

Week 1: Lecture Notes

Topics:

Introduction to Cryptography.

Classical Cryptosystem.

Cryptanalysis on substitution cipher.

Playfair Cipher.

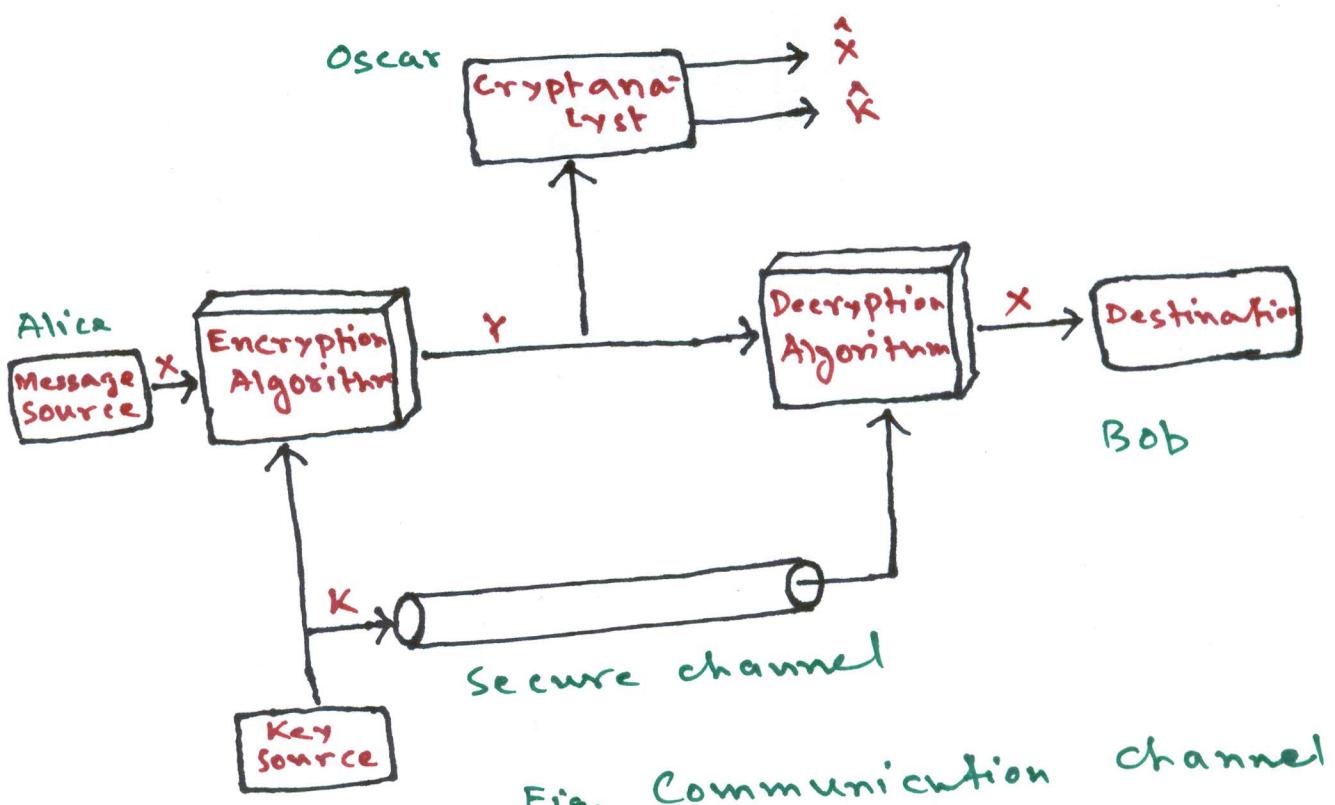
Block Cipher.

Introduction to Cryptography :

- Cryptography is the science or art of secret writing.
- The fundamental objective of cryptography is to enable two people (Alice and Bob) to communicate over an insecure channel in such a way that an opponent (Oscar) can not understand what is being said.
- **Plaintext:** the information that

Alice wants to send to Bob.

- Alice encrypts the plaintext, using a predetermined key, and send the resulting ciphertext to Bob over the public channel.
- Upon receiving the ciphertext
 - Oscar can not determine what the plaintext was.
 - But Bob knows the encryption key, can decrypt the ciphertext and get the plaintext.



- **Cryptology**: two competing areas:
 - Cryptography — Art of converting information to a form that will be unintelligible to an unintended recipient, carried out by Cryptographer.
 - Cryptanalysis — Art of breaking cryptographic systems, carried out by Cryptanalyst.
- Two main types of cryptography in use today:
 - Symmetric or secret key cryptography.
 - Asymmetric or public key cryptography.

Conventional Encryption :

- Also termed single key or symmetric encryption .

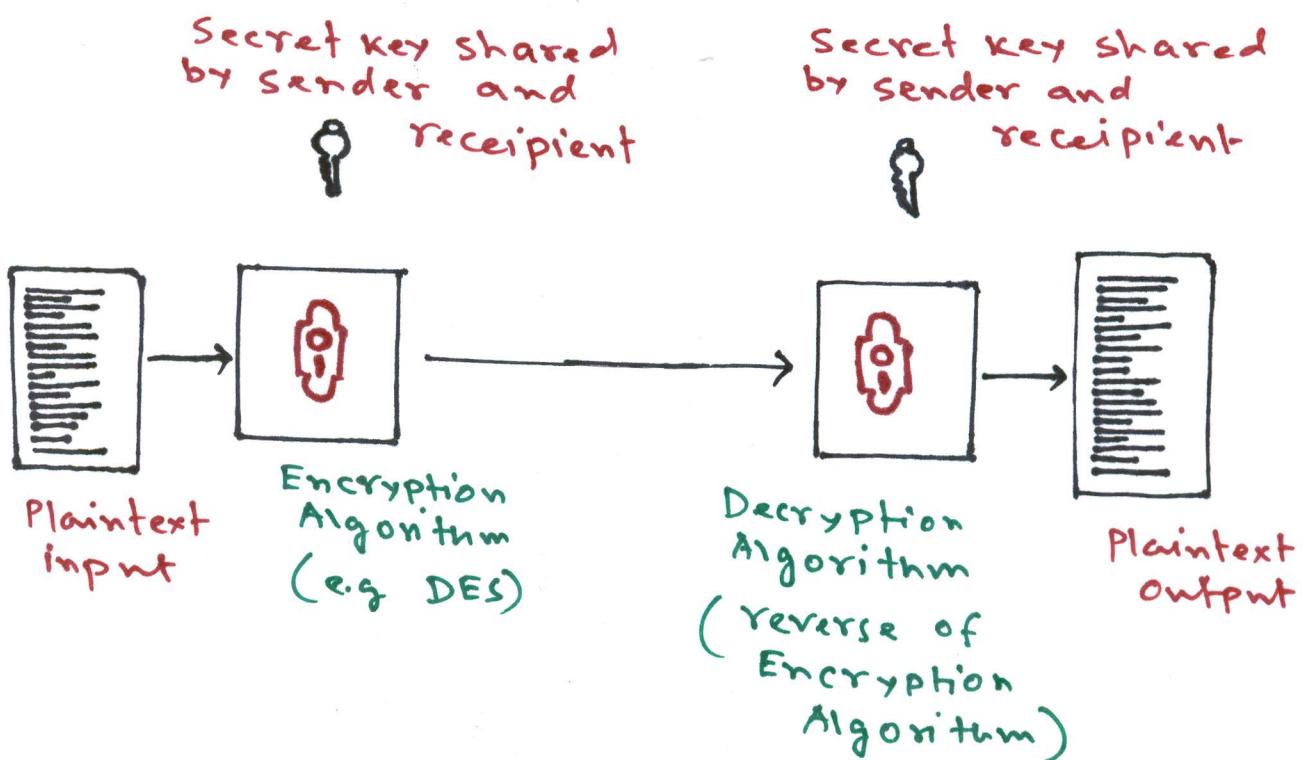


Fig: Simplified model of conventional encryption .

Cryptosystem

- Cryptosystem is a five tuple (P, C, X, E, D)

- Plaintext Space (P): set of all possible plaintext.
 - Ciphertext Space (C): set of all possible ciphertext.
 - Key Space (K): set of all possible keys.
 - E : set of all encryption rules
 - D : set of all possible decryption rules.
- For each $k \in K$, there is an encryption rule $e_k \in E$ and corresponding decryption rule $d_k \in D$ such that
- $$d_k(e_k(x)) = x \text{ for every plaintext } x \in P.$$
- A practical cryptosystem should satisfy
- Each encryption function e_k

and each decryption function d_k should be efficiently computable.

- An opponent, upon seeing the ciphertext string y , should be unable to determine the key k that was used or the plaintext string x .
- The process of attempting to compute the key k , given a string of ciphertext y , is called cryptanalysis
 - If the opponent can determine k , then he can decrypt y just as Bob would, using d_k .
 - Determining k should be as difficult as determining the plaintext string x , given the ciphertext string y

Classical Cryptosystem

Shift cipher

- $Z_{26} = \{0, 1, 2, \dots, 24, 25\}$

- $P = C = K = Z_{26}$

- For $k \in K$,

$$e_k(x) = (x+k) \bmod 26 \text{ for } x \in P$$

$$d_k(y) = (y-k) \bmod 26 \text{ for } y \in C$$

- Caesar Cipher is a particular case
 $(k=3)$

Example

- Plaintext is ordinary English text.
- Correspondence between alphabetic characters and integer : $A=0, B=1, \dots, Y=24, Z=25$

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Q	R	S	T	U	V	W	X	Y	Z
16	17	18	19	20	21	22	23	24	25

Encryption

- Key K = 11
 - Plaintext is "We will meet at midnight"
 - corresponding sequence of integers:
22, 4 22 8 11 11 12 4 4 19 0 19 12 8 3 13 8 6 7
19
 - We add 11 (key) to each value (rounding modulo 26):

7 15 7 19 22 22 23 15 15 4 11 4 23 19 14 24
19 17 18 9

- Convert the sequence of integers to alphabetic characters.
Ciphertext is:
“ HPHTWWXPPPELEXTOTYTRSE ”

Decryption

- ciphertext : "HPHTWWXPPELEXTOYTRSE"
- convert the ciphertext to sequence of integers :

7 15 7 19 22 22 23 15 15 4 11 9 23 19 19
24 19 17 18 4

- Subtract 11 from each value (reducing modulo 26) :

22 4 22 8 11 11 12 4 4 19 0 19 12 8 3
13 8 6 7 19

- convert the sequence of integers to alphabetic characters :

Plaintext is "we will meet at midnight"

Caesar Cipher

- Caesar cipher is the earliest known (and the simplest). It involves

replacing each letter of the alphabet with the letter standing three places further down. This is then wrapped around on itself when the end is reached. For example

Key :

$K=3$

Plaintext : meetmeaftertheparty

Ciphertext : PHHWPHDIWHVWKHSUWB

Shift cipher is not secure

- Brute-force cryptanalysis easily performed on the shift cipher by trying all 25 possible keys.
- Given a ciphertext string, Oscar successively try the decryption process with $K=0, 1, 2$ etc. until get a meaningful text.

- Ciphertext: JBCRC~~L~~CRWCRVNBJENBW
RWN
 - $k=0 \rightarrow$ jberelqrwervnbjenbwrrwn
 - $k=1 \rightarrow$ iabqbkpqr**b**qumaidmarqvm
 - $k=2 \rightarrow$ hzapajopuaptlzhclzupul
 - $k=3 \rightarrow$ gyzozi~~not~~oskygbkytotk
 - $k=4 \rightarrow$ fxyhymnsynrjxfajxsnsj
 - $k=5 \rightarrow$ ewxm~~xg~~fmrxmqiweziwr*mri*
 - $k=6 \rightarrow$ dvwlwfklqwlp*h*vdyhvqlqh
 - $k=7 \rightarrow$ cuvkvejkprkojucxjupkpg
 - $k=8 \rightarrow$ btujudijoujnf*t*bwftojof.
 - $k=9 \rightarrow$ astitchintimesavesnine
- The key $k=9$.

Substitution Cipher

- $P = C =$ set of 26-letter English alphabet
- $$P = \{a, b, c, \dots, y, z\}$$
- $$C = \{A, B, C, \dots, Y, Z\}.$$

- $X = \text{set of all possible permutations of 26 alphabet characters.}$
- For each permutation $\phi \in X$
 $e_\phi(x) = \phi(x)$ for $x \in P$
 $d_\phi(y) = \phi^{-1}(y)$ for $y \in C$,
 where ϕ^{-1} is the inverse permutation of ϕ .
- Encryption function is the permutation ϕ :

a	b	c	d	e	f	g	h	i	j	k	l	m	n
X	N	Y	A	H	P	O	G	Z	Q	W	B	T	S

o	p	q	r	s	t	u	v	w	x	y	z		
F	L	R	C	V	M	U	E	K	J	D	I		

- Decryption function is the inverse permutation ϕ^{-1}

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
d	l	r	y	v	o	h	e	z	x	w	p	t	b	g

P	Q	R	S	T	U	V	W	X	Y	Z				
f	j	q	u	m	u	s	k	a	c	i				

- Key : $K = \emptyset$

- ciphertext :

MGIZVYZLGHCMHJMYXSNHAYCDL
MHA

- Find the plain text ???

Vigenere Cipher

- Polyalphabetic cipher : use different monoalphabetic substitutions while moving through the plaintext.
- Let m be a positive integer
- $\mathcal{P} = \mathcal{C} = \mathcal{X} = (\mathbb{Z}_{26})^m$
- For $K = (K_1, K_2, \dots, K_m) \in \mathcal{K}^m$

$$e_K(x_1, \dots, x_m) = (x_1 + K_1, \dots, x_m + K_m)$$

$$d_K(y_1, \dots, y_m) = (y_1 - K_1, \dots, y_m - K_m)$$
- All above operations are performed in \mathbb{Z}_{26} .

- Example

- Correspondence between alphabetic characters and integers:

$$A=0, B=1, \dots, Y=24, Z=25$$

- $m = 6$
- Keyword is "CIPHER", this corresponds to numerical equivalent $k = (2, 8, 15, 7, 4, 17)$
- Plaintext: "this cryptosystem is not secure"
- Encryption: add modulo 26

19	7	8	18	2	17	24	15	19	14	18	24
2	8	15	7	4	17	2	8	15	7	4	17
21	15	23	25	6	8	0	23	8	21	22	15
18	19	4	12	8	18	13	14	19	18	9	2
2	8	15	7	4	17	2	8	15	7	4	17
20	1	19	19	12	9	15	22	8	25	8	19
20	17	9									
2	8	15									
22	25	19									

- ciphertext: "VPX~~Z~~GIAXIVWPUBTTM~~J~~PWIZ ITWZT"

- Transposition techniques: so far all the ciphers we have looked at involved only substitution. A very different kind of mapping is achieved using transposition.
- In its simplest form, the rail fence technique involves writing down the plaintext as a sequence of columns and the ciphertext is read off as a sequence of rows. For example, if we use a rail fence of depth 2 with the plaintext 'meet me after the party is over' we get:

m	e	m	a	t	r	h	p	r	y	s	v	r
e	t	e	f	e	t	r	a	t	i	o	e	

- ciphertext is 'mematrhprysvrete feteatiae' which is simply the first row concatenated with the second.

Transposition/Permutation Cipher

- Let m be a positive integer
- $P = C = (\mathbb{Z}_{26})^m$
- $X =$ set of all possible permutations of $\{1, 2, \dots, m\}$
- For each permutation $\pi \in X$
 $e_\pi(x_1, \dots, x_m) = (x_{\pi(1)}, \dots, x_{\pi(m)})$
 $d_\pi(y_1, \dots, y_m) = (y_{\pi^{-1}(1)}, \dots, y_{\pi^{-1}(m)})$
- π' being the inverse permutation of π

Example :

- $m = 6$
- Key is the following permutation π :

x	1	2	3	4	5	6
$\pi(x)$	3	5	1	6	4	2

- inverse permutation π^{-1}

x	1	2	3	4	5	6
$\pi^{-1}(x)$	3	6	1	5	2	4

- Plaintext: "defendthehilltopatsunset"
- partition the plaintext into group of six letters:
 defend | thehil | ltopat | sunset
- rearrange according to π :
 fnddee | eitlhh | oaltp | nestsu
- Ciphertext: " FNDDEEEITLHHOALTPT
NESTSU "
- Decryption can be done using π^{-1}

Frequency Analysis

- Suppose we have a long ciphertext, the challenge is to decipher it.
- Let us know the text is in English and has been encrypted using a monoalphabetic substitution cipher.

- searching all possible keys is impractical as the key space size is 26!
- In English, e is the most common letter, followed by t, then a, and so on as shown in the figure

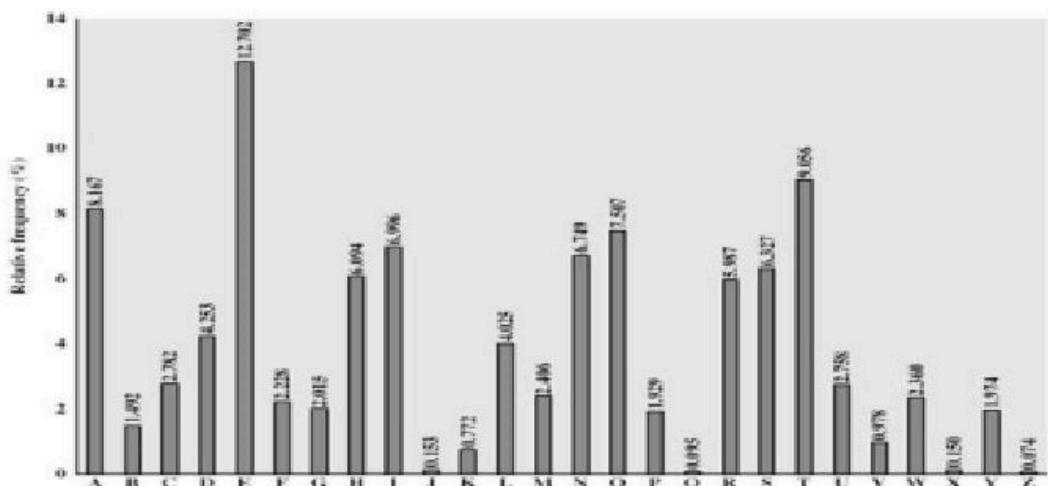


Figure 3: Relative Frequency of letters in the English Language.

- examine the ciphertext in question, and work out the frequency of each letter.
- if most common letter in the ciphertext is, for example, J then it would

seems likely that this is a substitution for e

- if the second most common letter in the ciphertext is P, then this is probably a substitution for t, and so on.
- however, regularities of the language may be exploited, e.g., relative frequency.
- frequency analysis requires logical thinking, intuition, flexibility and guesswork.

Playfair Cipher

- Use the key word CHARLES (Charles Wheatstone invented the cipher)
- Draw up a 5×5 matrix with the keyword first removing any repeating letters as follows:

c	h	a	r	t
e	s	b	d	f
g	i/j	k	m	n
o	p	q	t	u
v	w	x	y	z

- Plaintext: "meet me at the bridge"
 - split the sentence into digrams removing spaces, 'x' used to make even number of letters:
me et me at th eb ri dg ex
- — Repeating plaintext letters that are in the same pair are separated with a filler letter such as 'x'
'ballon' would be treated as
ba lx lo on
- Two plaintext letters in the same row are each replaced by the letter to the right, with the first element of the row

circularly following the last.

eb is replaced by sd

ng is replaced by gi (or gi as preferred)

- Two plaintext letters that fall in the same column are each replaced by the letter beneath, with the top element of the column circularly following the last.

dt would be replaced by my

ty would be replaced by yr

- Otherwise, each plaintext letter in a pair is replaced by the letter that lies in its own row and the column occupied by the other plaintext letter.

me becomes gd

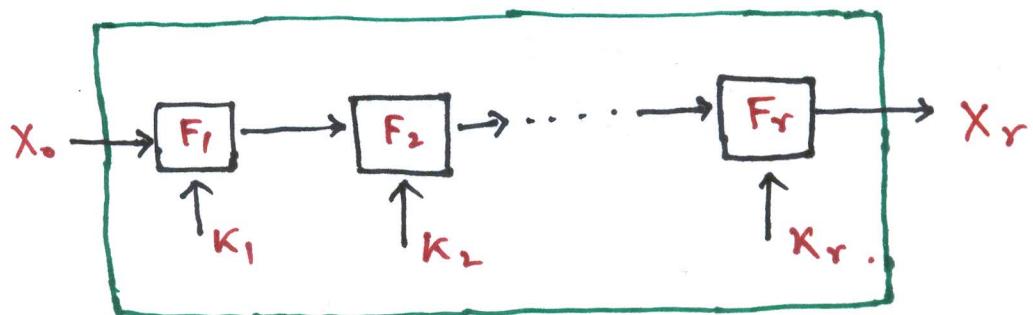
- Ciphertext therefore is:

"gd do gd rq pr sd hm em bv"

Block cipher

- Divided message (plaintext) into fixed sized blocks M_1, \dots, M_i, \dots
- Encrypt each message block separately to obtain $C_i = E_K(M_i)$ where
 - K is the secret symmetric key
 - $E_K()$ is the encryption function
- The cipher is $C_1, C_2, \dots, C_i, \dots$
- Decrypt each cipher block separately to obtain $M_i = D_K(C_i)$ where $D_K()$ is the decryption function.

r-round block cipher



- K_1, K_2, \dots, K_r is the list of round keys derived from the secret key using a publicly known key scheduling algorithm

- $x_i = F_i(x_{i-1}, k_i)$

Encryption

- Let X be plaintext.
- Let F be the round function for all the round.
- The encryption operation is carried out as follows:

$$\begin{aligned}
 x_0 &\leftarrow X \\
 x_1 &\leftarrow F(x_0, k_1) \\
 x_2 &\leftarrow F(x_1, k_2) \\
 &\vdots \\
 x_r &\leftarrow F(x_{r-1}, k_r)
 \end{aligned}$$

- Ciphertext $Y = x_r$.

Decryption

- F is injective function if its second argument is fixed
- There exists a function F' such that $F'(F(x, Y), Y) = x$. Then the decryption operation is carried out as follows:

$$\begin{aligned}
 x_r &\leftarrow Y \\
 x_{r-1} &\leftarrow F'(x_r, k_r) \\
 &\vdots \\
 x_1 &\leftarrow F'(x_2, k_2) \\
 x_0 &\leftarrow F'(x_1, k_1) \\
 x &\leftarrow x_0.
 \end{aligned}$$

Substitution-Permutation Network

- plaintext: $l m$ -bit binary string,
 $X = (x_1, x_2, \dots, x_m)$
- We can regard X as the concatenation
of m l bit substring:
 $X = X_{(1)} || X_{(2)} || \dots || X_{(m)}$ and for
 $1 \leq i \leq m$, we have that
 $X_{(i)} = (X_{(i-1)l+1}, \dots, X_{il})$
- S-box is a permutation $\pi_S : \{0,1\}^l \rightarrow \{0,1\}^l$
- $\pi_P : \{1, 2, 3, \dots, lm\} \rightarrow \{1, 2, 3, \dots, lm\}$
- k_1, k_2, \dots, k_{r+1} is the list of round
keys derived from the secret key
 k .

- The encryption algorithm is as follows:

Algorithm SPN

input : $x, \pi_s, \pi_p, (k_1, \dots, k_{r+1})$

output : Y

1. for $\omega^0 \leftarrow x$
2. for $i \leftarrow 1$ to $r-1$ do
3. $u^i \leftarrow \omega^{i-1} \oplus k_i$
4. for $j \leftarrow 1$ to m
5. do $v_{(ij)}^i \leftarrow \pi_s(u_{(ij)}^i)$
6. $\omega^i \leftarrow (v_{\pi_p(1)}, \dots, v_{\pi_p(m)}^i)$
7. end do
8. $u^r \leftarrow \omega^{r-1} \oplus k_r$.
9. for $j \leftarrow 1$ to m
10. do $v_{(jj)}^r \leftarrow \pi_s(u_{(jj)}^r)$
11. $Y \leftarrow v^r \oplus k_{r+1}$

- Let $l=m=r=4$ and π_s be defined as follows, where the input (i.e z) and the output (i.e $\pi_s(z)$) are written in hexadecimal notation,
 $(0=(0,0,0,0), 1=(0,0,0,1), \dots, 9=(1,0,0,1), A=(1,0,1,0), \dots, F=(1,1,1,1))$

z	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
$\pi_s(z)$	E	4	D	1	2	F	B	8	3	A	6	C	5	9	0	7

- Let π_p be defined as follows:

z	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\pi_p(z)$	1	5	9	13	2	6	10	14	3	7	11	15	4	8	12	16

- Key Scheduling Algorithm: $K = (K_1, \dots, K_{32})$.

For $1 \leq i \leq 5$, define K_i to consist of 16 consecutive bits of K , beginning with K_{4i-3}

- Example: Suppose the key is

$$K = 0011 \ 1010 \ 1001 \ 0100 \ 1101 \ 0110 \\ 0011 \ 1111$$

- The round keys are as follows:

$$K_1 = 0011 \ 1010 \ 1001 \ 0100$$

$$K_2 = 1010 \ 1001 \ 0100 \ 1101$$

$$K_3 = 1001 \ 0100 \ 1101 \ 0110$$

$$K_4 = 0100 \ 1101 \ 0110 \ 0011$$

$$K_5 = 1101 \ 0110 \ 0011 \ 1111$$

- Suppose that the plaintext is

$$X = 0010 \quad 0110 \quad 1011 \quad 0111$$

- Then the encryption of X proceeds as follows:

$$\omega^0 = 0010 \quad 0110 \quad 1011 \quad 0111$$

$$K_1 = 0011 \quad 1010 \quad 1001 \quad 0100$$

$$u' = 0001 \quad 1100 \quad 0010 \quad 0011$$

$$v^1 = 0100 \quad 0101 \quad 1101 \quad 0001$$

$$\omega^1 = 0010 \quad 1110 \quad 0000 \quad 0111$$

$$K_2 = 1010 \quad 1001 \quad 0100 \quad 1101$$

$$u^2 = 1000 \quad 0111 \quad 0100 \quad 1010$$

$$v^2 = 0011 \quad 1000 \quad 0010 \quad 0110$$

$$\omega^2 = 0100 \quad 0001 \quad 1011 \quad 1000$$

$$K_3 = 1001 \quad 0100 \quad 1101 \quad 0110$$

$$u^3 = 1101 \quad 0101 \quad 0110 \quad 1110$$

$$v^3 = 1001 \quad 1111 \quad 1011 \quad 0000$$

$$\omega^3 = 1110 \quad 0100 \quad 0110 \quad 1110$$

$$K_4 = 0100 \quad 1101 \quad 0110 \quad 0011$$

$$u^4 = 1010 \quad 1001 \quad 0000 \quad 1101$$

$$v^4 = 0110 \quad 1010 \quad 1110 \quad 1001$$

$$K_5 = 1101 \quad 0110 \quad 0011 \quad 1111$$

$$Y = 1011 \quad 1100 \quad 1101 \quad 0110$$

- Y is the ciphertext.