**Lec\_5**

**Класифікація кодів і їх характеристика**

**Способи подання кодів**

*Передача інформації по лініях і каналах зв 'язку пов 'язана з певними труднощами. Головним чином* — *це дія завад на сигнали, що несуть інформацію. Тому для усунення помилок, які виникають при передачі сигналів під впливом завад, повідомлення кодують, оскільки кодування є найефективнішим способом їх захисту.*

*Для кращого розуміння процесу кодування повідомлень треба знати системи числення, деякі основні операції над елементами поля GF(q), способи подання кодів, основні теореми кодування для каналів. Про все це йтиметься в цьому розділі. Крім того, тут викладено класифікацію кодів, наведено характеристики їх, описано основні принципи оптимального кодування.*

5.1. КЛАСИФІКАЦІЯ КОДІВ І ХАРАКТЕРИСТИКИ ЇХ

*Кодування* — це процес перетворення повідомлення на впорядкований набір символів, елементів, знаків. При кодуванні кожному повідомленню ставиться у відповідність зумовлена кодова комбінація — набір символів (елементів, знаків) з деякої скінченної кількості їх, яка називається *алфавітом.*

Розрізняють дві групи кодів: ***безнадмірні***(некоректувальні, первинні, прості) та ***надмірні***(завадостійкі). Перші не дають змоги виявити та виправити спотворені елементи в своїх комбінаціях, другі — забезпечують можливість виявлення та/або виправлення елементів кодових комбінацій, спотворених унаслідок дії завад.

У надмірних кодах комбінації можуть мати інформаційні та перевірні елементи. Обидві групи кодів поділяються на ***рівномірні***та ***нерівномірн****і,* тобто коди зі сталою та змінною кількістю розрядів.

Надмірні коди також бувають ***неперервними***(рекурентними) і ***блоковими****.* В неперервних кодах процес кодування та декодування має неперервний характер, у блокових — кожному повідомленню відповідає кодова комбінація (блок) зі скінченної кількості елементів. Блоки кодуються та декодуються окремо.

Блокові коди, в свою чергу, можуть бути ***подільними***та ***неподільними****.* До перших належать коди, що будуються доповненням інформаційних елементів перевірними; до других — коди, в яких немає чітко зумовлених інформаційних і перевірних елементів.

Подільні блокові коди бувають ***систематичними*** та ***несистематичними***. *Систематичним подільним блоковим* кодом називається такий код, у комбінаціях якого перші *k* позицій (розрядів) зайнято інформаційними елементами, а решту *r=n-k* позицій — перевірними, де *п* — загальна кількість позицій в кодовій комбінації. До *несистематичних подільних блокових* кодів належать коди, в яких інформаційними елементами не зайнято усі *k* перших позицій.

Різновидом подільних систематичних блокових кодів є ***циклічні***коди.

При виборі кодів для передачі інформації керуються вимогами до вірогідності інформації, що передається, та швидкості передачі, які визначаються такими характеристиками кодів:

• кількістю *k* інформаційних елементів;

• кількістю *r* перевірних елементів (для коректувальних кодів);

• довжиною (розрядністю) *п* — кількістю елементів (символів), які входять до складу кодової комбінації (n=*к+r);*

• основою (алфавітом) *q;*

• потужністю *Nd* — кількістю дозволених кодових комбінацій, що використовуються для передачі повідомлень;

• повною кількістю N кодових комбінацій, тобто кількістю всіх можливих комбінацій, яка дорівнює *qn* (для двійкових кодів *N = 2n);*

• надмірністю

 (для неподільних кодів)

або

*Rнад=*1-k/n=r/n(для подільних кодів при *Nd = 2k N= 2n);*

• відносною швидкістю *R,* яка характеризує ступінь використання в надмірному коді інформаційних можливостей його потужності, причому вагою кодової комбінації (для двійкового коду визначається кількістю одиниць у ній);

 або *R=*k/n=1-Rнад

• мінімальною кодовою відстанню *d =* min *dij,* тобто мінімальною відстанню між парами кодових комбінацій



де *аli, alj,*— елементи, що знаходяться в *l-му* місці в *і-* та j-й комбінаціях. Це значить, що *dij* визначається кількістю однойменних розрядів з різними значеннями;

• імовірністю *Рнв п* невиявленої помилки, тобто ймовірністю такої події, за якої прийнята кодова комбінація відрізняється від переданої, а властивості коду не дають змоги визначити факт наявності помилки;

• імовірністю *Рв* п виявленої помилки, тобто ймовірністю, за якої прийнята кодова комбінація відрізняється від переданої і завдяки властивостям коду встановлюється факт наявності помилки в кодовій комбінації;

• імовірністю Рвп виправленої помилки, тобто ймовірністю такої події, за якої прийнята кодова комбінація відрізняється від переданої і завдяки властивостям коду виправляється помилка в кодовій комбінації;

• імовірністю *Рп* виникнення помилки, тобто ймовірністю такої події, за якої прийнята кодова комбінація відрізняється від переданої;

• кратністю v помилки, що визначається кратністю vB виявлених та vBП виправлених помилок;

• ефективністю



де *Pi* - ймовірність виявленої або виправленої помилки залежно від властивостей коду.

Ступінь захисту інформації від помилок відповідним способом кодування залежить головним чином від мінімальної кодової відстані *dmin* даного коду [12].

Розрізняють три види кодової відстані: Хеммінга, Лі та матричну. Перша знайшла найбільше поширення в теорії кодування. Кодова відстань Хеммінга нероздільно пов'язана з поняттям *ваги w кодової комбінації*— кількістю її елементів, які не дорівнюють нулю.

Кодова відстань Хеммінга *d* між двома комбінаціями однієї довжини *п* визначається як кількість однойменних розрядів (позицій), які мають неоднакові елементи. Так, для двійкових кодів, оскільки в двійковій арифметиці додавання однакових елементні дає 0, а неоднакових — 1, відстань Хеммінга між двома кодовими комбінаціями можна визначити порозрядним додаванням їх за модулем 2 та подальшим підрахунком кількості ненульових елементів, тобто визначенням ваги *w* такої суми.

Загальна кількість кодових комбінацій завдовжки *п* дорівнює 2n, а кількість тих з них, які віддалені від заданої на відстань *d*, — кількості сполучень з *n* no *d:*



Щоб визначити кодову комбінацію, яка віддалена від заданої на відстань *d,* до цієї комбінації можна додати будь-яку Комбінацію вагою *w=d* (з *d* одиницями *i* n-d нулями). Додавання — порозрядне за модулем 2.

Для виявлення всіх помилок кратністю VB кодова відстань миє становити *d >=* VB + 1, а виправлення помилок кратністю Vвп – *d >=* 2VBП + 1. Щоб виправити та виявити всі помилки, має виконуватися умова

d>= VB +VBП + 1

Через те, що загалом кожний елемент (розряд) комбінації недвійкового (багатопозиційного) коду може мати на відміну від двійкового й понад однієї позиції *(т >=*1) з алфавіту *q,* кодова відстань при цьому визначається виразом



де *т* — кількість позицій в кожному розряді (поодиноковому часовому інтервалі, що відповідає тривалості одного елемента) кодової комбінації.

У метриці Хеммінга кодова відстань, як і для двійкового коду, визначається кількістю однойменних розрядів з різними позиціями (символами):



У метриці Лі



де



У модульній метриці , тобто слід виконувати віднімання за модулем *q.*

Відзначимо, що коли значення кодової відстані для двійкового коду в різних матрицях збігаються, для недвійкового коду при *q =3* значення *d* в метриках Хеммінга та Лі також збігаються. При *q >* 3 значення *d* у різних метриках різняться.

При конкретній реалізації недвійкового коду з використанням позицій тих або інших ознак сигналу кодова відстань визначається відповідною метрикою.

СПОСОБИ ПОДАННЯ КОДІВ

Код кожного виду має свій найраціональніший спосіб подання, що випливає з його властивостей. Проте відомо також кілька загальних способів подання кодів, які є досить універсальними і можуть застосовуватися для опису широких класів кодів. До цих способів належать подання кодів у вигляді:

1) таблиць кодових комбінацій;

2) кодового дерева;

3) геометричної моделі;

4) матриці.

*Перший спосіб* полягає в поданні коду у вигляді таблиці всіх його комбінацій. Наприклад, п'ятиелементний двійковий блоковий код зі сталою вагою, в кожній комбінації якого містяться три одиниці, задається так:

**Номер кодової Комбінація двійкового**

**комбінації блокового коду з вагою З**

1 00111

2 01011

3 01101

4 01110

5 10011

6 10101

7 10110

8 11001

9 11010

10 11100

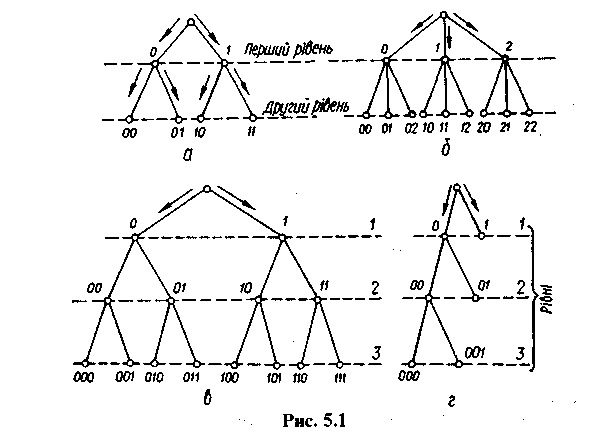
Цей спосіб застосовується для подання будь-яких блокових кодів, але не може бути використаний для неперервних кодів.

*Другий спосіб* подання кодів полягає в зображенні комбінацій коду у вигляді кодового дерева, коли комбінації розміщуються в його вузлах. Під ***кодовим деревом***розумітимемо графічний образ, який складається з точок і ліній, що розходяться від них і також закінчуються точками. Останні називатимемо ***вузлами****,* а лінії, які їх з'єднують, — ***ребрами****.* Перший вузол, від якого починається розходження ребер, називається ***коренем дерева****,* а кількість ребер, які треба пройти від кореня до будь-якого вузла – ***рівнем****,* або ***порядком****,* цього вузла.

Максимальна кількість вузлів, які зустрічаються під час руху вздовж кодового дерева в напрямку від кореня до вершини, визначає ***висоту h***кодового дерева. Вона дорівнює максимальній довжині комбінації коду, побудованому за допомогою цього дерева.

Вузли кодового дерева розташовуються на різних рівнях. Кожний рівень дерева рівномірного коду може мати *qi* вузлів, де *q* — основа коду, *і* — номер рівня *(і =* 1,2,...,n, тут *п* — довжина коду). Для рівномірного двійкового простого коду кількість вузлів на останньому рівні *п* дорівнює кількості N комбінацій коду, тобто *2n = N.*

Вузли, що не з'єднуються з наступними рівнями, називаються ***кінцевими****;* вони відповідають комбінаціям коду.

Основою коду обмежується максимальна кількість ребер, яка може виходити з кожного вузла дерева, а максимальною довжиною кодової комбінації — максимальна кількість рівнів кодового дерева. Кожному вузлу приписується значення розрядів комбінації, що відповідає напрямкам руху вздовж ребер від кореня дерева до вузла. Ребра, що йдуть від кореня до вузлів першого рівня, визначають значення першого зліва розряду кодової комбінації, а ті, що з'єднують вузли першого та другого рівнів — значення другого зліва розряду і т. д. На рис. 5.1 покачано приклади кодових дерев: рівномірних двоелементного двійкового (рис. 5.1, *а),* двоелементного трійкового (рис. 5.1,5), триелементного двійкового (рис. 5.1, в) та нерівномірного двійкового (рис. 5. 1, *г).* 

Цей спосіб застосовується для зображення як блокових, так і неперервних кодів.

За допомогою кодових дерев легко зобразити ***префіксні***коди, що мають властивість префікса й можуть бути утворені послідовним викреслюванням останнього розряду кодової комбінації, причому жодна з комбінацій даного префіксного коду не може бути префіксом його комбінації. Наприклад, префіксами кодової комбінації 10111001 будуть 1, 10, 101, 1011, 10111, 101110, 1011100, 10111001, тобто для однозначного її декодування жодна з комбінацій цього коду не повинна мати перелічені вище комбінації.

Та частина, яка доповнює префіксний код до повної кодової комбінації, утворює суфікс, тобто кожна кодова комбінація складається з префікса та суфікса.

Префіксні коди можна утворити за допомогою кодового дерева, в якого немає вершини і кожний його кінцевий вузол відповідає комбінації префіксного коду.

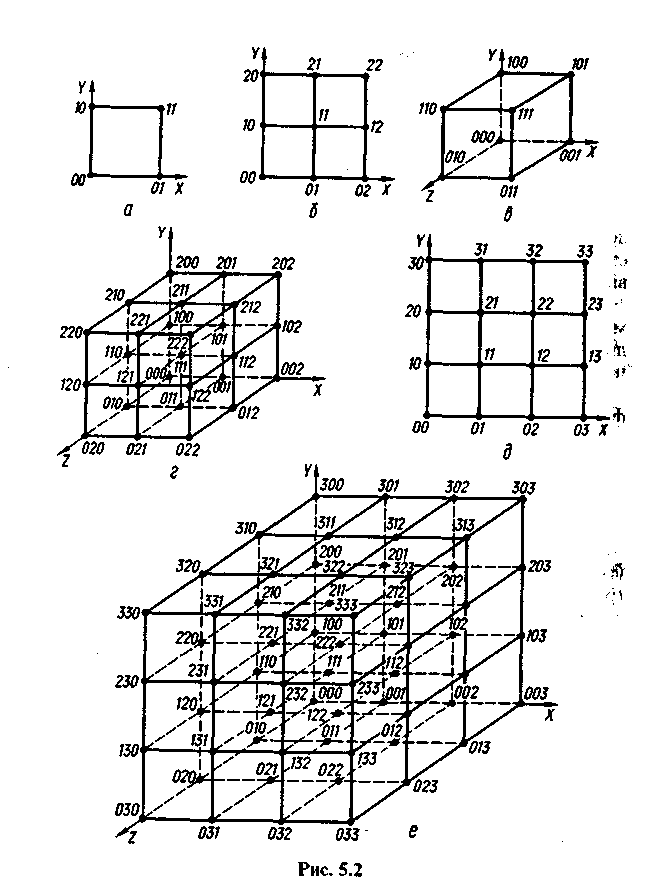
*Третій спосіб* подання кодів полягає в зображенні комбінацій коду точками дискретного и-вимірного векторного простору. Так, кожну комбінацію рівномірного блокового коду (з основою q і довжиною n) *V=(Vn,Vn-1,...,V2,* V1) можна розглядати як вектор або точку деякого n-вимірного векторного простору з координатами *Vn, Vn-1,* ..., *V2, V1.* Якщо значення *q* скінченне, а будь-яка координата вектора є цілим додатним числом від 0 до *q-*1, то зазначений код можна розглядати як дискретний n-вимірний простір, що складається з N=qn точок, які відповідають кінцям усіх можливих векторів.

Цей n-вимірний простір дістав назву *кодового.* Кількість просторових вимірювань кодового простору для коду з будь-якою основою дорівнює довжині n коду, а кількість градацій по кожній з осей (напрямків вимірювання) визначається основою коду і становить *q*-1.

Якщо для дискретного n-вимірного простору, що тут розглядається, ввести поняття кодової відстані *d* між точками *Vi* та *Vj,* то матимемо

** (5.2)

Цілком природно, що для простору з відстанню (5.2), як і для будь-якого іншого кодового простору, *.*



Одним з основних параметрів коду з довільною основою *q,* що визначають його завадостійкість, є *мінімальна кодова відстань dmin.* На відміну від кодової відстані *d,* що визначає кількість станів, які мають пройти якісні ознаки кодової комбінації, щоб опинитися в стані, який відповідає порівнюваній кодовій комбінації, мінімальна кодова відстань характеризує не дві окремо взяті комбінації, а код у цілому, і визначається мінімальною кількістю якісних ознак, за якими відрізняються одна від одної будь-яка пара комбінацій цього коду.

Для визначення кодової відстані між комбінаціями коду з основою *q* треба виконати їх порозрядне віднімання за модулем q. Кодова відстань дорівнює вазі комбінації, що складається з різниці значень комбінацій, між якими визначається ця відстань.

З'єднавши кожну точку простору, що розглядається, прямими лініями з усіма точками, віддаленими на відстань *,* дістанемо геометричну фігуру сіткової структури. Цю фігуру називають *геометричною моделлю* n-елементного *q-*коду*.*

Точки дискретного простору, які містить ця геометрична фігура, називаються її *вершинами,* а лінії, що їх з'єднують, — ребрами*.*

На рис. 5.2 зображено геометричні моделі деяких кодів. Моделлю будь-якого двозначного набору якісних ознак (двоелементного коду) є фігура двовимірного простору — квадрат (рис. 5.2, *а)* або фігура, що складається з квадратів (рис. 5.2, *б,* д): моделлю будь-якого тризначного набору якісних ознак (триелементного коду) — фігура тривимірного простору — куб (рис. 5.2, в) або фігура, що складається з кубів (рис. 5.2, *г, е).*

Побудова моделі чотиризначного набору якісних ознак (чотириелементного коду), тобто фігури чотиривимірного простору, можлива, якщо тривимірний куб або кожну з його вершин змістити в новому напрямку. Взагалі в цьому разі n-вимірyий куб повинен мати 2n вершин, *п\*2n* ребер, n(n-1)\*2n-3 граней, а найвіддаленіша від певної його вершини точка має знаходитися на відстані *п* ребер.

Розглянемо більш докладно властивості, що випливають з геометричної фігури, яка є моделлю n-елементного двійкового коду й дістала назву *п-вимірного куба.* Відстань між будь-якими його вершинами, тобто між двома кодовими комбінаціями, pгідно з (5.2) можна визначити як кількість розрядів, якими вони ріpняться. Так, відстань між комбінаціями 010 і 100 дорівнює двом, оскільки вони різняться елементами в двох розрядах — першому та другому.

Відмінність елементів однойменних розрядів комбінацій двійкового коду легко визначити, застосувавши до них операцію додавання за модулем 2. Як відомо з п. 5.2, результат такого додавання тільки тоді дорівнює 1, коли числа, що додаються, різняться. З урахуванням цього відстань між будь-якими двома комбінаціями n-елементного двійкового коду (вершинами n-вимірного куба) визначається виразом

** (5.3)

Відстань між комбінаціями *Vi* та *Vj* можна знайти також через кількість одиниць у деякій комбінації *Vl* [або через її вагу *w*(Vl)*],* здобуту після додавання за модулем 2 комбінацій *Vi* та V 

(5.4)

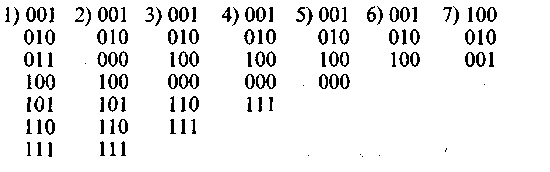
Зручність геометричної моделі зображення будь-якого коду полягає в тому, що кожна її вершина відповідає одній комбінації коду, а відстань між комбінаціями *Vi* та *V,* згідно з (5.2) дорівнює кількості ребер, які треба пройти найкоротшим шляхом з першини *Vi* до вершини *V:.*

Недоліком геометричної моделі є те, що при довжині коду *п* > З зобразити її у звичайному тривимірному просторі неможливо. Тому вона застосовується лише для рівномірних блокових кодів з метою наочного зображення та полегшення аналізу їхніх властивостей.

*Четвертий спосіб* подання кодів у вигляді матриці з 2n рядками та n стовпцями можливий тільки для рівномірних n-елементних двійкових блокових кодів. Якщо матрицею подається сукупність ненульових комбінацій коду, то кількість рядків дорівнюватиме *2n* - 1.

З урахуванням того, що матриця n-елементного коду складається з *2n* - 1 комбінацій, записаних у вигляді рядків, особливість такого запису полягає в тому, що додавання за модулем 2 будь-якої кількості рядків цієї матриці приводить до появи дозволеної комбінації коду, в тому числі й нульової. Якщо останню відкинути, то дістанемо нову матрицю коду, але вже з меншою кількістю рядків. Повторивши аналогічну операцію додавання рядків матриці за модулем 2, можна знову дістати нульову комбінацію коду. Ця операція повторюється доти, поки не буде здобута матриця з лінійно незалежними рядками, додавання яких за модулем 2 вже не приведе до утворення нульової комбінації коду.

Наприклад, щоб побудувати матрицю триелементного двійкового простого коду, треба, записавши у вигляді матриці всі 2n - 1 комбінацій простого коду, крім нульової, послідовно додати їх за модулем 2, виключаючи кожного разу ті комбінації, які в сумі з попередніми утворюють нульову комбінацію:



Друга колонка тут формується так: якщо до третього рядка першої колонки додати суму її першого та другого рядків, то утвориться нульова комбінація, яку виключаємо при запису третьої колонки. Якщо до п'ятого рядка другої колонки після цього додати суму її першого та четвертого рядків, то дістанемо нульову комбінацію в третій колонці. Цю комбінацію також виключаємо, записуючи четверту колонку, і т. д. Таким чином, відкинувши всі нульові комбінації, матимемо шосту колонку з кодовими комбінаціями, що містять тільки по одній одиниці. Це й буде матриця даного коду (сьома колонка).

Квадратна матриця, діагональ якої складається з одиниць, а решта її елементів — нулі, називається *одиничною.* Якщо рядки такої n-елементної матриці додавати за модулем *2,* то підбором відповідної комбінації їх можна дістати всі комбінації n-елементного коду. Тому такі матриці ще називаються *визначальними.*

Якщо напрямок головної діагоналі матриці проходить справа наліво, то матриця називається *транспонованою.* Для розглянутого коду це буде матриця (шоста колонка)



Загалом визначальна матриця n-елементного коду записується так:



Розглянутий приклад утворення визначальної матриці й здобута одинична матриця можуть бути використані тільки для побудови всіх комбінацій двійкового простого коду з мінімальною кодовою відстанню *dmin* = 1.

Матричний спосіб може бути застосований також для побудови коректувальних кодів з *dmin >* 1, здатних виявляти та виправляти помилки. Проте при цьому твірна (породжувальна) матриця складається з двох підматриць — уже відомої одиничної (інформаційної) та додаткової (перевірної).

За допомогою інформаційної одиничної підматриці *Ек* утворюють інформаційну частину кодової комбінації коректувального коду, яка складається з *к* інформаційних елементів і визначне розмір підматриці *(к* х *к),* що відповідає розмірам визначальної квадратної матриці двійкового простого коду, оскільки кількість *N* комбінацій коректувального коду дорівнює кількості *N* комбінацій початкового двійкового простого коду, які треба закодувати цим коректувальним кодом.

Зa допомогою додаткової перевірної підматриці *Сr,k*, утворюють перевірну частину комбінації коректувального коду, що складається з *r* перевірних елементів. Тому додаткова перевірна підматриця має розмір *r* х *к.*

Таким чином, загальний розмір твірної матриці коректувальних кодів дорівнює n х *к,* оскільки *п = к + r.*