Information Theory

Mohamed Hamada

Software Engineering Lab
The University of Aizu

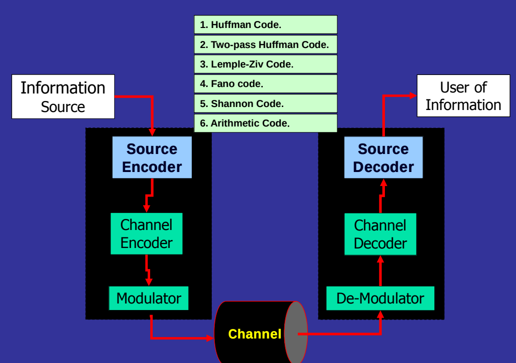
Email: hamada@u-aizu.ac.jp
URL: http://www.u-aizu.ac.jp/~hamada

Today's Topics

- **Channel Coding / Decoding**
- **Hamming Method:**
  - **- Hamming Distance**
  - **- Hamming Weight**
- **Hamming (4, 7)**

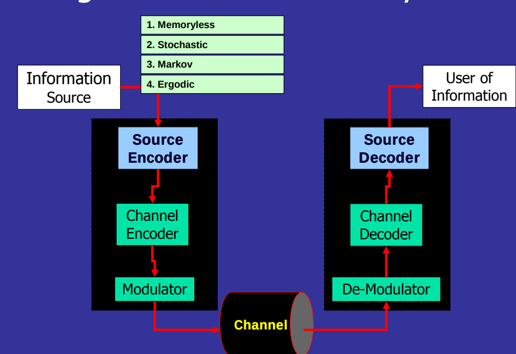
1

Digital Communication Systems



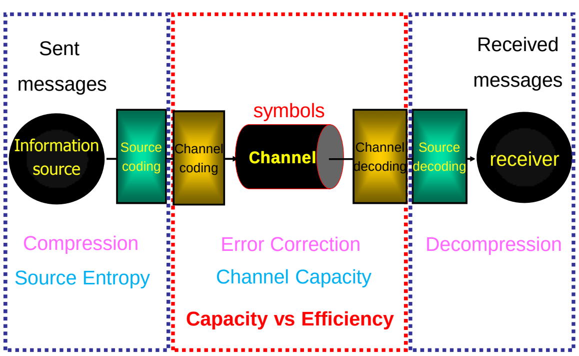
2

Digital Communication Systems



3

INFORMATION TRANSFER ACROSS CHANNELS



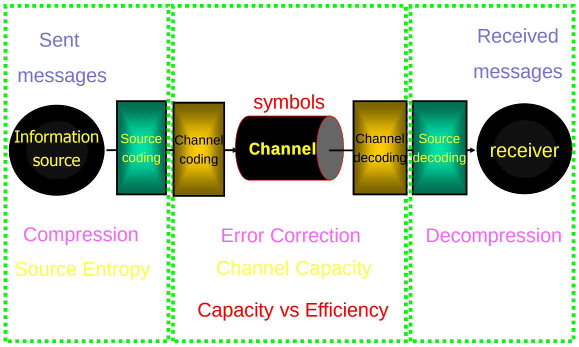
4

Channel Coding/Decoding

Hamming Method

5

INFORMATION TRANSFER ACROSS CHANNELS



6

Channel Coding/Decoding

The purpose of channel coding / decoding is to detect and correct errors in noisy channels

7

Error correction in our life

Cell phone	Hard disk	Optical disk	Fiber
10^{-5}–10^{-6}	10^{-14}	10^{-12}	10^{-12}

8

Channel Coding/Decoding

Error detection and correction

In a noisy channel errors may occur during the transmission of data from information source to destination, so we need a method to detect these errors and then correct them.

9

Channel Coding/Decoding

Hamming Method

- **It was the first complete error-detecting and error-correcting procedure.**
- **It represents one of the simplest and most common method for the transmission of information (in the presence of noise).**
- **It assumes that the source transmits binary messages (i.e. The information source alphabet is { 0, 1 })**
- **It uses the parity checker method to detect an error**
- **It assumes that the channel is a *binary symmetric channel* (BSC)**

10

Hamming Codes

- (4,7) Hamming Code detects all one- and two-bit errors
- Corrects all 1-bit errors
- Magic: Any two different codewords differ in at least 3 places!

0000000	0001011	0010111	0011100
0100110	0101101	0110001	0111010
1000101	1001110	1010010	1011001
1100011	1101000	1110100	1111111

11

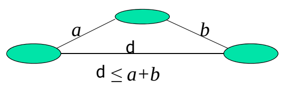
Hamming Distance

- Number of places in which two bit strings differ
-

1	0	0	0	1	0	1
1	0	0	1	1	1	0

= Hamming distance 3

- Acts like a distance:



Definitions

- Hamming distance between $\underline{x}$ and $\underline{y}$ is

  $d_H := d(\underline{x}, \underline{y})$ is the # of positions where $x_i \neq y_i$

- The minimum distance of a code C is
  - $d_{min} = \min \{ d(\underline{x}, \underline{y}) \mid \underline{x} \in C, \underline{y} \in C, \underline{x} \; \underline{y} \}$

- Hamming weight of a vector $\underline{x}$ is
  - $w(\underline{x}) := d(\underline{x}, \underline{0})$ is the # of positions where $x_i \neq 0$

Example

- *Hamming distance* d(1001, 0111) = 3
- *Minimum distance* (101, 011, 110) = 2
- *Hamming weight* w(0110101) = 4

Performance

A code with minimum distance d_{min} is capable of correcting t errors if

$$d_{min} \geq 2t + 1.$$

Proof: If $\leq t$ errors occur, then since $d_{min} \geq 2t + 1$ an incorrect code word has at least t+1 differences with the received word.

Hamming codes

We assume that the sequence of symbols generated by the information source is divided up into blocks of K symbols.

- Minimum distance 3
- Construction
  - G = $\left(I_m \quad \text{All k-tuples of Hamming weight} > 1 \right)$
  - where m = $2^k - k - 1$

I_m is identity matrix

Example: Hamming (7, 4) codes

We assume that the sequence of symbols generated by the information source is divided up into blocks of 4 symbols.

- Generating matrix
  - G = $\left(I_4 \quad \mathbf{P} \right)$

P is a 4x3 matrix determined by:

$C_1 = u_2 + u_3 + u_4$

$C_2 = u_1 + u_3 + u_4$

$C_3 = u_1 + u_2 + u_4$

Where + is modulo 2:
0+0=1+1=0 and
1+0=0+1=1
and u_i are I_4 elements

Example: Hamming (7, 4) codes

We assume that the sequence of symbols generated by the information source is divided up into blocks of 4 symbols. Codewords have length 7

- Generating matrix

$$G = \begin{pmatrix} I_4 & P \end{pmatrix} = \begin{array}{cccc|ccc} u_1 & u_2 & u_3 & u_4 & C_1 & C_2 & C_3 \\ \hline 1&0&0&0&0&1&1 \\ 0&1&0&0&1&0&1 \\ 0&0&1&0&1&1&0 \\ 0&0&0&1&1&1&1 \end{array}$$

I_4 = 1000, 0100, 0010, 0001; P = 011, 101, 110, 111

Where + is modulo 2:
0+0=1+1=0 and
1+0=0+1=1

and u_i are I_4 elements

$C_1=u_2+u_3+u_4$

$C_2=u_1+u_3+u_4$

$C_3=u_1+u_2+u_4$

18

Hamming (7, 4) Syndrome decoding

H is the parity check matrix

H^T is the Transpose matrix of H

P^T is the Transpose matrix of P

Let $G = [\, I_k \; P \,]$

For Hamming(7, 4) code: n=7 and k=4

Step 1. construct $H = \begin{pmatrix} P^T & I_{n-k} \end{pmatrix}$

Step 2. Arrange the columns of H in order of increasing binary values

Step 3. Determine the syndrome $S= y.\, H^T$ (y is the received message)

Step 4. If S=0 then no error occurs during transmisstion of information

Step 5. If $S\neq 0$ then S gives a binary representation of the error position (we assume only one error ocuured)

19

Example:

Suppose that y=(1111011) is received

$$G = \begin{array}{cc} I_k & P \\ \begin{pmatrix} 1000 & 011 \\ 0100 & 101 \\ 0010 & 110 \\ 0001 & 111 \end{pmatrix} \end{array} \quad \xrightarrow{\text{Step 1}} \quad H = \begin{array}{cc} P^T & I_{n-k} \\ \begin{pmatrix} 0111 & 100 \\ 1011 & 010 \\ 1101 & 001 \end{pmatrix} \end{array}$$

n=7 and k=4

Step 2 → $H = \begin{pmatrix} 0001111 \\ 0110011 \\ 1010101 \end{pmatrix}$

Step 3 → $S= y.\, H^T$ **= (101) = $(5)_{10}$**

Step 5

→ **An error had occurred at position 5 in the received message y =(1111011)**

→ **The correct sent message is then = (1111111)**

20