N(t)$ – кількість виробів, невідновлених протягом інтервалу часу від 0 до t ; $n(t)$ – кількість виробів, для яких відновлення, що почалося у початковий момент, закінчилось до моменту t .

Середній час відновлення виробу τ – математичне сподівання часу відновлення:

$$\tau = \int_0^{\infty} x \alpha \Phi(x).$$

Статистичне визначення середнього часу:

$$\tau = \frac{1}{N} \{ \eta^{(1)} + \eta^{(2)} + \dots + \eta^{(N)} \} = \frac{1}{N} \sum_{i=1}^N \eta^{(i)},$$

де N – загальна кількість виробів; $\eta^{(i)}$ – час відновлення i -го виробу.

Коефіцієнт готовності K_r – імовірність знаходження виробу в робочому стані у довільний момент часу, «достатньо віддалений» від початку відліку. Для будь-яких функцій розподілу часу між відмовами $F(t)$ і часу відновлення $\Phi(t)$, які мають кінцеві значення T і τ відповідно, завжди буде

$$K_r = \frac{T}{T + \tau}.$$

Статистичне визначення коефіцієнта готовності:

$$K_r = \frac{N_{\infty}}{N_0},$$

де N_0 – загальна кількість виробів у початковий момент часу; N_{∞} – кількість виробів, працездатних у дозвільний «достатньо віддалений» момент часу.

Тракти передавання даних – відновлювані системи, роботу яких після усунення відмови може бути поновлено.

Тракт передавання даних у збільшеному вигляді можна зобразити як два послідовні з'єднання основних елементів: АПД і каналів зв'язку (рис. 15.4). Тому для визначення показників його надійності потрібно знати надійність АПД і каналів зв'язку.

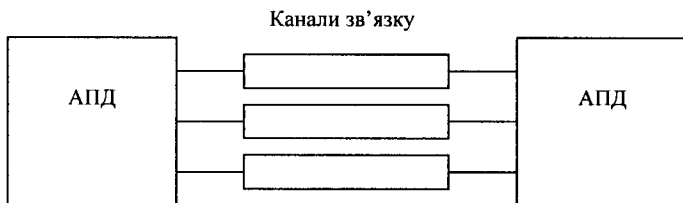


Рис. 15.4. Узагальнена функціональна схема ТПД

Апаратура передавання даних – це різновид радіоелектронної апаратури. Розрахунок надійності виконують різними методами, які докладно описано в літературі. Тому головну увагу приділимо розрахунку каналів зв'язку і ТПД.

15.4. Розрахунок надійності каналів передавання даних

Існує декілька різновидів методів розрахунку надійності, які відрізняються один від одного вихідними даними, метою розрахунку і стадією проектування, на якій цей розрахунок виконують. Суть розрахунку полягає у визначенні числових значень основних показників надійності (напрацювання на відмову і коефіцієнт готовності) за відомими λ -характеристиками.

Розрахунки надійності можна умовно поділити на дві групи: 1) приблизний або орієнтовний; 2) повний остаточний.

Щоб виконати приблизний розрахунок, треба знати усереднені значення інтенсивності відмов типових елементів певного типу в кожній групі N_i (у групу об'єднуються елементи, які мають однакову інтенсивність відмов). У ряді випадків для орієнтовного розрахунку надійності системи досить знати загальну кількість елементів у системі та дані про надійність аналогічної системи, отримані з досліду експлуатації.

Для повного розрахунку надійності потрібно мати дані про реальні режими роботи елементів системи і про залежність інтенсивності відмов елементів від температурних, електричних та інших режимів і навантажень.

Канали зв'язку, використовувані для передавання даних, мають довжину від декількох кілометрів до багатьох тисяч. Найбільш складні – канали міжміського зв'язку, які складаються із багатокілометрових ділянок кабелю і великої кількості проміжної апаратури, розосередженої на відстані сотень і тисяч кілометрів.

Надійність кінцевої і проміжної апаратури зв'язку можна розраховувати так само, як і надійність АПД. Подібний розрахунок надійності лінійних споруд зв'язку має значні труднощі через специфічність їх конструкцій та умов роботи. Тому найдоцільніше визначити надійність каналів зв'язку за результатами систематичної обробки даних про відмови каналів, отриманих шляхом спеціальних вимірів або на підставі узагальнення досвіду експлуатації. Аналіз статистичних даних показав, що відмови у телефонних каналах, утворені в кабельних магістралях зв'язку, можна поділити на три групи: тривалі (від декількох десятків хвилин до декількох годин), середньої тривалості (від 3 до 30 хв), короткочасні (менше 3 хв).

Тривалі відмови виникають дуже рідко як результат пошкодження кабелю та загальних станційних пристроїв, і для них характерний вихід із ладу всіх каналів магістралі.

Відмови середньої тривалості виникають зазвичай у разі пошкоджень окремих вузлів станційного обладнання, а також у результаті дій технологічного персоналу.

Численними є короткочасні відмови, які виникають у результаті пошкодження апаратури, дій обслуговуючого персоналу або імпульсних завад. Короткочасні відмови, які виникають через перерви та імпульсні завади, для розрахунків вважаються корельованими.

Розраховуючи надійність трактів різної структури (див. рис. 15.1), доводиться розглядати системи, що являють собою послідовно-паралельне з'єднання багатьох елементів (рис. 15.5).

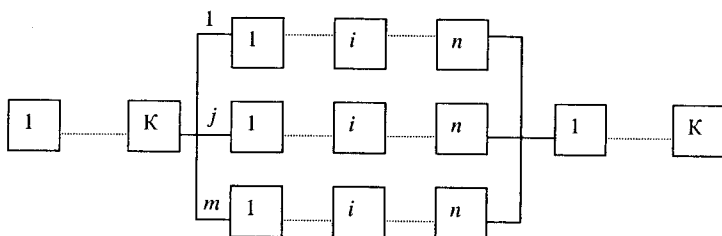


Рис. 15.5. Система з послідовно-паралельним з'єднанням пристроїв

Якщо функції розподілу часу безвідмовної роботи елементів показникові, то стандартний прийом розрахунку таких систем полягає у послідовному спрощенні початкової структури. Це спрощення полягає в тому, що сукупність послідовно з'єднаних елементів замінюється одним еквівалентним елементом, показники надійності якого розраховують відповідним чином.

Для розрахунків показників надійності елементів будемо користуватись такими позначеннями: $F(x)$ – функція розподілу часу безвідмовної роботи системи; $F_i(x)$ – функція розподілу часу безвідмовної роботи i -го елемента; $\Phi_i(x)$ – функція розподілу часу відновлення i -го елемента; β_i – коефіцієнт готовності i -го елемента; $\alpha_i = 1 - \beta_i$.

Очевидно, що

$$\beta_i = \frac{1}{\frac{\lambda_i}{\int_0^{\infty} x \alpha \Phi_i(x) dx + \frac{1}{\lambda_i}}}.$$

Якщо $\Phi_i(x) = 1 - e^{-\mu_i x}$, тоді

$$\beta_i = \frac{\frac{1}{\lambda_i}}{\frac{1}{\mu_i} + \frac{1}{\lambda_i}} = \frac{\mu_i}{\lambda_i + \mu_i}.$$

Розглянемо систему, яка складається з n послідовно з'єднаних елементів (рис. 15.6), і знайдемо її показники надійності.

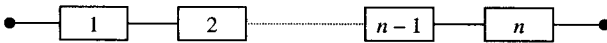


Рис. 15.6. Система з n послідовно з'єднаних елементів

Імовірність того, що система буде працювати безвідмовно протягом часу більше x , дорівнює (через незалежність відмов елементів)

$$1 - F(x) = \prod_{i=1}^n e^{-\lambda_i x} = e^{-x \sum_{i=1}^n \lambda_i},$$

звідки

$$1 - F(x) = 1 - e^{-x \sum_{i=1}^n \lambda_i} = 1 - e^{-\lambda x}, \quad (15.1)$$

де $\lambda = \sum_{i=1}^n \lambda_i$, тобто інтенсивність відмов системи дорівнює сумі інтенсивностей відмов елементів, що входять до неї. Коефіцієнт готовності системи знаходять із виразу (через незалежність відмов)

$$K_r = \prod_{i=1}^n \beta_i. \quad (15.2)$$

Із рівнянь (15.1) і (15.2) випливає, що систему, яка складається із послідовно з'єднаних елементів, можна замінити одним еквівалентним елементом з показниками надійності:

$$\beta = \prod_{i=1}^n \beta_i; \quad (15.3)$$

$$\lambda = \sum_{i=1}^n \lambda_i, \quad (15.4)$$

де β_i – коефіцієнт готовності i -го еквівалентного елемента; λ_i – інтенсивність відмов i -го еквівалентного елемента.

Отже, систему, яку показано на рис. 15.5, можна звести до системи, яка складається з $m + 1$ еквівалентних елементів (рис. 15.7). Показники надійності для i -го еквівалентного елемента такої системи розраховують за формулами (15.3) і (15.4). Схему, яку показано на рис. 15.7, можна перетворити ще раз, замінивши сукупність m паралельних елементів одним еквівалентним з показниками надійності $\beta^*_{2, \dots, m+1}, \lambda^*_{2, \dots, m+1}$.

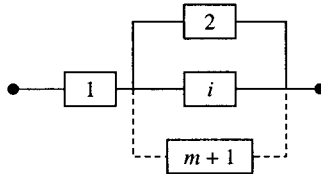


Рис. 15.7. Система з $m + 1$ еквівалентних елементів

Для пояснення методу пошуку показників надійності розглянемо систему, яка складається з L елементів. Припустимо, що визначено поняття відмови системи і потрібно знайти коефіцієнт готовності та середній час між відмовами системи. Кожний складовий елемент системи може знаходитись в одному із двох станів: справному і несправному. Тому загальна кількість станів дорівнює 2^L . Будемо вважати, що відмови елементів незалежні й усю множину станів системи Ω (вміщує 2^L елементів) розіб'ємо на дві підмножини: Ω_1 і Ω_2 .

Першу підмножину Ω_1 , яка являє собою сукупність станів системи для виконання покладених на неї функцій, назвемо множиною робочих станів. Друга підмножина Ω_2 , яка являє собою сукупність станів системи для невиконання покладених на неї функцій, назвемо множиною відмовних станів системи. Підмножина Ω_1 містить такі стани системи, коли вихід з ладу певного елемента спричиняє відмову всієї системи, тобто здійснюється перехід із робочого стану системи у відмовний. Множину таких

станів, яка є частиною підмножини Ω_1 , позначимо через Ω^*_1 ($\Omega^*_1 \subset \subset \Omega_1$) і назвемо множиною передвідмовних станів.

Нехай S_i позначає i -й стан системи, T_x – середній час роботи елемента з номером x , а τ_x – середній час відновлення елемента з номером x . Імовірність того, що у довільний момент часу елемент під номером x справний, очевидно, дорівнює $T_x/(T_x + \tau_x)$, а ймовірність того, що він несправний, – $\tau_x/(T_x + \tau_x)$.

Обидві ці можливості станів S_i можна записати одним виразом, якщо ввести величину X_{ix} :

$$X_{ix} = \begin{cases} T_x, & \text{якщо елемент під номером } x \text{ у стані } S_i \text{ справний;} \\ \tau_x, & \text{якщо елемент під номером } x \text{ у стані } S_i \text{ несправний.} \end{cases}$$

Тоді ймовірність стану S_i визначають так:

$$P(S_i) = \prod_{x=1}^l \frac{X_{ix}}{T_x + \tau_x}. \quad (15.5)$$

Коефіцієнт готовності представляється як ймовірність знаходження системи у множині станів Ω_1 і, оскільки всі ці стани незалежні (через незалежність відмов елементів), його можна подати у вигляді

$$K_r = \sum_{i \in \Omega_1} \prod_{x=1}^l \frac{X_{ix}}{T_x + \tau_x}. \quad (15.6)$$

Для визначення середнього часу між відмовами системи введемо поняття циклу роботи системи. Назвемо циклом роботи системи відрізок часу між двома послідовними відмовами (рис. 15.8). Математичне сподівання довжини цього відрізка позначимо через l .

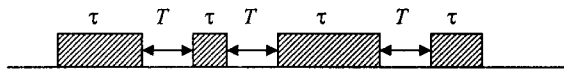


Рис. 15.8. Часова діаграма станів системи

Очевидно, що $l = T + \tau$, де τ – середній час відновлення системи. Тепер розглянемо досить великий відрізок часу і його довжину візьмемо за одиницю. Цей відрізок дорівнює середній довжині одного циклу, помноженій на середню кількість циклів, тобто

$$lN = 1, \quad (15.7)$$

де N – середня кількість циклів.

За один цикл система один раз переходить із робочого стану у відмовний, тому середня кількість циклів N дорівнює середній кількості переходів системи з множини передвідмовних станів Ω^*_1 у множину Ω_2 . Оскільки N являє собою середню кількість переходів, тоді, щоб її визначити, використаємо стандартний спосіб знаходження середнього: визначимо всі стани системи, із яких можливий безпосередній перехід у множину Ω_2 , і обчислимо кількість переходів із кожного такого стану в Ω_2 . Відповідно до початкового розбиття множини станів Ω ці стани містяться тільки у множині Ω^*_1 . Потім складемо суму із добутків імовірностей цих станів на відповідну їм кількість переходів. Ця сума і визначає середню кількість переходів системи у відмовну множину, тобто N .

Нехай S_i означає i -й передвідмовний стан системи, з якого в результаті відмови елемента під номером x вона переходить у j -й відмовний стан S_j . Очевидно, що середня кількість переходів із стану S_i в S_j дорівнює середній кількості відмов елемента x . Назвемо циклом елемента з номером x відрізок часу між двома моментами i відновлення елемента. Середня довжина цього циклу дорівнює $T_x + \tau_x$, і середня кількість циклів елемента з номером x за одиничний часовий інтервал буде $1/(T_x + \tau_x)$. Звідси середня кількість циклів дорівнює середній кількості відмов елемента під номером x , оскільки всередині одного циклу елемент відмовляє один раз. Імовірність стану S_i визначено за рівнянням (15.5). Тому середня кількість переходів N_{ij} із передвідмовного стану S_i у відмовний знаходимо як

$$N_{ij} = \frac{1}{T_x + \tau_x} \prod_{y \neq x} \frac{X_{ix}}{T_y + \tau_y} = \frac{1}{x_{ix}} \prod_{y=x}^l \frac{X_{ix}}{T_y + \tau_y} = \frac{1}{T_x} \prod_{y=x}^l \frac{X_{iy}}{T_y + \tau_y}.$$

Оскільки $N = \sum_{i,y} N_{ij}$, тоді

$$N = \sum_{i \in \Omega_1^*} \frac{1}{T_x} \prod_{y=1}^l \frac{X_{iy}}{T_y + \tau_y}.$$

Значимо, що завжди має місце співвідношення

$$K_r = \frac{T}{T + \tau} = \frac{T}{l}. \quad (15.8)$$

Із рівнянь (15.5), (15.7), (15.8) отримаємо

$$T = \frac{K_r}{N} = \frac{K_r}{\sum_{i \in \Omega_1^*} \frac{1}{T_x} \prod_{y=1}^l \frac{X_{iy}}{T_y + \tau_y}}. \quad (15.9)$$

Вирази (15.6) і (15.9) отримано в загальних припущеннях щодо характеру відмов елементів системи і її структури, тому вони мають загальний вигляд. Щоб визначити T , треба знайти всі передвідмовні стани системи і обчислити кількість переходів із кожного такого стану у відмовний, а потім скласти суму із добутоків імовірностей цих станів на відповідну їм кількість переходів. Ця сума і визначить середню кількість переходів системи у відмовний стан, тобто N . За формулою (15.9) можна знайти T .

Тепер повернімося до схеми, зображеної на рис. 15.7, і визначимо показники надійності еквівалентного елемента, що відповідає m паралельно ввімкненим елементам. Коефіцієнт готовності $\beta_{2, \dots, m+1}^*$ такого елемента дорівнює

$$\beta_{2, \dots, m+1}^* = 1 - \prod_{i=2}^{m+1} \alpha_i.$$

Середній час між відмовами цього елемента $T_{2, \dots, m+1}^* = 1/\lambda_{2, \dots, m+1}^*$ можна знайти описаним методом. Передвідмовними станами в системі, яка складається із m паралельно ввімкнених елементів, будуть стани, коли справний тільки один з елементів, наприклад, під номером k ($k = 2, \dots, m+1$), а останні $(m-1)$ елементів несправні. Імовірність такого стану, очевидно, дорівнює

$$\beta_k^* = \prod_{\substack{s=2 \\ s \neq k}}^{m+1} \alpha_s^*.$$

Перехід у відмовний стан відбувається кожного разу, коли відмовляє елемент під номером k . Середня кількість таких переходів дорівнює сумі середньої кількості відмов елемента під номером k , тобто λ_k^* . Звідси середня кількість переходів у відмовний стан

$$N = \lambda_k^* \beta_k^* \prod_{\substack{s=2 \\ s \neq k}}^{m+1} \alpha_s^*,$$

а середній час між відмовами

$$T_{2, \dots, m+1}^* = \frac{1 - \prod_{i=2}^{m+1} \alpha_i^*}{\sum_{k=2}^{m+1} \lambda_k^* \beta_k^* \prod_{\substack{s=2 \\ s \neq k}}^{m+1} \alpha_s^*}. \quad (15.10)$$

Тепер показники надійності системи, яку зображено на рис. 15.7, остаточно визначають як

$$K_r = \beta_1 \beta_{2, \dots, m+1}^*; T = 1/\lambda_1 + \lambda_{2, \dots, m+1}^*. \quad (15.11)$$

Вирази (15.11) визначають також показники надійності схеми, показаної на рис. 15.5.

15.5. Розрахунок надійності тракту передавання дискретної інформації

Розглянутий метод розрахунку надійності тракту застосуємо для випадку, коли ТПД складається із двох паралельних каналів передавання даних, коли надійність одного каналу не задовольняє задані вимоги. Інформація у такі канали надходить одночасно. У разі виходу з ладу одного з каналів або виникнення помилок в інформації, яка передається цим каналом, інформація продовжує надходити до користувача припиняється тільки тоді, коли одночасно виходять із ладу обидва канали. Таку подію називають відмовою тракту. Умовну схему для розрахунку надійності ТПД показано на рис. 15.9.

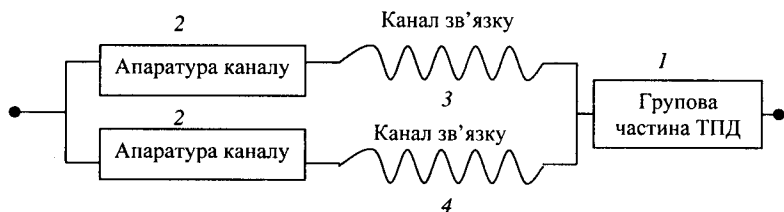


Рис. 15.9. Умовна схема для розрахунку надійності двоканального ТПД

Надалі з метою компактності запису будемо позначати коефіцієнт готовності i -го елемента через β_i , а коефіцієнт втрат $\alpha_i = 1 - \beta_i$. Показники надійності елемента 1 обчислюють за формулами (15.3) і (15.4) як показники надійності двох послідовно з'єднаних підкомплектів (пристрої групового керування та синхронізації каналів), а показники надійності елементів 2 як показники надійності двох послідовно з'єднаних підкомплектів ПЗП і ППС СПДІ.

Для остаточного розрахунку надійності двоканального тракту відповідно до еквівалентної схеми (рис. 15.9) потрібно ще знати показники надійності елементів 3 і 4 , тобто оцінити надійність кожного каналу зв'язку. Ця оцінка не буде однозначною і залежить від того, як організо-

вано обидва канали зв'язку: за однією системою ущільнення, за різними системами, але по одній магістралі або по географічно рознесених магістралях. Для спрощення методу припустимо, що ймовірність появи в тракті двох і більше перерв усередині одного інтервалу передавання інформації майже дорівнює нулю.

Припустимо, що ТПД функціонує тривалий час t і за цей час зафіксовано n перерв. Нехай $H(x)$ – функція розподілу тривалості перерв, а T – середній час між перервами будь-якої тривалості у тракті. Оскільки завжди $T \gg \tau$, де $\tau = \int_0^{\infty} x dH(x)$, то $T \approx t/n$. Із загальної кількості перерв n те-

пер треба вилучити перерви, тривалість яких менша за θ (критерій відмови). Імовірність того, що тривалість перерви буде менша ніж θ , є $H(\theta)$. Тому кількість перерв, тривалість яких менша за θ , становить від загальної кількості перерв величину $nH(\theta)$, а кількість перерв, тривалість яких більша ніж θ , дорівнює $n - nH(\theta)$. Звідси знаходимо середній час між перервами, тривалість яких більша за θ , тобто середній час між відмовами

$$T^* = \frac{t}{n - nH(\theta)} = \frac{T}{1 - H(\theta)}.$$

Надійність тракту будемо розраховувати за умови організації каналу по географічно рознесених магістралях. Розрахунок виконаємо відповідно до умовної схеми (рис. 15.9). Для цього заздалегідь обчислимо показники надійності елементів 3 і 4. Потім елементи 2 і 3 схеми (див. рис. 15.9) перетворюємо і замінюємо одним еквівалентним елементом 2* з показниками надійності:

$$\lambda_{2^*} = \lambda_2 + \lambda_3; \beta_{2^*} = \beta_2\beta_3; \alpha_{2^*} = 1 - \beta_2,$$

а елементи 2 і 4 – одним елементом 3* з показниками надійності:

$$\lambda_{3^*} = \lambda_2 + \lambda_4; \beta_{3^*} = \beta_2\beta_4; \alpha_{3^*} = 1 - \beta_3.$$

Після цього умовна схема для розрахунку надійності тракту набуває вигляду, як показано на рис. 15.10.

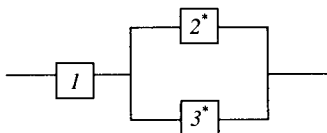


Рис. 15.10. Еквівалентна схема для розрахунку надійності двоканального тракту

Ця схема припускає подальші перетворення. Паралельно з'єднані елементи 2* і 3* замінюють еквівалентним елементом з показниками надійності:

$$\beta_{2,3} = 1 - \alpha_2 \cdot \alpha_3; \quad \lambda_{2,3} = \frac{\lambda_3 \cdot \alpha_3 \cdot \beta_3 + \lambda_2 \cdot \alpha_2 \cdot \beta_2}{1 - \alpha_2 \cdot \alpha_3}. \quad (15.12)$$

Вираз (15.12) отримуємо з рівняння (15.10). Тепер показники надійності схеми (див. рис. 15.10) визначаємо так:

$$K_r = \beta_1 \beta_{2,3}; \quad (15.13)$$

$$T = \frac{1}{\lambda_1 + \lambda_{2,3}}. \quad (15.14)$$

Приклад. Розрахуємо показники надійності тракту завдовжки 2000 км, схему організації якого показано на рис. 15.11.

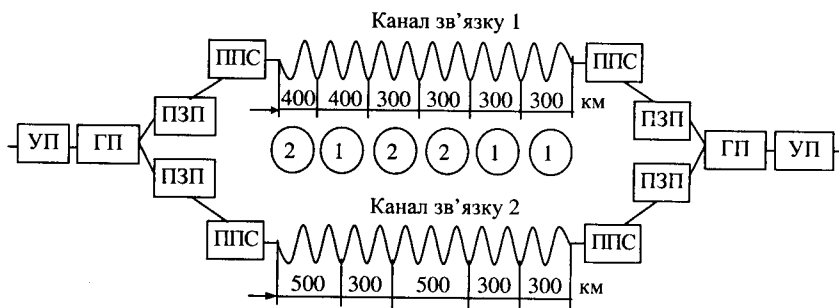


Рис. 15.11. Схема організації двоканального тракту, який проходить по різних магістралях: ГП – груповий пристрій; УП – узгоджувальний пристрій

Канали зв'язку складаються із декількох магістральних ділянок (один із шести ділянок, другий – із п'яти). Будемо враховувати, що кінцеві станції з'єднуються прямими каналами. Вихідні дані (умови) для розрахунку надійності окремих пристроїв АПД наведено у табл. 15.1.

Обсяг передаваної інформації по тракту з урахуванням уведеної надмірності ПЗП дорівнює 480 000 біт за кожні 20 с. Швидкість передавання – 48 000 біт/с.

Таблиця 15.1

| Показник | УП | ГП | ПЗП | ППС |
|-----------|--------|--------|-------|-------|
| K_r | 0,9997 | 0,9999 | 0,995 | 0,995 |
| T , год | 5000 | 5000 | 300 | 500 |

Розрахунок будемо виконувати в такій послідовності:

1. Визначаємо критерій відмови за формулою (15.13). Отримаємо, що $\theta = 10$ с.

2. Визначаємо показники надійності елементів 1 і 2 (див. рис. 15.9):

$$\beta_1 = K_{r_{УП}}^2 \cdot K_{r_{ГП}}^2 = (0,9999)^2 \cdot (0,9997)^2 \approx 0,9994;$$

$$T_1 = \frac{1}{\lambda_1} = \frac{1}{2\lambda_{УП} + 2\lambda_{ГП}} = \frac{1}{2 \cdot 0,2 \cdot 10^{-4} + 2 \cdot 0,2 \cdot 10^{-4}} = 1250 \text{ год};$$

$$\beta_2 = K_{r_{ПЗП}}^2 \cdot K_{r_{ППС}}^2 = (0,995)^2 \cdot (0,995)^2 \approx 0,982;$$

$$T_2 = \frac{1}{\lambda_2} = \frac{1}{2\lambda_{ПЗП} + 2\lambda_{ППС}} = \frac{1}{2 \cdot 3 \cdot 10^{-3} + 2 \cdot 2 \cdot 10^{-3}} = 100 \text{ год}.$$

3. Визначаємо надійність каналу зв'язку 1. Перший канал зв'язку складається із шести ділянок різної довжини. Кожна ділянка може мати свою ВЧ систему ущільнення з різними статистичними даними за відмовами другої групи. Для цього прикладу використовуємо два типи систем ущільнення (І і II), які показано на рис. 15.12, і визначаємо показники надійності за розглянутою методикою.

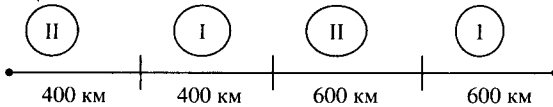


Рис. 15.12. Схема ділянок каналу I

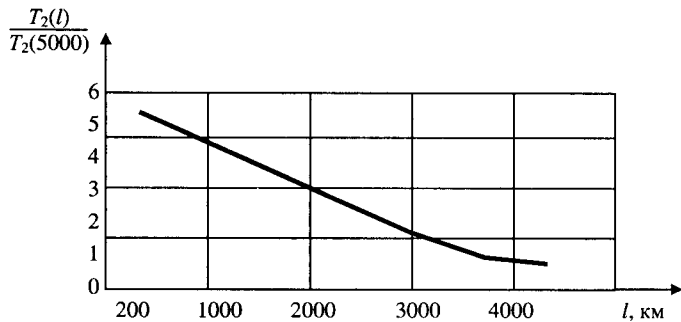
Класифікацію відмов ТПД встановлено три групи відмов за тривалістю відмови та причинами її виникнення. Для розрахунку надійності канал зв'язку умовно зображають у вигляді послідовного з'єднання п'яти елементів (рис. 15.13), кожний з яких відповідає одному з видів відмов.

Статистично зареєстровані (виміряні) середні тривалості інтервалів між сусідніми відмовами, а також середні тривалості відновлення відмов показали залежність середніх цих характеристик від довжини тракту.

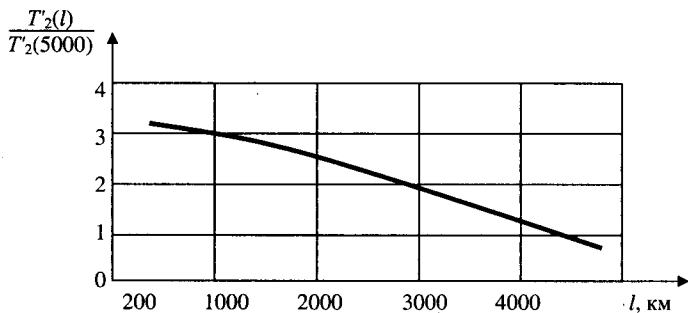
Центральним науково-дослідним інститутом зв'язку Росії встановлено приблизний характер залежності середнього часу між двома некорельованими відмовами другої групи (T_2) від довжини лінії (рис. 15.14, а), а також залежність для корельованих відмов (T'_2) – рис. 15.14, б.



Рис. 15.13. Еквівалентна схема для розрахунку надійності каналу зв'язку



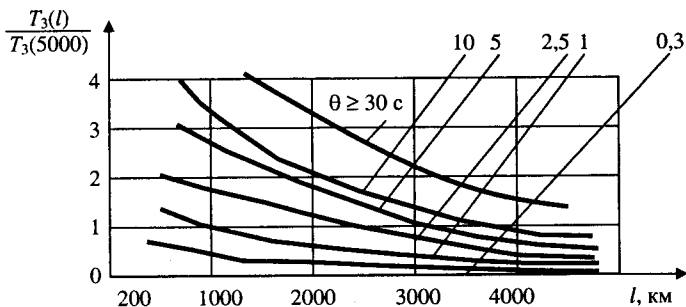
a



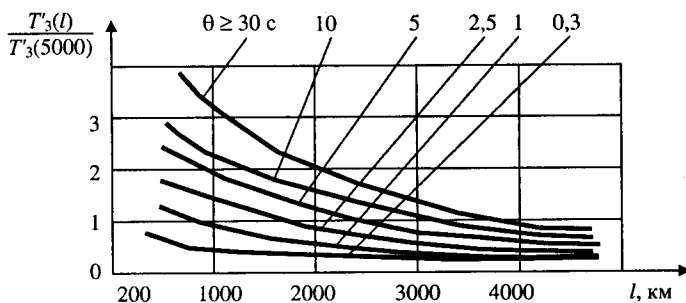
б

Рис. 15.14. Графік залежності середнього часу між відмовами другої групи від довжини магістралі: *a* – некорельованими; *б* – корельованими

Для найбільш чисельних короткочасових відмов тривалістю від часток секунди до 3 хв включно приблизний для T_3 і T'_3 вигляд мають відповідні залежності (рис. 15.15).



a



б

Рис. 15.15. Графік залежності середнього часу між короткотерміновими відмовами від довжини магістралі: *a* – некорельованими; *б* – корельованими

Середня тривалість короткотермінових відмов суттєво залежить від критерію відмови і майже не змінюється від довжини магістралі. Приблизну залежність середньої тривалості короткотермінових перерв τ_3 від критерію θ показано на рис. 15.16.

Розрахунок T і τ для визначення відмов кожної групи виконано за формулою (15.15). Відповідно до розглянутої методики та статистичних графіків (рис. 15.14–15.16) маємо:

$T_1 = 10950$ год; $\tau_1 = 2,5$ год; $T_2 = 74$ год; $T'_2 = 410$ год; $\tau_2 = \tau'_2 = 0,5$ год;
 $T_3 = 45$ год; $T'_3 = 275$ год; $\tau_3 = \tau'_3 = 0,011$ год.

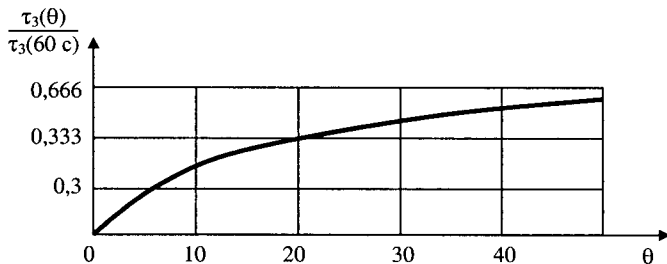


Рис. 15.16. Графік залежності середньої тривалості короткотермінових відмов від критерію відмови

Показники надійності першої ділянки (див. рис. 15.12) визначають так:

$$K_r^{(I)} = \frac{T_1}{(T_1 + \tau_1)} \cdot \frac{T_2}{(T_2 + \tau_2)} \cdot \frac{T'_2}{(T'_2 + \tau'_2)} \cdot \frac{T_3}{(T_3 + \tau_3)} \cdot \frac{T'_3}{(T'_3 + \tau'_3)} =$$

$$= \frac{10950 \cdot 74 \cdot 410 \cdot 45 \cdot 275}{(10950 + 2,5) \cdot (74 + 0,5) \cdot (410 + 0,5) \cdot (45 + 0,011) \cdot (275 + 0,011)} =$$

$$= 0,9912;$$

$$\lambda^{(I)} = \frac{1}{T_1} + \frac{1}{T_2} + \frac{1}{T'_2} + \frac{1}{T_3} + \frac{1}{T'_3} = \frac{1}{10950} + \frac{1}{74} + \frac{1}{410} + \frac{1}{45} + \frac{1}{410} =$$

$$= 42,94 \cdot 10^{-3} \frac{1}{\text{год}};$$

$$T^{(I)} = \frac{1}{\lambda^{(I)}} = 23,28 \text{ год.}$$

Еквівалентна схема для розрахунків надійності другої ділянки має такий самий вигляд, як і для першої. Для спрощення прикладу припустимо, що системи ущільнення I і II приблизно однакові, тобто статистичні дані за короткотерміновими відмовами також приблизно однакові. Тоді показники надійності для другої ділянки будуть такими самими, як і для першої:

$$K_r^{(II)} = 0,9912; \lambda^{(II)} = 42,94 \cdot 10^{-3}; T^{(II)} = 23,28 \text{ год.}$$

Відповідно до еквівалентної схеми третьої ділянки (аналогічної схеми на рис. 15.14) так само, як і для двох перших ділянок знаходимо:

$$T_1 = 7300 \text{ год}; \tau_1 = 2,5 \text{ год}; T_2 = 71 \text{ год}; \tau_1 = 0,5 \text{ год}; T'_2 = 400 \text{ год}; \tau'_2 = 0,5 \text{ год}; T_3 = 37,5 \text{ год}; \tau_3 = 0,011 \text{ год}; T'_3 = 250 \text{ год}; \tau'_3 = 0,011 \text{ год.}$$

На основі цих величин визначаємо показники надійності третьої ділянки:

$$K_r^{(III)} = \frac{7300 \cdot 71 \cdot 400 \cdot 37,5 \cdot 250}{(7300 + 2,5) \cdot (71 + 0,5) \cdot (400 + 0,5) \cdot (37,5 + 0,011) \cdot (250 + 0,011)} = 0,9909;$$

$$\lambda^{(III)} = \frac{1}{T_1} + \frac{1}{T_2} + \frac{1}{T_2'} + \frac{1}{T_3} + \frac{1}{T_3'} = 47,23 \cdot 10^{-3} \frac{1}{\text{год}};$$

$$T^{(III)} = \frac{1}{\lambda^{(III)}} = 21,17 \text{ год.}$$

Остання, четверта, ділянка має однакову з третьою ділянкою довжину і системи ущільнення з приблизно однаковими статистичними характеристиками, тому її показники надійності такі:

$$K_r^{(IV)} = K_r^{(III)} = 0,9909; T^{(IV)} = T^{(III)} = 21,17 \text{ год.}$$

Тепер визначаємо показники надійності каналу зв'язку 1 (див. рис. 15.11) в цілому:

$$K_{г.к.з}^{(I)} = K_r^{(I)} K_r^{(II)} K_r^{(III)} K_r^{(IV)} = 0,9646;$$

$$\lambda_{г.к.з}^{(I)} = \lambda^{(I)} \lambda^{(II)} \lambda^{(III)} \lambda^{(IV)} = 180,34 \cdot 10^{-3} \frac{1}{\text{год}};$$

$$T_{к.з}^{(I)} = \frac{1}{\lambda_{к.з}^{(I)}} = 5,54 \text{ год}; E_{к.з}^1 = \frac{1 - K_{г.к.з}^{(I)}}{K_{г.к.з}^{(I)}} T_{к.з}^{(I)} = 0,2 \text{ год.}$$

4. Визначимо надійність каналу зв'язку II. Подамо його у вигляді трьох послідовних еквівалентних ділянок (рис. 15.17).

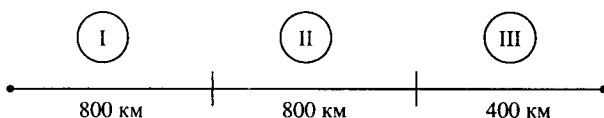


Рис. 15.17. Схема ділянок каналу зв'язку II

Розрахунок показників надійності для кожної із цих ділянок аналогічний розрахунку, виконаному для подібних ділянок каналу зв'язку 1. Нехай відповідно до методики і позначень, прийнятих раніше, отримано (з використанням графіків, подібних до графіків на рис. 15.14–15.16) такі значення T_i і τ_i :

$$T_1 = 5495 \text{ год}; \tau_1 = 2,5 \text{ год}; T_2 = 68 \text{ год}; T_2' = 385 \text{ год}; \tau_2 = \tau_2' = 0,5 \text{ год}; T_3 = 32,5 \text{ год}; \tau_3 = 0,011 \text{ год}; T_3' = 225 \text{ год}; \tau_3' = 0,011 \text{ год.}$$

На підставі цих значень визначаємо показники надійності першої ділянки:

$$K_r^{(I)} = \frac{5495 \cdot 68 \cdot 385 \cdot 32,5 \cdot 225}{(5495 + 2,5) \cdot (68 + 0,5) \cdot (385 + 0,5) \cdot (32,5 + 0,011) \cdot (225 + 0,011)} = 0,9904;$$

$$\lambda^{(I)} = 53,07 \cdot 10^{-3} \frac{1}{\text{год}};$$

$$T^{(I)} = \frac{1}{\lambda^{(I)}} = 18,84 \text{ год.}$$

Оскільки довжини другої та першої ділянок однакові, показники надійності другої ділянки такі самі, як і першої, тобто

$$K_r^{(II)} = 0,9904; T^{(II)} = 18,84 \text{ год.}$$

Третя ділянка має таку саму довжину, як і перша ділянка першого каналу зв'язку – 400 км. Тоді для третьої ділянки $K_r^{(III)} = 0,9912$; $T^{(III)} = 23,28$ год.

Тепер показники надійності каналу зв'язку II у цілому будуть мати такі значення:

$$K_{г.к.з}^{[2]} = K_r^{(I)} K_r^{(II)} K_r^{(III)} = 0,9722;$$

$$\lambda^{[1]} = \frac{1}{T^{(I)}} + \frac{1}{T^{(II)}} + \frac{1}{T^{(III)}} = 149,08 \cdot 10^{-3} \frac{1}{\text{год}};$$

$$T_{к.з}^{[2]} = 6,7 \text{ год};$$

$$\tau_{к.з}^{[2]} = \frac{1 - K_{г.к.з}^{[2]}}{K_{г.к.з}^{[2]}} T_{к.з}^{[2]} = 0,19 \text{ год.}$$

5. Розраховуємо показники надійності елементів 2* і 3* (див. рис. 15.10) за формулами (15.12) і (15.13). Результати розрахунку наведено у табл. 15.2.

Таблиця 15.2

| Елемент | λ , 1/год | β | α |
|---------|------------------------|---------|----------|
| 2* | $190,34 \cdot 10^{-3}$ | 0,9472 | 0,0528 |
| 3* | $159,08 \cdot 10^{-3}$ | 0,9547 | 0,0453 |

6. Розраховуємо показники надійності тракту в цілому за формулами (15.13) та (15.14) і отримуємо

$$K_{г.тр} = 0,997; T_{г.тр} = 61 \text{ год.}$$

Із виконаних розрахунків видно, що показники надійності тракту завдовжки 2000 км з критерієм відмови 10 с досить високі.

15.6. Вплив ненадійності трактів на ефективність передавання інформації

15.6.1. Загальні положення

Одна із важливіших характеристик СПДІ – надійність таких систем. У цьому розділі ми розглянули задачі і деякі методи їх розв’язання, що пов’язані з оцінюванням показників надійності СПДІ (див. підрозд. 15.5), а також виконали розрахунок надійності тракту передавання інформації за наявності двох територіально (географічно) рознесених каналів (трактів). Отримані показники дають переважно кількісну оцінку, тому впровадження високонадійних пристроїв, програмних продуктів, засобів захисту від впливу зовнішніх завад, а також кращої якості лінійних кабельних ліній з’єднання можна оцінити як збільшення (зменшення) показників надійності. Тому теорія надійності і займається методами, які підвищують надійність системи. Для діючих систем відомі показники їх надійності; нехай вони задовольняють критерій відмов. Виникає питання: якщо в робочому стані канал передавання інформації в деякі моменти часу припиняє передавання через технічні відмови, відновлення якого потребує часових витрат, у цьому інтервалі не використовується пропускна здатність каналу, то як урахувати відмови та відновлення для визначення основної характеристики системи передавання, тобто час доставляння повідомлення від джерела до користувача (або швидкість передавання інформації):

Відомо, що в мережах загального користування канал передавання інформації надається тимчасово користувачу для передавання його повідомлень. Процеси за часом надання тимчасових послуг на масові запити на ці послуги називають процесами масового обслуговування за часом, а дисципліна, яка вивчає динамічні процеси масового обслуговування, – теорією масового обслуговування.

Згідно з основними положеннями цієї теорії канал передавання інформації можна представити як обслуговуючий апарат із функцією розподілу $W(t) = P(\gamma < t)(t)$, де γ – час обслуговування повідомлення (час передавання повідомлення каналом) або функція тривалості обслуговування повідомлення.

Параметри надійності технічних пристроїв ТПД можна згорнути у функцію, яка описує обслуговування повідомлення абсолютно надійним трактом. Цю процедуру будемо називати збільшенням надійнісних параметрів, результатом дії якої є отримання нової функції розподілу часу обслуговування.

Збільшення надійнісних параметрів зводиться до пошуку функції розподілу $\Phi_i(t) = \Phi(w_i(t), F_1(t), F_2(t))$ та її характеристик. Процес обслуговування повідомлення (перебування на приладі) може припинитися внаслідок різних причин. Так, з моменту початку обслуговування повідомлення і до його закінчення може кілька разів перериватися процес обслуговування через усіяні причини, зокрема такі:

- виникнення відмов і потреба відновлень;
- виникнення збоїв як програмних, так і апаратних, з якими пов'язані витрати часу на діагностику, відновлення, прийняття рішень і т. ін.;
- надходження в систему пріоритетних повідомлень, що потребують негайного припинення обслуговування повідомлення, що знаходиться на обслуговуючому пристрої.

Причини, що викликають переривання процесу обслуговування повідомлення, можуть справляти як пасивний, так і активний вплив на подальший хід обслуговування повідомлення.

Пасивний вплив зумовлює тимчасове припинення обслуговування з наступним його продовженням з моменту його переривання. До пасивного впливу належать переривання типу запитів на введення-висновок, відмов пристроїв, деякі програмні збої і т. ін.

Активний вплив змушує припинити обслуговування, відновлення й обслуговування цього повідомлення спочатку. До активних переривань належать: втручання оператора, збій процесора, збої різних зовнішніх пристроїв, деякі програмні збої та ін.

Нехай задано:

$F_1(t)[H_1(t)]$ – функцію розподілу тривалості інтервалу безупинного обслуговування до моменту пасивного (активного) переривання обслуговування;

$F_2(t)[H_2(t)]$ – функцію розподілу інтервалу відновлення (усунення відмови, збою і т. ін.), що слідує за пасивним (активним) перериванням;

$w(t)$ – функція розподілу часу обслуговування повідомлення надійним приладом;

$\Phi(t)$ – функцію розподілу «сумарного» часу обслуговування (повний час перебування повідомлення на приладі до завершення обслуговування – ϕ).

Визначимо «сумарний» час перебування повідомлення на приладі $\phi(t)$ з моменту початку обслуговування до завершення для трьох випадків:

- 1) обслуговування переривається через пасивні переривання;
- 2) обслуговування переривається через активні переривання;
- 3) обслуговування переривається через пасивні й активні переривання.

15.6.2. Розподіл часу обслуговування у разі пасивних переривань

Сумарний час обслуговування з урахуванням відмов і відновлень визначимо таким способом. Нехай $x < t$ імовірність того, що час перебування повідомлення на приладі лежить у межах $[x, x + \Delta x] \in dw(x)$. Якщо відомо, що час перебування повідомлення на приладі дорівнює t , тоді через $P_k(x)$ позначимо ймовірність того, що виникне саме k переривань обслуговування. Для того, щоб сумарний час, витрачений на обслуговування повідомлення при фіксованих x і k не перевищував t , необхідно і достатньо, щоб за час $(t - x)$ прилад був відновлений k раз, тобто були обслуговані k переривань, імовірність чого є величина $F_2^*(t - x)$, де $F_2(t)$ – функція розподілу часу відновлення приладу (функція обслуговування пасивного переривання), тоді $\Phi(t)$ – функція розподілу загального часу обслуговування повідомлення з урахуванням переривань, і реакція на переривання буде мати вигляд

$$\Phi(t) = \int \sum_{k=0}^{\infty} F_2^{*k}(t-x) P_k(x) dw(x),$$

де * – символ стілттьєсівської згортки.

Переходячи до перетворень Лапласа, одержуємо

$$\varphi(s) = \int_0^{\infty} e^{-st} P(f_1(s), t) dt, \quad (15.15)$$

де

$$P(z, t) = \sum_{k=0}^{\infty} z^k P_k(t),$$

$$f_1(s) = \int_0^{\infty} e^{-st} dF_1(t).$$

Якщо потік пасивних переривань пуассонівський, то мають місце такі вирази:

$$F_1(t) = 1 - e^{-\lambda_1 t};$$

$$P_k(t) = \frac{(\lambda_1 t)^k}{k!} e^{-\lambda_1 t};$$

$$P(z, t) = e^{-\lambda_1 t(1-z)};$$

$$\begin{aligned}\varphi(s) &= \int_0^{\infty} e^{-st} e^{-\lambda_1 t(1-f_1(s))} dW(t); \\ P(f_1(s), t) &= e^{-\lambda_1 t(1-f_1(s))}; \\ \varphi(s) &= w(s + \lambda_1 - \lambda_1 f_1(s)).\end{aligned}$$

Звідси

$$\bar{\varphi} = \bar{w}(1 + \lambda_1 \bar{f}_1), \quad (15.16)$$

де $\bar{\varphi}$ – середнє значення сумарного часу обслуговування повідомлення на приладі; \bar{f} – середнє значення часу відновлення приладу; \bar{w} – середнє значення часу обслуговування повідомлення надійним приладом; λ_1 – інтенсивність виникнення відмов.

Отже, вираз (15.16) устанавлює зв'язок часу обслуговування повідомлення ненадійним приладом за відповідних значень інтенсивності відмов і тривалості відновлення (ремонту).

15.6.3. Розподіл часу обслуговування у разі активних переривань

Нехай $H_1(t)$ – функція розподілу часу виникнення активного переривання, $H_2(t)$ – функція розподілу часу реакції на переривання (відновлення). Тоді функція розподілу сумарного часу обслуговування повідомлення з урахуванням активних відмов і відновлень буде мати вигляд

$$\begin{aligned}\Phi(t) &= \int_0^t [1 - H_1(x)] dW(x) + \\ &+ \int_0^t [H_2(t-x) * \Phi(t-x)] [1 - w(x)] dH_1(x).\end{aligned} \quad (15.17)$$

Дійсно, для того щоб сумарний час обслуговування повідомлення не перевищував часу t (імовірність чого дорівнює $\Phi(t)$), необхідно і достатньо щоб або за час обслуговування повідомлення, що не перевищує t , не надійшло жодного активного переривання (імовірність чого дорівнює першому доданку в правій частині рівняння (15.16)), або щоб перше активне переривання виникло в проміжку $[x, x + \Delta x]$ $x \leq t$ і час реакції на переривання (обслуговування переривання) плюс нова реалізація сумарного часу обслуговування із самого початку не перевищували $(t-x)$ (імовірність чого дорівнює другій складовій правій частині рівняння (15.17)).

Нехай $\varphi(s)$, $h_1(s)$ і $h_2(s)$ – перетворення Лапласа – Стілтєса розподілів $\Phi(t)$, $H_1(t)$ і $H_2(t)$ відповідно. Тоді, розв'язавши рівняння (15.17) за допомогою перетворень Лапласа – Стілтєса, одержимо

$$\varphi(s) = \frac{w(s) - \delta(s)}{1 - h_2(s)[h_1(s) - \Delta(s)]}, \quad (15.18)$$

$$\text{де } \delta(s) = \int_0^{\infty} e^{-st} H_1(t) dw(t); \quad \Delta s = \int_0^{\infty} e^{-st} w(t) dH_1(t).$$

Для окремого випадку, коли потік активних переривань пуассонівський $H_1(t) = 1 - e^{-\lambda_2 t}$, визначимо $\varphi(s)$ за формулою (15.18). Вирази для $\delta(s)$, $\Delta(s)$ і $h_1(s)$ мають вигляд

$$\delta(s) = w(s) - w(s + \lambda_2); \quad (15.19)$$

$$\Delta(s) = \frac{\lambda_2}{\lambda_2 + s} w(s + \lambda_2);$$

$$h_1(s) = \frac{\lambda_2}{\lambda_2 + s}. \quad (15.20)$$

Підставляючи рівняння (15.20) у вираз (15.19), одержимо $\varphi(s)$ у вигляді

$$\varphi(s) = \frac{w(s + \lambda_2)}{1 - \lambda_2 h_2(s) \frac{1 - w(s + \lambda_2)}{s + \lambda_2}}. \quad (15.21)$$

Продиференціювавши (15.21) за s і поклавши $s = 0$, одержимо середню тривалість обслуговування $\bar{\varphi}$ повідомлення з урахуванням тимчасових витрат на відновлення:

$$\varphi(s) = \frac{1 + \lambda_2 \bar{h}_2 (1 - w(\lambda_2)) - (1 - \lambda_2) w(\lambda_2) - \lambda_2 \bar{w}(\lambda_2)}{\lambda_2 w(\lambda_2)}, \quad (15.22)$$

$$\text{де } w(\lambda_2) = \int_0^{\infty} e^{-\lambda_2 t} dw(t); \quad \bar{w}(\lambda_2) = \int_0^{\infty} e^{-\lambda_2 t} t dw(t); \quad \bar{h}_2 = - \int_0^{\infty} t dH_2(t).$$

Отже, вираз (15.22) являє собою середнє значення часу перебування повідомлення на приладі у разі відмов і відновлень приладу, причому час перебування повідомлення на приладі переривається, якщо прилад виходить з ладу. Після відновлення приладу повідомлення має обслуговуватися заново, тобто раніше витрачений час на обслуговування не враховується.

15.6.4. Розподіл часу обслуговування у разі активних і пасивних переривань

Нехай відомо, що обслуговування повідомлення потребує часу x , протягом якого можуть надходити як пасивні, так і активні переривання. Для $x < t$ імовірність того, що в процесі обслуговування не станеться жодного активне переривання, дорівнює $1 - H_1(t)$. Як і у разі виведення формули (15.15), неважко переконатися, що ймовірність повного обслуговування повідомлення за цей час x (з урахуванням реакції на пасивні переривання) дорівнює

$$\int_0^x \sum_{k=0}^{\infty} F_2^{*k}(x-y)P_k(y)dw(y),$$

де $dw(y)$ – імовірність того, що час перебування повідомлення на приладі лежить у межах $[y, y + \Delta y]$; $y < x$ – імовірність виникнення саме k пасивних переривань у цьому інтервалі; $F_2^{*k}(x-y)$ – імовірність того, що за час $(x-y)$ буде виконано k відновлень, що є потрібною і достатньою умовою завершення обслуговування повідомлення за час, що не перевищує x . Зазначимо, що якщо протягом часу $x < t$ відбулося активне переривання (імовірність чого дорівнює $dH_1(t)$, тоді ймовірність того, що за час, який залишився, $(t-x)$ буде відновлено прилад і завершено обслуговування повідомлення, якщо весь процес обслуговування повідомлення почнеться заново, дорівнює $H_2(t-x)\Phi(t-x)$.

Неважко також бачити, що ймовірність того, що до моменту виникнення активного переривання повідомлення не встигне повністю обслужитися, дорівнює відповідно

$$1 - \int_0^x \sum_{k=0}^{\infty} F_2^{*k}(x-y)P_k(y)dw(y).$$

Отже, функція розподілу $\Phi(t)$ з урахуванням як пасивних, так і активних переривань і відповідних їм відновлень задовольняє таке функціональне рівняння:

$$\begin{aligned} \Phi(t) = & \int_0^t [1 - H_1(x)] dx \left\{ \int_0^x F_2^{*k}(x-y)P_k(y)dw(y) \right\} + \\ & + \int_0^{\infty} H_2(t-x) \cdot \Phi(t-x) [1 - \int_0^x F_2^{*k}(x-y)P_k(y)dw(y)] dH_1(x). \end{aligned} \quad (15.23)$$

Дійсно, перший доданок у (15.23) враховує той випадок, коли обслуговування повідомлення не переривається активним перериванням, а другим доданок, мабуть, враховує той випадок, коли під час обслуговування повідомлення обов'язково виникне хоча б одне активне переривання.

Функціональне рівняння (15.23) щодо розподілу $\Phi(t)$ зважається за допомогою перетворення Лапласа – Стілтєса. Розглянемо для прикладу випадок, коли моменти виникнення переривань як пасивного, так і активного впливу утворюють найпростіші потоки з інтенсивностями λ_1 і λ_2 відповідно. У цьому разі $F_1(x) = 1 - e^{-\lambda_1 x}$, $H_1(x) = 1 - e^{-\lambda_1 x}$, $P_k(y) = \frac{(\lambda_1 y)^k}{k!} e^{-\lambda_1 y}$.

Тоді рівняння (15.23) набуде вигляду

$$\begin{aligned} \Phi(t) = & \int_0^t e^{-\lambda_2 x} dx \left\{ \int_0^x F_2^{*k}(x-y) \frac{(\lambda_1 y)^k}{k!} e^{-\lambda_1 y} dw(y) \right\} + \\ & + \lambda_2 \int_0^t H_2(t-x) \cdot \Phi(t-x) \left[1 - \int_0^x F_2^{*k}(x-y) \frac{(\lambda_1 y)^k}{k!} e^{-\lambda_1 y} dw(y) \right] e^{-\lambda_2 x} dx. \end{aligned} \quad (15.24)$$

Переходячи до перетворень Лапласа – Стілтєса, рівняння (15.24) запишемо у такому вигляді:

$$\varphi(s) = \frac{w(z)/z}{1 + \lambda_2 h_2(s) \left(\frac{w(z)}{z} - \frac{1}{\lambda_2 + s} \right)}, \quad (15.25)$$

де $z = s + \lambda_1 + \lambda_2 + \lambda_1 f_2(s)$.

Виконавши над рівнянням (15.25) дії, як і над (15.21), одержимо середнє значення сумарного часу $\bar{\varphi}$ перебування повідомлення на приладі, підданому впливу активних і пасивних переривань:

$$\bar{\varphi} = \frac{[1 - \beta(z_1)][1 + \lambda_2 \bar{h}_2]}{\lambda_2 \beta(z_1)}, \quad (15.26)$$

де $z = \lambda_2 + \lambda_1 - \lambda_1 f_2(\lambda_2)$.

Отримані формули (15.16), (15.22) і (15.26) дозволяють знайти середні значення тривалості обслуговування повідомлення ненадійним приладом. Відхилення від середнього значення обчислюють за відомими процедурами.

16. Елементи систематичного аналізу передавання даних системою передавання дискретної інформації

16.1. Основні положення системного аналізу

Сучасні телекомунікаційні системи являють собою складні комплекси, які складаються з різних функціонально залежних елементів. Ці системи характеризуються не тільки великою кількістю елементів, але й ієрархічністю структури, наявністю між елементами прямих, зворотних і перехресних зв'язків. До цього ми вивчали окремі пристрої та процеси їх функціонування у створеній на їх основі системі.

Основне призначення будь-якої ТС зв'язку – це забезпечення високонадійного, достовірного і своєчасного обміну повідомленнями – діалогового режиму взаємодії між джерелами інформації та її споживачем. З позиції призначення ТС тепер потрібно розглянути її функціонування в цілому, визначити алгоритм її функціонування з урахуванням взаємодії та властивостей складових її елементів і знайти шляхи її оптимізації. Вирішення таких завдань потребує системного підходу (системного аналізу).

Принцип системного підходу ґрунтується на представленні об'єкта як складної системи з урахуванням її специфічних зв'язків і властивостей. Систему визначають як цілісне утворення, яке складається із зв'язаних між собою елементів. Тому система має власні властивості, які не впливають безпосередньо із властивостей її елементів. Головна властивість системи – її закінченість, яку розглядають як цілісність. Концепція цілісності є основою методології системного підходу. Специфіка складної системи не вичерпується особливостями елементів, які її складають, а зв'язана з характером взаємозв'язків між елементами. Сукупність стійких зв'язків між елементами, що визначають цілісність і основні властивості системи, утворюють структуру системи.

Властивості системи передусім визначаються її цільовим призначенням (цілями функціонування), яке трактується як сукупність функцій, які виконує ця система. Щоб отримати бажаний результат, належить виконати певну сукупність операцій, спрямованих на досягнення поставленого завдання. Ці операції реалізуються використанням деяких ресурсів системи. Такими операціями, які виконує СПДІ, є кодування, модуляція, підсилення сигналу, демодуляція, декодування, селекція сигналів тощо, а ресурсами системи – потужність сигналу, смуга частот каналу (пропускна здатність) і нагромаджувачі повідомлень для тимчасового їх зберігання до моменту звільнення каналу від попереднього повідомлення.

Система передавання дискретної інформації як технічна система має ряд специфічних особливостей. Найсуттєвіші з них – це об'єкт (продукт) передавання і середовище (умова). Об'єкт передавання в СПДІ – інформація, яка надходить на її вхід у вигляді потоку повідомлень (даних) – потоку певних кодових послідовностей.

Середовище СПДІ – це не тільки лінія (середовище поширення сигналу), а й апаратура, що використовується для передавання сигналів від передавача та заважає впливу завад і спотворень, які перешкоджають якісному передаванню інформації каналами зв'язку. Потреба в запобіганні шкідливому впливу завад суттєво ускладнює СПДІ.

Для дослідження того чи того явища або технічного об'єкта (у цьому разі – СПДІ) створюють модель об'єкта. Модель являє собою відображення системи, її образ, використовуваний для аналізування реальної системи. Залежно від завдань і цілей моделювання об'єкта можна створювати на різних рівнях абстракції. Модель використовують для теоретичних і експериментальних досліджень системою. У процесі цих досліджень модель можна удосконалювати у напрямі якомога повнішого відображення властивостей реальної системи.

Модель – це частковий або повний опис системи у вигляді схем, рисунків, математичних співвідношень, імітаційних програм для комп'ютерів і т. ін. Математична модель технічної системи являє собою систему математичних співвідношень, які відображають структуру системи, алгоритм її функціонування, статистичні характеристики каналу, сигналу і завади, технічні й економічні показники системи.

Стохастичний характер завад і непередбаченість виникнення повідомлень і сигналів обумовлюють широке використання ймовірнісних моделей.

Система передавання дискретної інформації – це технічна система, призначена для передавання повідомлень від джерела інформаційних повідомлень (абонента системи) до споживача цих повідомлень (абонента системи). Абоненти системи є користувачами СПДІ, яка надає у тимчасове (на час передавання повідомлень) користування свої ресурси, тобто надає послуги користувачам – обслуговує користувача. Ефективність роботи СПДІ визначається її зайнятістю щодо надання послуг багатьом користувачам – масовістю користування. Системний підхід до розгляду функціонування СПДІ в часі і полягає в тому, що тут СПДІ виконує функції, зумовлені системою в разі масовості запитів на тимчасове користування її ресурсами. І не випадково, що для кількісного аналізу характеристик діяльності СПДІ використовують математичні методи випадкових процесів масового обслуговування, якими займається теорія масового обслуговування.

Розглянемо деякі процеси, які описують поведінку функціонування за часом СПДІ як системи масового обслуговування, на вхід якої надходить випадковий потік повідомлень, які підлягають передаванню каналом зв'язку. У СПДІ передбачено наявність обмеженої буферної пам'яті.

16.2. Моделювання дуплексного режиму роботи системи передавання дискретної інформації

Дуплексний режим роботи СПДІ передбачає наявність двох каналів зв'язку: один для передавання інформації в одному напрямі, другий – у зворотному. Розглянемо систему масового обслуговування типу $M|D|1|N$ (у позначеннях Кандела – Башарина), яка описує дуплексний режим функціонування СПДІ за часом, де M – пуассонівський стаціонарний потік з інтенсивністю λ ; D – стала величина τ (постійний час обслуговування одного повідомлення – час передавання пакета, байта, біта); 1 – один комплект передавально-приймальної апаратури СПДІ разом з каналом зв'язку; N – ємність буферної пам'яті.

Визначимо алгоритм оцінювання потрібної ємності пам'яті. Схему моделі дуплексного режиму СПДІ для одного напрямку зобразимо у вигляді рис. 16.1.

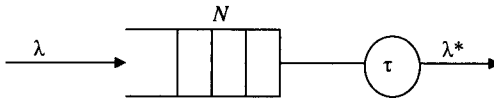


Рис. 16.1. Схематичне зображення функціонування СПДІ в одному напрямі

Оцінювання потрібної ємності буферної пам'яті. Організація черг у комп'ютерах, що функціонують під керуванням багатозадачної операційної системи, під час взаємодії комп'ютерів – термінальних устаткувань в інформаційно-комп'ютерній мережі, досягається виділенням буферної пам'яті. У буферній пам'яті нагромаджуються повідомлення (вимоги), що мають потребу в обслуговуванні тими самими ресурсами (пристроями). Обмін даними між пристроями комп'ютера чи комплексу мережі, передавання даних по лініях зв'язку здійснюються байтами. Байт складається з однакової (постійної) кількості бітів (двійкових розрядів). Швидкість оброблення одного байта пристроєм-каналом постійна. Тоді тривалість займання пристрою повідомленням за постійної швидкості оброблення (пропускної здатності) знаходиться в прямій залежності від обсягу повідомлення (кількості байтів у повідомленні).

Очевидно, що потрібна ємність буферної пам'яті залежить як від пропускної здатності відповідного ресурсу, так і від інтенсивності надходження повідомлень до цього ресурсу й обсягу повідомлення.

Процес нагромадження й оброблення повідомлень – це процес масового обслуговування, що описує поведження однолінійної системи масового обслуговування з обмеженою чергою, допускаючи відмову вимозі в постановці в чергу. Теоретично ймовірність відмови буде прагнути до нуля з ємністю буферної пам'яті, що прагне до нескінченності.

Стаціонарний режим функціонування систем масового обслуговування характеризується чергуванням інтервалів зайнятості (періодів) та інтервалів вільного стану.

Період зайнятості системи починається з моменту надходження повідомлення в обслуговуючий апарат, якщо застає апарат вільним і готовим негайно приступити до його обслуговування (тобто довжина черги дорівнює нулю). Період зайнятості закінчується в той момент, коли апарат завершує обслуговування деякого повідомлення, якщо в цей момент немає жодного повідомлення, що очікує обслуговування (тобто довжина черги знову дорівнює нулю).

Протягом інтервалу часу між закінченням одного періоду зайнятості і початком наступного апарат має чекати моменту надходження повідомлення.

Тривалість періоду зайнятості залежить від кількості місць у черзі. Максимальний період зайнятості досягається організацією нескінченної черги.

Нехай кількість місць у черзі обмежено числом N , тоді існує ймовірність відмовлення повідомленню в постановці в чергу. Ймовірність відмовлення $P_{\text{відм}}$ системою можна знайти, використовуючи ту очевидну обставину, що за великої тривалості процесу середня відносна кількість незареєстрованих повідомлень збігається за ймовірністю із ймовірністю відмовлення, тобто

$$P_{\text{відм}} = 1 - \frac{\mu[\eta]/\tau}{\lambda\{\mu[\eta] + \mu[\sigma]\}}, \quad (16.1)$$

де $\mu[\eta]$ – математичне сподівання тривалості періоду зайнятості приладу; $\mu[\sigma]$ – математичне сподівання тривалості інтервалу вільного стану приладу; $\mu[\eta]/\tau$ – середня кількість повідомлень, обслужених за період зайнятості приладу; λ – інтенсивність надходження повідомлень у систему.

Система знаходиться в стаціонарному режимі, тобто виконується умова

$$\rho = \lambda\tau < 1.$$

Щоб оцінити ємність буферної пам'яті, треба задати допустиму ймовірність відмовлення $P'_{\text{відм}}$. Тоді, з одного боку, з виразу (16.1) можна визначити середню тривалість періоду зайнятості системи, що забезпечує допустимий рівень відмовлення $P'_{\text{відм}}$:

$$\mu[\eta] = \frac{1 - P'_{\text{відм}}}{\frac{1}{\rho} - (1 - P'_{\text{відм}})} \mu[\sigma], \quad (16.2)$$

а з другого, – можна визначити середню тривалість періоду зайнятості залежно від кількості місць у черзі як функцію $\mu'[\eta] = F(\lambda, \tau, N)$.

Оскільки $\mu'[\eta] = F(\lambda, \tau, N)$ залежить від ємності буферної пам'яті, то потрібну ємність буферної пам'яті можна визначити таким способом: спочатку за формулою (16.2) обчислити середню тривалість періоду зайнятості системи, що забезпечує обслуговування вимог із заданою ймовірністю відмовлення, потім, послідовно збільшуючи ємність буферної пам'яті для $N = 1, 2, \dots, j$, визначити таке значення $\mu'[\eta]_j$, за якого виконуться умови

$$\mu'[\eta]_j \geq \mu[\eta].$$

Тоді число j і визначить ємність буферної пам'яті системи, що забезпечує обслуговування потоку повідомлень із заданою ймовірністю відмовлення. Визначення ємності буферної пам'яті пов'язано з обчисленням періоду зайнятості.

Період зайнятості системи з обмеженою буферною пам'яттю. Протягом періоду зайнятості системи може бути обслужено n повідомлень: $n = 1, 2, \dots$. Дослідження періоду зайнятості проведемо методом вкладених ланцюгів Маркова. Якщо в деякий момент часу в черзі знаходиться $0 \leq j < N$ повідомлень, то будемо вважати, що система знаходиться в стані E_j . Стан E_0 характеризує відсутність черги і наявність повідомлення в обслуговуючому апараті.

Уведемо такі позначення:

P_n – ймовірність того, що за період зайнятості обслужено рівно n повідомлень;

$\pi_0(n)$ – ймовірність того, що через n кроків система перейде в стан E_0 ;

$\pi_j(n)$ – ймовірність того, що через n кроків система перейде в стан E_j ;

$w_i(t)$ – ймовірність надходження i повідомлень за інтервал часу t .

Зведемо систему рівнянь, що описує стан системи на $(n + 1)$ -му кроці через стани системи на n -му кроці й імовірності переходу:

$$\begin{aligned}\pi_j(n+1) &= \sum_{i=0}^{j+1} \pi_i(n) w_{j+1-i}(\tau), \quad j = 0, 1, \dots, N-1; \\ \pi_N(n+1) &= \sum_{i=0}^N \pi_i(n) \left[1 - \sum_{k=0}^{N-1} w_k(\tau)\right].\end{aligned}\quad (16.3)$$

Узявши суму лівої та правої частин системи рівнянь (16.3), одержимо

$$\sum_{j=0}^N \pi_j(n+1) = \sum_{j=0}^N \pi_j(n) - \pi_0(n) w_0(\tau). \quad (16.4)$$

Пояснимо фізичний зміст виразу $\pi_0(n) w_0(\tau)$. За визначенням $\pi_0(n)$ – імовірність того, що через n кроків система перейде в стан E_0 , тобто в обслуговуючому апараті є повідомлення і немає черги. Отже, добуток імовірностей являє собою імовірність того, що період зайнятості закінчується під час обслуговування n -го повідомлення.

Рівняння (16.4) перепишемо з новими позначеннями:

$$q_{n+1} = q_n - p_{n+1}, \quad n = 0, 1, 2, \dots, \quad (16.5)$$

де
$$q_n = \sum_{j=0}^N \pi_j(n), \quad p_{n+1} = \pi_0(n) w_0(\tau). \quad (16.6)$$

Розв'язуємо рекурентні співвідношення (16.6) за методом твірних функцій, для чого i -те рівняння помножимо на s^i і складемо їх; тоді одержимо вираз

$$Q(s) - q_0 = sQ(s) - P(s), \quad (16.7)$$

де $Q(s) = \sum_{n=0}^{\infty} q_n s^n$; $P(s) = \sum_{n=0}^{\infty} p_n s^n$.

Розв'яжемо рівняння (16.7) відносно $P(s)$:

$$P(s) = q_0 - (1-s)Q(s). \quad (16.8)$$

За виразом (16.8) можна розподілити кількість повідомлень, обслугованих за період зайнятості системи, і його моменти. Імовірність p_n ($n = 1, 2, \dots$) обслуговування n повідомлень за період зайнятості можна знайти за формулою

$$p_n = \frac{1}{(n-1)!} \frac{d^{n-1}}{ds^{n-1}} Q(s) \Big|_{s=0} - \frac{1}{n!} \frac{d^n}{ds^n} Q(s) \Big|_{s=0}.$$

Математичне сподівання і дисперсію одержимо з таких виразів:

$$\begin{aligned}\mu\{P_n\} &= Q(s); \\ D\{P_n\} &= 2 \frac{d}{ds} Q(s) \Big|_{s=1} + Q(s) \Big|_{s=1} - [Q(s) \Big|_{s=1}]^2.\end{aligned}$$

Отже, визначивши $Q(s)$, можна знайти характеристики періоду зайнятості.

Для визначення твірної функції $Q(s)$ заздалегідь розкладемо її на складові:

$$Q(s) = Q_0(s) + Q_1(s) + \dots + Q_N(s).$$

Згідно з позначеннями, уведеними в рекурентному співвідношенні (16.5),

$$Q_j(s) = \sum_{n=0}^{\infty} \pi_j(n) s^n, \quad j = 0, 1, 2, \dots, N. \quad (16.9)$$

Помножимо ліву частину системи рівнянь (16.9) на матрицю перехідних імовірностей, елементи якої заздалегідь помножені на параметр s , тоді рівняння (16.9) у матричній формі набуває вигляду

$$sPQ = sPZ,$$

де $Q = (Q_0(s), Q_1(s), \dots, Q_N(s))^T$ – вектор-стовпець; $Z = (\sum_{n=0}^{\infty} \pi_0(n) s^n,$

$\sum_{n=0}^{\infty} \pi_1(n) s^n, \dots, \sum_{n=0}^{\infty} \pi_N(n) s^n)$ – вектор-стовпець; P – матриця перехідних імовірностей, елементи якої відповідають елементам співвідношень (16.3).

Виконуючи множення в обох частинах матричного рівняння (16.9) і використовуючи співвідношення (16.3), одержимо таку систему рівнянь:

$$\begin{aligned}\pi_0(0) &= Q_0(s) - s \sum_{i=0}^1 w_{1-i}(\tau) Q_i(s); \\ \pi_j(0) &= Q_j(s) - s \sum_{i=0}^{j+1} w_{j+1-i}(\tau) Q_i(s), \quad j = 1, 2, \dots, N-1; \\ \pi_N(0) &= Q_N(s) - s \sum_{i=0}^N (1 - \sum_{k=0}^{N-1} w_k(\tau)) Q_i(s).\end{aligned} \quad (16.10)$$

Період зайнятості починається з моменту надходження повідомлення в систему. Вектор початкових станів має вигляд

$$\pi = (\pi_0(0), \pi_1(0), \dots, \pi_N(0))^T = (1, 0, \dots, 0)^T. \quad (16.11)$$

Склавши рівняння (16.10) і використавши результати (16.3) та (16.11), одержимо твірну функцію $Q(s)$ у вигляді

$$Q(s) = \frac{1}{1-s} (1 - \varphi(s)),$$

де $\varphi(s) = s w_0(\tau) Q_0(s)$.

Визначивши $Q_0(s)$ із системи рівнянь (16.10), одержимо твірну функцію

$$\varphi(s) = -s w_0(\tau) \frac{\det[sP - E]_{N-1}}{\det[sP - E]_N}, \quad (16.12)$$

де P – матриця перехідних імовірностей; E – одинична матриця.

Значення $\frac{d^n}{ds^n} Q(s) \Big|_{s=0}$ можна знайти з виразів:

$$\begin{aligned} \frac{d}{ds} Q(s) \Big|_{s=0} &= 1 - \varphi(s) \Big|_{s=0} - \frac{d}{ds} \varphi(s) \Big|_{s=0}; \\ \frac{d^n}{ds^n} Q(s) \Big|_{s=0} &= n! \left[1 - \varphi(s) \Big|_{s=0} - \frac{d}{ds} \varphi(s) \Big|_{s=0} \right] - \\ &- \sum_{i=0}^{n-1} \frac{n!}{(n+1-i)!} \frac{d^{n+1-i}}{ds^{n+1-i}} \varphi(s) \Big|_{s=0}. \end{aligned} \quad (16.13)$$

Підставивши значення (16.13) у вираз (16.8), одержимо формулу для визначення P_n у вигляді

$$P_n = \frac{1}{n!} \frac{d^n}{ds^n} \varphi(s) \Big|_{s=0}. \quad (16.14)$$

Підставивши $s = 1$ у вираз (16.12), одержуємо $\varphi(s) \Big|_{s=1} = 1$.

Застосовуючи правило Лопітала до рівняння (16.12), одержуємо формулу для визначення $\mu\{P_n\}$:

$$Q(s) \Big|_{s=1} = \lim_{s \rightarrow 1} \frac{1 - \varphi(s)}{1 - s} = \frac{d}{ds} \varphi(s) \Big|_{s=1}.$$

Визначимо дисперсію, для цього знайдемо $\frac{d}{ds} Q(s)$:

$$\frac{d}{ds} Q(s) = \frac{1}{(1-s)^2} \left[1 - \varphi(s) - (1-s) \frac{d}{ds} \varphi(s) \right].$$

Оскільки $\varphi(s)|_{s=1} = 1$, значення $\left. \frac{d}{ds} Q(s) \right|_{s=1}$ знайдемо, розкриваючи невизначеність за правилом Лопіталія:

$$\left. \frac{d}{ds} Q(s) \right|_{s=1} = \lim_{s \rightarrow 1} \frac{1}{(1-s)^2} \left[1 - \varphi(s) - (1-s) \frac{d}{ds} \varphi(s) \right] = \left. \frac{1}{2} \frac{d^2}{ds^2} \varphi(s) \right|_{s=1}.$$

Вираз для визначення дисперсії в остаточному вигляді запишемо так:

$$D\{P_n\} = \left. \frac{d}{ds} \varphi(s) \right|_{s=1} + \left. \frac{d^2}{ds^2} \varphi(s) \right|_{s=1} - \left(\left. \frac{d}{ds} \varphi(s) \right|_{s=1} \right)^2.$$

Розподіл кількості повідомлень у період зайнятості. Наведемо остаточні результати для розподілу кількості повідомлень, обслужених за період зайнятості, і його моменти (математичне сподівання і дисперсія). Імовірності $w_i(\tau)$ визначають з виразу

$$w_i(\tau) = \frac{(\lambda\tau)^i}{i!} e^{-\lambda\tau}.$$

Покладемо, що τ дорівнює одній тимчасовій одиниці.

Для визначення $\varphi(s)$ потрібно знайти визначники матриць $[sP - E]_N$ і $[sP - E]_{N-1}$ і їх значення підставити у вираз (16.12). Тоді розподіл кількості повідомлень можна знайти із формули (16.14). Визначення ймовірностей P_n для різної кількості місць у черзі зведено в табл. 16.1.

Таблиця 16.1

| | $N = 1$ | $N = 2$ | $N = 3$ | $N = 4$ | | $N = k$ |
|--------|---|-------------------------|--|--|--|---|
| PP_1 | $e^{-\lambda}$ | $e^{-\lambda}$ | $e^{-\lambda}$ | $e^{-\lambda}$ | | $e^{-\lambda}$ |
| PP_2 | $(1 - e^{-\lambda})e^{-\lambda}$ | $\lambda e^{-2\lambda}$ | $\lambda e^{-2\lambda}$ | $\lambda e^{-2\lambda}$ | | $\lambda e^{-2\lambda}$ |
| PP_3 | $(1 - e^{-\lambda})^2 e^{-\lambda}$ | $P_{3,2}$ | $3 \frac{\lambda^2}{2!} e^{-3\lambda}$ | $3 \frac{\lambda^2}{2!} e^{-3\lambda}$ | | $3 \frac{\lambda^2}{2!} e^{-3\lambda}$ |
| PP_4 | $(1 - e^{-\lambda})^3 e^{-\lambda}$ | $P_{4,2}$ | $P_{4,3}$ | $4^2 \frac{\lambda^3}{3!} e^{-4\lambda}$ | | $4^2 \frac{\lambda^3}{3!} e^{-4\lambda}$ |
| ... | | | | | | |
| ... | | | | | | |
| PP_k | $(1 - e^{-\lambda})^{k-1} e^{-\lambda}$ | $P_{k,2}$ | $P_{k,3}$ | $P_{k,4}$ | | $\frac{(\lambda k)^{k-1}}{k!} e^{-\lambda k}$ |

Цю таблицю жирною лінією розділено на дві частини. Верхню частину ($N \geq n - 1$) можна обчислити за формулою

$$P_n = \frac{(\lambda n)^{n-1}}{n!} e^{-\lambda n}. \quad (16.15)$$

Таке розділення таблиці зумовлено обмеженням за кількістю місць у черзі. Верхня частина таблиці відображає ті випадки, коли кількість повідомлень, обслужених за період зайнятості, не більша за кількість місць у черзі, при цьому відмовлень в обслуговуванні повідомлень, що надходять, немає. Як видно з виразу (16.15), імовірності обслуговування визначеної кількості повідомлень за період зайнятості не залежать від кількості місць у черзі N . Цей факт інтуїтивно зрозумілий, бо коли немає відмовлень повідомленням в обслуговуванні, зазначені ймовірності не залежать від кількості місць у черзі. Для окремого випадку $N = \infty$ (системи з необмеженою чергою) формула (16.15) збігається із формулою Бореля.

Нижню частину таблиці ($n - 1 > N \geq 1$) для різних N обчислено за формулою (16.14). Прості формули для обчислення $P_{n,N}$ знайти не вдалося. Для окремих випадків $N = 1, 2, 3$ будуть обчислені характеристики далі в прикладах.

Середню кількість повідомлень, обслужених за період зайнятості, після обчислення першої похідної $\varphi(s)$ можна знайти за виразом

$$\mu\{P_n\} = \frac{d}{ds} \varphi(s) \Big|_{s=1} = \sum_{i=0}^N \frac{[\lambda(i-N)]^i}{i!} e^{-\lambda(N-i)}, \quad N = 1, 2, \dots \quad (16.16)$$

Підставляючи рівняння (16.16) у вираз (16.1) і зважаючи на те, що величина $\mu[\sigma] = 1/\lambda$, для найпростішого потоку знайдемо $P_{\text{відм}}$ у вигляді

$$P_{\text{відм}} = 1 - \frac{1}{\lambda\tau + \frac{1}{\sum_{i=0}^N e^{(N-i)\lambda\tau} \frac{[(i-N)\lambda\tau]^i}{i!}}}$$

Приклади.

1. Для випадку $N = 1$

$$\varphi(s) = \frac{se^{-\lambda}}{s(e^{-\lambda} - 1) + 1}.$$

Похідна n -го порядку $\varphi(s)$

$$\frac{d}{ds^n} \varphi(s) = \frac{n!(1 - e^{-\lambda})^{n-1} e^{-\lambda}}{[s(e^{-\lambda} - 1) + 1]^{n+1}}.$$

Імовірності P_n обчислюємо за виразом

$$P_n = \frac{1}{n!} \frac{d^n}{ds^n} \varphi(s) \Big|_{s=0} = (1 - e^{-\lambda})^{n-1} e^{-\lambda}.$$

2. Для випадку $N = 2$

$$\varphi(s) = \frac{s^2(e^{-2\lambda} - e^{-\lambda}) + se^{-\lambda}}{s^2(e^{-2\lambda} + \lambda e^{-\lambda} - e^{-\lambda}) + s(e^{-\lambda} - \lambda e^{-\lambda} - 1) + 1}.$$

Визначимо ймовірності $P_{3,2}$; $P_{4,2}$; $P_{5,2}$:

$$P_{3,2} = \frac{1}{3!} \frac{d^3}{ds^3} \varphi(s) \Big|_{s=0} = e^{-2\lambda} + e^{-3\lambda}(\lambda^2 - \lambda - 1);$$

$$P_{4,2} = \frac{1}{4!} \frac{d^4}{ds^4} \varphi(s) \Big|_{s=0} = e^{-2\lambda} + (\lambda - 2)e^{-3\lambda} + (\lambda^3 - 2\lambda^2 - \lambda + 1)e^{-4\lambda};$$

$$P_{5,2} = \frac{1}{5!} \frac{d^5}{ds^5} \varphi(s) \Big|_{s=0} = e^{-2\lambda} + (\lambda + \lambda^2)e^{-3\lambda} + (2\lambda^3 - \lambda^2)e^{-4\lambda} + (\lambda^4 - 3\lambda^3 + 4\lambda^2 - \lambda - 1)e^{-5\lambda}.$$

3. Для випадку $N = 3$

$$\varphi(s) = (s^3(e^{-3\lambda} + \lambda e^{-2\lambda} - e^{-2\lambda}) + s^2(e^{-2\lambda} - \lambda e^{-2\lambda} - e^{-\lambda}) + se^{-\lambda}) / (s^3(e^{-3\lambda} - e^{-2\lambda} + \lambda e^{-2\lambda} + \frac{\lambda^2}{2!}e^{-2\lambda}) + s^2(e^{-2\lambda} - \lambda e^{-2\lambda} + \frac{\lambda^2}{2!}e^{-\lambda}) + s(e^{-\lambda} - 2\lambda e^{-\lambda} - 1) + 1).$$

Імовірність $P_{4,3}$ обчислюємо за виразом

$$P_{4,3} = \frac{1}{4!} \frac{d^4}{ds^4} \varphi(s) \Big|_{s=0} = (1 + \lambda - \lambda^2)e^{-3\lambda} + (3\frac{\lambda^3}{2!} + 5\frac{\lambda^2}{2!} - 2\lambda + 1)e^{-4\lambda}.$$

4. Визначити ємність буферної пам'яті системи обслуговування, на вхід якої надходить найпростіший потік інтенсивністю $\lambda = 0,4$ 1/с, тривалість обслуговування $\tau = 1$ с, імовірність відмовлення в обслуговуванні має не перевищувати $P_{\text{відм}}^6 = 10^{-6}$. Як легко переконатися, тут виконуються потрібні умови існування стаціонарного режиму (у середньому за одиницю часу надходить у систему менше повідомлень, ніж може обслужити система).

Обчислимо середню тривалість періоду зайнятості системи, що забезпечує обслуговування повідомлень з імовірністю відмовлень 10^{-6} за формулою

$$\mu[\eta] = \frac{1 - P_{\text{відм}}^6}{1 - \lambda(1 - P_{\text{відм}}^6)} \approx 1,666664.$$

Послідовно обчислюючи при $N = 1, 2, \dots$ середню тривалість періоду зайнятості системи за формулою (16.1), одержуємо, що якщо $N = 4$, то

$$\sum_{i=0}^4 e^{(4-i)0,4} \frac{[(i-4)0,4]^i}{i!} \approx 1,668914.$$

Отже, система із заданими параметрами $\lambda = 0,4$ 1/с, $\tau = 1$ с і обчисленою ємністю буферної пам'яті $N = 4$ забезпечує обслуговування повідомлень з імовірністю відмовлення $P_{\text{відм}}^6 = 10^{-6}$.

Оцінювання використовуваної ємності буферної пам'яті. У попередньому розділі було розглянуто алгоритм одностороннього передавання даних за обмеженої ємності буферної пам'яті для тимчасового зберігання повідомлення до моменту звільнення каналу передавання від попереднього повідомлення. Якщо ресурси буферної пам'яті необмежені, виникає потреба в оцінюванні використовуваної пам'яті та затримки повідомлень у черзі (у буферній пам'яті) маршрутизаторів (комутаторів) магістральних ліній телекомунікаційної мережі. Вивчаючи поведінку одного з вихідних каналів маршрутизатора, можна уявити, що цей канал сумісно з маршрутизатором і АПД може знаходитись в одному зі станів – зайнятий процесом передавання і вільний від передавання, причому маршрутизатор періодично буде організовувати зв'язок із сусіднім маршрутизатором мережі. Таким чином, маршрутизатор періодично звертається до спостережуваного каналу для передавання через нього повідомлень, що надійшли на його адресу.

Процес організації передавання повідомлень через вибраний канал передавання даних здійснюється протоколом керування інформаційним каналом (див. розд. 12), який можна подати як однолінійну систему масового обслуговування, на вхід якої надходять заявки (повідомлення) у вигляді пуассонівського стаціонарного потоку з параметром λ . Ці повідомлення нагромаджуються в буферній пам'яті – чекають у черзі звільнення каналу (обслуговуючого апарата), або надання його маршрутизатором. Маршрутизатор згідно з протоколом керування устанавлює зв'язок з вибраним сусіднім маршрутизатором. Отже, такий процес взаємодії можна подати як систему масового обслуговування типу $M|G|1|\infty$. Особливість такої системи полягає в тому, що обслуговування розпочинається не

з моменту надходження заявки, а з моменту встановлення зв'язку. Ілюстрацію цього процесу показано на рис. 16.2. На ньому використано такі позначення: τ_j – інтервал часу, який витрачається на передавання k_{j-1} повідомлень, що надійшли за цикл c_{j-1} (час обслуговування черги); w_j – час, який витрачає маршрутизатор на виконання робіт, не зв'язаних з керуванням передавання повідомлень каналом (знаходиться в стані «перемикання»). Момент закінчення часу «перемикання» – момент часу надходження обслуговуючого апарата до черги.

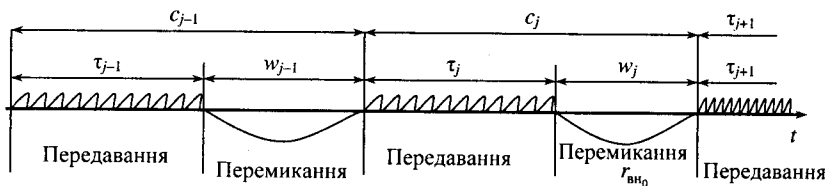


Рис. 16.2. Послідовність інтервалів використання каналу передавання даних СПДІ

Нехай r_{in} – імовірність того, що після прибуття обслуговуючого апарата у черзі знаходиться n повідомлень за умови, що у попередньому прибутті було i повідомлень.

Тоді ймовірність того, що апарат матиме в черзі n повідомлень r_n у момент прибуття до неї, буде такою:

$$\pi_n = \sum_{i=1}^{\infty} r_{in} \pi_i.$$

Зазначимо, що щільність розподілу часу обслуговування i повідомлень є n -кратна згортка функції $b(t)$ – функція цілісності обслуговування повідомлення (час передавання повідомлення каналом зв'язку). Позначимо її через $b_i(t)$. Щільність розподілу суми тривалості обслуговування i повідомлення i часової затримки перед обслуговуванням наступної групи повідомлень є згортокою функцій $b_i(t)$ і $w(t)$, яку позначимо: $Z_r [b_i(t), w(t)]$. З урахуванням уведених скорочень знаходимо умовну ймовірність r_{in} за виразом:

$$r_{in} = \int_0^{\infty} \frac{(\lambda \tau)^n}{n!} e^{-\lambda t} Z_r [b_i(t), w(t)] dt. \quad (16.17)$$

Уведемо твірну функцію для π_n і перетворення Лапласа $b(t)$ та $w(t)$:

$$b^*(s) = \int_0^{\infty} e^{-st} b(t) dt; \quad W(s) = \int_0^{\infty} e^{-st} w(t) dt; \quad (16.18)$$

$$P(z) = \sum_{n=0}^{\infty} \pi_n z^n. \quad (16.19)$$

Виконаємо підстановку формул (16.17) і (16.18) у рівняння (16.19) і отримаємо вираз для визначення функції $P(Z)$ у вигляді

$$P(Z) = \sum_{n=0}^{\infty} \sum_{i=0}^{\infty} r_{in} \pi_i z^n = \sum_{n=0}^{\infty} \sum_{i=0}^{\infty} \pi_i z^n \int_0^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \mathfrak{Z}_r [b_i(t), w(t)] dt. \quad (16.20)$$

Спростимо вираз (16.20), виконуючи дію підсумовування за n :

$$\sum_{n=0}^{\infty} \int_0^{\infty} \frac{(\lambda Z t)^n}{n!} e^{-\lambda t} \mathfrak{Z}_r [b_i(t), w(t)] dt = \int_0^{\infty} e^{-(1-Z)\lambda t} \mathfrak{Z}_r [b_i(t), w(t)] dt. \quad (16.21)$$

Оскільки в рівнянні (16.21) вираз під інтегралом – це перетворення Лапласа згортки функцій $b_i(t)$, $w(t)$ з комплексною змінною $\lambda(1-Z)$, то

$$P(Z) = \int_0^{\infty} e^{-(1-Z)\lambda t} w(t) dt \sum_{i=0}^{\infty} \pi_i \int_0^{\infty} e^{-(1-Z)\lambda t} \mathfrak{Z}_r [b_i(t)] dt. \quad (16.22)$$

Ураховуючи те, що перетворення Лапласа згортки функцій – це добуток перетворень Лапласа згорткових (складових) функцій, вираз (16.22) переписуємо у вигляді

$$P(Z) = w^* [\lambda(1-Z)] P \{ b^* [\lambda(1-Z)] \}. \quad (16.23)$$

Зазначимо, що обчислення або пошук характеристик оригіналів за рівняннями (16.23) та (16.22) можна спростити, якщо їх прологарифмувати і використати логарифмічне диференціювання, тобто

$$[\ln f(x)]' = \frac{f'(x)}{f(x)}.$$

Очевидніше буде, коли прологарифмуємо вираз (16.22) і визначимо першу похідну за z та знайдемо її значення в точці $Z = 1$:

$$\frac{d}{dz} \ln P(z) \Big|_{z=1} = \frac{d}{dz} \left[\ln \int_0^{\infty} e^{-\lambda(1-z)t} w(t) dt \right]' + \frac{d}{dz} \left[\ln \sum_{i=1}^{\infty} \pi_i \int_0^{\infty} e^{-\lambda(1-z)t} \mathfrak{Z}_r [b_i(t)] dt \right] \Big|_{z=1};$$

$$\left. \frac{\frac{d}{dz} P^*(z)}{P(z)} \right|_{z=1} = \sum_{n=0}^{\infty} n \pi_n; \quad \left. \frac{\frac{d}{dz} \int_0^{\infty} \lambda t e^{-\lambda(1-z)t} w(t) dt}{\int_0^{\infty} e^{-\lambda(1-z)t} w(t) dt} \right|_{z=1} = \lambda \int_0^{\infty} t w(t) dt = \lambda \bar{w};$$

$$\ln \sum_{i=0}^{\infty} \pi_i \int_0^{\infty} e^{-\lambda(1-z)t} 3_r [b_i(t)] dt = \ln \sum_{i=0}^{\infty} \pi_i \{b^*[\lambda(1-z)]\}^i;$$

$$\frac{d}{dz} \ln \sum_{i=0}^{\infty} \pi_i \{b^*[\lambda(1-z)]\}^i = \frac{d}{dz} \ln P\{b^*[\lambda(1-z)]\} =$$

$$= \left. \frac{\frac{d}{dz} P\{b^*[\lambda(1-z)]\}}{P\{b^*[\lambda(1-z)]\}} \right|_{z=1} = \lambda \bar{b} \sum_{i=0}^{\infty} n \pi_i.$$

У результаті зведення цих результатів, отримаємо

$$M\{\pi_n\} = \lambda \bar{w} + \lambda \bar{b} M\{\pi_n\},$$

звідки

$$M\{\pi_n\} = \frac{\lambda \bar{w}}{1 - \lambda \bar{b}},$$

де \bar{w} – середнє значення розподілу інтервалу «перемикання»; \bar{b} – середнє значення розподілу тривалості обслуговування (передавання одного повідомлення).

Для знаходження дисперсії $M\{\pi_n\}$ використовуємо теорему Філера, згідно з якою

$$D(M) = \left. \frac{d^2}{dz^2} P(z) \right|_{z=1} + \left. \frac{d}{dz} P(z) \right|_{z=1} - \left\{ \left. \frac{d}{dz} P(z) \right|_{z=1} \right\}^2.$$

Отже, отримано твірну функцію розподілу кількості заявок, які надходять на обслуговуючий апарат, що приступає до обслуговування черги, а також характеристики цього розподілу, тобто математичне сподівання та дисперсію, які і є оцінками використовуваної ємності пам'яті.

16.3. Моделювання півдуплексного режиму системи передавання дискретної інформації

16.3.1. Побудова математичної моделі

Півдуплексним режимом роботи СПДІ передбачається наявність одного каналу зв'язку між взаємодіючими абонентськими станціями (термінальне устаткування – кінцеве обладнання даних) або між сусідніми (нехай A і B) маршрутизаторами (комутаторами) магістральних ліній комп'ютерних мереж. Канал зв'язку періодично використовується станціями для передавання, наприклад, у напрямі $A \rightarrow B$, або $B \rightarrow A$, тобто канал періодично буде належати тій чи тій станції. Очевидно, коли станція A передає повідомлення на станцію B , на ній може виникнути потреба в передаванні повідомлень, адресованих станції A . Ці повідомлення тимчасово зберігаються (накопичуються) до моменту закінчення інтервалу зайнятості каналу зв'язку станцією A . Станція B отримала від станції A протокольну команду «Кінець передавання» – сигнал звільнення каналу формує запит на дозвіл для передавання накопичених повідомлень і після отримання підтвердження (цю процедуру називають перемикання каналу з A на B) передає накопичені повідомлення на станцію A .

Процес функціонування системи обміну повідомленнями між станціями A і B з використанням одного каналу зв'язку можна описати математичною моделлю масового обслуговування циклічної дії.

Розглянемо систему масового обслуговування без втрат (без відмовлень в обслуговуванні), яка складається з одного обслуговуючого апарата, на вхід якого надходять два незалежні пуассонівські потоки заявок (повідомлень – інформації, яка підлягає передаванню) з інтенсивностями λ_1 і λ_2 . Тривалість обслуговування заявок, незалежних одна від одної і від потоку, що надходить, випадкові величини з довільними функціями розподілу $B_i(t)$,

$i = 1, 2$ і середнім $\beta_{i1} = \int_0^1 t dB_i(t) < +\infty$, щільність розподілу $b_i(t) = dB_i(t)$.

Перейдемо до розгляду моделі. Режим роботи обслуговуючого апарата передбачає «переходи» від черги до черги, за час яких заявки (повідомлення), які знаходяться в пам'яті станції, не обслуговуються (не передаються каналом зв'язку).

Нехай у деякий момент часу апарат повернувся до першої черги з переходу (рис. 16.3, б, точка A_1) і останнє обслуговуване перед цим переходом повідомлення – це заявка другого потоку (черги). Тоді з цього моменту апарат буде обслуговувати заявки першого потоку. Причому обслужено буде лише ті заявки першого потоку, які були в системі в момент закінчення переходу γ_2 . Потім апарат робить перехід γ_2 до другої черги, піс-

ля якого вже будуть обслуговуватись тільки заявки другого потоку, причому знову буде обслужено тільки заявки цього потоку, які були в системі, коли апарат приступав до обслуговування. Потім настає перехід апарата і т. д.

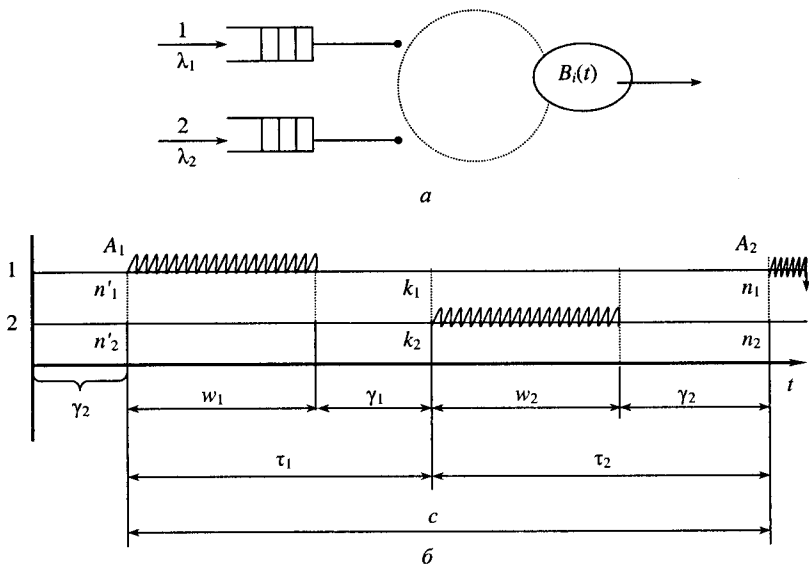


Рис. 16.3. Схематичне зображення моделі півдуплексного режиму (а) та діаграма зміни станів системи (A_1 і A_2 – точки регенерації) (б)

Тривалість переходів, які виникають після обслуговування заявок i -го потоку, – це незалежні від тривалості обслуговування та вхідного потоку випадкові величини з функцією розподілу $\Gamma_i(t)$, $i = 1, 2$ і середні $\gamma_{ij} =$

$$= \int_0^{\infty} t d\Gamma_i(t) < +\infty.$$

У моменти виходу апарата на обслуговування заявок деякого потоку (моменти закінчення операцій з установалення зв'язку) стани системи (кількість заявок у системі) характеризуються вектором $n = (n_1, n_2)$, де n_i – кількість заявок i -го потоку в системі в цей момент.

Розглянемо стаціонарні ймовірності $P_{n_1, n_2}^{(i)}$ наявності в системі n_1, n_2 заявок відповідно у 1-му і 2-му потоках (чергах) у моменти виходу апарата (коли канал зв'язку розпочинає передавання накопичених повідомлень) на обслуговування i -го потоку.

Функції $B_i(t)$ і $\Gamma_i(t)$ передбачаються абсолютно неперервними. Відповідні щільності ймовірностей позначають через $b_i(t)$ і $g_i(t)$. Нехай $\beta_i(s) = \int_0^{\infty} e^{-st} dB_i(t) = \int_0^{\infty} e^{-st} b_i(t) dt$ і $\gamma_i(s) = \int_0^{\infty} e^{-st} d\Gamma_i(t) = \int_0^{\infty} e^{-st} g_i(t) dt$ – перетворення Лапласа – Стільтєса розподілів $B_i(t)$ і $\Gamma_i(t)$ відповідно.

Із опису системи очевидно, що з точністю до циклічної перенумерації будь-який потік можна враховувати першим. Тому природно зосередити увагу на вивченні розподілу $P^{\{1\}}_{n_1, n_2}$. Для зручності запису опустимо індекс $\{1\}$ у цьому розподілі, тобто його позначимо через P_{n_1, n_2} .

Надалі будемо розглядати стани системи лише в моменти, коли апарат приступає до обслуговування заявок 1-го потоку (див. рис. 16.3, точки A_1, A_2, \dots). Проміжок часу між двома такими сусідніми моментами за сталого режиму будемо називати циклом і позначати через c . Згідно з описом системи протягом довільного циклу апарат чергує в суворій послідовності обслуговування 1-го і 2-го потоків з «переходами». Кількість заявок, обслужених в одному циклі, позначимо через $m = (m_1, m_2)$, де m_1 – кількість обслужених заявок 1-го потоку, m_2 – 2-го потоку.

16.3.2. Дослідження моделі методом укладених ланцюгів Маркова

Основний зміст укладених ланцюгів полягає в тому, що для випадкового процесу, який описує поведінку розглядуваної моделі, конструюється зручний марковський ланцюг, досліджуваний аналітичними методами, звичними для ланцюгів Маркова, і отримані результати (наприклад, стаціонарні ймовірності станів) переобчислюються для величин початкової системи.

Отже, метод складається із таких чотирьох кроків.

Крок 1. Визначення деякої відповідної послідовності моментів часу (t_n) , $0 \leq t_1 < t_2 < \dots < \infty$, що дозволяє конструювати вкладений ланцюг Маркова.

Крок 2. Обчислення перехідних ймовірностей для вкладеного ланцюга Маркова.

Крок 3. Обчислення стаціонарного розподілу або інших величин для вкладеного ланцюга Маркова.

Крок 4. Перерахунок отриманих за кроком 3 результатів у пошукові величини системи, причому в разі потреби і для точок, що відрізняються від t_n .

Розглядаючи стани системи в моменти початку циклів, отримаємо вкладений ланцюг Маркова M . Перехідні ймовірності цього ланцюга позначимо через $P(n_1, n_2 | n'_1, n'_2)$. Тоді стаціонарні ймовірності цього ланцюга задовольняють

$$P_{n_1, n_2} = \sum_{n'_1 \geq 0} \sum_{n'_2 \geq 0} P_{n'_1, n'_2} P(n_1, n_2 | n'_1, n'_2). \quad (16.24)$$

Знайдемо перехідні ймовірності $P(n | n')$, дотримуючись позначень на діаграмі, показаній на рис. 16.3. Визначимо довжину інтервалу циклу c (розподіл цього інтервалу або його щільність). Інтервал c складається з двох інтервалів часу – τ_1 і τ_2 . Нехай у момент A обслуговуючий апарат починає обслуговування n'_1 повідомлень, на кожне з яких апарат витрачає відрізок часу тривалістю $X_i (i = 1, \dots, n'_1)$ – щільність розподілу інтервалу $b_i(t)$; $X_1, \dots, X_{n'_1}$ – невід'ємні, незалежно розподілені випадкові величини зі щільністю розподілу $b_1(t), \dots, b_{n'_1}(t)$.

Щільність розподілу W_1 як щільність розподілу суми $X_1 + \dots + X_{n'_1}$ досить складна і має вигляд

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} b_1(u_1) b_2(u_2 - u_1) b_3(u_3 - u_2) \dots b_{n'_1}(x - u_{n'_1-1}) du_1 \dots du_{n'_1-1},$$

причому передбачається, що $b_i(x)$ за визначенням дорівнює нулю для від'ємних x . Цей інтеграл, названий згортою функцій $b_1(x), \dots, b_{n'_1}(x)$, не будемо в явному вигляді використовувати, оскільки оперувати завжди легше перетвореннями Лапласа, які визначаються рівнянням

$$b_i(s) = \int_0^{\infty} e^{-st} b_i(t) dt.$$

Для наочності згортку з i -функцій будемо позначати як $b^{*i}(t)$.

У такому разі, очевидно, перехідні ймовірності $P(n_1, n_2 | n'_1, n'_2)$ визначаються із виразу

$$P(n_1, n_2 | n'_1, n'_2) = \sum_{k_2 \geq n'_2} \int_0^{\infty} \int_0^{\infty} b^{*n'_1}(\tau_1) * g_1(\tau_1) \frac{(\lambda_2 \tau_1)^{k_2 - n'_2}}{(k_2 - n'_2)!} e^{-\lambda_2 \tau_1} \times \quad (16.25)$$

$$\times b^{*k_2}(\tau_2) * g_2(\tau_2) \left\{ \frac{[\lambda_1(\tau_1 + \tau_2)]^{n_1}}{n_1!} e^{-\lambda_1(\tau_1 + \tau_2)} \frac{(\lambda_2 \tau_2)^{n_2}}{n_2!} e^{-\lambda_2 \tau_2} d\tau_1 d\tau_2, \right.$$

де величина у фігурних дужках – щільність імовірності того, що якщо в першій черзі знаходиться n'_1 заявок, а у другій n'_2 заявок у момент, коли обслуговуючий апарат починає обслуговування першої черги, то потрібно мати час τ_1 для того, щоб закінчити обслуговування в першій черзі і

здійснити перехід до другої черги, в якій треба обслужити k_2 заявок і витратити час τ_2 для обслуговування другої черги та повернення до першої (час переходу від другої до першої черги – фаза устанавлення зв'язку, див. розд. 13). У фігурних дужках множник $\frac{(\lambda_2 \tau_1)^{k_2 - n'_2}}{(k_2 - n'_2)!} e^{-\lambda_2 \tau_1}$ визначає імо-

вірність надходження $(k_2 - n'_2)$ заявок за час обслуговування заявок n'_1 першої черги і перехід до другої, тому береться сума, яка розпочинається з n'_2 .

Для розв'язання рівняння (16.24) спочатку введемо твірну функцію

$$G(x_1, x_2) = \sum_{n_1 \geq 0} \sum_{n_2 \geq 0} p_{n_1 n_2} x_1^{n_1} x_2^{n_2}. \quad (16.26)$$

Підставимо вираз (16.25) у рівняння (16.24), а (16.24) – у (16.26) і отримаємо

$$\begin{aligned} G(x_1, x_2) &= \sum_{n_1 \geq 0} \sum_{n_2 \geq 0} p_{n_1 n_2} x_1^{n_1} x_2^{n_2} = \sum_{n_1 \geq 0} \sum_{n_2 \geq 0} p_{n_1 n'_2} p(n_1, n_2 | n'_1, n'_2) x_1^{n_1} x_2^{n_2} = \\ &= \sum_{n_1 \geq 0} \sum_{n_2 \geq 0} \sum_{k_2 \geq n'_2} \int_0^\infty \int_0^\infty p_{n_1 n'_2} x_1^{n_1} x_2^{n_2} \left\{ b^{*n'_1}(\tau_1) * g_1(\tau_1) \frac{(\lambda_2 \tau_{12})^{k_2 - n'_2}}{(k_2 - n'_2)!} \times \right. \\ &\quad \left. \times e^{-\lambda_2 \tau_1} b^{*k_2}(\tau_2) * g_2(\tau_2) \right\} \times \\ &\quad \times \frac{[\lambda_1(\tau_1 + \tau_2)]^{n_1}}{n_1!} e^{-\lambda_1(\tau_1 + \tau_2)} \cdot \frac{(\lambda_2 \tau_2)^{n_2}}{n_2!} e^{-\lambda_2 \tau_2} d\tau_1 d\tau_2. \end{aligned} \quad (16.27)$$

Спростимо вираз (16.27), змінивши порядок підсумовування.

1. Виконаємо підсумовування за n_1 :

$$\begin{aligned} \sum_{n_1 \geq 0} x_1^{n_1} \frac{[\lambda_1(\tau_1 + \tau_2)]^{n_1}}{n_1!} e^{-\lambda_1(\tau_1 + \tau_2)} &= \\ = e^{-\lambda_1(\tau_1 + \tau_2)} e^{\lambda_1 x_1(\tau_1 + \tau_2)} &= e^{-\lambda_1(1-x_1)\tau_1} e^{-\lambda_1(1-x_1)\tau_2}. \end{aligned}$$

2. Виконаємо підсумовування за n_2 :

$$\sum_{n_2 \geq 0} x_2^{n_2} \frac{(\lambda_2 \tau_2)^{n_2}}{n_2!} e^{-\lambda_2 \tau_2} = e^{-\lambda_2(1-x_2)\tau_2}.$$

3. Знайдемо суму за $k_2 \geq n'_2$ для виразу $\sum_{k_2 \geq n'_2} \left\{ \frac{(\lambda_2 \tau_{12})^{k_2 - n'_2}}{(k_2 - n'_2)!} e^{-\lambda_2 \tau_1} \right\}$:

$$\sum_{k_2 \geq n_2} \left\{ \frac{(\lambda_2 \tau_2)^{k_2 - n_2'}}{(k_2 - n_2')!} e^{-\lambda_2 \tau_2} \right\} = \left[1 + \lambda_2 \tau_2 + \frac{(\lambda_2 \tau_2)^2}{2!} + \frac{(\lambda_2 \tau_2)^3}{3!} + \dots \right] \cdot e^{-\lambda_2 \tau_2} = e^{\lambda_2 \tau_2} \cdot e^{-\lambda_2 \tau_2} = 1.$$

З урахуванням цих перетворень вираз (16.27) набуває такого вигляду:

$$\begin{aligned} G(x_1, x_2) &= \sum_{k_1 \geq n_1'} \sum_{k_2 \geq n_2'} \int_0^{\infty} \int_0^{\infty} p_{n_1' n_2'} e^{-\lambda_1(1-x_1)\tau_1} e^{-[\lambda_1(1-x_1) + \lambda_2(1-x_2)]\tau_2} \times \\ &\quad \times b_1^{*n_1'}(\tau_1) * g_1(u_1) * b_2^{*k_2}(\tau_2) * g_2(\tau_2) d\tau_1 d\tau_2 = \\ &= \sum_{k_1 \geq n_1'} \sum_{k_2 \geq n_2'} p_{n_1' n_2'} \int_0^{\infty} \int_0^{\infty} e^{-\lambda_1(1-x_1)\tau_1} \cdot e^{-[\lambda_1(1-x_1) + \lambda_2(1-x_2)]\tau_2} \times \\ &\quad \times b_1^{*n_1'}(\tau_1) * g_1(\tau_1) * b_2^{*k_2}(\tau_2) * g_2(\tau_2) d\tau_1 d\tau_2. \end{aligned}$$

Позначимо: $\vartheta_1 = (1-x_1)\lambda_1$; $\vartheta_2 = \lambda_1(1-x_1) + \lambda_2(1-x_2)$. Тоді

$$\begin{aligned} G(x_1, x_2) &= \sum_{k_1 \geq n_1'} \sum_{k_2 \geq n_2'} p_{n_1' n_2'} \int_0^{\infty} e^{-\vartheta_1 \tau_1} \times \\ &\quad \times b_1^{*n_1'}(\tau_1) * g_1(\tau_1) d\tau_1 \int_0^{\infty} e^{-\vartheta_2 \tau_2} b_2^{*k_2}(\tau_2) * g_2(\tau_2) d\tau_2 = \\ &= \int_0^{\infty} e^{-\vartheta_1 \tau_1} g_1(\tau_1) d\tau_1 \int_0^{\infty} e^{-\vartheta_2 \tau_2} g_2(\tau_2) d\tau_2 \sum_{k_1 \geq n_1'} \sum_{k_2 \geq n_2'} p_{n_1' n_2'} \times \\ &\quad \times \left[\int_0^{\infty} e^{-\vartheta_1 \tau_1} b_1(\tau_1) d\tau_1 \right]^{n_1'} \left[\int_0^{\infty} e^{-\vartheta_2 \tau_2} b_2(\tau_2) d\tau_2 \right]^{k_2}. \end{aligned} \tag{16.28}$$

Перші два інтеграли – це перетворення Лапласа функцій переходів від першої черги до другої черги і від другої до першої з параметрами $s = \vartheta_1 = (1-x_1)\lambda_1$ і $s = \vartheta_2 = (1-x_1)\lambda_1 + \lambda_2(1-x_2)$, а в квадратних дужках – перетворення Лапласа з тими самими параметрами функцій щільності ймовірностей обслуговування повідомлень першої та другої черг відповідно.

Тоді вираз (16.28) у скороченому вигляді буде мати запис:

$$G(x_1, x_2) = \gamma_1(\vartheta_1) \gamma_2(\vartheta_2) G(\beta_1(\vartheta_1), \beta_2(\vartheta_2)).$$

Моменти розподілу $p_{n_1 n_2}$ зручно знаходити так. Уводимо функцію

$$\begin{aligned} F(x_1, x_2) &= \ln G(x_1, x_2); \\ F(x_1, x_2) &= \ln [\gamma_1(\vartheta_1) \gamma_2(\vartheta_2) G(\beta_1(\vartheta_1), \beta_2(\vartheta_2))] = \\ &= \ln \gamma_1(\vartheta_1) + \ln \gamma_2(\vartheta_2) + F(\beta_1(\vartheta_1), \beta_2(\vartheta_2)). \end{aligned}$$

Очевидно, що моменти розподілу легко знаходити, знаючи відповідні похідні функції $F(x_1, x_2)$ у точках $x_1 = x_2 = 1$:

$$\left. \frac{\partial F(x_1, x_2)}{\partial x_i} \right|_{x_1=x_2=1} = F'_{x_i}(1) = n_{i1}^{\{1\}}, \quad i = 1, 2;$$

$$\left. \frac{\partial^2 F(x_1, x_2)}{\partial x_i^2} \right|_{x_1=x_2=1} = F''_{x_i x_i}(1) = n_{i2}^{\{1\}} - n_{i1}^{\{1\}} - (n_{i1}^{\{1\}})^2,$$

де через $n_{i1}^{\{1\}}$ і $n_{i2}^{\{1\}}$ позначено відповідно перший і другий моменти розподілу числа повідомлень i -го потоку в системі, коли апарат приступає до обслуговування першого потоку.

Отже, використовуючи твірну функцію $G(x_1, x_2)$ і характеристики сумісного розподілу, можна отримати оцінки використовуваної ємності пам'яті.

16.4. Затримки повідомлень каналами системи передавання дискретної інформації

16.4.1. Загальні положення. Простір станів системи

У попередніх розділах ми розглядали роботу дискретного каналу СПДІ у дуплексному та півдуплексному режимах роботи під керуванням відповідних мережевих протоколів і відображення механізмів їх дії у вигляді випадкових процесів масового обслуговування, які описують поведінку у часі тієї чи тієї схеми масового обслуговування. Під час розгляду таких моделей було знайдено аналітичні процедури оцінювання ресурсного показника – ємності використовуваної пам'яті для тимчасового зберігання до моменту звільнення каналу від попереднього повідомлення.

Таким чином, виникають тимчасові затримки передаваних повідомлень. Нагадаємо, що один з основних показників роботи СПДІ – своєчасність доставляння повідомлень споживачу, або їх швидкість передавання.

Очевидно, що тимчасове зберігання повідомлень у буферній пам'яті привносить часові затримки. Такі затримки, які створюють канали, особливо можуть мати суттєві значення в мережах з комутацією пакетів для організації трактів обміну інформацією з багатоступеневою маршрутизацією.

У цьому розділі розглянемо методи визначення величин затримок у моделях, які описують дуплексний і півдуплексний режими роботи СПДІ.

У процесі функціонування СПДІ може знаходитись у двох станах: у зайнятому – виконує роботу з передавання повідомлень, та у вільному – система простояє (простій). Якщо розглядати СПДІ в робочому стані, то цей стан можна визначати за кількістю повідомлень, які треба передати каналом. Стан системи визначають за кількістю повідомлень – заявок. Процес переходу та повернення в деякий стан j показано на рис. 16.4.

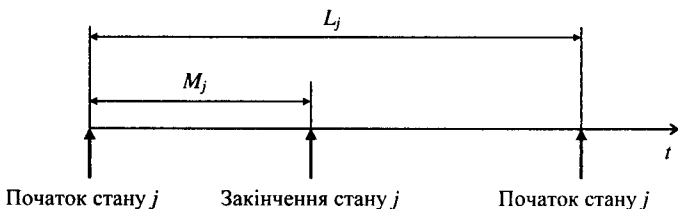


Рис. 16.4. Діаграма переходу та повернення в стан j системи

Починаючи з моменту переходу в стан j , система залишається в цьому стані час M_j . Потім система виходить зі стану j , проходить ряд інших станів і повертається в стан j через час L_j після її попереднього переходу в стан j . Для інтервалу, зображеного на рис. 16.4, частка часу перебування в стані j дорівнює M_j/L_j . Якщо є декілька таких послідовних інтервалів і для k -го інтервалу час перебування в стані j дорівнює M_{jk} , а час між послідовними переходами в цей стан дорівнює L_{jk} , тоді частка часу перебування в стані j дорівнює $\sum_k M_{jk} / \sum_k L_{jk}$. У теорії відновлення отримано асимптотичні значення цієї ймовірності. Якщо π_j – стаціонарна ймовірність стану j , $m_j = E(M_j)$, $l_j = E(L_j)$ – математичні сподівання випадкових величин M_j і L_j відповідно, тоді

$$\pi_j = \lim_{t \rightarrow \infty} p_j(t) = \frac{m_j}{l_j}. \quad (16.29)$$

Стани визначають так, щоб можна було оцінити тривалість перебування заявки в системі залежно від стану, у якому система знаходилась у момент надходження заявки. Оскільки потік заявок пуассонівський, тоді ймовірність нової надісланої заявки застати систему в стані j дорівнює π_j – стаціонарній ймовірності стану j .

16.4.2. Період зайнятості системи

Поняття «період зайнятості системи передавання даних» під час обчислення потрібної ємності буферної пам'яті було розглянуто у підрозд. 16.2. Повернемось до визначення цієї характеристики моделей, розглянутих у підрозд. 16.2.4 та 16.3.1, з метою використання її для обчислення затримок повідомлень СПДІ.

У досліджуваних моделях, на вхід яких надходить пуассонівський потік, періоди зайнятості і незайнятості утворюють дві чергові та незалежні послідовності однаково розподілених випадкових величин. Розподіл тривалості періодів зайнятості може залежати від дисципліни вибору із черги тільки в тому разі, коли дисципліна передбачає додаткове обслуговування (наприклад, повторення передавання повідомлення у разі виявлення помилок, які треба виправити).

Зручно для графічного зображення стан системи оцінювати за допомогою обсягу роботи, яка вимірюється в одиницях часу і складається із суми тривалостей обслуговування заявок, що знаходяться в черзі, і тривалості обслуговування заявки, що знаходиться в апараті. Типовий приклад реалізації такого процесу показано на рис. 16.5.

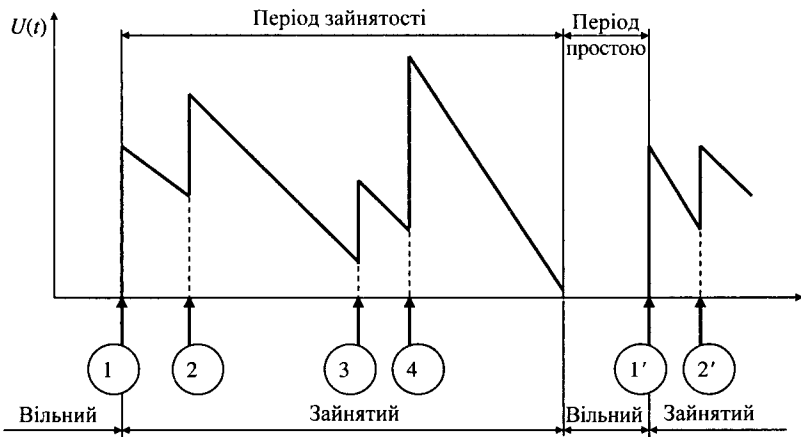


Рис. 16.5. Приклад реалізації циклу роботи і простоювання апарата

«Стрибки вгору» виникають у моменти надходження нової заявки, а їх тривалість виникає з тривалістю обслуговування відповідної заявки (у кружечках – номери заявок, що надходять). Коли величина роботи додатна, вона збігається зі швидкістю -1 до того моменту, поки не виникне додаткова робота, яка може бути пов'язана з перервами в обслуговуванні або періодом вимушеного простою обслуговуючого апарата. Коли вели-

чина роботи набуває значення 0, вона зберігає це значення до моменту надходження нової заявки (на рис. 16.5 – це $1'$ в кружечку). Із графіка видно, що тривалість періоду зайнятості не залежить від порядку вибору із черги повідомлень, оскільки він повністю визначається моментами надходження заявок і моментами їх обслуговування.

Тепер перейдемо до детального розгляду періоду зайнятості. Прослідкуємо за процесом утворення періоду зайнятості, який показано на рис. 16.6, у вигляді поведінки функції $U(t)$ – наявності роботи обслуговуючого апарата. Якщо значення цієї функції дорівнює 0, тоді система переходить у стан простою, або за введеного термінологією – обслуговуючий апарат «переключається» на інші роботи, не пов'язані з обслуговуваною ним чергою до каналу передавання.

Уведемо поняття *породжувальна заявка* і породжувальний нею *елементарний період зайнятості* (ЕПЗ). Породжувальна заявка – це заявка, яка породжує період зайнятості обслуговуванням тих і тільки тих заявок, які надходять у систему за час її обслуговування.

Як відомо, тривалість періоду зайнятості не залежить від порядку обслуговування заявок, що надходять (рис. 16.6, *a*).

Припустимо, що система вільна до моменту τ_1 , коли заявка c_1 розпочинає період зайнятості тривалістю Y . Час обслуговування цієї заявки дорівнює x_1 . Очевидно, що ця заявка залишить систему в момент $\tau_1 + x_1$. За час її обслуговування в систему можуть надійти інші заявки, які продовжать період зайнятості. Як видно з рис. 16.5, у розглядуваному прикладі за час обслуговування заявки c_1 надійде три інші заявки (c_2 , c_3 і c_4). Змінимо порядок обслуговування і введемо дисципліну обслуговування у зворотному порядку, тобто остання c_4 надійшла, а обслуговується першою після залишення апарата заявка c_1 .

За час обслуговування c_4 можуть надійти нові заявки (c_5 і c_6), які у порядку надходження обслуговуються одразу ж за c_4 , коли остання залишить апарат. Таким чином заявка c_4 мовби розпочинає новий період зайнятості. Такі періоди називали елементарними періодами зайнятості, що породжуються заявкою c_4 .

Елементарний період зайнятості, утворений заявкою c_4 , буде мати тривалість X_4 , точно таку саму, як час обслуговування цієї заявки і всіх нових заявок, які надходять у систему (при цьому заявки c_2 і c_3 не враховуються).

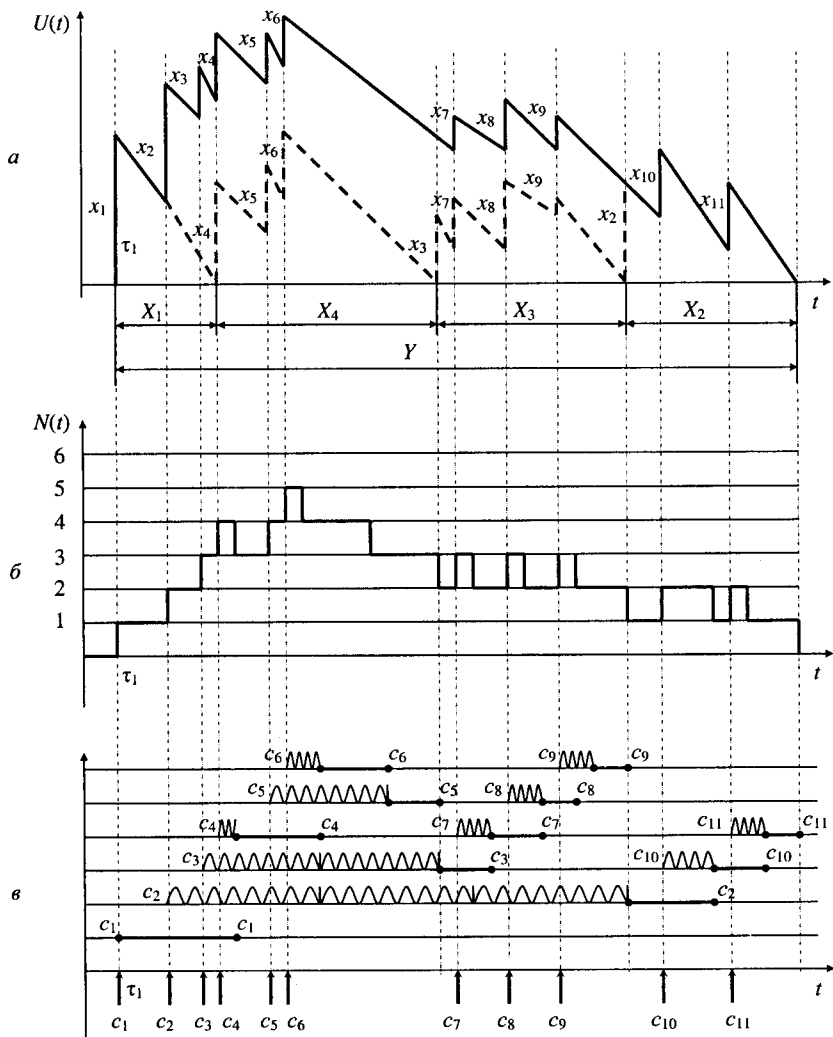


Рис. 16.6. Період зайнятості за зворотного порядку обслуговування:
 а – розгортання періоду зайнятості; б – графік зміни кількості заявок;
 в – сліди перебування заявок у системі

Отже, на рис. 16.6 показано ЕПЗ, породжений заявкою c_4 , протягом якого обслуговуються за визначенням порядком заявки c_4 , c_6 і c_5 . У момент $t_1 + x_1 + X_4$ цей ЕПЗ закінчується і продовжується той самий зворотний порядок обслуговування, який використовується тепер для поверненої для обслуговування заявки c_3 . Очевидно його можна розглядати як породжувальний свій особистий ЕПЗ тривалістю X_3 , за який у зворотному порядку (тобто c_3 , c_7 , c_8 і c_9) обслуговуються ті заявки, що надійшли за час його обслуговування. Нарешті, коли система знову стає порожньою, повертається заявка c_2 і розпочинається її ЕПЗ (довжиною X_2), яким завершується основний період зайнятості і протягом якого обслуговуються заявки c_2 , c_{10} , c_{11} .

Із рис. 16.6, *a* видно, що контур зміни будь-якого ЕПЗ точно повторює контур головного періоду зайнятості над тим самим проміжком часу. Передісторії кожної заявки показано на рис. 16.6, *b*, а загальну кількість заявок у системі в будь-який момент часу за розглядуваної дисципліни обслуговування – на рис. 16.6, *b*.

Виконаємо аналіз процесу масового обслуговування. Зосередивши увагу на ЕПЗ, помітимо, що кожний із них статистично подібний до головного періоду зайнятості, породженого заявкою c_1 . Дійсно, породжувальна заявка c_1 породжує стільки ЕПЗ, скільки надійде заявок за час її обслуговування x_1 (c_2 , c_3 і c_4); аналогічно кожна з них породжує ЕПЗ. Заявки, які надходять на породжувальні інтервали ($x_1 \rightarrow c_1$, $x_2 \rightarrow c_2$, $x_3 \rightarrow c_3$, $x_4 \rightarrow c_4$) підпорядковані одному й тому самому закону – пуассонівському стаціонарному потоку; причому всі заявки незалежні і мають один і той самий розподіл часу обслуговування. Отже, випадкові величини $\{X_k\}$ незалежні, однаково розподілені і мають ту саму функцію розподілу, що й головний період зайнятості Y .

Тривалість Y періоду зайнятості являє собою суму випадкових величин: $Y = x_1 + X_{\bar{v}+1} + X_{\bar{v}} + \dots + X_3 + X_2$.

Позначимо функцію розподілу періоду зайнятості:

$$G(y) = P(Y \leq y).$$

Із вихідних даних відомо, що x_1 має функцію розподілу $B(x)$, а X_k – функцію розподілу $G(y)$. Уведемо перетворення Лапласа для щільності розподілу, зв'язаного з Y :

$$G^*(s) = \int_0^{\infty} e^{-sy} dG(y) = E[e^{-sy}].$$

Використаємо техніку умовних розподілів, яка дозволяє записувати ймовірність, пов'язану зі складними подіями, через умовні ймовірності цієї події за відповідних умов, для яких можна обчислити умовну ймовірність. Якщо ці умови несумісні і достатні, тоді безумовну ймовірність знаходять за формулою повної імовірності як суму добутків умовних імовірностей на імовірності умов. У цьому разі розглянемо Y залежно від двох умов: тривалості обслуговування заявки s_1 і кількості нових заявок, що надійшли за час її обслуговування. Потім обчислимо такі умовні перетворення:

$$\begin{aligned} E\left[e^{-sy} \mid x_1 = x, \bar{v} = k\right] &= E\left[e^{-s(x+X_{k+1}+\dots+X_2)}\right] = \\ &= E\left[e^{-sx} e^{-sX_{k+1}} \dots e^{-sX_2}\right]. \end{aligned} \quad (16.30)$$

Оскільки тривалості ЕПЗ взаємно незалежні, рівність (16.30) можна переписати в такому вигляді:

$$E\left[e^{-sy} \mid x_1 = x, \bar{v} = k\right] = E\left[e^{-sx}\right] E\left[e^{-sX_{k+1}}\right] \dots E\left[e^{-sX_2}\right].$$

Оскільки x – задана постійна величина, тоді $E[e^{-sx}] = e^{-sx}$, і, крім того, якщо ЕПЗ розподілені однаково з відповідними перетвореннями $G^*(s)$, то

$$E\left[e^{-sy} \mid x_1 = x, \bar{v} = k\right] = e^{-sx} [G^*(s)]^k.$$

Оскільки \bar{v} – це кількість заявок, що надійшли за час x , яка за умов задачі розподілена за законом Пуассона із середнім значенням числа λx , то можна знехтувати умовою, яка накладається на \bar{v} :

$$\begin{aligned} E\left[e^{-sy} \mid x_1 = x\right] &= \sum_{k=0}^{\infty} \left[e^{-sy} \mid x_1 = x, \bar{v} = k \right] P[\bar{v} = k \mid x_1 = x] = \\ &= \sum_{k=0}^{\infty} e^{-sx} [G^*(s)]^k \frac{(\lambda x)^k}{k!} e^{-\lambda x} = e^{-x[s+\lambda-\lambda G^*(s)]}. \end{aligned}$$

Так само можна «зняти» умову, яка накладається на x_1 , інтегруючи інтегрувальною функцією $B(x)$, і зрештою отримати $G^*(s)$ у вигляді

$$G^*(s) = \int_0^{\infty} e^{-x[s+\lambda-\lambda G^*(s)]} dB(x). \quad (16.31)$$

У виразі (16.31) легко бачити перетворення щільності розподілу часу обслуговування, обчисленого для значення, що в дужках у показнику степеня, тобто

$$G^*(s) = B^*[s + \lambda - \lambda G^*(s)]. \quad (16.32)$$

На жаль, $G^*(s)$ знаходиться в обох частинах цього функціонального рівняння, що робить зазвичай неможливим отримання явного розв'язку. Однак рівняння (16.32) дозволяє визначити моменти періоду зайнятості, які найчастіше використовують для розрахунків систем і мереж передавання інформації.

16.4.3. Розподіл часу затримки повідомлень

Розподіл тривалості періоду зайнятості має важливе значення для вивчення залежності розподілу тривалості перебування заявки в системі, що й визначає затримку повідомлення трактом передавання інформації СПДІ.

Розглядувані моделі дуплексного та півдуплексного режимів роботи будуть мати різні функціональні рівняння для знаходження характеристик періоду зайнятості.

Замінімо деякі позначення, використовувани для аналізу часу заявки перебування в системі. Згадаймо, що $\lambda\tau_{\text{обсл}} = \lambda\bar{b} = \rho$ – навантаження на систему, $\rho < 1$, $Y = P \circ \beta(s) = \gamma(s)$, де P – тривалість обслуговування одиночної заявки, що надійшла у вільний стан (простій) системи, а $\beta(s)$ – перетворення Лапласа часу обслуговування цієї заявки і збігається з періодом зайнятості, оскільки він утворений тривалістю обслуговування тієї самої заявки. Тоді з імовірністю ρ заявка застає апарат зайнятим, і отже, очікує.

Позначимо: $E\{e^{-sy} \mid \text{зайнято}\}$. Для знаходження цієї величини виразимо період зайнятості у вигляді послідовності залежних випадкових інтервалів і обчислимо перетворення Лапласа – Стілтєса розподілу тривалості перебування заявки, що надійшла в один із цих інтервалів.

У нових позначеннях нехай T_0 – тривалість обслуговування початкової заявки. Інтервалу T_1 , протягом якого обслуговуються заявки, що надійшли за час обслуговування початкової заявки, присвоїмо номер 1. Позначимо через T_j суму тривалостей обслуговування всіх заявок, які надійшли в інтервалі номер $j - 1$. Послідовність таких інтервалів зображено на рис. 16.7, на якому через N_j позначено кількість надходжень в j -му інтервалі.

Позначимо функцію розподілу T_j і її перетворення Лапласа – Стілтєса відповідно через $H_j(t)$ і $\eta_j(s)$. Виразимо $\eta_j(s)$ у термінах перетворення Лапласа – Стілтєса для T_{j-1} , тобто через $\eta_{j-1}(s)$, використовуючи ідею отримання (виведення) перетворення Лапласа – Стілтєса для періоду зайнятості (див. п. 16.4.2). Використовуючи властивість згортки перетворень Лапласа – Стілтєса, отримуємо

$$E\{e^{-sT_j} \mid T_{j-1} = t, N_{j-1} = n\} = [\gamma(s)]^n,$$

$$\eta_i(s) = E(e^{-sT_j}) = \int_0^\infty e^{-\lambda t} \sum_{n=0}^\infty \frac{[\lambda t \gamma(s)]^n}{n!} dH_{j-1}(t) = \eta_{j-1}(\lambda - \lambda \gamma(s)). \quad (16.33)$$

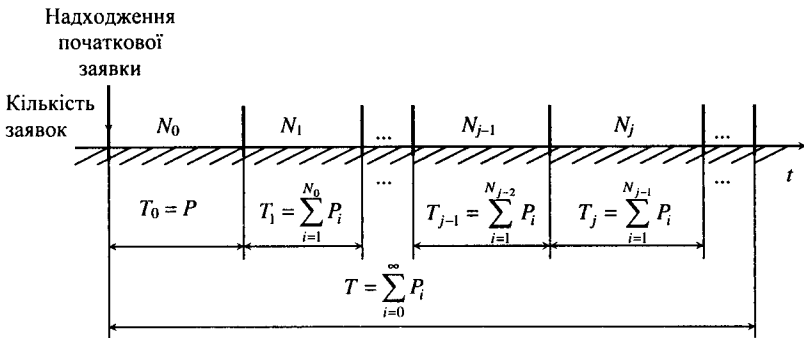


Рис. 16.7. Діаграма послідовності інтервалів у період зайнятості

Отже, період зайнятості виражено як нескінченна послідовність інтервалів довжиною T_j кожний, але за умови $\rho < 1$ з імовірністю 1 є лише кінцеве число таких j , для яких $T_j > 0$, тобто $\lim_{j \rightarrow \infty} \eta_j(s) = 1$.

Нехай система знаходиться в деякий момент часу в стані j , якщо в цей момент реалізується j -й інтервал періоду зайнятості. Завдяки властивостям пуассонівського потоку ймовірність того, що нова заявка, яка надійшла, застане систему в стані j , дорівнює стаціонарній імовірності π_j цього стану.

Використовуючи рівняння (16.29) та рис. 16.7, отримуємо

$$\pi_j = P\{j \text{ зайнято}\} = \frac{E(T_j)}{E(T)}. \quad (16.34)$$

Позначимо через $E\{e^{-sF} | j\}$ перетворення Лапласа – Стілтєса тривалість перебування заявки в системі за умови, що вона надійшла в систему в стані (інтервалі) j . Тривалість перебування цієї заявки залежить від номера інтервалу, оскільки залежить від його довжини. Як показано на рис. 16.8, після надходження заявка спочатку очікує закінчення j -го інтервалу, а потім – закінчення обслуговування всіх заявок, які надійшли в цьому інтервалі раніше від неї. Визначимо $E\{e^{-sF} | T_j = t, Y = y, N = n\}$ перетворення Лапласа – Стілтєса часу перебування заявки в системі за умови, що тривалість часу T_j буде дорівнювати змінній t , тривалість часу Y очікування закінчення j -го інтервалу – змінній y , кількість N заявок, що надійшли в інтервалі $T_j - Y$, – змінній n .

Використовуючи властивість згортки перетворень Лапласа – Стілтєса, знаходимо

$$E\{e^{-sF} | T_j = t, Y = y, N = n\} = e^{-sy} [\gamma(s)]^{n+1}.$$

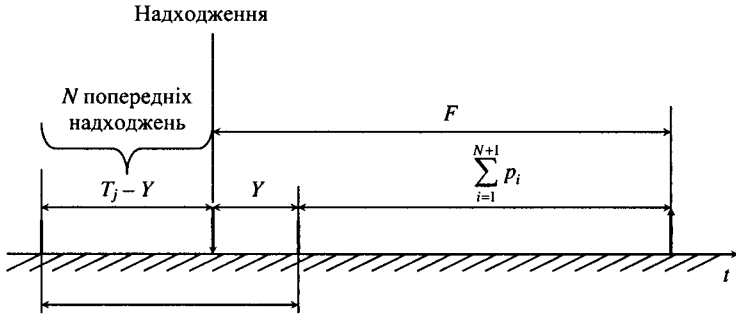


Рис. 16.8. Діаграма поведінки заявки, що надійшла в інтервал T_j (надходження)

Припустивши, що за час $t - y$ надійде n заявок, змінимо умову N так:

$$\begin{aligned} E\{e^{-sF} | T_j = t, Y = y\} &= e^{-sy} [\gamma(s)]^{n+1} \sum_{n=0}^{\infty} \frac{(\lambda(t-y))^n}{n!} e^{-\lambda(t-y)} = \\ &= e^{-sy} \gamma(s) e^{-\lambda(t-y)} \sum_{n=0}^{\infty} \frac{[\lambda(t-y)\gamma(s)]^n}{n!} = \\ &= e^{-sy} \gamma(s) e^{-\lambda(t-y)[1-\gamma(s)]}. \end{aligned} \quad (16.35)$$

Для того щоб «зняти» умови t і y й отримати безумовну величину $E\{e^{-sF}\}$ (перетворення Лапласа – Стілтєса) – часу перебування поміченої заявки (час її надходження показано на рис. 16.9), розглянемо на осі часу моменти виникнення подій відповідно до закону потоку Пальма, тобто потік, у якому події розділено інтервалами часу T_1, T_2, \dots , які незалежні між собою і однаково розподілені випадкові величини із заданою функцією розподілу

$$\Phi(x) = P(X \leq x).$$

Визначимо щільність розподілу, а також однакова для всіх проміжків:

$$\varphi(x) = \frac{d}{dx} \Phi(x).$$

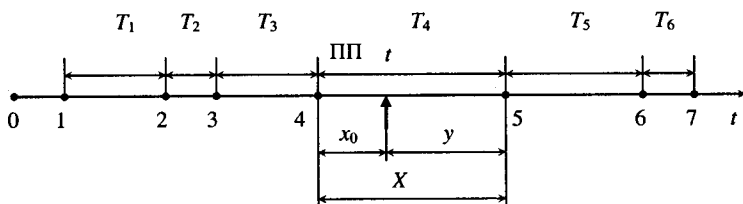


Рис. 16.9. Реалізація потоку Пальма

Виберемо тепер випадково момент t , у який виникає помічена подія (ПП), не зв'язана з перенумерованими подіями протоколу.

Потрібно визначити закон розподілу того проміжку часу між сусідніми подіями (тут він T_4^* , а надалі опустимо індекс 4), на який випала точка ПП, що зображає момент t (момент виникнення поміченої події). Позначимо через X увесь проміжок часу, який займає інтервал T_4^* між подіями потоків 4 і 5.

Незвичайний результат, який буде отримано, полягає в тому, що величина X не розподілена за законом $\Phi(x)$. Для розглядуваного прикладу це означає, що проміжок часу між виникненням ПП і 5-ї точки потоку не буде типовим проміжком. Фактично в цьому і полягає вирішення розглядуваного парадоксу: довгий проміжок імовірніше «захопить» момент появи випадкової ПП, ніж короткий. У випадку пуассонівського процесу надходження ця обставина, як бачимо, приводить до того, що виділений проміжок виявляється удвічі довший від типового.

Нехай залишковий час Y описується функцією розподілу

$$\hat{\Phi}(x) = P[Y \leq x]$$

зі щільністю

$$\hat{\phi}(x) = \frac{d}{dx} \hat{\Phi}(x).$$

Аналогічно нехай виділений проміжок часу X описується функцією розподілу

$$\Phi_x(x) = P[X \leq x]$$

і відповідною щільністю $\phi_x(x)$.

Тут віддамо перевагу інтуїтивному висновку, який ґрунтується на фізичному описі розглядуваної задачі. Основне зауваження полягає в тому, що довгі проміжки часу між моментами виникнення подій (заявок) у потоці захоплюють на осі часу більші ділянки, ніж короткі проміжки, і тому

Ймовірніше, що випадково вибрана точка t (ППП) попаде на довгий проміжок. Якщо погодитись з цим, то ймовірність вибору проміжку часу довжини x має бути пропорційною цій довжині, а також відносній частоті виникнення таких проміжків і дорівнювати $\varphi(x)dx$.

Отже, для виділеного інтервалу можна записати

$$\varphi_x(x)dx = Kx\varphi(x)dx, \quad (16.36)$$

де ліва частина являє собою $P[x < X < x + dx]$, а права частина виражає лінійну значущу функцію відносно довжини проміжку і містить множник K , що нормує щільність. Інтегруючи обидві частини рівності (16.36), знаходимо, що $K = \frac{1}{m_1}$, де m_1 – середній час між подіями потоку заявок, що надійшли.

Отже, показано, що щільність розподілу виділеного проміжку часу зв'язана зі щільністю розподілу типових проміжків часу рівністю

$$\varphi_x(x) = \frac{x\varphi(x)}{m_1}.$$

Перейдемо до визначення щільності розподілу залишкового часу $\hat{\varphi}(x)$. Імовірність того, що залишковий час Y не перебільшує y за умови $X = x$, визначаємо за рівністю

$$P[Y \leq y | X = x] = \frac{y}{x}, \text{ якщо } 0 \leq y \leq x.$$

Це правильно, оскільки точка, випадково вибрана у проміжку, має бути рівномірно розподілена на цьому проміжку. Отже, можна записати сумісну щільність розподілу величин X і Y :

$$\begin{aligned} P[y < Y < y + dy, x < X \leq x + dx] &= \\ = \left(\frac{dy}{x}\right) \left(\frac{x\varphi(x)dx}{m_1}\right) &= \frac{\varphi(x)dydx}{m_1}, \text{ якщо } 0 \leq y \leq x. \end{aligned} \quad (16.37)$$

Інтегруючи за x , отримаємо функцію $\hat{\varphi}(y)$, яка являє собою безумовну щільність розподілу Y :

$$\hat{\varphi}(y)dy = \int_{x=y}^{\infty} \frac{\varphi(x)dydx}{m_1},$$

що безпосередньо приводить до кінцевої рівності

$$\hat{\varphi}(y) = \frac{1 - \Phi(y)}{m_1}. \quad (16.38)$$

Рівняння (16.38) дає сумісний розподіл залишкового часу Y і повного проміжку X , а вираз (16.37) – сумісну щільність розподілу X і Y . Тут в загальному випадку знайдено розподіл залишкового часу без дотримання уведених позначень процесу обслуговування.

Для використання цього результату зробимо деякі заміни та повернемося до формули (16.35). Формулу (16.37) з урахуванням раніше прийнятих позначень і рис. 16.8 перепишемо в такому вигляді:

$$P[y \leq Y \leq y + dy, y \leq T_j \leq t + dt] = \frac{dH_j(t)}{E(T_j)}, \quad 0 \leq y \leq t \leq \infty.$$

За допомогою цього сумісного розподілу можна «зняти» умови T_j і Y у рівності (16.35), тоді

$$\begin{aligned} E\{e^{-sF} | j\} &= \gamma(s) \int_{t=0}^{\infty} \int_{y=0}^{\infty} e^{-\lambda[1-\gamma(s)]} e^{-y[\lambda\gamma(s)-\lambda+s]} \frac{dH_j(t)dy}{E(T_j)} = \\ &= \frac{\gamma(s)}{E(T_j)[\lambda\gamma(s) - \lambda + s]} \int_{t=0}^{\infty} \{e^{-\lambda[1-\gamma(s)]} - e^{-st}\} dH_j(t). \end{aligned}$$

Другий доданок під знаком інтегралу дорівнює $\eta_j(s)$, перший доданок – $\eta_{j+1}(s)$, що впливає в рівняння (16.33) після заміни j на $j + 1$.

Тому

$$E\{e^{-sF} | j\} = \frac{\gamma(s)[\eta_{j+1}(s) - \eta_j(s)]}{E(T_j)[\lambda\gamma(s) - \lambda + s]}. \quad (16.39)$$

Із рівнянь (16.34) і (16.35) випливає, що

$$\begin{aligned} E\{e^{-sF} | \text{зайнято}\} &= \sum_{j=0}^{\infty} \pi_j E\{e^{-sF} | j\} = \\ &= \frac{\gamma(s)}{[\lambda\gamma(s) - \lambda + s]E(T)} \sum_{j=0}^{\infty} [\eta_{j+1}(s) - \eta_j(s)]. \end{aligned}$$

Сума дорівнює $1 - \eta_0(s)$, а оскільки нульовий інтервал складається із часу обслуговування лише однієї заявки, то $\eta_0(s) = \gamma_0(s)$.

Із формального погляду $\sum_{j=0}^{\infty} [\eta_{j+1}(s) - \eta_j(s)] = -\eta_0(s)$, але це не так,

тому що $\rho = \lambda\bar{\beta} < 1$ – стаціонарний режим, причому j має кінцеву величину j_0 з імовірністю 1. Отже, частина нескінченної суми, починаючи з проміжків $j \geq j_0$, буде дорівнювати одиниці, а взагалі ця сума дорівнюватиме $1 - \eta_0(s)$. Знайдемо $E(T)$ – математичне сподівання періоду зайнятості з

виразу (16.32), використовуючи процедури диференціювання $G^*(s)$ і пошук значення першої похідної в точці $s = 0$ як $G'(0)$, тобто

$$E(T) = \frac{\bar{\beta}}{1-\rho} = \frac{E(B)}{1-\rho}.$$

Знаходимо також

$$E\{e^{-sF} \mid \text{зайнято}\} = \frac{(1-\rho)\gamma(s)[1-\gamma(s)]}{E(B)[\lambda\gamma(s) - \lambda + s]}.$$

До цього ми розглядали (або спостерігали) заявку, яка надійшла в період зайнятості. Але треба «зняти» умову, що заявка надійде тільки в період зайнятості, оскільки вона може надійти і в період незайнятості, тобто у період простою. Таким чином, вилучаючи цю умову, отримаємо

$$\begin{aligned} \beta(s) &= E[e^{-sF}] = \\ &= (1-\rho)E[e^{-sF} \mid \text{надходження заявки у період простою}] + \\ &+ \rho E[e^{-sF} \mid \text{надходження заявки у період зайнятості}] = \\ &= (1-\rho)\gamma(s) + \rho E\{e^{-sF} \mid \text{зайнято}\} = \frac{(1-\rho)s\gamma(s)}{\lambda\gamma(s) - \lambda + s}. \end{aligned}$$

Для того щоб отримати моменти тривалості перебування заявки в системі, обчислюємо $\beta'(0)$ і $\beta''(0)$. Оскільки похідна від $\beta(s)$ в нулі має вигляд $0/0$, то слід двічі використати правило Лопітала. У результаті отримаємо

$$\begin{aligned} E(F) &= E(B) + \frac{\lambda E(B^2)}{2(1-\rho)}; \\ E(F^2) &= E(B^2) + \frac{\rho E(B^2)}{1-\rho} + \frac{\lambda E(B^3)}{3(1-\rho)} + \frac{[\lambda E(B^2)]^2}{2(1-\rho)^2}. \end{aligned}$$

Тривалість очікування заявки у черзі легко визначити, тому що тривалість перебування дорівнює сумі двох випадкових незалежних величин – тривалості очікування і тривалості обслуговування. Оскільки $\beta(s) = \alpha(s)\gamma(s)$, то

$$\alpha(s) = E(e^{-sw}) = \frac{(1-\rho)s}{\lambda\gamma(s) - \lambda + s}.$$

Перший і другий моменти тривалості очікування відповідно дорівнюють:

$$E(W) = \frac{\lambda E(B^2)}{2(1-\rho)};$$
$$E(W^2) = \frac{\lambda E(B^3)}{3(1-\rho)} + \frac{[\lambda E(B^2)]^2}{2(1-\rho)^2}.$$

Отже, очевидно, що визначати характеристики ємності буферної пам'яті для тимчасового зберігання інформаційних повідомлень до моменту звільнення каналу передавання інформації та затримки повідомлень системами передавання дані мережевими трактами дуже важко.

Для кращого усвідомлення методів дослідження такого класу систем наведемо узагальнений алгоритм:

1. Визначитись, що розуміється під станом системи, і встановити (знайти) імовірність знаходження системи в будь-якому стані.

2. Для стаціонарного режиму визначити поняття періодів зайнятості і незайнятості (простою) системи.

3. Обґрунтувати доцільність для теоретичного аналізу вибору дисципліни (порядку) обслуговування повідомлень у разі їх накопичення. Причому вибрати дисципліну обслуговування бажано таку, щоб можна було помітити деяку подібність окремих процесів обслуговування із загальним. (У схемі досліджень використовувалась схема Такача, за якою повідомлення, що надійшло останнім, обслуговується першим. Це дало можливість побудувати ЕПЗ, обвідна незавершених робіт яких точно збігається з обвідною загального періоду зайнятості, що надає право користуватись однаково розподіленою випадковою величиною інтервалу обслуговування).

4. Визначити період зайнятості системи обслуговування повідомлень. Період зайнятості складається з послідовності проміжків часу. Кожен проміжок утворюється кількістю надходжень повідомлень у попередньому проміжку. Це дозволяє визначити ймовірність знаходження системи в якомусь проміжку як частки часу надходження системи в його стані. Визначити час перебування заявки, що надійшла в якийсь момент часу в систему. Для цього розглядати період зайнятості як суму згаданих проміжків часу і момент часу надходження заявки в якийсь j проміжок. Моменти початків послідовності проміжків можна розглядати як моменти подій, які виникають згідно із законом потоку Пальма. Згідно з парадоксом, що випадкова точка попала в момент часу t на якийсь інтервал X , цей інтервал особливий та має характеристики, що відрізняються від характеристик інших інтервалів. Визначені характеристики цього інтервалу

використовуються для пошуку функції розподілу часу перебування заявки в системі – час затримки повідомлення системою СПДІ та час очікування в пам'яті системи.

16.5. Метод оптимізації процедур інформаційної підтримки адаптивної маршрутизації

16.5.1. Постановка завдання

Одна з головних проблем, які виникають під час створення й експлуатації телекомунікаційних мереж – проблема маршрутизації та керування потоками (трафіком) передаваних повідомлень. Мета будь-якого класу алгоритмів маршрутизації – забезпечення найкращої сукупності шляхів між джерелом інформації (адресантом) і споживачем (адресатом) за заданих вхідних потоків і конфігурації мережі. Найчастіше найкращими вважаються шляхи мінімальної затримки. Оскільки вхідні потоки і конфігурація мережі змінюються з часом (наприклад, через відмови ліній та вузлів), протокол маршрутизації має бути «адаптивним», тобто повинен пристосовувати маршрути до змінних у часі мережеских умов. Такий спосіб визначення маршруту доставляння повідомлення адресату називають адаптивною маршрутизацією, завдання якої полягає у визначенні процедури, що динамічно коригує маршрутні таблиці відповідно до змін станів у мережі та її вузлах.

Для формування маршрутних таблиць у вузлах потрібна інформація про миттєві зміни станів мережі та її трафіку. На жаль, вузли не можуть з точністю до наносекунд мати інформацію навіть про стан сусідніх вузлів, не кажучи вже про всю мережу. Це потребувало б непомірних накладних витрат, які ще називають системними витратами. Отже, вузли змушені приймати рішення за застарілою на півсекунди і більше інформацією.

Це стосується розділу будь-якої мережевої технології керування інформаційними потоками. До змінних, які відображають стан мережі і становлять інтерес для обчислення маршрутів, належать: топологічна зв'язність, структура трафіку, затримка і т. ін. Для того щоб визначити локальну зв'язність, кожний вузол контролює стан суміжних ліній і вузлів, відслідковуючи трафік у лінії або (якщо немає трафіку чи в примусовому порядку) періодично опитуючи сусідів. Навантаження локального трафіку контролюється за допомогою вимірювання або середніх (чи то миттєвих) довжин черг, або середніх потоків у вихідних лініях, або вхідного трафіку, який надходить із зовнішніх джерел.

Процес функціонування комутаційного вузла мережі за адаптивної маршрутизації за часом будемо розглядати як процес виконання певних робіт маршрутизатором мережі згідно з моделлю, яку показано на рис. 16.10.

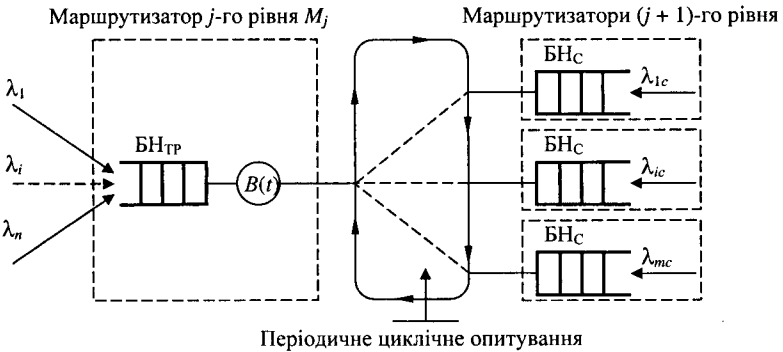


Рис. 16.10. Процес взаємодії маршрутизатора M_j з маршрутизаторами $(j + 1)$ -го рівня: БН_{ТР} і БН_С – нагромаджувачі для повідомлень відповідно до інформаційного трафіку та трафіку сигналізації про зміни стану маршрутизаторів

Маршрутизатор j -го рівня M_j відповідно до базової дисципліни періодичного опитування виконує операцію «опитування», згідно з якою нагромаджена інформація у буфері опитуваного вузла $(j + 1)$ -го рівня передається каналом зв'язку у маршрутизатор-ініціатор M_j . Так маршрутизатор-ініціатор проводить періодично сеансові опитування всіх сусідніх вузлів $(j + 1)$ -го рівня M_{j+1} , з якими у нього є канал зв'язку. На взаємодію маршрутизатора-ініціатора з одним із сусідніх маршрутизаторів витрачається процесорний час і час зайняття каналу зв'язку. Таким чином, маршрутизатор-ініціатор періодично опитує всі сусідні маршрутизатори, утворюючи сеанс (цикл) взаємодії. Витрати часу зайняття його на взаємозв'язок називають *системними витратами* або *системним навантаженням* (системним трафіком).

Описаний процес демонструється часовою діаграмою (рис. 16.11), на якій $\sigma^*_j = (\sigma_1, \dots, \sigma_i, \dots, \sigma_n)$ – цикл опитування (σ_i – час опитування i -го маршрутизатора); λ_i – інтенсивність надходження інформаційних повідомлень (заявок на обслуговування); λ_i^c – інтенсивність виникнення повідомлень про зміни станів відносно інформації про них, яка опитувалась у попередньому циклі; τ – інтервал обміну інформаційними повідомленнями. Отже, $(\tau + \sigma)$ – період опитування (інтервал між сусідніми моментами початку опитування).

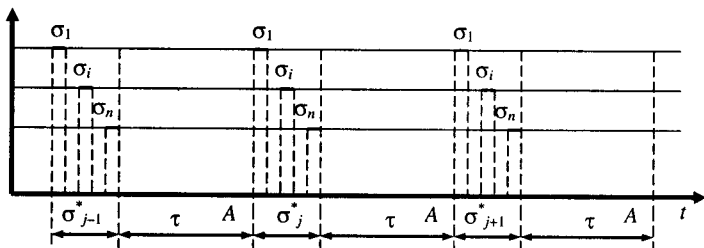


Рис. 16.11. Часова діаграма опитування маршрутизаторів

Принципові труднощі, з якими доводиться стикатися під час визначення періоду виконання циклів опитування маршрутизаторів, виникають через дві суперечливі вимоги, яким має відповідати організація ефективного керування інформаційним трафіком. Ефективне використання маршрутизатора передбачає мінімізацію його витрат на опитування сусідніх маршрутизаторів, оскільки чим рідше виникають цикли опитування, тим менше відволікається він на роботу із системним трафіком. Проте надто рідкісні опитування сусідів призводять до порушення своєчасності забезпечення інформаційної підтримки основного алгоритму маршрутизації (запізнення призводять до старіння інформації). Це і спонукає до пошуку такого періоду опитування, за якого витрати часу маршрутизатором на опитування сусідніх маршрутизаторів були мінімальні і при цьому забезпечувався б мінімальний вплив їх на затримку передавальних трафікових інформаційних повідомлень.

16.5.2. Модель процесу передавання повідомлення системою маршрутизатора – каналу зв'язку

Розглянемо поведінку у часі довільного маршрутизатора j -го рівня у процесі виконання роботи з вибору напряму передавання інформаційних потоків повідомлень (заявок на тимчасове використання ресурсів як маршрутизатора, так і каналів передавання даних до вибраного маршрутизатора $(j + 1)$ -го рівня). Припустімо, що загальний потік заявок $\lambda = \sum_{i=1}^n \lambda_i$, який надходить від n маршрутизаторів $(j - 1)$ -го рівня, описується пуассонівським потоком подій (заявок на обслуговування).

Адаптивні алгоритми маршрутизації, застосовувані в реальних мережах, класифікують за інформацією, яка використовується ними для прийняття маршрутних рішень. Один із головних параметрів будь-якого методу маршрутизації – затримка заявок у вузлах і каналах зв'язку.

Для визначення впливу на час затримки заявок системою «маршрутизатор – канал зв'язку» (М – К) розглянемо типову ситуацію функціонування її в мережі за часом на сумісній часовій діаграмі (рис. 16.12). Уведемо деякі додаткові позначення:

B – тривалість передавання повідомлення каналом зв'язку (час обслуговування заявки) без упровадження процедури опитування – випадкова величина з довільною функцією розподілу $B(t)$ і щільністю розподілу ймовірностей $dB(t)$;

H – тривалість передавання інформаційного повідомлення каналом зв'язку в разі строго регламентованих перерв, зумовлених вимушеними тимчасовими припиненнями передавання його на час σ^* – час опитування маршрутизаторів $(j + 1)$ -го рівня – випадкова величина з довільною функцією розподілу $H(t)$ і щільністю розподілу ймовірностей $dH(t)$;

T – випадковий момент часу надходження поміченого повідомлення з інформаційного потоку;

θ – інтервал часу між T і початком передавання його каналом зв'язку – випадкова величина з довільною функцією розподілу $\theta(t)$. Ця функція описує час очікування заявки-повідомлення у БНтр маршрутизаторі (час перебування заявки у черзі).

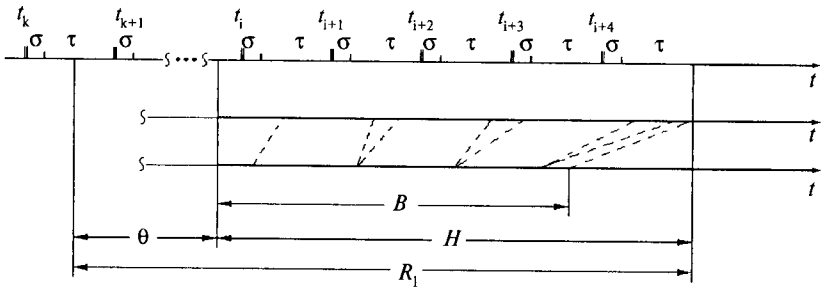


Рис. 16.12. Діаграма сумісного часу опитування і передавання

Розглянемо тривалість R_1 часу перебування інформаційного повідомлення-заявки у системі М – К від моменту його надходження в маршрутизатор M_j до моменту доставки до вибраного M_i $(j + 1)$ -го рівня. Для визначення R_1 використаємо таку модель. Моменти t_i звернення маршрутизатора M_j на опитування маршрутизаторів $(j + 1)$ -го рівня являють собою послідовність подій, які виникають у момент часу, інтервали між якими однакової довжини $(\sigma + \tau)$ – регулярний потік подій. Незалежно від моментів виникнення подій цього потоку візьмемо на осі часу точку T , яка визначає момент надходження заявки-повідомлення в маршрутизатор M_j (зрозуміло, що момент T надходження заявки не залежить від моментів

початку циклу опитування). У такому разі θ і H – невід’ємні незалежно розподілені випадкові величини.

Перетворення Лапласа щільності розподілу суми θ і H за визначенням має вигляд

$$M[\exp\{-s(\theta + H)\}] = M(e^{-s\theta} e^{-sH}) = \varphi(s)\eta(s) = r_1(s),$$

де $r_1(s)$ – перетворення Лапласа розподілу часу перебування заявки у системі $M - K$; $\varphi(s)$ – перетворення Лапласа розподілу часу перебування заявки у маршрутизаторі (час очікування у черзі); $\eta(s)$ – перетворення Лапласа розподілу часу зайняття каналу зв’язку передавання повідомлення.

Перейдемо до визначення $\eta(s)$. Розглянемо сумарний час зайнятості системи $M - K$ передаванням повідомлення, яке надійшло в момент T , урахувавши при цьому те, що протягом H будуть виникати інтервали-перерви σ маршрутизатора M_j від передавання цього поміченого повідомлення на селективно вибраний маршрутизатор $(j + 1)$ -го рівня.

Знайдемо вигляд функції розподілу інтервалу $H - H(t)$. Імовірність того, що час перебування заявки в системі $M - K$ знаходиться в інтервалі $[x, x + \Delta x]$, дорівнює $dB(x)$. Припустимо, що час перебування заявки дорівнює t і за цей час виникло k разів відхилення на опитування з імовірністю $P_k(x)$. Для того щоб сумарний час заняття $M - K$ цим повідомленням за фіксованих x і k не перевищував t , необхідно і достатньо, щоб за час $(t - x)$ маршрутизатор відхилявся k разів на опитування, імовірність чого $F_1^{*k}(t - x)$, де $F_1(t)$ – функція розподілу часу опитування зі щільністю розподілу $f_1(t) = \delta(t - \sigma)$. Тоді $H(t)$ – функція розподілу загального часу перебування заявки у маршрутизаторі-каналі з урахуванням відхилень останнього на опитування матиме вигляд

$$H(t) = \int_0^t \sum_{k=0}^{\infty} F_1^{*k}(t - x) P_k(x) dB(x),$$

де $*$ – символ стільтєсової згортки.

Перейдемо до перетворення Лапласа, тоді отримаємо

$$\eta(s) = \int_0^{\infty} e^{-st} dH(t) = \int_0^{\infty} e^{-st} \sum_{k=0}^{\infty} F_1^{*k}(t - x) P_k(x) dB(x). \quad (16.40)$$

Вираз $\sum_{k=0}^{\infty} F_1^{*k}(t-x)P_k(x)$ замінимо твірною функцією $P[f_1(s), t] = \sum_{k=0}^{\infty} f_1^k(s)P_k(t)$, де $f_1(s) = \int_0^{\infty} e^{-st} f_1(t) dt$, тоді вираз (16.40) набуде вигляду

$$\eta(s) = \int_0^{\infty} e^{-st} P[f_1(s), t]. \quad (16.41)$$

Функцію $P[f_1(s), t]$ знаходимо, виходячи з таких міркувань. Імовірність розподілу кількості k відхилень маршрутизатора від виконання роботи з передавання інформаційного повідомлення (обслуговування заявки) на інтервалі t

$$P_k(t) = \begin{cases} 1, & \text{якщо } k = \left\lfloor \frac{t}{\tau} \right\rfloor; \\ 0, & \text{якщо } k \neq \left\lfloor \frac{t}{\tau} \right\rfloor, \end{cases} \quad (16.42)$$

де $\lfloor * \rfloor$ – ціла частина відношення t/τ .

З урахуванням виразу (16.42) твірна функція $P[f_1(s), t]$ набуде вигляду

$$P[f_1(s), t] = \sum_{k=0}^{\infty} f_1^k(s) P_k(t) = f_1^{\left\lfloor \frac{t}{\tau} \right\rfloor}(s) = e^{\left\lfloor \frac{t}{\tau} \right\rfloor \ln f_1(s)}.$$

Перетворення Лапласа щільності розподілу $f_1(t)$ буде таким:

$$f_1(s) = \int_0^{\infty} e^{-st} f_1(t) dt = \int_0^{\infty} e^{-st} \delta(t - \sigma) dt = e^{-s\sigma},$$

тоді

$$P[f_1(s), t] = e^{\left\lfloor \frac{t}{\tau} \right\rfloor \ln e^{-s\sigma}} = e^{-s\sigma \left\lfloor \frac{t}{\tau} \right\rfloor}. \quad (16.43)$$

Підставимо рівняння (16.43) у вираз (16.41), отримаємо $\eta(s)$ у вигляді

$$\eta(s) = \int_0^{\infty} e^{-s \left(t + \sigma \left\lfloor \frac{t}{\tau} \right\rfloor \right)} dB(t). \quad (16.44)$$

Для отримання верхньої оцінки вираз (16.44) набуде вигляду

$$\eta(s) = \int_0^{\infty} e^{-st \left(1 + \frac{\sigma}{\tau}\right)} dB(t) = \beta(s\alpha), \quad (16.45)$$

де $\alpha = 1 + \frac{\sigma}{\tau}$.

Щоб отримати середні характеристики тривалості передавання інформаційного повідомлення каналом зв'язку між маршрутизаторами, обчислюємо $-\eta'(0)$ і $\eta''(0)$:

$$\begin{aligned} \bar{\eta} = -\eta'(0) &= \left(1 + \frac{\sigma}{\tau}\right) \int_0^{\infty} t^* dB(t) = \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta}; \\ \eta'' = \eta''(0) &= \left(1 + \frac{\sigma}{\tau}\right)^2 \int_0^{\infty} t^2 dB(t) = \left(1 + \frac{\sigma}{\tau}\right)^2 \bar{\beta}^2, \end{aligned} \quad (16.46)$$

де $\bar{\beta}$ і $\bar{\beta}^2$ – відповідно перші початкові моменти розподілу часу перебування повідомлення у каналі зв'язку без урахування витрат на тривалість і частоту опитування.

Тепер визначимо $\phi(s)$. Наведений опис процесу функціонування системи М – К (вхідний трафіковий інформаційний потік плюс маршрутизатор M_j зі своїм буферним нагромаджувачем БНтр плюс канал зв'язку для передавання повідомлень по визначеному маршруту через M_j ($i + 1$)-го рівня) відображається адекватно динамічним процесом масового обслуговування моделі $M|G|1$, стаціонарний режим якої визначається нерівністю $\rho = \lambda \bar{x} < 1$, де \bar{x} – середнє значення інтервалу обслуговування заявки.

В однолінійній системі, на яку надходить пуассонівський потік, періоди зайнятості і незайнятості утворюють дві чергуючі та незалежні послідовності однаково розподілених випадкових величин. Розподіл тривалості періоду зайнятості відіграє суттєву роль у визначенні залежності розподілу тривалості перебування (затримки) заявки в черзі або системі. Для кращого розуміння впливу на період зайнятості $M|G|1$ періодичного виникнення перерв на опитування розглянемо часову діаграму, показану на рис. 16.13.

Розглянемо процес утворення періоду зайнятості системи М – К як суму інтервалів ЕПЗ. Період зайнятості розпочинається обслуговуванням заявки, що надійшла першою, і тих заявок, що надійшли за час її обслуговування, утворюючи ЕПЗ₁. Протягом тривалості ЕПЗ₁ надійдуть нові заявки, які утворюють ЕПЗ₂, і т. д. Тоді повний період зайнятості буде складатись із взаємозв'язаних інтервалів ЕПЗ, за яких допускається можливість нескінченного числа проміжків часу. Зрозуміло, що для тих проміж-

ків, які лежать поза періодом зайнятості, тривалість ЕПЗ; дорівнює нулю. Якщо $\rho < 1$ відомо з імовірністю одиниця, існує кінцеве число i_0 , для якого проміжок ЕПЗ_{*i*} і тривалості подальших проміжків дорівнюють нулю.

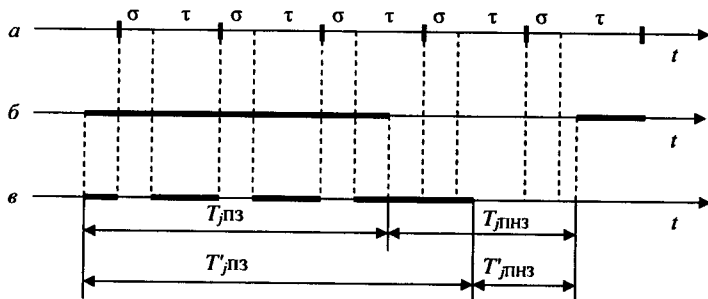


Рис. 16.13. Часова діаграма періоду зайнятості системи $M|G|1$:
a – послідовність періодів опитування; *б* – послідовність чергування періодів зайнятості та незайнятості для моделі $M|G|1$; *в* – послідовність чергування періодів зайнятості та незайнятості через вплив перерв на опитування;
 $T_{jлз}$ і $T'_{jлз}$ – відповідно *j*-ті періоди зайнятості і незайнятості моделі $M|G|1$;
 $T_{лз}$ і $T'_{лз}$ – відповідні періоди $M|G|1$ на періодичні перерви на опитування

Позначимо тривалість інтервалу ЕПЗ_{*i*} через x_i . Можна отримати рекурентне співвідношення між перетвореннями Лапласа інтервалів $x_i(s)$ і $x_{i-1}(s)$ через відомі λ та $\beta(sa)$ (див. вираз (16.45)). Далі будемо спостерігати за «поміченою» заявкою, яка надійде на *i*-й ЕПЗ(x_i), причому використаємо із теорії альтернативних процесів відновлення граничну стаціонарну ймовірність стану як відношення тривалості ЕПЗ(x_i) до повної тривалості періоду зайнятості, а також умови надходження заявки у період зайнятості $\lambda \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta} = \rho$ та період незайнятості $1 - \lambda \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta} = 1 - \rho$. Для «зняття» умови, що помічена заявка надійшла в *i*-й ЕПЗ, виконаємо підсумовування за *i*, якщо $i = 0, 1, \dots, \infty$. Отримаємо $\phi(s)$ у вигляді

$$\phi(s) = \frac{s \left[1 - \lambda \left(1 + \frac{\sigma}{\tau} \right) \bar{\beta} \right]}{s - \lambda + \lambda \beta \left[s \left(1 + \frac{\sigma}{\tau} \right) \right]}.$$

Середні значення затримки інформаційних повідомлень у маршрутизаторі (буферному нагромаджувачі) знайдемо, обчисливши $-\varphi'(0)$ і $\varphi''(0)$. Оскільки похідна від $\varphi(s)$ у точці $s = 0$ має вигляд $0/0$, то правило Лопітала потрібно застосовувати двічі. У результаті матимемо:

$$\bar{\varphi} = -\varphi'(0) = \frac{\lambda \left(1 + \frac{\sigma}{\tau}\right)^2 \bar{\beta}^2}{2 \left[1 - \lambda \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta}\right]}; \quad (16.47)$$

$$\bar{\varphi}^2 = \varphi''(0) = \frac{\lambda \left(1 + \frac{\sigma}{\tau}\right)^3 \bar{\beta}^3}{3 \left[1 - \lambda \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta}\right]} + \frac{\left[\lambda \left(1 + \frac{\sigma}{\tau}\right)^2 \bar{\beta}^2\right]^2}{2 \left[1 - \lambda \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta}\right]^2}. \quad (16.48)$$

За формулами (16.47) і (16.48) можна обчислювати середнє значення та дисперсію затримки повідомлення заявки у маршрутизаторі. Їх можна використати як інформацію для підтримки параметрів маршрутної матриці.

Тепер знаходимо $r_1(s)$ перетворення Лапласа часу затримки заявки інформаційного трафікового повідомлення маршрутизатором M_j і каналом зв'язку із селективно вибраним маршрутизатором $(j + 1)$ -го рівня:

$$r_1(s) = \frac{s \left[1 - \lambda \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta}\right]}{s - \lambda + \lambda \bar{\beta} \left[s \left(1 + \frac{\sigma}{\tau}\right)\right]} \bar{\beta} \left[s \left(1 + \frac{\sigma}{\tau}\right)\right], \quad (16.49)$$

Рівняння (16.49) виражається через відомі величини, отримані з початкової постановки задачі: функцію розподілу часу зайнятості каналу передавання трафікового інформаційного повідомлення (час обслуговування заявки), λ – параметр вхідного потоку і σ – часові витрати на обмін службовою інформацією, а також через поки ще не визначений параметр τ , який разом з σ являє собою частоту опитування. Вираз (16.49) за формою є рівнянням Полячека – Хінчина для $M|G|1$, якщо $\tau \rightarrow \infty$.

Середні значення повної затримки інформаційної заявки-повідомлення системою $M - K$ – час доставки повідомлення від моменту його надходження у M_j до моменту рестрації вибраним маршрутизатором M_j $(j + 1)$ -го рівня, визначаємо з рівняння (16.49) так само, як і з формули (16.47) та (16.48), тобто

$$\begin{aligned} \bar{r}_1 &= -r_1'(0) = \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta} + \frac{\lambda \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta}^2}{2 \left[1 - \lambda \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta}\right]}; \\ \bar{r}_1^2 &= r_1''(0) = \left(1 + \frac{\sigma}{\tau}\right)^2 \bar{\beta}^2 + \frac{\lambda \left(1 + \frac{\sigma}{\tau}\right)^3 \bar{\beta} \bar{\beta}^2}{1 - \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta}} + \\ &+ \frac{\lambda \left(1 + \frac{\sigma}{\tau}\right)^3 \bar{\beta}^3}{3 \left[1 - \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta}\right]} + \frac{\left[\lambda \left(1 + \frac{\sigma}{\tau}\right)^2 \bar{\beta}^2\right]^2}{2 \left[1 - \lambda \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta}\right]^2}. \end{aligned} \quad (16.50)$$

Знаходимо границю \bar{r}_1 як функцію від τ , якщо $\tau \rightarrow \infty$:

$$\lim_{\tau \rightarrow \infty} \bar{r}_1(\tau) = \bar{\beta} + \frac{\lambda \bar{\beta}^2}{2(1\lambda\bar{\beta})}. \quad (16.51)$$

Результат (16.51) означає, що при $\tau \rightarrow \infty$ система М – К функціонує без перерв (витрат на опитування). Це відомий результат для $M|G|1$. Функція (16.50) $\bar{r}_1(t)$ монотонно спадає і при $\tau \rightarrow \infty$ наближається до границі, визначеної з виразу (16.51).

Отже, отримані результати (16.46), (16.47) і (16.50) визначають функціональну залежність передавання повідомлення каналом, часу очікування у маршрутизаторі та часу затримки повідомлення системою М – К (час доставки) від тривалості періоду опитування $(\tau + \sigma)$.

16.5.3. Вплив періоду опитування на затримку інформації про стан M_j маршрутизатора

Розглянемо вплив періоду опитування на затримку своєчасного інформування маршрутизатора M_j про зміни станів опитуваних маршрутизаторів M_i $(j + 1)$ -го рівня. Очевидно, що інформування M_j про миттєві зміни параметрів станів за ініціативою M_i $(i = 1, \dots, m)$ може призвести як до перевантаження каналів, так і до повного завантаження M_j роботою з неперервного відновлення маршрутних таблиць і розрахунками з визначення маршрутів трафікових інформаційних повідомлень.

Зазначимо, що значення параметра σ визначається протокольними процедурами (кадрами та командами опитування), кількістю вимірюваних параметрів стану опитуваного вузла та кількістю опитуваних вузлів. Тому для визначення впливу частоти опитування на затримки повідомлень R_2 про зміни параметрів стану обмежимося визначенням часу затримки повідомлення R_2 .

Розглянемо типову ситуацію періодичного опитування, яку зображено на рис. 16.14.

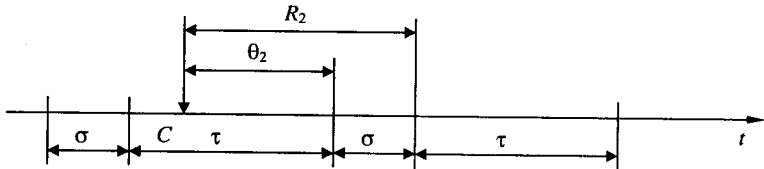


Рис. 16.14. Часова діаграма періодичного опитування сусіднього маршрутизатора:

C – момент виникнення повідомлення про зміну значення контрольованого параметра; θ_2 – інтервал часу між C і початком наступного періоду опитування;

R_2 – час затримки повідомлення про зміну стану та його обробкою для матриці маршрутизації

У загальному випадку щільність розподілу того інтервалу (j -го періоду опитування), на який випадково випала точка C , визначають так:

$$f^*(t) = \begin{cases} \frac{t}{m_t} f(t), & \text{якщо } t > 0; \\ 0, & \text{якщо } t < 0, \end{cases} \quad (16.52)$$

де $f^*(t)$ – щільність розподілу ймовірностей випадкових величин (інтервалів між двома сусідніми подіями – початками періодів опитування);

$m_t = M[\tau + \sigma] = \int_0^{\infty} t f(t) dt = \tau + \sigma$ – математичне сподівання інтервалу $\tau + \sigma$.

Для цього випадку, коли інтервал строго постійний між початками періодів опитування, щільність розподілу інтервалу, на який випала точка C , буде мати вигляд

$$f^*(t) = \frac{t}{m_t} f(t) = \frac{t}{(\tau + \sigma)} \sigma [t - (\tau + \sigma)] = \sigma [t - (\tau + \sigma)]. \quad (16.53)$$

На підставі виразів (16.52), (16.53) знаходимо закон розподілу часу $(\theta_2 + \sigma) = R_2$ від випадкової точки C до моменту закінчення циклу опитування і отримуємо:

$$r_2(t) = \frac{1 - R_2(t)}{\tau + \sigma} = \frac{1 - \int_0^t \delta[t - (t + \sigma)]}{\tau + \sigma} = \left\{ \begin{array}{l} \frac{1}{\tau + \sigma}, \quad \text{якщо } t \in (0, t + \sigma); \\ 0, \quad \text{якщо } t \notin (0, t + \sigma). \end{array} \right\}$$

Перетворення Лапласа $r_2(s)$ буде таким:

$$r_2(s) = \frac{1}{(\tau + \sigma)s} (1 - e^{-s(\tau + \sigma)}).$$

Як і для \bar{r}_1 знаходимо \bar{r}_2 :

$$\bar{r}_2 = -r_2(0) = \frac{\tau + \sigma}{2}. \quad (16.54)$$

Із формули (16.54) видно, що \bar{r}_2 як функція від τ , тобто $\bar{r}_2(\tau)$, лінійно зростає, якщо $\tau \rightarrow \infty$ і $\sigma = \text{const}$. Цю функцію можна розглядати як функцію від змінного аргументу σ , значення якого залежать від числа m – кількості опитуваних маршрутизаторів та інтенсивності зміни параметрів станів опитуваних вузлів. За кількістю $\sigma = \text{const}$ можна спростити операції запропонованого методу.

Будемо вважати оптимальним таке значення періоду опитування $(\tau + \sigma)$, за якого величина τ забезпечує мінімум сумарних часових затрат на затримку трафікових інформаційних повідомлень і мінімум сумарних затрат на своєчасність інформування сусідніх вузлів про зміни параметрів, які вимірюють стан вузла. Затрати на затримки повідомлень про зміни спостережуваних параметрів (станів) відповідно до рівняння (16.54) збільшуються зі зростанням періоду $(\tau + \sigma)$ і пропорційні середньому часу очікування \bar{r}_2 тобто часу від моменту виникнення повідомлення про зміни станів до завершення обробки його опитувальним маршрутизатором. Якщо ввести коефіцієнт пропорційної зміни середнього значення затримки, тоді він буде мати зміст штрафу за одиницю часу затримки повідомлень про зміни стану.

Функція $\bar{r}_1(\tau)$ монотонно спадає зі зростанням τ , а функція $\bar{r}_2(\tau)$ лінійно зростає, і, отже, функція $\bar{r}_1(\tau) + \bar{r}_2(\tau)$ має єдиний мінімум, який є розв'язком рівняння:

$$F(\tau) = \bar{r}_1 + \bar{r}_2 = \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta} + \frac{\lambda \left(1 + \frac{\sigma}{\tau}\right)^2 \bar{\beta}^2}{2 \left(1 - \lambda \left(1 + \frac{\sigma}{\tau}\right) \bar{\beta}\right)} + \frac{\tau + \sigma}{2}. \quad (16.55)$$

Оптимальне значення τ для неперервної та диференційованої функції $F(\tau)$ можна знайти, якщо прирівняти нулю першу похідну $F(\tau)$ за τ , а потім із отриманого рівняння знайти τ .

16.5.4. Урахування ненадійності та відновлення каналів зв'язку

Дотепер проводився системний аналіз поведінки у часі маршрутизатора мережі за адаптивної маршрутизації та надійних каналів зв'язку. У ТПД зв'язку виникають збойні ситуації, які призводять до порушення нормального режиму передавання даних. Боротьба з цими порушеннями і їх впливами на достовірність передавання інформації проводиться впровадженнями засобів завадостійкого кодування та зворотного вирішального зв'язку, який діє через протокольні процедури повторення передавання інформації, що зрештою зводиться до додаткових часових затрат на використання каналних ресурсів. Зазначимо, що темпи (інтенсивності) завантаження каналу зв'язку (особливо на магістральних лініях) трафіковими потоками і темпи виникнення «збойних ситуацій» значно відрізняються, тому їх ще називають «рідкісними подіями». Визначимо вплив цих подій на динамічний процес обміну потоками в ланках мережі.

Звернувши увагу на функціональні залежності $\varphi(s)$ і $r_1(s)$, помітимо, що в них одним з аргументів є перетворення Лапласа часу обслуговування (часу передавання інформаційного повідомлення).

Для визначення впливу ненадійності та відновлення каналу зв'язку на ефективність використання каналу для передавання трафікових повідомлень уведемо такі характеристики:

$H_1(t)$ – функція розподілу тривалості часу до моменту виникнення відмови наступного передавання повідомлення;

$H_2(t)$ – функція розподілу тривалості часу, витрачаного на відновлення робочого стану каналу – ліквідація наслідків відмови;

$V(t)$ – функція розподілу часу обслуговування заявки – тривалість часу передавання повідомлення каналом;

$\Phi(t)$ – функція розподілу «сумарного» часу обслуговування заявки – повний час зайняття каналу передаванням повідомлення з урахуванням ліквідації відмови.

Нехай $H_1(t)$ – функція розподілу часу виникнення відмови, $H_2(t)$ – функція розподілу часу реакції на відмову (відновлення), тоді функція $\Phi(t)$ розподілу сумарного часу обслуговування повідомлення з урахуванням відмов і відновлень набуде вигляду

$$\begin{aligned} \Phi(t) = & \int_0^t [1 - H_1(t-x)] dB(x) + \\ & + \int_0^t [H_2(t-x) * \Phi(t-x)] [1 - B(x)] dH_1(x). \end{aligned} \quad (16.56)$$

Дійсно, для того щоб сумарний час обслуговування повідомлення не перевищував час t (імовірність чого дорівнює $\Phi(t)$), необхідно і достатньо, щоб або за час обслуговування повідомлення, що не перевищував t , не виникло ні однієї відмови (імовірність чого дорівнює першому доданку в правій частині рівняння (16.56)), або щоб перша відмова виникла у проміжку $[x, x + \Delta x]$, $x \leq t$ і час реакції на відмову (відновлення робочого стану) плюс нова реалізація сумарного часу обслуговування із самого початку не перевищували $(t-x)$ (імовірність чого дорівнює другій складовій правій частині рівняння (16.56)).

Нехай $\varphi(s)$, $h_1(s)$ і $h_2(s)$ – перетворення Лапласа – Стілтєса функцій розподілів $\Phi(t)$, $H_1(t)$, $H_2(t)$ відповідно. Тоді, розв'язавши рівняння (16.56) за допомогою перетворень Лапласа – Стілтєса, дістаємо

$$\varphi(s) = \frac{B(s) - \delta(s)}{1 - h_2(s)[h_1(s) - \Delta(s)]}, \quad (16.57)$$

$$\text{де } \delta(s) = \int_0^{\infty} e^{-st} H_1(t) dB(t), \quad \Delta(s) = \int_0^{\infty} e^{-st} B(t) dH_1(t).$$

Для окремого випадку, коли потік відмов пуассонівський $H_1(t) = 1 - e^{-\lambda t}$, визначимо $\varphi(s)$. Вирази для $\delta(s)$, $\Delta(s)$ і $h_1(s)$ мають вигляд

$$\delta(s) = w(s) - w(s - \lambda); \quad \Delta(s) = \frac{\lambda}{\lambda + s} w(s + \lambda); \quad h_1(s) = \frac{\lambda}{\lambda + s}. \quad (16.58)$$

Підставляючи рівняння (16.58) у вираз (16.57), дістаємо

$$\varphi(s) = \frac{\beta(s + \lambda)}{1 - \lambda h_2(s) \frac{1 - \beta(s + \lambda)}{s + \lambda}}. \quad (16.59)$$

Продиференціювавши формулу (16.59) за s і поклавши $s = 0$, одержуємо середню тривалість обслуговування $\bar{\varphi}$ повідомлення з урахуванням тимчасових витрат на відновлення:

$$\bar{\varphi} = -\varphi'(0) = \frac{(1 + \lambda \bar{h}_2)[1 - \beta(\lambda)]}{\lambda \beta(\lambda)}, \quad (16.60)$$

$$\text{де } \beta(\lambda^*) = \int_0^{\infty} e^{-\lambda t} dB(t); \quad \bar{h}_2 = - \int_0^{\infty} t dH_2(t).$$

Отже, вираз (16.60) являє собою середнє значення часу передавання інформаційного повідомлення у разі випадкових відмов і відновлень, причому після відновлення відмовної ситуації повідомлення заново передається каналом, тобто раніше витрачений час на обслуговування не враховується. Такі події виникають під час передавання інформації каналами зв'язку, у яких використовують для підвищення достовірності передавання засоби завадостійкого кодування разом з процедурою вирішального зворотного зв'язку.

Запропонований метод складається із трьох незалежних складових: 1) окремо визначається вплив затрат часу на тривалість передавання повідомлення каналом зв'язку періодичністю опитування (див. вираз (16.45)); 2) час затримки повідомлення маршрутизатором (16.49) і функція $F(\tau)$ (16.55) дають змогу отримати оптимальне значення періоду опитування; 3) перетворення Лапласа (16.57) дозволяє оцінити вплив ненадійності каналу на ефективне передавання повідомлення.

І нарешті, підставлянням виразу (16.57) в рівняння (16.45) і (16.55) отримано загальний результат – алгоритм оптимального пошуку величини періоду сканування за ненадійної роботи каналу.

Контрольні питання

1. Загальне завдання системного аналізу передавання даних.
2. Принципи синхронного та асинхронного передавання цифрових даних.
3. Різновид асинхронного методу передавання цифрових даних – старт-стопне передавання. Позитивні якості та недоліки.
4. Синхронний метод передавання. Формування кадру передавання. Загальна структура кадру. Вимоги до ознак початку та кінця кадру.
5. Заборонені кодові комбінації для передавання інформаційних символів.
6. Принцип забезпечення «прозорості» передавання інформації кадрами. Визначення протоколу передавання даних.
7. Визначення байтзорієнтованого протоколу двійкового синхронного зв'язку.
8. Структура байтзорієнтованого протоколу. Призначення керувальних кадрів і склад загального формату інформаційного кадру.

9. Протокол BSC. Розміщення пакета в кадрі.
10. Функціонування системи обміну даними за протоколом BSC у разі безпомилкового передавання сигналів.
11. Алгоритм обміну даними протоколом BSC у разі виникнення помилок у каналі.
12. Класифікація та призначення керувальних байтів у протоколі BSC.
13. Операція байтстафіngu. Процедура розрізнення кодових комбінацій.
14. Позитивні якості та недоліки байтзорієнтованих протоколів.
15. Принципи побудови бітзорієнтованих протоколів. Призначення прапорців у протоколі HDLC.
16. Операція бітстафіngu. Процедура введення спеціального біта для забезпечення прозорості кадру.
17. Структура кадру протоколу HDLC.
18. Призначення керувальних байтів у кадрі HDLC. Змінювання адресних полів. Керувальні перевірни послідовності.
19. Алгоритм обміну даними в режимі півдуплексного зв'язку.
20. Обмін даними в дуплексному режимі.
21. Особливості обміну даними в разі виникнення помилок у каналі зв'язку.
22. Способи підвищення достовірності передавання дискретної інформації.
23. Класифікація методів.
24. Поняття про надмірне кодування дискретної інформації.
25. Принципи виявлення та виправлення помилок.
26. Постановка завдання завадостійкого кодування.
27. Класифікація завадостійкості кодів.
28. Лінійні (систематичні) коди. Коди Хеммінга.
29. Властивості лінійних (систематичних) кодів.
30. Побудувати лінійний (систематичний) код за заданою породжувальною матрицею.
31. Принцип неперервного згорткового кодування.
32. Згорткове кодування із синдромною корекцією.
33. Згорткові коди з використанням послідовного декодування.
34. Загальні поняття надійності передавання дискретної інформації.
35. Класифікація відмов (відказів).
36. Характеристики надійності трактів передавання інформації.
37. Розрахунок надійності каналів передавання даних.
38. Загальні положення впливу ненадійності трактів на ефективність передавання інформації.
39. Урахування ненадійності та відновлення у разі пасивних переривань.
40. Урахування ненадійності та відновлення у разі активних переривань.
41. Основні положення системного аналізу функціонування СПДІ.
42. Моделювання дуплексного режиму роботи СПДІ.
43. Метод оцінювання ємності буферної пам'яті.
44. Період зайнятості системи $M/D/1/N$.
45. Розподіл кількості повідомлень у періоді зайнятості.
46. Метод оцінювання використовуваної ємності буферної пам'яті.
47. Побудувати математичну модель півдуплексного режиму СПДІ.
48. Метод вкладених ланцюгів Маркова і його складові етапні кроки.

49. Загальні положення оцінювання затримок повідомлень у вузлах маршрутизації.
50. Простір станів СПДІ за півдуплексного режиму функціонування.
51. Період зайнятості системи. Графічне зображення тривалості періоду зайнятості.
52. Розподіл часу затримки повідомлень СПДІ.

Список рекомендованой литературы

1. *Андреев В. С.* Теория нелинейных электрических цепей: Учеб. пособие для вузов. – М.: Радио и связь, 1982. – 280 с.
2. *Банкет В. Л., Дорофеев В. М.* Цифровые методы в спутниковой связи. – М.: Радио и связь, 1988. – 240 с.
3. *Покровский Н. Б.* Телефония. Теория и техника передачи речи: Учеб. для вузов. – Л.: ВКАС, 1970. – 286 с.
4. *Вемян Г. В.* Передача речи по сетям электросвязи. – М.: Радио и связь, 1985. – 272 с.
5. *Быков Ю. С.* Теория разборчивости речи и повышение эффективности радиотелефонной связи. – Л.: Энергоиздат, 1959. – 351 с.
6. *Горелов Г. В., Ромашкова О. Н., Чан Туан Ань.* Качество управления речевым трафиком в телекоммуникационных сетях. – М.: Радио и связь, 2001. – 112 с.
7. *Лазарев В. Г.* Интеллектуальные цифровые сети. – М.: Финансы и статистика, 1996. – 224 с.
8. *Теория передачи сигналов: Учеб. для вузов.* / А. Г. Зюко, Д. Д. Кловский, М. В. Назаров, Л. М. Финк. – М.: Радио и связь, 1986. – 302 с.
9. *Назаров М. В., Прохоров Ю. Н.* Методы цифровой обработки и передачи речевых сигналов. – М.: Радио и связь, 1985. – 176 с.
10. *Теория электрической связи: Учеб. для вузов.* / А. Г. Зюко, Д. Д. Кловский, В. И. Коржик, М. В. Назаров. – М.: Радио и связь, 1998. – 432 с.
11. *Помехоустойчивость и эффективность систем передачи информации* / Под ред. А. Г. Зюко. – М.: Радио и связь, 1985. – 279 с.
12. *Тихонов В. И., Харисов В. Н.* Статистический анализ и синтез радиотехнических устройств и систем: Учеб. пособие для вузов. – М.: Радио и связь, 1991. – 608 с.
13. *Финк Л. М.* Теория передачи дискретных сообщений. – М.: Сов. радио, 1970. – 727 с.
14. *Левин Б. Р.* Теоретические основы статистической радиотехники. – М.: Радио и связь, 1989. – 653 с.
15. *Емельянов Г. А., Шварцман В. О.* Передача дискретной информации: Учеб. для вузов. – М.: Радио и связь, 1982. – 424 с.
16. *Витерби А. Д., Омура Д. К.* Принципы цифровой связи: Пер. с англ.; Под ред. К. Ш. Зигангирова. – М.: Радио и связь, 1982. – 526 с.
17. *Мизин И. А., Богатырев В. А., Кулешов А. П.* Сети коммутации пакетов. – М.: Радио и связь, 1986. – 408 с.
18. *Якубайтис Э. А.* Открытые информационные сети. – М.: Радио и связь, 1991. – 208 с.

19. Шварц М. Сети связи: протоколы, моделирование и анализ. – М.: Наука, 1992. – Ч.1. – 336 с.
20. Шварц М. Сети связи: протоколы, моделирование и анализ. – М.: Наука, 1992. – Ч.2. – 272 с.
21. Протоколы и методы управления в сетях передачи данных: Пер. с англ. / Под ред. Ф. Ф. Куо. – М.: Радио и связь, 1985. – 480 с.
22. Проектирование и техническая эксплуатация сетей передачи дискретных сообщений / Под ред. Г. П. Захарова. – М.: Радио и связь, 1988. – 360 с.
23. Мартин Дж. Вычислительные сети и распределенная обработка данных. – М.: Финансы и статистика, 1985. – Вып. 1. – 256 с.
24. Мартин Дж. Вычислительные сети и распределенная обработка данных. – М.: Финансы и статистика, 1986. – Вып. 2. – 269 с.
25. Протоколы информационно-вычислительных сетей / Под ред. И. А. Мизина и А. П. Кулешова. – М.: Радио и связь, 1990. – 504 с.
26. Стеклов В. К., Беркман Л. Н. Телекомунікаційні мережі: Підруч. для студ. вищ. навч. закл. за напрямом «Телекомунікації». – К.: Техніка, 2001. – 392 с.
27. Конвей Р. В., Максвелл В. Л., Миллер Л. В. Теория расписаний. – М.: Наука, 1975. – 395 с.
28. Бертсекас А., Галлагер Р. Сети передачи данных / Под ред. Б. С. Цыбакова. – М.: Мир, 1989. – 544 с.
29. Кокс Д. Р., Смит В. Л. Теория восстановления / Под ред. Ю. К. Беляева. – М.: Сов. радио, 1967. – 299 с.
30. Скрипченко С. І. Математична модель дослідження півдуплексного режиму обміну інформацією мережами передачі даних / Реєстрація, зберігання і обробка даних. – Т. 5. – № 2. – 2003. – С. 18–25.
31. Вінницький В. П., Поліщук В. Г. Метод оптимізації процедур інформаційної підтримки адаптивної маршрутизації. / Реєстрація, зберігання і обробка даних. – Т. 5. – № 2. – 2003. – С. 51–65.
32. Виницкий В. П., Хиленко В. В. Методы системного анализа и автоматизации проектирования телекоммуникационных сетей. – К.: Интерлинк, 2002. – 192 с.

Зміст

| | |
|---|-----------|
| Вступ | 3 |
| Частина I. Телекомунікаційні системи телефонного та вокодерного зв'язку. Термінальне устаткування неперервної дії..... | 7 |
| 1. Основні положення фізичної акустики | 7 |
| 1.1. Звук і його основні властивості..... | 7 |
| 1.2. Енергетична характеристика звуку | 9 |
| 1.3. Часові та спектральні характеристики звуків. Класифікація звуків | 12 |
| 1.4. Рівні інтенсивності звуку та звукового тиску | 15 |
| 1.4.1. Загальні положення | 15 |
| 1.4.2. Рівні спектрів складних звуків | 16 |
| 1.4.3. Загальний рівень звуку, що виникає від дії декількох джерел | 17 |
| 1.4.4. Шуми та їх характеристики..... | 19 |
| 2. Мовний сигнал і його сприйняття | 22 |
| 2.1. Механізм утворення мовних звуків та їх основні властивості..... | 22 |
| 2.2. Інтегральні характеристики мови | 25 |
| 2.3. Механізм слухового сприйняття | 29 |
| 2.4. Основні властивості та характеристики слуху..... | 32 |
| 2.4.1. Поріг чутності та больовий поріг | 32 |
| 2.4.2. Рівень відчуття | 33 |
| 2.4.3. Критична смуга слуху | 34 |
| 2.4.4. Ефективний рівень відчуття..... | 35 |
| 2.4.5. Гучність та адаптація слуху | 35 |
| 2.4.6. Маскування звуків | 36 |

| | |
|--|-----------|
| 3. Перетворення акустичної енергії мови в електричну та навпаки | 39 |
| 3.1. Призначення і класифікація перетворювачів | 39 |
| 3.2. Загальний принцип будови і дії обернених перетворювачів | 39 |
| 3.3. Основні системи перетворювачів..... | 41 |
| 3.3.1. Перетворювачі електромагнітної системи | 41 |
| 3.3.2. Перетворювачі електродинамічної системи | 42 |
| 3.3.3. Перетворювачі конденсаторної системи | 43 |
| 3.3.4. Перетворювачі п'єзоелектричної системи | 44 |
| 3.3.5. Вугільні перетворювачі | 45 |
| 3.4. Експлуатаційні характеристики основних пристроїв термінального устаткування для передавання та приймання мовних звуків..... | 46 |
| 3.4.1. Мікрофони | 46 |
| 3.4.2. Телефони | 49 |
| 3.4.3. Лорингофони..... | 51 |
| 3.4.4. Гучномовець | 52 |
| 4. Телефонні апарати..... | 54 |
| 4.1. Загальні положення. Класифікація телефонних апаратів та їх схем | 54 |
| 4.2. Основні елементи схем телефонних апаратів | 59 |
| 4.2.1. Номеронабирач | 60 |
| 4.2.2. Індуктор | 63 |
| 4.2.3. Поляризований дзвінок змінного струму | 63 |
| 4.3. Підвищення зручності користування телефонним апаратом | 64 |
| 4.4. Параметри і характеристики телефонних апаратів..... | 67 |
| 4.4.1. Електроакустичні параметри..... | 67 |
| 4.4.2. Електричні параметри..... | 69 |

| | |
|--|------------|
| 5. Оцінювання якості передавання мови | 71 |
| 5.1. Загальні положення. Вимоги до якості передавання | 71 |
| 5.2. Метод порівняння за гучністю..... | 73 |
| 5.3. Метод артикуляції..... | 75 |
| 5.4. Метод вимірювання еквівалента загасання розбірливості мови AEN..... | 80 |
| 5.5. Оцінювання натуральності передавання мови | 81 |
| 5.6. Електроакустичний розрахунок мовного тракту..... | 82 |
| 5.6.1. Постановка завдання і вихідні дані | 82 |
| 5.6.2. Загальний метод розрахунку розбірливості за параметрами тракту..... | 85 |
| 5.6.3. Особливість розрахунку розбірливості мови в умовах сильного шуму..... | 91 |
| 5.7. Методика розрахунку розбірливості за параметрами тракту (послідовність виконання розрахунків) | 93 |
| 5.7.1. Постановка завдання..... | 93 |
| 5.7.2. Вихідні дані | 93 |
| 5.7.3. Послідовність розрахунку та розрахункові вирази | 94 |
| 6. Методи перетворення мовного сигналу | 96 |
| 6.1. Загальні відомості та види перетворення сигналу | 96 |
| 6.1.1. Трансформація | 100 |
| 6.1.2. Компандування | 101 |
| 6.2. Частотне компандування мови. Основні відомості та загальний принцип дії вокодера | 102 |
| 6.3. Принципи амплітудного та часового перетворення мови | 108 |
| 6.4. Перетворення мовного сигналу в дискретну форму..... | 109 |
| 6.4.1. Амплітудно-імпульсна модуляція..... | 110 |
| <i>Контрольні запитання.....</i> | <i>114</i> |

Частина II. Термінальне устаткування дискретної інформації каналами телекомунікаційних мереж.....117

| | |
|---|------------|
| 7. Передавання дискретної інформації каналами телекомунікаційних мереж | 117 |
| 7.1. Основні поняття та визначення | 117 |
| 7.2. Перетворення повідомлень у сигнали | 119 |
| 7.3. Кодування джерелом інформації | 122 |
| 7.4. Канали передавання дискретної інформації | 124 |
| 7.5. Вимоги до системи передавання дискретної інформації | 128 |
| 7.6. Математичний опис загального алгоритму роботи системи передавання дискретної інформації | 129 |
| 7.7. Основні характеристики каналів зв'язку та їх вплив на передавання даних..... | 131 |
| 7.7.1. Основні характеристики каналів зв'язку | 131 |
| 7.7.2. Спотворення дискретних сигналів у каналах зв'язку | 132 |
| 7.7.3. Характеристичні крайові спотворення..... | 134 |
| 8. Методи перетворення дискретних сигналів для передавання каналами зв'язку – методи модуляції | 136 |
| 8.1. Загальний підхід до опису сигналів..... | 136 |
| 8.2. Амплітудна модуляція. Спектр модульованого сигналу за амплітудної модуляції..... | 142 |
| 8.3. Частотна модуляція. Спектр модульованого сигналу за частотної модуляції | 153 |
| 8.4. Фазова модуляція. Спектр модульованого сигналу за фазової модуляції..... | 157 |
| 8.5. Багатопозиційні методи модуляції | 164 |

| | |
|---|------------|
| 8.6. Квадратурна амплітудна модуляція..... | 168 |
| 9. Методи реєстрації дискретних сигналів | 172 |
| 9.1. Реєстрація сигналів методом стробування | 173 |
| 9.2. Реєстрація сигналів інтегральним методом | 176 |
| 9.3. Аналоговий інтегральний метод реєстрації..... | 177 |
| 9.4. Дискретний інтегральний метод реєстрації..... | 178 |
| 10. Методи синхронізації та фазування в системі передавання дискретної інформації..... | 182 |
| 10.1. Потреба в синхронізації та фазуванні..... | 182 |
| 10.2. Поелементна синхронізація..... | 183 |
| 10.3. Способи підтримки синхронізму | 188 |
| 10.4. Основні вимоги до пристроїв синхронізації | 190 |
| 10.5. Синхронізація безпосереднім діянням на частоту задавального генератора..... | 191 |
| 10.6. Синхронізація з посереднім діянням на частоту задавального генератора | 194 |
| 10.7. Системний аналіз окремих процесів перетворення сигналів та їх взаємодія в єдиній системі передавання дискретної інформації..... | 201 |
| 11. Моделювання передавання дискретної інформації дискретним каналом | 209 |
| 11.1. Загальні положення..... | 209 |
| 11.2. Узагальнена модель дискретного каналу | 211 |
| 11.3. Аналітичний статистичний аналіз експериментальних даних | 215 |
| 11.3.1. Поняття та завдання математичної статистики | 215 |
| 11.3.2. Побудова варіаційного ряду | 217 |
| 11.3.3. Знаходження статистичних характеристик..... | 219 |
| 11.3.4. Поняття про оцінювання параметрів..... | 221 |
| 11.3.5. Перевірка статистичної гіпотези..... | 226 |

| | |
|--|------------|
| 11.3.6. Принципи побудови моделі дискретного каналу | 230 |
| 11.3.7. Методика аналізу експериментальних даних – зареєстрованих помилок та установлення їх закономірності виникнення | 237 |
| 12. Телекомунікаційні системи передавання зображень..... | 245 |
| 12.1. Основні принципи побудови систем факсимільного зв'язку | 245 |
| 12.2. Принцип передавання зображень | 247 |
| 12.3. Структура факсимільних апаратів і системи факсимільного зв'язку | 249 |
| 12.4. Модеми сучасних факсимільних апаратів | 252 |
| 12.5. Кодування, передавання інформації і скорочення надмірності..... | 255 |
| 12.5.1. Факсимільний сигнал..... | 255 |
| 12.5.2. Передавання інформації і скорочення надмірності..... | 259 |
| <i>Контрольні питання</i> | <i>261</i> |

Частина III. Системний аналіз, надійність, завадостійкість телекомунікаційної системи під час керування обміном інформацією

| | |
|---|------------|
| 13. Загальні принципи керування дискретним (інформаційним) каналом..... | 263 |
| 13.1. Багаторівневий підхід. Протокол. Інтерфейс. Стек протоколів | 263 |
| 13.2. Модель відкритих систем OSI..... | 268 |
| 13.3. Керування інформаційним каналом | 271 |
| 13.4. Методи керування передаванням даних канального рівня (інформаційного каналу). Протоколи керування | 273 |
| 13.4.1. Асинхронні протоколи | 274 |

| | |
|---|------------|
| 13.4.2. Синхронні байтзорієнтовані та бітзорієнтовані протоколи | 274 |
| 13.4.3. Байтзорієнтований протокол двійкового синхронного зв'язку BSC | 275 |
| 13.4.4. Бітзорієнтований протокол двійкового синхронного зв'язку HDLC..... | 281 |
| 14. Методи підвищення достовірності передавання дискретної інформації..... | 302 |
| 14.1. Класифікація методів | 302 |
| 14.2. Надмірне кодування дискретної інформації | 303 |
| 14.3. Принципи виявлення і виправлення помилок надлишковими кодами – завадостійким кодуванням | 306 |
| 14.4. Загальна характеристика завдання завадостійкого кодування | 310 |
| 14.5. Класифікація завадостійкості кодів | 313 |
| 14.6. Лінійні (систематичні) коди. Коди Хеммінга | 315 |
| 14.7. Властивості та завдання побудови лінійних (систематичних) кодів | 317 |
| 14.8. Побудова лінійного систематичного коду | 324 |
| 14.9. Циклічні коди..... | 325 |
| 14.10. Неперервне згорткове кодування..... | 331 |
| 14.10.1. Згорткове кодування із синдромним коригуванням..... | 337 |
| 14.10.2. Згорткові коди з використанням послідовного декодування | 339 |
| 15. Надійність функціонування системи передавання дискретної інформації..... | 341 |
| 15.1. Надійність систем передавання дискретної інформації | 341 |
| 15.2. Класифікація відмов | 344 |
| 15.3. Характеристики надійності трактів передавання даних | 345 |

| | |
|---|------------|
| 15.4. Розрахунок надійності каналів передавання даних | 349 |
| 15.5. Розрахунок надійності тракту передавання дискретної інформації | 356 |
| 15.6. Вплив ненадійності трактів на ефективність передавання інформації | 365 |
| 15.6.1. Загальні положення..... | 365 |
| 15.6.2. Розподіл часу обслуговування у разі пасивних переривань..... | 367 |
| 15.6.3. Розподіл часу обслуговування у разі активних переривань..... | 368 |
| 15.6.4. Розподіл часу обслуговування у разі активних і пасивних переривань | 370 |
| 16. Елементи системного аналізу передавання даних системою передавання дискретної інформації | 372 |
| 16.1. Основні положення системного аналізу | 372 |
| 16.2. Моделювання дуплексного режиму роботи системи передавання дискретної інформації..... | 374 |
| 16.3. Моделювання півдуплексного режиму системи передавання дискретної інформації | 387 |
| 16.3.1. Побудова математичної моделі | 387 |
| 16.3.2. Дослідження моделі методом укладених ланцюгів Маркова..... | 389 |
| 16.4. Затримки повідомлень каналами системи передавання дискретної інформації | 393 |
| 16.4.1. Загальні положення. Простір станів системи..... | 393 |
| 16.4.2. Період зайнятості системи | 395 |
| 16.4.3. Розподіл часу затримки повідомлень | 400 |
| 16.5. Метод оптимізації процедур інформаційної підтримки адаптивної маршрутизації..... | 408 |
| 16.5.1. Постановка завдання | 408 |
| 16.5.2. Модель процесу передавання повідомлення системою маршрутизатор – канал зв'язку..... | 410 |

| | |
|---|------------|
| 16.5.3. Вплив періоду опитування на затримку інформації про стан M_j маршрутизатора..... | 417 |
| 16.5.4. Урахування ненадійності та відновлення каналів зв'язку | 420 |
| <i>Контрольні питання.....</i> | <i>422</i> |
| Список рекомендованої літератури | 425 |

Навчальне видання

Вінницький Владислав Петрович

Поліщук Василь Григорович

Термінальне устаткування
та передавання інформації
в телекомунікаційних системах

*За загальною редакцією
члена-кореспондента НАН України М. Ю. Льченка*

Підручник

Редактори: *Л. М. Костенко, Р. М. Шульженко*

Коректор *Т. В. Рибитва*

Комп'ютерна верстка *М. Є. Пігурнов*

Дизайн обкладинки *О. С. Єременко*

Темплан 2004 р., поз. 1-1-006

Оригінал-макет підготовлено ІВЦ "Видавництво «Політехніка»"

Підп. до друку 26.11.2004. Формат 60×84¹/₁₆. Папір офс. Гарнітура Times.
Спосіб друку – офсетний. Ум. друк. арк. 25,34. Обл.-вид. арк. 42,15. Наклад 1000 пр.

Інформаційно-видавничий центр "Видавництво «Політехніка»" НТУУ «КПІ»

Свідоцтво про держреєстрацію ДК № 1665 від 28.01.2004 р.

03056, Київ-56, вул. Політехнічна, 14, корп. 15,
тел./факс (044) 241-68-78, 241-66-64, e-mail: izdatelstvo@ntu-kpi.kiev.ua

Друк СПД «Кіслова В.М.»

зам. 4-35