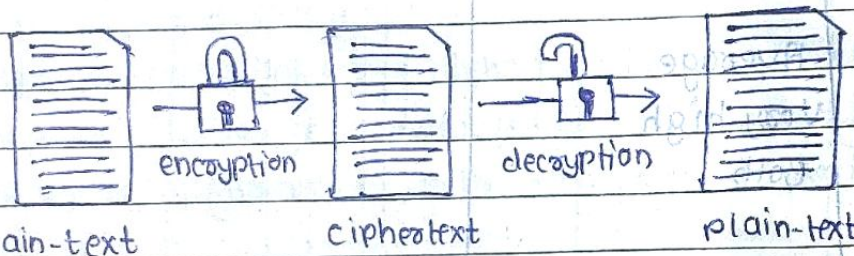
Marks - 16

3) Cryptography

Page No. 62
Date

3.1 Introduction: Cryptography is used to hide information. It is not only used by spies but for phone, fax, email communication, bank transactions, bank account security, PINs, passwords, credit card transactions on the web.

- **Cryptography:** Cryptography is a method of storing and transmitting data in a form so that it can be no more interpreted or understood. It is a science of protecting information by encoding it into unreadable format.



Encryption and decryption

How does cryptography work?

A cryptographic algorithm, or cipher, is a mathematical function used in the encryption and decryption process. A cryptographic algorithm works in combination with a key - a word, a number, or phrase - to encrypt the plaintext. The same plaintext encrypts to different cipher text with different keys.

Purpose and application of cryptography:

1. Secure communication
2. Identification and Authentication
3. secret sharing
4. Electronic commerce
5. certification

6. Key recovery
7. Remote access
8. ~~Remote~~ Cell phone
9. Access control

Cryptography definition:

- Algorithm: Set of mathematical rules used in encryption and decryption
- Cryptology: Originated for the Greek kryptos logos, meaning hidden words.
- Plaintext :- The message to be encrypted.
- Key: It is the object used to encrypt~~ed~~ the plaintext.
- Cipher text: It is the encrypted text
- Encryption: The process of converting plain text into cipher text using a appropriate key.
- Decryption: The process of converting cipher text into plain text using a appropriate key.
- Cryptosystem: Hardware and software implementation of cryptography that transforms a message to cipher text and back to plain text
- Cryptographers: People who indulge in cryptography are known as cryptographers
- Cryptanalysis: The art or science of decrypting a cipher text without knowing the authorised key is known as cryptanalysis.
- Cryptanalyst: People who indulge in cryptanalysis. Could be etical or fraudsters.
- Cipher: The method of encryption or decryption is known as cipher
- Encipher: Act of transforming data into an unreadable format.
- Decipher: Act of transforming data into an readable format
- Key clustering :- Instance when two different keys generate same cipher from the same plaintext.

- Key space: Possible values used to construct keys.
- Work factor: Estimated time, effort, and resources necessary to break the cryptosystem.

- Substitution Technique: It is the very basic technique, which makes use of simple letter substitution.

① Caesar Cipher: It is a type of substitution cipher in which each letter in the plaintext is replaced by a letter by some fixed number of positions down the alphabet.

Plain Text	A	B	C	D	E	F	G	H	I	J	K	L	M
Cipher Text	D	E	F	G	H	I	J	K	L	M	N	O	P
Plain Text	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher Text	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Using this scheme, you can tell the key is 3, because all letters are replaced by letters which are succeeding to them by 3.

Algorithm to break caesar cipher:

- Read each alphabet in cipher text message, and search for it in the second row of the table.
- When a match is found, replace the alphabet as specified in above table.
- Repeat the process for all alphabets.
- Simple and easy to memorise.
- This is exceeding weak cryptography by today's standards.

② Modified Version of Caesar cipher: Let us assume that cipher text alphabets corresponding to the original plain text alphabets may not necessarily be three places down the order, but can be any places down the order. Thus out of 26 alphabets, an alphabet can be replaced through any other alphabet in English. For each alphabet, it can have 25 possibilities, thus making it more complex. Trying out all the possibilities in this cipher is known as Brute-Force attack.

Algorithm to break Modified Caesar cipher:

① Let k be a number equal to 1.
② Read the complete cipher text message.
③ Replace each alphabet in the cipher text message with an alphabet that has positions down the order.
④ Increment k by 1.
⑤ If k is less than 26, then go to step 2, Otherwise, stop the process.
⑥ The original message is one of the 25 possibilities produced by the above steps.

③ Mono-alphabetic cipher :- The predictability of Caesar cipher is high. So, rather than using uniform alphabets substitution, mono-alphabetic cipher uses random substitution.

Ex :- 'A' can be replaced by any other letter (random), same 'B' can be replaced by any other random letter and same case for all alphabets. There is no relationship of replacement between letters. Mathematically there are 4×1026 possibilities. This can take years to crack.

④ Polyalphabetic :- A polyalphabetic substitution cipher involves the use of two or more cipher alphabets. Instead of there being a one to one relationship between each letter and its substitute, there is one to many relationship between each letter and its substitutes.

↓ \ →	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Ex:-

Plain text : A T T A C K A T D A W N

Key: : L E M O N L E M O N L E

Cipher Text: L X F O P V E F R N H R

(5) Vernam cipher: The Vernam cipher, also called one time pad, is implemented using a random set of non-repeating characters as the input cipher text.

Algorithm for Vernam cipher:

- Treat each plain text alphabet as a number in an increasing sequence : eg: A = 0, B = 1, ... Z = 25
- Do same for each character of the input cipher text.
- Add each number corresponding to the plain text alphabet to the corresponding input cipher text alphabet number.
- If the sum thus produced is greater than 26, then subtract 26 from it.
- Translate each number of the sum back to the corresponding alphabet. This gives the output of cipher text.

1. Plain Text	H	E	L	L	O	Message
	7	4	11	11	14	
2. One Time Pad	X	M	C	K	L	Key
	23	12	2	10	11	
3. Total	30	16	13	21	25	Message + key
4. Subtract 26 if >25	4	16	13	21	25	Message + key (mod 26)
5. Cipher	E	Q	N	V	Z	

3.2 Transposition Techniques: It is the modified version of substitution technique because this not only substitutes letters but also makes some sort of permutation over the plain text in order to generate cipher text.

Rail Fence Technique:- In Rail fence ciphers, techniques are essentially Transposition ciphers and generated by rearrangement of characters in the plain text.

Let the Plain text be "COMPREHENSIVELY"

(a) Then this plaintext is arranged as Dual slope Rail Fence, Straight, with 3 rungs (levels)

C				R				N				E			
	O		P		E		E		S		V		L		
		M				H				I				Y	

And cipher text read horizontally as: "CRNEOPEESVLMHIY"

(b) The same could be arranged as Dual slope Rail Fence, Straight with 4 rungs (levels)

C						H						E			
	O				E		E				V		L		
		M		R				N		I				Y	
			P						S						

And cipher text read horizontally as: "CHEOEEVLMRNIYPS"

(c) The same could be arranged as Single slope rail fence, with 4 rungs (levels).

			P				E				V			
		M				H				I				Y
	O				E				S				L	
C				R				N				E		

And cipher text read horizontally as :- "PEVMHIYOESLCRNE".

Route cipher :- In a route cipher, the plaintext is first written out in a grid of given dimensions, and then read them off in a pattern given in the key. For example, using the same plain text that we used for rail fence:

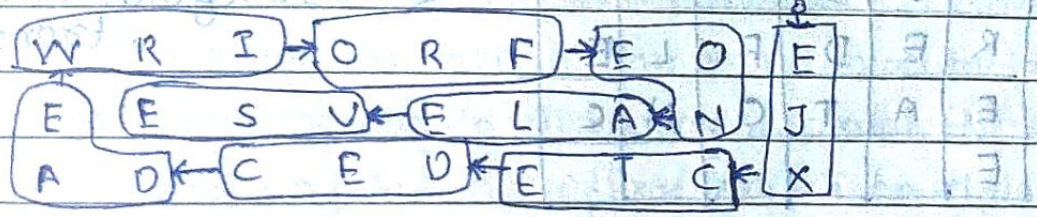
"WE ARE DISCOVERED . FLEE AT ONCE"

W	R	I	O	R	F	E	O
E	E	S	V	E	L	A	N
A	D	C	E	D	E	T	C

E
J
X

The key might specify "spiral inwards, clockwise, starting from the top right". That would give a cipher text of :

"EJXCTEDECDAEWRIORFEONALEVSE"



Columnar Transposition:- In a columnar transposition, the message is written out in rows of a fixed length, and then read out again, column by column, and the columns are chosen in some scrambled order.
Ex:- For word ZEBRAS, in this case order would be "6 3 2 4 1 5".

Suppose we use keyword zebras and the message "WE ARE DISCOVERED FLEE AT ONCE"

6	3	2	4	1	5
W	E	A	R	E	D
I	S	C	O	V	E
R	E	D	F	L	E
E	A	T	O	N	C
E	Q	J	K	E	U

← Regular columnar transposition

Providing five nulls (QJKEU) at the end, the cipher text is read of as:

EVLNE ACDTJ ESEAQ ROFOK DEECU WIREE

6	3	2	4	1	5
W	E	A	R	E	D
I	S	C	O	V	E
R	E	D	F	L	E
E	A	T	O	N	C
E					

← Irregular columnar transposition

It is read as:

EVLNA CDTES EAROF ODEEC WIREE

Steganography:- Stegnography is the art and science of writing hidden messages in such a way that no one, apart from the sender and intended receiver, suspects the existance of message. It works by replacing bits of useless or unused data in regular computer files.

The word steganography is of greek origin and it means "concealed writing". Cryptography protects the content of a message, steganography can be said to protect both messages and communicating parties.

Steganography embeds a secret message in a cover message, this process is usally parameterised by a stego key, and the detection or reading of an embedded information is possible only having this key.

cover_medium + hidden_data + stego_key = stego_medium

Steganography techniques:

- Physical steganography :- Hidden messages within wax tablets, Hidden messages on paper written in secret links, Messages at the back of postage stamps.
- Digital Steganography :- Concealing messages within the lowest bit of noisy images or sound files, Pictures embedded in video material, Concealed messages in tampered executable files.
- Printed Steganography :- A message, plaintext, may be first encrypted in a cipher text. Then cipher text is converted randomly, producing stego text. Only the recipient who knows the technique can decrypt the message.

Terminologies used in steganography :-

- The payload is the data to be covertly communicated
- The carrier is the signal stream or data file into which the payload is hidden; which differs from the channel.
- The resulting signal, stream or data file which has the payload encoded into it is sometimes refered to as a package, stego file or covert message.
- The percentage of bytes, samples, or other signal elements which are modified to encode the payload is referred to as the encoding density.
- In a set of files, those files considered likely to contain a payload are called suspects. If suspect was identified through some type of statistical analysis, it might be referred to as a candidate.

Countermeasures or detection :- Detection of physical steganography requires careful physical examination, including the use of magnification, developer chemicals and ultraviolet light. In computing, detection of steganographically encoded packages is called as steganalysis. The simplest method to detect modified files is to compare them to known originals.

Applications of steganography :-

- Usage in modern printers.
- Digital pictures, which contains large amount of data, are used to hide messages on the internet and on the other communication media.
- Digital watermarking.

3.3 Hashing :- Hash functions are mathematical algorithms that generate message summary or digest to confirm the identity of a specific message and to confirm that there have not been any changes to the content.

- Hash Algorithms are publicly known function that create hash value, also known as message digest, by converting variable length messages into a single fixed length value.
- The message digest is the digest of author's message that is to be compared with the receiver's locally calculated digest of the same message, if both hashes are identical after transmission, the message has arrived without modification.

- Idea of Message Digest : Assume we need to calculate the message digest of the number 7391753, then we multiply each digit in the number by next digit (and discarding the first digits of the multiplication operation.

7391753

$7 \times 3 = 21 = 1$

$1 \times 9 = 09 = 9$

$9 \times 1 = 09 = 9$

$9 \times 7 = 63 = 3$

$3 \times 5 = 15 = 5$

$5 \times 3 = 15 = 5$

- Secure hash function :- Hashing is used to index and retrieve items in a database because it is faster to find them using the shorter hashed key than to find it using the original value.

The hash function is used to index the original value or key and then used later each time the data associated with the value or key is to be retrieved. Thus hashing is always a one-way operation. There is no need to reverse engineer the hash function by analysing hash values. A good hash function should not produce the same hash value from two different inputs. If it does, this is known as collision. A hash function that offers a low risk of collision is considered acceptable.

Simple hash functions use following methods:

1. Division-remainder method
2. Folding method
3. Radix transformation method
4. Digit rearrangement method.

- MD-5 :- In cryptography, MD-5 (Message-Digest algorithm 5) is widely used hash function with a 128-bit hash value and is also commonly used to check integrity of files. An MD-5 hash is typically expressed as a 32-bit hexadecimal number.

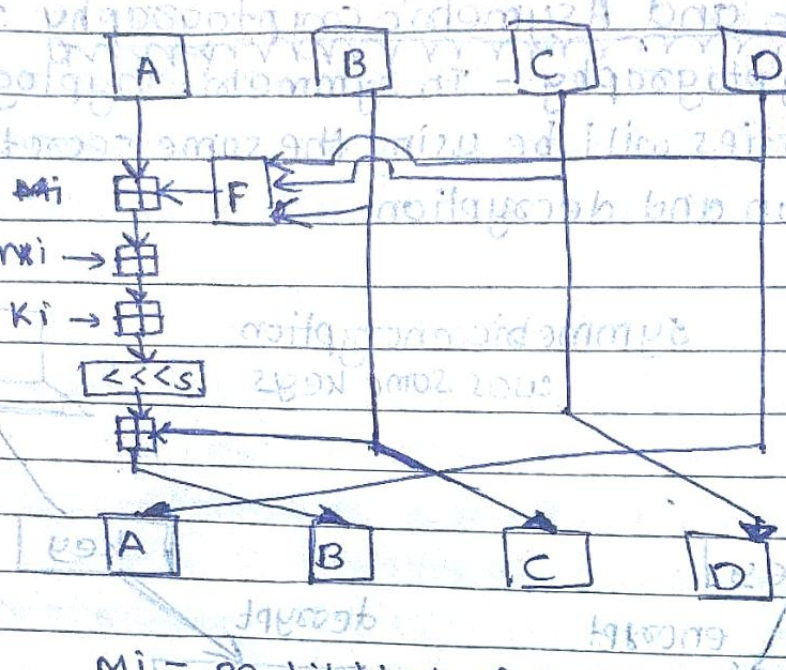
In MD-5:

1. The input message is broken up into chunks of 512-bit blocks.
2. The message is padded so that its length is divisible by 512.

The padding works as follows:

A single bit, 1, is appended to the end of the message. This is followed by adding zeroes, as to bring the length of the message upto 64 bits fewer than multiple of 512.

Following figure shows one operation within a round

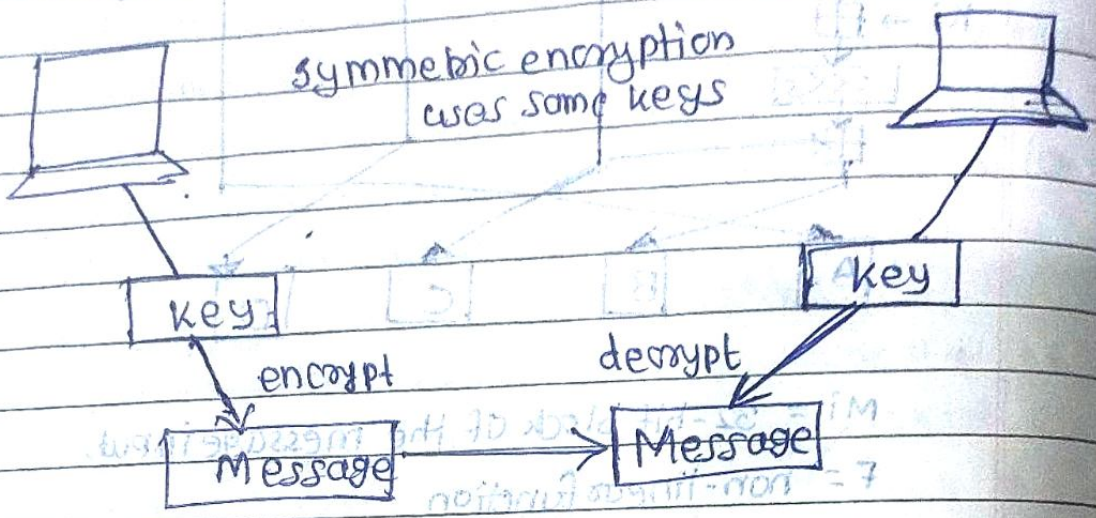


Mi = 32-bit block of the message input
F = non-linear function
Ki = 32-bit constant, different for each operation
<<<s = left bit rotation by s places, s varies each operation
⊞ = addition modulo 2^{32}.

SHA (Secure Hash Algorithm): Secure Hash Algorithm-1 (SHA-1) which is used to compute hash values for calculating a 160 bit hash code based on the plain text message. The hash code is then encrypted with DSS or RSA and append to the original message. The receiver uses the sender's public key to decrypt and recover hash code. Using the same encryption algorithm, the receiver then generates a new hash code from the same message. If two hash code are identical, then the message and the sender is authentic.

3.4 Symmetric and Asymetric cryptography :

Symmetric cryptography :- In symmetric cryptography, both parties will be using the same secret key for encryption and decryption.



Each pair of users who want to exchange data using symmetric key encryption must have their own set of keys as shown in above figure.

Example:

If A and B want to communicate, both need to obtain the copy of the same key. If A wants to communicate with B, C and D, then he needs to have 3 seperate keys, one for each friend.

Because both users use the same key to encrypt and decrypt messages, symmetric cryptosystems can provide confidientiality, but they cannot provide authentication or non-repudiation. There is no way to prove who actually sent a message, if two people are using the exact same key.

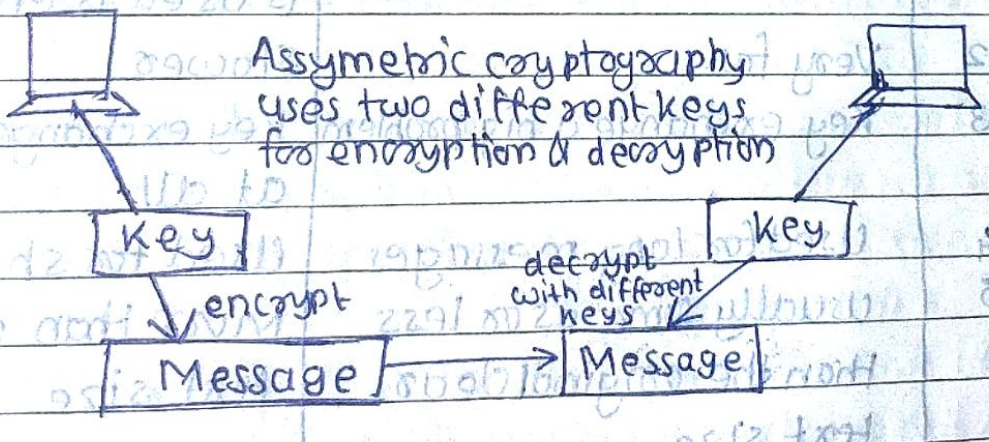
Merits :

① Much faster than assymetric systems.
② Hard to break if using a large key size

Demerits:
1. Key distribution is improper.
2. Scalability
3. Limited security

Types of Symmetric Cryptography Ciphers :-
1. Stream cipher :- Each bit or byte encrypted or decrypted individually
   - Simple substitution ciphers
   - Used for a single message
2. Block cipher :- A block cipher is a type of symmetric key encryption algorithm that transforms a fixed length of block of plain-text data of the same length. Encrypt data one bit or one byte a time. Used, if data is a constant stream of information. Iterated block cipher is when ciphering is repeatedly done.

Assymetric cryptography:



Public key cryptography or assymetric cryptography, is any cryptographic system that uses pairs of keys which may be disseminated widely, and private keys which are known only to the owner.

Merits:

① Better key distribution than symmetric systems
② Better scalability than symmetric systems
③ Can provide confidientiality, authentication and non-repudiation

Demerits:

① slower than assymetric systems

Examples of assymetric key algorithms:

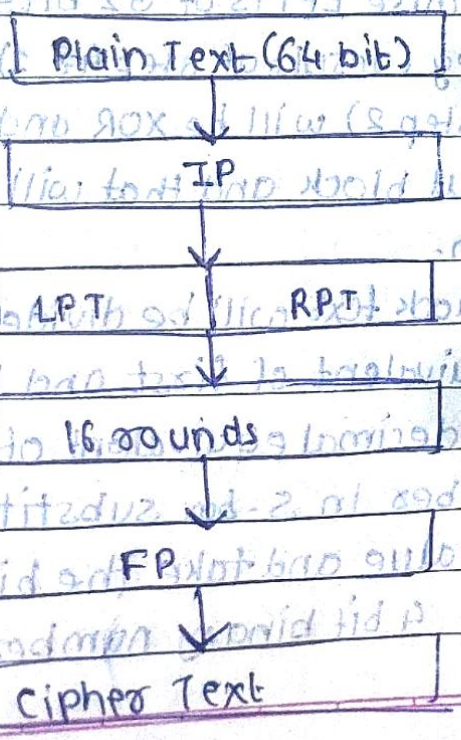
① RSA
② Elliptic Curve Cryptography (ECC)
③ Diffie-Hellman
④ El-Gamal
⑤ Digital Signature Standard (DSS)

	Symmetric key Cryptography	Assymetric key cryptography
1	Same key is used for encryption decryption	One key is used for encryption and other one is used is encryption
2	Very fast	Slower
3	Key exchange a big problem	Key exchange no problem at all
4	Used for long messages	Used for short message
5	Usually same as or less than the original clear text size	More than original clear text size

DES (Data Encryption Standard) :- The Data Encryption Standard (DES) was developed in 1970's by the National Bureau of standards with the help of national Security Agency.

Algorithm:

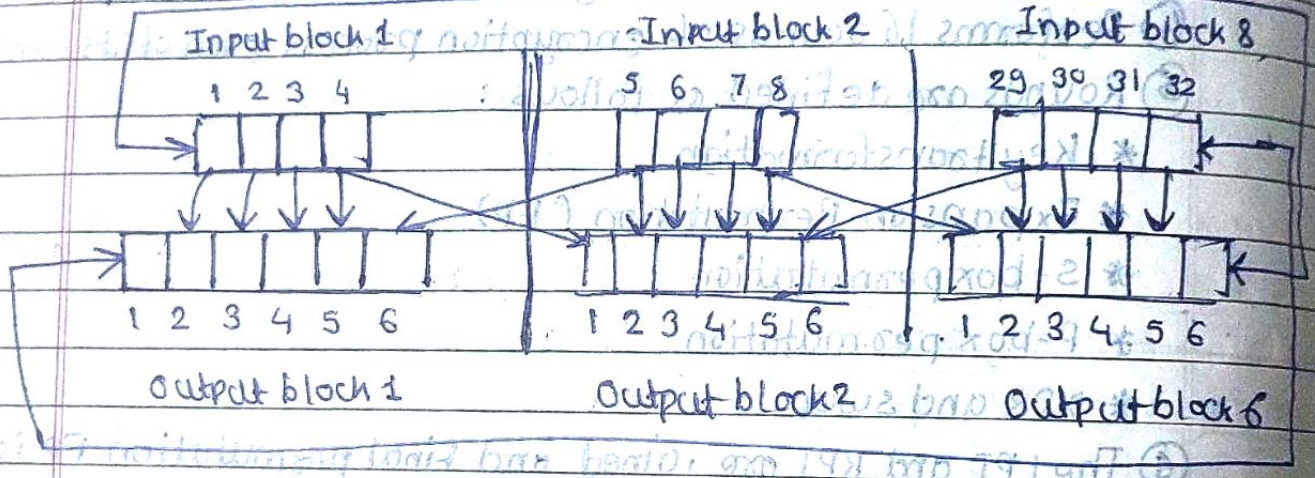
① 64 bit plain text blocks are handed over to the Initial Permutation (IP) function.
② IP is performed on plain text
③ IP produces 2 halves ; say LPT and RPT, both 32 bit each
④ Performs 16 rounds of encryption process each of its own key
⑤ Rounds are defined as follows :
* Key transformation
* Expansion Permutation (EP)
* S-box permutation
* P-box permutation
* XOR and swap
⑥ The LPT and RPT are joined and Final permutation FP is performed on the combined block.
⑦ The result of the process produces 64 bit cipher text.



Step 1: Key Transformation:
- Shifting the key position by considering Round table
- Compare the compression table to get the subkey of 48 bits

Step 2: Expansion Permutation (EP) :- In this step, the 32 bit RPT is expanded to 48 bits as it of key length. The process is as shown under:

The 32 bit text is divided into 8 blocks of 4 bit each. Then by adding 2 bits extra i.e first bit of block 1 is last bit of block 8 and the last bit of block 8 is the first bit of the 7th block the 48 bit text is obtained.



Steps: S-Box permutation:- This step reduces 48 bits RPT into 32 bits because LPT is of 32 bits

A] The 48 bits key (result of step 1) and the 48 bits of RPT (result of step 2) will be XOR and the output will be 48 bits input block and that will be given input for S-box substitution.

B] The 48 bit block text will be divided into 8 blocks of 6 bits each

C] Decimal equivalent of first and last bit denotes row number and decimal equivalent of bit 2, 3, 4, 5 denote column number in S-box substitution table.

D] Check the value and take the binary equivalent of number

E] The result is 4 bit binary number.

0010 Column number 2

Ex:

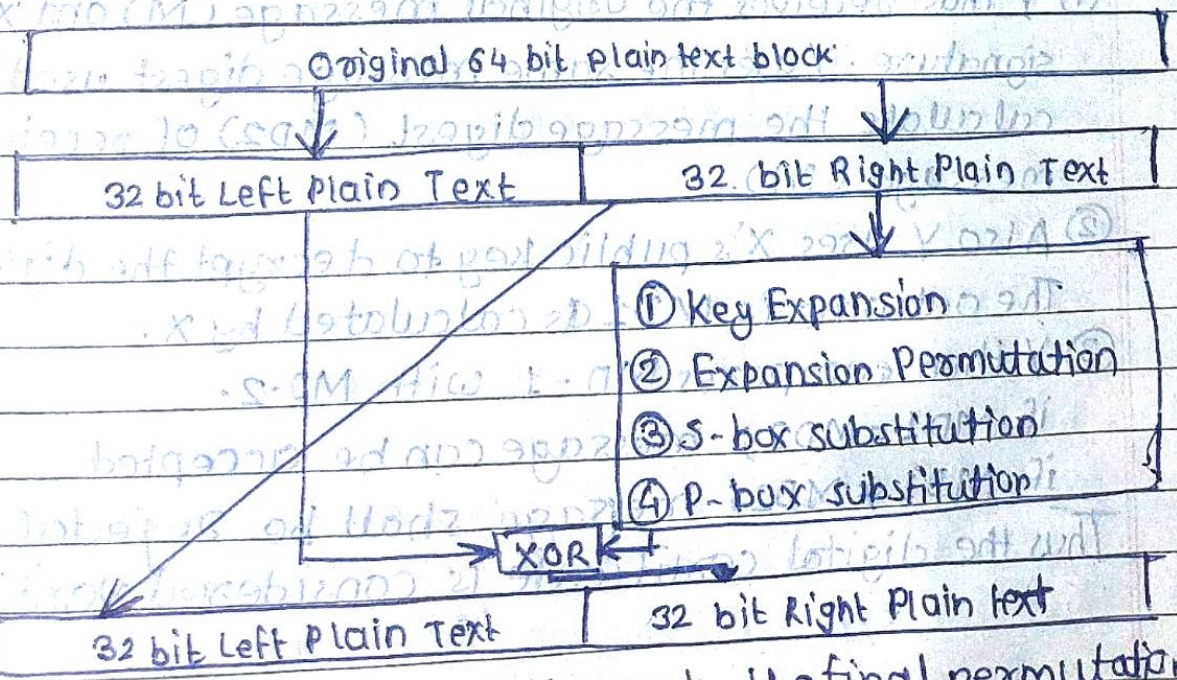
1	0	0	1	0	1

11 → Row number 3

Check third row and second column in SBox-1 substitution table. It is given as 1 in the table. So binary equivalent is 0001. Thus input of 100101 is reduced to 0001.

STEP 4:- P-box permutation:- In this step, the output of S-box is permuted using a P-box. This mechanism involves a simple permutation i.e replacement of each bit with another bit as specied in P-box table without any expansion or compression. This is called P-box permutation.

STEP 5:- XOR and swap: The untouched LPT of 32 bits is XORed with output of P-box permutation. The result of this operation becomes new right half. The old right half becomes the new left half in the process of swapping.



Final Permutation:- At the end of 16 rounds, the final permutation. This is the simple transposition based on the final permutation table. The output of final permutation is the 64 bit encrypted block.

Digital Signature:- A digital signature is a mathematical scheme for presenting the authenticity of digital messages or documents. A valid digital signature gives a recipient reason ~~that~~ to believe that the message was created by a known sender (authentication), that the sender cannot deny having sent to message (non-repudiation), and that the message was not altered to transist (integrity).

steps for process:

sender's side:-

① If X is the sender, the SHA-1 algorithm is used to calculate the message digest (MD1) of original message.

② This MD1 is further encrypted using RSA with X's private key. The output is called the Digital signature (DS) of X.

③ Further, the original message (M) along with Digital signature (DS) is sent to the receiver.

Receiver's side:-

① Y thus receives the original message (M) and X's digital signature. Y uses the same message digest used by X to calculate the message digest (MD2) of received message (M).

② Also Y uses X's public key to decrypt the digital signature. The outcome is MD1 as calculated by X.

③ Y then compares MD-1 with MD-2.

if MD1 = MD2; message can be accepted

if MD1 ≠ MD2; message shall be rejected.

Thus the digital certificate is considered verified.

Properties of digital signature: (describe in detail)

① Integrity
② Authentication
③ Non-repudiation

Advantages of digital signature:

① Imposter prevention
② Message integrity
③ Legal requirements
④ Better than ink on paper signature
⑤ Can be used to conduct business thousands of miles away.

Disadvantages :- The disadvantages of using digital signature involve the primary avenue for any business : money. This is because the business may have to spend more money than usual to work with digital signatures including buying certificates from certification authorities and getting the verification software.

Physical signature	Digital signature
① Physical signature is just written on paper	① Digital signature encompasses crucial parameters of identification
② Can be copied	② Impossible to copy
③ Does not give privacy to any content	③ Enables encryption and thus privacy
④ Cannot protect the content	④ Protects the content