INTERNATIONAL JOURNAL OF ENHANCED RESEARCH IN SCIENCE TECHNOLOGY & ENGINEERING

Vol.1 Issue 1, OCT-2012 ISSN NO: 2319-7463

Codes and Group Codes

Shipra Gupta

M. Phil, Department of Mathematics, University of Delhi

Codes And Group Codes: In today's modern world of communication, data items are constantly being transmitted from point to point. We have various examples in the form of television transmission, radio transmission or a telephone call. The Basic problem in transmission of data is that the received data may differ from the transmitted data. Distortion or error can be caused by a number of factors such as noise, weather, and electrical problem and so on.

Another entirely different problem is that of modifying the data being sent so that only the intended recipient is able to reconstitute the actual data.

Coding theory has developed techniques for detecting and sometimes correcting error in the transmitted data.
The American mathematician Claude E. Shannon (1916-2001), who worked at a Bell Laboratories, published a paper in 1948 that described a mathematical theory of communication and thereby founded the field of information theory. Shortly thereafter, Richard Hamming and his colleagues at Bell Laboratories laid the foundation for error-correcting codes.

PRELIMINARY DEFINITIONS

The basic unit of information, called a MESSAGE is a finite sequence of characters from a finite alphabet: binary alphabet.

BINARY ALPHABET

- By a binary alphabet we mean the set B = { 0, 1 }

So every character or symbol to be transmitted is represented in binary form.

WORD

A sequence of letters from binary alphabet is called a word, i.e, it is a sequence of m 0's and 1's.

CODE

A code is a collection of words that are used to represent distinct messages.

CODE WORD

A word in a code is called a code word.

ERROR CORRECTION

Suppose a code word is transmitted from its origin to its destination, in the course of transmission, interferences such as noises might cause some of the 1 's to be received as 0's and some of the 0's to be received as 1 's. Consequently the received word might no longer be the transmitted one, and it is our desire to recover the transmitted word, if at all possible. This is what we mean by error correction.

The set B is a group under the binary operation `+', as defined in table 1.

Table 1

+	0	1
0	0	1
1	1	0

Vol.1 Issue 1, OCT-2012 ISSN NO: 2319-7463

Now if we think of B as the group Z_2, then '+' is merely the mod 2 addition. Therefore, B^m = B×B× ……..×B (m times) is a group under the operation (+) defined by:
$(x_1 , x_2 , \ldots\ldots\ldots\ldots,x_m)$ (+) $(y_1 , y_2 \ldots\ldots\ldots\ldots..Y_m) = (x_1+y_1, x_2+y_2 \ldots\ldots\ldots\ldots\ldots Xm+y_m)$
[Result used: If G_1, G2 are groups, then $G = G_1 \times G_2$ is a group with binary operation defined by (a , b) (c , d) = (ac , bd)]

$$Let\ B_m = b_1 b_2 \ldots\ldots\ldots\ldots b_m : b_i \in \{0,1\}\}$$

With x (+) y = the sequence of length m with 1's where x and y differ and 0's where x and y are same.
For example 1: x = 00101, y = 10110, x (+) y = 10011
Then, B^m forms a group with respect to the composition (+) defined above.
its identity is $\overline{0} = 00 \ldots\ldots\ldots\ldots\ldots\ldots 0$ and every element is its own inverse. (m times)
Also, $0(B^m) = 2^m$ (because each element can take two values 0,1)
Figure 1 shows the basic process of sending a word from one point to another point over a transmission channel. An element $x \in B^m$ is sent through the transmission channel and is received as an element $x_t \in B^m$

Fig. 1 Word $x \in B^m$ transmitted — Transmission channel — Word $x_t \in B^m$ received

If an error occurs then, $x \neq x_t$ and x_t could be any element of B^m.
The basic task of the transmission process is to reduce the likelihood of the error. We do it as follows:
Choose n, m ∈ Z_+ , n > m and a one-one function, called the (m,n) encoding function

$$e : B^m \rightarrow B^n$$

we view it as a means of representing every word in B^m as a word in B^ni.e, if $b \in B^m$, then e(b) is called the code word representing b.

The additional 0's and 1's can provide the means to detect or correct errors produced in the transmission channel the function e is one-one so that different words in B^m will be assigned different code words.

Word $b \in B^m$ → Encoded word $x = e(b) \in B^n$ → → → Transmission channel → Word $x_t \in B^n$
To be sent recieved

If the transmission channel is noiseless then $x = x_t$ for all x in B^n and
x = e (b) is received for each $b \in B^m$.
Since e is a known function b may be identified.
If errors occur then we say that code word e(b) has been transmitted with k or fewer errors if x and x_t differ in at least 1 but not more than k positions,i.e, e detects k or fewer errors if x = e(b) is transmitted with k or fewer errors.

WEIGHT OF A WORD
Let x be any word in B^m. The weight of x, denoted by w(x), is defined to be the number of 1's in x.
For example2: let x = 01100 y = 11111 z = 00000
then w(x) = 2 w(y) = 5 w(z)= 0

Example 3: **Parity check code**
The following encoding function $e: B^m \rightarrow B^{m+1}$ is called the parity (m, m+1) check code :$if\ for\ b = b_1 b_2 \ldots\ldots\ldots b_m \in B^m$,

$$e(b) = b_1 b_2 \ldots\ldots b_m b_{m+1}$$

$$\text{where, } b_{m+1} = \begin{cases} 0\ if\ w(b)\ is\ even \\ 1\ if\ w(b)\ is\ odd \end{cases}$$

We observe that $b_{m+1}$is zero iff the number of l's in b is an even number i.e, every codeword e(b) has even weight .A single error in the transmission of a code word will change the received word to a word of odd weight and therefore can be detected. In the same way we see that any odd number of errors can be detected.
For a concrete illustration of this encoding function we take m = 3. Then
e(000) = 0000

Vol.1 Issue 1, OCT-2012 ISSN NO: 2319-7463

e(001) = 0011
e(010) = 0101
e(011) = 0110
e(100) = 1001
e(101) = 1010
e(110) = 1100
e(111) = 1111

suppose now that b = 111. Then x = e (b) = 1111.
If the transmission channel transmits x as x_t = 1101, then w(x_t) = 3, and we know that odd number of errors (at least one) has occurred.
Remark: we see that if the received word has even weight, then we cannot conclude that the code word is transmitted correctly, since this encoding function does not detect an even number of errors. Despite this limitation Parity check code is widely used.
Let x and y be words in B^m. The HAMMING DISTANCE d(x,y) between x and y is the weight of x (+) y, i.e,

d(x,y) = w (x (+) y)

For example 4: x = 110110 y = 000101

x(+)y= 110011

w(x (+) y) = 4 = d(x,y)

Properties of the distance function :
1. d(x,y) = d(y,x)
2. d(x,y) = 0
3. d(x,y) = 0 iff x = y
4. d(x,z) = d(x,y) + d(y,z)

i.e, distance function satisfies the properties of a metric function. The minimum distance of an encoding function e :$B^m \rightarrow B^n$ is the minimum distance between all distinct pairs of code words, i.e. min$\{ d(e(x) , e(y) \mid x , y B^m \}$
Example5: consider the following (2,5) encoding function e :
e (00) = 00000
e (10) = 00111 code words
e (01) = 01110
e (11) = 11111
by computing the minimum distance between all distinct pairs of code words we get the minimum distance to be 2.
We have the following theorem:
An (m, n) encoding function e: $B^m \rightarrow B^n$ can detect k or fewer errors if and only if its minimum distance is at least k+1. So in the above example e will be able to detect maximum of only one error.

GROUP CODES

We now make use of the fact that (B^n, (+)) is a group.
An (m,n) encoding function e :$B^m \rightarrow B^n$ is called a group Code if

$$e\ (B^m) = \{ e(b) \mid b \in B^m \} = range(e) = N$$

is a subgroup of B_n.
By definition of subgroup, N is a subgroup of B^n if:

a) the identity of B^n is in N
b) if x andy belong to N then x(+)y ∈ N
c) if x is in N then its inverse is also in N

Property c) need not be checked because every element in B^n is its own inverse. Moreover since $B^n$is abelian, every subgroup of$B^n$is a normal subgroup.
Example6: Consider the (3,6) encoding function e:$B^3 \rightarrow B^6$ defined by

e(000) = 000000
e(001) = 001100
e(010) = 010011
e(011) = 011111
e(100) = 100101
e(101)= 101001

Vol.1 Issue 1, OCT-2012 ISSN NO: 2319-7463

e(110) = 110110
e(111) = 111010

We show that this encoding function is a group code. For that we must show that the set of all code words N= {000000,001100,010011,011111,100101,101001,110110,111010} is a subgroup of B^6.

a) identity of B^6 is in N, and
b)if x and y are elements of N, then x(+)y is an element of N.
Hence N is a subgroup of B^6 and the given encoding function is a group code.
Theorem: Let e: $B^m \to B^n$ be a group code. The minimum distance of e is the minimum weight of a nonzero code word.
Proof: Let d = d(x,y) be the minimum distance of the group code,where x and y are distinct code words. Let m be the minimum weight of a nonzero code word.
i.e, m = w(z), where z is a nonzero codeword
Since e is a group code, therefore x (+) y is a nonzero code word
Thus, d = d(x,y) = w(x(+)y) ≥ m ……..(1)
On the other hand, since 0 and z are distinct code words
m = w(z) = w(z(+)0) = d(z,0) ≥d……...(2)
From (1) and (2), we get that
m = d.
We now site an example to show the advantage of group codes
Example7: By the above theorem the minimum distance of the group code in example 6 is the minimum weight of a nonzero code word i.e, 2. But if we check this directly we would require 28 different Calculations.

DECODING AND ERROR CORRECTION
Consider an (m,n) encoding function e : $B^m \to B^n$
We have received the word

$$x = e(b) \in B^n , b \in B^m \text{ as } x_t$$

Our problem is to identify the word b,the original message. An onto function d :$B_n \to B_m$ is called an (n,m) decoding function associated with e if
d (x_t) = b' ∈ B_n
is such that when the transmission channel has no noise then
b' = b
i.e, d•e = 1_{B^m}, the identity function on B^m.
The decoding function is required to be onto so that every received word can be decoded to give a word in B^m.
Example8 : consider the parity check code (defined earlier)
We now define the decoding function d: $B^{m+1} \to B^m$.
$If\ y = y_1 y_2 \dots\dots\dots y_m y_{m+1} \in B_{m+1}$, then
d(y) = y_1y_2…….y_m.
We observe that if b = b_1b_2…… $b_m \in B^m$, then
(d • e)(b) = d(e(b)) = b,
So d • e = 1_{B^m}.
For a concrete example we take m = 4.
Then we obtain d(10010) = 1001 and
d(11001) = 1100.
Let e be an (m,n) encoding function and let d be an (n,m) decoding function associated with e. We say that the pair (e,d) corrects k or fewer errors if whenever x = e(b) is transmitted correctly or with k or fewer errors and x_t is received, then d(x_t) = b. Thus x_t is decoded as b.
Given an (m,n) encoding function $e: B^m \to B^n$,we often need to determine an (n,m) decoding function d:$B^n \to B^m$ associated with e. We now discuss a method called the maximum likelihood technique, for determining a decoding function d for a given e. Since B^m has 2^m elements, there are 2^m code words in B^n. We first list the code words in a fixed order:

$$X^{(1)}X^{(2)}\dots\dots\dots\dots X^{(2m)}.$$

Vol.1 Issue 1, OCT-2012 ISSN NO: 2319-7463

If the received word is x_t, we compute $d(x^{(i)},x_t)$ for $1 \leq i \leq 2^m$
And choose the first code word,say $x^{(s)}$, such that it is closest to x_t and first in the list. If $x^{(s)}$ = e(b),we define maximum likelihood decoding function d associated with e by $d(x_t) = b$.
We observe that d depends on the particular order in which code words are listed.
Theorem: Let e be an (m,n) encoding function and d a maximum likelihood decoding function associated with e. Then (e,d) can correct k or fewer errors if and only if minimum distance of e is at least 2k+1.
We now discuss a simple and effective technique for determining a maximum likelihood decoding function associated with a given group code :

STEP1: determine all the cosets of N =e(B^m) in B^n.
STEP2: for each coset pick the word of smallest weight. This word of smallest weight is called the coset leader.
STEP3: let s be a coset leader .compute $x = x_t (+) \varepsilon$ .If x = e(b)
We let d (x_t) = b .That is we decode x_t as b.
We now consider a question based on the above procedure.
Q) Consider the (2,4) group encoding function e :$B^2 \rightarrow B^4$,
Defined by

e(00) = 0000 e(10) =1101
e(01) = 0011 e(11) = 1110

Decode the following word relative to maximum likelihood Decoding function : 1011 (received word).
Solution: B^4= set of all binary sequences of length 4.
Therefore,$0(B^4) = 2^4 =16$.
We know that (B^4,(+)) is a group.
N = {0000,0011, 1101, 1110}
Then N is a subgroup of (B^4,(+)),
since N is finite we need to check only the closure property.
It is clear from the following composition table

(+)	0000	0011	1101	1110
0000	0000	0011	1101	1110
0011	0011	0000	1110	1101
1101	1101	1110	0000	0011
1110	1101	1101	0011	0000

Also since (B^4, (+)) is abelian ,(N ,(+)) is a normal subgroup of (B^4,(+)).
We now compute distinct cosets of N in B^4.
We know that order of (B^4/N) = $0(B^4)/0(N) = 16/4 = 4$.
Therefore there will be 4 cosets of N in B^4:
N(+)0000 = {0000 , 0011 , 1101 , 1110 }
N(+)1000 = {1000 , 1011 , 0101 , 0110 }
N(+)0100 = {0100 , 0111 , 1001 , 1010)
N(+)0010 = {0010 , 0001 , 1111 , 1100 }
Now x_t= 1011 $\in$N(+)1000 ={1000, 1011, 0101, 0110}
Now ε = leader of the coset N(+)1000
= word of minimum weight in the set
{1000, 1011, 0101, 0110}
= 1000
therefore, $x = x_t(+)\varepsilon$
= 1011(+)1000
= 0011
= e(01)
= e(b)
therefore, 01 is the decoded word.

Vol.1 Issue 1, OCT-2012 ISSN NO: 2319-7463

REFERENCES

[1] Error – Correcting Codes by Prof. Peterson (1965) The MIT Press.
[2] Discrete Mathematical Structures by Bernard Kolman Robart C. Busby Sharon Cutler Ross Pearson Education 2008.
[3] Elements of Discrete Mathematics by C.L.Liu Mac-Graw Hill Education, 1986.
[4] Humphreys, O.F. and Prest, M.Y. numbers, Groups and Codes, New York, Cambridge University Press 1990.