Coding Theory

Error pattern :- Suppose a word $c = c_1 c_2 \cdots c_n \in Z_2^n$ is transmitted from a point A through a transmission channel T.

Due to the noise present in the channel transmission it will alter the word pattern. ('because of disturbances)

Thus $r = r_1 . r_2 \cdots r_n \in Z_2^n$ be the word received at point B. where each c_i & r_i are 0's & 1's

Note :- Here $r_j \neq c_j$ for some j, then we say that error has occured.

If $r_i = c_i \ \forall i$ except k values $(k<n)$, we say that r differs from c in k places.

$\therefore$ The word 'r' is denoted by $T(c)$. Some times it is convenient to write 'r' as $\boxed{r = c + e}$ where 'e' is a word in Z_2^n

$\therefore$ The word 'e' is reffered as the error pattern.

*Probability:-

If 'p' is the probability of an individual signal in a word is received incorrectly.

Thus probability that $r_i \neq c_i$ is $\boxed{P}$

& probability that $r_i = c_i$ is $\boxed{(1-P)^{n-1}}$

$\therefore$ This is the probability that r differs from c exactly one places.

Illy The probability that r - differs from k - places is $\boxed{p^k \cdot (1-p)^{n-k}}$

Problem:-

1) The word $c = 1010110$ is transmitted through a binary ~~symmetric~~ systematic channel. If $e = 0101101$ is the error pattern. Find the word r - received. If $P = 0.05$ is the probability that a signal is incorrectly received, find the probability with which 'r' is received.

Soln :- Given, $c = 1010110$

$e = 0101101$ | $n = 7$

$r = ?$ | $k = ?$

$P = 0.05$

$\therefore$ The received word $r = c + e$

$\therefore r = c + e = (1,0,1,0,1,1,0) + (0,1,0,1,1,0,1)$

$= (1,1,1,1,0,1,1)$

$\therefore \underline{r = 1111011}$

We observe that 'r'-differs from 'c' in the second, fourth, fifth and seventh place.

$\therefore$ Totally $k = 4$ place.

Thus the probability with which 'r' is received is

$$p^k (1-p)^{n-k} = p^4 (1-p)^{7-4}$$

$$= (0.05)^4 (1-0.05)^3$$

$$= (0.05)^4 (0.95)^3$$

$$= \underline{0.000005}$$

2) The word $c = 1010110$ is sent through a binary symmetric channel. If $p = 0.02$ is the probability of incorrect receipt of a signal, find the probability that c is received as $r = 1011111$. Determine the error pattern.

Soln:- Given, $c = 1010110$

$r = 1011111$

$p = 0.02$

$e = ?$

$n = 7$

$k = ?$

By comparing c & r, it differs in two places

$\therefore \quad k = 2$

$$\therefore p^k (1-p)^{n-k} = p^2 (1-p)^{7-2}$$

$$= (0.02)^2 (1-0.02)^5$$

$$= (0.02)^2 (0.98)^5$$

$$= 0.00036$$

$\therefore$ The error pattern 'e' is given by $r = c + e$

i.e, $1011111 = 1010110 + e_1 e_2 e_3 e_4 e_5 e_6 e_7$

$\therefore \quad e = r - c$

$\Rightarrow 1 = 1 + e_1 \Rightarrow e_1 = 0$

$0 = 0 + e_2 \Rightarrow e_2 = 0$

$1 = 1 + e_3 \Rightarrow e_3 = 0$

$1 = 0 + e_4 \Rightarrow e_4 = 1$

$1 = \cancel{0}1 + e_5 \Rightarrow e_5 = 0$

$1 = 1 + e_6 \Rightarrow e_6 = 0$

$1 = 0 + e_7 \Rightarrow e_7 = 1$

$\therefore e = 0001001$

R.W. Hamming

Hamming distance :

Let $x = x_1 x_2 \dots x_n$ & $y = y_1 y_2 \dots y_n \in Z_2^n$

Then the no of 1's such that $x_i = y_i$, $1 \leq i \leq m$ is the Hamming distance b/w x & y.

It is denoted by $d(x, y)$

In other words $d(x, y)$ is the no of positions in which x & y differ.

[Hamming distance b/w two words of the same size is the number of difference b/w the corresponding bits.

x & y $\longrightarrow$ hamming distance $d(x, y)$]

→ called Hamming matrix.

Eg:- $x = 01001$

$y = 11010$

$\therefore$ x & y differ in 3-places.

Hence $d(x,y) = 3$.

Alternatively, $x+y = 10011$

so that $wt(x+y) = 3$

$\therefore d(x,y) = 3$.

A decoding Scheme:-

The parity check matrix H provides a decoding scheme that corrects single errors in transmission if the following condition hold.

i) H does not contain a column of 0's.

ii) No two columns of H are identical.

When H satisfies these two conditions, the following algorithm is used to decode the received word.

$T(c) = r = r_1, r_2 \ldots r_n \in Z_2^n$

Unit-5 Group Codes.

Let $E : Z_2^m \rightarrow Z_2^n$, $n > m$, be an encoding function

and $C = E[Z_2^m] = \{ E(w) \mid w \in Z_2^n \}$ be the set of codes.

Then C is called group code if C is a subgroup of

Z_2^n.

Eg:- Consider the encoding function, $E : Z_2^2 \rightarrow Z_2^6$

of the triple repetition codes.

$\Rightarrow$ For this code ; we have

$E(0\ 0) = 000000$ $\qquad$ $E(1\ 0) = 101010$

$E(0\ 1) = 010101$ $\qquad$ $E(11) = 111111$

$\therefore C = \{000000, 010101, 101010, 111111\}$

Hamming Matrices:-

For a specified positive integer K,

Let $m = (2^K - 1 - K)$ and $n = (2^K - 1)$

Consider a group code from Z_2^m to Z_2^n given by the generator matrix $G = [I_m / B^T]$, where B is a $K \times m$ matrix.

Then the associated parity-check matrix $H = [B / I_K]$ is called Hamming matrix.

and code being considered is called Hamming code $(2^k-1, 2^k-1-k)$.

Theorem - 1

1) In a group code, the minimum distance b/w al distinct code words is the minimum of the weights of the non-zero elements of the code.

Proof :- Consider two elements $a, b \in G$ with $a \neq b$

Such that $d(a,b)$ is minimum.

Let 'c' be an element in G with minimum weight.

Since $a, b \in G$ & G is a subgroup

then $a + b \in G$

Also $d(a,b) = wt(a+b)$,

Since c has minimum weight, we have

$wt(c) \leq wt(a+b)$

i.e, $wt(c) \leq d(a,b)$

Next we note that $c + 0 = c$ where '0' is the identity in G

$\therefore d(c,0) = wt(c)$

code

On the other hand $d(c,0) \geq d(a,b)$

because $d(a,b)$ is minimum.

$\therefore wt(c) \geq d(a,b)$

Thus $wt(c) \leq d(a,b)$ & $wt(c) \geq d(a,b)$

$\therefore d(a,b) = wt(c)$

Hence proved.

Prove that "Let C be a group code in Z_2^n. If $r \in Z_2^n$ is received word and r is decoded as the code word c^*, then $d(c^*,r) \leq d(c,r)$, $\forall c \in C$.

Theorem-2

- Let C be a group code in Z_2^n. If $r \in Z_2^n$ is a received word and r is decoded as the code word c^*, then $d(c^*,r) \leq d(c,r)$ $\forall$ $c \in C$.

Proof:- Let $x + C$ be a coset containing r, where x is the element of minimum weight in the coset.

then, $r = x + c^*$ for some $c^* \in C$

$$\Rightarrow r + c^* = (x + c^*) + c^* = x + (c^* + c^*)$$
$$= x + 0$$
$$= x$$

so that, $d(c^*,r) = wt(c^* + r)$
$= wt(r + c^*) = wt(x)$

For every $c \in C$ we have.

$$c + r = c + (x + c^*)$$
$$= x + (c + c^*) \in x + C \quad (\because c + c^* \in C)$$

Consequently $wt(c+r) > wt(x)$

Because in $x + C$, $wt(x)$ is minimum.

i.e, $d(c, r) \geq wt(x)$

Thus we proved that

$$d(c^*, r) = wt(x) \leq d(c, r)$$

Hence proved

1) Suppose the encoding function $E: Z_2^m \to Z_2^{m+1}$ is defined as follows

$$E(w) = E(w_1, w_2, \ldots w_m) = w_1, w_2 \ldots w_m, w_{m+1}$$

where $w_{m+1} = \begin{cases} 0 & \text{if } w \text{ contains even no of 1's} \\ 1 & \text{if } w \text{ contains odd no of 1's} \end{cases}$

and corresponding decoding function $D: Z_2^{m+1} \to Z_2^m$ is

$$D(r) = D(r_1, r_2 \ldots r_m, r_{m+1}) = r_1, r_2 \ldots r_m$$

a) Find the code words assigned E to the following messages in Z,

000, 001, 011, 100, 110, 101, 111, 010

Find the decoded words assigned D to the following received words in Z_2^4

0000, 0001, 0101, 1111, 1010, 1100, 1101, 1001.

Soln:- a) By using the definition of E, we find that

$E(000) = 0000$ $E(001) = 0011$

$E(011) = 0110$ $E(100) = 1001$

$E(110) = 1100$ $E(101) = 1010$

$E(111) = 1111$ $E(010) = 0101$

b) By using the definition of D we find that

$D(0000) = 000$ $D(0001) = 000$

$D(0101) = 010$ $D(1111) = 111$

$D(1010) = 101$ $D(1100) = 110$

$D(1101) = 110$ $D(1001) = 100$

2) The parity check matrix for an encoding function $E: Z_2^3 \longrightarrow Z_2^6$ is given by

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

a) Determine the associated generator matrix.

b) Does this code correct all single errors in transmission.

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

which is of the form $H = [A^T / I_3]$

Accordingly,

$$A^T = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \therefore A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Hence the associated generator matrix is

$$G = [I_3 | A] = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

b) We observe that two columns of are identical [namely that 2nd & 5th].

But if any two columns are identical.

H does not provide a decoding scheme that corrects single error is transmission.

3) An encoding $E : Z_2^2 \rightarrow Z_2^5$ is given by the generator matrix $G = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$

a) Determine all the code words. What can be said about the error-detection capability of this code? What about its error-correction capability?

b) Find the associated parity check matrix H.

c) Use H to decode the received words: 11101, 11011.

Soln:- WKT, $G = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$ is of the form

$G = [I_2 | A]$, where

$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ & $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

a) We find that

$[E(00)] = [0\ 0]G = [0\ 0]\begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$

$= [0\ 0\ 0\ 0\ 0]$

$[E(0\ 1)] = [0\ 1]G = [0\ 1]\begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$

$= [0\ 1\ 0\ 1\ 1]$

$$[E(1\,0)] = [1\ 0]\begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

$$= [1\ 0\ 1\ 1\ 0]$$

$$[E(1\,1)] = [1\ 1]\begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

$$= [1\ 1\ 1\ 0\ 1]$$

These matrix equations show that the code words are

$E(0\,0) = 00000$

$E(0\,1) = 01011$

$E(1\,0) = 10110$

$E(1\,1) = 11101$

From these, we find that

$d(E(00) \cdot E(0\,1)) = 3$

$d(E(0\,0), E(1\,1)) = 4$ $\quad d(E(0\,0), E(1,0)) = 3$

$d(E(0\,1), E(1\,1)) = 3$

$d(E(0,1)\ E(1\,0)) = 4$

$d(E(1\,0)\ E(1\,1)) = 3$

Thus $\min(E) = 3$

$\therefore$ the code can detect all errors of weight ≤ 2 & can correct all single errors.

) The parity check matrix. H associated with G is given by

$$H = [A^T \mid I_3] = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

We observe that H does not contain a column of 0's and further not two columns of H are identical

∴ H corrects single errors in transmission.

c) For $r = 11101$, the syndrome of r is

$$H[1\ 1\ 1\ 0\ 1]^T = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

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

Since this is a zero matrix., the decoded message is got by retaining the first two components of r.

The decoded message is therefore 1 1.

⇒ For $r = 11011$, the syndrome of r is

$$H[1\ 1\ 0\ 1\ 1]^T = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

We observe that this matrix is identical with the first column of H.

$\therefore$ We change the first component of r (from 1 to 0) to get 01011. This is the code word. The first two components of this code word, namely 01, is the original message.

Decoding with coset leaders

In the following two Examples, we illustrate a procedure for the decoding of received words in group codes.

Example 1 *A group code C is defined by the generator matrix*

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

Decode the following received words using the cosets of C:

$$11110,\ 11011,\ 10100,\ 10101.$$

▶ We note that the given G is a 2×5 matrix. Therefore, the encoding function E is from Z_2^2 to Z_2^5 and is defined by $[E(w)] = [w]G$ for every $w \in Z_2^2$. We find that

$$[E(00)] = \begin{bmatrix} 0 & 0 \end{bmatrix} G = [0\,0\,0\,0\,0],$$
$$[E(01)] = \begin{bmatrix} 0 & 1 \end{bmatrix} G = [0\,1\,0\,1\,1],$$
$$[E(10)] = \begin{bmatrix} 1 & 0 \end{bmatrix} G = [1\,0\,1\,1\,0],$$
$$[E(11)] = \begin{bmatrix} 1 & 1 \end{bmatrix} G = [1\,1\,1\,0\,1].$$

Thus, we have

$$E\left(Z_2^3\right) = C = \{00000, 01011, 10110, 11101\}.$$

This is a subgroup of Z_2^5. We now select some $x_1 \in Z_2^5$ such that $x_1 \notin C$ and $wt(x_1)$ is minimum, and construct the coset $x_1 + C$. Let us take $x_1 = 10000$, so that

$$\begin{aligned} x_1 + C &= \{x_1 + c \mid c \in C\} \\ &= \{10000, 11011, 00110, 01101\} \end{aligned}$$

Now, we select some $x_2 \in Z_2^5$ such that $x_2 \notin C$, $x_2 \notin x_1 + C$ and $wt(x_2)$ is minimum, and construct the coset $x_2 + C$. Let us take $x_2 = 01000$, so that

$$x_2 + C = \{x_2 + c \mid c \in C\} = \{01000, 00011, 11110, 10101\}.$$

Similarly, we construct the cosets $x_3 + C$, $x_4 + C$ and so on by choosing x_k such that $x_k \notin C$, $x_k \notin x_{k-1} + C$ with $wt(x_k)$ is minimum, $k = 3, 4, \ldots$ We stop the process when the union of C and the distinct cosets so constructed is equal to Z_2^5. Since $o(C) = 4 = 2^2$ and $o\left(Z_2^5\right) = 2^5 = 32$, the number of distinct cosets of C in Z_2^5 is, by Lagrange's Theorem,

$$\left(Z_2^5 : C\right) = \frac{o\left(Z_2^5\right)}{o(C)} = \frac{2^5}{2^2} = 2^3 = 8.$$

The elements of C and those of the cosets of C are shown in the following table, called the *Decoding Table*.

Decoding Table for the Code of Example 1

00000	01011	10110	11101
10000	11011	00110	01100
01000	00011	11110	10101
00100	01111	10010	11001
00010	01001	10100	11111
00001	01010	10111	11100
11000	10011	01110	00101
10100	10111	00010	01001

We observe that the elements in the first column of the above table are $x_0 = 0, x_1, x_2, \ldots, x_7$, here $x_1, x_2, \ldots, x_7$ are chosen as described above, and the rows containing these are the elements of the corresponding cosets of C. Here, $x_1, x_2 \ldots x_7$ are called the *coset leaders*.

The decoding table enables us to decode any received word $r \in Z_2^5$ by adopting the following rule: Find the column containing r and identify the code word c in C that belongs to this column. Decode this c by deleting the last three components of c; the resulting word w is the message sent.

We observe that $r = 11110$ appears in the third column, and the corresponding $c = 10110$. Therefore, the corresponding message is $w = 10$.

The word $r = 11011$ appears in the second column, and the corresponding $c = 01011$. Therefore, the corresponding message is $w = 01$.

Similarly, we find that the messages corresponding to the received words 10100, 10101 are 00, 11. ∎