

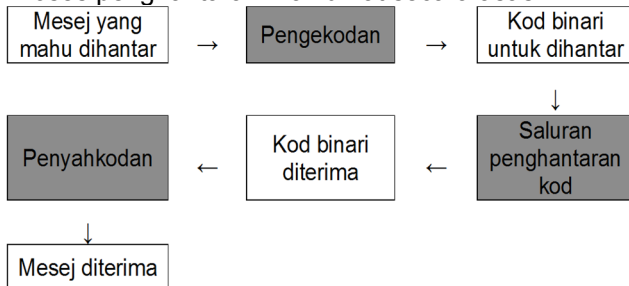
TOPIK 3: KOD DAN KRIPTOGRAFI

Pengenalan Kepada Sistem Penghantaran Data

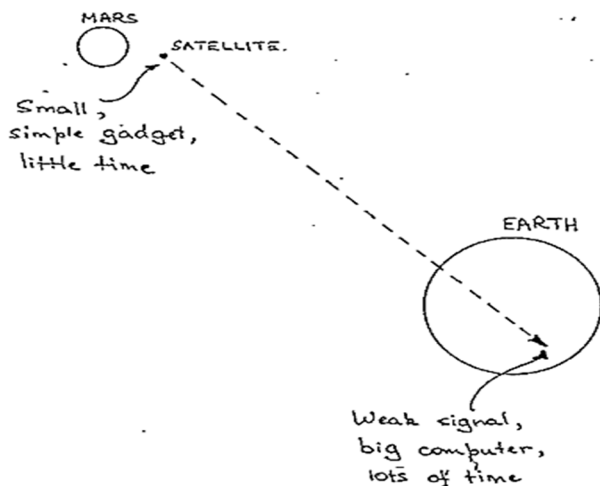
- Semua sistem yang bekerja dengan data secara digital menggunakan peranti elektronik adalah menggunakan teori pengkodan kerana semua akan mengalami gangguan yang dinamakan *noise*.
- Oleh itu, masalah teori pengkodan adalah antara masalah yang sangat asas dan yang paling kerap penyimpanan dan penghantaran maklumat.
- Keputusan teori pengkodan membenarkan untuk mewujudkan sistem yang boleh dipercayai daripada sistem yang tidak boleh dipercayai untuk menyimpan dan / atau menyampaikan maklumat.
- Kaedah teori pengkodan adalah aplikasi konsep yang sangat asas dan kaedah-kaedah (abstrak) algebra.

Penghantaran Data

- Proses penghantaran maklumat secara asas:



Kapal Angkasa *Mariner* 1969



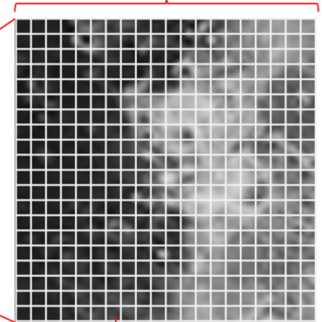
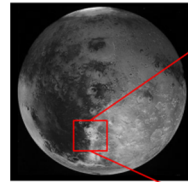
- 1965, Amerika Syarikat menghantar *Mariner* 4 yang merupakan kapal angkasa yang pertama untuk merakamkan imej Marikh.
- Pada masa itu, setiap imej yang dihantar ke bumi mengambil masa 8 jam. Terdapat banyak ralat kod yang berlaku.
- Misi *Mariner* 6 seterusnya menghasilkan imej yang lebih jelas dan tepat dengan menggunakan pembetulan ralat kod.
- Misi *Mariner* 6 telah menambahbaik misi yang sebelumnya.

- Penghantaran imej dalam misi *Mariner* 6

- Penghantaran imej melibatkan penghantaran grid-grid halus pada setiap imej yang dipanggil piksel.
- Kualiti setiap piksel akan dinilai melalui darjah kehitaman (*degree of blackness*) antara 0 hingga 63.

- Contoh imej Marikh:

Imej marikh dibahagikan kepada grid-grid kecil (piksel) untuk memudahkan penghantaran imej ke bumi



Contoh: Darjah Kehitaman bagi piksel ini = 43

- Imej marikh yang dihantar dalam misi *Mariner* 6 adalah imej *grayscale*.
- Nilai darjah kehitaman bagi setiap piksel dihantar ke bumi menggunakan sistem binari.

0	↔	000000
1	↔	000001
2	↔	000010
3	↔	000011
4	↔	000100
5	↔	000101
6	↔	000110
7	↔	000111
8	↔	001000
9	↔	001001
...		
43	↔	101011
...		
63	↔	111111

Seperti rajah di atas, piksel yang mempunyai darjah kehitaman 43 dihantar sebagai 101011

$$43 = 101011$$

- Dalam kes 6 *Mariner*, setiap gambar yang telah dipecahkan kepada 700 x 832 grid.
- Justeru, jika darjah kehitaman setiap grid dikodkan menggunakan 6 digit binari, setiap imej akan mempunyai urutan:

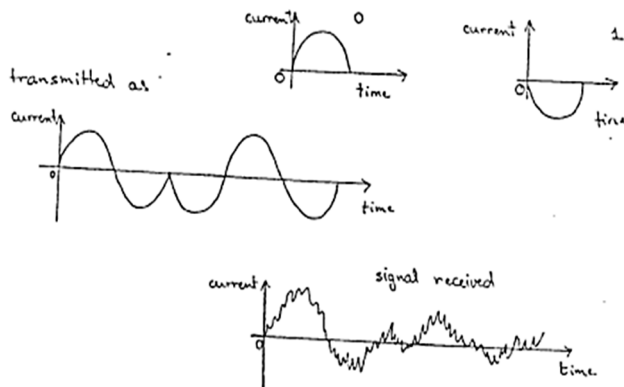
$$6 \times 700 \times 832 = 3\,494\,400 \text{ digit binari}$$

- Namun, darjah kehitaman setiap grid sebenarnya dikodkan menggunakan 32 digit
- Nota Padat MTE3114 – Aplikasi Matematik | 12

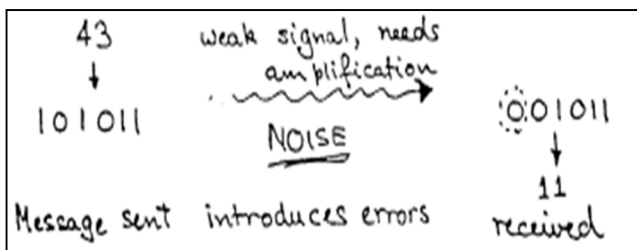
binari untuk mendapatkan imej yang lebih tepat. Justeru, setiap imej akan mempunyai urutan:

$$32 \times 700 \times 832 = 18\,636\,800 \text{ digit binari}$$

- Namun, semasa penghantaran data, berlaku ralat yang menyebabkan imej yang diterima tidak tepat dari imej sebenar.
- Ralat berlaku kerana faktor gangguan, jarak dan saluran penghantaran data.
- Rajah di bawah adalah graf penghantaran data bagi nilai '0' dan '1' bagi setiap darjah kehitaman.



- Ralat yang merupakan gangguan ini dipanggil *noise*.



- Ralat yang merupakan gangguan ini dipanggil *noise*.
- Jika banyak ralat yang berlaku, ianya akan memberi kesan kepada kualiti imej. Justeru, pembetulan kod ralat perlu dilakukan.

Pembetulan Kod Ralat (*error correcting codes*)

- Proses pengekodan mesej biasanya bermula dengan penukaran teks "biasa" kepada turutan nombor sistem binari seperti di bawah:

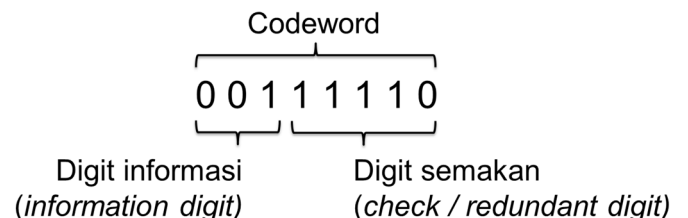
= 00000 A = 00001 B = 00010 C = 00011 D = 00100
 E = 00101 F = 00110 G = 00111 H = 01000 I = 01001
 J = 01010 K = 01011 L = 01100 M = 01101 N = 01110
 O = 01111 P = 10000 Q = 10001 R = 10010 S = 10011
 T = 10100 U = 10101 V = 10110 W = 10111 X = 11000
 Y = 11001 Z = 11010 , = 11011 . = 11100 ? = 11101
 : = 11110 ; = 11111

- Sebagai contoh, setiap mesej dihantar terdiri daripada hanya 3 digit binari. Terdapat 8 mesej, yang mewakili integer 0-7.
- Kita tambahkan 5 angka tambahan pada setiap mesej seperti menjadikannya sebagai *codewords* yang panjangnya 8 digit. (Pada masa ini kita tidak akan cuba untuk menerangkan bagaimana kita memilih digit tambahan.)

0	=	000
1	=	001
2	=	010
3	=	011
4	=	100
5	=	101
6	=	110
7	=	111

000	00000
001	10110
010	10101
011	00011
100	10011
101	00101
110	00110
111	10000

- Bayangkan bahawa mesej yang diterima adalah 00111110
- Codeword* yang terdekat adalah 00110110. Ia berbeza hanya satu tempat – digit yang kelima paling mungkin *codeword* telah dihantar adalah 00110110.
- Ini adalah benar untuk semua kes-kes apabila hanya satu kesilapan yang berlaku – pembetulan ralat kod tunggal (*single error correcting codes*).



Bilangan digit *codeword*, $n = 8$

Bilangan digit informasi, $k = 3$

Bilangan digit semakan, $r = 5$

$$\text{Kadar informasi (Information rate), } R = \frac{k}{n}$$

Kod Pengulangan (*repetition codes*)

- Kod pengulangan adalah satu cara untuk pembetulan ralat.
- Oleh itu, sekiranya satu mesej untuk dikodkan, pengulangan bagi setiap digit akan diulang sebanyak n kali.
- Contoh:

Jika $n = 5$, n ialah pengulangan bagi setiap digit

S	10011	11111	00000	00000	11111	11111
U	10101	11111	00000	11111	00000	11111
S	10011	11111	00000	00000	11111	11111
I	01001	00000	11111	00000	00000	11111
E	00101	00000	00000	11111	00000	11111

- Menyahkan pengulangan
 - Kira bilangan nombor 1 dalam kod pengulangan.

- Jika bilangan nombor 1 ≥ 3 , tuliskan 11111
- Jika bilangan nombor 1 ≤ 2 , tuliskan 00000

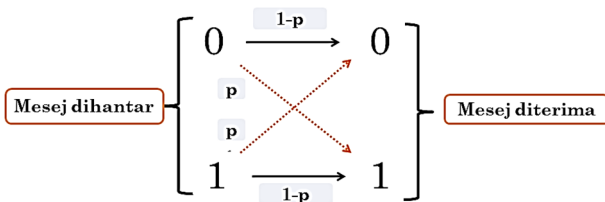
- Contoh:

11011 00110 11000 10000 10111

↓ ↓ ↓ ↓ ↓

1 0 0 0 1

- Simetri saluran binari (*Binary symmetric channel*)



- Jika $p = \frac{1}{100}$, maka kebarangkalian satu digit tunggal yang tidak betul diterima adalah 0.01
- Oleh itu, kebarangkalian satu digit tunggal yang diterima betul adalah 0.99

Mesej asal	Mesej dihantar	Mesej yang mungkin diterima		Mesej yang dinyahkod
0	000	000	001	0
		010	100	
1	111	101	011	1
		110	111	

- Contoh:
 - Anggap mesej yang dihantar adalah 000.
 - Mesej yang mungkin diterima iaitu: 000, 001, 010, 100

$$Pr(000) = 0.99 \times 0.99 \times 0.99 = 0.970299$$

$$Pr(001) = 0.99 \times 0.99 \times 0.01 = 0.009801$$

$$Pr(010) = 0.99 \times 0.01 \times 0.99 = 0.009801$$

$$Pr(100) = 0.01 \times 0.99 \times 0.99 = 0.009801$$

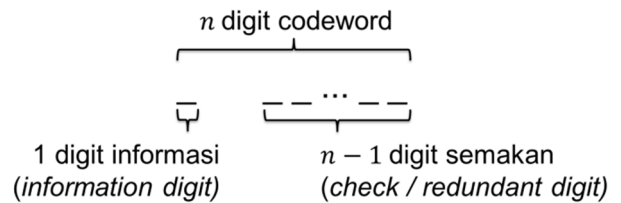
- Oleh itu, kebarangkalian-penyahkodan mesej sebagai 0:

$$Pr(0) = Pr(000) + Pr(001) + Pr(010) + Pr(100)$$

$$Pr(0) = 0.970299 + (3 \times 0.009801)$$

$$Pr(0) = 0.999702$$

- Kadar informasi (*Information rate*) bagi kod pengulangan



$$\text{Kadar informasi, } R = \frac{1}{n}$$

Kod Semakan Pariti Tunggal (*single parity check codes*)

- Jika kod pengulangan mempunyai 1 digit mesej (*message digit*), kod semakan pariti pula hanya mempunyai satu digit semakan (*check digit*).
- Kita boleh mendapatkan digit semakan dengan menambah digit maklumat mod 2.

$$\begin{array}{r|rr} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$

- Contoh:

C \longleftrightarrow 00011

5 digit mesej

Ingat!

Kita boleh mendapatkan **digit semakan** dengan menambah **digit mesej** (mod 2)

- Umumnya codewords terdiri daripada:

$C_1, C_2, C_3, C_4, C_5, C_6$

Dimana, digit mesej = C_1, C_2, C_3, C_4, C_5
digit semakan = C_6

- Mencari digit semakan

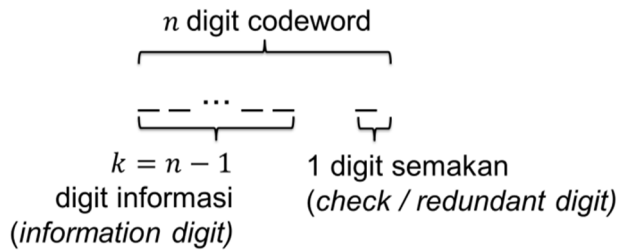
$$\begin{aligned} \text{Digit semakan, } C_6 &= C_1 + C_2 + C_3 + C_4 + C_5 \pmod{2} \\ &= 0 + 0 + 0 + 1 + 1 \pmod{2} \\ &= 0 \pmod{2} \end{aligned}$$

- Codework ditulis sebagai 000110
- Contoh:

J \longleftrightarrow 01010 0 \longrightarrow 010100
 L \longleftrightarrow 01100 0 \longrightarrow 010100
 S \longleftrightarrow 10011 1 \longrightarrow 100111

5 digit mesej 1 digit semakan

- Kadar informasi (*Information rate*) bagi kod pengulangan

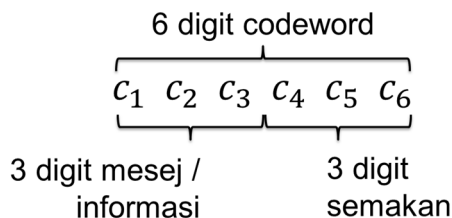


$$\text{Kadar informasi, } R = \frac{k}{n} = \frac{n-1}{n}$$

Kod Pengulangan	Kod Semakan Pariti
Kadar informasi rendah $R = \frac{1}{n}$ Contoh : $R = \frac{1}{6} = 0.16$	Kadar informasi tinggi $R = \frac{n-1}{n}$ Contoh : $R = \frac{6-1}{6} = 0.83$
Baik dalam membetulkan ralat sehingga $\frac{n-1}{2}$ ralat	Hanya mengesan ralat kod, tidak boleh betulkan ralat

Kod Linear

- Diberi 6 digit codeword:



- Pengiraan digit semakan:

Diberi mesej / informasi sebagai $c_1 \ c_2 \ c_3$ maka

$$c_4 = c_1 + c_2 \pmod{2}$$

$$c_5 = c_1 + c_3 \pmod{2}$$

$$c_6 = c_2 + c_3 \pmod{2}$$

untuk mendapatkan codeword

$$\mathbf{C} = [c_1 \ c_2 \ c_3 \ c_4 \ c_5 \ c_6]$$

$$\text{Maka, } \begin{bmatrix} c_4 \\ c_5 \\ c_6 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

- Contoh:

Diberi digit mesej adalah [010], cari digit semakan dan codeword.

$$\begin{bmatrix} c_4 \\ c_5 \\ c_6 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} c_4 \\ c_5 \\ c_6 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Digit semakan = [101]
Codeword = [010101]

Persamaan semakan pariti

- Persamaan semakan pariti:

$$c_1 + c_2 + c_4 = 0 \pmod{2}$$

$$c_1 + c_3 + c_5 = 0 \pmod{2}$$

$$c_2 + c_3 + c_6 = 0 \pmod{2}$$

- Dalam bentuk matrik:

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{H}\mathbf{C}^T = \mathbf{0}$$

Diberi codeword

$$\mathbf{C} = [c_1 \ c_2 \ c_3 \ c_4 \ c_5 \ c_6]$$

Pada saluran dengan kebisingan dan ralat (*noise and error*)

$$\mathbf{E} = [e_1 \ e_2 \ e_3 \ e_4 \ e_5 \ e_6]$$

Menghasilkan perkataan yang diterima (*received word*)

$$\mathbf{R} = [r_1 \ r_2 \ r_3 \ r_4 \ r_5 \ r_6]$$

$$\mathbf{R} = \mathbf{C} + \mathbf{E} \longrightarrow r_i = c_i + e_i$$

- Contoh:

$$\mathbf{C} = [100110], \mathbf{E} = [000101]$$

$$\mathbf{R} = [100110] + [000101]$$

$$= [100011]$$

- Kenalpasti ralat melalui sidrom (*syndrome*):

Bagi perkataan yang diterima $R = [r_1 \ r_2 \ r_3 \ r_4 \ r_5 \ r_6]$, dapat dikenalpasti sindrom (*syndrome*) $s = [s_1 \ s_2 \ s_3]$ bagi R :

$$s_1 = r_1 + r_2 + r_4 \pmod{2}$$

$$s_2 = r_1 + r_3 + r_5 \pmod{2}$$

$$s_3 = r_2 + r_3 + r_6 \pmod{2}$$

$$\text{maka, } \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \end{bmatrix}$$

$$s^T = HR^T.$$

- Contoh:

Diberi codeword = [001011], cari sindrom, s

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- Ingat kembali, *received word* R adalah hasil tambah *codeword* C dan *error* E .

$$R = C + E$$

Jadi, $C = R - E$ dan

$$\begin{aligned} 0 &= HC^T = H(R - E)^T \\ &= H(R^T - E^T) \\ &= HR^T - HE^T \end{aligned}$$

Maka, $s^T = HR^T = HE^T$

- Oleh itu, *received word* R mempunyai *syndrome* yang terdiri daripada *error* E .

- Daripada pembuktian tersebut, jika R dan E mempunyai *syndrome* yang sama, kita boleh menterbalikkannya untuk membentuk persamaan lain:

$$\text{Jika } HR^T = HE^T$$

Maka, $R - E = C$

Slepian's Standard Array

- Boleh digunakan untuk mencari *words* berdasarkan *syndrome* yang telah diberikan.

Syndrome	Words							
000	000000	001011	010101	011110	100110	101101	110011	111000
001	000001	001010	010100	011111	100111	101100	110010	111001
010	000010	001001	010111	011100	100100	101111	110001	111010
011	001000	000011	011101	010110	101110	100101	111011	110000
100	000100	001111	010001	011010	100010	101001	110111	111100
101	010000	011011	000101	001110	110110	111101	100011	101000
110	100000	101011	110101	111110	000110	001101	010011	011000
111	001100	000111	011001	010010	101010	100001	111111	110100

Kod Linear Secara Umum

- Kod Linear/Sekumpulan Kod:

Satu kod yang mana *codeword* adalah set vector C memenuhi satu persamaan $HC^T = 0$ yang mana H ialah semakan-pariti (kod linear dengan 3×6 matrik semakan-pariti).

- Kod Semakan-pariti Tunggal :
Dalam kod semakan-pariti tunggal, digit c_1, c_2, c_3, c_4, c_5 dan c_6 *codeword* $[c_1 \ c_2 \ c_3 \ c_4 \ c_5 \ c_6]$ adalah memenuhi persamaan semakan pariti.
 $c_6 = c_1 + c_2 + c_3 + c_4 + c_5 \pmod{2}$
- Ianya sama juga dengan
 $c_1 + c_2 + c_3 + c_4 + c_5 + c_6 = 0 \pmod{2}$
- Kita boleh menuliskannya sebagai $HC^T = 0$.
 $H = [111111]$.
- Secara umumnya, jika terdapat kod semakan-pariti tunggal yang panjangnya adalah n , kemudian matrik semakan-pariti adalah sepadan dengan matrix kod
 $1 \times n$ matrix, $H = [111 \dots 1]$

Block length

Satu *code* yang mana setiap *codeword* ialah susunan nombor yang tetap, n . Dalam kod linear, *block length* ialah nombor lajur dalam H .

Syndrom

Syndrom, s ialah satu *received word* yang diperoleh daripada $S^T = HR^T$

Cosets

Satu *coset* yang mengandungi semua *word* yang mempunyai *syndrome*.

Weight

Weight dalam *word* ialah bilangan *word*.

Coset leader

Satu *word* yang mengandungi paling kurang *weight*.

Menyahkod *word* yang diterima

1. Kira *syndrom*, s .
2. Cari *coset leader*, E
3. Kira $C = R - E$

- Kegunaan kod linear

Pengekoden yang lebih cepat dan kurang ruang penyimpanan.

Pengiraan jelas dan mudah.

Pola kesilapan mudah diperjelaskan

Jarak minimal; mudah untuk mengira jika *Codeword* adalah satu kod linear.

Kod linear mempunyai spesifikasi yang mudah untuk menentukan kod yang bukan linear. Biasanya semua *codewords* telah disenaraikan.

Kod Hamming (Hamming codes)

- Kod Hamming telah dipelopori oleh Richard Hamming pada tahun 1950.
- Kod ini adalah kod linear yang membetulkan satu kesalahan

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

$$\text{dan } \mathbf{s}^T = \mathcal{H} \mathbf{R}^T = \mathcal{H} \mathbf{E}^T.$$

- Jika ralat, $\mathbf{E} = [e_1 \ e_2 \ e_3 \ e_4 \ e_5 \ e_6]$, ianya boleh ditulis dalam persamaan:

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = e_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + e_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + e_3 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + e_4 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + e_5 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + e_6 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

- Jika mana-mana kolum \mathcal{H} adalah 0, maka kesalahan pada posisi tersebut tidak dapat dikesan
- Jika ada antara 2 kolum \mathcal{H} adalah sama, kita tidak dapat membezakan antara 2 kesilapan itu.
- Maka kod linear hanya boleh membetulkan satu kesilapan jika kolum-kolum \mathcal{H} berbeza dan $\neq 0$.
- Menyahkod perkataan:
 - Untuk nyahkod perkataan yang diterima, \mathbf{R} , maka sindrom, \mathbf{s} dikira.

- Jika $\mathbf{s} = 0$, kesilapan tidak berlaku.
- Jika $\mathbf{s} \neq 0$ dan sama dengan lajur \mathcal{H} , maka satu kesilapan berlaku pada kedudukan itu.
- Jika $\mathbf{s} \neq 0$ dan tidak sama dengan lajur mana-mana \mathcal{H} , prosedur menyahkod gagal.

- Kegagalan dan kesilapan menyahkod hanya boleh berlaku apabila dua atau lebih saluran kesilapan berlaku.
- Contoh:

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

Contoh 1

Diberi perkataan diterima $\mathbf{R} = [101000101]$, nyahkodkan \mathbf{R} .

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

$$= [1100]$$

Maka sindrom, $\mathbf{s} = [1100]$

Jadi, \mathbf{s}^T ada pada kolum kelima \mathcal{H} , maka $\mathbf{E} = [000010000]$

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

$$\mathbf{E} = 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0$$

$$\text{Justeru, } \mathbf{C} = \mathbf{R} - \mathbf{E} = [101010101]$$

Proses nyahkod berjaya.

Contoh 2

Diberi $\mathbf{R} = [101001101]$, nyahkodkan \mathbf{R} .

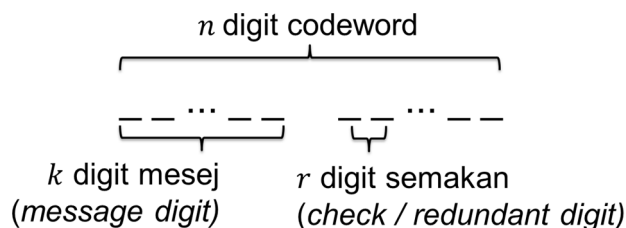
$$\mathbf{s} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$= [0100]$$

Maka sindrom, $\mathbf{s} = [0100]$

Namun, \mathbf{s}^T tidak terdapat dalam \mathcal{H} , maka proses nyahkod gagal.

- Kadar informasi:



$$\text{Kadar informasi, } R = \frac{2^r - 1 - r}{2^r - 1} = 1 - \frac{r}{2^r - 1}$$

Kriptografi Kekunci Awam

Pengenalan kepada Kriptografi

- Kriptografi - berasal daripada perkataan Greek *kryptós*, "tersembunyi", dan *gráphein*, "untuk menulis".
- Kriptografi, secara umum adalah ilmu dan seni untuk menjaga kerahsiaan berita (Bruce Schneier, n.d)
- Istilah-istilah kriptografi:

Plaintext

mesej yang dapat dibaca secara kasar.

Ciphertext

mesej yang tidak dapat dibaca dan difahami secara kasar.

Enkripsi atau penyulitan

Proses yang dilakukan untuk mengubah *plaintext* ke dalam *ciphertext*

Dekripsi atau penyahsulitan

Proses digunakan untuk membuat *ciphertext* kembali menjadi *plaintext*.

Cryptographer

Pakar di dalam bidang kriptografi.

Cryptanalysis

Seni dan ilmu untuk memecahkan *ciphertext* menjadi *plaintext* tanpa melalui cara yang seharusnya (dekripsi).

Cryptanalyst

Orang yang melakukan *cryptanalysis*

- Sigh (2002)

"Jika dua orang hendak bertukar mesej rahsia" menggunakan telefon pengirim hendaklah mengekod mesej tersebut. Untuk mengekod mesej pengirim hendaklah menggunakan kunci rahsia. Kunci rahsia akan digunakan oleh penerima mesej untuk mendikod mesej tersebut. Dipendekkan, sebelum dua orang boleh bertukar mesej rahsia, mereka terlebih dahulu telah berkongsi rahsia (kunci rahsia).

Proses pendaraban dan pemfaktoran nombor sebagai asas Kriptografi

Darabkan
222 156 x 4

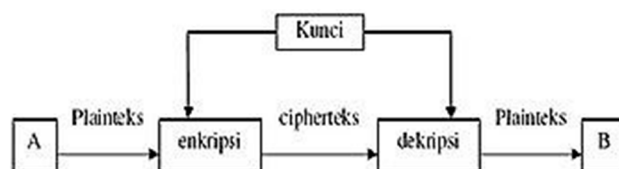
Faktorkan
888 624

- Proses pengekodan diibaratkan sebagai proses mendarab di mana ianya menggunakan suatu **kekunci iaitu 4**. Ianya lebih mudah dilakukan.
- Namun, proses nyahkod semula mesej adalah jauh lebih kompleks dan mengambil masa yang lama. Ia boleh diibaratkan sebagai proses pemfaktoran. **Faktor bagi 888 624 bukan semata-mata kekunci 4**.

Kriptografi Simetri VS Asimetri

Kriptografi Simetri	Kriptografi Asimetri
<ul style="list-style-type: none"> Penggunaan kekunci awam dan kekunci peribadi yang sama untuk tujuan enkripsi dan dekripsi maklumat 	<ul style="list-style-type: none"> Penggunaan kekunci awam dan peribadi yang berbeza untuk tujuan enkripsi dan dekripsi maklumat

- Kriptografi simetri



- Kriptografi asimetri



Tujuan Kriptografi

Confidality (kerahsiaan) iaitu pesanan/mesej yang dikirimkan tetap rahsia dan tidak diketahui oleh pihak lain.

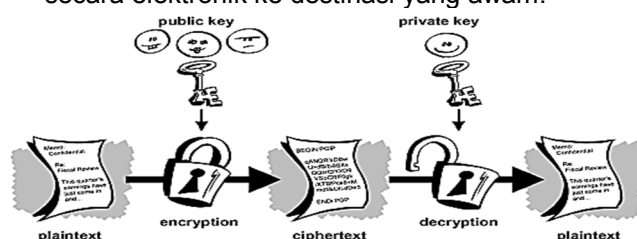
Data integrity (keutuhan data) iaitu mampu mengenali kewujudan manipulasi data yang tidak sah (oleh pihak lain).

Authentication (ketulenan) iaitu berhubungan dengan identifikasi/ketulenan data.

Non-repudiation (anti-penyangkalan) iaitu mencegah suatu pihak untuk menyangkal tindakan yang dilakukan sebelumnya.

Kriptografi Kekunci Awam

- Kriptografi Kekunci Awam membolehkan orang ramai untuk menghantar mesej dengan selamat secara elektronik ke destinasi yang awam.



Penggunaan Modular Aritmetik dalam Kriptografi Kekunci Awam

- Juga dipanggil sebagai jam arithmetic dihasilkan oleh K.F.Gauss (1777-1855)
- Untuk sebarang nombor natural arithmetic modular n adalah berdasarkan mengelas set integer tersebut mengikut baki yang diperolehi apabila integer dibahagikan dengan n .
- Modular arithmetik telah digunakan dalam banyak sistem kod untuk menyamarkan maklumat dalam pelbagai cara.
- Kriptografi menggunakan arithmetik modular 11:

P	0	1	2	3	4	5	6	7	8	9	10
$C = P^3$	0	1	8	27	64	125	216	343	512	729	1000
$C = P^3 \text{ modulo } 11$	0	1	8	5	9	4	7	2	6	3	10

- Penambahan dan pendaraban modular n bagi set $Z_n = \{0, 1, 2, \dots, n-1\}$:

+	0	1	2	3	x	0	1	2	3
0	0	1	2	3	0	0	0	0	0
1	1	2	3	0	1	0	1	2	3
2	2	3	0	1	2	0	2	0	2
3	3	0	1	2	3	0	3	2	1

- Dua integer a dan b adalah modulo n yang kongruen apabila $a - b$ adalah gandaan n dan ditulis sebagai:

$$a \equiv b \pmod{n}$$

- Hasil darab modulo:

Cari $X = 36 * 53 * 91 * 17 * 22 \pmod{29}$

Maka,

$$\begin{aligned} 36 &\equiv 7 \pmod{29} \\ 53 &\equiv 24 \pmod{29} \\ 91 &\equiv 4 \pmod{29} \\ 17 &\equiv 17 \pmod{29} \\ 22 &\equiv 22 \pmod{29} \end{aligned}$$

Justeru, ianya boleh ditulis semula:

$$\begin{aligned} X &= 36 * 53 * 91 * 17 * 22 \pmod{29} \\ &= 7 * 24 * 4 * 17 * 22 \pmod{29} \\ &= 168 * 68 * 22 \pmod{29} \\ &= 23 * 10 * 22 \pmod{29} \\ &= 230 * 22 \pmod{29} \\ &= 27 * 22 \pmod{29} \\ &= 594 \pmod{29} \\ &= 14. \end{aligned}$$

Semakan boleh dibuat pada kalkulator saintifik
 $36 * 53 * 91 * 17 * 22 = 64\,936\,872$ dan
 $64\,936\,872 \pmod{29} = 14$.

- Pengiraan modulo melibatkan kuasa:

Cari $X = 11^{43} \pmod{13}$

$$\begin{aligned} 11^2 \pmod{13} &= 121 \pmod{13} = 4 \\ 11^4 \pmod{13} &= 4^2 \pmod{13} = 16 \pmod{13} = 3 \\ 11^8 \pmod{13} &= 3^2 \pmod{13} = 9 \pmod{13} = 9 \\ 11^{16} \pmod{13} &= 9^2 \pmod{13} = 81 \pmod{13} = 3 \\ 11^{32} \pmod{13} &= 3^2 \pmod{13} = 9 \pmod{13} = 9. \end{aligned}$$

Proses tidak diteruskan memandangkan

$$11^{64} > 11^{43}$$

Namun,

$$\begin{aligned} 11^{43} &= 11^{32} * 11^{11} \\ &= 11^{32} * 11^8 * 11^3 \\ &= 11^{32} * 11^8 * 11^2 * 11 \end{aligned}$$

Maka

$$\begin{aligned} 11^{43} \pmod{13} &= 11^{32} * 11^8 * 11^2 * 11 \pmod{13} \\ &= 9 * 9 * 4 * 11 \pmod{13} \\ &= 81 * 44 \pmod{13} \\ &= 81 * 44 \pmod{13} \\ &= 3 * 5 \pmod{13} \\ &= 15 \pmod{13} \\ &= 2 \end{aligned}$$

Theorem kecil Fermat

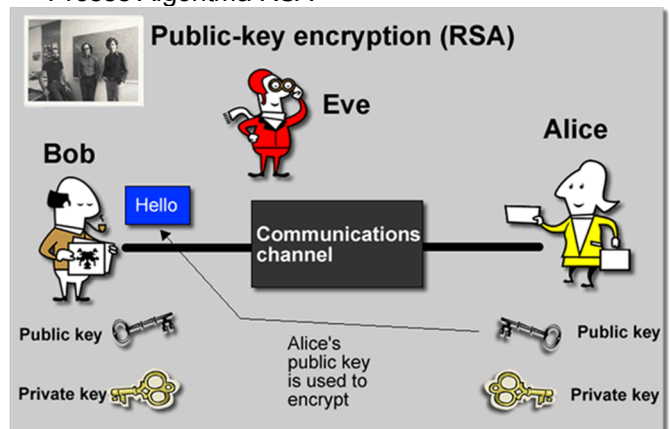
- Sering digunakan dalam teori nombor untuk menguji nombor perdana yang besar.

- $n = p \times q$ dimana p dan q adalah nombor perdana
- Jadi, $(n) = (p-1)(q-1)$
- Untuk setiap integer a hendaklah merupakan nombor yang tidak boleh dibahagikan dengan p dan q .
- Oleh itu, $a^{\varphi(n)}$

Algoritma Rivest-Shamir-Adleman (RSA)

- RSA dicipta pada tahun 1978 dan dipatenkan pada 1983.
- RSA adalah singkatan dari nama perintis-perintis iaitu Ron Rivest, Adi Shamir, dan Leonard Adleman dari Massachusetts Institute of Technology.
- RSA adalah algoritma yang paling efektif kerana memiliki kecepatan yang lebih lambat berbanding algoritma simetrik lainnya.

- Proses Algoritma RSA



- Misalnya, Bob ingin mengirim sebuah pesan peribadi (*private message*) melalui media transmisi yang tidak selamat (*insecure*).
- Sebagai contoh, Bob ingin memesan n (*plaintext*) kepada Alice dan Bob perlu menukar pesanan n kepada c (*chiphertext*) melalui sistem RSA supaya pesanan dapat dirahsiakan daripada orang lain.
- Kemudian Alice perlu menukar c (*chiphertext*) kepada n (*plaintext*) kembali untuk mendapat pesanan tersebut.
- Mereka melakukan langkah-langkah berikut untuk membuat pasangan kunci awam dan kunci peribadi.

- Pengiraan enkripsi RSA:
 - Parameter yang digunakan di sini berupa bilangan kecil.

Parameter	
$p = 61$	bilangan prima pertama (harus dijaga kerahasiannya atau dihapus secara hati-hati)
$q = 53$	bilangan prima kedua (harus dijaga kerahasiannya atau dihapus secara hati-hati)
$N = pq = 3233$	modulus (diberikan kepada awam)
$e = 17$	eksponen publik (diberikan kepada awam)
$d = 2753$	eksponen pribadi (dijaga kerahasiannya)

- Kunci awam yang digunakan adalah (e, N) .
- Kunci peribadi yang digunakan adalah d .

Fungsi pada **enkripsi** ialah:

$$\begin{aligned} \text{encrypt}(n) &= n^e \bmod N \\ &= n^{17} \bmod 3233 \end{aligned}$$

dimana n adalah teks biasa (*plaintext*)

Fungsi pada **dekripsi** ialah:

$$\begin{aligned} \text{decrypt}(c) &= c^d \bmod N \\ &= n^{2753} \bmod 3233 \end{aligned}$$

dimana c adalah teks *cipher* (*ciphertext*)

- Contoh pengiraan:

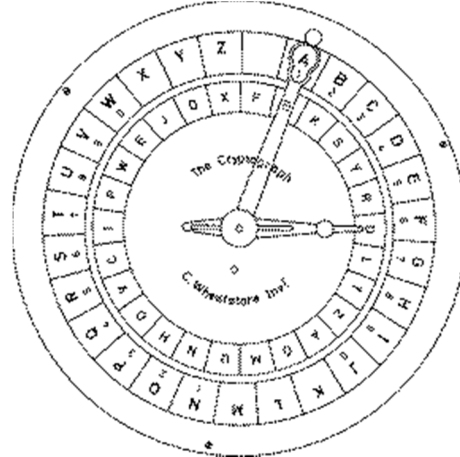
$$\begin{aligned} \text{encrypt}(123) &= 123^{17} \bmod 3233 \\ &= 855 \\ \text{decrypt}(855) &= 855^{2753} \bmod 3233 \\ &= 123 \end{aligned}$$

Alat-alat Kriptografi

- Silinder Jefferson (1790an)



- Cakera Wheatstone (mula direka pada 1817, terhasil pada 1860an).



- Mesin Rotor Enigma (Perang Dunia Ke-2)



Rotor:

