CHAPTER 5 CODING THEORY

Messages are to be sent from one place to another. On the way, these messages may be distorted by random disturbances, or noises. So the message received may be different from what is sent. In coding, we include some additional information which would enable the receiver to detect error and to correct error.

The basic unit of information is word. We confine our study to binary words, which are strings of 0's and 1's. We assume all the words to be transmitted are of the same length, m.
Binary word of length m, $w \in B^m$, where B^m = set of all strings of length m formed using symbols in B = {0,1}.

An (m,n) encoding function, where n > m, is a function e: $B^m \rightarrow B^n$ n > m means additional bits are included for checking purposes.

The weight of a binary word w, written as |w|, is the number of 1's in w.

The even parity check code e: $B^m \rightarrow B^{m+1}$ is defined as follows:

The resulting code word e(w) then has an even number of 1's. |e(w)| is even for all w.

If one error occurs in a certain bit b_i, the word received would have an odd number of 1's. The receiver knows that an error has occurred.

We say that an encoding function e detects k or fewer errors if whenever e(w) is transmitted with k or fewer errors, the word received as not a code word.
The parity check code above has k = 1. If there are 2 or more errors, these may escape detection.

The Hamming distance between 2 binary words of the same length is defined as the number of positions (corresponding bits) where the 2 words are different.

$$\delta(x, y) = |x \oplus y|$$

Example 1

$\delta(x, y) = 3$. In $\oplus$, different bits produce 1 and same bits produce 0.

Theorem

(1) $\delta(x, y) = \delta(y, x)$

(2) $\delta(x, y) \geq 0$

(3) $\delta(x, y) = 0$ if and only if $x = y$

(4) $\delta(x, y) \leq \delta(x, z) + \delta(z, y)$

The minimum distance of an encoding function is the minimum of the distance between every pair of code words.

Example 2

Consider this (2,5) encoding function e:

e(00) = 00000 e(01) = 01110 e(10) = 00111 e(11) = 11111

Tabulate the distances between all pairs of code words.

What is the minimum distance ?

	00000	01110	00111	11111
$c_o = 00000$				
$c_1 = 01110$				
$c_2 = 00111$				
$c_3 = 11111$				

Theorem

An (m,n) encoding function can detect k or fewer errors if its minimum distance is at least k + 1.

In the preceding example, the minimum distance is 2. This e enables us to detect 1 error.

Group Codes

$(B^n, \oplus)$ is a commutative group. An (m,n) encoding function e: $B^m \rightarrow B^n$ is called a group code if the set of all code words form a subgroup of B^n. ie Ran(e) = e(B^m) = {e(w): w $\in$ B^m} is a subgroup of B^n.

Example 3

Consider the code words c_2 = 00111 and c_3 = 11111 in the last example.

Example 4

Consider this encoding function e:

$e(00) = 00000 = c_o$ $\quad e(01) = 01100 = c_1$

$e(10) = 10011 = c_2$ $\quad e(11) = 11111 = c_3$

Let $C = \{c_o, c_1, c_2, c_3\} = Ran(e) = e(B^2)$

Show that C is a subgroup of B^5.

	C_o	C_1	C_2	C_3
C_o				
C_1				
C_2				
C_3				

Theorem

Let e: $B^m \mapsto B^n$ be a group code. Then the minimum distance of e is the minimum weight of all nonzero code words.

$\delta(c_i, c_j) = |c_i \oplus c_j|$
For a group code, $c_i \oplus c_j = c_k$ for some k.
$\min \delta(c_i, c_j) = \min |c_k|$.

Parity check matrix

A group code e: $B^m \mapsto B^n$ can be generated using a parity check matrix of the form $H = \begin{pmatrix} H_{mr} \\ I_r \end{pmatrix}$,

where $r = n - m$, H_{mr} is an $m \times r$ Boolean matrix, and I_r is the $r \times r$ identity matrix.

$$H = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1r} \\ h_{21} & h_{22} & \cdots & h_{2r} \\ \vdots & \vdots & \cdots & \vdots \\ h_{m1} & h_{m2} & \cdots & h_{mr} \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

m rows
r columns

(each h_{ij} = 0 or 1)

The mod 2 Boolean product of matrices is defined by this rule:

$$(b_1, b_2, \ldots b_m) \begin{bmatrix} h_{1j} \\ h_{2j} \\ \ldots \\ h_{mj} \end{bmatrix} = b_1h_{1j} \oplus b_2h_{2j} \oplus \ldots \oplus b_mh_{mj}$$

$$\begin{cases} 0 \text{ if number of 1 is even.} \\ 1 \text{ if number of 1 is odd.} \end{cases}$$

Example 5

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

Obtain an (m,n) encoding function e_H as follows.
For $b = (b_1, b_2, \ldots b_m) \in B^m$,

$$e_H(b) = b_1, b_2, b_3 \ldots, b_m \underbrace{c_1, c_2, c_3 \ldots, c_r}_{\text{check bits}}$$

where

$$c_1, c_2, c_3 \ldots, c_r = (b_1, b_2, b_3 \ldots, b_m) \begin{bmatrix} h_{11} & h_{12} & . & . & h_{1r} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ h_{m1} & h_{m2} & . & . & h_{mr} \end{bmatrix}$$

Then this e_H is a group code.

Example 6

Obtain a (2,5) group code using this parity check matrix:

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

Solution:

{00000, 01010, 10110, 11100} form a subgroup of B^5. This e_H is a group code.

Decoding

Suppose we receive a word $x \in B^n$. We wish to know what is most likely the original word sent. We look for an element $b \in B^m$ such that x is as near as possible to the code word e(b). Then we decode x as b, and write d(x) = b. This d is an (n,m) decoding function, associated with the (m,n) encoding function e.

Suppose the set of code words can be arranged in some order. $C = e(B^m) = \{c_o, c_1, c_2, \ldots, c_k\}$.

When we receive a word x, we compare its distance from every code word. Let c_s be the first code word whose distance from x is minimum: $\delta(c_s, x) \leq \delta(c_i, x)$ for all i.

If $e(b) = c_s$, then $d(x) = d(c_s) = b$.

A decoding function, d, obtained this way is called a maximum likelihood decoding function.

Example 7

Refer to the group code of the last example.

$e(00) = 00000 = c_o$ $e(01) = 01010 = c_1$

$e(10) = 10110 = c_2$ $e(11) = 11100 = c_3$

How would the following words received be decoded according to maximum likelihood technique?

(i) 01010

(ii) 10111

Solution:

(i) 01010 = c_1,

(ii) x = 10111,

Decoding Table

For a group code, we may construct a decoding table by computing all the distinct cosets of the subgroup, C, of code words. For each coset, we choose an element of least weight, ε, called the coset leader. The coset may be written as εC. For a word received, x, we locate it in the decoding table. The column heading indicates which code word is nearest to this x.

c	00000	01010	10110	11100
00001 ⊕ c				
00010 ⊕ c				
00100 ⊕ c				
10000 ⊕ c				
00011 ⊕ c				
00101 ⊕ c				
10001 ⊕ c				

For word x received, such as 10111, locate it in the decoding table.

Column heading for this column is c_2 =

d(10111) =

Example 8

Obtain a (3,5) group code based on the parity check matrix

$$H = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$ Construct a decoding table.

Decode the following words received:

(i) 01111 (ii) 01011 (iii) 01100

Solution:

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

w	e(w)
000	
001	
010	
011	
100	
101	
110	
111	

Decoding table

00000	00110	01001	01111	10011	10101	11010	11100
00001							
00010							
10000							

(i) d(01111) =
(ii) d(01011) =
(iii) d(01100) =

Theorem

For an (m, n) group code generated by the parity check matrix H, two words x, y $\in B^n$ are in the same coset if and only if their mod 2 Boolean products with H are equal: $x \otimes H = y \otimes H$.

Based on this theorem, we may do decoding without having to search the entire decoding table. For the coset leader ε of each coset, we find $\varepsilon \otimes H$, called a syndrome. For a word received, x, find its syndrome $x \otimes H$. Choose the ε with $\varepsilon \otimes H = x \otimes H$. This ε and x are in the same coset. Suppose c is the code word with $x = c \oplus \varepsilon$, and ε has been chosen of minimum weight, then c is the code word nearest to x. As $\varepsilon \oplus \varepsilon = 0$ and $x = c \oplus \varepsilon$, we get $c = x \oplus \varepsilon$. $d(x) = d(x \oplus \varepsilon)$.

Example 9

Using a table of syndromes, decode the 3 words received in the last example.

Solution:

$$(\qquad)\begin{bmatrix}1 & 1\\ 0 & 1\\ 1 & 0\\ 1 & 0\\ 0 & 1\end{bmatrix} =$$

$$(\qquad)\begin{bmatrix}1 & 1\\ 0 & 1\\ 1 & 0\\ 1 & 0\\ 0 & 1\end{bmatrix} =$$

$$(\quad) \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} =$$

$$(\quad) \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} =$$

$$(\quad) \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} =$$

$$(\qquad)\begin{bmatrix}1 & 1\\ 0 & 1\\ 1 & 0\\ 1 & 0\\ 0 & 1\end{bmatrix} =$$

ε	ε ⊗ H
00000	
00001	
00010	
10000	

(i) $(01111)\begin{bmatrix}1 & 1\\ 0 & 1\\ 1 & 0\\ 1 & 0\\ 0 & 1\end{bmatrix} =$ $\varepsilon =$

$d(01111) =$

(ii) $(01011)\begin{bmatrix}1 & 1\\ 0 & 1\\ 1 & 0\\ 1 & 0\\ 0 & 1\end{bmatrix} =$ $\varepsilon =$

$d(01011) =$

(iii)

$$(01100)\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \qquad \varepsilon =$$

$$d(01100) =$$

Extra Example 1:

a) Show that the (2, 5) encoding function $e : B^2 \to B^5$ defined by

$e(00) = 00000 = C_0$
$e(01) = 01010 = C_1$
$e(10) = 10110 = C_2$
$e(11) = 11100 = C_3$
is a group code.

b) Decode the words 11010 and 01100 relative to a maximum likelihood decoding function.

Extra Example 2:

The parity check matrix of a group code is $H = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

a) Determine the (3, 6) group code function $e_H : B^3 \rightarrow B^6$.

b) Find the minimum distance of the encoding function and determine the number of error(s) that can be detected.

c) Decode 101110 and 010111 relative to a maximum likelihood decoding function.