Chương 5. Mã tuyến tính nhị phân

Phép toán nhị phân

- ĐN phép toán cộng (+) và nhân (.) trên bảng ký tự nhị phân 0, 1 như sau:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

- 1 = - 1 → ‘cộng’ giống ‘trừ’

Mã tuyến tính nhị phân

ĐN: Một mã K là ***mã tuyến tính (linear code)*** nếu:

- Tổng a + b của hai codeword bất kỳ cũng là một codeword.
- Tích k.a (với k = const và codeword a) cũng là một codeword.

ĐN: Mã nhị phân K là ***mã nhị phân tuyến tính (binary linear code)*** nếu tổng a + b của hai codeword bất kỳ cũng là một codeword.

Ví dụ mã nhị phân tuyến tính

1. Mã lặp K_N = {‘00…0’, ‘11…1’}.
2. Mã kiểm chẵn lẻ (tổng số bit 1 là chẵn)

Information Bits	Code Word	Information Bits	Code Word
000	0000	110	1100
100	1001	101	1010
010	0101	011	0110
001	0011	111	1111

Figure 2: The even-parity code of length 4

Hamming weight

<u>Định nghĩa</u>:

- ***Trọng số Hamming (Hamming weight)*** của một codeword **a**, ký hiệu **w(a)**, là số lượng các bit khác 0 của **a**.
- Với mỗi mã **K**, trọng số Hamming nhỏ nhất của các codeword khác **0** = '00…0' được gọi là ***trọng số nhỏ nhất (minimum weight)*** của **K**, ký hiệu **w(K)**.

<u>Ví dụ</u>: Từ mã a = '11000' có w(a) = 2; a là một codeword của mã kiểm chẵn lẻ độ dài 5.

Mọi codeword a của mã kiểm chẵn lẻ K này có w(a) bằng 2 hoặc 4. Nên w(K) = 2.

Ma trận kiểm tra tính chẵn lẻ

ĐN: Một ma trận nhị phân H được gọi là ***ma trận kiểm chẵn lẻ (parity check matrix)*** của mã khối K độ dài n nếu với mọi từ mã x của mã K ta có $Hx^T = 0$.

Ví dụ: Một ma trận kiểm chẵn lẻ của mã kiểm chẵn lẻ là H = [1 1 … 1].

Sửa lỗi

Định lý: *Một mã nhị phân tuyến tính K sửa được ít nhất một lỗi khi và chỉ khi mọi ma trận kiểm chẵn lẻ của K có các cột đôi một khác nhau và khác 0.*

Chứng minh: (xem sách).

→ Mã kiểm chẵn lẻ không sửa được lỗi.

Mã Hamming

ĐN: Một mã nhị phân tuyến tính độ dài $2^m - 1$ được gọi là ***mã Hamming*** nếu nó có một ma trận kiểm chẵn lẻ H kích thước m x $2^m - 1$ thoả mọi từ nhị phân khác 0 độ dài m đều là một cột của H.

Ví dụ

Information Symbol	Code word	Information Symbol	Code word
0000	0000000	0110	0110011
1000	1000011	0101	0101010
0100	0100101	0011	0011001
0010	0010110	1110	1110000
0001	0001111	1101	1101001
1100	1100110	1011	1011010
1010	1010101	0111	0111100
1001	1001100	1111	1111111

Figure 6: Code words of the Hamming code of length 7

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

information rate $R = 4/7$.

Tính chất của một mã Hamming

1. Độ dài các từ mã $n = 2^m - 1$.
2. Số bit mang thông tin: $k = 2^m - m - 1$.
3. → Tỷ lệ thông tin $R = k/n = \ldots$.
4. Khoảng cách nhỏ nhất của mã: $d = 3$.
5. → sửa được ít nhất một lỗi.

Giải mã

- Khi chuỗi nhận được là $w = w_1 w_2 \ldots w_n$,
- Tính $s = Hw^T$.
- Nếu s = 00…0: không có lỗi.
- Nếu $s \neq 00\ldots0$. Vị trí của s trong ma trận H chính là vị trí bị lỗi.
- Ta gọi s là ***syndrome***.

- → '*Giải mã bằng syndrome*'

Ví dụ

- Truyền 1111111, nhưng nhận w = 1110111.

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

- Syndrome là $s = Hw^T$.

$$\begin{aligned} s_1 &= w_4 + w_5 + w_6 + w_7 = 1, \\ s_2 &= w_2 + w_3 + w_6 + w_7 = 0, \\ s_3 &= w_1 + w_3 + w_5 + w_7 = 0. \end{aligned}$$

- s = (100). Vị trí bị lỗi là vị trí số 4.
- Sửa 111**0**111 là 111**1**111.

Không phát hiện được lỗi

- K là mã tuyến tính.
- Giả sử truyền từ mã $v \in K$, nhận w ($w \neq v$) cũng là từ mã $\in K$
- → có lỗi nhưng không phát hiện được.

- Tính xác suất không phát hiện được lỗi?

- Đặt e = w – v (= w + v).

1. e = 0 ⇔ w = v : không có lỗi.
2. e ≠ 0: Nếu e là codeword thì không phát hiện được lỗi.
   - Giả sử w(e) = i (truyền v có i bit bị lỗi)
   - → Xác suất xảy ra = $p^i q^{n-i}$.

- Đặt A_i = #{từ mã x có w(x) = i}.
- Xác suất không phát hiện được lỗi

$$P_{und} = \sum_{i=1}^{n} A_i p^i q^{n-i}.$$

Ví dụ

0000	0000000	0110	0110011
1000	1000011	0101	0101010
0100	0100101	0011	0011001
0010	0010110	1110	1110000
0001	0001111	1101	1101001
1100	1100110	1011	1011010
1010	1010101	0111	0111100
1001	1001100	1111	1111111

$$P_{\text{und}} = \sum_{i=1}^{n} A_i p^i q^{n-i}. \qquad P_{\text{und}} = 7p^3q^4 + 7p^4q^3 + p^7. \qquad p = 0.01$$

$$P_{\text{und}} = 7(0.01)^3(0.99)^4 + 7(0.01)^4(0.99)^3 + (0.01)^7 \approx 7 \times 10^{-6},$$

Cách tính P_{und}.

$$P_{\text{und}} = \sum_{i=1}^{n} A_i p^i q^{n-i}.$$

$$P_{\text{und}} = q^n \sum_{i=1}^{n} A_i p^i q^{-i} = q^n \sum_{i=1}^{n} A_i \left(\frac{p}{q}\right)^i.$$

$$A_0 = 1$$

$$P_{\text{und}} = q^n \left[\sum_{i=0}^{n} A_i \left(\frac{p}{q}\right)^i - 1\right] = q^n \left[A\left(\frac{p}{q}\right) - 1\right].$$

$$A(x) = \sum_{i=0}^{n} A_i x^i$$ ← **weight enumerator *of the code K*.**

Ví dụ

0000	0000000	0110	0110011
1000	1000011	0101	0101010
0100	0100101	0011	0011001
0010	0010110	1110	1110000
0001	0001111	1101	1101001
1100	1100110	1011	1011010
1010	1010101	0111	0111100
1001	1001100	1111	1111111

$$A(x) = 1 + 7x^3 + 7x^4 + x^7.$$

$$\begin{aligned} P_{\text{und}} &= q^7\left[1 + 7\left(\frac{p}{q}\right)^3 + 7\left(\frac{p}{q}\right)^4 + \left(\frac{p}{q}\right)^7 - 1\right] \\ &= 7p^3q^4 + 7p^4q^3 + p^7. \end{aligned}$$

Tóm tắt

- Mã tuyến tính nhị phân
- Hamming weight w(a)
- Parity check matrix
- Mã Hamming
- Xác suất không phát hiện được lỗi P_{und}.

Homework

- Đọc lại và làm bài tập
  - Chương 5 [1]
  - Chương 1 [2]
- Đọc trước chương 6+7 [1]

Bài tập

- 'Palindrome' = chuỗi đối xứng (đọc xuôi đọc ngược như nhau).
- VD: 'was it a rat I saw'
- Xét mã nhị phân đối xứng K. Hỏi K có thể là một mã tuyến tính không? Nếu có, K có thể phát hiện được bao nhiêu lỗi.

Bài tập

- Lập mã K nhị phân độ dài 7 như sau:
  - Bit thứ ba kiểm chẵn lẻ 2 bit đầu
  - Bit thứ sáu kiểm chẵn lẻ bit 4 và bit 5.
  - Bit thứ bảy kiểm chẵn lẻ toàn bộ từ mã.
- Tính số lỗi mà K có thể:
  - Phát hiện được
  - Sửa được

Bài tập

- Tính $P_{err}(K)$ của mã Hamming K (7,4) với p = 0.01.
- Tổng quát, Tính $P_{err}(K)$ của mã Hamming K độ dài $2^m - 1$ theo p và m.
- Tính $P_{und}(K)$ của mã Hamming K độ dài $2^m - 1$