

МІНІСТЕРСТВО ОСВІТИ І НАУКИ
МОЛОДІ ТА СПОРТУ УКРАЇНИ
ДЕРЖАВНИЙ ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД
«НАЦІОНАЛЬНИЙ ГІРНИЧИЙ УНІВЕРСИТЕТ»



В.Л. Кожевников
А.В. Кожевников

ТЕОРІЯ ІНФОРМАЦІЇ ТА КОДУВАННЯ

Дніпропетровськ
НГУ
2012

УДК 621.391
ББК 32.811
К58

Рекомендовано редакційною радою НГУ як навчальний посібник для студентів вищих навчальних закладів, які навчаються за напрямом «Автоматизація та комп'ютерно-інтегровані технології» (протокол № 5 від 15.05.2012 р.).

Рецензенти:

І.В. Жуковицький, д-р техн. наук, професор, завідувач кафедри електронно-обчислювальних машин (Національний університет залізничного транспорту ім. академіка Лазаряна).

О.М. Ахметшин, д-р фіз.-мат. наук, професор кафедри автоматизованих систем обробки інформації (Дніпропетровський національний університет ім. Олеся Гончара)

Кожевников, В.Л.

К58 Теорія інформації та кодування [Текст]: навч. посібник / В.Л. Кожевников, А.В. Кожевников. – Д.: Національний гірничий університет, 2011. – 108 с.

Розглянуто основи статистичної теорії інформації, пропускну здатності каналів зв'язку, а також теорії кодування повідомлень в системах автоматики та передачі технологічної інформації.

Посібник укладений відповідно до програми дисципліни «Теорія інформації та кодування» для студентів, що навчаються за напрямом підготовки 050202 «Автоматизація та комп'ютерно-інтегровані технології», а також студентів інших спеціальностей, які вивчають цю дисципліну.

УДК 621.391
ББК 32.811

© В.Л. Кожевников, А.В. Кожевников, 2012
© Державний ВНЗ «Національний гірничий університет», 2012

ПЕРЕДМОВА

Дисципліна «Теорія інформації та кодування» належить до групи професійно орієнтованих і займає важливе місце у підготовці бакалаврів за напрямом 050202 «Автоматизація та комп'ютерно-інтегровані технології». Завдання предмета – формування знань з основи статистичної теорії інформації, пропускну здатності каналів зв'язку, а також теорії кодування повідомлень в системах автоматики та передачі технологічної інформації, набуття навичок визначення інформативності дискретних та безперервних джерел повідомлень, пропускну здатності каналів зв'язку та узгодження характеристик сигналів і каналів, синтезу оптимальних, елементарних завадостійких, лінійних, циклічних та каналних кодів. Навчальний матеріал дисципліни вимагає від студентів знань з вищої математики, алгоритмічних мов й програмування, основ електротехніки та електроніки, а також допомагає вивчати предмети: основи збирання, обробки і передачі інформації, комп'ютерні мережі та дипломне проектування.

Робота відповідає лекційному курсу дисципліни, який впродовж низки років викладався авторами у Національному гірничому університеті. Посібник може бути використаний для студентів заочно-дистанційної і екстернатної форми навчання. З цією метою під час викладення теоретичного матеріалу наводяться приклади розв'язання задач, наприкінці кожного розділу містяться контрольні питання, у додатках – підсумкові тести.

Посібник розроблений викладачами Національного гірничого університету В.Л. Кожевниковим, канд. техн. наук, професором кафедри автоматизації і комп'ютерних систем, А.В. Кожевниковим, канд. техн. наук, доцентом кафедри автоматизації і комп'ютерних систем. Він складається з двох розділів: перший написаний А.В. Кожевниковим, другий – В.Л. Кожевниковим.

Враховуючи навчальне призначення, у посібнику не робиться бібліографічних посилань на запозичення з відомих джерел. Список літератури містить тільки вказівки на матеріали, що рекомендуються як додаткові для поглиблення знань та розширення навичок у галузі теорії інформації і кодування.

Посібник створений для бакалаврів, що навчаються за напрямом 050202 «Автоматизація та комп'ютерно-інтегровані технології», а також для спеціалістів і магістрів, які продовжують навчання на основі означеного напрямка, зокрема 05020201 «Автоматизоване управління технологічними процесами».

Написання навчального посібника досить клопітка робота, при виконанні якої неминучі помилки. Автори з вдячністю отримують зауваження та побажання, які сприятимуть подальшому покращенню змісту книги.

ВСТУП

Існування й розвиток різноманітних систем – технічних, біологічних, соціальних та інших – тією чи іншою мірою визначається процесами обміну між частинами цих систем матерією, енергією й інформацією.

Остання проявляє себе в умовах, коли причини й наслідки процесів мають імовірнісні зв'язки. Якщо процеси детерміновані, інформаційна складова в них вважається відсутньою. Такий підхід до визначення кількості інформації має назву статистичного. При цьому загальні уявлення про інформацію та її цінність відсутні – до уваги приймається лише кількісний бік, що відображає невизначеність наслідків процесів (дослідів). Варто зауважити, що в інформаційних системах технологічного призначення поняття цінності інформації не використовується. Натомість вводяться пріоритети повідомлень, які базуються на техніко-економічних критеріях обслуговування заявок, що формуються в інформаційній системі.

Надалі розглядаються основи теорії інформації щодо задач збору, перетворення, передачі, зберігання й надання інформації в технологічних системах гірничо-металургійні виробництва, енергосистеми, телекомунікації. При вирішенні задач виділяються три основні частини інформаційних підсистем: джерела повідомлень, канали зв'язку й одержувачі повідомлень. У загальному випадку і підсистеми охоплюють як людські, так і машинні складові – джерелами і одержувачами повідомлень можуть бути оператори та (або) керуючі ЕОМ.

Далі для систем, що розглядаються, під терміном «інформація» буде матися на увазі та частина повідомлення, яка заздалегідь невідома одержувачеві.

Отже, необхідними умовами появи інформації при статистичному підході до визначення її кількості є наявність досліду та невизначеність його наслідків. Переваги цього підходу полягають в об'єктивності. Кількість інформації визначається на основі аналізу ймовірнісних характеристик джерела інформації і не залежить від оцінюючого та універсальності. Процедура визначення кількості інформації та одиниці виміру не залежать від природи її джерела.

Важливою складовою процесів обробки, передачі й зберігання інформації є кодування повідомлень. Метою кодування за відсутності перешкод є подання повідомлень в більш компактному вигляді або стиснення інформації, а за їх наявності – виявлення та виправлення помилок, що виникають в повідомленнях під впливом перешкод, тобто забезпечення достовірності інформації.

Основні положення статистичної теорії інформації були створені Клодом Шенноном до 1948 р. Їм же були розроблені основи кодування повідомлень в каналах зв'язку за відсутності перешкод та їх наявності. Бурхливий розвиток теорії кодування у другій половині ХХ століття призвів до формування її як окремої науки. Сучасна теорія інформації і кодування дозволяє вирішувати задачі збору, перетворення, передачі, зберігання й надання інформації найбільш економічними й ефективними способами.

1. ОСНОВИ СТАТИСТИЧНОЇ ТЕОРІЇ ІНФОРМАЦІЇ

Навчальні цілі: формування знань щодо основ статистичної теорії інформації та набуття навичок визначення інформативності дискретних та безперервних джерел повідомлень, пропускної здатності каналів зв'язку та узгодження характеристик сигналів і каналів.

1.1. Кількість інформації й ентропія

Поняття ентропії, як міри вихідної (апріорної) невизначеності результатів дослідження, було впроваджено Шенноном за аналогією з термодинамічною ентропією.

На основі цього кількість інформації, що одержують у результаті проведення дослідження, визначається як

$$I = H - H_o, \quad (1.1)$$

де H – апріорна ентропія, H_o – залишкова ентропія (невизначеність наслідків) дослідження.

У цьому поданні кількість інформації, отримана в результаті дослідження, – це міра знятої невизначеності.

При передачі повідомлень у каналах без перешкод $H_o = 0$ і $I = H$.

У додатку до систем зв'язку термін «кількість інформації» увів Ральф Хартлі в 1928 р. Він запропонував визначати цю величину як

$$I = -\log p = -\log \frac{1}{N} = \log N, \quad (1.2)$$

де p – ймовірність появи одного з N можливих рівноймовірних наслідків дослідження.

Логарифмічна міра кількості інформації може бути обґрунтована такими міркуваннями:

1. Кількість інформації більше в досліді з більшим числом рівноймовірних наслідків:

$$N_1 > N_2 \Rightarrow I(N_1) > I(N_2). \quad (1.3)$$

2. Якщо дослід має один результат, $N = 1$, то

$$I(N = 1) = 0. \quad (1.4)$$

3. Кількість інформації в досліді, що складається з окремих етапів, в яких мають місце взаємозалежні наслідки $N_1, N_2 \dots$

$$I(N_1, N_2, \dots) = I(N_1) + I(N_2) + \dots \quad (1.5)$$

Тоді в загальному випадку кількість інформації визначиться як

$$I = -a \cdot \log_b p. \quad (1.6)$$

Постійні a й b встановлюють масштаб величини I . По Хартлі $a = 1$ $b = 2$. У цьому випадку, якщо

$$p = \frac{1}{2}, \text{ те } I = -\log_2 \frac{1}{2} = \log_2 2 = 1 \text{ біт.} \quad (1.7)$$

Ця одиниця кількості інформації використовується, як правило, у системах зв'язку. Відомі також аналогічні одиниці виміру нат ($b = e \approx 2,73\dots$) і діт ($b = 10$). Надалі, якщо значення основи $b = 2$, воно у функції логарифма не вказується.

1.2. Ентропія дослідів з нерівноймовірними взаємозалежними наслідками

Припустимо, що дослід – це генерація дискретним джерелом випадкових послідовностей елементів (повідомлень) зі значеннями $1, 2, 3, \dots, m$. Число відмінних елементів називають основою алфавіту джерела. Статистичні властивості джерела незмінні, тому ймовірності появи елементів $p_1, p_2, p_3, \dots, p_m$ у послідовностях досить великої довжини n також можна вважати незмінними.

Відповідно до теореми про множення ймовірностей для дослідів з випадковими нерівноймовірними й взаємозалежними наслідками середня ймовірність появи будь-якої послідовності сигналів джерела довжиною $n \rightarrow \infty$

$$p_c = p_1^{n_1} \cdot p_2^{n_2} \cdot p_3^{n_3} \cdot \dots \cdot p_m^{n_m} = \prod_{i=1}^m p_i^{n_i}, \quad (1.8)$$

де $n_1, n_2, n_3, \dots, n_m$ – числа елементів зі значеннями $1, 2, 3, \dots, m$ в послідовності довжиною $n = n_1 + n_2 + n_3 + \dots + n_m$. При цьому

$$p_1 = \frac{n_1}{n}, p_2 = \frac{n_2}{n}, p_3 = \frac{n_3}{n}, \dots, p_m = \frac{n_m}{n}, n \rightarrow \infty. \quad (1.9)$$

Зазначимо, що кількість можливих послідовностей N в цьому випадку визначається як усі розміщення із повтореннями з m по n , тобто $N = m^n$.

Тому що для даного джерела $p_c = \text{const}$, до нього можливо застосувати вираз у вигляді (1.2)

$$H_c = -\log p_c = -\log \prod_{i=1}^m p_i^{n_i} = -\log \prod_{i=1}^m p_i^{n \cdot p_i} = -n \cdot \sum_{i=1}^m p_i \cdot \log p_i, \text{ біт/послід. (1.10)}$$

Величина

$$H = \frac{H_c}{n} = -\sum_{i=1}^m p_i \cdot \log p_i, \text{ біт/елемент. (1.11)}$$

називається середньою питомою ентропією багатоелементного джерела.

У розрахунках інформаційних характеристик систем широко використовується поняття часткової ентропії

$$H_i = -p_i \cdot \log p_i, \text{ біт/результат. (1.12)}$$

При реалізації конкретного результату досліду кількість інформації становить

$$I_i = H_{\Pi i} = \log p_i, \text{ біт/результат, (1.13)}$$

а $H_{\Pi i}$ називають повною ентропією на « i »-й результат досліду.

1.3. Дослідження властивостей часткової ентропії

Якщо « i »-й результат досліду детермінований, то $p_i = 1$ і згідно (1.12)

$$H_{i1} = -1 \cdot \log 1 = 0. (1.14)$$

У випадку, коли $p_i = 0$ (результат не реалізуємо) $H_{i0} = -0 \cdot \log 0$ – часткова ентропія безпосередньо невизначена.

Заміна змінних для рівноймовірних наслідків досліду $p_i = \frac{1}{k}$ й розкриття невизначеності за правилом Лопіталя дає результат

$$\lim_{p_{i0} \rightarrow 0} H_{i0} = \lim_{k \rightarrow \infty} \left(\frac{1}{k} \cdot \log k \right) = \lim_{k \rightarrow \infty} \left(\frac{\frac{1}{k} \cdot \log e}{1} \right) = 0. (1.15)$$

Похідна функції ентропії має вигляд

$$\begin{aligned} \frac{dH_i}{dp_i} &= \frac{d}{dp_i}(-p_i \log p_i) = -\left(1 \cdot \log p_i + p_i \frac{1}{p_i} \log e\right) = \\ &= -\log p_i - \log e = -\log(e \cdot p_i). \end{aligned} \quad (1.16)$$

Оскільки

$$\begin{cases} dH_i / dp_i = 0 \text{ при } p_i = e \\ dH_i / dp_i > 0 \text{ при } 0 < p_i < \frac{1}{e} \\ dH_i / dp_i < 0 \text{ при } \frac{1}{e} < p_i < 1, \end{cases} \quad (1.17)$$

часткова ентропія має один максимум, при $p_i = \frac{1}{e} \approx 0,36$, а її значення складає

$$H_{i \max} = -\frac{1}{e} \log \frac{1}{e} = 0,5 \text{ біт/результат.}$$

Графік залежності $H_i = f(p_i)$ наведений на рис. 1.1.

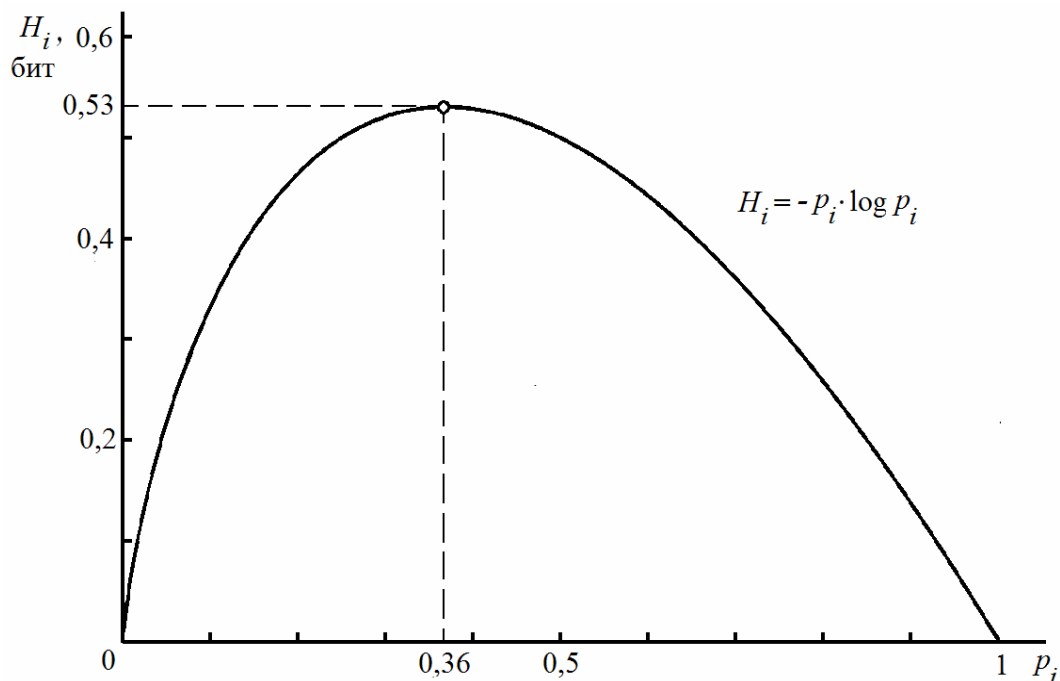


Рис. 1.1

При виконанні розрахунків інформаційних характеристик джерел дискретних повідомлень зручно користуватися табульованими значеннями цієї функції, які наведені в додатку 1.

1.4. Дослідження ентропії двійкового джерела. Умова максимуму ентропії дискретного джерела при $m > 2$

Приклад визначення ентропії за формулою Шеннона (1.11) покажемо для найпростішого досліду на основі джерела повідомлень, який створює потік (послідовність) із двох можливих наслідків – повідомлень 0,1, що з'являються випадково й незалежно один від одного з безумовними ймовірностями p_0, p_1 .

Середнє значення ентропії на повідомлення розглянутого джерела становить

$$H = -p_0 \log p_0 - p_1 \log p_1 = -p_0 \log p_0 - (1 - p_0) \log(1 - p_0). \quad (1.18)$$

Сума безумовних імовірностей $p_1 + p_0 = 1$. При граничних значеннях $p_1 = 0$ ($p_0 = 1 - 0 = 1$) і $p_1 = 1$ ($p_0 = 1 - 1 = 0$), у відповідності із властивостями часткової ентропії, ентропія джерела дорівнює нулю. Похідна функції ентропії має вигляд

$$\begin{aligned} \frac{dH}{dp_0} &= - \left[1 \cdot \log p_0 + p_0 \cdot \frac{1}{p_0} \cdot \log e - 1 \cdot \log(1 - p_0) - (1 - p_0) \cdot \frac{1}{1 - p_0} \cdot \log e \right] = \\ &= - [\log p_0 - \log(1 - p_0)]. \end{aligned} \quad (1.19)$$

Оскільки

$$\begin{cases} dH / dp_0 = 0 \text{ при } p_0 = 0.5 \\ dH / dp_0 > 0 \text{ при } 0 < p_0 < 0.5 \\ dH / dp_0 < 0 \text{ при } 0.5 < p_0 < 1, \end{cases} \quad (1.20)$$

ентропія двійкового джерела досягає максимуму, коли воно створює рівноймовірні повідомлення.

Максимальна величина ентропії становить

$$H_{\max} = -\frac{1}{2} \cdot \log \frac{1}{2} - \left(1 - \frac{1}{2}\right) \cdot \log \left(1 - \frac{1}{2}\right) = 1 \text{ біт/повідомлення.} \quad (1.21)$$

Середню ентропію джерела можна зобразити також суму часткових ентропій H_0 і H_1 .

Отримані результати зображені в графічній формі на рис. 1.2.

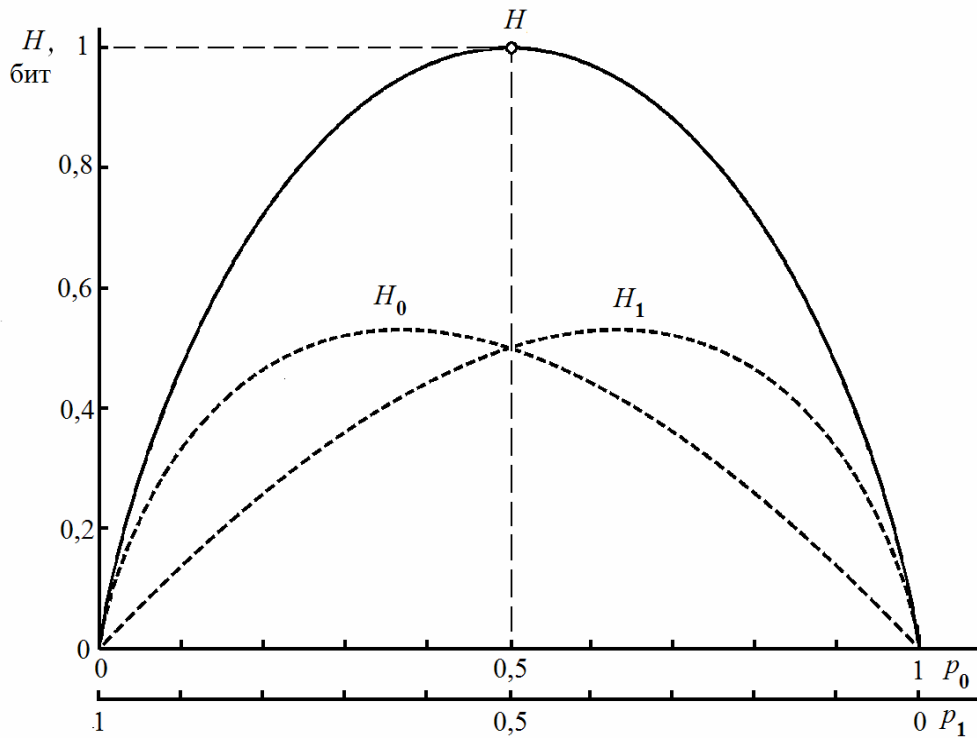


Рис. 1.2

Якщо $m > 2$, максимум ентропії джерела також буде мати місце за умови того, що вироблені їм повідомлення рівноймовірні й взаємозалежні. Тоді $p_i = \frac{1}{m} = const$ й згідно (1.11)

$$H_{\max} = -m \cdot \frac{1}{m} \cdot \log \frac{1}{m} = \log m, \text{ біт/сигнал.} \quad (1.22)$$

Справедливість (1.22) доводиться на основі методу невизначених множників Лагранжа. При цьому розв'язується задача оптимізації цільової функції

$$W(p_i) = -\sum_{i=1}^m p_i \log p_i, \quad (1.23)$$

при обмеженні $\sum_{i=1}^m p_i - 1 = 0$. Функція Лагранжа та її часткові похідні для цієї задачі мають вигляд

$$L(p_i, \lambda) = -\sum_{i=1}^m p_i \log p_i + \lambda \left(\sum_{i=1}^m p_i - 1 \right);$$

$$\frac{\partial L(p_i, \lambda)}{\partial p_i} = -\log p_i - 1 + \lambda. \quad (1.24)$$

Оптимум цільової функції досягається, коли всі часткові похідні функції Лагранжа дорівнюють нулю, що можливо тільки за умови рівності значень $p_i = \frac{1}{m} = const$. Про те, що оптимум є максимумом, свідчать умови (1.20) для випадка $m=2$.

1.5. Ентропія типових дискретних розподілів

Розв'язування багатьох задач передачі інформації засновано на моделях дослідів, у яких імовірності наслідків мають аналітичні вираження. Найпростіша модель являє собою рівномірний розподіл – дослід з m рівноймовірними взаємозалежними наслідками.

Аналітичне вираження для рівномірного дискретного розподілу й середньої питомої ентропії джерела при рівноймовірно розподілених повідомленнях має вигляд

$$p_i = p(x_i = l) = \begin{cases} \frac{1}{m} & \text{при } 1 \leq x_i \leq m \\ 0 & \text{при } x_i < 1, x_i > m, \end{cases} \quad (1.25)$$

а його полігон зображений на рис. 1.3.

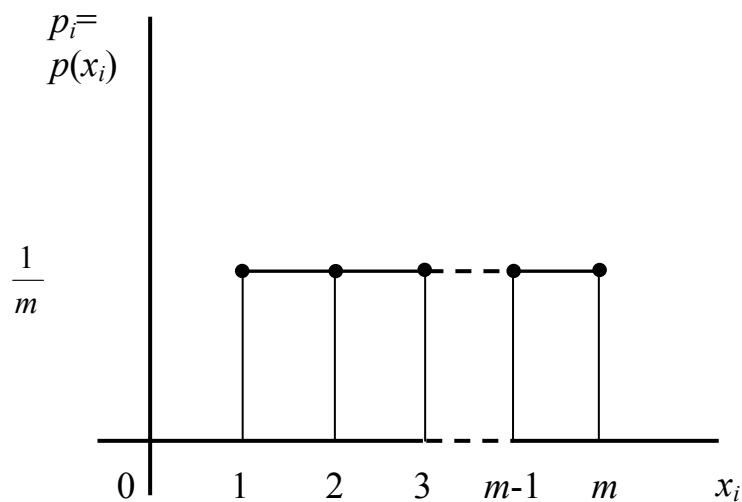


Рис. 1.3

При цьому середня питома ентропія джерела визначається за формулою (1.22).

Біноміальному розподілу, або розподілу Бернуллі підпорядковується випадкова величина, що визначає число l появ випадкової події в послідовності з m незалежних дослідів. Розподіл описується залежностями

$$p_i = p(x_i = l) = \begin{cases} C_m^l p^l (1-p)^{m-l} & \text{при } 0 \leq x_i \leq m \\ 0 & \text{при } x_i < 0, x_i > m. \end{cases} \quad (1.26)$$

Тут p – ймовірність появи випадкової події в одному досліді.
Полігон біноміального розподілу наведений на рис. 1.4.

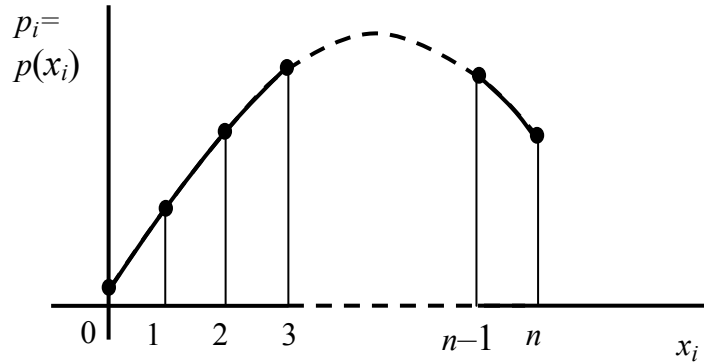


Рис. 1.4

Середня питома ентропія джерела обчислюється шляхом підстановки ймовірностей (1.26) у вираз (1.11) і визначається за формулою

$$\begin{aligned} H &= -\sum_{l=0}^m \left\{ C_m^l p^l (1-p)^{m-l} [\log C_m^l + l \log p + (m-l) \log(1-p)] \right\} = \\ &= -m[p \log p + (1-p) \log(1-p)] - \sum_{l=1}^{m-1} C_m^l p^l (1-p)^{m-l} \log C_m^l, \text{ біт/сигнал.} \end{aligned} \quad (1.27)$$

Задача визначення ймовірності появи помилок (трансформації елементів) кратності l у випадковій послідовності двійкових елементів довжиною n , де p – ймовірність трансформації елементів цієї послідовності, коли елемент 0 замінюється на 1, або навпаки, 1 переходить у 0. При цьому умовні ймовірності взаємної трансформації елементів рівні й становлять

$$P(0/1) = P(1/0) = p. \quad (1.28)$$

Коли ймовірність появи помилок мала, $p \rightarrow 0$, розподіл Бернуллі переходить у розподіл Пуассона, для якого вираження $p(x_i = l)$ й H приймають відповідно вигляд

$$p_i = p(x_i = l) = \begin{cases} \frac{\lambda^l e^{-\lambda}}{l!} & \text{при } x_i > 0 \\ 0 & \text{при } x_i < 0, \end{cases} \quad (1.29)$$

$$H = - \sum_{l=0}^{+\infty} \frac{\lambda^l e^{-\lambda}}{l!} (l \log \lambda - \log(l!) - \lambda) = \lambda \log \left(\frac{e}{\lambda} \right) + \sum_{l=0}^{+\infty} \frac{\lambda^l e^{-\lambda}}{l!} \log(l!), \text{ біт/сигнал, (1.30)}$$

де $\lambda = p \cdot n$ – інтенсивність, або середня кількість помилок при довжині послідовності n .

З інших типових розподілів варто назвати геометричний, гіпергеометричний та ін.

1.6. Ентропія типових неперервних розподілів

Повідомлення про стан багатьох процесів є неперервними (аналоговими) випадковими величинами, які в межах своєї зміни можуть приймати дійсні значення. Кількість таких значень нескінченно (утворює континуум). Джерела неперервних повідомлень x характеризуються функціями щільності ймовірності $w(x)$.

Для визначення ентропії в цьому випадку перейдемо від неперервного розподілу до дискретного, проквантувавши випадкову величину із кроком (інтервалом) Δx (рис. 1.5)

$$\Delta x = \frac{X_{\max} - X_{\min}}{m}, \quad (1.31)$$

де m – число дискретних значень x у межі зміни від X_{\min} до X_{\max} (деякі розподіли мають нескінченні межі зміни від $-\infty$ до $+\infty$).

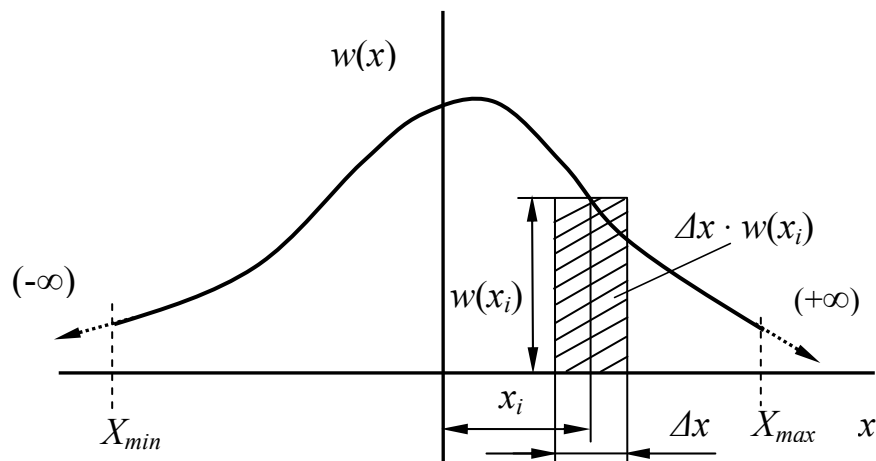


Рис. 1.5

Виберемо значення дискретної величини x_i на кроці i таким чином, щоб виконувалася умова

$$\int_{\Delta x_i} w(x) dx = w(x_i) \cdot \Delta x_i, \quad (1.32)$$

Ліва частина цього рівняння являє собою площу під кривою розподілу на інтервалі Δx_i (індекс i позначає лише номер інтервалу й $\Delta x_i = \Delta x = const$), а права – еквівалентну площу прямокутника з висотою $w(x_i)$. Ця площа дорівнює ймовірності влучення випадкової величини x в інтервал Δx_i :

$$p(x_i) = w(x_i) \Delta x. \quad (1.33)$$

Використання замість неперервного подання імовірнісного розподілу випадкової величини дискретного дозволяє в подальших перетвореннях використовувати вираження (1.11) для середньої ентропії. У припущенні, що функція $w(x)$ визначена на всій числовій осі, величина ентропії визначається як

$$H = - \sum_{i=-\infty}^{\infty} w(x_i) \cdot \Delta x \cdot \log[w(x_i) \cdot \Delta x] = - \sum_{i=-\infty}^{\infty} w(x_i) \cdot \log w(x_i) \cdot \Delta x - \sum_{i=-\infty}^{\infty} \log \Delta x \cdot w(x_i) \cdot \Delta x. \quad (1.34)$$

Перехід до межі при $\Delta x \rightarrow 0$ дає вираз

$$\lim_{\Delta x \rightarrow 0} H = - \int_{-\infty}^{\infty} w(x) \cdot \log w(x) dx - \int_{-\infty}^{\infty} \log \Delta x \cdot w(x) dx. \quad (1.35)$$

Перший доданок називається диференціальною ентропією

$$H_d = - \int_{-\infty}^{\infty} w(x) \cdot \log w(x) \cdot dx \quad (1.36)$$

і повністю визначається виглядом функції $w(x)$.

Другий доданок – це міра невизначеності переходу від неперервної моделі до дискретного із квантом Δx , або залишкова ентропія,

$$H_0 = - \int_{-\infty}^{\infty} \log \Delta x \cdot w(x) \cdot dx = - \log \Delta x \int_{-\infty}^{\infty} w(x) dx = - \log \Delta x, \quad (1.37)$$

оскільки $\int_{-\infty}^{\infty} w(x) \cdot dx = 1$.

Те, що в (1.37) і далі присутня мала величина Δx , яка являє собою крок дискретизації функції $w(x)$, відбиває той факт, що в практичних додатках

неперервна величина завжди відображається в дискретному вигляді через обмежену точність виміру.

У такий спосіб

$$H = H_d - H_0 = H_d - \log(\Delta x) = H_d + \log(1/\Delta x). \quad (1.38)$$

Очевидно, що при $\Delta x \rightarrow 0$, $H_0 \rightarrow \infty$ і $H(x)$ також нескінченно велика.

Для рівномірного розподілу щільність імовірностей якого дається за допомоги виразу

$$w(x) = \frac{1}{X_{\max} - X_{\min}} \quad (1.39)$$

і зображена на рис. 1.6.

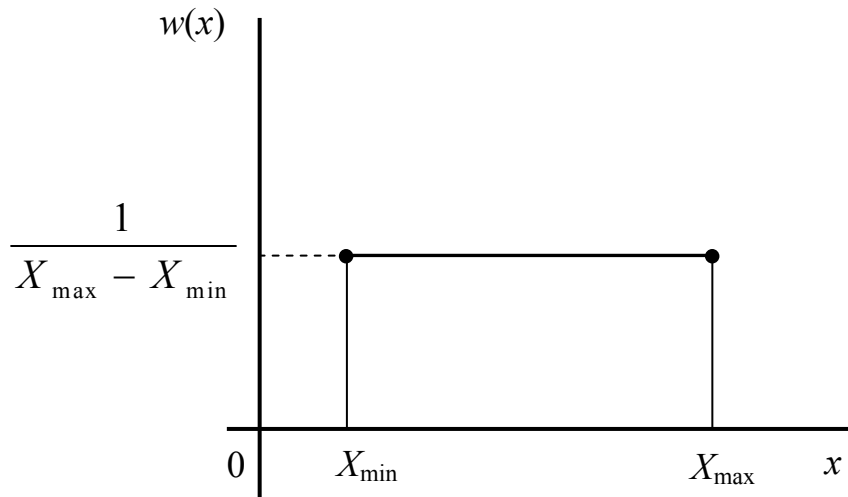


Рис. 1.6

Згідно (1.36), (1.38)

$$\begin{aligned} H_d &= - \int_{-\infty}^{\infty} \frac{1}{X_{\max} - X_{\min}} \cdot \log \frac{1}{X_{\max} - X_{\min}} \cdot dx = \\ &= - \frac{1}{X_{\max} - X_{\min}} \cdot \log \frac{1}{X_{\max} - X_{\min}} \cdot x \Big|_{X_{\min}}^{X_{\max}} = \log(X_{\max} - X_{\min}), \end{aligned} \quad (1.40)$$

$$H = \log(X_{\max} - X_{\min}) - \log \Delta x = \log \frac{X_{\max} - X_{\min}}{\Delta x}. \quad (1.41)$$

Найбільше широко на практиці поширений нормальний розподіл

$$w(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}, \quad (1.42)$$

де σ – середньоквадратичне відхилення випадкової величини x від математичного очікування a . Щільність імовірностей нормального розподілу зображена на рис. 1.7.

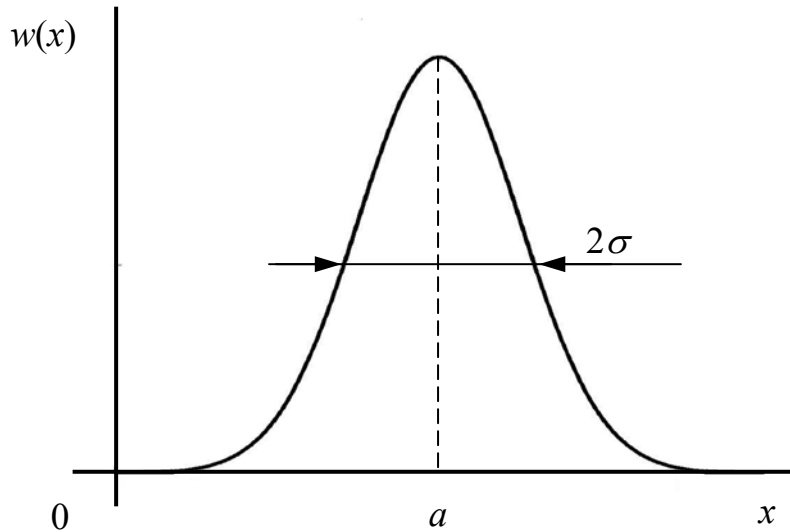


Рис. 1.7

Такому розподілу підпорядковуються випадкові помилки вимірів, що виникають внаслідок дії великої кількості незалежних факторів, кожний з яких має на помилку малий вплив. Нормальний розподіл характеризується диференціальною ентропією

$$H_d = \log(\sigma\sqrt{2\pi e}). \quad (1.43)$$

У теорії надійності часто застосовується експонентний розподіл

$$w(x) = \lambda e^{-\lambda x}, \quad (1.44)$$

де λ – постійна експонентного спаду випадкової величини x , яка стосовно задач надійності має сенс величини, оберненої до середнього терміну служби об'єкта.

Вираз для диференціальної ентропії експонентного розподілу має вигляд

$$H_d = -\frac{\lambda}{\ln 2} \int_0^{+\infty} e^{-\lambda x} (\ln \lambda - \lambda x) dx = \frac{e^{-\lambda x}}{\ln 2} (\ln \lambda - \lambda x - 1) \Big|_0^{+\infty} = \frac{\ln \frac{e}{\lambda}}{\ln 2} = \log \frac{e}{\lambda}. \quad (1.45)$$

До інших відомих неперервних розподілів відносяться Пірсона, Коші, трикутне (Сімпсона) та ін.

1.7. Ентропія ергодичних послідовностей

Якщо властивості джерела повідомлень, установлені в п.1.2, містять у собі взаємозалежність появи елементів у послідовностях цих сигналів, то таке джерело називають ергодичним.

Для нього ймовірність появи чергового елемента послідовності x_j

$$p(x_j / x_i^{(1)}, x_h^{(2)}, \dots, x_v^{(r)}, x_q^{(r+1)}) = p(x_j / x_i^{(1)}, x_h^{(2)}, \dots, x_v^{(r)}) \quad (1.46)$$

залежить тільки від того, які « r » елементів йому передують, а величина « r » називається порядком ергодичного джерела.

Якщо ймовірнісні властивості джерела незмінні, то послідовність, що воно виробляє, також називають ергодичною.

Черговий набір з r елементів визначає стан джерела S_k , з якого він при надходженні елемента x_j перейде в стан S_l .

Число можливих станів S_k ергодичного джерела – це кількість всіх можливих сполучень з m по r ,

$$N = m^r, \quad (1.47)$$

де m – основа алфавіту джерела.

Нехай двійкове ергодичне джерело характеризується величиною $r = 2$. Тоді можливе число станів джерела $N = 2^2 = 4$, а самі стани S_k – це двійкові комбінації: $S_1 - 00$, $S_2 - 01$, $S_3 - 10$, $S_4 - 11$.

У прикладі послідовності повідомлень джерела 01001110110... такі комбінації виділені на рис. 1.8.

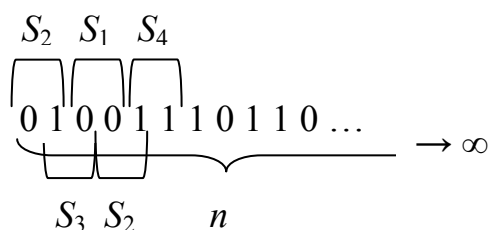


Рис. 1.8

Тут джерело послідовно переходить зі стану S_2 в S_3 , далі в S_1 , S_2 , S_4 і т.д.

Якщо розглядати ряд аналогічних послідовностей досить великої довжини (у загальному випадку залежної від величин m і r), то ймовірнісні

характеристики джерела – безумовні ймовірності $p(0)$, $p(1)$ і умовні ймовірності $p(x_j / x_i, x_h)$, $x_j, x_i, x_h \in \{0,1\}$ для будь-якої такої послідовності будуть тими самими.

Коли джерело перебуває в стані S_l , то є невизначеність, з якого стану S_k воно у нього перейшло. Міра цієї невизначеності – ентропія, що визначається згідно (1.11) та становить

$$H(S_k) = - \sum_{l \neq k} p(S_l / S_k) \cdot \log p(S_l / S_k). \quad (1.48)$$

Усреднюючи ентропію (1.48) по всіх станах S_k з імовірностями $p(S_k)$, одержимо середню ентропію ергодичного джерела

$$H = - \sum_{k=1}^N \sum_{l \neq k} p(S_k) \cdot p(S_l / S_k) \cdot \log p(S_l / S_k). \quad (1.49)$$

1.8. Ентропія найпростіших ергодичних джерел

1. $r=0$. Це значить, що елементи послідовності взаємонезалежні. Джерело має лише один стан S_0 з $p(S_0)=1$ і вертається в нього після надходження чергового елемента x_j . Тоді $p(S_0 / S_0) = p(x_j)$ й (1.49) перетвориться до вигляду

$$H = - \sum_{k=1}^1 \sum_{j=1}^m 1 \cdot p(x_j) \cdot \log p(x_j) = - \sum_{j=1}^m p(x_j) \cdot \log p(x_j), \quad (1.50)$$

що відповідає (1.11).

2. $r=1$. Тут імовірність переходу з одного стану джерела в інше $p(S_l / S_k) = p(x_j / x_i)$, можливе число таких переходів дорівнює m , як і кількість станів джерела. Ймовірність знаходження джерела в стані S_k , $p(S_k) = p(x_k) = p(x_i)$.

Тоді

$$\begin{aligned} H &= - \sum_{i=1}^m \sum_{j=1}^m p(x_i) \cdot p(x_j / x_i) \cdot \log p(x_j / x_i) = \\ &= - \sum_{i=1}^m \sum_{j=1}^m p(x_i, x_j) \cdot \log p(x_j / x_i). \end{aligned} \quad (1.51)$$

3. $r = 2$. Ймовірність переходу $p(S_l / S_k) = p(S_{ih} / S_{hj}) = p(x_j / x_i, h_j)$, а число станів джерела S_k – це число пар x_i, x_h , яке дорівнює m^2 .

При цьому $\sum_{i=1}^m \sum_{h=1}^m p(S_{ih}) = \sum_{i=1}^m \sum_{h=1}^m p(x_i, x_h) = 1$.

Число переходів S_{ih} / S_{hj} дорівнює основі алфавіту m (кількості можливих значень елемента j)

С урахуванням отриманих співвідношень

$$H = - \sum_{i=1}^m \sum_{h=1}^m \sum_{j=1}^m p(x_i, x_h, x_j) \cdot \log p(x_j / x_i, x_h). \quad (1.52)$$

На основі властивостей ергодичних джерел К. Шенон показав можливість передачі інформації з як завгодно великою точністю в каналах з перешкодами

1.9. Ентропія двопозиційних ергодичних джерел

Формули для ентропії ергодичних джерел дискретних повідомлень (1.50)–(1.52) передбачають їх однозначний характер. В той самий час у цифрових інформаційних системах телемеханіки (управління на відстані) поширені двопозиційні повідомлення.

Згідно ДОСТ 26.205–88 джерелами двопозиційних повідомлень в системах телемеханіки є ланцюги введення двопозиційних команд телеуправління (ТУ) та повідомлень спорадичної телесигналізації (ТСС) при використанні методу забезпечення достовірності – вирішального зворотного зв'язку (перезапиту). Прикладом двопозиційних команд (повідомлень) може бути: «Увімкнути»/«Вимкнути» i об'єкт. Для забезпечення своєчасної доставки повідомлень в багатоканальній системі передачі інформації за допомогою цього методу повідомлення формується та передається тільки в моменти зміни стану об'єкту ТУ (ТСС). Ця передача при відсутності помилки та наступного перезапиту є однократною, а прийняте повідомлення буде зберігатися на прийомній стороні до чергової зміни стану об'єкту ТУ (ТСС). Звідси випливає, що з потоку повідомлень від одного й того самого об'єкту виключається виникнення послідовних однотипних повідомлень, наприклад вигляду: «Увімкнути» i об'єкт. Це приводить до наявності статистичної залежності повідомлень в ймовірнісній моделі системи передачі інформації. Тим не менш, в системах телемеханіки двопозиційні повідомлення, що відповідають різним станам одного об'єкта, вважаються різними та кодуються різними кодовими комбінаціями. За цією причиною ентропії джерел однозначних та двопозиційних повідомлень з однаковими ймовірнісними розподілами мусять бути різними.

1. $r = 0$. Оскільки за досить тривалий проміжок часу число переходів j -го об'єкта з одного стану в інший i в зворотному напрямку в середньому однакове, для одиничного опитування ймовірності появи повідомлень $p(x_{j1})$

«Увімкнути» та $p(x_{j2})$ «Вимкнути» однакові, тобто $p(x_{j1}) = p(x_{j2}) = 0.5 p(x_j)$. Тоді для середньої питомої ентропії такого джерела повідомлень можна записати:

$$H = -\sum_{j=1}^m \sum_{k=1}^2 p(x_{jk}) \log p(x_{jk}) = -\sum_{j=1}^m p(x_j) \log [0.5 p(x_j)] = 1 - \sum_{j=1}^m p(x_j) \log p(x_j). \quad (1.53)$$

Тут другий доданок являє собою середню питому ентропію джерела m незалежних несумісних дискретних повідомлень, а перший доданок (додатковий біт) несе інформацію про конкретний варіант двопозиційного повідомлення («Увімкнути» або «Вимкнути»).

2. $r=1$. Якщо позначити як ймовірність появи при двох послідовних опитуваннях повідомлення i -го об'єкту і повідомлення «Увімкнути» j -го об'єкту як $p(x_i, x_{j1})$, ймовірність повідомлення i -го об'єкту і повідомлення «Вимкнути» j -го об'єкту як $p(x_i, x_{j2})$, а відповідні умовні ймовірності як $p(x_{j1}/x_i)$ та $p(x_{j2}/x_i)$, то, керуючись тими самими міркуваннями, що й у випадку джерела незалежних повідомлень, можна записати:

$$p(x_i, x_{j1}) = p(x_i, x_{j2}) = 0.5 p(x_i, x_j); \quad (1.54)$$

$$p(x_{j1}/i) = \frac{p(x_i, x_{j1})}{p(x_i)} = p(x_{j2}/x_i) = \frac{p(x_i, x_{j2})}{p(x_i)} = 0.5 p(x_j/x_i) = \frac{0.5 p(x_i, x_j)}{p(x_i)}. \quad (1.55)$$

Тоді ентропія джерела, що розглядається згідно з урахуванням двопозиційного характеру повідомлень, становитимете

$$\begin{aligned} H &= -\sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^2 p(x_{j1}, x_i) \log [p(jk/i)] = -\sum_{i=1}^m \sum_{j=1}^m p(x_i, x_j) \log [0.5 p(x_j/x_i)] = \\ &= 1 - \sum_{i=1}^m \sum_{j=1}^m p(x_i, x_j) \log [p(x_j/x_i)]. \end{aligned} \quad (1.56)$$

Також показано, що для довільного значення порядку ергодичного джерела r у співвідношеннях, які визначають ентропії джерел залежних повідомлень, так само як і у виразах (1.53), (1.56), додатковий біт несе інформацію про конкретний вигляд двопозиційного повідомлення.

1.10. Ентропія об'єднання дослідів. Умовна ентропія

Об'єднанням дослідів називаються спільні досліди із взаємозалежними наслідками. Приклад об'єднання – дія джерела й приймача сигналів при передачі сигналів через канал з перешкодами.

Джерело виробляє сигнали $x_1, x_2, \dots, x_i, \dots, x_m$, а приймач приймає $y_1, y_2, \dots, y_i, \dots, y_m$, які можуть відповідати сигналам джерела чи ні. При цьому задані

розподіли безумовних імовірностей $p(x_1), p(x_2), \dots, p(x_i), \dots, p(x_m)$, а також умовних імовірностей $y_j/x_i, x_i/y_j$.

Як відомо, безумовні й умовні ймовірності в спільних дослідах зв'язані виразом

$$p(x_i, y_j) = p(y_j, x_i) = p(x_i) \cdot p(y_j / x_i) = p(y_j) \cdot p(x_i / y_j) \quad (1.57)$$

Процес передачі сигналів відображається графом каналу рис. 1.9

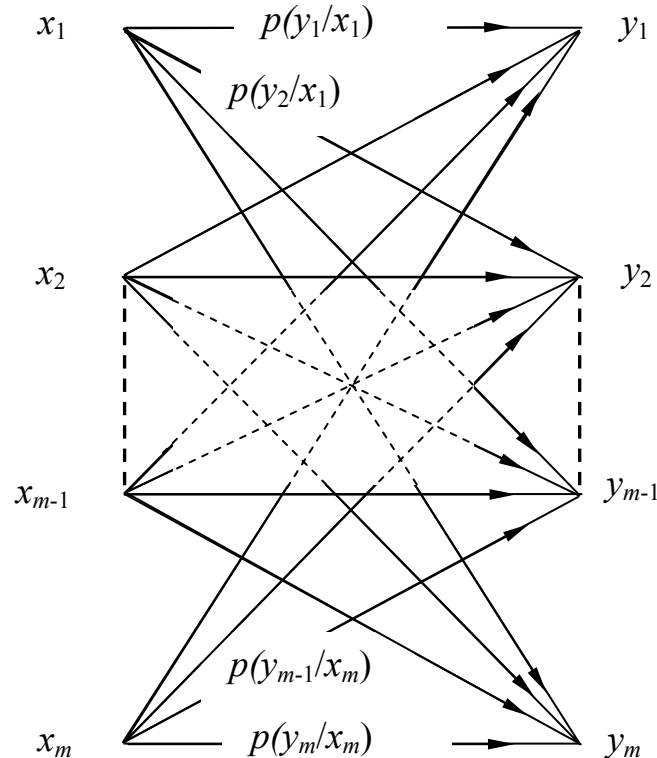


Рис. 1.9

або каналною матрицею

$$M(y/x) = \begin{vmatrix} y_1/x_1 & y_2/x_1 & \dots & y_{m-1}/x_1 & y_m/x_1 \\ y_1/x_2 & y_2/x_2 & \dots & y_{m-1}/x_2 & y_m/x_2 \\ \dots & \dots & \dots & \dots & \dots \\ y_1/x_{m-1} & y_2/x_{m-1} & \dots & y_{m-1}/x_{m-1} & y_m/x_{m-1} \\ y_1/x_m & y_2/x_m & \dots & y_{m-1}/x_m & y_m/x_m \end{vmatrix}. \quad (1.58)$$

Дія перешкод приводить до того, що передаючи сигнал x_i , у послідовності після нього будемо одержувати різні сигнали y_j , що й відображається графом і каналною матрицею.

Вибором якісних ознак сигналів (частоти, фази та ін.) з урахуванням характеристик перешкод забезпечуються найбільші значення умовних

імовірностей на головній діагоналі матриці, чим i досягається найбільш достовірною передачею сигналів.

Оскільки згідно (1.11), ентропія об'єднання визначається як

$$\begin{cases} H(X, Y) = -\sum_i \sum_j p(x_i, y_j) \cdot \log p(x_i, y_j) \\ H(Y, X) = -\sum_i \sum_j p(y_j, x_i) \cdot \log p(y_j, x_i) \\ H(X, Y) = H(Y, X) \end{cases} \quad (1.59)$$

з урахуванням (1.57)

$$\begin{aligned} H(X, Y) &= -\sum_i \sum_j p(x_i) \cdot p(y_j / x_i) \cdot \log [p(x_i) \cdot p(y_j / x_i)] = \\ &= -\sum_i p(x_i) \cdot \log p(x_i) \cdot \sum_j p(y_j / x_i) - \\ &\quad -\sum_i p(x_i) \cdot \sum_j p(y_j / x_i) \cdot \log p(y_j / x_i) = H(X) + H(Y / X). \end{aligned} \quad (1.60)$$

У перетвореннях (1.60) використані заміни подвійних сум вкладеними, а умовні ймовірності, як відомо задовольняють співвідношенню

$$\sum_j p(y_j / x_i) = 1. \quad (1.61)$$

У формулі (1.60) перший доданок – це ентропія джерела $H(X)$, а $H(Y/X)$ – середня умовна ентропія – міра невизначеності того, що ансамбль прийнятих сигналів Y відповідає ансамблю переданих X . Ця ентропія складається із часткових умовних ентропій

$$H(Y / x_i) = -\sum_j p(y_j / x_i) \cdot \log p(y_j / x_i), \quad (1.62)$$

тобто часткова умовна ентропія – це міра невизначеності прийнятого ансамблю Y щодо переданого сигналу x_i , а ентропія $H(Y / X)$ виходить шляхом їхнього усереднення з урахуванням розподілу безумовних ймовірностей $p(x_i)$

$$\begin{aligned} H(Y / X) &= \sum_i p(x_i) \cdot H(Y / x_i) = \\ &= -\sum_i p(x_i) \cdot \sum_j p(y_j / x_i) \cdot \log p(y_j / x_i) = \\ &= -\sum_i \sum_j p(y_j, x_i) \cdot \log p(y_j / x_i). \end{aligned} \quad (1.63)$$

На підставі (1.57), (1.59) і (1.60) також можна записати

$$H(Y, X) = H(Y) + H(X/Y). \quad (1.64)$$

Умовні ймовірності $p(y_j/x_i)$ можуть приймати граничні значення $p(y_j/x_i)=1$, якщо прийняті сигнали такі самі, як і передані (або існує детермінований зв'язок між прийнятими і переданими сигналами) й $p(y_j/x_i) = p(y_j)$, якщо який-небудь зв'язок між переданими й прийнятими сигналами відсутній.

У першому випадку

$$H(X, Y) = H(X) - \sum_i p(x_i) \cdot \sum_j 1 \cdot \log 1 = H(X), \quad (1.65)$$

або $H(Y, X) = H(Y)$.

У другому

$$\begin{aligned} H(X, Y) &= H(X) - \sum_i p(x_i) \cdot \sum_j p(y_j/x_i) \cdot \log p(y_j/x_i) = \\ &= H(X) - \sum_i p(x_i) \cdot \sum_j p(y_j) \log p(y_j) = H(X) + H(Y). \end{aligned} \quad (1.66)$$

Ентропія об'єднання тут досягає максимуму, а передача інформації від джерела до приймача відсутня.

1.11. Пропускна здатність дискретного каналу без перешкод

У цьому каналі інформація, що вироблюється джерелом, без втрат доставляється приймачу.

Пропускна здатність каналу – це максимальна кількість інформації, що може бути передана по каналу зв'язку в одиницю часу

$$C_{\Pi} = \lim_{T \rightarrow \infty} \frac{J(T)_{\max}}{T}. \quad (1.67)$$

Тут $J(T)_{\max}$ – максимально можлива кількість інформації, вироблена джерелом за час T .

Вище було показано, що $J(T)_{\max}$ досягається для джерела, виробляючого рівноймовірні й взаємозалежні повідомлення.

Далі розглядається випадок, коли повідомлення являють собою послідовності довжиною n при алфавіті з основою m . Максимальне число різних слів (повідомлень) становить $N_{\max} = m^n$; а на інтервалі часу T буде створено $N(T)_{\max} = m^{n(T)}$ повідомлень.

Згідно (1.2) вираження (1.67) можна записати у вигляді

$$C_{\Pi} = \lim_{T \rightarrow \infty} \frac{\log N(T)_{\max}}{T} = \lim_{T \rightarrow \infty} \frac{\log m^{n(T)}}{T}. \quad (1.68)$$

Нехай елементи послідовності відображаються сигналами з тривалістю τ , тоді $n(T) = \frac{T}{\tau}$ й

$$C_{\Pi} = \lim_{T \rightarrow \infty} \frac{\log m^{\frac{T}{\tau}}}{T} = \lim_{T \rightarrow \infty} \frac{\frac{T}{\tau} \log m}{T} = \frac{1}{\tau} \log m. \quad (1.69)$$

Величина $U = \frac{1}{\tau}$, c^{-1} називається швидкістю передачі сигналів. У системах зв'язку для неї використовується також одиниця виміру Бод, c^{-1} . Оскільки згідно (1.22) для джерела, повідомлення якого становлять символи з основою алфавіту m , максимальне значення ентропії дорівнює $H_{\max} = \log m$, остаточно пропускна здатність дискретного каналу без перешкод становить

$$C_{\Pi} = U \cdot H_{\max}, \text{ біт/с.} \quad (1.70)$$

У системах передачі інформації реальні значення ентропії джерел $H < H_{\max}$ і швидкість передачі інформації

$$C = U \cdot H < C_{\Pi}.$$

Збільшити значення C можливо, використовуючи більш швидкісні канали зв'язку (наприклад, замість електричних провідних – волоконнооптичні) або застосовуючи так зване оптимальне кодування повідомлень джерела, що дозволяє збільшити його ентропію.

1.12. Пропускна здатність каналу з перешкодами (статистичний підхід)

Використовуючи поняття умовної ентропії як залишкової невизначеності при проведенні сполучення наслідків в об'єднанні дослідів (дія джерела, каналу й приймача повідомлень), кількість інформації, що досягла приймача, за час T може бути записана у вигляді

$$J = H(Y) - H(Y/X) \text{ або } J = H(X) - H(X/Y). \quad (1.71)$$

Тоді, з огляду на (1.70), у загальному випадку пропускна здатність каналу з перешкодами складе

$$C_{\Pi}^{\Pi} = U \cdot \max(H) = U \cdot \max\{H(Y) - H(Y/X)\} = U \cdot \max\{H(X) - H(X/Y)\}. \quad (1.72)$$

Таким чином, для визначення C_{Π}^{Π} необхідно знайти максимум функції $\{H\}$. Рішення цієї задачі покажемо на прикладі так званого двійкового симетричного каналу без пам'яті.

У цьому випадку джерело створює сигнали x_0, x_1 з безумовними ймовірностями $p(x_0), p(x_1)$, а симетричний канал характеризується рівністю умовних ймовірностей

$$p(y_0/x_1) = p(y_1/x_0) = p \text{ и } p(y_0/x_0) = p(y_1/x_1) = 1 - p. \quad (1.73)$$

Термін «без пам'яті» визначає той факт, що джерело створює взаємозалежні сигнали, тут x_0, x_1 і $p(x_0) + p(x_1) = 1$.

Граф розглянутого каналу наведений на рис. 1.10.

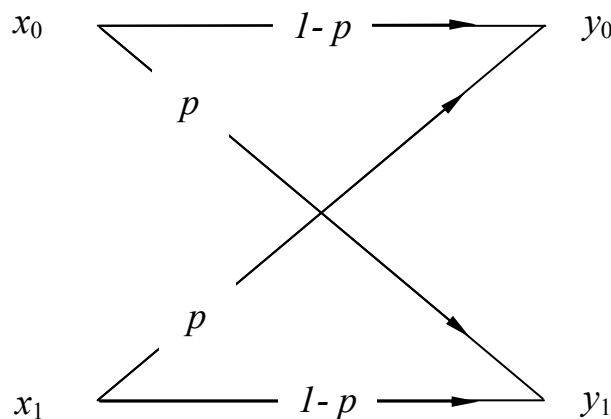


Рис. 1.10

Очевидно, що максимум функції (1.73) досягається при $H(Y)_{\max}, H(X)_{\max}$, якщо при цьому умовні ентропії $H(Y/X), H(X/Y)$ залежать лише від властивостей каналу й не залежать від властивостей джерела й приймача.

Відповідно до формули повної ймовірності

$$p(y_j) = \sum_{i=1}^m p(y_j/x_i) \cdot p(x_i). \quad (1.74)$$

Тоді

$$\begin{aligned}
p(y_0) &= p(y_0/x_0) \cdot p(x_0) + p(y_0/x_1) \cdot p(x_1) = \\
&= (1-p) \cdot p(x_0) + p \cdot p(x_1) = p(x_0) - p \cdot p(x_0) + p \cdot p(x_1) = \quad (1.75) \\
&= p(x_0) + p[p(x_1) - p(x_0)].
\end{aligned}$$

Джерело створює максимум інформації, коли $p(x_1) = p(x_0)$ й у цьому випадку $p(y_0) = p(x_0) = p(y_1)$.

Отже, максимум ентропії $H(Y)_{\max}$ досягається при максимумі $H(X)_{\max}$.

Згідно (1.21) для двійкового джерела $H_{\max} = 1$ біт/сигнал, тобто в співвідношенні (1.72) $H(Y) = 1$ біт/сигнал.

Умовна ентропія за виразом (1.58) у розглянутому прикладі зводиться до вигляду

$$\begin{aligned}
H(Y/X) &= -\{p(x_0) \cdot [p(y_0/x_0) \cdot \log p(y_0/x_0) + p(y_1/x_0) \cdot \log p(y_1/x_0)] + \\
&+ p(x_1) \cdot [p(y_0/x_1) \cdot \log p(y_0/x_1) + p(y_1/x_1) \cdot \log p(y_1/x_1)]\} = \\
&= -\{p(x_0) \cdot [(1-p) \cdot \log(1-p) + p \cdot \log p] + \\
&+ p(x_1) \cdot [p \cdot \log p + (1-p) \cdot \log(1-p)]\} = \quad (1.76) \\
&= -[p(x_0) + p(x_1)] \cdot [p \cdot \log p + (1-p) \cdot \log(1-p)] = \\
&= -p \cdot \log p - (1-p) \cdot \log(1-p).
\end{aligned}$$

Отриманий результат показує, що $H(Y/X)$ залежить тільки від властивостей даного каналу й не залежить від розподілів безумовних імовірностей, з якими джерело створює сигнали.

Таким чином, (1.72) остаточно здобуває вигляд

$$C_{\Pi}^{\Pi} = U \cdot [1 + p \cdot \log p + (1-p) \cdot \log(1-p)], \text{ біт/с.} \quad (1.77)$$

Графік цієї залежності зображений на рис. 1.11.

При граничних значеннях імовірності помилкового прийому сигналів $p = 0$, $p = 1$ пропускна здатність каналу досягає максимуму, $C_{\Pi}^{\Pi} = |U|$, а при $p = 0,5$ $C_{\Pi}^{\Pi} = 0$.

Варто помітити, що якщо $p > 0,5$, має місце негативний режим роботи каналу, коли сигнал x_0 розпізнається як y_1 , а x_1 , як y_0 . Крім того, зменшення пропускної здатності каналу до нуля при $p = 0,5$ – наслідок використання статистичної моделі каналу.

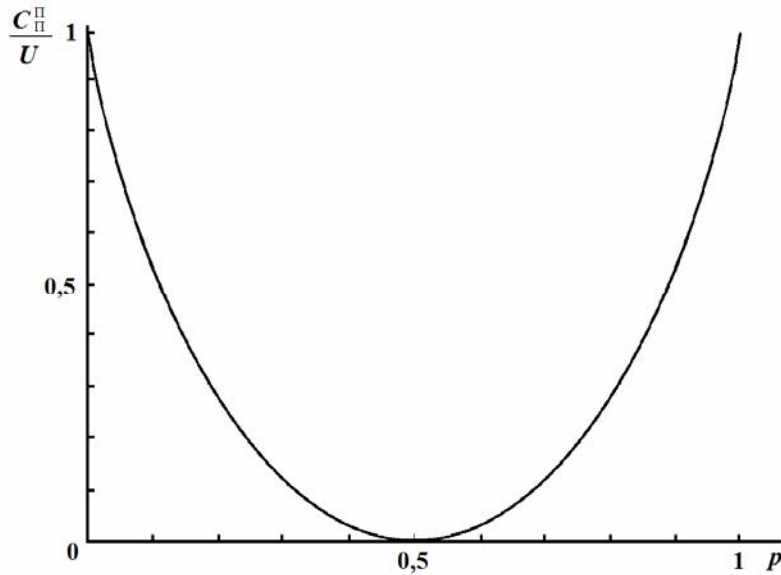


Рис. 1.11

Як буде викладено нижче, пропускна здатність існує для каналів із як завгодно більшим впливом перешкод на передані сигнали.

Становить інтерес різновид розглянутого каналу, у якого приймач за певних умов не може розпізнати прийнятий сигнал як y_0 або y_1 . Це може бути, наприклад, у тому випадку, якщо для передачі сигналів x_0 і x_1 використовуються різнополярні імпульси постійного струму.

Граф цього каналу показаний на рис. 1.12.

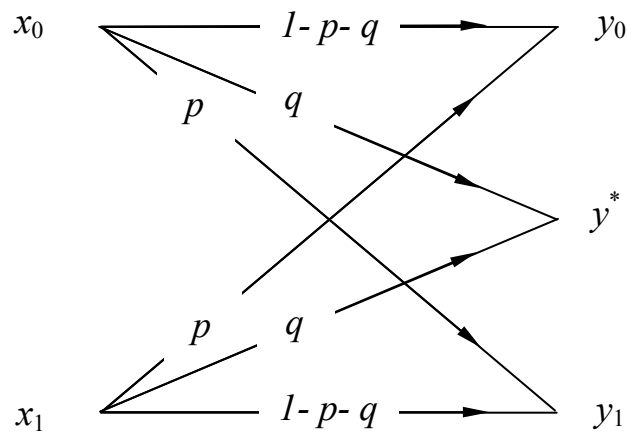


Рис. 1.12

Канал, у якого прийнятий сигнал може не розпізнатися (приймач при цьому переходить у стан y^* з імовірністю q), називається каналом зі стиранням. Для нього пропускна здатність визначається як

$$C_{\Pi}^{\Pi} = U \cdot \{(1 - q) \cdot [1 - \log(1 - q)] + (1 - p - q) \cdot \log(1 - p - q) + p \cdot \log p\}, \text{ біт/с. (1.78)}$$

1.13. Передача інформації в каналі з перешкодами. Пропускна здатність каналу. Енергетичний підхід

Вище ця задача була розглянута, виходячи зі статистичних характеристик об'єкта – безумовних і умовних імовірностей, що визначають передачу інформації в каналі зв'язку.

Носієм інформації в сигналі є його енергія, тому К. Шеннон вирішив задачу передачі інформації в каналі із цієї позиції.

Розглянемо неперервний випадковий процес $x(t)$, що є сигналом джерела інформації. У каналі зв'язку на цей процес накладається перешкода $x_{\Pi}(t)$; обидві ці складові підсумуються за миттєвим значенням, так що в каналі створюється перекручений сигнал

$$x'(t) = x(t) + x_{\Pi}(t). \quad (1.79)$$

Перейдемо від неперервного подання величини $x'(t)$ до дискретного за часом і за множиною значень. Такий перехід здійснюється шляхом квантування за часом із кроком (квантом) Δt , а за множиною значень – із квантом Δx . На рис. 1.13 наданий графік залежності величини $x'(t)$, тут кружками показані відліки її значень у моменти часу t_i , а прямокутниками – відповідні величини, які квантовані за рівнем.

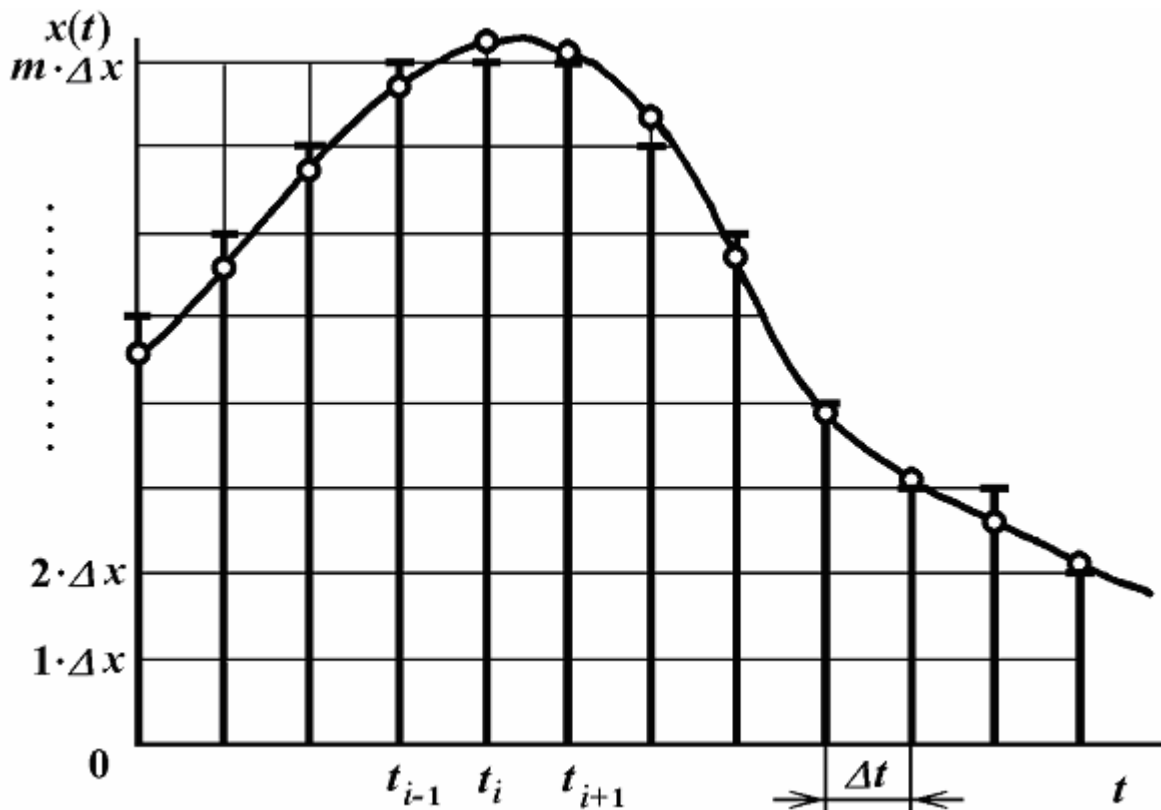


Рис. 1.13

За теоремою Котельникова, якщо спектр $x'(t)$ обмежений зверху частотою F_B , то такий неперервний процес точно відображається послідовністю його відліків у моменти часу t_i , що мають із інтервал

$$\Delta t = \frac{1}{2F_B}. \quad (1.80)$$

Квантування відліків $x'(t_i)$ за рівнем обмежує кількість можливих значень сигналу величиною

$$m = E \left[\frac{x_{\max}}{\Delta x} \right] + 1, \quad (1.81)$$

де x_{\max} – максимально можливе значення величини сигналу, E – символ округлення до його більшого результату ділення. Вираз записано в припущенні, що мінімально можливе значення величини сигналу $x_{\min} = 0$.

Внаслідок дискретизації на інтервалі часу T створюється $n(T) = \frac{T}{\Delta t}$ відліків випадкової величини $x'(t_i)$, які при виконанні умови (1.80) будуть взаємозалежними. Тоді на інтервалі T може бути отримано максимум

$$N(T) = m^{n(T)} \quad (1.82)$$

варіантів дискретних повідомлень у процесі $x'(t)$. З урахуванням (1.80) останній вираз перетвориться до вигляду

$$N(T) = m^{2TF_B}. \quad (1.83)$$

Враховуючи сформульовані вище припущення про властивості сигналу, можна стверджувати, що таке джерело створює кількість інформації, максимальне значення якого становить

$$\dot{J}_C(T) = \log N(T) = \log m^{2TF_B} = 2TF_B \cdot \log m. \quad (1.84)$$

З метою спрощення наступних перетворень припускається, що функції щільностей імовірностей значень сигналу $w(x)$ та перешкоди $w(x_{\Pi})$ відомі й однакові, наприклад, є рівномірними з параметрами, що збіглись.

Тоді величина m^2 у рівнянні (1.84) пропорційна середній потужності сигналу (перешкоди) на інтервалі T і цей вираз може бути записаний у вигляді

$$\dot{J}_C(T) = TF_B \cdot \log(AP_C). \quad (1.85)$$

Перешкода, що діє в цій самій смузі частот, що й сигнал, за аналогією, буде характеризуватися кількістю шумової інформації

$$\dot{J}_\Pi(T) = TF_B \cdot \log(AP_\Pi). \quad (1.86)$$

Тому що перешкода адитивна, у каналі зв'язку передається потужність $P_K = P_C + P_\Pi$, з якою пов'язана кількість інформації

$$\dot{J}_K(T) = TF_B \cdot \log[A(P_C + P_\Pi)]. \quad (1.87)$$

Кількість інформації, що доставляється на вхід приймача від джерела інформації, у свою чергу становить

$$\dot{J}(T) = \dot{J}_K(T) - \dot{J}_\Pi(T) = TF_B \cdot \log\left(1 + \frac{P_C}{P_\Pi}\right). \quad (1.88)$$

З останнього виразу випливає, що передача необхідної кількості інформації можлива при будь-якому, як завгодно малому, відношенні $\frac{P_C}{P_\Pi}$, тому що при цьому значення логарифма, що входить у вираз є позитивним.

Дана формула справедлива й у тих випадках, коли смуга частот обмежена знизу значенням $F_H > 0$, тоді замість F_B використовується значення $\Delta F = F_B - F_H$, і формула набуває вигляд

$$\dot{J}(T) = T \cdot \Delta F \cdot \log\left(1 + \frac{P_C}{P_\Pi}\right). \quad (1.89)$$

Пропускна здатність розглянутого каналу визначається як

$$C_\Pi = \lim_{T \rightarrow \infty} \frac{\dot{J}(T)}{T} = \Delta F \cdot \log\left(1 + \frac{P_C}{P_\Pi}\right), \quad (1.90)$$

тобто пропорційна ширині смуги частот ΔF , однакової для сигналу, перешкоди й каналу.

На практиці поширені випадки каналів зв'язку з вузькосмуговим сигналом і широкосмуговою перешкодою, при яких співвідношення між шириною смуг частот сигналу й каналу становить $\Delta F_C \ll \Delta F_\Pi$. Дослідимо

пропускну здатність такого каналу. Для цього перетворимо вираження (1.90) при $\Delta F = F_B$, зобразивши потужність перешкоди як

$$P_{\Pi} = p_0 F_B, \quad (1.91)$$

де p_0 – питома потужність перешкоди в смузі частот 1 Гц, Вт/Гц. Тоді зазначений вираз набуває вигляд

$$C_{\Pi}^{\Pi}(F_B) = F_B \cdot \log\left(1 + \frac{P_C}{p_0 F_B}\right) = F_B \cdot \log\left(1 + \frac{F_0}{F_B}\right) = \frac{\log\left(1 + \frac{F_0}{F_B}\right)}{\frac{1}{F_B}}, \quad (1.92)$$

тут F_0 – смуга частот, де потужність перешкоди дорівнює потужності сигналу.

Аналізована залежність $C_{\Pi}^{\Pi} = C_{\Pi}^{\Pi}(F_B)$ має невизначеності в граничних випадках при F_B , рівному 0, і $+\infty$. Для їхнього розкриття скористаємося правилом Лопітала. Функція, що зображує відношення похідних чисельника й знаменника у виразі (1.92), має вигляд

$$f(F_B) = \frac{F_0 \cdot F_B}{(F_0 + F_B) \cdot \ln 2}. \quad (1.93)$$

Відповідно шукані межі складуть

$$\lim_{F_B \rightarrow 0} C_{\Pi}^{\Pi}(F_B) = \lim_{F_B \rightarrow 0} f(F_B) = 0, \quad (1.94)$$

$$\lim_{F_B \rightarrow \infty} C_{\Pi}^{\Pi}(F_B) = \lim_{F_B \rightarrow \infty} f(F_B) = \frac{|F_0|}{\ln 2}, \text{ біт/с}, \quad (1.95)$$

а графік залежності (1.92) буде мати вигляд, зображений на рис. 1.14.

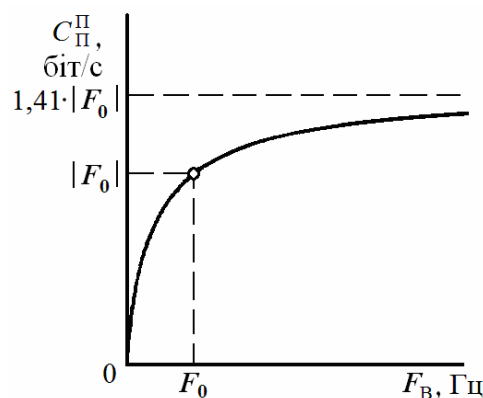


Рис. 1.14

На практиці смуга частот каналу із широкосмуговою перешкодою обмежується величиною порядку $F_B = F_0$. При цьому відповідне значення пропускну́ї здатності каналу становить

$$C_{\Pi}^{\Pi}(F_0) = |F_0|, \text{ біт/с.} \quad (1.96)$$

Тут символ модуля означає чисельну рівність значень частоти F_0 й пропускну́ї здатності каналу.

1.14. Об'єми сигналів і каналів, їхнє узгодження

Характеристики джерела повідомлень і каналу зв'язку, якими є: часи дії T , смуги частот ΔF (що залежать від швидкості формування сигналів джерелом і передачі їх по каналу відповідно), а також логарифмічні функції відношення потужностей сигналу й перешкоди $\log\left(\frac{P_C}{P_{\Pi}}\right)$, можуть бути різними.

Така невідповідність може призводити до втрат інформації, переданої від джерела в канал. Вихід із цього протиріччя може бути здійснений на основі подання про об'єми сигналу й каналу. Вираз для кількості інформації у формі (1.89), що є добутком трьох співмножників, може розглядатися як аналог геометричного об'єму паралелепіпеда з розмірами основи T , ΔF і висотою

$$h = \log\left(1 + \frac{P_C}{P_{\Pi}}\right) \quad (\text{рис. 1.15, а}).$$

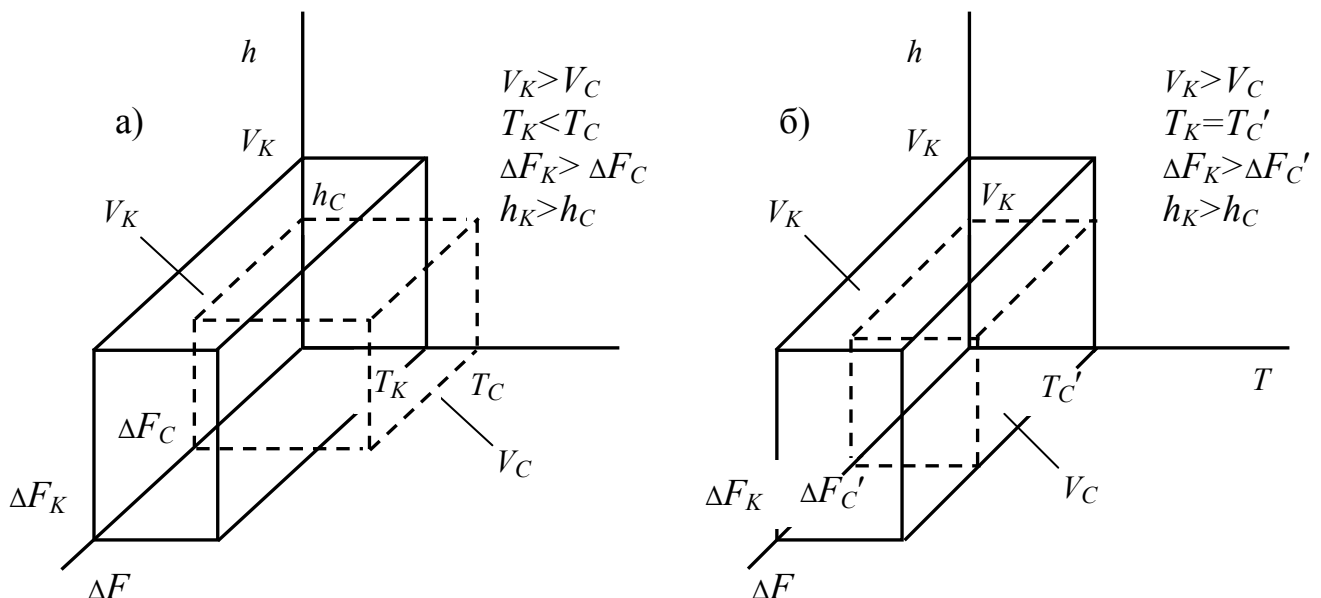


Рис. 1.15

У цьому поданні передача інформації без втрат здійснена за умови, що об'єм каналу V_K не менше об'єму сигналу V_C , тобто

$$V_K \geq V_C, \text{ біт/с.} \quad (1.97)$$

Тут

$$\begin{cases} V_C = T_C \cdot \Delta F_C \cdot h_C \\ V_K = T_K \cdot \Delta F_K \cdot h_K, \end{cases} \quad (1.98)$$

а індекси C або K всіх співмножників означають, що вони характеризують сигнал або канал відповідно. У системах передачі технологічної (СП) інформації (системах телемеханіки) $\frac{P_C}{P_{\Pi}} \gg 1$ і, отже,

$$h_C \approx \log\left(\frac{P_C}{P_{\Pi}}\right). \quad (1.99)$$

Аналогічно для каналу можна записати

$$h_K \approx \log\left(\frac{P_K}{P_{\Pi}}\right), \quad (1.100)$$

де P_K – максимальна потужність сигналу в каналі, забезпечувана технічними характеристиками приймально-передавальної апаратури. При цьому передбачається, що в номінальному режимі роботи СП $P_C < P_K$, тобто апаратура допускає, при необхідності, деяке збільшення номінальної потужності сигналу.

Нерівність (1.97) зветься загальною умовою узгодження сигналу й каналу і є необхідною для здійснення узгодження.

Відповідними достатніми умовами є нерівності системи

$$\begin{cases} T_K \geq T_C \\ \Delta F_K \geq \Delta F_C \\ h_K \geq h_C, \end{cases} \quad (1.101)$$

які зветься частковими умовами узгодження сигналу й каналу.

Якщо виконується загальна умова узгодження сигналу й каналу, але не виконується одна або дві із часткових, то використовується трансформація параметрів сигналу на основі інваріантів перетворень

$$\begin{cases} T_C \cdot \Delta F_C = const \\ T_C \cdot h_C = const \\ \Delta F_C \cdot h_C = const, \end{cases} \quad (1.102)$$

Необхідне узгодження може здійснюватися й трансформаціями параметрів каналу. Однак це неможливо в СПП, де каналотворююча апаратура обслуговує множину пар «джерело – приймач» повідомлень із різнорідними індивідуальними характеристиками.

Так, згідно з рис. (1.15, а) $V_K > V_C$, але час формування повідомлення T_C тут у півтора рази більше часу T_K надання каналу для його передачі, тобто не виконується перша часткова умова узгодження. Застосовуючи перший інваріант перетворення у вигляді $T_C \cdot \Delta F_C = T'_C \cdot \Delta F'_C$ й вважаючи $T'_C = T_K$, можна одержати

$$\Delta F'_C = \Delta F_C \cdot \frac{T_C}{T_K}.$$

Тепер об'єм сигналу за усіма трьома вимірами уписується в об'єм каналу (рис. 1.15, б).

Практично таке узгодження використовується, наприклад, під час передачі телефонних сигналів багатьох абонентів по спільній лінії зв'язку – високочастотному електричному або оптоволоконному кабелю. Тут можливий перенос спектра сигналу на необхідну ділянку частот каналу, як це наведено на рис. 1.16. Перенос здійснюється внаслідок процесу модуляції високочастотної гармоніки (несучої частоти) сигналом.

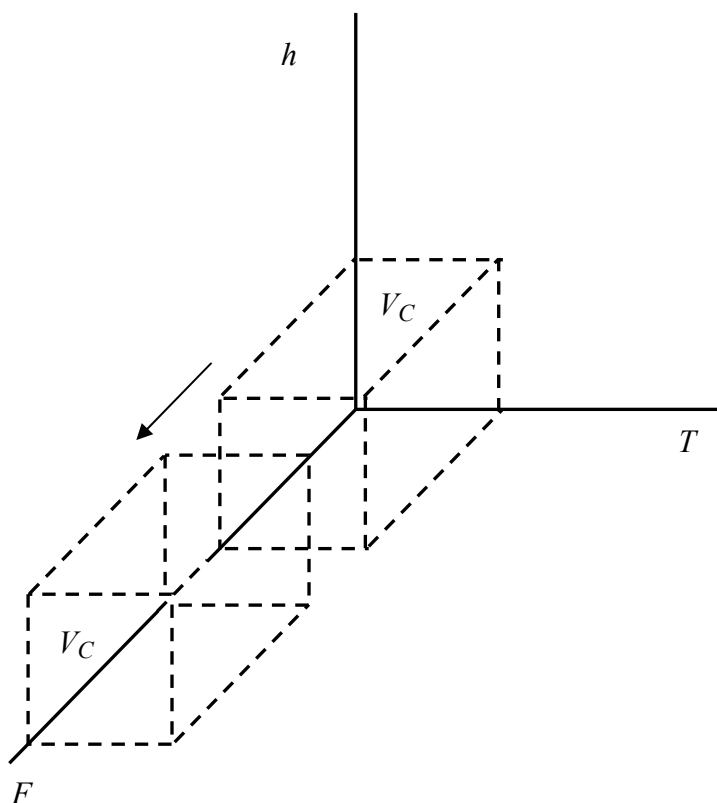


Рис. 1.16

Негативною стороною розглянутого узгодження є затримки в доставці повідомлень одержувачеві, які виникають через необхідність попереднього запису їх у пам'ять передавача в реальному часі, передачі в канал у прискореному режимі, а також зворотному перетворенні сигналу в приймачі.

З такими самими затримками пов'язана передача зображень у малокадровому телебаченні. Тут досить великий об'єм відеоінформації, одержуваної за короткий час експонування – запису в пам'ять (рис. 1.17), може передаватися по низькочастотній лінії зв'язку, але це потребує виконання нерівності $T_K \gg T_C$. Очевидно, що передача чергового кадру зображення можлива тільки після звільнення каналу, тобто з інтервалом T'_C , якщо до моменту закінчення передачі даного кадру в пам'яті джерела вже записаний наступний кадр.

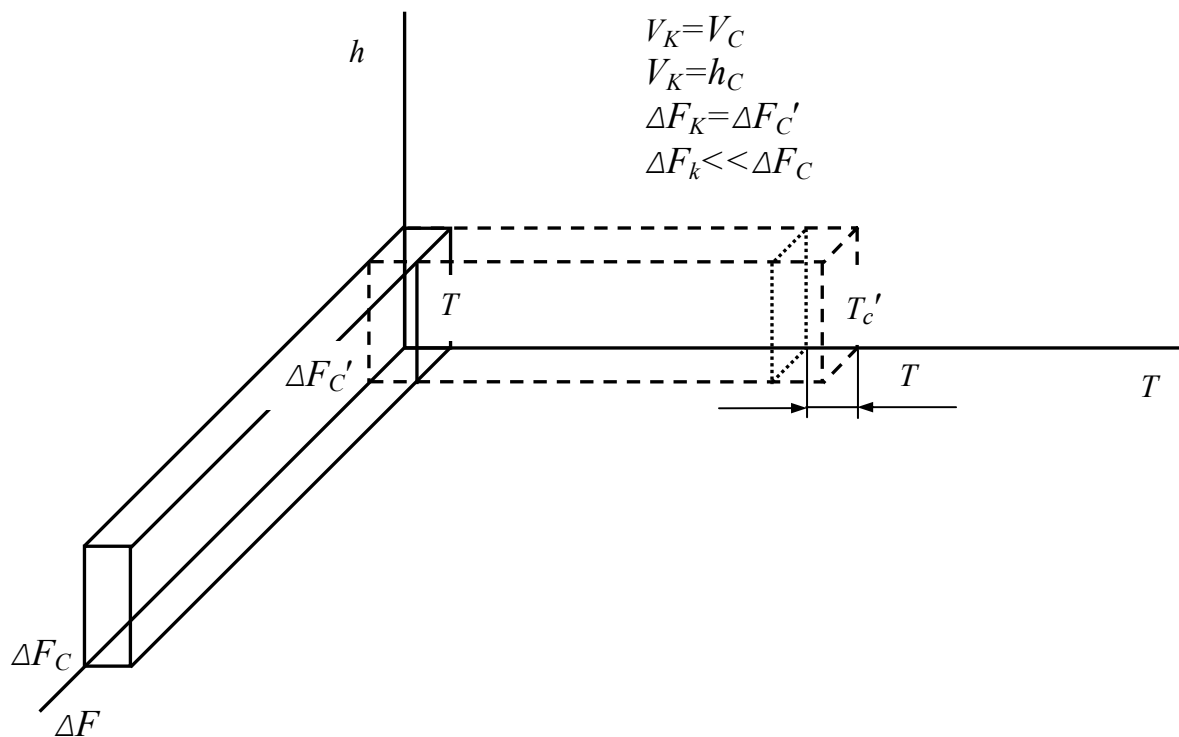


Рис. 1.17

1.15. Надлишковість повідомлень

Характеристика, що розглядається, є наслідком нерівноймовірності й взаємозалежності появи як елементів, так і самих повідомлень.

Поняття інформаційної надлишковості проявляється як у технічних системах, так і в інших галузях, наприклад структурах мов людського спілкування.

Наприклад, англійський алфавіт містить $m = 26$ літер. Якби в словах зустрічалися з однаковою ймовірністю та були статистично незалежні, то ентропія джерела, у якому повідомленнями виступали літери такого алфавіту, була б максимальною й становила

$$H_{\max} = \log 26 \approx 4,70, \text{ біт/літеру.}$$

Однак літери в словах з'являються нерівноймовірно й взаємозалежно одна від одної. Прикладом нерівноймовірності появи літер є більша поширеність голосних ніж приголосних, а взаємозалежності – різна ймовірність появи сполучень з двох літер (біграмм) при різному порядку слідування літер. Розрахунок ентропії розглянутого джерела, зроблений за допомоги репрезентабельної виборки слів з 8-мі літер дає значення 2,35 біт/літеру.

При урахуванні статистики проходження слів у технічних текстах ентропія знижується до величини 2 біт/літеру.

Надлишковість повідомлення визначається як

$$D = 1 - \frac{H}{H_{\max}}, \quad (1.103)$$

де H – фактичне значення ентропії повідомлення, $H_{\max} = \log m$ – максимально можливе значення ентропії джерела символів алфавіту.

Якщо повідомлення джерела виникають нерівноймовірно й взаємозалежно при порядку джерела $r = 1$, то згідно (1.51)

$$H = -\sum_{i=1}^m \sum_{j=1}^m p(x_i, x_j) \cdot \log p(x_j / x_i).$$

Відсутність взаємозалежності приводить до збільшення ентропії (1.11)

$$H_0 = -\sum_{i=1}^m p_i \cdot \log p_i$$

Співвідношення (1.103) може бути перетворене до вигляду

$$D = 1 - \frac{H}{H_0} \cdot \frac{H_0}{H_{\max}} = 1 - (1 - D_s) \cdot (1 - D_p), \quad (1.104)$$

де $D_s = 1 - \frac{H}{H_0}$, $D_p = 1 - \frac{H_0}{H_{\max}}$ – носять назви статистичної й ймовірнісної надлишковостей відповідно. При невеликих значеннях D_s , D_p вираз (1.104) може бути записаний як $D \approx D_s + D_p$.

Величина

$$\mu = \frac{H}{H_{\max}} \quad (1.105)$$

зветься коефіцієнтом стиску повідомлення.

Надлишковість у інформаційних об'єктах відіграє двояку роль. Первинна надлишковість джерел – негативну, тому що знижує пропускну здатність каналів зв'язку й збільшує необхідні об'єми пам'яті обчислювальних пристроїв. Таку надлишковість зазвичай, якщо є можливість, прагнуть знизити шляхом оптимального кодування повідомлень або стиску даних. Однак спеціально організована надлишковість у кодах, що виявляють й виправляють помилки, дозволяє досягти необхідної вірогідності передачі повідомлень, що особливо важливо для систем телемеханіки.

Контрольні питання і завдання

1. Дайте визначення поняттю інформація.
2. Поясніть переваги й недоліки статистичної міри інформації.
3. Назвіть одиниці виміру інформації.
4. Наведіть приклади джерел дискретних і неперервних повідомлень.
5. Наведіть співвідношення для визначення кількості інформації, що несе дискретне повідомлення із заданою ймовірністю, поясніть у яких межах може змінюватися ця величина.
6. Поясніть, з яких міркувань обґрунтовується логарифмічна міра кількості інформації.
7. Повідомлення відображує результат витягування карти з преферансної колоди, що містить $n = 32$ карти. Обчисліть кількість інформації, що міститься в повідомленнях, які полягають у тому, що карта має: певний колір, певну масть, певне значення.
8. В умовах попереднього завдання визначте мінімальну, максимальну та середню кількість інформації, яку може нести повідомлення «так» чи «ні» у відповідь на запитання, чи має карта певне значення.
9. В умовах попереднього завдання визначте мінімальну кількість запитань, на які даються відповіді «так» чи «ні», що дозволяє гарантовано встановити значення карти. Обґрунтуйте, яким чином при цьому повинні ставитися запитання.
10. В урні містяться n білих та m чорних кульок. Випадковим чином послідовно виймаються дві кульки. Запишіть вирази для ймовірності події, яка полягає у тому, що одна кулька білі, а інша – чорна та для кількості інформації, яку несе повідомлення про цю подію. Розгляньте випадки для схем виймання кульок з поверненням та без повернення.
11. По цілі здійснюється $k = 4$ незалежних пострілів. Ймовірність влучення у ціль при кожному пострілі становить $p = 0,5$. Знайдіть кількість інформації, яку несе повідомлення про стан цілі після обстрілу.
12. Ймовірність того, що випадково обраний мешканець міста є студентом, складає $P(A)$, ймовірність того, що випадково обраний мешканець є юнаком – $P(B)$. Ймовірність того, що мешканець – юнак при умові, що він студент – $P(B/A)$. Скільки інформації несе повідомлення, що випадково обраний юнак є студентом?

13. Поясніть зв'язок між повною, частковою ентропією повідомлення та середньою питомою ентропією джерела повідомлень. Назвіть одиниці виміру цих величин та межі у яких вони можуть змінюватися.

14. Сформулюйте властивості часткової ентропії.

15. Сформулюйте властивості середньої питомої ентропії джерела дискретних повідомлень.

16. Назвіть умову максимуму ентропії джерела дискретних повідомлень.

17. Джерело дискретних повідомлень генерує незалежні дискретні повідомлення з ймовірностями $\{0,5, 0,25, 0,125, 0,0625, 0,0625\}$. Знайдіть ентропію джерела.

18. Доведіть, що ентропія геометричного ймовірнісного розподілу

$$p_i = p(x_i = l) = \begin{cases} P(1-P)^i & \text{при } x_i \geq 0 \\ 0 & \text{при } x_i < 0 \end{cases} \quad \text{з параметром } P \text{ обчислюється за}$$

формулою: $-\frac{P \cdot \log P - (1-P) \cdot \log(1-P)}{P}$. При доказі скористуйтеся тим, що

$$\sum_{i=0}^{+\infty} (1-P)^i = \frac{1}{P}, \text{ а } \sum_{i=0}^{+\infty} i(1-P)^i = \frac{1-P}{P^2}.$$

19. Доведіть, що ентропія нормального ймовірнісного розподілу

$$w(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2} \quad \text{з параметрами } (a, \sigma) \text{ обчислюється за формулою:}$$

$H_d = \log(\sigma\sqrt{2\pi e})$. При доказі скористуйтеся тим, що $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$, а

$$\int_{-\infty}^{+\infty} x^2 e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

20. Дайте визначення порядку ергодичного джерела повідомлень.

21. Дайте тлумачення поняття двопозиційного повідомлення. Приведіть формулу для обчислення ентропії джерела незалежних двопозиційних повідомлень та поясніть її складові.

22. Знайдіть ентропію джерела повідомлень про результат кидання двох монет.

23. Ймовірність $p(x_i, y_j)$ сумісної появи повідомлень об'єднання джерел $X - x_i$ та $Y - y_j$ визначається матрицею

$$\begin{vmatrix} 0,5 & 0 & 0 & 0 \\ 0 & 0,125 & 0,125 & 0 \\ 0 & 0 & 0 & 0,25 \end{vmatrix}.$$

Знайдіть ентропію об'єднання джерел повідомлень $H(X,Y)$, безумовні ймовірності появи повідомлень кожного з джерел та значення їхніх ентропій. Поясніть співвідношення між ними.

24. Дайте визначення каналної матриці. Поясніть, яким властивостям повинні задовольняти її елементи.
25. В умовах попереднього завдання побудуйте таблиці умовних імовірностей $p(y_j/x_i)$ та $p(x_j/y_i)$ та знайдіть значення умовних ентропій $H(Y/X)$, $H(X/Y)$. Поясніть співвідношення між ними.
26. Для джерела взаємозалежних двійкових повідомлень задані значення однієї безумовної $p(0)$ і однієї умовної $p(0/0)$ ймовірностей появи повідомлень. Знайдіть решту умовних імовірностей та умовну ентропію джерела, якщо попереднім повідомленням є 0.
27. Наведіть співвідношення, яке зв'язує швидкість передачі сигналів U та інформації C при відсутності перешкод. Назвіть одиниці виміру цих величин.
28. При відсутності перешкод швидкість розмови по телефону становить $V = 20$ слів/хв, середня довжина слова – $k = 6$, а середня питома ентропія – $H = 2$ біт/літеру. Знайти швидкість передачі інформації.
29. Джерело інформації виробляє $m = 16$ рівноймовірних і незалежних повідомлень, кожне з яких кодується $k = 4$ двійковими символами. Тривалість двійкового символу становить $\tau = 1$ мс. Знайдіть швидкість передачі двійкових сигналів та інформації при відсутності перешкод.
30. Для передачі по каналу зв'язку неперервного випадкового сигналу він дискретизується у часі з інтервалом $\Delta t = 10$ мс, що забезпечує незалежність відліків, а кожний відлік квантується з перетворенням його в дискретний каналний сигнал з кількістю рівнів квантування $m = 64$. Знайдіть пропускну здатність каналу.
31. Назвіть чинники, які обмежують пропускну здатність каналу передачі інформації при відсутності перешкод.
32. Поясніть поняття симетричності двійкового каналу.
33. Доведіть, що при малих значеннях імовірності помилкового прийому сигналів p пропускну здатність двійкового симетричного каналу з перешкодами може бути обрахована за співвідношенням: $C_{\Pi}^{\Pi} = U \cdot [1 + p(\log p - 1)]$. Обчисліть її значення, якщо $U = 10^5$ Бод, $p = 2^{-7}$.
34. Побудуйте сімейство залежностей пропускну здатності двійкового симетричного каналу з перешкодами C_{Π}^{Π} від імовірності помилкового прийому сигналів p при варіації швидкості передачі сигналів U . Визначіть імовірності, при яких досягаються мінімальні та максимальні значення пропускну здатності.
35. Побудуйте сімейство залежностей пропускну здатності двійкового симетричного каналу з перешкодами C_{Π}^{Π} від швидкості передачі сигналів U при варіації імовірності помилкового прийому сигналів p .
36. Наведіть вираз для пропускну здатності каналу з перешкодами C_{Π}^{Π} , яке отримано при використанні енергетичного підходу та сформулюйте використані при цьому припущення.

37. Побудуйте сімейство залежностей пропускної здатності каналу з перешкодами C_{Π}^{Π} від відношення потужностей сигналу та перешкоди $\frac{P_C}{P_{\Pi}}$ при варіації ширини смуги частот каналу ΔF .

38. По каналу зв'язку з шириною смуги частот $\Delta F = 10^3$ Гц, у якому діє широкосмугова перешкода – білий шум з питомою спектральною щільністю $p_0 = 2 \cdot 10^{-6}$ Вт/Гц, передається сигнал потужністю $P_C = 4$ Вт. Обчисліть пропускну спроможність каналу. Визначте верхню межу пропускної здатності каналу з такими енергетичними характеристиками.

39. Знайдіть необхідну пропускну спроможність і ширину полоси частот монофонічного радіомовного каналу, якщо частота дискретизації аудіосигналу становить $f = 44,1$ кГц, кожний відлік звукового сигналу кодується $m = 16$ бітами, а відношення потужностей сигналу та перешкоди становить $\frac{P_C}{P_{\Pi}} = 10^3$.

40. Знайдіть необхідну пропускну спроможність і ширину полоси частот телевізійного каналу, якщо розмір телевізійного растру становить $x \cdot y = 800 \cdot 625$ крапок, частота кадрової розгортки $f = 25$ Гц, колір кожної крапки растру кодується $m = 8$ бітами, а відношення потужностей сигналу та перешкоди становить $\frac{P_C}{P_{\Pi}} = 10^3$.

41. Наведіть співвідношення для об'ємів сигналу та каналу, поясніть сенс їх складових та назвіть одиниці їх виміру.

42. Наведіть загальну та часткові умови узгодження сигналу й каналу.

43. Наведіть практичні приклади узгодження сигналів та каналів.

44. Виконайте узгодження сигналу і каналу з параметрами $T_K = 1$ с $T_C = 5$ с, $\Delta F_K = 100$ кГц, $\Delta F_C = 10$ кГц, $h_K = h_C$. Поясніть, яким чином технічно може бути виконано це узгодження.

45. Дайте визначення величинам статистичної та ймовірнісної надлишковостей повідомлень, поясніть, які ймовірнісні характеристики джерела повідомлень впливають на їх значення.

46. Дискретне джерело генерує незалежні повідомлення з імовірностями $\{0,5, 0,25, 0,125, 0,125\}$. Знайдіть надлишковість повідомлень.

47. Матриці умовних імовірностей повідомлень двійкового джерела $p(x_j/x_i)$ мають вигляд

$$\begin{vmatrix} 0,5 & 0,5 \\ 0,5 & 0,5 \end{vmatrix} \quad \begin{vmatrix} 0,25 & 0,75 \\ 0,75 & 0,25 \end{vmatrix} \quad \begin{vmatrix} 0,75 & 0,25 \\ 0,25 & 0,75 \end{vmatrix} \quad \begin{vmatrix} 0,9 & 0,1 \\ 0,9 & 0,1 \end{vmatrix}.$$

Знайдіть величини статистичної та ймовірнісної надлишковостей повідомлень.

48. Наведіть співвідношення для визначення коефіцієнту стиску повідомлення, поясніть у яких межах може змінюватися ця величина.

2. ОСНОВИ ТЕОРІЇ Й ПРАКТИКИ КОДУВАННЯ ПОВІДОМЛЕНЬ У СИСТЕМАХ ПЕРЕДАЧІ ТЕХНОЛОГІЧНОЇ ІНФОРМАЦІЇ

Навчальні цілі: формування знань щодо основ теорії кодування повідомлень в системах автоматики і системах передачі технологічної інформації та набуття навичок синтезу оптимальних, елементарних завадостійких, лінійних, циклічних та канальних кодів.

2.1. Загальні положення. Класифікація кодів

Кодування в галузі систем передачі інформації (СПІ) – це відображення множини дискретних повідомлень на множину елементів (літер), які у свою чергу відображаються елементарними сигналами переносниками інформації.

Згадані множини елементів називаються кодами або кодovими комбінаціями повідомлень, при цьому кожному повідомленню відповідає свій код.

Кодом також називають правила, за якими утворюються кодovі комбінації.

Кодування може виражатися також у переході від одного коду до іншого, побудованому за своїми власними правилами (наприклад, від нормального двійкового до двійково-десятькового).

Кодування повідомлень може переслідувати цілі:

а) одержання необхідної множини повідомлень при обмеженій кількості використовуваних елементів (об'ємі алфавіту, мінімум $m = 2$);

б) досягнення максимальної ефективності СПІ шляхом зниження надлишковості переданих повідомлень;

в) забезпечення необхідної вірогідності передачі інформації в каналах з перешкодами.

Коди в різних галузях додатків можуть класифікуватися за різними ознаками. У СПІ їми можуть бути:

1. Логіко-математичні основи побудови коду – коди алгебраїчні, комбінаторні й ін.

2. Стійкість до перекручувань (трансформацій) елементів коду під час передачі або зберіганні повідомлень – звичайні (незавадостійкі) та коди що виявляють, а також виправляють помилки.

3. Структура кодovих комбінацій – блокові й неперервні (ланцюгові) коди.

4. Ознака довжин кодovих комбінацій – рівномірні, у яких ця величина незмінна, і нерівномірні, у яких використовуються кодovі комбінації змінної довжини.

5. Основа коду – число використовуваних елементів алфавіту – двійкові, трійкові й ін.

6. Місце здійснення операції кодування в структурі СПІ – коди джерел повідомлень, коди видів повідомлень, канальні коди й ін.

7. Спосіб передачі елементів коду сигналами – послідовні (часові) і паралельні (частотні) або передані по багатопроводовій лінії (шині).

Є коди, що мають додаткові назви за прізвищами їхніх авторів – Хеммінга, Бергера, Хаффмана та ін.

Помітимо, що нами не розглядаються системи кодування (шифрування), які забезпечують обмеження несанкціонованого доступу до інформації.

На закінчення необхідно визнати, що найбільш потужний і ефективний код, створений не людиною, а еволюцією живої матерії на основі молекул спадкоємної речовини – ДНК, що забезпечує як збереження, так і видозміну життєвих форм на Землі.

2.2. Постановка задачі кодування в каналах без перешкод. Нерівність Крафта

З метою досягнення максимальної ефективності кодування повинні бути вирішені задачі:

1. Розробка такої структури кодових комбінацій, щоб їх можна було передавати без розділових елементів.

Це можливо, якщо короткі кодові комбінації не будуть початками довгих (такі коди називаються нероздільними).

2. Визначення довжин кодових комбінацій відповідно до ймовірностей появи повідомлень – більше ймовірні короткими кодовими комбінаціями, менш ймовірні – довгими.

У галузі зв'язку практична реалізація цієї ідеї знайшла в коді Морзе, де літери, що часто зустрічаються в англійських текстах передаються короткими сигналами, і навпаки (наприклад, буква «e» → «.» (крапкою), «s» → «- - -») (трьома тире).

Пояснимо, як може виглядати нероздільний код на наступному прикладі.

Нехай є три групи повідомлень – одне високоймовірне, декілька – із середньою ймовірністю появи, а решта, більша частина повідомлень – малоймовірні. Встановимо також, що основа коду $m = 4$ (літери a, b, c, d).

Закодуємо високоймовірне повідомлення S_1 літерою «a». Тоді середньоїмовірні повідомлення не повинні на першій позиції в кодових комбінаціях містити літеру «a». Припустимо, що такі кодові комбінації починаються з літери «b». Отже, кодові комбінації низькоймовірних повідомлень будуть починатися з літер «c» і «d».

Згідно цього одержимо коди повідомлень, що наведені в табл. 2.1.

Таблиця 2.1

$S_1 - a$	$S_2 - ba$	$S_3 - bb$	$S_4 - bc$	$S_5 - bd$
$S_6 - caa$	$S_7 - cab$	$S_8 - cac$	$S_9 - cad$	$S_{10} - cba$
$S_{11} - cbb$	$S_{12} - cbc$	$S_{13} - cbd$	$S_{14} - cca$	$S_{15} - ccb$
$S_{16} - ccc$	$S_{17} - ccd$			

Решта з 12-ти комбінацій коду, які починаються з літери «d», для повідомлень $S_{18} \dots S_{29}$ одержуємо аналогічно, як і для $S_{16} \dots S_{17}$, використовуючи повторення й перестановки на других і третіх позиціях кодів.

У загальному випадку співвідношення між основою коду m і кількістю кодових комбінацій різної довжини для нероздільного коду повинні задовольняти нерівності Крафта – Макміллана.

Отже, при основі коду m число l_1 одноелементних кодових комбінацій $l_1 \leq m$. Якщо $l_1 < m$, з'являється можливість одержання двохелементних комбінацій, число яких $l_2 \leq (m - l_1)m = m^2 - ml_1$.

Аналогічно для кодових комбінацій довільної довжини можна записати

$$l_3 \leq [(m - l_1)m - l_2] - m = m^3 - m^2l_1 - ml_2$$

·
·
·

$$l_z \leq m^z - m^{z-1}l_1 - m^{z-2}l_2 - \dots - m^2l_{z-2} - ml_{z-1}.$$

Ділення правої й лівої частини цього виразу на m^z й перегрупування його дає нерівність

$$m^{-1}l_1 + m^{-2}l_2 + \dots + m^{-z+1}l_{z-1} + m^{-z}l_z \leq 1. \quad (2.1)$$

Загальна кількість комбінацій такого коду дорівнює кількості повідомлень, що кодуються, і становитиме $N = \sum_{s=1}^z l_s$, де z – максимальна довжина кодової комбінації. Вираз (2.1) може бути зображений у вигляді

$$\sum_{j=1}^z \sum_{k=1}^{l_j} m^{-j} \leq 1, \quad (2.2)$$

де кожному члену подвійної суми відповідає одне повідомлення.

В припущенні, що повідомленню з номером i відповідає кодова комбінація з довжиною n_i , розташування членів суми (2.2) в порядку нумерації повідомлень дає остаточну нерівність

$$\sum_{i=1}^N m^{-n_i} \leq 1. \quad (2.3)$$

За допомоги останньої можна стверджувати, що якщо чисельні характеристики m й n_i усіх (або використовуваної частини) комбінацій даного

коду такі, що задовольняються (2.3), то код буде нероздільним. Вище наведена необхідність виконання цієї нерівності для нероздільного коду.

Нерівність Крафта – Макміллана не встановлює зв'язку між ймовірностями появи повідомлень і довжинами кодових комбінацій, що їм відповідають. Рішення цієї задачі було знайдено К. Шенноном у вигляді доказу наступної теореми.

2.3. Основна теорема кодування в каналах без перешкод і її наслідок

Основна теорема кодування в каналах без перешкод має вигляд: середня довжина кодової комбінації n_c не може бути менше частки від ділення середньої ентропії джерела $H(x)$ на максимальну інформативність елемента коду, $\log m$

$$n_c \geq \frac{H(x)}{\log m}. \quad (2.4)$$

При цьому середня довжина n_c для заданого розподілу ймовірностей появи повідомлень $p(x_i)$ і довжинах їхніх кодів n_i визначається як

$$n_c = \sum_{i=1}^N n_i p(x_i). \quad (2.5)$$

Для доказу теореми використовується нерівність

$$\sum_{i=1}^N p(x_i) \log \frac{p(x_i)}{q(x_i)} \geq 0. \quad (2.6)$$

У даній нерівності $q(x_i)$ – це будь-який розподіл імовірностей із числом складових, що дорівнює N . Справедливість (2.6) доводиться на основі методу невизначених множників Лагранжа. При цьому розв'язується задача оптимізації цільової функції

$$W(q(x_i)) = \sum_{i=1}^N p(x_i) \log \frac{p(x_i)}{q(x_i)}, \quad (2.7)$$

при обмеженні $\sum_{i=1}^N q(x_i) - 1 = 0$. Функція Лагранжа та її часткові похідні для цієї задачі мають вигляд

$$L(q(x_i), \lambda) = \sum_{i=1}^N p(x_i) \log \frac{p(x_i)}{q(x_i)} + \lambda \left(\sum_{i=1}^N q(x_i) - 1 \right); \quad (2.8)$$

$$\frac{\partial L(q(x_i), \lambda)}{\partial q(x_i)} = -\frac{p(x_i)}{q(x_i)} + \lambda.$$

Оптимум цільової функції досягається, коли всі часткові похідні функції Лагранжа дорівнюють нулю, що можливо тільки за умови $q(x_i) = p(x_i)$. При цьому $W(q(x_i)) = 0$. В тому, що оптимум є мінімумом, тобто цільова функція (2.7) є невід'ємною і виконується нерівність (2.6), можна переконатися, узявши в якості ймовірнісного розподілу $q(x_i)$ будь-який відмінний від $p(x_i)$, наприклад, рівномірний

$$\sum_{i=1}^N p(x_i) \log \frac{p(x_i)}{q(x_i)} = \sum_{i=1}^N p(x_i) \log (Np(x_i)) = \log N \sum_{i=1}^N p(x_i) - \sum_{i=1}^N p(x_i) \log (p(x_i)) = \log N - H(X) \geq 0. \quad (2.9)$$

Нерівність (2.6) можна записати у вигляді

$$\sum_{i=1}^N p(x_i) \log p(x_i) \geq \sum_{i=1}^N p(x_i) \log q(x_i). \quad (2.10)$$

Ліва частина цієї нерівності – це середня ентропія $H(X)$ зі зворотним знаком. Тому (2.10) записується у вигляді

$$H(X) \leq -\sum_{i=1}^N p(x_i) \log q(x_i). \quad (2.11)$$

Задамо розподіл $q(x_i)$ у вигляді

$$q(x_i) = \frac{m^{-n_i}}{\sum_{j=1}^N m^{-n_j}}. \quad (2.12)$$

Тут чисельник дробове число, а знаменник – це сума дробових чисел при $j=1\dots N$, яка згідно (2.3) дорівнює одиниці або менше її. Тоді права частина (2.11) набуває вигляду

$$-\sum_{i=1}^N p(x_i) \log \frac{m^{-n_i}}{\sum_{j=1}^N m^{-j}} = \sum_{i=1}^N n_i p(x_i) \log m + \sum_{i=1}^N p(x_i) \log \sum_{j=1}^N m^{-j}. \quad (2.13)$$

Перший доданок правої частини складає

$$\sum_{i=1}^N n_i p(x_i) \log m = \log m \sum_{i=1}^N n_i p(x_i) = n_c \log m.$$

У другому доданку співмножник згідно нерівності Крафта є недодатним

$$\log \sum_{j=1}^N m^{-n_j} \leq \log(1) = 0,$$

тобто цей доданок менше або дорівнює нулю. Отже, якщо їм зневажити, нерівність (2.11) лише підсилиться й остаточно можна записати

$$H(x) \leq n_c \log m,$$

що й відповідає твердженню теореми (2.4). Доведена теорема носить назву теорема Шеннона.

Її наслідком є часткові нерівності для визначення довжин кодів повідомлень вигляду

$$n_i \geq \frac{H(x_i)_\Pi}{\log m} = \frac{-\log p(x_i)}{\log m}. \quad (2.14)$$

Справедливість цих залежностей можна підтвердити шляхом їх підстановки у рівняння (2.5), внаслідок чого буде отримано нерівність (2.4).

Оскільки довжини n_i є цілочисельними, (2.11) слід записати у вигляді

$$n_i = E \left[\frac{-\log p(x_i)}{\log m} \right], \quad (2.15)$$

де E – символ округлення результату до найближчого більшого цілого.

Після розв'язання цієї задачі залишається невирішеним питання алгоритмів побудови нероздільних кодових комбінацій при заданій їхній кількості N й розподілі ймовірностей їхньої появи $p(x_i)$. Тут відомі два теоретичних алгоритми й кілька прикладних для побудови нероздільних кодів. Розглянемо реалізацію теоретичних алгоритмів.

2.4. Алгоритм Шеннона – Фено

Під час реалізації алгоритму спочатку повідомлення розташовуються згідно з порядком спадання їхніх імовірностей (табл. 2.2).

Таблиця 2.2

Повідомлення x_i	Ймовірність $p(x_i)$	Номер кроку ділення	Код Шеннона – Фено			Нормальний двійковий код	
			I	II	III		
x_1	1/2	I — II — III —	0			0	0
x_2	1/4		1	0		0	1
x_3	1/8		1	1	0	1	0
x_4	1/8		1	1	1	1	1

Далі алгоритм виконується покроково.

На першому кроці вся множина повідомлень розділяється на « m » груп I-го порядку з максимально близькими за можливостями сумарними ймовірностями груп. За прикладом $m = 2$ і відповідно у першій групі буде повідомлення x_1 , в другій – x_2 , x_3 і x_4 . На першому кроці встановлюються значення кодових символів коду Шеннона – Фено (КШФ) – вище риси ділення – один символ, наприклад, «0», нижче інший – «1».

Другий крок виконується одночасно вище й нижче риси ділення I-го порядку. При цьому ділення повинні задовольняти вимогам відповідних сумарних імовірностей, як і на першому кроці.

Наступні кроки алгоритму виконуються аналогічно до тих пір, поки в кожній групі не залишається по одному повідомленню.

За прикладом на першому кроці кодування повідомлення x_1 закінчується. На другому одержуємо групи повідомлень – x_2 з одного боку та x_3 , x_4 з іншого і аналогічно встановленому вище правилу призначаються другі елементи КШФ.

Розглянутий варіант алгоритму є наближеним. Для реальних імовірностних розподілів і впорядкованих повідомлень сумарні ймовірності груп під час чергового ділення можуть розрізнятися досить суттєво. Тому при кожному діленні бажано розглядати всі можливі варіанти ділення множини повідомлень на групи і вибирати з них спосіб з максимально близькими сумарними ймовірностями груп. Однак, враховуючи, що кількість варіантів розбиття множини з N елементів на дві непорожні групи становить $2^{N-1} - 1$, виконання цієї вимоги при розрахунках вручну є трудомістким.

При розподілі ймовірностей за негативними ступенями числа m (тут $\frac{1}{2^1}$, $\frac{1}{2^2}$ і $\frac{1}{2^3}$) значення довжин КШФ виходять цілочисельними, без округлення, а інформаційне навантаження на елемент кодової комбінації є максимальним.

$$H(x)_{\text{КШФ}} = -\sum_{i=1}^N \frac{1}{n_i} p(x_i) \log p(x_i) =$$

$$= -\frac{1}{1} \cdot \frac{1}{2} \log \frac{1}{2} - \frac{1}{2} \cdot \frac{1}{4} \log \frac{1}{4} - 2 \cdot \frac{1}{3} \cdot \frac{1}{8} \log \frac{1}{8} = 1 \text{ біт/ел.}$$

Отриманий за прикладом КШФ є оптимальним, тому що його надлишковість згідно (1.98) дорівнює нулю.

Порівняємо ефективність цього коду з нормальним двійковим (НДК), довжини кодових комбінацій якого незмінні.

Довжина НДК становить

$$n = E[\log N] = \log 4 = 2$$

(про це свідчать і дані табл. 2.2), інформаційне навантаження на елемент кодової комбінації

$$H(x)_{\text{НДК}} = -\frac{1}{n} \sum_{i=1}^N p(x_i) \log p(x_i) = -\frac{1}{2} \left[\frac{1}{2} \log \frac{1}{2} + \frac{1}{4} \log \frac{1}{4} + \frac{2}{8} \log \frac{1}{8} \right] = \frac{7}{8} \text{ біт/ел.},$$

а його надлишковість

$$D = 1 - \frac{H(x)_{\text{НДК}}}{H(x)_{\text{max}}} = 1 - \frac{\frac{7}{8}}{1} = \frac{1}{8}.$$

Таким чином, проведені розрахунки підтверджують ефективність нерівномірного нероздільного коду в порівнянні з рівномірним.

2.5. Алгоритм Хаффмана

Як і в алгоритмі КШФ, повідомлення розташовуються згідно з порядком спадання ймовірностей їхньої появи (рис. 2.1). Алгоритм Хаффмана може бути виконаний у табличній або графічній формі. На практиці у ролі складової частини алгоритмів стиску растрових графічних даних широко використовується алгоритм Хаффмана з фіксованою таблицею. Під час його реалізації повідомлення ранжуються за зменшенням імовірностей їх появи, а потім повідомленню з заданим рангом ставиться у відповідність кодова комбінація, що обумовлена таблицею.

При невеликому числі повідомлень наочним є графічний варіант побудовання коду Хаффмана у вигляді кодового дерева. Тут імовірності повідомлень призначаються вихідними вершинами графа алгоритму. На першому кроці « t » вершин з найменшими ймовірностями з'єднуються дугами з першою проміжною вершиною із сумарною ймовірністю об'єднання. Так, згідно рис. 2.1 при $t = 2$ це ймовірності $p(x_6)$ й $p(x_7)$ появи повідомлень x_6 і x_7 й сумарною ймовірністю 0,06. Дуги графа позначаються елементами коду, для $t = 2$ верхня дуга, наприклад, «0», нижня – «1».

Другий крок об'єднання повторює дії першого із включенням у число вихідних першої проміжної вершини і виключенням вершин, що об'єднані на попередньому кроці. У результаті утворюється друга проміжна вершина, що розглядається черговою, що приєднується до числа інших вихідних.

Наступні кроки алгоритму виконуються аналогічно до тих пір, поки всі вихідні ймовірності не будуть об'єднані. Остаточний код повідомлення отримується як послідовність елементів коду по дугах від кінцевої вершини до вершини даного повідомлення.

Як зазначено на рис. 2.1 код Хаффмана (КХ) є нероздільним, а довжини кодових комбінацій відповідають обчисленим за співвідношенням (2.14).

Так, наприклад, для повідомлення x_1

$$n_1 = E \left[\frac{-\log 0,25}{\log 2} \right] = E \left[-\log \frac{1}{4} \right] = 2,$$

а для x_7

$$n_7 = E[-\log 0,03] = E \left[-\log \frac{1}{33,3} \right] = 6.$$

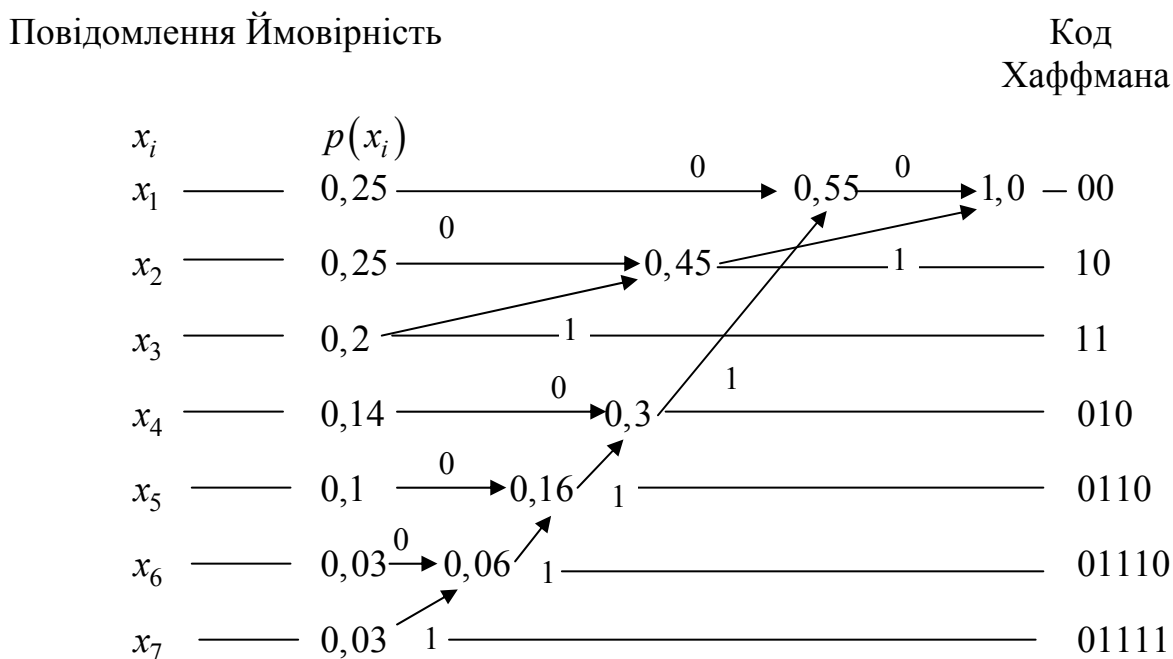


Рис. 2.1

Алгоритм Хаффмана дає $n_7 = 5$; це відображає той факт, що знаменник дроби в останній формулі практично не відрізняється від числа $2^5 = 32$, а вихідна формула для обчислення n_i є наближеною; у варіанті визначення n_i за алгоритмом може бути їхній зсув в одну або в іншу сторону на невелике значення.

Як і КШФ, КХ має меншу надлишковість, ніж відповідний йому за числом кодових комбінацій нормальний двійковий.

2.6. Незавадостійкі (звичайні) коди

2.6.1. Числоімпульсний (унітарний) код

Він відображає повідомлення A_i числом елементарних сигналів N_i , що формуються на певному інтервалі часу τ (пачка імпульсів), як це показано на рис. 2.3.

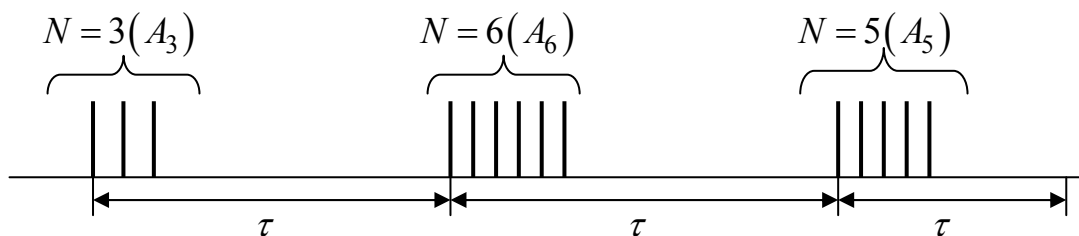


Рис. 2.3

Комбінації цього коду можуть мати числовий зміст, відображаючи, наприклад, значення параметрів технологічного процесу.

2.6.2. Нормальний двійковий код

Найбільше поширення одержав варіант нормального двійкового коду (НДК) 8-4-2-1 за десятковою вагою четвірки двійкових розрядів, починаючи зі старшого, $2^3 = 8$, $2^2 = 4$, $2^1 = 2$ і $2^0 = 1$. У загальному випадку десяткова вага n -розрядної комбінації НДК складає

$$N_{10} = \sum_{i=1}^n \alpha_i 2^{i-1}, \quad (2.16)$$

де i – номер розряду, починаючи з молодшого, $i = 1$ і α_i – значення i розряду НДК, 0 або 1.

НДК структурується за байтами – послідовними 8-розрядними частинами n -розрядної комбінації.

Іноді цей код розглядається як код на всі сполучення при $m = 2$. Дійсно, вся множина його комбінацій містить у собі підмножину комбінацій, що

містять одну «1», інші нулі, а їхня кількість дорівнює C_n^1 . Наступною є підмножина НДК на сполучення із двох «1» і $n - 2$ нулів з кількістю комбінацій C_n^2 , на сполучення по 3 і т.д., до останньої, єдиної комбінації із всіх «1», на сполучення C_n^n .

Таким чином, загальна кількість комбінацій НДК, крім нульової становить

$$N = \sum_{i=1}^n C_n^i = \sum_{i=0}^n C_n^i - 1 = (1+1)^n - 1 = 2^n - 1. \quad (2.17)$$

НДК у багатьох завадостійких кодах відображає повідомлення джерела, а організована надлишковість такого коду задається (обчислюється) за його інформаційною частиною, НДК повідомленням.

Для ряду таких кодів застосуємо матричний метод генерації їхніх комбінацій. Зокрема, і сам k -розрядний НДК породжує одинична діагональна матриця, що повернута на кут 90° за годинниковою стрілкою,

$$G_{(k,k)} = \begin{pmatrix} 00 & \dots & 01 \\ 00 & \dots & 10 \\ \cdot & & \cdot \\ \cdot & & \cdot \\ 01 & \dots & 00 \\ 10 & \dots & 00 \end{pmatrix}, \quad (2.18)$$

яка складається з k рядків і k стовпців. У самій матриці міститься k комбінацій коду C_k^1 . Інші $2^n - 1$ комбінацій НДК одержують за допомогою порозрядного підсумовування рядків $G_{(k,k)}$ по двох, усього C_k^2 , по трьох – C_k^3 і т.д. до підсумовування всіх рядків – C_k^k , що дає кодову комбінацію, яка складається з одиниць.

У СПІ технологічного призначення комбінації НДК можуть мати крім числового змісту, по вазі (2.16), так і символічне, для відображення дискретних повідомлень кількістю до $2^k - 1$.

2.6.3. Код Грея

Комбінації коду Грея (КГ) отримуються шляхом лінійних перетворень розрядів вихідного НДК

$$\begin{cases} \alpha_n^\Gamma = \alpha_n^H \\ \alpha_i^\Gamma = \alpha_i^H \oplus \alpha_{i+1}^H. \end{cases} \quad (2.19)$$

Таким чином, старший розряд у КГ α_n^Γ повторює старший розряд НДК, α_n^H , а інші є сумами за модулем 2 поточного і старшого відносно поточного розрядів НДК. Зворотний перехід (перекодування) здійснюється згідно залежностей

$$\begin{cases} \alpha_n^H = \alpha_n^\Gamma \\ \alpha_i^H = \alpha_i^\Gamma \oplus \alpha_{i+1}^H. \end{cases} \quad (2.20)$$

Особлива властивість КГ може бути показана, наприклад, на послідовності комбінацій НДК і відповідних їм комбінацій КГ

НДК ...0101 0110 0111 1000
КГ ...0111 0101 0100 1100.

З їх розгляду випливає, що при переході від однієї комбінації КГ до сусідньої в ній змінюється значення лише одного розряду, тоді як у відповідних комбінаціях НДК можуть змінюватися кілька розрядів. Кодами Грея відображаються повідомлення джерел вимірювальних перетворювачів координат, такти перемикачів у так званих кінцевих автоматах, згадана особлива властивість КГ використовується також при каналному кодуванні, завдяки чому додатково поліпшується завадостійкість СПП.

2.6.4. Двійководесятковий код

В двійководесятковий код (ДДК) десяткові розряди числа відображаються четвірками двійкових розрядів тетрад (рис. 2.4). Тут показана структура ДДК із трьома десятковими розрядами; кожна тетрада представляє собою одну з десяти комбінацій НДК – від 0000 до 1001 (інші 6 комбінацій у коді не використовуються).

$$\begin{array}{ccc} \text{Старша тетрада} & & \text{Молодша тетрада} \\ \underbrace{\alpha_{43}\alpha_{33}\alpha_{23}\alpha_{13}}_{N_{10}\cdot 10^2} & \underbrace{\alpha_{42}\alpha_{32}\alpha_{22}\alpha_{12}}_{N_{10}\cdot 10^1} & \underbrace{\alpha_{41}\alpha_{31}\alpha_{21}\alpha_{11}}_{N_{10}\cdot 10^0} \end{array}$$

Рис. 2.4

Тетради, починаючи з молодшої мають десяткові множники ваги 1, 10, 100 і т.д. Вага ж ДДК – сума десяткових ваг чисел, що записані в тетрадах. Так, наприклад, ДДК 1000.0110.0001 має десяткову вагу

$$N_{10} = 8 \cdot 100 + 6 \cdot 10 + 1 \cdot 1 = 861.$$

Розглянутий код знаходить застосування в пристроях ручного уведення й відображення десяткових чисел на цифрових відтворюючих приладах.

На закінчення відзначимо, що хоча ДДК і має надлишковість, але вона марна, тому що не служить для підвищення вірогідності передачі повідомлень у каналах з перешкодами.

2.7. Загальні властивості кодів, що виявляють і виправляють помилки

Для таких кодів властива надлишковість, яка проявляється у тому, що в їх кодових комбінаціях окрім інформаційних розрядів, які кодують первинні повідомлення, присутні контрольні (перевірочні) розряди. Останні використовуються для контролю правильності інформаційних і виправлення наявних в них помилок. Коди, в яких існує чітке визначення місцезнаходження інформаційних і контрольних розрядів, зветься систематичними.

2.7.1. Геометричне подання комбінацій двійкових кодів

Залежно від розрядності НДК його комбінацій можуть відображатися в такий спосіб:

а) при числі розрядів $n=1$ геометрична модель – лінійний граф з вершинами 0 та 1, з'єднані дугою одинично довжини $d=1$ (рис. 2.5, а);

б) якщо $n=2$, маємо плоский граф з вершинами 00, 01, 11 і 10; його суміжні вершини розташовані на відстані $d=1$ (рис. 2.5, б);

в) для трирозрядного коду геометрична модель є просторовою, так званий одиничний куб (рис. 2.5, в), або розгортку – плоский граф (рис. 2.5, г);

г) при $n=4$ існує штучна просторова модель коду – гіперкуб; якщо $n > 4$, подання двійкового коду можливе лише у формі графів.

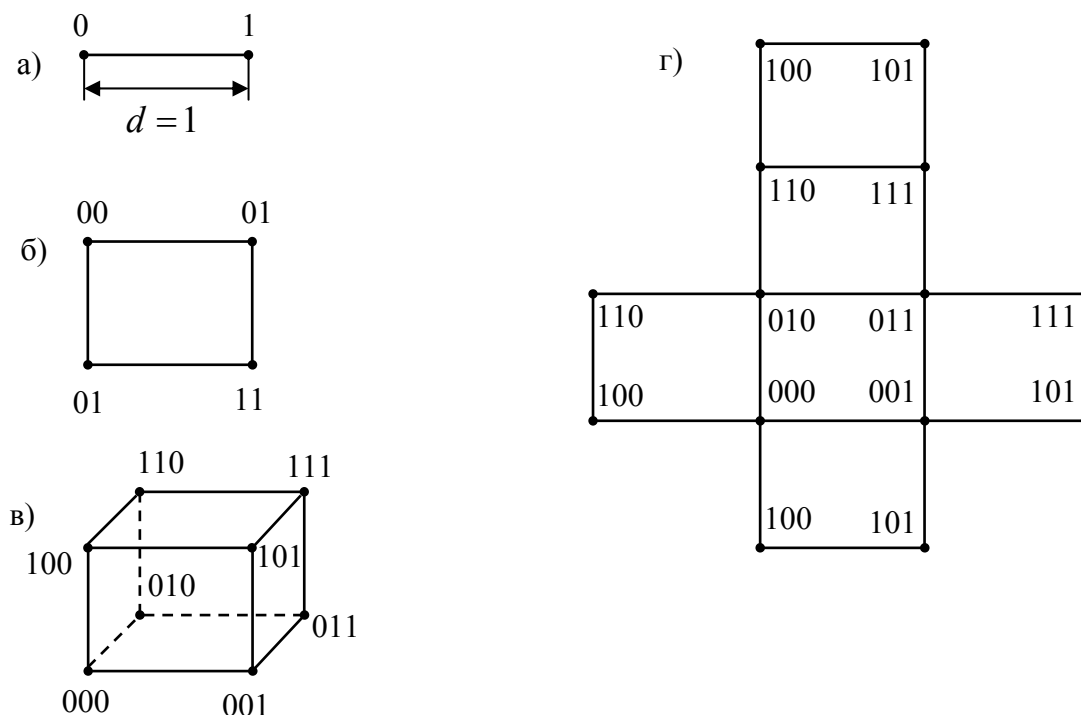


Рис. 2.5.

Властивості коду щодо виявлення й виправлення помилок ґрунтуються на понятті відстані між його комбінаціями (відстані Хеммінга). Під відстанню d між двома комбінаціями мається на увазі кількість їх відповідних розрядів, що мають різні значення. Для двійкових кодів ця відстань може бути знайдена як число одиниць у порозрядній сумі за модулем 2 кодових комбінацій, що розглядаються. Наприклад,

$$\oplus \begin{array}{r} 011010 \text{ I комбінація} \\ 101011 \text{ II комбінація} \\ \hline 110001 \quad d=3. \end{array}$$

2.7.2. Виявлення помилок у кодових комбінаціях

Для досягнення поставленої мети вся множина комбінацій НДК розділяється на дві підмножини – дозволених, із числом комбінацій N_0 , і заборонених – із числом N_3 , так що $N = 2^n = N_0 + N_3$.

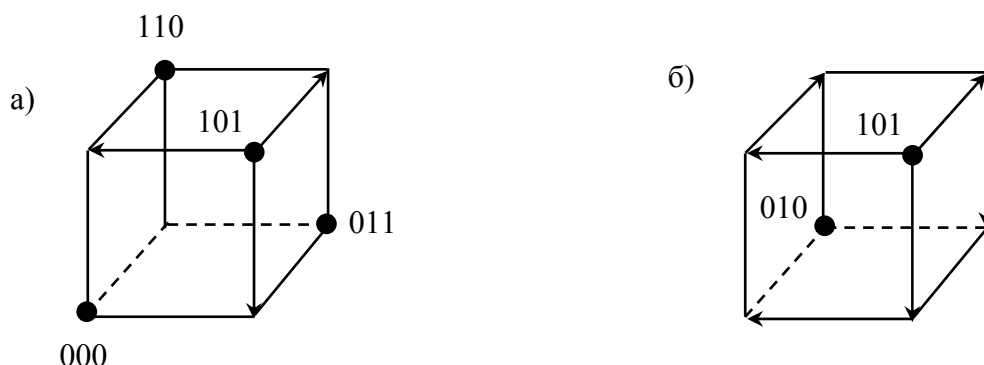
Дозволеними кодовими комбінаціями передаються повідомлення, а заборонені – віддаляють дозволених друг від друга на відстані, які принаймні складають величину

$$d_{\min} = r + 1, \quad (2.21)$$

де r – кратність виявлених помилок, на яку розраховується даний код. Тобто величина d_{\min} визначається як мінімальна відстань між дозволеними комбінаціями і носить тут назву кодової відстані.

Так, при $n=3$, $r=1$ і $d_{\min}=2$ на моделі (рис. 2.5, в) дозволеними комбінаціями можуть бути, наприклад, ті, які позначені на рис. 2.6, а) жирними крапками.

Рис. 2.6



Стрілками показано, як одна з дозволених комбінацій – 101 однократними помилками переводиться в одну із заборонених – 100, 001 або 111. У цьому прикладі $N_0 = N_3 = 4$; виявлення однократних помилок досягається ціною надлишковості $D = 0,33$.

Взагалі надлишковість коду згідно (1.103) становить

$$D = 1 - \frac{H}{H_{\max}} = 1 - \frac{\log N_0}{\log N}, \quad (2.22)$$

яка для систематичного коду може бути записана у вигляді

$$D = 1 - \frac{\log N_0}{\log N} = 1 - \frac{\log 2^k}{\log 2^n} = 1 - \frac{k}{n} = \frac{m}{n}, \quad (2.23)$$

де k та m – кількість інформаційних та контрольних розрядів відповідно

Вимога виявлення дворазових помилок на моделі рис. 2.6, а залишає серед дозволених лише 2 комбінації, які розташовуються на одній з головних діагоналей куба, наприклад 010 і 101. Очевидно, що для них $d_{\min} = 3$, що відповідає умові (2.21) при $r = 2$. Стрілками показано, як одна з дозволених комбінацій – 101 двократними помилками переводиться в одну із заборонених – 000, 011 або 110. Надлишковість цього коду природно більше, ніж коду при $r = 1$ і становить 0,67.

Загальна умова існування коду, що виявляє помилки заданої кратності r полягає у тому, що останні не повинні переводити одну дозволена комбінація в будь-яку іншу дозволена. При цьому належність комбінації до підмножини заборонених свідчить про наявність в ній принаймні однієї помилки.

На стороні прийому декодер будується таким чином, що він сприймає лише дозволена кодові комбінації. На появу будь-якої недозволена комбінації декодер видає узагальнений сигнал «помилка».

2.7.3. Виявлення й виправлення помилок у кодових комбінаціях

Як і у варіанті лише виявлення помилок вся множина комбінацій НДК розділяється на підмножини дозволених і заборонених комбінацій. Будь-яка пара комбінацій із числа дозволених повинна знаходитись на відстані, яка принаймні складає величину

$$d_{\min} = s + r + 1, \quad (2.24)$$

де s – кратність виправлених помилок із числа виявлених r .

Найбільш ощадливими кодами цього класу є такі, у яких $r = s$ і тоді (2.22) приймає вигляд:

$$d_{\min} = 2s + 1. \quad (2.25)$$

Загальна умова існування розглянутого коду полягає у тому, що заборонені підмножини комбінацій, що відповідають двом будь-яким з дозволених, не повинні перетинатися.

Приклад такого коду при $n = 3$ й $s = 1$ ($d_{\min} = 3$) наведений на рис. 2.6, б. У якості дозволених тут призначені комбінації 010 і 101. Першій з них при однократних помилках відповідають комбінації 110, 000, 011, а другій – 001, 111 і 100; жодна із заборонених кодових комбінацій однієї підмножини не міститься в іншій.

При одержанні забороненої кодової комбінації декодер приймача на основі встановленого d_{\min} визначає найближчу дозволена комбінацію, що і буде результатом виправлення комбінації з помилками.

Загальної теорії розглянутих кодів не існує, також немає для них єдиних алгоритмів кодування й декодування. Однак, як це було згадано раніше, розроблені коди, які забезпечують будь-яку необхідну вірогідність передачі повідомлень у каналах зв'язку з перешкодами.

Далі розглядаються такі коди, у тому числі ті, що знайшли застосування в СПІ технологічного призначення.

2.8. Найпростіші коди, що виявляють помилки

В основу цих кодів покладені елементарні логічні міркування, що не вимагають спеціального математичного обґрунтування. У той самий час, деякі з них, наприклад, код «пар», який є найпростішим із класу лінійних систематичних, мають теоретичне обґрунтування здатності виявляти й виправляти помилки.

Властивості частини цих кодів визначаються в припущенні випадковості й взаємозалежності появи помилок у кодових комбінаціях.

При ймовірності помилки прийому елемента коду p , ймовірність його правильного прийому дорівнює $1 - p$. Ймовірність того, що кодова комбінація довжиною n прийнята правильно становить

$$q_k = (1 - p)^n, \quad (2.26)$$

а помилково

$$p_k = 1 - q_k = 1 - (1 - p)^n. \quad (2.27)$$

У СПІ технологічного призначення нормується припустиме значення ймовірності помилки, яке дорівнює 10^{-4} . При невеликій довжині кодових комбінацій у цих умовах ймовірність помилкового прийому кодової комбінації може бути визначена як

$$p_k \approx n \cdot p. \quad (2.28)$$

2.8.1. Код «пар» або код із захистом за паритетом

Є систематичним. У цьому коді k -розрядна інформаційна частина в НДК доповнюється контрольним розрядом, у якому записується «0» або «1». В комбінації цього коду довжиною $n = k + 1$ завжди міститься парне число одиниць. Іноді використовується варіант такого коду – «непар», з непарним числом одиниць. Властивості цих варіантів коду рівноцінні.

Цей код має $d_{\min} = 2$, виявляє всі помилки непарної кратності і не виявляє помилки парної кратності.

При випадкових і взаємнезалежних помилках ймовірність виникнення l помилок у кодовій комбінації довжиною n підпорядковується розподілу Бернуллі (1.26).

Сумарна ймовірність помилок, що виявляються становить

$$p_B = \sum_{i=0}^{+\infty} p^{2i+1}, \quad (2.29)$$

а тих, що не виявляються

$$p_H = \sum_{i=1}^{+\infty} p^{2i}. \quad (2.30)$$

При ймовірності трансформації розряду коду $p = 10^{-4}$ й менше, а також невеликих n третій співмножник у формулі Бернуллі близький до одиниці. Тоді суми (2.29), (2.30) можуть бути записані відповідно як

$$p_B \approx C_n^1 p + C_n^3 p^3 + \dots, \quad (2.31)$$

$$p_H \approx C_n^2 p^2 + C_n^4 p^4 + \dots. \quad (2.32)$$

Другі доданки в цих формулах менше перших, тому ними (і іншими в сумах) можемо зневажити.

Уведемо поняття ефективності коду що виявляє помилки

$$\eta = \frac{p_B}{p_H}. \quad (2.33)$$

Для коду «пар» вказана величина становить

$$\eta \approx \frac{C_n^1 p}{C_n^2 p^2} = \frac{\frac{n!}{(n-1)! \cdot 1!} \cdot p}{\frac{n!}{(n-2)! \cdot 2!} \cdot p^2} = \frac{2}{p(n-1)}. \quad (2.34)$$

Оцінимо отриманий результат для значень $p = 10^{-4}$ і $n = 9$.

$$\eta = \frac{2}{10^{-4}(9-1)} = \frac{2500}{1}.$$

Це значить, що приблизно на 2500 помилок, що виявляються, у кодї «пар» буде припадати одна невиявлена.

Надлишковість коду згідно (2.23) становить

$$D = 1 - \frac{n-1}{n} = \frac{1}{n}. \quad (2.35)$$

Код «пар» використовується в каналах з низьким рівнем перешкод і передачі сигналів на невеликі відстані.

2.8.2. Коди з постійною вагою (коди на одне сполучення)

Є несистематичними. Дані коди довжиною n мають постійну вагу (число одиниць у комбінації) w і утворюються за правилами сполучень. Число комбінацій такого коду складає

$$N_0 = C_n^w = \frac{n!}{(n-w)!w!}. \quad (2.36)$$

Деяке поширення має код C_n^1 , число комбінацій якого $N_0 = n$. Сам по собі цей код неефективний, але може являти собою інформаційну частину інверсного коду (див. п.2.8.3).

Більшу популярність мають коди C_5^2 , $N_0 = 10$ і C_7^3 , $N_0 = 35$. Перший з них відображає десяткові цифри, а другий – букви алфавітів і символи.

Як приклад приведемо комбінації коду C_5^2 :

«1» – 00011, «2» – 00101, «3» – 01001, «4» – 10001, «5» – 00110
«6» – 01010, «7» – 10010, «8» – 01100, «9» – 10100, «10» – 11000.

Ефективність коду на одне сполучення може бути визначена в такий спосіб. Очевидно, що код C_n^w виявляє всі помилки, які дають комбінації із числом одиниць, не рівним w .

Помилки найменшої кратності, що не виявляються – парний зсув двох розрядів, в одному йз яких відбувається зсув $1 \rightarrow 0$, а в іншому – $0 \rightarrow 1$.

Ймовірність такої події, на підставі викладеного вище визначається як

$$p_H \approx C_w^1 \cdot C_{n-w}^1 \cdot p^2, \quad (2.37)$$

а ймовірність виявлення помилки згідно (2.28) складає

$$p_B \approx p_k - p_H \approx n \cdot p - C_w^1 \cdot C_{n-w}^1 \cdot p^2. \quad (2.38)$$

Остаточно ефективність коду на одне сполучення становить

$$\eta = \frac{p_B}{p_H} \approx \frac{n \cdot p - C_w^1 \cdot C_{n-w}^1 \cdot p^2}{C_w^1 \cdot C_{n-w}^1 \cdot p^2}, \quad (2.39)$$

що вище, ніж у коду «пар» з тим же числом розрядів n . Але коди на одне сполучення мають більшу надлишковість. Так, у кодї «пар» число дозволених комбінацій складає біля половини, а, наприклад, для коду C_5^2 – третину.

2.8.3. Кореляційний, інверсний і коди з контрольними сумами

2.8.3.1. Кореляційний код (КК)

Отримується із НДК заміною одного розряду останнього на пару розрядів за правилом: $0 \rightarrow 01, 1 \rightarrow 10$. Таким чином, комбінація НДК 001011 перетвориться в комбінацію 010110011010 КК.

Довжина даного коду $n = 2k$, де k розрядів НДК являють собою код повідомлення або інформаційну частину кореляційного коду.

Помилки нижньої кратності, що не виявляються, у КК – це одночасна трансформація двох його розрядів, що відображують будь-який розряд НДК.

Аналогічно співвідношенням (2.37) – (2.39) для КК можна записати

$$\begin{aligned} p_H &\approx C_k^1 \cdot p^2 \\ p_B &\approx n \cdot p - C_k^1 \cdot p^2 \\ \eta &= \frac{n \cdot p - C_k^1 \cdot p^2}{C_k^1 \cdot p^2}. \end{aligned} \quad (2.40)$$

Надлишковість КК складає $D = 0,5$. Власно кореляційний код у цей час не становить інтересу, але є основою для порівняння з інверсним кодом, котрий завдяки більш раціональному алгоритму кодування при тій же надлишковості набагато перевершує КК по здатності виявлення помилок.

2.8.3.2. Інверсний код (ІК)

Є систематичним, у якому НДК повідомлення доповнюється контрольною частиною тої самої довжини, що й інформаційна, тому довжина ІК становить $n = 2k$.

Контрольна частина утвориться за правилом: якщо інформаційна частина (НДК) містить парне число одиниць, то контрольна повторює її, а при непарному числі одиниць в інформаційній частині, контрольна частина утворюється внаслідок порозрядної інверсії інформаційної частини повідомлення.

Наприклад, НДК 0101 відповідає ІК 01010101, а НДК 0111 – ІК 01111000.

Аналіз показує, що розглянутий код має $d_{\min} = 4$. Помилками нижньої кратності, що не виявляються, у комбінаціях КК є подвійні зсуви пари розрядів, що відповідають тим самим розрядів НДК повідомлення. Наприклад, у комбінації 01111000 трансформуються розряди, відзначені «*». У результаті одержуємо комбінацію 11100001, що відноситься до підмножини дозволених, але відображує помилкове повідомлення.

Співвідношення (2.40) для ІК набувають вигляду

$$\begin{cases} p_H \approx C_k^2 \cdot p^4 \\ p_B \approx n \cdot p - C_k^2 \cdot p^4 \\ \eta = \frac{n \cdot p - C_k^2 \cdot p^4}{C_k^2 \cdot p^4} \end{cases} \quad (2.41)$$

Надлишковість ІК складає $D \approx 0,5$ (комбінація, що складається з одних нулів у цьому коді відсутня). Порівняння p_{H0} КК і ІК показує, що в останнього ця ймовірність на кілька порядків нижче, що суттєво для таких повідомлень, як команди управління в технологічних системах.

2.8.3.3. Коди з контрольними сумами (коди Бергера)

У цих кодах k -розрядна інформаційна частина доповнюється $\mu = E[\log k]$ розрядами «контрольної суми», що може визначатися різними способами залежно від властивостей каналу зв'язку. Так, для асиметричного каналу, де співвідношення ймовірностей трансформацій сигналів $p(0/1) > p(1/0)$, у контрольній частині коду записується інверсія двійкового числа, вага якого дорівнює числу одиниць інформаційної частини.

Так, наприклад, якщо код повідомлення 1101101 ($k = 7, \mu = 3$), то код Бергера (КБ) буде отриманий у такий спосіб. Число одиниць в інформаційній частині дорівнює 5, а відповідне двійкове число, що відповідає вазі – 101. Отже, контрольна частина буде 010, а код у цілому 1101101010.

При декодуванні прийнятої кодової комбінації аналогічно, як і при кодуванні обчислюється контрольна частина й вона порівнюється з переданою. Перехід будь-якого числа одиниць у нулі в КБ приводить до того, що десяткове

значення обчисленої контрольної частини буде менше, ніж відповідне значення інвертованої переданої.

Нехай отримана вище кодова комбінація прийнята з потрібною помилкою й має вигляд 1000100010 . Обчислена за інформаційного контрольна частина становитиме 011 (3), а інвертована передана – 101 (5). Таким чином помилки виявлені. У загальному випадку нерівність контрольних частин є ознакою помилок.

При появі помилок вигляду $0 \rightarrow 1$ ефективність КБ зменшується, зокрема за рахунок невиявляємих помилок нижньої кратності – зсуву двох розрядів інформаційної частини коду.

КБ одержали деяке поширення в локальних комп'ютерних мережах. У каналах зв'язку великої довжини, зокрема каналах систем телемеханіки ці коди не використовуються.

2.9. Лінійні коди

2.9.1. Загальні відомості. Постановка задачі кодування з виявленням і виправленням помилок

Клас розглянутих кодів був відкритий на початку 1940-х років Хеммінгом, коли виникла потреба підвищення вірогідності обчислень на перших ЕОМ.

Математичний апарат цих кодів – лінійна алгебра, засобами якої можливо забезпечувати задане d_{\min} між кодовими комбінаціями дозволеної підмножини. Для цього в НДК повідомлень вводяться контрольні розряди, значення яких обчислюються підсумовуванням за модулем 2 певних інформаційних розрядів.

Хеммінг довів, що для $d_{\min} = 3$ при виправленні однократних помилок необхідно виконати умову

$$2^{n-k} = 2^{\mu} \geq n + 1. \quad (2.42)$$

Логіка цього виразу полягає в наступному: число кодових комбінацій контрольних розрядів не повинне бути менше, ніж довжина коду n (одночасно це число варіантів розміщення перекрученого розряду), а також варіант одержання неспотвореної кодової комбінації. Найбільш ощадливі лінійні коди виходять, якщо (2.40) буде рівністю $2^{\mu} = n + 1$, звідки число контрольних розрядів

$$\mu = E[\log(n + 1)]. \quad (2.43)$$

Для цілих n , μ і $k = n - \mu$ одержимо наступний ряд чисел

n	3	7	15	31	...
μ	2	3	4	5	...
k	1	4	11	26	...

Для більших значень d_{\min} отримані наступні значення чисел контрольних розрядів, зокрема

$$d_{\min} = 4 \quad \mu \geq E[\log(2 \cdot n)],$$

$$d_{\min} = 5 \quad \mu \geq E\left[\log\left(1 + n + \frac{n(n-1)}{2}\right)\right] \text{ і т.д.}$$

Є також узагальнені оцінки нижніх значень μ , наприклад, оцінка Хеммінга, яка може бути застосована при непарних значеннях d_{\min}

$$\mu \geq E\left[1 + \sum_{i=1}^{(d_{\min}-1)/2} C_n^i\right]. \quad (2.44)$$

По Хеммінгу контрольні розряди розміщуються на певних позиціях усередині n -розрядної кодової комбінації. При деяких перевагах такого рішення в нього є й недоліки, яких позбавлений різновид лінійних – систематичні коди (ЛСК), у яких контрольні розряди відокремлені й слідуєть за інформаційними.

2.9.2. Основні властивості ЛСК, матричний метод кодування повідомлень

Надалі структуру ЛСК уявляємо у вигляді блоку

$$x_1 x_2 \dots x_{k-1} x_k y_1 y_2 \dots y_{\mu-1} y_{\mu},$$

де x_i, y_j – відповідно « i » інформаційний і « j » контрольний розряди.

Контрольні розряди обчислюються як суми за модулем 2

$$y_j = c_{1j} \cdot x_1 \oplus \dots \oplus c_{ij} \cdot x_i \oplus \dots \oplus c_{kj} \cdot x_k, \quad (2.45)$$

$$c_{ij} \in \{0,1\}.$$

Загального методу вибору значень співмножників c_{ij} для обчислення конкретного розряду y_j не існує, але відомі критерії, яким повинні задовольняти суми (2.43).

Тому що інформаційна частина ЛСК – це НДК, який породжується матрицею (2.18), а контрольна – лінійні перетворення (2.43), ЛСК у цілому породжується твірною матрицею

$$G_{(n,k)} = \left| G_{(k,k)}, R_{(k,n-k)} \right|, \quad (2.46)$$

де $G_{(k,k)}$ – інформаційна, а $R_{(k,n-k)}$ – контрольна підматриці.

Рядки матриці $G_{(n,k)}$ є вихідними кодовими комбінаціями даного ЛСК.

Вихідні комбінації ЛСК мають такі властивості:

1. Нульова комбінація в твірній матриці відсутня, але в кодї є дозволеною.

2. Вага рядку (число одиниць у ньому) твірної матриці повинна бути не менш d_{\min} .

3. Порозрядна сума за модулем 2 будь-яких сполучень кодових комбінацій ЛСК також є комбінацією цього ЛСК.

4. Порозрядна сума за модулем 2 будь-яких сполучень кодових комбінацій ЛСК не повинна перетворюватися в нульову. Ця вимога забезпечує лінійну незалежність комбінацій ЛСК на основі відповідного вибору сум (2.45).

Як приклад розглянемо синтез ЛСК $G_{(7,4)}$, що при $d_{\min} = 3$ буде мати $k = 4$ інформаційних і $\mu = 3$ контрольних розрядів.

Один з варіантів обчислення значень контрольних розрядів по (2.45) дає наступні їхні значення

$$\begin{cases} y_1 = x_1 \oplus x_2 \oplus x_3 \\ y_2 = x_2 \oplus x_3 \oplus x_4 \\ y_3 = x_1 \oplus x_3 \oplus x_4. \end{cases} \quad (2.47)$$

Тоді твірна матриця даного ЛСК $G_{(7,4)}$ буде мати вигляд

$$G_{(7,4)} = \left| \begin{array}{ccccccc} 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ x_1 & x_2 & x_3 & x_4 & y_1 & y_2 & y_3 \end{array} \right|. \quad (2.48)$$

Тут y_1, y_2, y_3 для першого рядка матриці згідно (2.45) $y_1 = 0 \oplus 0 \oplus 0 = 0$, $y_2 = 0 \oplus 0 \oplus 1 = 1$ і $y_3 = 0 \oplus 0 \oplus 1 = 1$, у контрольних частинах інших рядків матриці значення y_j одержуються аналогічно.

Усього цей код, згідно (2.17), має $2^4 - 1 = 15$ ненульових кодових комбінацій, з яких 4, де в інформаційній частині є одна одиниця, записані у вигляді рядків твірної матриці. Решта комбінацій отримується внаслідок підсумовування за модулем 2 будь-яких сполучень рядків твірної матриці, як це було показано у п. 2.6.2 для НДК.

Оскільки d_{\min} ЛСК збігається з мінімальним числом одиниць у його комбінаціях, що є рядками твірної матриці (2.48), а в інформаційній частині рядка міститься одна одиниця, то в контрольних частинах рядків матриці $G_{(n,k)}$ повинно бути не менш $d_{\min} - 1$ одиниць.

Другим критерієм правильності вибору форм (2.46) є те, що у контрольній частині порозрядної суми за модулем 2 двох будь-яких рядків матриці $G_{(n,k)}$ повинно бути не менш $d_{\min} - 2$ одиниць.

Якщо в (2.34) усі коефіцієнти $c_{ij} = 1$, ЛСК переходить в код «пар» з єдиним контрольним розрядом $y_1 = \sum_{i=1}^k x_i$.

Автори – Голей, Малер, Макдональд, Варшамов і ін. запропонували різновиди ЛСК, які мають власні корисні коригувальні особливості.

2.9.3. Методи декодування ЛСК з виявленням або виявленням і виправленням помилок

У класі лінійних кодів ці методи ґрунтуються, зокрема на еквівалентності операцій додавання й вирахування в полі за модулем 2. Це приводить до того, що в сумах (2.45) доданки можуть переноситися з однієї частини в іншу без зміни знака. Вказані методи будуть розглянуті на конкретних прикладах.

2.9.3.1. Метод перевірочних рівнянь

Для коду $G_{(7,4)}$ безпосередньо з (2.47) за відсутності помилок можна одержати

$$\begin{cases} x'_1 \oplus x'_2 \oplus x'_3 \oplus y'_1 = 0 & I \\ x'_2 \oplus x'_3 \oplus x'_4 \oplus y'_2 = 0 & II \\ x'_1 \oplus x'_3 \oplus x'_4 \oplus y'_3 = 0 & III. \end{cases} \quad (2.49)$$

Знак «'» позначає, що мова йде про значення розрядів ЛСК, що зафіксовані на стороні прийому.

Особливо варто мати на увазі, що якщо мова йде про виявлення помилок (тут $r = 2$), то відрізнитися від переданих можуть не більше ніж два прийнятих розряди, а при виправленні – тільки один.

Ознакою наявності помилок є незадоволення хоча б одного з рівнянь (2.49).

Якщо поставлено задачу виправлення помилки, то перевіряється виконання всіх рівнянь системи (2.49). При цьому можливі варіанти помилок, при яких не задовольняються відповідні рівняння, мають вигляд:

Помилка в $x'_1 - I$ и III ;

Помилка в $x'_2 - I$ и II ;

Помилка в $x'_3 - I, II$ и III ;

Помилка в $x'_4 - II$ и III .

Помилка в одному з контрольних розрядів – I або II , або III не потребує виправлення, яке має сенс тільки щодо інформаційних розрядів.

Перевірки рівнянь (2.49) називають перевітками на парність. Якщо число одиниць у сумі парне, то це є частковим підтвердження правильності прийнятого коду даним рівнянням.

Якщо твірна матриця відома, то за її допомоги можуть бути отримані перевірочні рівняння (2.49) на основі перевірочної матриці, що складається з одиничної підматриці розміру $(\mu \times \mu)$ й приписаної ліворуч підматриці $(k \times \mu)$, що являє собою контрольну підматрицю $R_{(k,n-k)}$ коду, яку повернуто на 90° за годинниковою стрілкою. Для наведеного вище приклада перевірочна матриця має вигляд

$$M_{(7,4)} = \begin{vmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ x_1 & x_2 & x_3 & x_4 & y_1 & y_2 & y_3 \end{vmatrix}. \quad (2.50)$$

2.9.3.2. Метод синдромів

У цьому методі вводиться поняття вектора помилки, якщо вона однократна – то це n розрядна кодова комбінація, що містить одиницю на позиції трансформованого розряду, а на інших $n - 1$ позиціях – нулі.

Для розглянутого ЛСК $G_{(7,4)}$ вектор помилки $V_{\Pi} = 0100000$ буде відповідати трансформації елемента x_2 , а $V_{\Pi} = 0000001$ – y_3 . Перекручену кодову комбінацію ЛСК можна уявити як порозрядну суму за модулем 2 переданої кодової комбінації V й коду помилки

$$V' = V \oplus V_{\Pi}. \quad (2.51)$$

Тепер виконання перевірки на задоволення умов (2.47) для комбінації коду V' може проводитися роздільно для V (умови задовольняються) і V_{Π} (не задовольняються).

З використанням перевірконої матриці (2.50) у процесі обчислення сум (2.49) одержуються елементи синдрому S – коду помилки, що однозначно відповідає позиції перекрученого розряду

$$S = M \cdot V' = M \cdot (V \oplus V_{\Pi}) = M \cdot V \oplus M \cdot V_{\Pi} = 0 \oplus M \cdot V_{\Pi} = M \cdot V_{\Pi}. \quad (2.52)$$

Отриманий вираз свідчить про те, що вектор розрядів синдрому збігається зі стовцем перевірконої матриці, номером якого є номер помилкового розряду.

Так, розряди синдрому для $V_{\Pi} = 0001000$ (трансформований елемент x_4) будуть мати вигляд

$$\begin{cases} S_{\Pi 1} = x_{\Pi 1} \oplus x_{\Pi 2} \oplus x_{\Pi 3} \oplus y_{\Pi 1} = 0 \oplus 0 \oplus 0 \oplus 0 = 0 \\ S_{\Pi 2} = x_{\Pi 2} \oplus x_{\Pi 3} \oplus x_{\Pi 4} \oplus y_{\Pi 2} = 0 \oplus 0 \oplus 1 \oplus 0 = 1 \\ S_{\Pi 3} = x_{\Pi 1} \oplus x_{\Pi 3} \oplus x_{\Pi 4} \oplus y_{\Pi 3} = 0 \oplus 0 \oplus 1 \oplus 0 = 1. \end{cases} \quad (2.53)$$

Отже, синдром коду помилки 0001000 $S_{\Pi} = 011$. Для коду $G_{(7,4)}$ при однократних помилках усього буде 7 синдромів, які, як впливає з табл. 2.3, однозначно визначають позицію трансформованого розряду, а отже, і розряд, який варто виправити.

Таблиця 2.3

V_{Π}	1000000	0100000	0010000	0001000	0000100	0000010	0000001
S_{Π}	101	110	111	011	100	010	001

Недоліком цього методу є істотне збільшення кількості перевірок за синдромами при зростанні вимог не тільки до коригувальної, але й виявляючої здатності коду.

Так, якщо поставлено задачу виявлення не тільки однократних, але й двократних помилок, то додатково до синдромів (табл. 2.3) повинні бути обчислені й синдроми подвійних помилок, кількість яких дорівнюватиме

$$C_n^2 = C_7^2 = \frac{7!}{(7-2)!2!} = \frac{5!6 \cdot 7}{5!2} = 21.$$

На методі синдромів ґрунтується код Хеммінга, який при кодовій відстані $d_{\min} = 3$ забезпечує виправлення однократних помилок, а при $d_{\min} = 4$ – виправлення однократних та виявлення двократних помилок.

Розглянемо приклад побудови коду Хеммінга при $d_{\min} = 3$, $k = 4$, $n = 7$, $m = 3$. Будемо вважати, що в кодовій комбінації перші k розрядів є інформаційними, а ті, що залишилися – контрольними, хоча ця умова не є обов'язковою. Так, у оригінальному варіанті, який був запропонований

Хеммінгом, контрольні розряди розташовуються на позиціях з номерами, які є ступенями двійки. При декодуванні виконується три перевірки, внаслідок чого формується трирозрядний синдром помилки. Значення синдрому й відповідні номери помилкових розрядів наведені в табл. 2.4.

Таблиця 2.4

Десяткове число	Значення розрядів синдрому помилки S_{Π}			Помилковий розряд
	3	2	1	
1	0	0	1	y_1
2	0	1	0	y_2
3	0	1	1	x_1
4	1	0	0	y_3
5	1	0	1	x_2
6	1	1	0	x_3
7	1	1	1	x_4

Відповідність між значеннями розрядів синдрому та номером помилкового розряду встановлено з тих міркувань, що кожний контрольний розряд охоплюється тільки однією перевіркою й, отже, контрольним розрядам $y_1 - y_3$ повинні відповідати синдроми з одиничною вагою. Решта синдромів поставлена у відповідність інформаційним розрядам у природному порядку. На підставі даних таблиці може бути записана перевірна матриця, яка являє собою матрицю значень розрядів синдрому помилки, повернуту на кут 90° проти годинникової стрілки та впорядковану за зростанням номеру помилкового розряду у кодовій комбінації

$$M_{(7,4)} = \begin{vmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ x_1 & x_2 & x_3 & x_4 & y_1 & y_2 & y_3 \end{vmatrix}. \quad (2.54)$$

Тоді співвідношення для визначення розрядів синдрому помилки будуть мати вигляд

$$\begin{cases} S_{\Pi 1} = x_1 \oplus x_2 \oplus x_4 \oplus y_1 \\ S_{\Pi 2} = x_1 \oplus x_3 \oplus x_4 \oplus y_2 \\ S_{\Pi 3} = x_2 \oplus x_3 \oplus x_4 \oplus y_3 \end{cases} \quad (2.55)$$

Після отримання значень розрядів синдрому, помилковий розряд визначається за допомогою табл. 2.4. Помилковий розряд також відповідає стовпцю перевіркової матриці, який збігається з вектором синдрому помилки.

2.9.3.3. Метод мажоритарного декодування

Під час використання цього метода для кожного інформаційного розряду спільно розглядаються його значення, які отримуються на основі значень інших розрядів та всіх рівнянь вигляду (2.45), до яких він входить з ненульовим коефіцієнтом, а також значення розряду, яке безпосередньо присутнє у кодовій комбінації.

Розглянемо цей метод для коду $G_{(6,3)}$ з $k=3$, $\mu=3$, $n=3$ і $d_{\min}=3$ й контрольними розрядами, які обчислюються за співвідношеннями

$$\begin{cases} y_1 = x_1 \oplus x_2 \\ y_2 = x_2 \oplus x_3 \\ y_3 = x_1 \oplus x_3 \end{cases} \quad (2.56)$$

Твірна матриця, цього коду має вигляд

$$G_{(6,3)} = \begin{vmatrix} 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ x_1 & x_2 & x_3 & y_1 & y_2 & y_3 \end{vmatrix}. \quad (2.57)$$

З урахуванням можливості переносу змінних в (2.56) з однієї частини рівняння в іншу без зміни знака складаються групи незалежних перевірок

$$\begin{cases} x_1 = x'_2 \oplus y'_1 \\ x_1 = x'_3 \oplus y'_3 \\ x_1 = x'_1 \end{cases} I \quad \begin{cases} x_2 = x'_1 \oplus y'_1 \\ x_2 = x'_3 \oplus y'_2 \\ x_2 = x'_2 \end{cases} II \quad \begin{cases} x_3 = x'_2 \oplus y'_2 \\ x_3 = x'_1 \oplus y'_3 \\ x_3 = x'_3 \end{cases} III. \quad (2.58)$$

Тут, як і раніше, індексом «'» позначені елементи прийнятої кодової комбінації.

Рішення про результат перевірки у випадку тільки виявлення одно- або двократних помилок приймається по неоднорідності результатів хоча б в одній з перевірок I , II , III .

Наприклад, у випадку трансформації двох розрядів ЛСК 001011^{*} виникає комбінація 011010 і для неї згідно (2.58) незалежні перевірки мають вигляд

$$\begin{cases} x_1 = 1 \oplus 0 = 1 \\ x_1 = 1 \oplus 0 = 1 \\ x_1 = 0 \end{cases} I \quad \begin{cases} x_2 = 0 \oplus 0 = 0 \\ x_2 = 1 \oplus 1 = 0 \\ x_2 = 1 \end{cases} II \quad \begin{cases} x_3 = 1 \oplus 1 = 0 \\ x_3 = 0 \oplus 0 = 0 \\ x_3 = 1 \end{cases} III. \quad (2.59)$$

Оскільки в перевірках отримані неоднорідні результати, то в прийнятому коді є помилки.

Якщо поставлено задачу виправлення помилок (а для коду $G_{(6,3)}$ при $d_{\min} = 3$ вони можуть бути тільки однократними), то рішення про ймовірне значення інформаційного розряду приймається «по більшості» значень, отриманих у всіх перевірках (2.58), які стосуються цього розряду.

Припустимо, що в коді $G_{(6,3)}$ з твірною матрицею (2.57) комбінація 001011 перекручена в розряді x_2 , внаслідок чого прийнята кодова комбінація має вигляд 011011. Незалежні перевірки виконуються за рівняннями (2.58)

$$\begin{cases} x_1 = 1' \oplus 0' = 1 \\ x_1 = 1' \oplus 1' = 0 \\ x_1 = 0' \end{cases} \Rightarrow x_1 = 0; \quad \begin{cases} x_2 = 0' \oplus 0' = 0 \\ x_2 = 1' \oplus 1' = 0 \\ x_2 = 1' \end{cases} \Rightarrow x_2 = 1'; \quad \begin{cases} x_3 = 1' \oplus 1' = 0 \\ x_3 = 0' \oplus 1' = 1 \\ x_3 = 1' \end{cases} \Rightarrow x_3 = 1. \quad (2.60)$$

Таким чином, у прийнятій кодовій комбінації виправлена помилка в розряді x_2 й відновлений правильний код переданого повідомлення – 001011.

Лінійні коди в СПІ технологічного призначення практично не використовуються, однак методики їхнього синтезу й декодування досить корисні при вивченні більш ефективних кодів, в каналах зв'язку, де помилки в кодових комбінаціях мають тенденцію до групування, а декодування може здійснюватися альтернативними шляхами.

2.10. Циклічні коди

2.10.1. Загальні відомості. Двійкове й алгебраїчне подання комбінацій циклічних кодів

Циклічні коди (ЦК) були розроблені Хаффманом у середині 50-х років минулого століття на основі теорії полів Галуа двійкових багаточленів $GF(2^n)$. В цих полях коефіцієнти багаточленів можуть приймати значення $c_i \in \{0,1\}$, операція додавання елементів поля проводиться за модулем 2 і еквівалентна вирахуванню. Назву ЦК одержали через найпростіший метод одержання їхніх комбінацій – шляхом циклічного зсуву вихідної n -розрядної комбінації вліво з переносом зі старшого розряду в молодший.

Так, якщо вихідна комбінація даного ЦК 1100101, то послідовний зсув дає наступні його комбінації: 1001011, 0010111 і т.д.

Є також спосіб запису комбінацій ЦК у вигляді багаточленів фіктивної змінної x

$$F(x) = c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \dots + c_1x^1 + c_0x^0. \quad (2.61)$$

Відповідно з цим наведена вище комбінація ЦК 1100101 записується як $F(x)_1 = x^6 + x^5 + x^2 + 1$; її циклічні зсуви одержуємо послідовним множенням $F(x)_1$ на x з урахуванням того, що $x^{n-1} \cdot x = x^0 = 1$. Тоді $F(x)_2 = F(x)_1 \cdot x = x^6 + x^3 + x + 1$, $F(x)_3 \cdot x = x^4 + x^2 + x + 1$ і т.д. Підсумовуючи наведені кодові комбінації ЦК одержуємо зокрема, $F(x)_1 + F(x)_2 = x^5 + x^3 + x^2 + x$ або у двійковому вираженні 0101110.

Слід зазначити, що множини комбінацій ЦК, створених циклічним зсувом, для цілей завадостійкого кодування непридатні, тому що в них інформаційна й контрольна частини при зсувах зміщуються, і виділити їх зокрема на стороні прийому практично неможливо.

2.10.2. Методика побудови систематичного ЦК

В основу цієї методики покладена теорема, у якій доводиться, що такий ЦК – це множина багаточленів (кодових комбінацій) ступеня $n - 1$ й менш, які без остачі діляться на так званий твірний (породжуючий) багаточлен $P(x)$, що є одним зі співмножників (або добутку декількох співмножників) розкладання бінома $x^n + 1$.

Таке подання комбінацій ЦК дозволяє одержати їх у систематичній формі

$$x_1 x_2 \dots x_k x_{k+1} x_{k+2} \dots x_{k+\mu}, \quad (2.62)$$

де $x_1 \dots x_k$ – інформаційні, а $x_{k+1} \dots x_{k+\mu}$ – контрольні розряди комбінації, $n = k + \mu$, при цьому число контрольних розрядів μ – це ступінь твірного багаточлена $P(x)$.

Твірний багаточлен повинен бути неприводимим, або добутком декількох неприводимих багаточленів. Останні задовольняють наступним вимогам:

багаточлен є дільником багаточлена $x^{2^n-1} + 1$.

багаточлен ділиться без остачі тільки сам на себе, або на одиницю;

коефіцієнтом при нульовому ступені є одиниця;

число членів з ненульовими коефіцієнтами є непарним (за винятком двочлену $x + 1$).

Таблиця неприводимих багаточленів ступеню до 10-го включно приведена у додатку. Окрім того, для ЦК $F_{(n,k)}$, твірний багаточлен повинен також задовольняти умовам:

ступінь багаточлена μ дорівнює $n - k$;

вага багаточлена (кількість ненульових коефіцієнтів) повинна бути не менш необхідної кодової відстані d_{min} ;

Послідовність дій по утворенню комбінацій систематичного ЦК така:

1. За вимогами до здатності коду по виявленню й виправленню помилок й довжини його інформаційної частини « k » вибирається відповідний багаточлен

$P(x)$ ступеня « μ ». Далі інформаційний багаточлен $G(x)$ множиться на одночлен x^μ . У двійковому поданні ця дія є приписуванням « μ » нулів праворуч до інформаційної частини $G(x)$.

2. Добуток $G(x) \cdot x^\mu$ ділиться на $P(x)$. У результаті виникає ціла $Q(x)$ (багаточлен того ж ступеню, що $G(x)$) й дробова $\frac{R(x)}{P(x)}$ складові частини:

$$\frac{G(x) \cdot x^\mu}{P(x)} = Q(x) + \frac{R(x)}{P(x)}. \quad (2.63)$$

3. Вираз (2.61) приводиться до загального знаменника й перегрупується з урахуванням того, що додавання й вирахування в полі за модулем 2 еквівалентні

$$Q(x) \cdot P(x) = G(x) \cdot x^\mu + R(x). \quad (2.64)$$

Згідно наведеної вище теореми тут у лівій частині міститься багаточлен циклічного коду, а отже у правій частині – також, але зі структурою систематичного коду – до багаточлена $G(x) \cdot x^\mu$ ступеня не менш μ і не більше $n-1$ додані члени ступеня не більш $\mu-1$ так званої остачі $R(x)$. У двійковому ж поданні на місце μ нулів у добутку $G(x) \cdot x^\mu$ приписується код остачі $R(x)$, що не може мати довжину більше μ .

2.10.3. Матричний метод побудови систематичного ЦК

З урахуванням того, що операції множення й ділення в поле за модулем 2 є лінійними, матричний метод побудови коду, докладно викладений в п. 2.9, може бути застосований і до циклічного коду, за винятком того, що контрольні частини твірної матриці, що $F_{(n,k)}$, одержуються у вигляді залишків $R(x)$

$$F_{(n,k)} = \left| G_{(k,k)}, R(x)_{(k,\mu)} \right|. \quad (2.65)$$

Пояснимо викладене вище на наступному прикладі. Якщо взяти $k=4$ й зажадати, щоб код мав $d_{\min} = 3$, то в якості $P(x)$ може бути використаний один з багаточленів третього ступеня в розкладанні бінома

$$x^n + 1 = x^7 + 1 = (x+1) \cdot (x^3 + x + 1) \cdot (x^3 + x^2 + 1). \quad (2.66)$$

Співмножники розкладання (2.66) є незвідними.

Отже, тут довжина ЦК буде $n = k + \mu = 4 + 3 = 7$. Виконаємо дії методики 2.10.2, виходячи з того, що вони застосовуються для рядків інформаційної підматриці $G_{(4,4)}$. В алгебраїчній формі результатом їх множення на x^μ буде вектор одночленів

$$G_{(4,4)} \cdot x^3 = \begin{vmatrix} 1 \\ x \\ x^2 \\ x^3 \end{vmatrix} \cdot x^3 = \begin{vmatrix} x^3 \\ x^4 \\ x^5 \\ x^6 \end{vmatrix}, \quad (2.67)$$

або у двійковій формі – матриця

$$G_{(4,4)} \cdot x^3 = \begin{vmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{vmatrix}. \quad (2.68)$$

Далі обчислюються залишки для рядків твірної матриці $F_{(7,4)}$. У якості $P(x)$ виберемо другий співмножник в (2.66), $P(x) = x^3 + x + 1$. Тоді в першому рядку $G_{(7,4)}$ остача $R(x)$ визначиться як:

$$\begin{array}{l} x^3 \quad \oplus \quad |x^3 + x + 1 \\ \underline{x^3 + x + 1} \quad | \quad 1 \\ \quad \quad \quad x + 1, \text{ або } R(x)_1 = 011, \end{array}$$

а в четвертому

$$\begin{array}{l} x^6 \quad \oplus \quad |x^3 + x + 1 \\ \underline{x^6 + x^4 + x^3} \quad | \quad x^3 + x + 1 \\ \quad \quad \oplus \quad \begin{array}{l} x^4 + x^3 \\ \underline{x^4 + x^2 + x} \\ \quad \quad x^3 + x^2 + x \\ \quad \quad \oplus \quad \begin{array}{l} x^3 + x + 1 \\ \underline{\quad \quad \quad} \\ \quad \quad \quad x^2 + 1, R(x)_4 = 101. \end{array} \end{array} \end{array}$$

Зручніше залишки обчислювати за «один прохід» двійковим діленням останнього рядка матриці (2.66) на двійковий еквівалент $P(x) = 1011$:

$$\begin{array}{r} \oplus \begin{array}{r} 1000000 \\ \underline{1011} \\ 011 \end{array} \quad \left| \begin{array}{l} \underline{1011} \\ 1011 \end{array} \right. \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad R(x)_1 \\ \quad \quad \quad \underline{110} \quad \quad \quad \text{не ділиться на } P(x), R(x)_2 \\ \oplus \begin{array}{r} 1100 \\ \underline{1011} \\ 111 \end{array} \quad \quad \quad R(x)_3 \\ \oplus \begin{array}{r} 1110 \\ \underline{1011} \\ 101. \end{array} \quad \quad \quad R(x)_4. \end{array}$$

Порівняння варіантів визначення залишків у двійковій і у формі багаточленів показує збіг їхніх результатів. Остаточна матриця ЦК має вигляд

$$F_{(7,4)} = \left(\begin{array}{ccccccc} 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \end{array} \right). \quad (2.69)$$

Вимоги до контрольної частини ЦК такі самі, як і для ЛСК – число одиниць у $R(x)$ рядку повинне бути не менш $d_{\min} - 1$, а в порозрядній сумі за модулем два для них – $d_{\min} - 2$. Відзначимо, що різним інформаційним частинам коду можуть відповідати однакові контрольні.

Усього даний ЦК має 15 ненульових комбінацій, які одержують аналогічно, як і для ЛСК $G_{(7,4)}$.

2.10.4. Властивості найпростіших ЦК по виявленню помилок і вибір необхідного твірного багаточлена

З викладеного вище випливає очевидний висновок про ознаку наявності в кодовій комбінації помилок кратності більше тієї, на яку розрахований код – багаточлен перекрученої комбінації на твірний багаточлен ділиться з остачею.

Як і для ЛСК уведемо поняття вектора (багаточлена) помилок, що порозрядно підсумовується з багаточленом вихідної комбінації цього ЦК.

У силу лінійності операцій над багаточленами ЦК (як і ЛСК) ознакою помилок буде те, що багаточлен помилок V_{Π} на утворюючий багаточлен не ділиться без остачі.

Звідси впливають висновки про виявлення помилок у ЦК, які породжуються найпростішими багаточленами $P(x)$:

1. ЦК із твірним багаточленом $P(x)$, що має у своєму складі більше одного члена, виявляє всі поодинокі помилки. Доказ цього елементарний – одночлен V_{Π} на багаточлен $P(x)$ без остачі поділити неможливо.

2. ЦК із твірним багаточленом $P(x) = (x+1)$ виявляє всі помилки непарної кратності.

Покажемо, що в комбінаціях даного ЦК завжди є парне число членів, а V_{Π} з непарним числом членів, як наслідок, на двочлен $P(x) = x+1$ без остачі поділити неможливо.

У розглянутому випадку багаточлен циклічного коду є добутком

$$F(x) = Q(x) \cdot (x+1) = Q(x) \cdot x + Q(x). \quad (2.70)$$

У цій сумі є $2p$ членів, з яких $2p_1$ з однаковими ступенями при порозрядному додаванні звернуться в нулі, тому в багаточлені ЦК буде $2p - 2p_1$ – парне число членів. Таким чином, висловлене твердження доведене.

3. ЦК виявляє будь-яку поодинокую або подвійну помилку, якщо довжина коду $n \leq l$, де l – найменше число, при якому біном $x^l + 1$ ділиться на утворюючий багаточлен $P(x)$ даного коду без остачі.

Виявлення поодинокі помилки забезпечується тим, що тут $P(x)$ буде багаточленом. Двочлен помилок запишемо у вигляді

$$V_{\Pi} = x^i + x^j = x^j (x^{i-j} + 1), \quad (2.71)$$

$n \geq i > j$, отже $(i-j) < n \leq l$, двочлен помилки $x^{i-j} + 1$ на $P(x)$ без остачі не поділиться й подвійна помилка буде виявлена.

За умовою Хеммінга (2.42) найбільш економічним при $d_{\min} = 3$ буде код, у якого число контрольних розрядів і довжина коду зв'язані співвідношенням $n = 2^{\mu} - 1$ або в нашому випадку $l = 2^{\mu} - 1$.

При цьому, в розкладанні бінома $x^l + 1 = x^{2^{\mu}-1} + 1$ буде хоча б один незвідний багаточлен ступеня μ , який забезпечує виявлення будь-якої поодинокі або подвійної помилки. У загальному випадку біном $x^{2^{\mu}-1} + 1$ є найменшим загальним кратним для усіх незвідних багаточленів $P(x)$ ступеня μ .

4. ЦК, у якого твірний багаточлен має вигляд

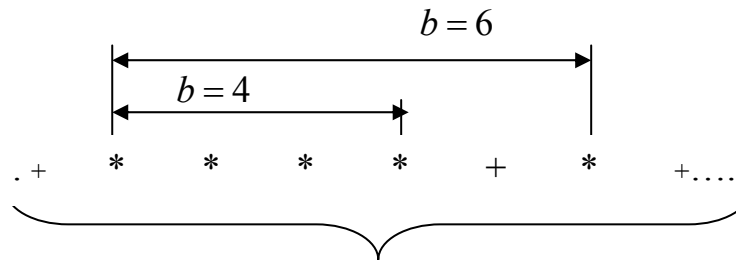
$$P(x) = (x+1) \cdot P'(x), \quad (2.72)$$

виявляє будь-яку поодинокую або подвійну, або потрійну помилку, якщо ступінь незвідного багаточлена $P'(x)$ такий, що виконується умова $2^\mu - 1 \geq n$.

Співмножник $P(x)$, $x+1$ забезпечує виявлення поодинокі або потрійної помилок, а $P'(x)$ – подвійної.

5. ЦК, у якого твірний багаточлен $P(x)$ має ступінь μ , виявляє пакет (групу) помилок довжиною $b \leq \mu$ в будь-якій частині кодової комбінації.

Пакети помилок можуть виглядати у такий спосіб, як це показано на рис. 2.7.



Кодова комбінація ЦК

Рис. 2.7

Пакет помилок починається й закінчується перекрученим розрядом коду, відзначеним як «*»; усередині пакета можуть перебувати неспотворені розряди «+».

Багаточлен пакетної помилки має вигляд $V_{\Pi} = x^i + \dots + x^\gamma + \dots + x^j$, і якщо $i > \gamma > j$, то

$$V_{\Pi} = x^j \cdot (x^{i-j} + \dots + x^{\gamma-j} + \dots + 1). \quad (2.73)$$

Тому що $i - j = b - 1$, багаточлен у дужках у виразі (2.71) на $P(x)$ ступені μ без остачі не ділиться й групова помилка виявляється.

Викладені положення про здатність ЦК до виявлення помилок дозволяють вирішити питання вибору необхідного твірного $P(x)$ багаточлена в низці практичних застосувань. Так, у системах передачі технологічної інформації й даних широко використовується код CRC-16 (cyclical redundancy check), циклічний надлишковий код з утворюючим багаточленом

$$P(x) = x^{16} + x^{12} + x^5 + 1 = (x+1) \cdot (x^{15} + x^{14} + x^{13} + x^2 + x^4 + x^3 + x^2 + x + 1). \quad (2.74)$$

Співмножник п'ятнадцятого ступеня в (2.74) – незвідний багаточлен, тому відповідний ЦК по властивості 4 виявляє будь-яку поодинокую або подвійну, або потрійну помилку в будь-якій частині кодової комбінації, або групову помилку з довжиною $b \leq 16$, якщо довжина коду $n \leq 2^{15} - 1 = 32767$.

Додаткове збільшення здатності до виявлення помилок досягається шляхом двоступінчастого кодування при передачі в одному інформаційному блоці декількох повідомлень, таких як адреса одержувача, вид операції (функціональна адреса), цифрові дані, тощо. При цьому частини блоку спочатку кодуються, наприклад, інверсним кодом, а потім отримана складена кодова комбінація захищається циклічним кодом, зокрема CRC-16.

Така структура ЦК використовується для телеуправління й телесигналізації в технологічних процесах електроенергетики, нафтохімії й інших, пов'язаних з підвищеною небезпекою наслідків помилок при контролі та управлінні.

При декодуванні на відсутність помилок послідовно перевіряються згадані частини блоку, а якщо в них помилки не виявлені, то й блок у цілому.

При виявленні помилок приймач передає джерелу коротке повідомлення «помилка», і останній повторює забраковане повідомлення. Такий спосіб підвищення вірогідності передачі інформації називається переопитуванням або вирішальним зворотним зв'язком. Він одержав практично виняткове поширення в системах телемеханіки й локальних системах передачі даних.

Крім розглянутих найпростіших, розроблена множина ЦК, що мають особливі коригувальні властивості – коди Файра, що виправляють пакети помилок, коди Боуза-Чоудхурі-Хоквінгема (БЧХ), для яких характерні добре розроблена методика побудови й здатність виправляти помилки будь-якої кратності, недвійкові коди Рида – Соломона (КРС), що використовуються у високошвидкісних магістральних каналах зв'язку й ін.

2.10.5. Виправлення помилок у ЦК

2.10.5.1. Метод синдромів

Метод може бути застосований для виправлення однократних помилок і використовує підхід, аналогічний тому, що був розглянутий в методі синдромів стосовно ЛСК (див. п. 2.9.3.2). Тут синдромами виступають залишки від ділення одночленів помилок V_{Π} на $P(x)$. Таке ділення зручно виконати у двійковій формі. Для коду $F_{(7,4)}$ з твірною матрицею (2.69) нижче наведена таблиця синдромів, відповідно до якої може бути виправлена поодинокі помилка, як в інформаційній, так і контрольній частинах.

Таблиця 2.5

$E(x)$	1000000	0100000	0010000	0001000	0000100	0000010	0000001
S_0	101	111	110	011	100	010	001

Так, наприклад, прийнята кодова комбінація $F'(x)=1011101$ ЦК з $P(x)=1011$. Ділення $F'(x)$ на $P(x)$ дає

$$\oplus \begin{array}{r|l} 1011101 & 1011 \\ 1011 & 1000 \\ \hline & 101. \end{array}$$

Відповідно до залишку 101 за допомоги табл. 2.3 визначається $V_{\Pi} = 1000000$ й виправлена комбінація коду становитиме

$$F(x) = F'(x) + V_{\Pi} = 1011101 \oplus 1000000 = 0011101.$$

2.10.5.2. Узагальнений алгоритм виправлення помилок

Поставлена задача корекції помилок ефективно вирішується з використанням наступного алгоритму, який включає такі кроки:

1. Як і в задачі виявлення помилок, прийнята кодова комбінація $F'(x)$ ділиться на $P(x)$. Якщо залишок $R(x) = 0$, то кодова комбінація прийнята вірно. У іншому випадку підраховується вага залишку w (число одиниць у двійковому коді $R(x)$), яка порівнюється з числом помилок, що s виправляються кодом.

2. Якщо $w \leq s$, то обчислюється порозрядна сума за модулем 2 $F'(x) + R(x)$, котра й буде виправленою комбінацією коду повідомлення $F(x)$.

3. Якщо $w > s$, здійснюється циклічний зсув $F'(x)$ вліво на один розряд, отримана кодова комбінація $F'(x)_1$ ділиться на $P(x)$ з обчисленням ваги w_1 остачі $R(x)_1$ й порівнюється з параметром s . У випадку, якщо $w_1 \leq s$ обчислюється аналогічно п. 2 сума $F'(x)_1 + R(x)_1$, отримана кодова комбінація циклічно зсувається вправо на один розряд і в результаті одержується виправлений код повідомлення.

4. Якщо після першого зсуву вага остачі $w_1 > s$, здійснюється наступний зсув $F'(x)_1$ вліво, ділення $F'(x)_2$ на $P(x)$, визначення ваги залишку w_2 . При $w_2 > s$ цикл операцій зсуву, ділення, визначення й аналізу залишку повторюється доти, поки на черговому циклі « q » не буде досягнута умова $w_q \leq s$.

Обчислений у цьому циклі залишок підсумовується з діленням (як у п. 2), і кодова комбінація $F'(x)_q + R(x)_q$ суми піддається $(q-1)$ послідовним зсувам вправо, у результаті чого й визначається виправлений код повідомлення.

Як приклад виправимо помилку ($s=1$) в кодовій комбінації $F'(x) = 1011101$, якщо $P(x) = x^3 + x + 1 = 1011$. Згідно алгоритму виконуються дії:

1. Ділення $F'(x)$ на $P(x)$

$$\oplus \begin{array}{r} 1011101|1011 \\ \underline{1011} \quad |1000 \end{array}$$

$$0000101 - P(x)_1 \Rightarrow w_1 = 2, w_1 > s.$$

2. Зсуву вліво $F'(x)$ й одержання $F'(x)_1 = 0111011$.

3. Ділення $F'(x)_1$ на $P(x)$

$$\oplus \begin{array}{r} 0111011|1011 \\ \underline{1011} \quad |0110 \end{array}$$

$$\oplus \begin{array}{r} 1011 \\ \underline{1011} \end{array}$$

$$0000001 - P(x)_2 \Rightarrow w_2 = 1, w_2 = s.$$

4. Обчислення суми

$$F'(x)_2 = F'(x)_1 + R(x)_2 = 0111011 \oplus 0000001 = 0111010.$$

5. Зсуву $F'(x)_2$ на $q-1=2-1=1$ розряд вправо й одержання виправленого кода повідомлення $F(x) = 0011101$.

6. Перевірки правильності обчислень діленням $F(x)$ на $P(x)$

$$\oplus \begin{array}{r} 0011101|1011 \\ \underline{1011} \quad |0011 \end{array}$$

$$\oplus \begin{array}{r} 1011 \\ \underline{1011} \end{array}$$

$$0 \Rightarrow F(x) \text{ не містить помилку.}$$

Відомі й інші методи виправлення помилок, зокрема мажоритарний і ін.

2.10.6. Згорткові коди (ЗК)

Ці коди також є лінійними, оскільки послідовність елементів на виході пристрою кодування (ПК) є результатом підсумовування за модулем 2 елементів вхідної послідовності.

ПК ЗК є кінцевим автоматом у якого поточний набір вихідних сигналів визначається як множиною поточних, так і попередніх вхідних. Основою такого ПК є K – розрядний регістр зсуву RG і n суматорів за модулем 2 (рис. 2.8). По тактах завантаження RG на нього надходить чергова k – розрядна група елементів вхідної послідовності, по кожному з тактів здійснюється послідовне опитування всіх суматорів, у результаті чого на виході ПК з'являється група з n елементів вихідної послідовності. Число комірок RG K називається величиною кодового обмеження, а відношення $\frac{k}{n}$ – ступенем кодування. В найпростішому випадку $k=1$. У розглянутому прикладі $n=2$ (суматор, що відповідає положенню I перемикача П є виродженим), $\frac{k}{n} = \frac{1}{2}$.

Обов'язковою умовою для ПК ЗК є те, що структура зв'язків суматорів повинна давати інверсні вихідні групи при інверсії вхідного елемента.

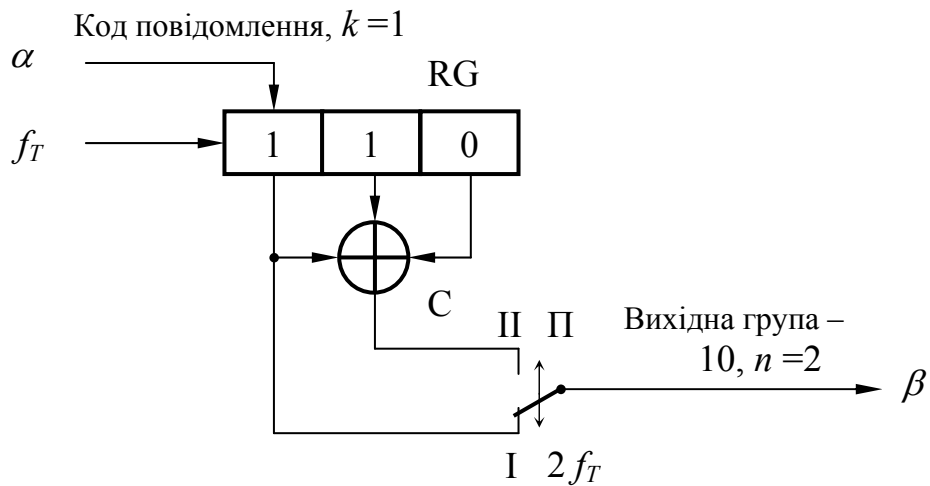


Рис. 2.8

Пам'ять ПК визначається його станами – умістом крайніх $K-1$ комірок RG (на рис. 2.8 вони розташовані ліворуч). При надходженні чергового вхідного елемента ПК може перейти лише в один із двох станів. На рис. 2.9 зображено граф кінцевого автомата ПК за рис. 2.8, що визначає його всілякі стани й переходи між ними. Суцільні стрілки відповідають черговому елементу вхідної послідовності 0, а штрихові – 1.

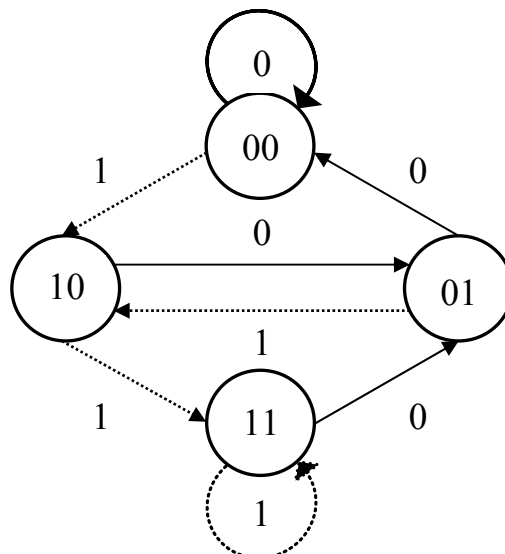


Рис. 2.9.

Формування вихідної групи ЗК ілюструється рис. 2.8, де на поточному такті в RG записана поточна частина коду повідомлення 110, якій відповідає група ЗК 10. Співвідношення для елементів групи β_{il} , β_{iII} і елементів вхідної послідовності α_i мають вигляд

$$\begin{cases} \beta_{iI} = \alpha_i \\ \beta_{iII} = \alpha_i \oplus \alpha_{i-1} \oplus \alpha_{i-2}. \end{cases} \quad (2.75)$$

Зв'язки ПК можуть бути також описані у вигляді багаточленів фіктивної змінної. Сигнали ПК в часовому вимірі, обумовленому тактами RG, описуються кодовими деревами й гратчастими діаграмами. Кодове дерево, що відповідає розглянутому прикладу, наведене на рис. 2.10.

Вертикальні гілки, спрямовані вгору, відповідають вхідному елементу ПК 0, а вниз – 1. Під кожною горизонтальною гілкою наведений стан RG на даному такті, а над гілкою – поточна група вихідної послідовності. З рис. 2.10 випливає, що після четвертого такту RG структура дерева стає регулярною. Так, коди гілок на четвертому такті при перегляді зверху вниз мають вигляд: 00-11-01-10-01-10-00-11, після чого послідовність кодів періодично повторюється. Аналогічна періодичність має місце й на всіх наступних тактах. У загальному випадку при нульовому початковому стані RG, регулярність дерева настає на $K+1$ -му такті.

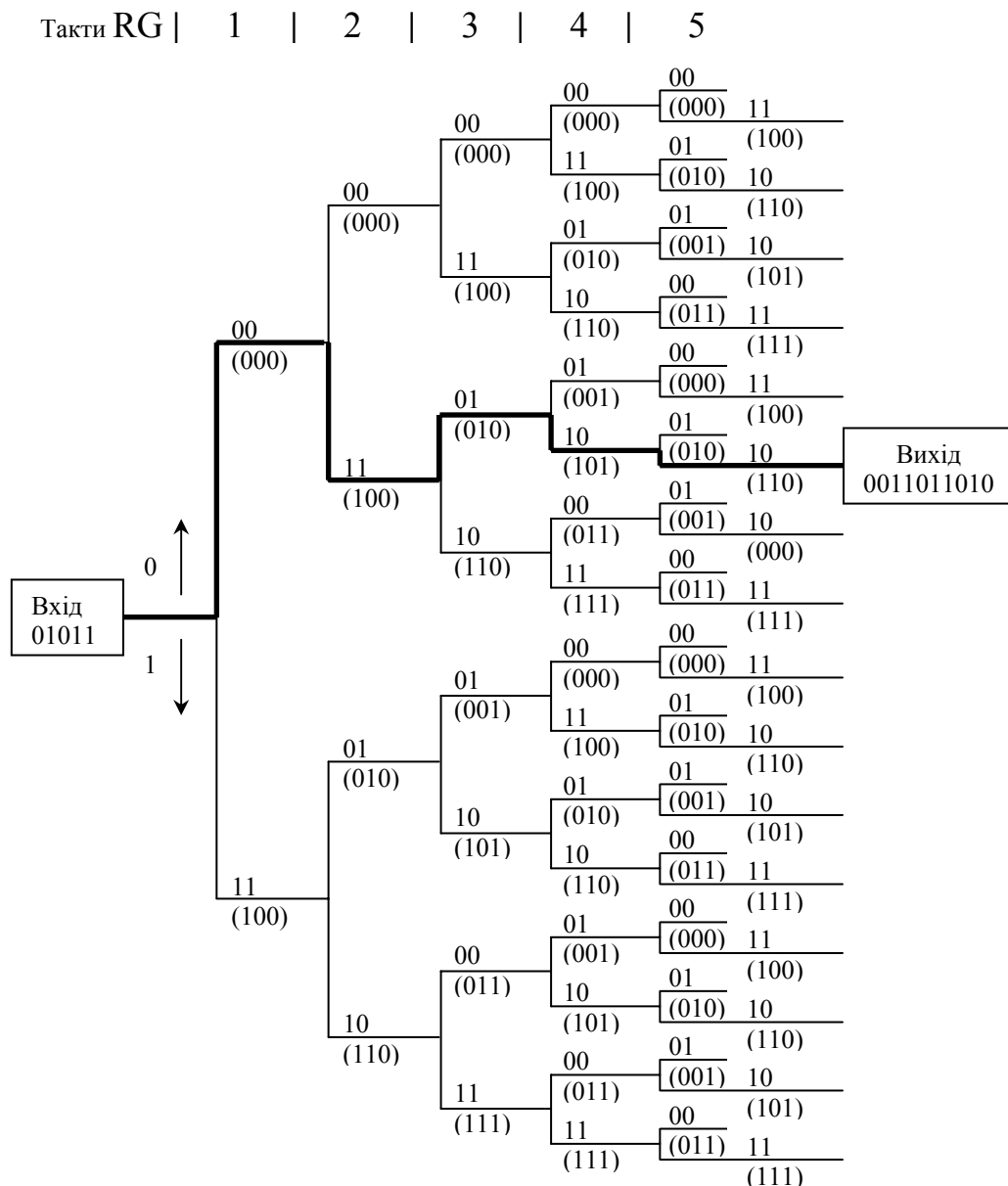


Рис. 2.10

Припустимо, що у вихідному стані RG обнулений і на вхід ПК надходить послідовність 01011. Після першого такту в RG буде записана комбінація 000, другого – 100, третього – 010, четвертого – 101, п'ятого – 110. Відповідно, вихідна послідовність буде: 0011011010.

Декодування комбінацій ЗК здійснюється тут групами по два елемента й залежно від того, у якому напрямку (вгору або вниз по кодовому дереву) відбувається перехід, встановлюється значення 0 або 1 поточного елемента повідомлення.

ЗК призначені для виправлення помилок, при цьому їхня ефективність є однією з найвищих серед коригувальних кодів, які використовуються в СПІ. Це досягається як за рахунок великої надлишковості, так і за рахунок удосконалювання алгоритмів кодування й декодування.

У розглянутому найпростішому коді виправлення поодинокі помилки пояснюється наступним прикладом за рис. 2.10. Для наведеної вихідної послідовності виникає помилка в групі 01 на такті 3 таким чином, що вона приймає вигляд 11. Розгалуження на цьому такті кодового дерева дає з групи 01 і 10, що не збігаються. Якби параметр n був би більше 2, то поодинокі помилку можна було б виправити за мінімумом відстані між перекрученою й дозволеними кодовими комбінаціями. У нашій випадку помилка може бути виправлена при відсутності помилок у наступній групі послідовності, що декодується, яка на такті 4 приймає значення 10 і виходить із гілки 01 такту 3. Отже, помилкова група 11 повинна бути виправлена на 01.

В 1967 р. Ендрю Віттербі розробив алгоритм декодування ЗК, що є оптимальним за виправною здатністю. Певний час його реалізація була обмежена складністю апаратної частини декодера, і тільки після появи мікро-ЕОМ високої швидкодії ЗК почали будуватися на їхній основі.

У цей час у СПІ знайшли застосування кодери-декодери (кодеки) ЗК із довжиною кодового обмеження до 10 і ступенем кодування $\frac{1}{2}$ й $\frac{1}{3}$.

2.10.7. Канальне кодування

Метою канального кодування є найбільш ефективно використання властивостей каналів зв'язку в процесах відображення елементів кодів повідомлень на сигнали, а також забезпечення функціонування власно каналоутворюючої апаратури. В деяких літературних джерелах до канальних кодів відносяться всі завадостійкі коди.

2.10.7.1. Гратчасті коди (ГК)

Ці коди забезпечують збільшення інформаційного навантаження на сигнал так, що він відображає не один елемент повідомлення, а декілька. Для цих цілей у локальних системах передачі даних на основі електричних ліній зв'язку широке поширення одержало використання частотних сигналів. Кожний з таких сигналів може мати кілька значень амплітуди і фази відносно опорного

коливання. Процес утворення таких сигналів зветься квадратурною амплітудною маніпуляцією (КАМ). Прикладом може служити КАМ-16, застосована в протоколі V.29 для максимальної швидкості передачі повідомлень 9600 с^{-1} . Гратчаста діаграма КАМ-16 наведена на рис. 2.11. Кожний з 16-ти сигналів тут являє собою комбінацію одного з 8-ми значень його фази та одного з 2-х значень амплітуди. Кожний з таких сигналів відображає свою чотирирозрядну комбінацію двійкових елементів послідовності коду повідомлення. Слід зазначити, що сусіднім крапкам в радіальному напрямку на діаграмі грат присвоюються коди Грея, які відрізняються друг від друга лише в одному розряді. Це дозволяє зменшити число помилок у прийнятій кодовій комбінації при змінах фази або амплітуди сигналу під впливом перешкод.

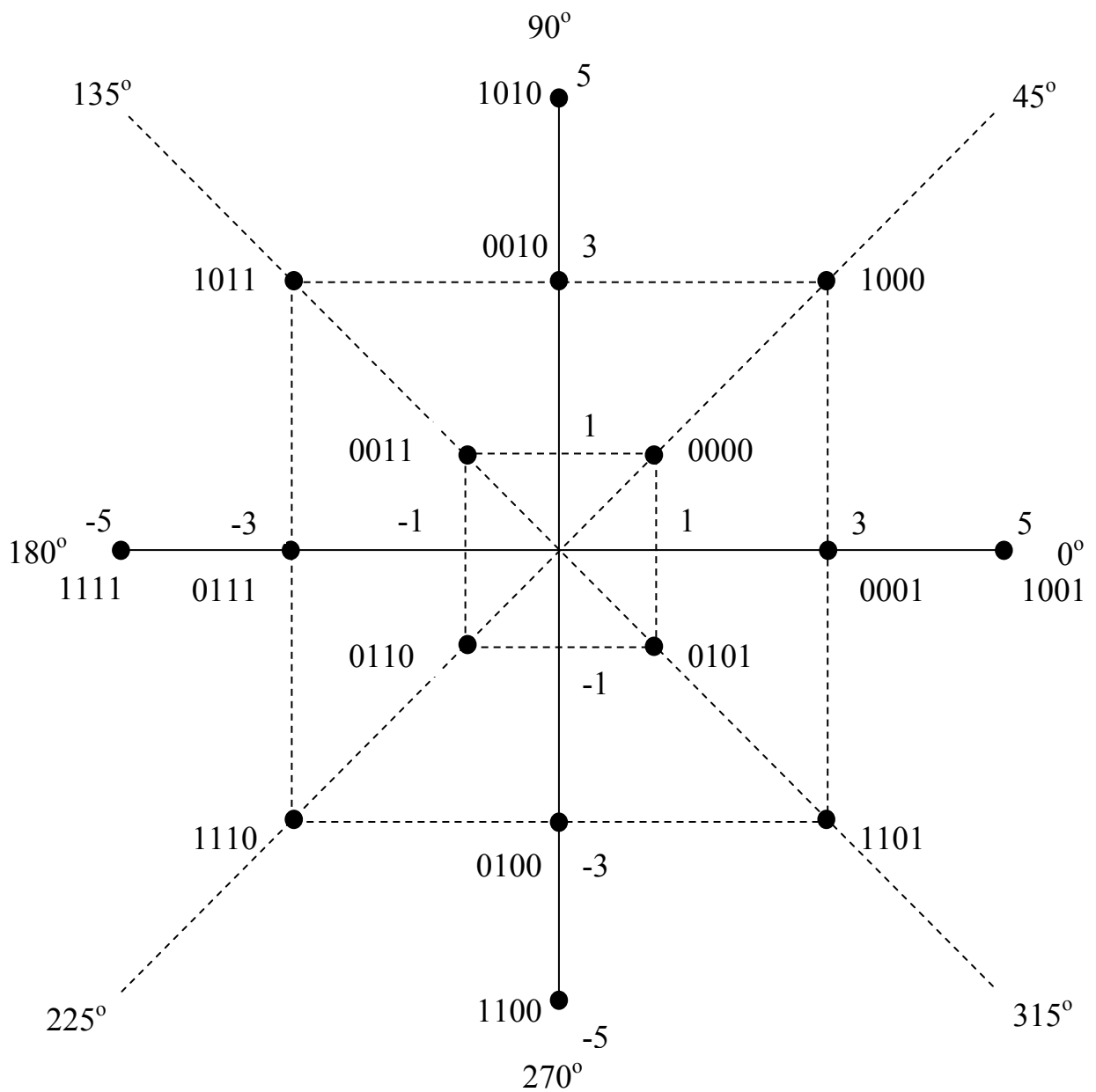


Рис. 2.11

Співвідношення амплітуд сигналів на тому самому, а також на сусідніх векторах ґрат задані таким чином, щоб різниці енергій «сусідніх» сигналів були близькими між собою, що забезпечує максимальну вірогідність передачі повідомлень по симетричних каналах зв'язку.

Помітимо, що в міру збільшення швидкості передачі за протоколами V.34, V.36, V.90 структури ґК істотно ускладнюються. Слід також зазначити, що при погіршенні якості каналу зв'язку – збільшенні частоти виникнення помилок, протоколи передбачають перехід на більш розріджені ґрати й меншу швидкість передачі повідомлень. Так, для згаданого протоколу V.29 швидкість передачі знижується до 4800 с^{-1} .

2.10.7.2. Шумові коди (ШК)

Ці коди одержали популярність у другій половині минулого століття як засіб захисту від несанкціонованого доступу до інформації і в криптографії одержали назву шифрів гамування. Однак при цьому виявилася й інша якість цих кодів – підвищена завадостійкість, як при зосередженій, так і розподіленій по смузі частот каналу перешкоді. Ця якість корисна при застосуванні силових ліній електропостачання як каналів зв'язку. На гірничо-металургійних виробництвах споживачами електроенергії можуть виступати: електроприводи з імпульсним регулюванням – видобувних комбайнів, конвеєрів і вентиляційних установок; дугові печі, зварювальні апарати, які у свою чергу є джерелами інтенсивних перешкод у широкому спектрі частот. ШК знайшли застосування в системах цифрового сотового зв'язку CDMA.

Основою ШК є псевдовипадкові двійкові послідовності (ПВП), у яких при досить великих періодах повторення практично досягається рівномірність і взаємнезалежність появи нулів і одиниць. Найбільш відомими є M-послідовності Хаффмана й Гордона – Мілза – Уолша (GMW). M-послідовності (послідовності максимальної довжини) утворюються рекурентними сумами за модулем 2 самих елементів послідовності. Довжина M-послідовності (період повторення) становить

$$N = 2^n - 1, \quad (2.76)$$

де n – натуральне число, що визначає довжину початкової фази ПВП. Початковою фазою може виступати будь-яка ненульова n -розрядна двійкова комбінація $\vec{\alpha} = \{\alpha_1, \dots, \alpha_n\}$, оскільки нульова комбінація з використанням будь-яких рекурентних сум за модулем 2 породжуватиме нульову. Структура рекурентних рівнянь пов'язана з поліноміальним поданням незвідних багаточленів, що є твірними для ЦК. Рекурентні суми завжди мають парну кількість доданків.

Саме рекурентний характер зв'язків значень елементів ПВП обумовлює її періодичність. Довжина M-послідовності N визначає загальну кількість ненульових n -розрядних двійкових комбінацій. Оскільки наступний елемент

ПВП формується виключно на основі комбінації з n попередніх, то через N кроків ця комбінація мусить повторитися, що викличе циклічне повторення всієї подальшої послідовності.

Твірні багаточлени ПВП мають бути не тільки незвідними, але й примітивними (додаток 1). Примітивний двійковий багаточлен порядку n має бути не тільки дільником багаточлена $x^{2^n-1} + 1$ (п. 2.10.2), але й не бути дільником будь-якого багаточлена $x^d + 1$, де d є дільником числа $2^n - 1$.

Покажемо, що багаточлен x^2+x+1 є примітивним. Число $2^n - 1$ у прикладі складає 3 є простим і не має дільників, тобто будь-який незвідний багаточлен порядку 2 є примітивним. З іншого боку, багаточлен, що розглядається, є дільником багаточлену x^3+1 :

$$\begin{array}{r} \oplus 1001 \underline{111} \\ \underline{111} \mid 11 \\ \oplus 111 \\ \underline{111} \\ 0 \end{array}$$

і задовільняє умові незвідності.

З використанням довільного примітивного багаточлена, n -го $P(x)$ ступеня може бути породжена псевдовипадкова двійкова послідовність, співвідношення для i -го члена α_i якої має вигляд

$$\begin{aligned} P(x) &= x^n + \dots + x^j + \dots + x^k + 1; \\ \alpha_i &= \alpha_{i-n} \oplus \dots \oplus \alpha_{i-j} \oplus \dots \oplus \alpha_{i-k} \text{ при } i > n. \end{aligned} \quad (2.77)$$

Тут j – ступінь довільного члена, коефіцієнт якого дорівнює одиниці, k – ненульовий ступінь останнього молодшого члена, коефіцієнт якого дорівнює одиниці.

Так, для $n=3$ твірному багаточлену $P(x) = x^3 + x + 1$ відповідає сума $\alpha_i = \alpha_{i-3} \oplus \alpha_{i-1}$. Період ПВП становить $N = 2^3 - 1 = 7$. Для початкової фази 101 період ПВП буде мати вигляд: 1010011.

Для того, щоб ПВП була максимально близькою до істинно випадкової, повинні виконуватися наступні умови:

1. Урівноваженості. Кількість одиниць у періоді ПВП повинне на одиницю перевищувати кількість нулів.

2. Групування. Половина сегментів послідовності повинна зображувати чергування нулів та одиниць: 01 і 10, четверта частина – пари нулів і одиниць: 00 і 11; восьма – трійок: 000 і 111 і т.д.

3. Кореляції. Дискретна автокореляційна функція (АКФ) на періоді ПВП має єдиний максимум при нульовому зсуві $R(0) = 1$ й $R(i) = -\frac{1}{N}$ для будь-яких інших величин зсувів i .

Пояснимо останню властивість на прикладі.

Під АКФ $R(\tau)$ для функції дійсної змінної $f(t)$, що має період T_n , мається на увазі величина

$$R(\tau) = \frac{1}{T_n} \int_{-\frac{T_n}{2}}^{\frac{T_n}{2}} f(t)f(t+\tau)dt. \quad (2.78)$$

Нормована на максимум АКФ має вигляд

$$\bar{R}(\tau) = \frac{\int_{-\frac{T_n}{2}}^{\frac{T_n}{2}} f(t)f(t+\tau)dt}{\int_{-\frac{T_n}{2}}^{\frac{T_n}{2}} f^2(t)dt} \quad (2.79)$$

і характеризує інтегральну на періоді ступінь зв'язку (корельованість) поточного значення функції з тим, що віддалений від нього з часовим лагом τ .

Якщо функція $f(t)$ описує двійкову послідовність, де нулю відповідає значення функції -1 , а одиниці $+1$, вираз (2.78) може бути спрощений і дискретна АКФ прийме вигляд

$$R(i) = \frac{N_C - N_{HC}}{N}, \quad (2.80)$$

де N_C , N_{HC} – кількість елементів, що збігаються та не збігаються у вихідній і зсунутій відносно неї на i кроків ПВП.

Порівняємо період розглянутої вище ПВП із самим собою. У цьому випадку $N_C = N = 7$, $N_{HC} = 0$ і, отже $R(0) = 1$. Якщо здійснити будь-який зсув періоду ПВП, наприклад, на $i = 2$ елемента, послідовність прийме вигляд: 1110100. Її порозрядне порівняння з вихідною послідовністю 1010011 дає значення $N_C = 3$, $N_{HC} = 4$ і, отже $R(i) = -\frac{1}{7}$. Полігон цієї АКФ наведений на рис. 2.12.

Вимога наявності єдиного максимуму АКФ на періоді необхідна для забезпечення незалежності роботи ряду каналів зв'язку, у яких шумові сигнали формуються на основі одного твірної багаточлена, але з різними початковими фазами. Лише примітивні твірні багаточлені задовільняють цій вимозі.

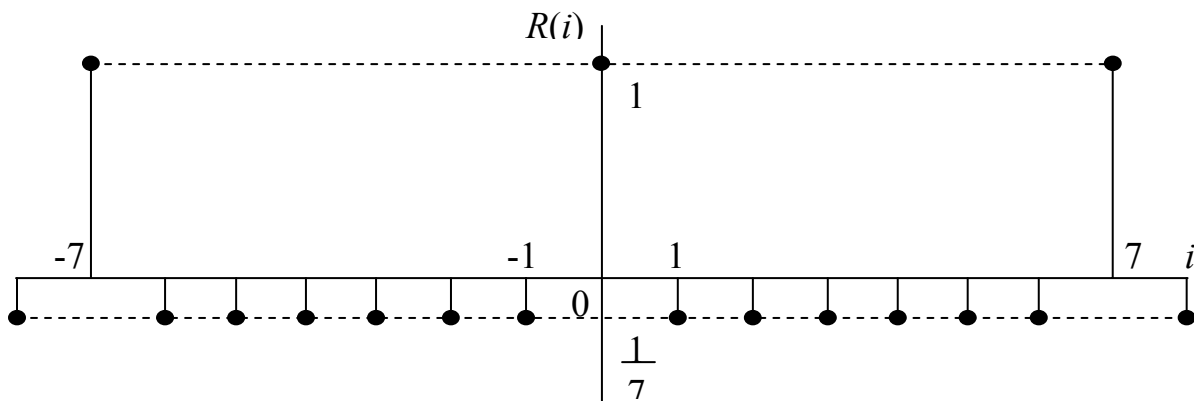


Рис. 2.12

В найпростішому варіанті процес кодування ШК полягає в накладанні на інформаційну бітову послідовність α ПВП ξ шляхом підсумовування за модулем 2 відповідних бітів інформаційної послідовності та ПВП. Процес декодування є цілком аналогічним. Якщо при кодуванні і декодуванні використовується одна і та сама ПВП з однією початковою фазою, на виході декодера при відсутності помилок отримується інформаційна послідовність γ , що збігається з первинною

$$\gamma_i = \xi_i \oplus (\xi_i \oplus \alpha_i) = (\xi_i \oplus \xi_i) \oplus \alpha_i = 0 \oplus \alpha_i = \alpha_i. \quad (2.81)$$

Покажемо, що побітове підсумовування за модулем 2 двійкової послідовності з довільною ймовірністю появи нулів p та ПВП з рівноймовірною появою нулів та одиниць дає також послідовність з рівноймовірною появою нулів та одиниць. За формулою повної ймовірності, ймовірність появи нуля у послідовності, що розглядається, становитиме

$$0,5p + 0,5(1 - p) = 0,5. \quad (2.82)$$

Таким чином, кодована за допомогою ШК інформаційна послідовність набуває властивостей ПВП.

2.10.7.3. Скремблерування даних

При послідовній передачі елементів кодів повідомлень виникає завдання синхронізації роботи приймача (Пр) з передавачем (Пд) СП. Процес передачі-прийому здійснюється за тактами (часовими інтервалами рівної тривалості), що задаються генераторами Пр та Пд. Синхронізація забезпечує прийом сигналів на тих самих тактах, на яких вони формуються. Підстроювання частоти

тактового генератора Пр виконується за змінами сигналів кода повідомлення, які надходять з лінії зв'язку в моменти формування тактових імпульсів генератора Пд.

Найкращі умови для підстроювання досягаються, коли з лінії зв'язку надходить послідовність нулів та одиниць, що чергуються. При суцільному потоці нулів або одиниць ознака, за якою здійснюється підналагодженням тактової частоти Пр – моменти переходів 0→1 або 1→0 відсутні і виникає можливість помилкового прийому повідомлень.

Наближене до рівномірного чергування нулів та одиниць у потоці повідомлень досягається шляхом його скремблерування (премішування) кодером. Вихідна послідовність кодера β формується шляхом додавання за модулем 2 поточного елемента вхідної послідовності α_i до деяких її попередніх елементів. На рис. 2.13 наведена структурна схема скремблера, у якого елемент вихідної послідовності кодера визначається як

$$\beta_i = \alpha_i \oplus (\alpha_{i-7} \oplus \alpha_{i-6}). \quad (2.83)$$

Сума у дужках відповідає елементу ПВП з твірним багаточленом $P(x) = x^7 + x^6 + 1$.

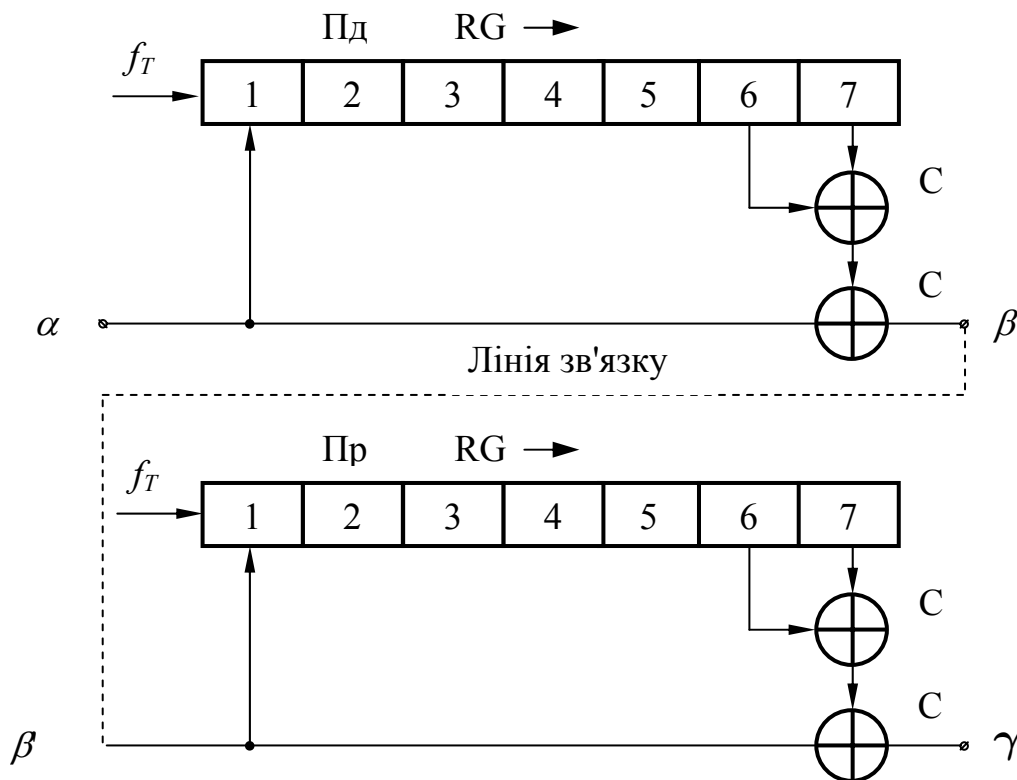


Рис. 2.13

На стороні Пр послідовність β' (можливо викривлена перешкодами у лінії зв'язку) декодується скремблером таким самим, як і на стороні Пд. При відсутності помилок вихідна послідовність Пр збігається з первинною

$$\begin{aligned} \gamma_i &= \beta_i \oplus (\alpha_{i-7} \oplus \alpha_{i-6}) = (\alpha_i \oplus (\alpha_{i-7} \oplus \alpha_{i-6})) \oplus (\alpha_{i-7} \oplus \alpha_{i-6}) = \\ &= \alpha_i \oplus [(\alpha_{i-7} \oplus \alpha_{i-6}) \oplus (\alpha_{i-7} \oplus \alpha_{i-6})] = \alpha_i. \end{aligned} \quad (2.84)$$

Якщо ж у послідовності β' наявні помилки, то на виході системи має місце їх розмноження. Так, у розглянутому прикладі, в послідовності γ помилка з'являється перший раз, коли вона надходить на вхід декодера і ще 2 рази, коли вона міститься у двох останніх комірках RG.

Другим недоліком скремблерування є неефективне розріджування послідовностей нулів та одиниць з довжиною, зрівняною з довжиною регістрів. Відома низка способів, які дозволяють усунути цей недолік. Найпростіший з них полягає у прекодуванні первинної послідовності α за кодом 4В/5В, що передбачає розбиття послідовності на четвірки елементів і заміну кожної із них п'ятіркою, в якій зустрічається не більше двох нулів поруч. Фрагмент кодової таблиці 4В/5В зображений в табл. 2.6.

Таблиця 2.6

4В	0000	0001	0010	0011	0100	0101	0110	0111
5В	11110	01001	10100	10101	01010	01011	01110	01111
4В	1000	1001	1010	1011	1100	1101	1110	1111
5В	10010	10011	10110	10111	11010	11011	11100	11101

Решта п'ятибітових комбінацій 4В/5В використовуються для кодування службових повідомлень, або є забороненими.

При передачі двійкових елементів у каналі зв'язку різнополярними імпульсами постійного струму розрідження потоків однотипних елементів також зменшує середню складову лінійного сигналу, що поліпшує швидкодію такого каналу.

А взагалі, вирішення питання синхронізації пов'язано з погіршенням перешкодостійкості та збільшенням надлишковості повідомлень.

Контрольні питання і завдання

1. Сформулюйте можливі цілі кодування повідомлень.
2. Поясніть, які принципи покладені в основу класифікації кодів.
3. Поясніть в чому полягає властивість нероздільності нерівномірного коду.
4. Поясніть, у чому полягає властивість нероздільності коду та чи є нероздільним код $\{11, 10, 01, 001, 000\}$.
5. Комбінації нерівномірного двійкового коду становлять: $\{a, ba, bc, bba, bbb, bbc\}$. Привести значення основи коду m . Перевірити виконання нерівності Крафта.
6. Сформулюйте основну теорему кодування в каналах без перешкод.
7. Викладіть алгоритм Хаффмена. Поясніть, за рахунок чого в даному коді відбувається зменшення середньої довжини кодової комбінації порівняно з рівномірними кодами.

8. Викладіть алгоритм Шеннона – Фено. Поясніть, чому варіант алгоритму, який передбачає впорядкування повідомлень за ймовірностями виникнення, є наближеним.
9. Джерело дискретних повідомлень генерує незалежні повідомлення з ймовірностями $\{0,5, 0,25, 0,125, 0,0625, 0,0625\}$. Знайдіть мінімальну середню довжину комбінації при кодуванні повідомлень джерела двійковим кодом. Побудуйте коди Шеннона – Фено та Хаффмана.
10. Уявіть у двійководесятковому коді рік власного народження.
11. Побудуйте таблицю коду Грея для 16-ти повідомлень.
12. Поясніть причину виникнення надлишковості в завадостійких кодах.
13. Дайте визначення систематичного завадостійкого коду.
14. Наведіть загальний вираз для надлишковості завадостійкого коду та його вигляд для систематичного коду.
15. Наведіть заборонені значення тетрад двійководесяткового коду та знайдіть його надлишковість.
16. Вважаючи, що систематичний код використовується для кодування 7-ми повідомлень, імовірність перекидання розряду двійкового коду становить $p = 10^{-3}$, кратність помилок у кодовій комбінації підпорядковується розподілу Бернуллі, а припустима ймовірність невиявлення помилок у комбінації складає $p_{H0} = 10^{-10}$, визначте: максимальну кратність помилок, яку повинен виявляти код, його кодову відстань, кількість контрольних розрядів та довжину кодової комбінації.
17. Поясніть принципи побудови кодів: «пар», КК, ІК, КБ. Закодуйте за їх допомоги комбінацію 111000.
18. Наведіть всі дозволені комбінації коду C_5^2 та знайдіть його надлишковість.
19. Для ЛСК $G_{(6,3)}$, твірна матриця якого має вигляд

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ x_1 & x_2 & x_3 & y_1 & y_2 & y_3 \end{pmatrix},$$

запишіть перевірочні співвідношення. Знайдіть всі дозволені кодові комбінації. виправте помилку у комбінації 101011 методом мажоритарного декодування.

20. Для ЛСК $G_{(7,4)}$ значення контрольних розрядів обчислюються за допомогою співвідношень

$$\begin{cases} y_1 = x_2 \oplus x_3 \oplus x_4 \\ y_2 = x_1 \oplus x_3 \oplus x_4, \\ y_3 = x_1 \oplus x_2 \oplus x_4 \end{cases}$$

запишіть перевірочну та твірну матриці. Доведіть наявність принаймні однієї помилки у комбінації 1011101. Складіть таблицю синдромів однократних

помилку. Використовуючи метод синдромів, виправте однократну помилку у комбінації 0111001.

21. Сформулюйте вимоги, яким задовольняють незвідні двійкові багаточлени.

22. Сформулюйте основні властивості ЦК по виявленню помилок та їх зв'язок з виглядом твірного багаточлена

23. Обчисліть, правильно чи з помилками одержаний ЦК $G_{(7,4)}$ 1001011 з твірним багаточленом $P(x)=x^3+x^2+1$.

24. Використовуючи незвідний багаточлен $P(x)=x^3+x^2+1$, побудуйте твірну матрицю та знайдіть всі дозволені комбінації ЦК, за допомогою якого може здійснюватися кодування 7-ми повідомлень з виявленням усіх помилок максимальною кратністю 3.

25. Для ЦК $G_{(6,3)}$ з твірним багаточленом $P(x)=x^3+x^2+1$ складіть таблицю синдромів однократних помилок та за її допомогою виправте помилку у комбінації 011111.

26. З застосуванням узагальненого алгоритма виправити помилку в ЦК $G_{(7,3)}$ 1010111 з твірним багаточленом $P(x)=x^4+x^2+x+1$ і кодовою відстанню $d_{\min} = 4$.

27. Наведіть структурну схему ПК ЗК, для якого величина кодового обмеження складає $K=3$, ступінь кодування $\frac{k}{n} = \frac{1}{2}$, а співвідношення для елементів групи

β_{ii}, β_{iii} і елементів вхідної послідовності α_i мають вигляд:

$$\begin{cases} \beta_{ii} = \alpha_i \oplus \alpha_{i-2} \\ \beta_{iii} = \alpha_i \oplus \alpha_{i-1} \oplus \alpha_{i-2} \end{cases}$$

пояснить принцип функціонування пристрою.

28. Наведіть граф кінцевого автомату, що являє собою двійковий регістр зсуву з числом комірок $K=4$.

29. Сформулюйте мету каналного кодування.

30. Пояснить, за рахунок чого відбувається збільшення інформаційного навантаження на сигнал у ГК.

31. Наведіть ескіз фазоамплітудної діаграми ГК. Пояснить, які параметри сигналу відповідають координатам діаграми.

32. Пояснить, як відбувається формування ШК.

33. Пояснить, чим відрізняються незвідні та примітивні багаточлени.

34. Доведіть, що незвідний багаточлен $x^4+x^3+x^2+x+1$ не є примітивним.

35. Пояснить причину виникнення періодичності М-послідовності Хаффмана.

36. Сформулюйте умови, які висуваються до ПВП.

37. Пояснить, що характеризує АКФ, наведіть співвідношення для дискретної АКФ ПВП.

38. Знайдіть рекурентне рівняння, за яким будуть визначатися елементи двійкової ПВП при твірному багаточлені x^4+x^2+1 . Знайдіть елементи періоду для початкової фази 1010.

39. Пояснить недоліки скремблерування.

40. Наведіть структурну схему пристрою скремблерування, що відповідає твірному багаточлену x^6+x+1 .

СПИСОК ЛІТЕРАТУРИ

1. Кузьмин, И.В. Основы теории информации и кодирования [Текст] / И.В. Кузьмин, В.А. Кедрус. – К.: Вища школ», 1977. – 280 с.
2. ГОСТ 26.205-88. Комплексы и устройства телемеханики. Общие технические условия [Текст]. – Взамен ГОСТ 26.205-83; введ. 2005.01.01. – М.: Государственный комитет СССР по стандартам. – 28 с.
3. Кожевников, В.Л. Основи збирання, обробки і передачі інформації. Теоретичні основи [Текст] / В.Л. Кожевников, А.В. Кожевников. – Д.: НГУ, 2007. – 108 с.
4. Кожевников, А.В. Засоби мультимедіа [Текст] / А.В. Кожевников, В.Л. Кожевников, В.О. Трусов. – Д.: НГУ, 2007. – 151 с.
5. Кожевников, А.В. Інформативність джерел двопозиційних повідомлень в системах телемеханіки [Текст] / А.В. Кожевников, В.Л. Кожевников // Збірник наукових праць. – 2008. – №31 – С. 112-117.
6. Цимбал, В.П. Теория информации и кодирования [Текст] / В.П. Цимбал. – К.: Вища школа, 1973. – 232 с.
7. Макаров, В.А. Теоретические основы телемеханіки [Текст] / В.А. Макаров. – Л.: ЛГУ, 1973. – 287 с.
8. Тутевич, В.Н. Телемеханика [Текст] / В.Н. Тутевич. – М.: Высшая школ», 1985. – 423 с.
9. Скляр, Б. Цифровая связь. Теоретические основы и практическое применение [Текст] / Б. Скляр. – М.: Издательский дом «Вильямс», 2003. – 1103 с.
10. Шляпоберский, В.И. Основы техники передачи дискретных сообщений [Текст] / В.И. Шляпоберский. – М.: Связь, 1973. – 480 с.
11. Кодирование информации (двоичные коды) [Текст] / Н.Т. Березюк, А.Г. Андрущенко, С.С. Мощицкий и др. – М.: Высшая школа, 1978. – 252 с.
12. Орлов, В.А. Теория информации в упражнениях и задачах [Текст] / В.А. Орлов, Л.И. Филиппов. – М.: Высшая школа, 1976. – 136 с.
13. Бовбель, Е.И. Элементы теории информации [Текст] / Е.И. Бовбель, И.К. Дайнеко, В.В. Изох. – Минск: БГУ, 1974. – 112 с.
14. Теория информации и кодирование [Текст] / Б.Б. Самсонов, Е.М. Плохов, А.И. Филоненков и др. – Ростов-на-Дону: Феникс, 2002. – 287 с.
15. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео [Текст] / Д. Ватолин, А. Ратушняк, М. Смирнов и др. – М.: ДИАЛОГ-МИФИ, 2002. – 384 с.
16. Боккер, П. Передача данных: техника связи в системах телеобработки данных [Текст]. Т.2. Устройства и системы: пер с нем. под ред. Д.Д. Кловского / П. Боккер. – М.: Радио и связь, 1981. – 256 с.
17. Мороз, Б.І. Методи раціональної організації обробки інформації в системах передачі даних [Текст] / Б.І. Мороз, В.П. Дюбко. – Д.: Академія митної служби України, 2007. – 251 с.
18. Безруков, В.В. Теорія інформації [Текст] / В.В. Безруков, В.Я. Кізяков, В.І. Профатілов. – Д.: ДДТУЗТ, 2001. – 110 с.

19. Золотарёв, В.В. Помехоустойчивое кодирование. Методы и алгоритмы [Текст]: справочник / В.В. Золотарёв, Г.В. Овечкин; под ред. чл.-кор. РАН Ю.Б. Зубарева. – М.: Горячая линия – Телеком, 2004. – 126 с.
20. Лидл, Р. Конечные поля [Текст]. Т.1.: пер. с англ. / Р. Лидл, Г. Нидеррайтер. – М.: Мир, 1988. – 430 с.

ДОДАТОК 1

Таблиця значень часткових ентропій $H = -p \cdot \log p$

p	Тисячні долі									
	0	1	2	3	4	5	6	7	8	9
0	0,0000	0,0100	0,0179	0,0251	0,0319	0,0382	0,0443	0,0501	0,0557	0,0612
0,01	0,0664	0,0716	0,0766	0,0814	0,0862	0,0909	0,0955	0,0999	0,1043	0,1086
0,02	0,1129	0,1170	0,1211	0,1252	0,1291	0,1330	0,1369	0,1407	0,1444	0,1481
0,03	0,1518	0,1554	0,1589	0,1624	0,1659	0,1693	0,1727	0,1760	0,1793	0,1825
0,04	0,1858	0,1889	0,1921	0,1952	0,1983	0,2013	0,2043	0,2073	0,2103	0,2132
0,05	0,2161	0,2190	0,2218	0,2246	0,2274	0,2301	0,2329	0,2356	0,2383	0,2409
0,06	0,2435	0,2461	0,2487	0,2513	0,2538	0,2563	0,2588	0,2613	0,2637	0,2662
0,07	0,2686	0,2709	0,2733	0,2756	0,2780	0,2803	0,2826	0,2848	0,2871	0,2893
0,08	0,2915	0,2937	0,2959	0,2980	0,3002	0,3023	0,3044	0,3065	0,3086	0,3106
0,09	0,3127	0,3147	0,3167	0,3187	0,3207	0,3226	0,3246	0,3265	0,3284	0,3303
0,1	0,3322	0,3341	0,3359	0,3378	0,3396	0,3414	0,3432	0,3450	0,3468	0,3485
0,11	0,3503	0,3520	0,3537	0,3555	0,3571	0,3588	0,3605	0,3622	0,3638	0,3654
0,12	0,3671	0,3687	0,3703	0,3719	0,3734	0,3750	0,3766	0,3781	0,3796	0,3811
0,13	0,3826	0,3841	0,3856	0,3871	0,3886	0,3900	0,3915	0,3929	0,3943	0,3957
0,14	0,3971	0,3985	0,3999	0,4012	0,4026	0,4040	0,4053	0,4066	0,4079	0,4092
0,15	0,4105	0,4118	0,4131	0,4144	0,4156	0,4169	0,4181	0,4194	0,4206	0,4218
0,16	0,4230	0,4242	0,4254	0,4266	0,4278	0,4289	0,4301	0,4312	0,4323	0,4335
0,17	0,4346	0,4357	0,4368	0,4379	0,4390	0,4401	0,4411	0,4422	0,4432	0,4443
0,18	0,4453	0,4463	0,4474	0,4484	0,4494	0,4504	0,4514	0,4523	0,4533	0,4543
0,19	0,4552	0,4562	0,4571	0,4581	0,4590	0,4599	0,4608	0,4617	0,4626	0,4635
0,2	0,4644	0,4653	0,4661	0,4670	0,4678	0,4687	0,4695	0,4704	0,4712	0,4720
0,21	0,4728	0,4736	0,4744	0,4752	0,4760	0,4768	0,4776	0,4783	0,4791	0,4798
0,22	0,4806	0,4813	0,4820	0,4828	0,4835	0,4842	0,4849	0,4856	0,4863	0,4870
0,23	0,4877	0,4883	0,4890	0,4897	0,4903	0,4910	0,4916	0,4923	0,4929	0,4935
0,24	0,4941	0,4947	0,4954	0,4960	0,4966	0,4971	0,4977	0,4983	0,4989	0,4994
0,25	0,5000	0,5006	0,5011	0,5016	0,5022	0,5027	0,5032	0,5038	0,5043	0,5048
0,26	0,5053	0,5058	0,5063	0,5068	0,5072	0,5077	0,5082	0,5087	0,5091	0,5096
0,27	0,5100	0,5105	0,5109	0,5113	0,5118	0,5122	0,5126	0,5130	0,5134	0,5138
0,28	0,5142	0,5146	0,5150	0,5154	0,5158	0,5161	0,5165	0,5169	0,5172	0,5176
0,29	0,5179	0,5182	0,5186	0,5189	0,5192	0,5196	0,5199	0,5202	0,5205	0,5208
0,3	0,5211	0,5214	0,5217	0,5220	0,5222	0,5225	0,5228	0,5230	0,5233	0,5235
0,31	0,5238	0,5240	0,5243	0,5245	0,5247	0,5250	0,5252	0,5254	0,5256	0,5258
0,32	0,5260	0,5262	0,5264	0,5266	0,5268	0,5270	0,5272	0,5273	0,5275	0,5277
0,33	0,5278	0,5280	0,5281	0,5283	0,5284	0,5286	0,5287	0,5288	0,5289	0,5291
0,34	0,5292	0,5293	0,5294	0,5295	0,5296	0,5297	0,5298	0,5299	0,5299	0,5300
0,35	0,5301	0,5302	0,5302	0,5303	0,5304	0,5304	0,5305	0,5305	0,5305	0,5306
0,36	0,5306	0,5306	0,5307	0,5307	0,5307	0,5307	0,5307	0,5307	0,5307	0,5307
0,37	0,5307	0,5307	0,5307	0,5307	0,5307	0,5306	0,5306	0,5306	0,5305	0,5305
0,38	0,5305	0,5304	0,5304	0,5303	0,5302	0,5302	0,5301	0,5300	0,5300	0,5299
0,39	0,5298	0,5297	0,5296	0,5295	0,5294	0,5293	0,5292	0,5291	0,5290	0,5289

<i>p</i>	Тисячні долі									
	0	1	2	3	4	5	6	7	8	9
0,4	0,5288	0,5286	0,5285	0,5284	0,5283	0,5281	0,5280	0,5278	0,5277	0,5275
0,41	0,5274	0,5272	0,5271	0,5269	0,5267	0,5266	0,5264	0,5262	0,5260	0,5258
0,42	0,5256	0,5255	0,5253	0,5251	0,5249	0,5246	0,5244	0,5242	0,5240	0,5238
0,43	0,5236	0,5233	0,5231	0,5229	0,5226	0,5224	0,5222	0,5219	0,5217	0,5214
0,44	0,5211	0,5209	0,5206	0,5204	0,5201	0,5198	0,5195	0,5193	0,5190	0,5187
0,45	0,5184	0,5181	0,5178	0,5175	0,5172	0,5169	0,5166	0,5163	0,5160	0,5157
0,46	0,5153	0,5150	0,5147	0,5144	0,5140	0,5137	0,5133	0,5130	0,5127	0,5123
0,47	0,5120	0,5116	0,5112	0,5109	0,5105	0,5102	0,5098	0,5094	0,5090	0,5087
0,48	0,5083	0,5079	0,5075	0,5071	0,5067	0,5063	0,5059	0,5055	0,5051	0,5047
0,49	0,5043	0,5039	0,5034	0,5030	0,5026	0,5022	0,5017	0,5013	0,5009	0,5004
0,5	0,5000	0,4996	0,4991	0,4987	0,4982	0,4978	0,4973	0,4968	0,4964	0,4959
0,51	0,4954	0,4950	0,4945	0,4940	0,4935	0,4930	0,4926	0,4921	0,4916	0,4911
0,52	0,4906	0,4901	0,4896	0,4891	0,4886	0,4880	0,4875	0,4870	0,4865	0,4860
0,53	0,4854	0,4849	0,4844	0,4839	0,4833	0,4828	0,4822	0,4817	0,4811	0,4806
0,54	0,4800	0,4795	0,4789	0,4784	0,4778	0,4772	0,4767	0,4761	0,4755	0,4750
0,55	0,4744	0,4738	0,4732	0,4726	0,4720	0,4714	0,4708	0,4702	0,4696	0,4690
0,56	0,4684	0,4678	0,4672	0,4666	0,4660	0,4654	0,4648	0,4641	0,4635	0,4629
0,57	0,4623	0,4616	0,4610	0,4603	0,4597	0,4591	0,4584	0,4578	0,4571	0,4565
0,58	0,4558	0,4551	0,4545	0,4538	0,4532	0,4525	0,4518	0,4511	0,4505	0,4498
0,59	0,4491	0,4484	0,4477	0,4471	0,4464	0,4457	0,4450	0,4443	0,4436	0,4429
0,6	0,4422	0,4415	0,4408	0,4401	0,4393	0,4386	0,4379	0,4372	0,4365	0,4357
0,61	0,4350	0,4343	0,4335	0,4328	0,4321	0,4313	0,4306	0,4298	0,4291	0,4283
0,62	0,4276	0,4268	0,4261	0,4253	0,4246	0,4238	0,4230	0,4223	0,4215	0,4207
0,63	0,4199	0,4192	0,4184	0,4176	0,4168	0,4160	0,4152	0,4145	0,4137	0,4129
0,64	0,4121	0,4113	0,4105	0,4097	0,4089	0,4080	0,4072	0,4064	0,4056	0,4048
0,65	0,4040	0,4031	0,4023	0,4015	0,4007	0,3998	0,3990	0,3982	0,3973	0,3965
0,66	0,3956	0,3948	0,3940	0,3931	0,3923	0,3914	0,3905	0,3897	0,3888	0,3880
0,67	0,3871	0,3862	0,3854	0,3845	0,3836	0,3828	0,3819	0,3810	0,3801	0,3792
0,68	0,3783	0,3775	0,3766	0,3757	0,3748	0,3739	0,3730	0,3721	0,3712	0,3703
0,69	0,3694	0,3685	0,3676	0,3666	0,3657	0,3648	0,3639	0,3630	0,3621	0,3611
0,7	0,3602	0,3593	0,3583	0,3574	0,3565	0,3555	0,3546	0,3537	0,3527	0,3518
0,71	0,3508	0,3499	0,3489	0,3480	0,3470	0,3460	0,3451	0,3441	0,3432	0,3422
0,72	0,3412	0,3403	0,3393	0,3383	0,3373	0,3364	0,3354	0,3344	0,3334	0,3324
0,73	0,3314	0,3305	0,3295	0,3285	0,3275	0,3265	0,3255	0,3245	0,3235	0,3225
0,74	0,3215	0,3204	0,3194	0,3184	0,3174	0,3164	0,3154	0,3144	0,3133	0,3123
0,75	0,3113	0,3102	0,3092	0,3082	0,3072	0,3061	0,3051	0,3040	0,3030	0,3020
0,76	0,3009	0,2999	0,2988	0,2978	0,2967	0,2956	0,2946	0,2935	0,2925	0,2914
0,77	0,2903	0,2893	0,2882	0,2871	0,2861	0,2850	0,2839	0,2828	0,2818	0,2807
0,78	0,2796	0,2785	0,2774	0,2763	0,2752	0,2741	0,2731	0,2720	0,2709	0,2698
0,79	0,2687	0,2676	0,2665	0,2653	0,2642	0,2631	0,2620	0,2609	0,2598	0,2587
0,8	0,2575	0,2564	0,2553	0,2542	0,2530	0,2519	0,2508	0,2497	0,2485	0,2474
0,81	0,2462	0,2451	0,2440	0,2428	0,2417	0,2405	0,2394	0,2382	0,2371	0,2359

p	Тисячні долі									
	0	1	2	3	4	5	6	7	8	9
0,82	0,2348	0,2336	0,2325	0,2313	0,2301	0,2290	0,2278	0,2266	0,2255	0,2243
0,83	0,2231	0,2219	0,2208	0,2196	0,2184	0,2172	0,2160	0,2149	0,2137	0,2125
0,84	0,2113	0,2101	0,2089	0,2077	0,2065	0,2053	0,2041	0,2029	0,2017	0,2005
0,85	0,1993	0,1981	0,1969	0,1957	0,1944	0,1932	0,1920	0,1908	0,1896	0,1884
0,86	0,1871	0,1859	0,1847	0,1834	0,1822	0,1810	0,1797	0,1785	0,1773	0,1760
0,87	0,1748	0,1736	0,1723	0,1711	0,1698	0,1686	0,1673	0,1661	0,1648	0,1636
0,88	0,1623	0,1610	0,1598	0,1585	0,1572	0,1560	0,1547	0,1534	0,1522	0,1509
0,89	0,1496	0,1484	0,1471	0,1458	0,1445	0,1432	0,1420	0,1407	0,1394	0,1381
0,9	0,1368	0,1355	0,1342	0,1329	0,1316	0,1303	0,1290	0,1277	0,1264	0,1251
0,91	0,1238	0,1225	0,1212	0,1199	0,1186	0,1173	0,1159	0,1146	0,1133	0,1120
0,92	0,1107	0,1093	0,1080	0,1067	0,1054	0,1040	0,1027	0,1014	0,1000	0,0987
0,93	0,0974	0,0960	0,0947	0,0933	0,0920	0,0907	0,0893	0,0880	0,0866	0,0853
0,94	0,0839	0,0826	0,0812	0,0798	0,0785	0,0771	0,0758	0,0744	0,0730	0,0717
0,95	0,0703	0,0689	0,0676	0,0662	0,0648	0,0634	0,0621	0,0607	0,0593	0,0579
0,96	0,0565	0,0552	0,0538	0,0524	0,0510	0,0496	0,0482	0,0468	0,0454	0,0440
0,97	0,0426	0,0412	0,0398	0,0384	0,0370	0,0356	0,0342	0,0328	0,0314	0,0300
0,98	0,0286	0,0271	0,0257	0,0243	0,0229	0,0215	0,0201	0,0186	0,0172	0,0158
0,99	0,0144	0,0129	0,0115	0,0101	0,0086	0,0072	0,0058	0,0043	0,0029	0,0014

Фрагмент таблиці незвідних багаточленів

μ	$P(x)$	$n = 2^\mu - 1$
1	$x+1^*$	1
2	x^2+x+1^*	3
3	x^3+x+1^* $x^3+x^2+1^*$	7
4	x^4+x+1^* $x^4+x^2+1^*$ $x^4+x^3+x^2+x+1$	15
5	$x^5+x^2+1^*$ $x^5+x^3+1^*$ $x^5+x^3+x^2+x+1^*$ $x^5+x^4+x^2+x+1^*$ $x^5+x^4+x^3+x+1^*$ $x^5+x^4+x^3+x^2+1^*$	31
6	x^6+x+1^* x^6+x^3+1 $x^6+x^5+1^*$ $x^6+x^4+x^2+x+1$ $x^6+x^4+x^3+x+1^*$ $x^6+x^5+x^2+x+1^*$ $x^6+x^5+x^3+x^2+1^*$ $x^6+x^5+x^4+x+1^*$ $x^6+x^5+x^4+x^2+1$	63
7	x^7+x+1^* $x^7+x^3+1^*$ $x^7+x^4+1^*$ $x^7+x^6+1^*$ $x^7+x^3+x^2+x+1^*$ $x^7+x^4+x^3+x^2+1^*$ $x^7+x^5+x^2+x+1^*$ $x^7+x^5+x^3+x+1^*$ $x^7+x^5+x^4+x^3+1^*$	127
8	$x^8+x^4+x^3+x^2+1^*$ $x^8+x^6+x^5+x^4+1^*$	255
9	$x^9+x^4+1^*$ $x^9+x^5+1^*$	511
10	$x^{10}+x^3+1^*$ $x^{10}+x^7+1^*$	1023

*) – примітивні багаточлени.

Однобальні тести для контролю залишкових знань з дисципліни

В наведених тестах лише одна відповідь на кожне питання є правильною

1. Функція залежності ентропії повідомлення від його ймовірності є:
 - 1) лінійною
 - 2) гіперболічною
 - 3) експоненційною
 - 4) логарифмічною
2. Величина, яка не є одиницею виміру об'єму інформації:
 - 1) байт
 - 2) бод
 - 3) ніт
 - 4) діт
3. З урни, в якій містяться 20 білих, по 15 червоних та синіх та 10 чорних куль, вилучається одна. Найменшу інформацію несе повідомлення, що вилучена куля має колір:
 - 1) чорний
 - 2) білий
 - 3) червоний або синій
 - 4) червоний
4. З урни, в якій містяться 20 білих, по 15 червоних та синіх та 10 чорних куль, вилучається одна. Найбільшу інформацію несе повідомлення, що вилучена куля має колір:
 - 1) чорний
 - 2) білий
 - 3) червоний або синій
 - 4) червоний
5. З преферансної колоди, в якій міститься 32 карти, вилучається одна. Максимальна кількість бітів інформації, яка несе відповідь «так» чи «ні» на питання про те, що вилучена карта є піковим тузом, складає:
 - 1) 1
 - 2) 2
 - 3) 3
 - 4) 5
6. Кількість інформації повідомлення про те, що у грі «морський бій» з полем 10×10 кліток влучення в чотирипалубний корабель відбулося на першому ході, дорівнює в бітах:
 - 1) 1.40
 - 2) 3.22
 - 3) 4.64
 - 4) 5.64
7. З гармати здійснюється два незалежних постріли. Ймовірність влучення в ціль при одному пострілі дорівнює p . Кількість інформації, про те, що в ціль влучено хоча б одним пострілом, складає:
 - 1) $\log p$
 - 2) $\log(1-p)$
 - 3) $\log[p(2-p)]$
 - 4) $\log[2p(1-p)]$
8. З гармати здійснюється два незалежних постріли. Ймовірність влучення в ціль при одному пострілі дорівнює p . Кількість інформації, про те, що в ціль влучено тільки одним пострілом, складає:
 - 1) $\log p$
 - 2) $\log(1-p)$
 - 3) $\log[p(2-p)]$
 - 4) $\log[2p(1-p)]$
9. Проводиться послідовність незалежних випробувань. Ймовірність появи події A при одному випробуванні складає p . Проведено k випробувань. Кількість інформації повідомлення про те, що в останньому іспиті мала місце подія A , якщо іспити проводяться до першої появи події A , тобто подія виникла у останньому випробуванні, складає:
 - 1) $\log[p(1-p)^{k-1}]$
 - 2) $\log p$
 - 3) $\log[p^{k-1}(1-p)]$
 - 4) $\log(1-p)$
10. Проводиться послідовність незалежних випробувань. Ймовірність появи події A при одному випробуванні складає p . Проведено k випробувань. Кількість інформації повідомлення про те, що в останньому іспиті мала місце подія A , якщо іспити проводяться незалежно від появи події, складає:
 - 1) $\log[p(1-p)^{k-1}]$
 - 2) $\log p$
 - 3) $\log[p^{k-1}(1-p)]$
 - 4) $\log(1-p)$

11. Повідомлення i з ймовірностями $\{p_i\}_{i=1,\dots,m}$ джерела, середня питома ентропія якого визначається виразом $H = -\sum_{i=1}^m p_i \log p_i$, не повинні задовільняти умові:

- 1) дискретності
- 2) неперервності
- 3) несумісності
- 4) статистичної незалежності

12. Середня питома ентропія джерела дискретних повідомлень може приймати множину значень:

- 1) $[0;1]$
- 2) $[0;+\infty)$
- 3) $(-\infty;+\infty)$
- 4) $[1;+\infty)$

13. Максимум часткової ентропії повідомлення досягається при його ймовірності:

- 1) 0,36
- 2) 0,44
- 3) 0,5
- 4) 1

14. Ентропія джерела дискретних повідомлень є максимальною, якщо ймовірності повідомлень підпорядковуються розподілу:

- 1) рівномірному
- 2) біноміальному
- 3) геометричному
- 4) Пуассона

15. Ентропія джерела N дорівнює $\log N$, якщо ймовірності повідомлень підпорядковуються розподілу:

- 1) рівномірному
- 2) біноміальному
- 3) геометричному
- 4) Пуассона

16. Джерело X генерує повідомлення $\{x_i\}_{i=1,\dots,n} = \frac{1}{i}$ з ймовірностями

$\{p_i\}_{i=1,\dots,n} = \frac{1}{n}$, а джерело Y – повідомлення $\{y_i\}_{i=1,\dots,n} = i$ з тими самими ймовірностями. Ентропії джерел та Y співвідносяться таким чином:

- 1) однакові
- 2) у джерела X менше
- 3) у джерела X більше
- 4) у джерела X може бути як менше, так і більше в залежності від значення n

17. Джерело X генерує повідомлення $\{x_i\}_{i=1,\dots,n} = i$ з ймовірностями $\{p_i\}_{i=1,\dots,n} = \frac{1}{n}$, а джерело Y – повідомлення $\{y_i\}_{i=1,\dots,2n} = i$ з ймовірностями

$\{p_i\}_{i=1,\dots,2n} = \frac{1}{2n}$. Ентропії джерел та Y співвідносяться таким чином:

- 1) однакові
- 2) у джерела X менше
- 3) у джерела X більше
- 4) у джерела X може бути як менше, так і більше в залежності від значення n

18. Ентропія джерела повідомлень з m літер алфавіта, вважаючи, що загальна кількість літер в алфавіті дорівнює k і всі повідомлення рівноймовірні, становить:

- 1) $m \log k$
- 2) $k \log m$
- 3) $\log(km)$
- 4) $\log(k+m)$

19. Ентропія джерела дискретних повідомлень при виникненні взаємозалежності повідомлень:

- 1) не змінюється
- 2) збільшується
- 3) зменшується
- 4) може як збільшуватися, так і зменшуватися

20. Ентропія джерела повідомлень з ймовірностями $\{p_i\}_{i=1,\dots,4} = \{0.5, 0.25, 0.25\}$ в бітах складає:

- 1) 1.25 2) 1.5 3) 1.75 4) 2
21. До неперервного не відноситься ймовірнісний розподіл:
1) рівномірний 2) нормальний 3) гіпергеометричний 4) експонентний
22. Ентропія складного повідомлення розглядається як ентропія системи джерела первинних повідомлень X та перешкод, які впливають на повідомлення у каналі зв'язку, Y . Для випадків, якщо X та Y взаємонезалежні, ентропія системи дорівнює:
1) $H(X)$ 2) 1 3) 0 4) $H(X) + H(Y)$
23. Ентропія складного повідомлення розглядається як ентропія системи джерела первинних повідомлень X та перешкод, які впливають на повідомлення у каналі зв'язку, Y . Для випадків, якщо повідомлення X однозначно викривляється перешкодою Y , ентропія системи дорівнює:
1) $H(X)$ 2) 1 3) 0 4) $H(X) + H(Y)$
24. Порядок дискретного ергодичного джерела повідомлень, у якого ймовірність наступного повідомлення залежить тільки від попереднього дорівнює:
1) 0 2) 1 3) 2 4) 3
25. Розмірністю швидкості передачі сигналів є:
1) біт 2) біт·с 3) біт/с 4) бод
26. Розмірністю швидкості передачі інформації є:
1) біт 2) біт·с 3) біт/с 4) бод
27. Розмірністю пропускної здатності каналу передачі інформації є:
1) біт 2) біт·с 3) біт/с 4) бод
28. При відсутності перешкод швидкість передачі інформації в системі джерело – канал – приймач при зростанні ентропії джерела:
1) не змінюється 2) збільшується 3) зменшується
4) може як збільшуватися, так і зменшуватися в залежності від швидкості
29. Найменша пропускна здатність симетричного каналу для двійкових повідомлень досягається при ймовірності помилкового приймання сигналу:
1) 0.3 2) 0.5 3) 0.8 4) 1
30. Найбільша пропускна здатність симетричного каналу для двійкових повідомлень досягається при ймовірності помилкового приймання сигналу:
1) 0.3 2) 0.5 3) 0.8 4) 1
31. Кількість співмножників у виразі для об'єму сигналу складає:
1) 2 2) 3 3) 4 4) 5
32. При збільшенні ширини смуги частот каналу зв'язку його пропускна здатність:
не змінюється збільшується прямо пропорційно
зменшується прямо пропорційно збільшується логарифмічно
33. При збільшенні відношення потужностей сигналу і перешкоди пропускна здатність каналу зв'язку:
не змінюється збільшується прямо пропорційно
зменшується прямо пропорційно збільшується логарифмічно
34. Величина надлишковості повідомлень може приймати множину значень:
1) $[0;1]$ 2) $[0;+\infty)$ 3) $[0;1)$ 4) $(0;1]$

35. При виникненні взаємозалежності повідомлень і збереженні розподілу безумовних ймовірностей повідомлень величина статистичної надлишковості:
 1) не змінюється 2) збільшується 3) зменшується
 4) може як збільшуватися, так і зменшуватися в залежності від розподілу
36. При виникненні взаємозалежності повідомлень і збереженні розподілу безумовних ймовірностей повідомлень, величина ймовірної надлишковості:
 1) не змінюється 2) збільшується 3) зменшується
 4) може як збільшуватися, так і зменшуватися в залежності від розподілу
37. Повідомлення джерела кодуються комбінаціями – a, b, c, ab, bc . Основа коду становить:
 1) 2 2) 3 3) 4 4) 5
38. Загальна кількість кодкових комбінацій n розрядного двійкового коду складає:
 1) n^2 2) $2n$ 3) 2^n 4) C_n^2
39. Джерело інформації генерує повідомлення $\{p_i\}_{i=1,\dots,4} = \{0.5, 0.25, 0.125, 0.125\}$, що кодуються двійковим кодом. Середня довжина кодової комбінації не може бути менше за:
 1) 1.25 2) 1.5 3) 1.75 4) 2
40. Для джерела дискретних повідомлень з питомою середньою ентропією $H = 12$ бітів, що кодується за допомогою кода з основою $m = 8$, середня довжина кодової комбінації не може бути менше ніж:
 1) 2.5 2) 3 3) 3,5 4) 4
41. Код Хаффмена можна охарактеризувати як:
 1) рівномірний 2) оптимальний 3) надлишковий 4) завадостійкий
42. До оптимальних не відноситься код:
 1) Хаффмена 2) Шеннона – Фено 3) з постійною вагою
 4) у якого середня довжина кодового слова дорівнює $\frac{H(X)}{\log m}$, де m – основа коду
43. Джерело інформації генерує 12 повідомлень, що кодуються нормальним двійковим кодом. Мінімальна достатня кількість розрядів коду складає:
 1) 2 2) 3 3) 4 4) 5
44. Кількість заборонених тетрад двійководосяткового коду складає:
 1) 2 2) 4 3) 6 4) 8
45. До завадостійких не відноситься код:
 1) числоімпульсний 2) інверсний 3) кореляційний
 4) з контрольними сумами
46. Властивістю завадостійкості володіє код:
 1) Хеммінга 2) Хаффмена 3) НДК 4) Грея
47. Відстань між кодovими комбінаціями 01110 та 10011 складає:
 1) 2 2) 3 3) 4 4) 5
48. Відстань між кодovими комбінаціями $abbb$ та $aabc$ складає:
 1) 1 2) 2 3) 3 4) 4
49. Кодова відстань НДК становить:

- 1) 1 2) 2 3) 3 4) 4

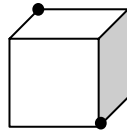
50. Найбільша кратність помилок, які може виявляти трирозрядний код з дозволеними комбінаціями 000, 110, 011, 101, складає:

- 1) не виявляє взагалі 2) 1 3) 2 4) 3

51. Найбільша кратність помилок, які може виправляти трирозрядний код з дозволеними комбінаціями 000, 110, 011, 101, складає:

- 1) не виправляє взагалі 2) 1 3) 2 4) 3

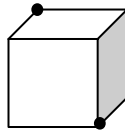
52. Припустимі кодові комбінації трирозрядного двійкового коду відмічені на геометричній моделі



Код виявляє всі помилки найбільшої кратності:

- 1) не виявляє взагалі 2) 1 3) 2 4) 3

53. Припустимі кодові комбінації трирозрядного двійкового коду відмічені на геометричній моделі



54. Код виправляє всі помилки найбільшої кратності:

- 1) не виправляє взагалі 2) 1 3) 2 4) 3

55. Максимальна кратність помилок, які може виявляти код з відстанню 5, складає:

- 1) 1 2) 2 3) 3 4) 4

56. Максимальна кратність помилок, які може виправляти код з відстанню 5, складає:

- 1) 1 2) 2 3) 3 4) 4

57. Мінімальна кодова відстань коду, який може виявляти двократні помилки, складає:

- 1) 2 2) 3 3) 4 4) 5

58. Мінімальна кодова відстань коду, який може виправляти двократні помилки, складає:

- 1) 2 2) 3 3) 4 4) 5

59. Серед зображених кодів найменшу надлишковість в середньому має код:

- 1) нормальний двійковий 2) Шеннона – Фено
3) який виявляє одну помилку 4) який виправляє одну помилку

60. Серед зображених кодів найбільшу надлишковість в середньому має код:

- 1) нормальний двійковий 2) Шеннона – Фено
3) який виявляє одну помилку 4) який виправляє одну помилку

61. Зі зростанням числа перевірочних (контрольних) розрядів систематичного завадостійкого двійкового коду його надлишковість:

- 1) не змінюється
2) зменшується зворотно пропорційно
3) збільшується прямо пропорційно

- 4) може як збільшуватися, так і зменшуватися в залежності від кількості розрядів
62. Код «пар» виявляє помилки кратності.
- 1) тільки однократні
 - 2) тільки однократні і двократні
 - 3) всі помилки парної кратності
 - 4) всі помилки непарної кратності
63. Кількість інформаційних розрядів коду «пар» становить k . Загальна довжина його кодової комбінації складає:
- 1) $k+1$
 - 2) $2k$
 - 3) $2k+1$
 - 4) $2(k+1)$
64. Кількість інформаційних розрядів КК становить k . Загальна довжина його кодової комбінації складає:
- 1) $k+1$
 - 2) $2k$
 - 3) $2k+1$
 - 4) $2(k+1)$
65. Кількість інформаційних розрядів ІК становить k . Загальна довжина його кодової комбінації складає:
- 1) $k+1$
 - 2) $2k$
 - 3) $2k+1$
 - 4) $2(k+1)$
66. Кількість дозволених кодових комбінацій n розрядного коду з постійною вагою w складає:
- 1) nw
 - 2) n^w
 - 3) w^n
 - 4) C_n^w
67. Систематичний код з кодовою відстанню 3 використовується для кодування 15-ти повідомлень. Мінімальна достатня кількість контрольних розрядів становить:
- 1) 1
 - 2) 2
 - 3) 3
 - 4) 4
68. Кількість контрольних розрядів ЛСК $G_{(n,k)}$ становить:
- 1) n
 - 2) k
 - 3) $n-k$
 - 4) $n+k$
69. Розмір інформаційної підматриці (кількість рядків \times кількість стовпців) твірної матриці ЛСК $G_{(n,k)}$ становить:
- 1) $n \times n$
 - 2) $k \times k$
 - 3) $n \times k$
 - 4) $k \times n$
70. Розмір контрольної підматриці (кількість рядків \times кількість стовпців) твірної матриці ЛСК $G_{(n,k)}$ становить:
- 1) $k \times k$
 - 2) $n-k \times n-k$
 - 3) $k \times n-k$
 - 4) $n-k \times k$
71. Розмір перевіркової матриці (кількість рядків \times кількість стовпців) ЛСК $G_{(n,k)}$ становить:
- 1) $n \times n$
 - 2) $n-k \times n-k$
 - 3) $n \times n-k$
 - 4) $n-k \times n$
72. Надлишковість ЛСК $G_{(6,3)}$ становить:
- 1) 0,3
 - 2) 0,4
 - 3) 0,5
 - 4) 0,6
73. Значення перевірочних (контрольних) розрядів у ЛСК визначаються через значення інформаційних розрядів за допомогою операції:
- 1) заперечення
 - 2) логічного додавання
 - 3) логічного множення
 - 4) додавання за модулем два
74. Про помилку (якщо вона має місце) лінійного систематичного коду $G_{(7,4)}$ 0101111, у якого контрольні розряди становлять $y_1=x_1 \oplus x_2 \oplus x_3$, $y_2=x_2 \oplus x_3 \oplus x_4$, $y_3=x_1 \oplus x_3 \oplus x_4$ свідчить розряд:
- 1) y_1
 - 2) y_2
 - 3) y_3
 - 4) помилка відсутня

75. Помилка, якщо вона має місце в лінійному систематичному коді $G_{(6,3)}$ 101011, у якого контрольні елементи становлять $y_1=x_1\oplus x_2$, $y_2=x_2\oplus x_3$, $y_3=x_1\oplus x_3$, знаходиться у інформаційному розряді з номером:
- 1) 1 2) 2 3) 3 4) помилка відсутня
76. Двійковим еквівалентом багаточлену x^6+x^4+x+1 є комбінація:
- 1) 1010011 2) 0101100 3) 101001 4) 010110
77. Двійковим еквівалентом залишку від ділення багаточлену x^3+1 на x^2+x+1 є комбінація:
- 1) 00 2) 01 3) 10 4) 11
78. Надлишковість ЦК $G_{(7,4)}$ становить:
- 1) 0,32 2) 0,43 3) 0,5 4) 0,77
79. Кодова відстань ЦК $G_{(7,3)}$ становить:
- 1) 2 2) 3 3) 4 4) 5
80. Кількість розрядів регістру зсуву ПК ЗК становить K . Регулярність кодового дерева виникає на такті роботи регістру:
- 1) 1 2) K 3) $K-1$ 4) $K+1$
81. Сигнали, які утворюються внаслідок квадратурної амплітудної маніпуляції, можуть мати різні значення:
- 1) амплітуди і частоти 2) фази і частоти
3) амплітуди і фази 4) амплітуди, фази і частоти
82. Сусіднім крапкам в радіальному напрямку на фазоамплітудній діаграмі ГК присвоюються коди:
- 1) Хеммінга 2) Хаффмена 3) НДК 4) Грея
83. ЦК із твірним багаточленом $P(x) = x^3 + x + 1$ виявляє всі помилки кратності:
- 1) тільки однократні 2) тільки двократні
3) тільки однократні і двократні 4) всі помилки кратності до 3 включно
84. ЦК із твірним багаточленом $P(x) = (x + 1)(x^3 + x + 1)$ виявляє всі помилки кратності:
- 1) тільки однократні 2) тільки двократні
3) тільки однократні і двократні 4) всі помилки кратності до 3 включно
85. Незвідний двійковий багаточлен не обов'язково задовільняє умові:
- 1) багаточлен без залишку ділиться тільки на себе і на одиницю
2) всі коефіцієнти багаточлену є нулями або одиницями
3) вільним членом є одиниця
4) порядок багаточлену є непарним
86. Незвідним не є двійковий багаточлен:
- 1) $x+1$ 2) x^2+x+1 3) x^3+1 4) x^3+x+1
87. Перші чотири члени псевдовипадкової двійкової послідовності 1010, згенеровані на основі незвідного багаточлену $P(x)=x^4+x+1$ і значення початкової фази 0101. Помилка має місце у позиції:
- 1) 1 2) 2 3) 3 4) 4
88. Період повторення псевдовипадкової двійкової послідовності, яка згенерована на основі примітивного багаточлену порядку n становить:
- 1) n 2) $2n$ 3) 2^n 4) 2^n-1

ПРЕДМЕТНИЙ ПОКАЖЧИК

- Автокореляційна функція (АКФ) 87
дискретна 87
Алгоритм
виправлення помилок ЦК 79
Шеннона – Фено 47
Хаффмана 48
Багаточлен
двійковий 71
незвідний 72
примітивний 86
твірний
– ПВП 85
– ЦК 72
багаточлену 54
Вага
коду 59
багаточлену 54
Вектор помилки 67
Відстань
кодова 55
між кодовими комбінаціями 54
Геометричне подання коду 53
Граф
двійкового симетричного каналу 25
– зі стиранням 27
кінцевого автомату 80
Гратчаста діаграма 84
Джерело повідомлень 4
двійкове 9
двопозиційне 19
дискретне 6
ергодичне 18
неперервне 13
Інформація 4
Ентропія
апріорна 5
двійкового джерела 9
диференціальна 14
двопозиційного джерела 19
ергодичного джерела 18
залишкова 5
ймовірнісних розподілів 11
об'єднання дослідів 22
повна 7
середня питома 7
умовна 22
часткова 7
Канал зв'язку 4
двійковий симетричний 25
– зі стиранням 27
Кінцевий автомат 80
Квадратурна амплітудна маніпуляція
(КАМ) 82
Класифікація кодів 41
Код 44
4В/5В 89
CRC-16 77
гратчастий 83
Грея 52
двійководесятковий (ДДК) 52
завадостійкий 41
згортковий 80
з захистом по паритету («пар») 57
з контрольною сумою (Бергера) 62
з постійною вагою (на одне
сполучення) 59
інверсний (ІК) 61
канальний 83
кореляційний (КК) 61
лінійний систематичний (ЛСК) 64
незавадостійкий (звичайний) 50
нерівномірний 41
нормальний двійковий (НДК) 50
оптимальний 48
рівномірний 41
роздільний 42
систематичний 53
Шеннона – Фено 47
шумовий 85
Хаффмана 49
Хеммінга 68
циклічний (ЦК) 71
числоімпульсний (унітарний) 50
Кодове дерево 82
Кратність помилок
що виявляються 55

що виправляються 56

Матриця

- канальна 21
- перевірочна 67
- твірна 65, 73

Метод

- виявлення і виправлення помилок
 - мажоритарного декодування 70
 - перевірочних рівнянь 66
 - синдромів 67, 78
- матричний побудови
- ЛСК 65
- ЦК 73

Міра інформації 5

Надлишковість

- коду 55
- систематичного 55
- повідомлення 36
 - ймовірнісна 36
 - статистична 36

Нерівність Крафта 43

Об'єм

- алфавіту 41
- каналу 32
- сигналу 32

Одиниця виміру

- кількості інформації 6
- пропускної спроможності каналу зв'язку 24
- швидкості передачі 24
 - інформації 24
 - сигналів 24

Пакет помилок 77

Підматриця

- інформаційна 65
- контрольна 65, 73

Пропускна здатність каналу зв'язку

- дискретного без прешкод 23
- двійкового симетричного 25
 - зі стиранням 27
- при енергетичному підході 30
 - і широкопasmовою прешкодою 32

Послідовність

- псевдовипадкова (ПВП) 85
- Хаффмена 85

Потужність

- сигналу 29
- перешкоди 29
 - широкопasmової 31

Розподіл ймовірнісний 11, 13

- дискретний 11
 - біноміальний 12
 - геометричний 38
 - Пуассона 12
 - рівномірний 11
- неперервний 13
 - експонентний 16
 - нормальний 16
 - рівномірний 15

Розряд кодової комбінації

- інформаційний 53
- контрольний 53

Основа коду 41

Синдром 67

Синхронізація 88

Скремблерування даних 88

Теорема Котельникова 29

Умова

- узгодження сигналу й каналу 33
 - загальна 33
 - часткові 33

Хеммінга 63

Цілі кодування 41

Швидкість передачі

- інформації 24
- сигналів 24

Ширина смуги частот

- сигналу 30
- каналу 30
 - з широкопasmовою прешкодою 31

ЗМІСТ

Передмова	3
Вступ	4
1. Основи статистичної теорії інформації	5
1.1. Кількість інформації й ентропія	5
1.2. Ентропія дослідів з нерівномірними взаємозалежними наслідками	6
1.3. Дослідження властивостей часткової ентропії	7
1.4. Дослідження ентропії двійкового джерела. Умова максимуму ентропії дискретного джерела при $m > 2$	9
1.5. Ентропія типових дискретних розподілів	11
1.6. Ентропія типових неперервних розподілів	13
1.7. Ентропія ергодичних послідовностей	17
1.8. Ентропія найпростіших ергодичних джерел	18
1.9. Ентропія двопозиційних ергодичних джерел	19
1.10. Ентропія об'єднань дослідів. Умовна ентропія	20
1.11. Пропускна здатність дискретного каналу без перешкод	23
1.12. Пропускна здатність каналу з перешкодами (статистичний підхід)	24
1.13. Передача інформації в каналі з перешкодами. Пропускна здатність каналу. Енергетичний підхід	28
1.14. Об'єми сигналів і каналів, їхнє узгодження	32
1.15. Надлишковість повідомлень	35
2. Основи теорії й практики кодування повідомлень у системах передачі технологічної інформації	41
2.1. Загальні положення. Класифікація кодів	41
2.2. Постановка задачі кодування в каналах без перешкод. Нерівність Крафта	42
2.3. Основна теорема кодування в каналах без перешкод і її наслідок	44
2.4. Алгоритм Шеннона – Фено	47
2.5. Алгоритм Хаффмана	48
2.6. Незавадостійкі (звичайні) коди	50
2.6.1. Числоімпульсний (унітарний) код	50
2.6.2. Нормальний двійковий код	50
2.6.3. Код Грея	51
2.6.4. Двійководесятковий код	52
2.7. Загальні властивості кодів, що виявляють і виправляють помилки	53
2.7.1. Геометричне подання комбінацій двійкових кодів	53
2.7.2. Виявлення помилок у кодових комбінаціях	54
2.7.3. Виявлення й виправлення помилок у кодових комбінаціях	55
2.8. Найпростіші коди, що виявляють помилки	56

2.8.1. Код «пар» або код із захистом по паритету.....	57
2.8.2. Коди з постійною вагою (коди на одне сполучення)	58
2.8.3. Кореляційні, інверсний і коди з контрольними сумами...	59
2.9. Лінійні коди	61
2.9.1. Загальні відомості. Постановка задачі кодування з виявленням і виправленням помилок	61
2.9.2. Основні властивості ЛСК, матричний метод кодування повідомлень.....	62
2.9.3. Методи декодування ЛСК з виявленням або виявленням і виправленням помилок.....	64
2.10. Циклічні коди (ЦК)	69
2.10.1. Загальні відомості. Двійкове й алгебраїчне подання комбінацій ЦК	69
2.10.2. Методика побудови систематичного ЦК.....	70
2.10.3. Матричний метод побудови систематичного ЦК.....	71
2.10.4. Властивості найпростіших ЦК по виявленню помилок і вибір необхідного твірного багаточлена	73
2.10.5. Виправлення помилок у ЦК.....	76
2.10.6. Згорткові коди	78
2.10.7. Канальне кодування	81
Список літератури	91
Додаток 1. Таблиця значень часткових ентропії.....	93
Додаток 2. Фрагмент таблиці незвідних багаточленів	96
Додаток 3. Однобальні тести для контролю залишкових знань з дисципліни	97
Предметний покажчик	104

Навчальне видання

Кожевников В'ячеслав Леонідович
Кожевников Антон В'ячеславович

ТЕОРІЯ ІНФОРМАЦІЇ ТА КОДУВАННЯ

Навчальний посібник

Редактор Є.М. Ільченко

Підписано до друку 08.06.2012 р. Формат 30x42/4
Папір офсетний. Ризографія. Ум. друк. арк. 6,3.
Обл.-вид. арк. 8,6. Тираж 100 пр. Зам. №

Підготовлено до друку та видруковано
у Національному гірничому університеті
Свідоцтво про внесення до Державного реєстру ДК № 1842
від 11.06.2004 р.

49005, м. Дніпропетровськ, просп. К. Маркса, 19