15. Принципи завадостійкого кодування. Основні характеристики завадостійких кодів. Класифікація завадостійких кодів. Математичний опис процесу кодування і декодування

Принципи завадостійкого кодування. Основні характеристики завадостійких кодів

Розглянемо загальну постановку задачі завадостійкого кодування повідомлень.

Від джерела інформації надходить послідовність елементарних повідомлень x_i. Кодування полягає в тому, що послідовність символів джерела заміняється послідовністю m-ічних кодових символів (надалі будемо розглядати тільки двійкові коди, для яких m=2). Зазначене перетворення є взаємнооднозначним, що й дозволяє здійснити в приймачі декодування, тобто відновити повідомлення по прийнятій кодовій комбінації.

У результаті дії перешкод деякі кодові символи в приймачі виділяються перекрученими - одиничні елементи комбінації сприймаються як нульові й навпаки.

Може бути передбачений і режим стирання символів, коли рішення про прийом деяких з них не може бути виконане з необхідною надійністю (упевненістю).

Завадостійке кодування повідомлень має своєю метою перетворення повідомлень x_i, i=1,…,m у послідовність кодових символів x_j, j=1, …, n, що обладає властивістю або виявлення присутності в ній помилок, або виявлення й виправлення цих помилок.

Відразу відзначимо, що не існує кодів кінцевої довжини, що дозволяють виявити й, тим більше, виправити всі можливі помилки. Можна побудувати лише коди, що виявляють або виправляють деяке число помилок певного виду (звичайно найбільш ймовірних або найнебезпечніших).

Визначена мета може бути досягнута тільки введенням надмірності, тобто збільшення числа кодових символів у переданій послідовності стосовно символів, що забезпечують однозначне подання переданих повідомлень (інформаційним символам). Наявність додаткових (надлишкових) символів дозволяє накласти на передані послідовності символів (кодові комбінації) додаткові зв'язки (умови) між ними, перевірка яких на прийомній стороні дає можливість виявити й виправити помилки. Вся сукупність можливих кодових комбінацій у цьому випадку називається завадостійким (коригувальним або надлишковим) кодом.

При використанні завадостійкого коду передаються в канал не всі кодові комбінації, що можна сформувати з наявного числа розрядів, а лише що мають певну властивість і що **називаються дозволеними.** Інші комбінації, що не використовуються, називаються **забороненими**. Введення додаткових розпізнавальних ознак в комбінації, що передаються дозволяє істотно підвищити правильність класифікації.

Завадостійкі коди поділяються на:

* коди, що **виявляють** помилки;

* коди, що **виправляють** помилки.

При використанні кодів, що виявляють помилки вся численність n-розрядних комбінацій розбивається на дві підмножини, що не перехрещуються. Одна підчисленність називається **дозволеною**, а інша - **забороненою.** Передаються тільки дозволені кодові комбінації, що мають певною властивістю. Якщо прийнята кодова комбінація відноситься до дозволених, то вважається, що помилки немає. На прийомній стороні відомо, які з комбінацій є дозволеними, а які забороненими. Тому, якщо передана комбінація в результаті помилки перетвориться в деяку заборонену комбінацію, то така помилка буде виявлена, а за певних умов і виправлена. При побудові кодів, що виправляють помилки вся численність кодових комбінацій розбивається на ряд підмножин, що не перехрещуються. В кожній з підмножин одна дозволена комбінація. При прийомі будь-якої комбінації з даної підмножин споживачу видається дозволена комбінація цієї підмножин.

Можливості по виявленню або виправленню помилок визначаються числом позицій, на яких відрізняються дозволені кодові комбінації, т. т. **кодовою відстанню.**

Кодова відстань між i-ю і j-ю кодовими комбінаціями (**відстань по Хеммингу**) визначається по формулі: $d_{ij} = \sum_{k=1}^{n}(x_{ik} \oplus x_{jk})$, де x_{ik} і x_{jk} - значення символів k - й позиції i - й і j - й кодових комбінацій. Число відмінних від 0 символів у кодовому слові x_i називається його **вагою** w_i**.**

Наприклад, між кодовими комбінаціями 1100101 і 1011100, що відрізняються символами в чотирьох розрядах l=(0111001), відстань дорівнює чотирьом.

Для будь-якого коду $1 \le d \le n$. Мінімальна відстань між дозволеними комбінаціями в даному коді називається **мінімальною кодовою відстанню** й позначається d_{min}.

Таким чином, якщо код має кодову відстань d_{min}, то це позначає, що дозволені комбінації цього коду відрізняються друг від друга не менш чим в d_{min} символах і, отже, тільки поява d_{min} і більше помилок у прийнятій комбінації можуть перевести її в іншу дозволену комбінацію.

З іншого боку, якщо число помилок у комбінації буде менше ніж d_{min}, то стає зрозумілим, що така комбінація помилок буде неодмінно приводити до забороненої комбінації.

В загальному випадку необхідна кодова відстань для забезпечення **виявлення всіх помилок кратності до** t_0 **включно** визначається вираженням $t_0 = d + 1$.

При **виправленні помилок кратності до** t_u **включно** кодова відстань повинна бути рівна $d = 2t_u + 1$. З цих виразів видно, що $t_0 = 2t_u$.

Необхідна кодова відстань при **виправленні помилок кратності до** t_u **включно і виявлення помилок кратності від** $t_u + 1$ **до** t_0 **повинна** бути рівна

$d \geq t_u + t_0 + 1$. Необхідна кодова відстань, а отже, і завадостійкість коду визначається **надмірністю коду**, т. є. числом введених перевірочних символів.

Визначимо кількість перевірочних розрядів, необхідну для виправлення t_u помилок. Для цього необхідно, щоб за допомогою перевірочних розрядів можна було описати наступні ситуації:

- помилка буде відсутня - 1-й випадок;
- одинока помилка - C_n^1 випадків;
- двократна помилка - C_n^2 випадків;

....

- помилка кратності t_u - $C_n^{t_u}$ випадків,

де C_n^i - число сполучень (комбінацій) з n по i. Таким чином, кількість перевірочних розрядів для виправлення помилок кратності t_u і менш визначається з наступної нерівності:

$$2^k \geq \sum_{i=0}^{t_u} C_n^i, \Rightarrow k \geq \log_2 \sum_{i=0}^{t_u} C_n^i.$$

Даним вираженням можна скористуватися і для знаходження числа перевірочних розрядів для виявлення помилок. Для цього необхідно використати той факт, що число помилок, що виявляються в два рази більше числа помилок, що виправляються.

Отже, для виявлення t_0 помилок кількість перевірочних одиничних елементів повинно задовольняти нерівності:

$$k \geq \log_2 \sum_{i=0}^{t_o/2} C_n^i.$$

Важливими показниками ефективності коду є **коефіцієнти помилок**, що виявляються $К_{ВП}$ та не виявляються $К_{НВ}$, під якими розуміють відношення числа кодових комбінацій з виявленими (невиявленими) помилками до числа всіх можливих комбінацій. Помилки не виявляються, коли кодова комбінація, що передається під впливом завад перетворюється в іншу дозволену. Число всіх дозволених комбінацій при m інформаційних розрядах рівно 2^m. Отже, при передачі будь якої кодової комбінації можливе число випадків невиявлених помилок буде рівно $N_{НП} = 2^m-1$. Оскільки число всіх можливих помилкових комбінацій рівно 2^n, де n - кількість розрядів у кодовій комбінації, то вираження для визначення величини $К_{НП}$ буде мати вигляд:

$$К_{НП} = \frac{2^m - 1}{2^n} = \frac{1}{2^{n-m}} - \frac{1}{2^n} = \frac{1}{2^k} - \frac{1}{2^n},$$

де $k = n - m$ - число надлишкових розрядів. Так як число всіх заборонених кодових комбінацій рівно $2^n - 2^m$, то коефіцієнт може бути визначений з вираження:

$$K_{вп} = \frac{2^n - 2^m}{2^n} = 1 - \frac{1}{2^k}.$$

Класифікація завадостійких кодів

Коди поділяються (рис.15.1) на:

* блочні (блокові);
* безперервні.

До **блочних** відносяться коди, у яких кожному повідомленню ставиться в однозначну відповідність блок з n символів. **Безперервні** коди представляють безперервну послідовність інформаційних і перевірочних розрядів.

Блочні коди діляться на: рівномірні;
нерівномірні.

Рівномірні коди діляться на: систематичні;
несистематичні.

Під **систематичними** розуміють код, в якому розряди можуть бути поділені на перевірочні і інформаційні. При цьому їхні місця в кодовій комбінації цілком визначені. Несистематичні коди цю властивість не мають.

Окрім цього, коди діляться на:

* лінійні;
* нелінійні.

Лінійні коди це такі, у яких сума по модулю 2 дозволених комбінацій дасть дозволену комбінацію того же коду. Нелінійні коди означену властивість не мають. Передані повідомлення визначаються m інформаційними символами. Надлишкові (контрольні, перевірочні) k символів доповнюють кодову комбінацію до n = m + k символів, у силу чого такі блокові коди позначають як (n,m) коди

Більшість кодів, що застосовуються на практиці, відноситься до лінійних. Найбільше поширення серед блокових кодів знайшли коди з перевіркою на парність, з повторенням символів, рівноважні коди, коди Хемминга, коди Рида-Малера й, найбільш великий клас кодів - циклічні коди. До останнього класу належать коди Боуза-Чоудхури-Хоквингема (БЧХ) і коди Рида-Соломона (РС), що одержали найбільше поширення в техніці військового зв'язку.

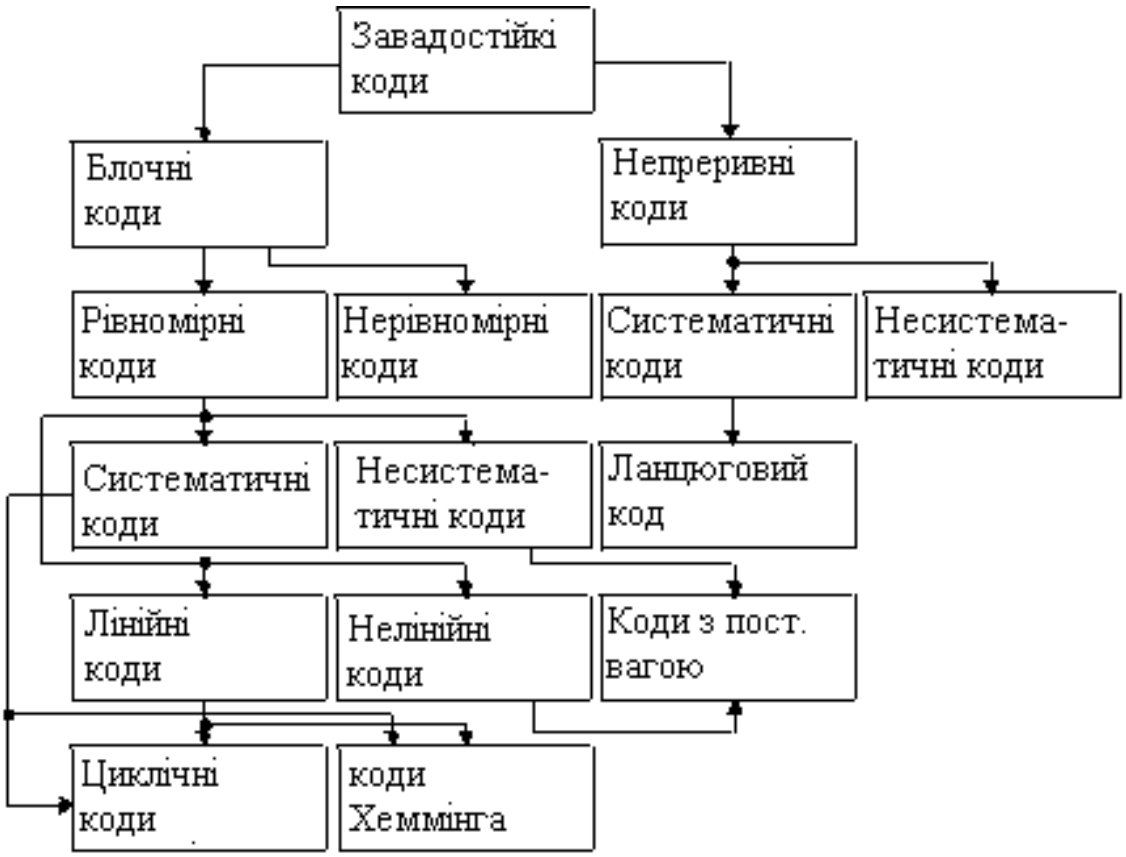

Рис.15.1

Математичний опис процесу кодування і декодування

При завданні коду звичайно вказують, які інформаційні елементи беруть участь в формуванні кожного з k перевірочних розрядів. Наприклад, для коду з n = 5, m =3, k = 2 кожний перевірочний розряд $П_i$ визначається складанням по модулю 2 по правилу:

$$П_1 = I_1 \oplus I_2;\ П_2 = I_2 \oplus I_3,$$

де I_i - інформаційні розряди.

Комбінація такого коду записується в вигляді $П_2, П_1, I_3, I_2, I_1$. При завданні коду можна вказати всі дозволені для цього коду комбінації. Для лінійних кодів засіб завдання можна значно спростити. Для m інформаційних розрядів число всіх дозволених кодових комбінацій буде рівно 2^m. Виберемо з усіх кодових комбінацій тільки лінійно незалежні. Під **лінійно незалежними кодовими комбінаціями** розуміють такі, сума по модулю 2 яких (в будь-якому поєднанні) не рівна нулю. Для наведеного прикладу такими комбінаціями можуть бути комбінації вигляду:

1) 01001	2) 01001	3) 01001
11010	10011	10011
10100	01110	00111

Всі інші кодові комбінації можна отримати складанням по модулю 2 лінійних незалежних комбінацій. Звичайно лінійно незалежні кодові комбінації записують в вигляді матриці розміром $n \times m$, яку називають **породжуючою матрицею** і позначають $G(n \times m)$. Найбільш часто породжуючі матриці записують так званій **канонічній формі**. При цьому перші або останні m

стовпчиків утворять одиничну матрицю. Інші стовпчики вказують правила формування перевірочних розрядів. Так, для наведеного прикладу породжуюча матриця в канонічній формі має вигляд:

$$G(5,3) = \begin{Vmatrix} 10100 \\ 11010 \\ 01001 \end{Vmatrix}.$$

З прикладу видно, що останні три стовпчика складають одиничну матрицю; другий стовпчик вказує, що при формуванні першого перевірочного розряду беруть участь перший і другій інформаційні розряди. Перший стовпчик вказує, що при формуванні другого перевірочного розряду беруть участь другій і третій інформаційні розряди. Позначають породжуючу матрицю, що написана в канонічній формі, наступним чином:

$$G(n,m) = \left\| R_{k\times m} \cdot I_m \right\| \quad \text{або} \quad G(n,m) = \left\| I_m \cdot R_{k\times m} \right\|.$$

Процес кодування математично записується в вигляді добутку матриці - рядка, що відображає кодову комбінацію, яка передається, на породжуючу матрицю. Припустимо, що кодова комбінація, яка передається, записана в вигляді вектору:

$$I = \left\| I_m I_{m-1} \ldots I_2 I_1 \right\|.$$

Тоді процес кодування запишеться в вигляді:

$$F = IG(n,m) = \left\| \underbrace{\Pi_k \Pi_{k-1} \ldots \Pi_1}_{\substack{\text{перевір.} \\ \text{розряди}}} \; \underbrace{I_m I_{m-1} I_{m-2} \ldots I_2 I_1}_{\substack{\text{інформ.} \\ \text{розряди}}} \right\|.$$

В результаті кодування отримаємо комбінацію, де перші m розрядів - інформаційні, а останні k розрядів - перевірочні Π_j.

Процес декодування математично описують добутком перевірочної матриці $H(n,m)$ і вектору-стовпчика, що відображає прийняту кодову комбінацію $F = F^T + E^T$, де E^T - вектор помилки

$$C = H(n,m)(F^T + E^T), \qquad (15.1)$$

де C - результат декодування (синдром);

$(\ )^T$ - знак транспонування.

Оскільки зв'язок між породжуючою і перевірочною матрицями визначається рівністю:

$$G(n,m)\, H^T(n,m) = H(n,m)\, G^T(n,m) = 0,$$

то рівність (12.1) буде мати вигляд:

$$C = H(n,m)\, E^T = E\, H^T(n,m).$$

Якщо позначити h_i – і - й стовпчик перевірочної матриці, то: $H(n,m) = \left\| h_1 h_2 \ldots h_n \right\|$.

Вектор помилки ($E = \|0010\ldots010\ldots0\|$) містить **1** на тих позиціях, символи на яких викривлені. Нехай ці позиції мають номери $n_1, n_2, \ldots n_\ell$. Тоді буде справедливо рівність:

$$C = H(n,m)\, E^T = \sum_{i=1}^{\ell} h_{\nu_i}.$$

Отже, якщо необхідно виявити ℓ помилок, то повинно бути виконана умова:

$$\sum_{i=1}^{\ell} h_{\nu_i} \neq 0$$

при будь-якому поєднанні ℓ викривлених символів. Якщо необхідно виправити ℓ помилок, то сума: $\sum_{i=1}^{\ell} h_{\nu_i}$ для будь-яких ℓ конкретних стовпчиків перевірочної матриці повинна бути цілком певною, і не співпадати з аналогічною сумою для інших ℓ стовпчиків. Так, наприклад, при виправленні одинокої помилки всі стовпчики перевірочної матриці повинні бути різноманітні, т. т. $h_i \neq h_j$ при будь-яких i і j, $i \neq j$. Якщо необхідно виправити одиноку помилку і виявити пакет, який складається з трьох і двох символів, то необхідно виконати наступні умови:

- $h_i \neq h_j$ при $i \neq j$ (для виправлення одинокої помилки);
- $h_i + h_{i+1} \neq 0; h_i + h_{i+1} \neq h_j$ при будь-яких i і j (для виявлення пакету з двох помилок);
- $h_i + h_{i+1} + h_{i+2} \neq 0; h_i + h_{i+1} + h_{i+2} \neq h_j$ при будь-яких i і j (для виявлення пакету з трьох помилок).

Матриці $H(n,m)$ і $G(n,m)$ можна змінити ролями. Тоді матриця $H(n,m)$ буде породжуючою, а $G(n,m)$ - перевірочною. Коди, взаємозв'язані між собою таким чином, називають **дуальними** (подвійними).