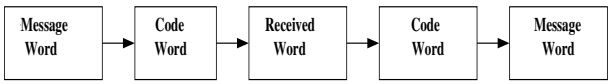
1 INTRODUCTION

Coding theory basically deals with how coding and decoding should take place such that error between the sending end and receiving end is minimized. Basically transmission of signal is not done through message words but through code words. Detection of error through message word is not possible as each received word is a message word, so unable to detect whether the received word is original message word or distorted word. If we transmit the message through code words detection of error is possible as each distorted word is not a code word.

In coding theory we are concerned with devising methods of encoding and decoding so that the errors which occur due to disturbance in channel, if not altogether eliminated, are minimized. One of the assumptions about the channel is that it does not increase or decrease the length of sequence that passes through it. Whole of this topic deals with how to convert message word to code word and from code word to message word. The process of coding theory is pictorially shown as follows.



2 MATHEMATICAL PRELIMANARIES

GROUP: A non empty set G is said to be Group if there exits one operation normally said "+". Then it should follow these properties.
(i) If a belongs to G, b belongs to G then $a + b$ should belongs to G.
(ii) $(a + b) + c = a + (b + c)$.
(iii) There should exits identity element such that $a + e = e + a = a$. Then e is the identity element.
(iv) There should exits additive inverse such that a+b=e

Example: set of vectors with 3 tuples in a binary field.

$$\{\ 000,001,010,011,100,101,110,111\ \}$$

ABELIAN GROUP: If $a + b = b + a$ then it is called Abelian group.

RING: A non empty set R is said to be Ring if it is defined over 2 operations. Let us say addition and multiplication. Then it should satisfy the following properties.
(i) All the properties of group over addition.
(ii) If a and b belongs to R then a*b should belongs to R.
(iii) It has to obey distributive property .

$$a * (b + c) = a * b + a * c$$

(iv) It has multiplicative identity.

$$a * e = e * a$$

Then e is the multiplicative identity element.
Example: set of all integers. {1,2,3,.....}

FIELD: A non empty set F is said to be field if it satisfies all the properties of a Ring along with the following property.
It should have multiplicative inverse.
Example: set of rational numbers. {1,2,3.5,0.55,.....} The multiplicative inverse of 2 is 0.5 and multiplicative inverse of 0.2 is 5. These are all consists of infinite elements.

GALOIS FIELD: A non empty set GF with finite numbers of elements satisfying all the properties of field. This field should contain number of elements as prime. GF_p:{ 0, 1....p-1} is a set of p elements over which addition and multiplication are defined by modulo function.

$$a + b = (a + b) \bmod p$$
$$a * b = (a * b) \bmod p$$

For example let p=3 then the elements in the set GF_p={0,1,2}. Then

$$2+1= 3 \bmod 3=0$$

sum of 2 and 1 is "0"

$$2*2= 4 \bmod 3 = 1$$

multiplication of 2 and 2 is 1.

3 ELEMENTARY PROPERTIES

Weight of any code word is defined as number of ones in the given code Word. It is represented as $wt(a)$, where a is the code word.
For example if a={001101010} then weight of a is $wt(a) = 4$.

Distance between two words is defined as the number of digits differ by each word.
For example if a={100101} and b={010110}
Then distance between a and b is $d(a,b) = 4$;

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

Proof: Let $a = \{a_1a_2a_3....a_n\}$ and $b = \{b_1b_2b_3....b_n\}$. For any i, $1 \leq i \leq n, a_i + b_i = 1$ iff $a_i \neq b_i$. Hence the pair (a_i, b_i) contributes 1 to $d(a,b)$.
Therefore, $d(a,b) = wt(a+b)$.
(ii) $d(a+c, b+c) = d(a,b) = wt(a+b)$
(iii) $d(a,b) \leq d(a,c) + d(c,b)$

Nearest neighbour decoding principle: If a word r is received and r belongs to Bn and r is not a code word, then the distance between the r and all the code words is considered. The code word corresponding to minimum distance $d(a,r) = d$ is considered as the code word for the received word r. If minimum distance d is with more than one code word then we say that decoding is failed.

Theorem: For a set of code words to detect all sets of k or fewer errors the minimum distance between any two code words is $k+1$ or more.
Proof: If the distance between any two code words is less than $k+1$ then if any error has occurred at k places then received word with error will be another code word. So the distance between all the code words should be more than $k+1$. The nonzero code word with less number of ones should be at least having weight of k+1.

Theorem: For a set of code words with minimum distance 2k+1 is capable of correcting any pattern of k or fewer errors.
Proof: If the distance of code word set is minimum of $2k+1$ then the received word with errors at maximum of k places will be nearer to only one code word . It implies that the received word is nearer to that particular code word. So the received word can be replaced by the nearest code word.

If the received word is having more than k errors then the received word will be nearer to another code word which is not its original code word. So If we correct that received word with the nearest code word also , the original code word corresponding to that received word will not be same. So in order to correct any received word in set of code words with minimum distance of $2k+1$ the error in the received word would not be more than k.

4 MATRIX ENCODING TECHNIQUES

For any (m,n) generator matrix G the encoding function $E : B_m \longrightarrow B_n$ is a monomorphism. It is given by $E(X) = X.G$ Here the function E is one to one . so it is called monomorphism. In a set of code words sum of any two code words is another code word. Then it is called group code. In group codes if error is a code word then it is not detected as the original code word plus error gives other code word
Encoding Process: For a given set of message words of length m the generator matrix is of size (m,n) where n is the length of required code word such that $n > m$. Each message word is of length m. So in the field of binary it has order of $2m$. The generator matrix will produce $2m$ code words which are uniquely determined by the each message word. Generator matrix is chosen as follows.

$$G = [IA]$$

of dimension (m,n)
Code word is obtained as follows:

$$[a_1a_2.....a_m] * [G] = [a_1a_2....a_ma_{m+1}.....a_n]$$

where m is length of message word and n is the length of code word.
For example the encoding matrix will be as follows:

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

Here the message words are of length 4 and code word is of length 7. Total number of code words will be 23 where as total of received words can be 27.

Code word corresponding to message word {1 0 0} is {1 0 0 1 0 1 1}.
Code check process: Decoding process is done by parity check matrix H. This parity check matrix is uniquely determined for a given generator matrix such that

$$H = [A^T I]$$

of dimension (n-m,n). For the above generator matrix G the corresponding parity check matrix will be as follows.

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

Now multiply one of the code words with the parity check matrix H. Then it has to map to zero vector. For example consider the code word {1 0 0 1 0 1 1} corresponding to the message word {1 0 0} and multiply this code word with matrix H. It result in Single column null vector. The received word to be code word the following condition is required.

$$[H] * [a_1 a_2 a_3 ... a_m ... a_n]^T = 0;$$

If $[H] * [a_1 a_2 a_3 ... a_m ... a_n]^T = s$ such that s resembles i^{th} coloumn of matrix H , then it implies that error occurred at i^{th} position of code word. If s doesn't resembles any one column of the matrix H matrix, it implies that more than one error has occurred. All the code words are said to be kernel for the transformation H.

5 POLYNOMIAL CODES

Getting basis vectors in linear codes such that the minimum distance of code words is desired one, can not be done easily. We have to choose proper basis vectors and has to check whether the code word set has minimum distance which we required. It is trial and error process which is not required. So we are taking certain polynomial as encoding polynomial which satisfies the requirement. Polynomial code words are obtained by multiplying message word polynomials with encoding polynomials. Each word of n tuple can be represented by a polynomial of at most order $n-1$.
Example: $\{a_0, a_1, a_2, ...a_n\}$ can be written as polynomial as follows.

$$\{a_0 + a_1 X + a_2 X^2 + ...a_n X^n\}$$

Then $g(X)$ is called encoding polynomial .The code word or code polynomial is given by $c(X) = a(X) * g(X)$. where $a(X)$ is the message word equivalent polynomial. If degree of $g(X) = r$, and degree of $a(X)$ is m , then degree of $c(X)$ equals sum of r and m. Number of terms in $c(X) = (r+1) + (m+1) - 1 = r + m + 1$ Let $g(X) = g_0 + g_1 X + ...g_n X^r$ then g_0 should not be zero. If $g_0 = 0$ then first entry of all the code words is zero which is not required. The polynomial code of length $n = r + m + 1$ generated by the encoding polynomial $g(X)$ is a subspace of F^n . Number of basis vectors for this subspace will be $m+1$ and number of code words are F^{m+1}.

Theorem: If the encoding polynomial $g(X)$ doesn't divide the polynomial of the form $X^k + 1$ such that for all $k < n$, then the code words has minimum distance of 3.
Proof: Each code word is divisible by $g(X)$ and $g_0 \neq 0$ .So $g(X)$ is not divisible by X. Let code word has at most two non zero terms

$$C(X) = X^i + X^j = X^i(1 + X^{i-j})$$

$g(X)$ should divide $c(X)$ But X^i and $g(X)$ are relatively prime. So $g(X)$ should divide $(1+X^{i-j})$, i.e. $g(X)|(1 + X^{i-j})$ which contradicts the hypothesis. So code words has minimum of distance of 3. An error vector $e = \{e_0 e_1 .....e_n\}$ will be undetected if its corresponding polynomial $e(X)$ is divisible by $g(X)$.

We say that exponent of a polynomial $g(X)$ is the least integer e such that $g(X)|(X^e + 1)$.
Theorem: In binary encoding polynomial $g(X) = (1+X)h(X)$ such that $h(X)$ has exponent $e > n$. Then any combination of two single or double errors are detected.
Proof: Let $e(X) = X^i + X^j = X^i(1 + X^{i-j})$ and $h(X)$ doesn't divide $(1 + X^{i-j})$. So $e(X)$ is undetected. If

$$\begin{aligned} e(X) &= X^i + X^{i+1} + X^j \\ &= X^i(1+X) + X^j \\ &= X^i(1+X)(1 + X^{j-i-1}) \end{aligned}$$

Then $h(X)$ divides $X^e + 1$ such that $e > n$. Then $h(X)$ can't divide $e(X)$ so the error is detected.
If

$$\begin{aligned} e(X) &= X^i + X^{i+1} + X^j + X^{j+1} \\ &= (1+X)X^i(1 + X^{j-i}) \end{aligned}$$

So the error is detected. If $g(X)$ is divisible by $(X+1)$ then every code word is of even length.

Let us consider $g(X) = 1 + X + X^3$ and message word of length 3 then length of code words will be 6. Then $g(X)$ does not divide the polynomial of the form $X^i + 1$ for all $i < 6$. So we can treat this polynomial as the encoding polynomial. The code words corresponding to the message word will be given as follows.

Message word	Code word
0 0 0	0 0 0 0 0 0
0 0 1	0 0 1 1 0 1
0 1 0	0 1 1 0 1 0
1 0 0	1 1 0 1 0 0
0 1 1	0 1 0 1 1 1
1 0 1	1 1 1 0 0 1
1 1 0	1 0 1 1 1 0
1 1 1	1 0 0 0 1 1

In this case the minimum distance between any two code words is 3.

6 GENERATOR AND PARITY CHECK MATRIX GENERAL CASE

The generator matrix need not be always in the form of $G = [IA]$. It can be in any form but of dimension (m, n). Polynomial code is a matrix code and it is a group code. For a encoding polynomial $g(X) = g_0 + g_1X + ...g_kX^k$ And for a message word of length m the length of code word is is $n = m + k$. Then the generating matrix will be as follows.

$$\text{G=}\begin{bmatrix} g_o & g_1 & g_2 & . & g_k & 0 & 0 & 0 & 0 \\ 0 & g_o & g_1 & g_2 & . & g_k & 0 & 0 & 0 \\ 0 & 0 & g_o & g_1 & g_2 & . & g_k & 0 & 0 \\ : & : & : & : & : & & : & : & : \\ 0 & 0 & 0 & 0 & g_o & g_1 & g_2 & . & g_k \end{bmatrix}$$

Each row of generator matrix will be a code word corresponding to the message words {1000000}, {010000}, {0010000} etc. It implies that the rows of generator matrix are basis for the subspace of code words. Let the basis be $\{b_1b_2b_3...b_n\}$ of code words , then the order of code words is 2^m. Let H be the parity check matrix. Each basis code word operated on it should map to zero. It implies that the generator matrix G^T operated on H should map to zero.

$$HG^T = 0 \qquad H \text{ is of dimension } (n - m, n)$$

$$GH^T = 0 \qquad G \text{ is of dimension } (m, n)$$

H^T will have $(n - m)$ columns operated on G and mapped to zero. Find out the kernel of this mapping . The dimension of this kernel is $(n - m)$. So there exists $(n - m)$ linearly independent vectors which map to zero. Those $(n - m)$ vectors are columns vectors of H.

7 HAMMING CODES

For a given r, we can always construct hamming codes such that message word length $m = 2r - r - 1$ and code word length $n = 2r - 1$. The code word will have r positions more compared to message word. The code word consists of $\{b_1b_2....b_n\}$ and message word is in the form of $\{a_1a_2.....a_m\}$. Replace all b_i by a_i in their original order such that the positions of b i.e. i is not equal to j^{th} power of 2. $\forall\ 0 \leq j \leq r - 1$
Then its code word will be as follows.

$$\{b_1b_2a_1b_4a_2....b_2^{r-1}\}$$

. In this code word there are r unknown variables. These are $b_i\ \forall\ 0 \leq j \leq r - 1$
Now we have to find out these unknown variables . For this purpose get a matrix M of size $(2r - 1, r)$ such that each i^{th} row of M is represented as binary equivalent of i. Then solve the equation $bM = 0$. We will get the unknown variables $\{b_1b_2b_4....b_2^{r-1}\}$.

8 BCH CODES

Here our interest is to get the generating polynomial for a code word of length n and the minimum distance of that set of code words is d. The corresponding $g(X)$ is found out as follows.

$$g(X) = LCM[M_1(X), M_2(X).....M_{d-1}(X)]$$

For finding out these $M_1(X), M_2(X).....M_{d-1}(X)$ we need the concept of extension field. The order of the extension field should be greater than (length of code word+1). For binary field, the extension field order is given by $2^r \geq n + 1$. Find the minimum r which satisfies this equation. Find a irreducible polynomial of degree r. Then the extension field K is given by

$$K = B[X]/ < X^r + X.... + 1 >$$

Let a be the primitive element of field K.

$$a = X + < X^r + X..... + 1 >$$

Then

$$M_1(X) = (X - a)(X - a^2)(X - a^4).....$$

$$M_3(X) = (X - a^3)(X - a^6)(X - a^{12})....$$

The minimum polynomial for a set of numbers will be same such that relation in between the set is given by

$$C_s = \{s, qs, ...q^{ms-1}s\}$$

Where ms is the smallest positive integer such that $s \equiv sq^{ms}(modn)$ The power of a are the elements from the set C_s.

9 DUAL CODES

From a known encoding polynomial we can get encoding matrix corresponding to the length of code word and set of code words. Using the same polynomial or encoding matrix we can generate another set of code words of same length but which are orthogonal to the set of previous code words. For getting another set of polynomials which are related by orthogonality we are using the concept of dual codes. For a given linear code $C[n, k]$ the generator matrix can be known by taking basis code words from the set of code words. The parity check matrix H is given for a generator matrix G by the following relation.

$$GH^T = 0$$

Dual code of a set of code words is given by the relation dot product. If C is code word then $C^\perp$ is dual iff $C.C^\perp = 0$. C is the set of code words are linear combinations of rows of G so all the columns of H^T operated on C will map to zero. It implies that columns of $C^\perp$ are linear combinations of columns of H^T . the set of dual code words are obtained by considering H as the generating matrix for the dual code words $C^\perp[n, n-k]$
The order of the code words $= 2^k$
Message word length should be k for code words.
Where as the order of dual code words $= 2^{n-k}$
Message word length for dual code is $= n - k$

10 CYCLIC CODES

A linear code C of length n over F is cyclic if any cyclic shift of a code word is again a code word, i.e. if $\{a_0a_2....a_{n-1}\}$ is in C then $\{a_{n-1}, a_0, a_1, ...a_{n-2}\}$ should be in C. The encoding polynomial which we are taking should be a factor of $X^n + 1$ and the encoding polynomial should be irreducible over that field.
Check polynomial: Let C be a cyclic code of length n over F with generator polynomial $g(X)$ of degree r. Let $h(X)$ be the polynomial of degree $(n-r)$ such that it satisfies the following equation.

$$X^n + 1 = g(X) * h(X)$$

Any code word $C(X)$ is of the form

$$C(X) + I = a(X) * g(X) + I$$

Where $I =< X^n + 1 >$ then $C(X) * h(X) + I = a(X)*g(X)*h(X)+I$ And therefore $C(X)*h(X) = 0, i.e. C(X)*h(X)$ is zero in $F[X]/I$. For this reason $h(X)$ is called the check polynomial of the code C. In matrix form H will be written as follows of dimension (n,n).

$$H = \begin{bmatrix} . & . & . & . & h_3 & h_2 & h_1 & h_0 \\ . & . & . & h_3 & h_2 & h_1 & h_0 & . \\ . & . & h_3 & h_2 & h_1 & h_0 & . & . \\ & & & & & & & \\ : & : & : & : & & & : & : \\ & & & & & & & \\ h_0 & . & . & . & . & h_3 & h_2 & h_1 \end{bmatrix}$$

11 POLYNOMIAL FACTORIZATION

For a given polynomial of the form X^n+1 partition the set $\{0, 1, ...n-1\}$ as cyclotomic sets basing on the given field. Each cyclotomic set is given by

$$C_s = \{s, qs, ...q^{ms-1}s\}$$

Such that ms is the least positive integer which satisfies the relation $S = q^{ms}s(modn)$. Then the polynomial can be factorized as follows

$$X^n + 1 = \prod_s M_s(X)$$

Such that $M_s(X)$ is the minimum polynomial for the cyclotomic set C_s.

$$M_s(X) = (X - \alpha^s)(X - \alpha^{qs})....(X - \alpha^{sq^{ms-1}})$$

Where α is the primitive element for the extension field K and K is given by

$$K = B[X]/ < X^r + ... + 1 >$$

$X^r + ... + 1$ is a irreducible polynomial of order r, and r is smallest integer such that $n|(2r-1)$.

12 BERLEKAMP'S ALGORITHM

For any given polynomial

$$f(X) = \sum_{i=0}^{m} a_i X^i$$

$\forall 0 \leq i \leq m$
Find a matrix Q of order m such that i^{th} row of Q is given by $X^{q(i-1)}$ Modulo $f(X)$ over a field of q elements. Then get the null space for $(Q-I)$ such that

$$(g_0 g_1 ..... g_{m-1}) * (Q - I) = 0$$

. Then find the $(g_o ..... g_{m-1})$,its equivalent polynomial $g(X) = g_0 + g_1 X + ...g_{m-1} X^{m-1}$. Then the polynomial can be written as follows.

$$f(X) = \sum_{i=0}^{m} a_i X^i = \prod_{s \in F} gcd(f(X), (g(X) - s))$$

Berlekamp's algorithm: A special case
Most of the time we are concerned with the factorization of $X^n + 1$. For such a polynomial factorization consider a set $S = \{0, 1, ...n-1\}$ And divide this set as cyclotomic sets relative to field of q elements.

$$C_s == \{s, qs, ...q^{ms-1}s\}$$

where ms is the least positive integer which satisfies the equation $s = q^{ms}s(modn)$. Then form of polynomials for corresponding to each cyclotomic set such that

$$P_i(X) = \sum_{j \in C_i} X^j$$

$j \in C_i$ Then the polynomial can be expressed as

$$X^n + 1 = gcd(X^n + 1, (g(X) - s))$$

$s \in F$. Where $g(X)$ is linear combination of $P_i(X)$ over the field F.

References

[1] Lekh R. Vermani, "Elements of algebraic coding theory", Chapman and Hall Mathematics Series. First edition

[2] Kenneth Hoffman, Ray Kunze, "Linear Algebra", Prentice Hall of India. Second edition