

**Unit-1**

**Introduction:**

**cryptography is technique of securing information and communications through use of codes so that only those person for whom the information is intended can understand it and process it. Thus preventing unauthorized access to information. The prefix “crypt” means “hidden” and suffix graphy means “writing”.**

**In Cryptography the techniques which are use to protect information are obtained from mathematical concepts and a set of rule based calculations known as algorithms to convert messages in ways that make it hard to decode it. These algorithms are used for cryptographic key generation, digital signing, verification to protect data privacy, web browsing on internet and to protect confidential transactions such as credit card and debit card transactions.**

**Techniques used For Cryptography:**

**In today’s age of computers cryptography is often associated with the process where an ordinary plain text is converted to cipher text which is the text made such that intended receiver of the text can only decode it and hence this process is known as encryption. The process of conversion of cipher text to plain text this is known as decryption.**

**Features Of Cryptography are as follows:**

**1.Confidentiality:**

**Information can only be accessed by the person for whom it is intended and no other person except him can access it.**

**2.Integrity:**

**Information cannot be modified in storage or transition between sender and intended receiver without any addition to information being detected.**

**3.Non-repudiation:**

**The creator/sender of information cannot deny his intention to send information at later stage.**

**4.Authentication:**

**The identities of sender and receiver are confirmed. As well as destination/origin of information is confirmed.**

**Consider two parties Alice and Bob. Now, Alice wants to send a message m to Bob over a secure channel. So, what happens is as follows. The sender’s message or sometimes called the Plaintext, is converted into an unreadable form using a Key k. The resultant text obtained is called the Ciphertext. This process is known as Encryption. At the time of receival, the Ciphertext is converted back into the plaintext using the same Key k, so that it can be read by the receiver. This process is known as Decryption.**

**Alice (Sender) Bob (Receiver)**

**C = E (m, k) ----> m = D (C, k)**

Here, C refers to the Ciphertext while E and D are the Encryption and Decryption algorithms respectively. Let’s consider the case of Caesar Cipher or Shift Cipher as an example. As the name suggests, in Caesar Cipher each character in a word is replaced by another character under some defined rules. Thus, if A is replaced by D, B by E and so on. Then, each character in the word would be shifted by a position of 3. For example:

Plaintext : ATTACKATDAWN

Ciphertext : LXFOPVEFRNHR

Note that even if the adversary knows that the cipher is based on Caesar Cipher, it cannot predict the plaintext as it doesn’t have the key in this case which is to shift the characters back by three places.

**Types Of Cryptography:**

**In general there are three types of cryptography:**

**1. Symmetric Key Cryptography:**

It is an encryption system where the sender and receiver of message use a single common key to encrypt and decrypt messages. Symmetric Key Systems are faster and simpler but the problem is that sender and receiver have to somehow exchange key in a secure manner. The most popular symmetric key cryptography system is Data Encryption System(DES).

**2. Hash Functions:**

There is no usage of any key in this algorithm. A hash value with fixed length is calculated as per the plain text which makes it impossible for contents of plain text to be recovered. Many operating systems use hash functions to encrypt passwords.

**3. Asymmetric Key Cryptography:**

Under this system a pair of keys is used to encrypt and decrypt information. A public key is used for encryption and a private key is used for decryption. Public key and Private Key are different. Even if the public key is known by everyone the intended receiver can only decode it because he alone knows the private key.

Modern cryptography is the cornerstone of computer and communications security. Its foundation is based on various concepts of mathematics such as number theory, computational-complexity theory, and probability theory.

## Characteristics of Modern Cryptography

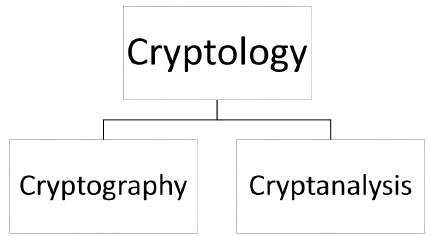
There are three major characteristics that separate modern cryptography from the classical approach.

|  |  |
| --- | --- |
| **Classic Cryptography** | **Modern Cryptography** |
| It manipulates traditional characters, i.e., letters and digits directly. | It operates on binary bit sequences. |
| It is mainly based on ‘security through obscurity’. The techniques employed for coding were kept secret and only the parties involved in communication knew about them. | It relies on publicly known mathematical algorithms for coding the information. Secrecy is obtained through a secrete key which is used as the seed for the algorithms. The computational difficulty of algorithms, absence of secret key, etc., make it impossible for an attacker to obtain the original information even if he knows the algorithm used for coding. |
| It requires the entire cryptosystem for communicating confidentially. | Modern cryptography requires parties interested in secure communication to possess the secret key only. |

## Context of Cryptography

Cryptology, the study of cryptosystems, can be subdivided into two branches −

* Cryptography
* Cryptanalysis



### **What is Cryptography?**

*Cryptography is the art and science of making a cryptosystem that is capable of providing information security.*

Cryptography deals with the actual securing of digital data. It refers to the design of mechanisms based on mathematical algorithms that provide fundamental information security services. You can think of cryptography as the establishment of a large toolkit containing different techniques in security applications.

### **What is Cryptanalysis?**

*The art and science of breaking the cipher text is known as cryptanalysis.*

Cryptanalysis is the sister branch of cryptography and they both co-exist. The cryptographic process results in the cipher text for transmission or storage. It involves the study of cryptographic mechanism with the intention to break them. Cryptanalysis is also used during the design of the new cryptographic techniques to test their security strengths.

**Note** − Cryptography concerns with the design of cryptosystems, while cryptanalysis studies the breaking of cryptosystems.

## Security Services of Cryptography

The primary objective of using cryptography is to provide the following four fundamental information security services. Let us now see the possible goals intended to be fulfilled by cryptography.

### **Confidentiality**

Confidentiality is the fundamental security service provided by cryptography. It is a security service that keeps the information from an unauthorized person. It is sometimes referred to as **privacy** or **secrecy**.

Confidentiality can be achieved through numerous means starting from physical securing to the use of mathematical algorithms for data encryption.

### **Data Integrity**

It is security service that deals with identifying any alteration to the data. The data may get modified by an unauthorized entity intentionally or accidently. Integrity service confirms that whether data is intact or not since it was last created, transmitted, or stored by an authorized user.

Data integrity cannot prevent the alteration of data, but provides a means for detecting whether data has been manipulated in an unauthorized manner.

### **Authentication**

Authentication provides the identification of the originator. It confirms to the receiver that the data received has been sent only by an identified and verified sender.

Authentication service has two variants −

* **Message authentication** identifies the originator of the message without any regard router or system that has sent the message.
* **Entity authentication** is assurance that data has been received from a specific entity, say a particular website.

Apart from the originator, authentication may also provide assurance about other parameters related to data such as the date and time of creation/transmission.

### **Non-repudiation**

It is a security service that ensures that an entity cannot refuse the ownership of a previous commitment or an action. It is an assurance that the original creator of the data cannot deny the creation or transmission of the said data to a recipient or third party.

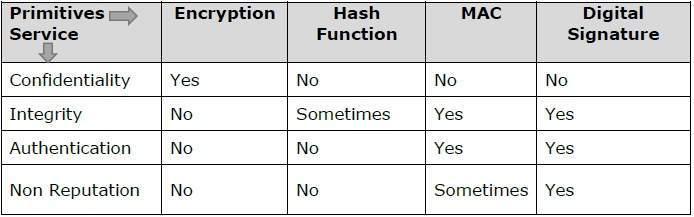
Non-repudiation is a property that is most desirable in situations where there are chances of a dispute over the exchange of data. For example, once an order is placed electronically, a purchaser cannot deny the purchase order, if non-repudiation service was enabled in this transaction.

## Cryptography Primitives

Cryptography primitives are nothing but the tools and techniques in Cryptography that can be selectively used to provide a set of desired security services −

* Encryption
* Hash functions
* Message Authentication codes (MAC)
* Digital Signatures

The following table shows the primitives that can achieve a particular security service on their own.



**Note** − Cryptographic primitives are intricately related and they are often combined to achieve a set of desired security services from a cryptosystem.

**Information theory**

**Information theory is based on probability theory and statistics**, where quantified information is usually described in terms of bits. Information theory often concerns itself with measures of information of the distributions associated with random variables.

Information theory is the scientific study of the [quantification](https://en.wikipedia.org/wiki/Quantification_(science)), [storage](https://en.wikipedia.org/wiki/Computer_data_storage), and [communication](https://en.wikipedia.org/wiki/Telecommunication) of [digital](https://en.wikipedia.org/wiki/Digital_data) [information](https://en.wikipedia.org/wiki/Information). The field was fundamentally established by the works of [Harry Nyquist](https://en.wikipedia.org/wiki/Harry_Nyquist) and [Ralph Hartley](https://en.wikipedia.org/wiki/Ralph_Hartley), in the 1920s, and [Claude Shannon](https://en.wikipedia.org/wiki/Claude_Shannon) in the 1940s.

The field is at the intersection of [probability theory](https://en.wikipedia.org/wiki/Probability_theory), [statistics](https://en.wikipedia.org/wiki/Statistics), [computer science](https://en.wikipedia.org/wiki/Computer_science), [statistical mechanics](https://en.wikipedia.org/wiki/Statistical_mechanics), [information engineering](https://en.wikipedia.org/wiki/Information_engineering_(field)), and [electrical engineering](https://en.wikipedia.org/wiki/Electrical_engineering).

A key measure in information theory is [entropy](https://en.wikipedia.org/wiki/Information_entropy). Entropy quantifies the amount of uncertainty involved in the value of a [random variable](https://en.wikipedia.org/wiki/Random_variable) or the outcome of a [random process](https://en.wikipedia.org/wiki/Random_process). For example, identifying the outcome of a fair [coin flip](https://en.wikipedia.org/wiki/Coin_flip) (with two equally likely outcomes) provides less information (lower entropy) than specifying the outcome from a roll of a [die](https://en.wikipedia.org/wiki/Dice) (with six equally likely outcomes). Some other important measures in information theory are [mutual information](https://en.wikipedia.org/wiki/Mutual_information), [channel capacity](https://en.wikipedia.org/wiki/Channel_capacity), [error exponents](https://en.wikipedia.org/wiki/Error_exponent), and [relative entropy](https://en.wikipedia.org/wiki/Relative_entropy). Important sub-fields of information theory include [source coding](https://en.wikipedia.org/wiki/Source_coding), [algorithmic complexity theory](https://en.wikipedia.org/wiki/Algorithmic_complexity_theory), [algorithmic information theory](https://en.wikipedia.org/wiki/Algorithmic_information_theory) and [information-theoretic security](https://en.wikipedia.org/wiki/Information-theoretic_security).

Applications of fundamental topics of information theory include source coding/[data compression](https://en.wikipedia.org/wiki/Data_compression) (e.g. for [ZIP files](https://en.wikipedia.org/wiki/ZIP_(file_format))), and channel coding/[error detection and correction](https://en.wikipedia.org/wiki/Error_detection_and_correction) (e.g. for [DSL](https://en.wikipedia.org/wiki/Digital_subscriber_line)). Its impact has been crucial to the success of the [Voyager](https://en.wikipedia.org/wiki/Voyager_program) missions to deep space, the invention of the [compact disc](https://en.wikipedia.org/wiki/Compact_disc), the feasibility of mobile phones and the development of the Internet. The theory has also found applications in other areas, including [statistical inference](https://en.wikipedia.org/wiki/Statistical_inference),

[cryptography](https://en.wikipedia.org/wiki/Cryptography), [neurobiology](https://en.wikipedia.org/wiki/Neurobiology), [perception](https://en.wikipedia.org/wiki/Perception), linguistics, the evolution and function[]](https://en.wikipedia.org/wiki/Information_theory#cite_note-7) of molecular codes ([bioinformatics](https://en.wikipedia.org/wiki/Bioinformatics)), [thermal physics](https://en.wikipedia.org/wiki/Thermal_physics), [molecular dynamics](https://en.wikipedia.org/wiki/Molecular_dynamics), [quantum computing](https://en.wikipedia.org/wiki/Quantum_computing), [black holes](https://en.wikipedia.org/wiki/Black_hole), [information retrieval](https://en.wikipedia.org/wiki/Information_retrieval), [intelligence gathering](https://en.wikipedia.org/wiki/Intelligence_(Information_Gathering)), [plagiarism detection](https://en.wikipedia.org/wiki/Plagiarism_detection), [pattern recognition](https://en.wikipedia.org/wiki/Pattern_recognition), [anomaly detection](https://en.wikipedia.org/wiki/Anomaly_detection) and even art creation.

**Number theory**

Cryptography allows its users, whether governments, military, businesses, or individuals, to maintain privacy and confidentiality in their communications. The goal of every cryptographic scheme is to be "crack proof" (i.e, only able to be decoded and understood by authorized recipients). Cryptography is also a means to ensure the integrity and preservation of data from tampering. Modern cryptographic systems rely on functions associated with advanced mathematics, including a specialized branch of mathematics termed [number theory](https://www.encyclopedia.com/science-and-technology/mathematics/mathematics/number-theory) that explores the properties of numbers and the relationships between numbers.

**Classical cryptography**

Cryptography is the technique which is used for doing secure communication between two parties in the public environment where unauthorized users and malicious attackers are present. In cryptography there are two processes i.e. encryption and decryption performed at sender and receiver end respectively.

Encryption is the processes where a simple multimedia data is combined with some additional data (known as key) and converted into unreadable encoded format known as Cipher. Decryption is the reverse method as that of encryption where the same or different additional data (key) is used to decode the cipher and it is converted in to the real multimedia data.

Cryptography techniques can be categorized according to their basic principles or protocols they follow. But here we are going to concentrate on the two types of cryptography technique: Classical Cryptography and Quantum Cryptography. These are explained as following below.

**1. Classical Cryptography:**

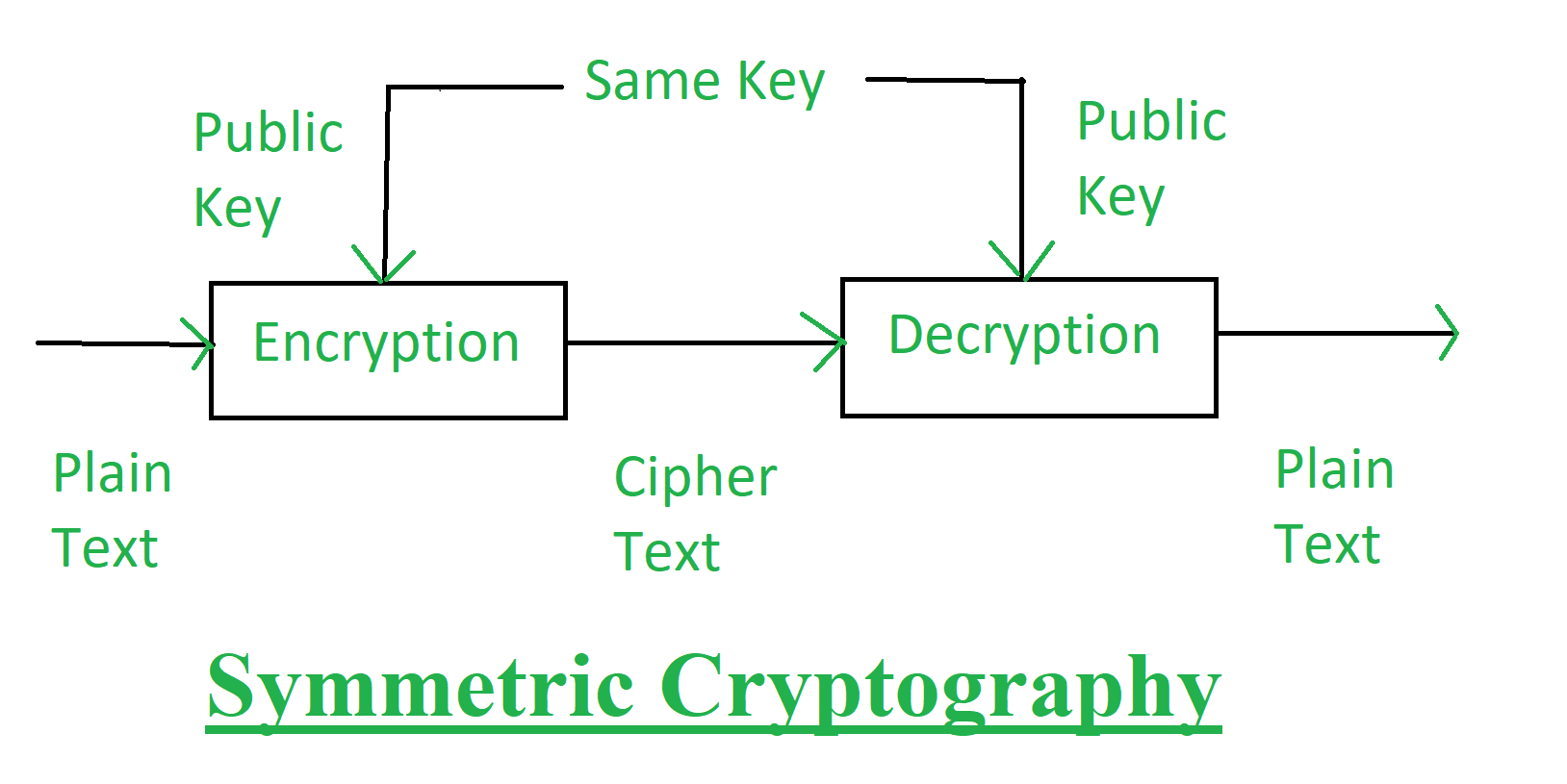
Classical cryptography is based on the mathematics and it relies on the computational difficulty of factorizing large number. The security of classical cryptography is based on the high complexity of the mathematical problem for the instance factorization of large number.

In the classical cryptography the original data i.e., the plain text is transformed into the encoded format i.e. cipher text so that we can transmit this data through insecure communication channels. A data string which known as key is used to control the transformation of the data from plain text to cipher text. This arrangement helps to keep data safe as it required the key for extracting the original information from the cipher text. Without the key no one can read the data. In this technique it is assumed that the only authorized receiver has the key.

**Classical Cryptography has two types of techniques:**

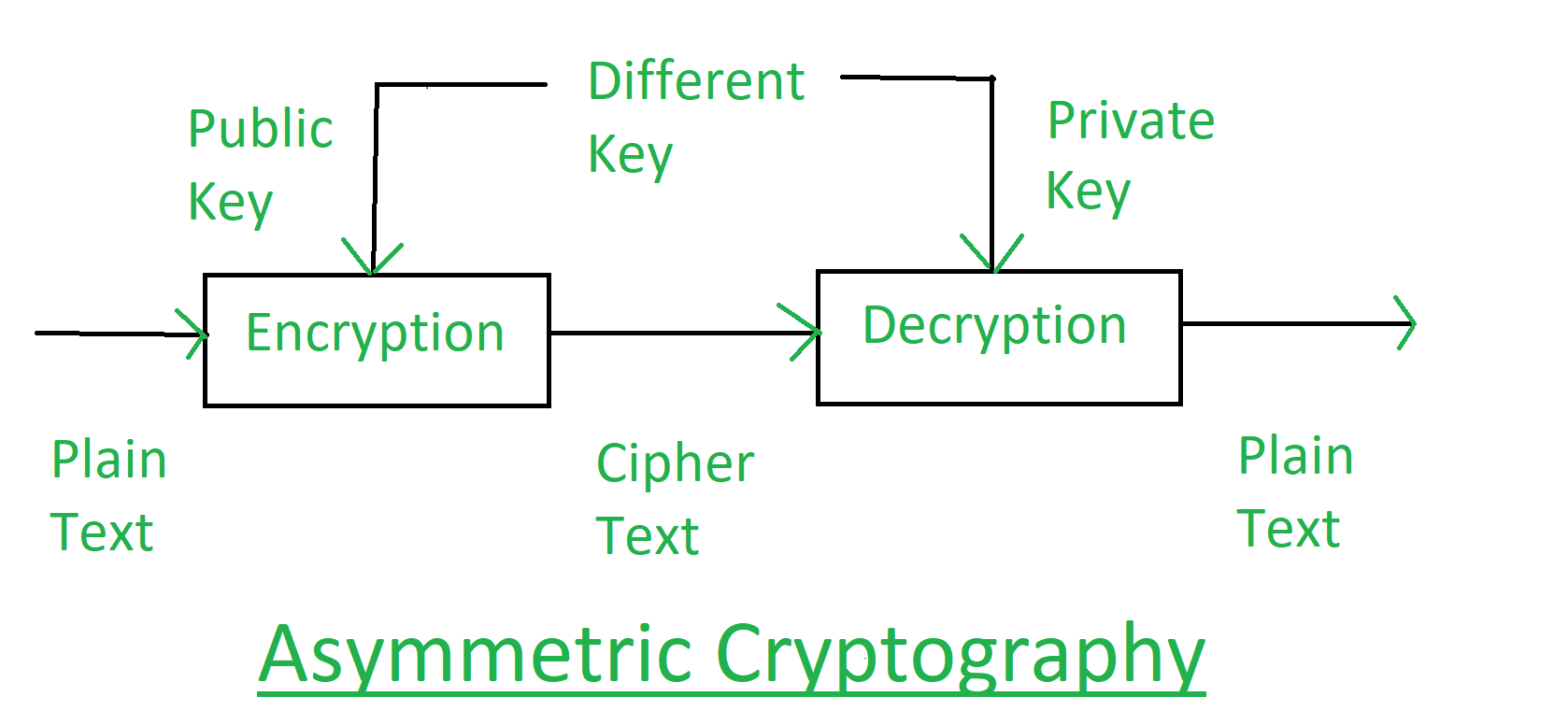
**Symmetric Cryptography:**

In the symmetric cryptography a single key is used for encrypting and decryption the data. This encryption key is private key. This is the limitation of this encryption technique that this private key must be distributed only among the authorized sender and receiver.



**Asymmetric Cryptography:**

In the asymmetric cryptography a pair of key, i.e., public key and private key is used for encryption and decryption. A sender can use its public key to encrypt the data and on receiver end receiver can decrypt the data by using its private key. This technique overcomes the problem of key distribution.



**Advantages of Classical Cryptography:**

* While employing the one-time pad, it is unbreakable.
* It is easy to do manually, no computer required.
* It protects the plain text from casual snooping.

**Disadvantages of Classical Cryptography:**

* While employing the one-time pad, it is cumbersome and requires a personal meetup to exchange the pads.
* If not employing the OTP, anyone who is even remotely interested in knowing what you wrote and knows about cryptography will be able to break the encryption.

**Shannon’s Theory**

Shannon’s Theory in 1949, Claude Shannon published a paper entitled “communication Theory of Secrecy Systems” in the bell systems technical journal. This paper had a great influence on the scientific study of cryptography.

**Hill cipher**

Hill cipher is a polygraphic substitution cipher based on linear algebra.Each letter is represented by a number modulo 26. Often the simple scheme A = 0, B = 1, …, Z = 25 is used, but this is not an essential feature of the cipher. To encrypt a message, each block of n letters (considered as an n-component vector) is multiplied by an invertible n × n matrix, against modulus 26. To decrypt the message, each block is multiplied by the inverse of the matrix used for encryption.

The matrix used for encryption is the cipher key, and it should be chosen randomly from the set of invertible n × n matrices (modulo 26).

**Examples:**

**Input** : Plaintext: **ACT**

**Key:** GYBNQKURP

**Output :** Ciphertext: **POH**

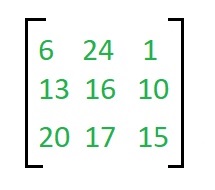
**Input**  : Plaintext: **GFG**

**Key**: HILLMAGIC

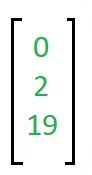
**Output :** Ciphertext: **SWY**

**Encryption**

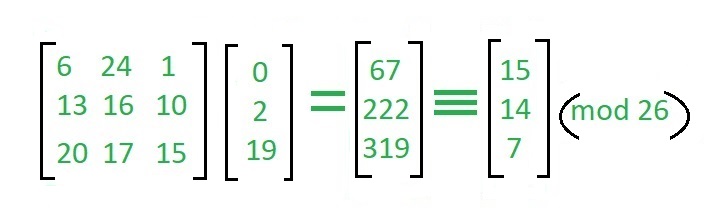
We have to encrypt the message ‘ACT’ (n=3).The key is ‘GYBNQKURP’ which can be written as the nxn matrix:



**The message ‘ACT’ is written as vector:**



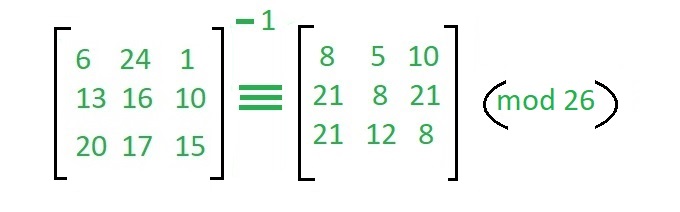
The enciphered vector is given as:



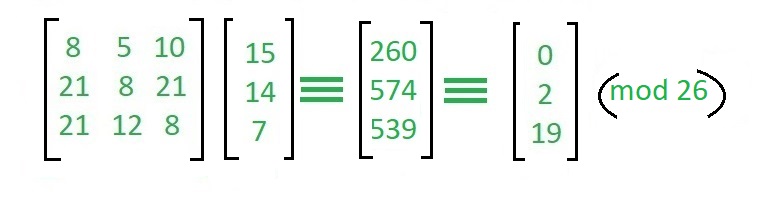
Which corresponds to ciphertext of ‘POH’

**Decryption**

To decrypt the message, we turn the ciphertext back into a vector, then simply multiply by the inverse matrix of the key matrix (IFKVIVVMI in letters).The inverse of the matrix used in the previous **example is:**



For the previous ciphertext ‘POH’



Which gives us back ‘ACT’

Assume that all the alphabets are in upper case.

Below is the implementation of the above idea for n=3.

**Unit – 2**

# Symmetric Cipher

Ciphers or algorithms can be either symmetric or asymmetric. Symmetric ones use the same key (called a secret key or private key) for transforming the original message, called plaintext, into ciphertext and vice versa. Symmetric ciphers are generally faster than their asymmetric counterparts, which use a closely-held private key as well as a public key shared between the two parties (hence public-key cryptography, or PKC). Examples of symmetric ciphers are Advanced Encryption Standard (AES), Data Encryption Standard (DES), Blowfish, and International Data Encryption Algorithm (IDEA).

**Modern block cipher**

A modern block cipher is a cipher which encrypts m-bit block of plaintext and decrypts m-bit block of ciphertext. For encryption or decryption, modern block cipher facilitate a K bit key and the decryption algorithm should be inverse of encryption algorithms and for both encryption and decryption similar key is used.

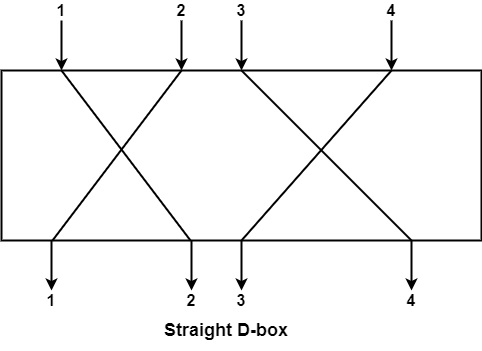
A block cipher works on a plaintext block of n bits to make a cipher text block of n bits. There are possible multiple plaintext blocks and, for the encryption to be reversible (i.e., for decryption to be applicable), each should create a unique cipher text block. Such transformation is known as reversible, or non-singular.

Block cipher modes of operation have been produced to delete the chance of encrypting identical blocks of text the similar method, the ciphertext formed from the previous encrypted block is used to the next block. A block of bits is known as an initialization vector (IV).

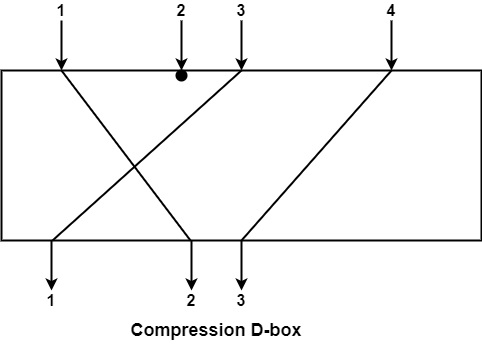
There are various components of Modern Block Cipher which are as follows −

* **D-boxes** − A D-box is a permutation box having similar features as traditional transposition ciphers. D-boxes transpose bits. There are three types of D-boxes which are as follows −

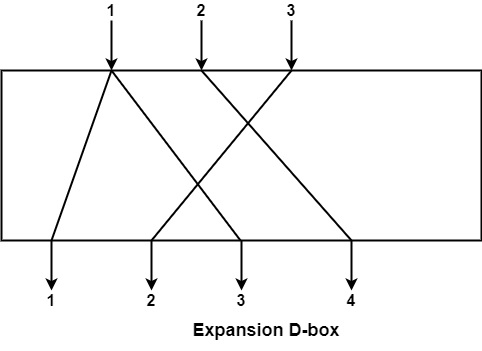
**Straight D-box** − It creates n inputs, permutes them and supports n outputs. In this, the second input after permutation is the first to be outputted. The first letter in input is permuted to second place, third on fourth place and fourth on third place. There are n! Possible way of mapping in D-box.



**Compression D-box** − This is a D-box with n inputs and m outputs, where m<n. There are various inputs are blocked and do not reach the output. Compression D-boxes are used when it is required to permute bits and at the similar time reduce the number of bits for the next stage.



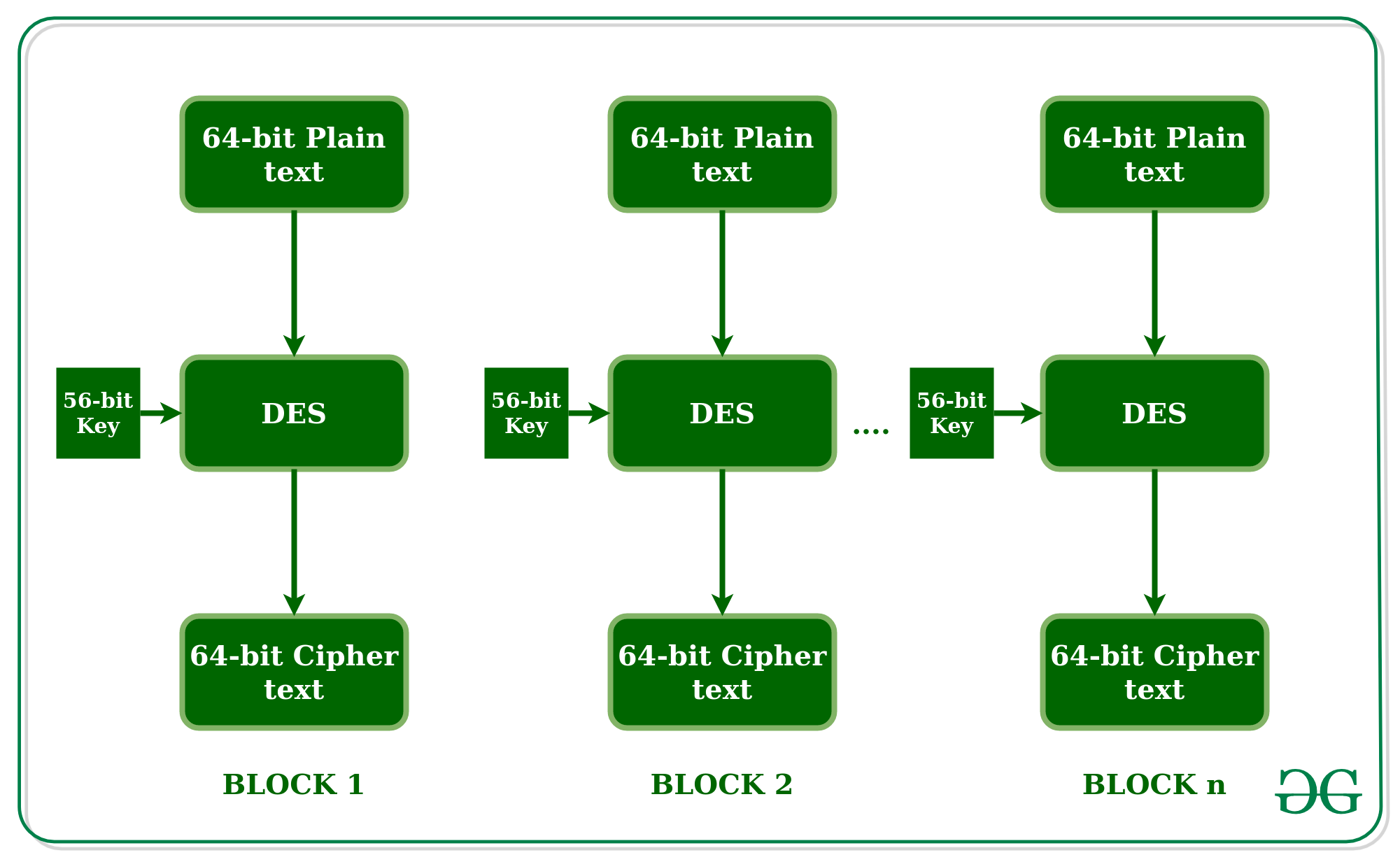
**Expansion D-box** − This is a D-box with n inputs and m outputs, where m >n i.e., there are various inputs are connected to more than one output it is used when it is required to transpose bits and the same increase the multiple bits for the next stage.



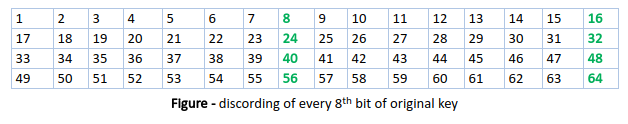
* **S-boxes** − These are substitution boxes same to the substitution cipher. The input to an S-box can be a n-bit word but the output can be a m-bit word, where m and n are not essentially the same.
* **Circular Shift** − It can also discovered in modern block ciphers, it can be such as leftshift or right-shift. In the circular left shift, shift each bit in n-bit word with m position to the left and the leftmost m-bits are deleted from the left and become the rightmost bits.

**Data encryption standard (DES)**

Data encryption standard (DES) has been found vulnerable to very powerful attacks and therefore, the popularity of DES has been found slightly on the decline. DES is a block cipher and encrypts data in blocks of size of 64 bits each, which means 64 bits of plain text go as the input to DES, which produces 64 bits of ciphertext. The same algorithm and key are used for encryption and decryption, with minor differences. The key length is 56 bits. The basic idea is shown in the figure:



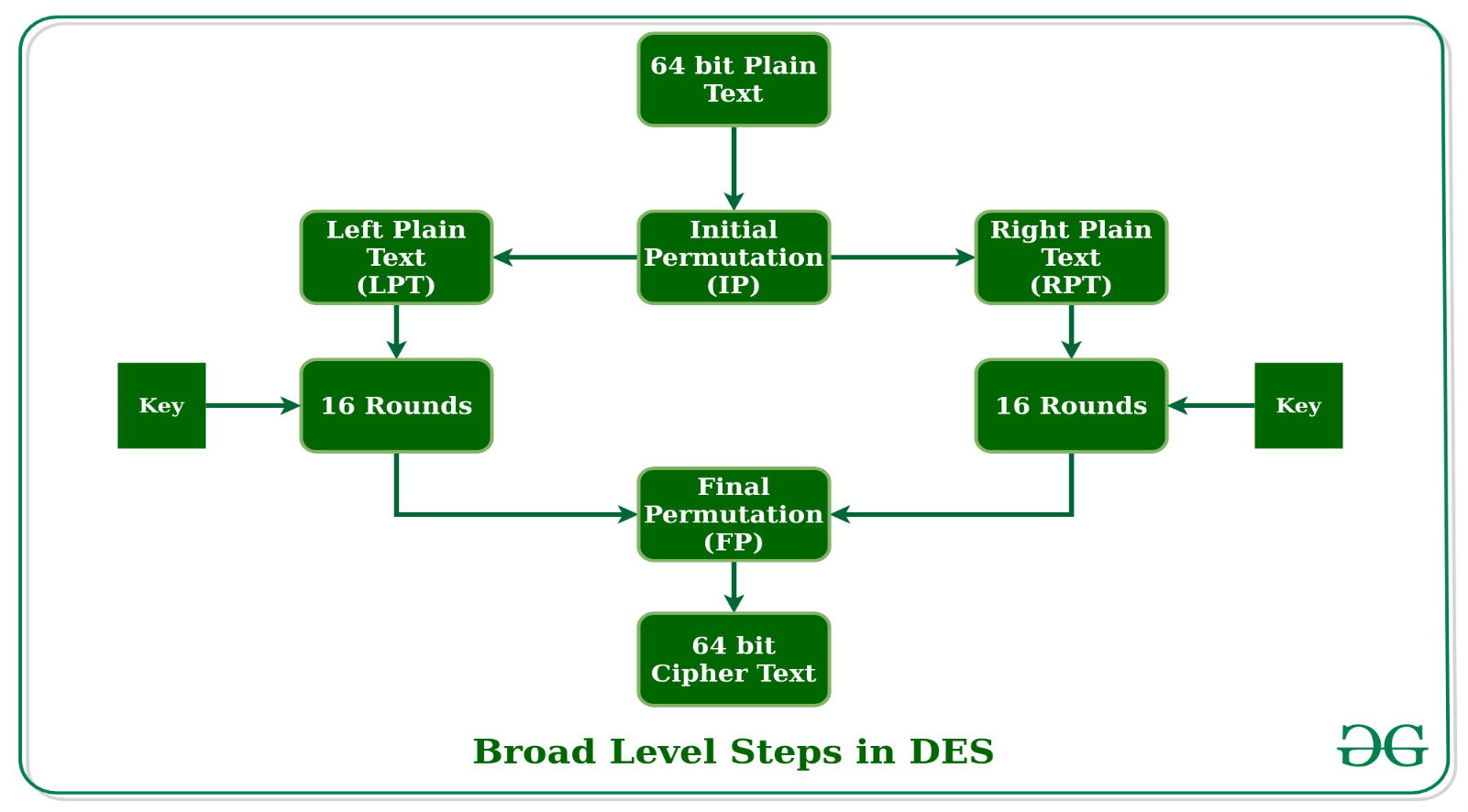
We have mentioned that DES uses a 56-bit key. Actually, the initial key consists of 64 bits. However, before the DES process even starts, every 8th bit of the key is discarded to produce a 56-bit key. That is bit positions 8, 16, 24, 32, 40, 48, 56, and 64 are discarded.



Thus, the discarding of every 8th bit of the key produces a 56-bit key from the original 64-bit key.

DES is based on the two fundamental attributes of cryptography: substitution (also called confusion) and transposition (also called diffusion). DES consists of 16 steps, each of which is called a round. Each round performs the steps of substitution and transposition. Let us now discuss the broad-level steps in DES.

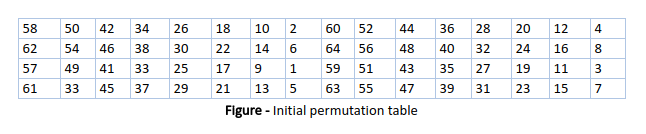
* In the first step, the 64-bit plain text block is handed over to an initial Permutation (IP) function.
* The initial permutation is performed on plain text.
* Next, the initial permutation (IP) produces two halves of the permuted block; saying Left Plain Text (LPT) and Right Plain Text (RPT).
* Now each LPT and RPT go through 16 rounds of the encryption process.
* In the end, LPT and RPT are rejoined and a Final Permutation (FP) is performed on the combined block
* The result of this process produces 64-bit ciphertext.



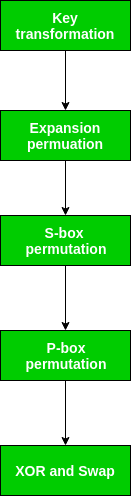
**Initial Permutation (IP):**

As we have noted, the initial permutation (IP) happens only once and it happens before the first round. It suggests how the transposition in IP should proceed, as shown in the figure. For example, it says that the IP replaces the first bit of the original plain text block with the 58th bit of the original plain text, the second bit with the 50th bit of the original plain text block, and so on.

This is nothing but jugglery of bit positions of the original plain text block. the same rule applies to all the other bit positions shown in the figure.



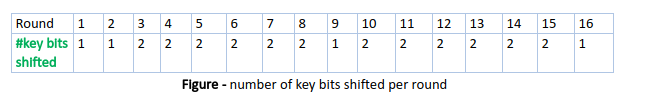
As we have noted after IP is done, the resulting 64-bit permuted text block is divided into two half blocks. Each half-block consists of 32 bits, and each of the 16 rounds, in turn, consists of the **broad-level steps outlined in the figure.**



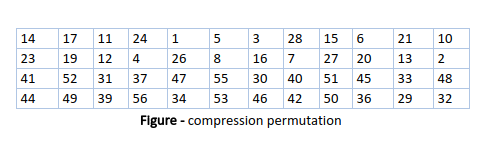
**Step-1: Key transformation:**

We have noted initial 64-bit key is transformed into a 56-bit key by discarding every 8th bit of the initial key. Thus, for each a 56-bit key is available. From this 56-bit key, a different 48-bit Sub Key is generated during each round using a process called key transformation. For this, the 56-bit key is divided into two halves, each of 28 bits. These halves are circularly shifted left by one or two positions, depending on the round.

For example: if the round numbers 1, 2, 9, or 16 the shift is done by only one position for other rounds, the circular shift is done by two positions. The number of key bits shifted per round is shown in the figure.



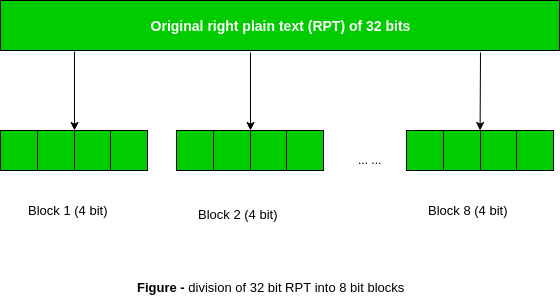
After an appropriate shift, 48 of the 56 bits are selected. for selecting 48 of the 56 bits the table is shown in the figure given below. For instance, after the shift, bit number 14 moves to the first position, bit number 17 moves to the second position, and so on. If we observe the table carefully, we will realize that it contains only 48-bit positions. Bit number 18 is discarded (we will not find it in the table), like 7 others, to reduce a 56-bit key to a 48-bit key. Since the key transformation process involves permutation as well as a selection of a 48-bit subset of the original 56-bit key it is called Compression Permutation.



Because of this compression permutation technique, a different subset of key bits is used in each round. That makes DES not easy to crack.

**Step-2: Expansion Permutation:**

Recall that after the initial permutation, we had two 32-bit plain text areas called Left Plain Text(LPT) and Right Plain Text(RPT). During the expansion permutation, the RPT is expanded from 32 bits to 48 bits. Bits are permuted as well hence called expansion permutation. This happens as the 32-bit RPT is divided into 8 blocks, with each block consisting of 4 bits. Then, each 4-bit block of the previous step is then expanded to a corresponding 6-bit block, i.e., per 4-bit block, 2 more bits are added.



This process results in expansion as well as a permutation of the input bit while creating output. The key transformation process compresses the 56-bit key to 48 bits. Then the expansion permutation process expands the 32-bit RPT to 48-bits. Now the 48-bit key is XOR with 48-bit RPT and the resulting output is given to the next step, which is the S-Box substitution.

**AES**

Advanced Encryption Standard (AES) is a specification for the encryption of electronic data established by the U.S National Institute of Standards and Technology (NIST) in 2001. AES is widely used today as it is a much stronger than DES and triple DES despite being harder to implement.

**Points to remember**

* AES is a block cipher.
* The key size can be 128/192/256 bits.
* Encrypts data in blocks of 128 bits each.

That means it takes 128 bits as input and outputs 128 bits of encrypted cipher text as output. AES relies on substitution-permutation network principle which means it is performed using a series of linked operations which involves replacing and shuffling of the input data.

**Working of the cipher :**

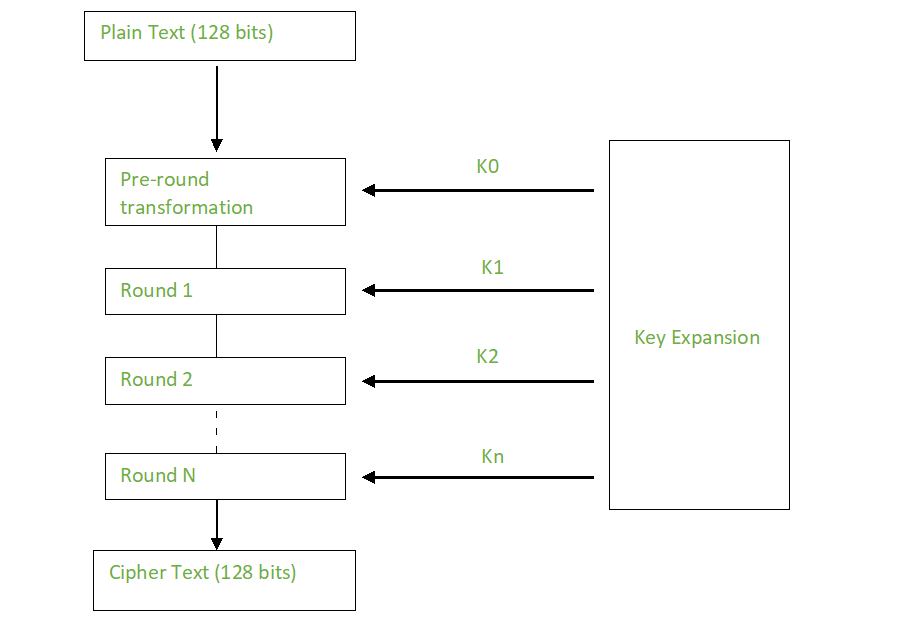
AES performs operations on bytes of data rather than in bits. Since the block size is 128 bits, the cipher processes 128 bits (or 16 bytes) of the input data at a time.

**The number of rounds depends on the key length as follows :**

* 128 bit key – 10 rounds
* 192 bit key – 12 rounds
* 256 bit key – 14 rounds

**Creation of Round keys :**

A Key Schedule algorithm is used to calculate all the round keys from the key. So the initial key is used to create many different round keys which will be used in the corresponding round of the encryption.



**Encryption** :

AES considers each block as a 16 byte (4 byte x 4 byte = 128 ) grid in a column major arrangement.

**[ b0 | b4 | b8 | b12 |**

**| b1 | b5 | b9 | b13 |**

**| b2 | b6 | b10| b14 |**

| b3 | b7 | b11| b15 ]

**Each round comprises of 4 steps :**

* SubBytes
* ShiftRows
* MixColumns

**Add Round Key**

The last round doesn’t have the MixColumns round.

The SubBytes does the substitution and ShiftRows and MixColumns performs the permutation in the algorithm.

**SubBytes :**

This step implements the substitution.

In this step each byte is substituted by another byte. Its performed using a lookup table also called the S-box. This substitution is done in a way that a byte is never substituted by itself and also not substituted by another byte which is a compliment of the current byte. The result of this step is a 16 byte (4 x 4 ) matrix like before.

**The next two steps implement the permutation.**

**ShiftRows :**

This step is just as it sounds. Each row is shifted a particular number of times.

* The first row is not shifted
* The second row is shifted once to the left.
* The third row is shifted twice to the left.
* The fourth row is shifted thrice to the left.

**(A left circular shift is performed.)**

[ b0 | b1 | b2 | b3 ] [ b0 | b1 | b2 | b3 ]

| b4 | b5 | b6 | b7 | -> | b5 | b6 | b7 | b4 |

| b8 | b9 | b10 | b11 | | b10 | b11 | b8 | b9 |

[ b12 | b13 | b14 | b15 ] [ b15 | b12 | b13 | b14 ]

**MixColumns :**

This step is basically a matrix multiplication. Each column is multiplied with a specific matrix and thus the position of each byte in the column is changed as a result.

**This step is skipped in the last round.**

[ c0 ] = [ 2 3 1 1 ] [ b0 ]

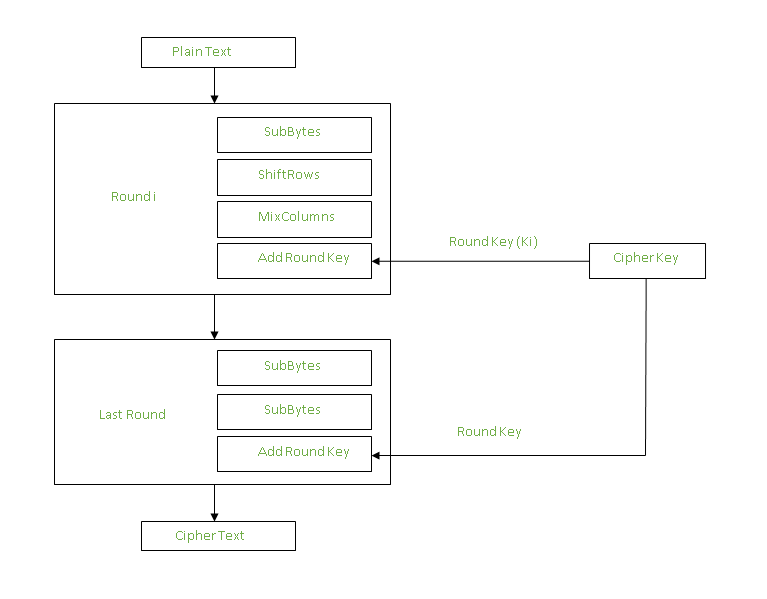
| c1 | = | 1 2 3 1 | | b1 |

| c2 | = | 1 1 2 3 | | b2 |

[ c3 ] = [ 3 1 1 2 ] [ b3 ]

**Add Round Keys :**

Now the resultant output of the previous stage is XOR-ed with the corresponding round key. Here, the 16 bytes is not considered as a grid but just as 128 bits of data.



After all these rounds 128 bits of encrypted data is given back as output. This process is repeated until all the data to be encrypted undergoes this process.

**Decryption :**

The stages in the rounds can be easily undone as these stages have an opposite to it which when performed reverts the changes.Each 128 blocks goes through the 10,12 or 14 rounds depending on the key size.

**The stages of each round in decryption is as follows :**

**Add round key**

Inverse MixColumns

**ShiftRows**

**Inverse SubByte**

The decryption process is the encryption process done in reverse so i will explain the steps with notable differences.

**Inverse MixColumns :**

This step is similar to the MixColumns step in encryption, but differs in the matrix used to carry out the operation.

[ b0 ] [ 14 11 13 9 ] [ c0 ]

| b1 | = | 9 14 11 13 | | c1 |

| b2 | | 13 9 14 11 | | c2 |

[ b3 ] [ 11 13 9 14 ] [ c3 ]

**Inverse SubBytes :**

Inverse S-box is used as a lookup table and using which the bytes are substituted during decryption.

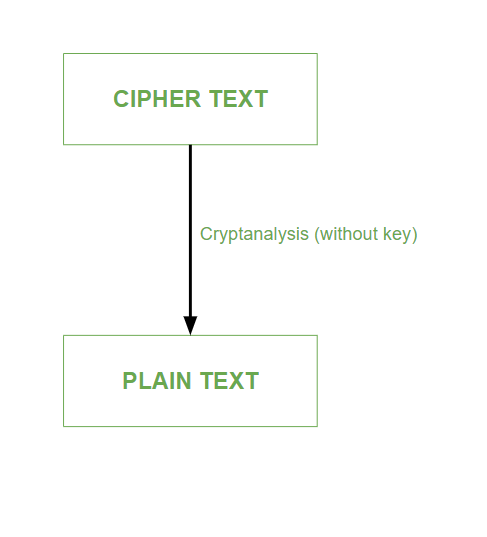
**Summary :**

AES instruction set is now integrated into the CPU (offers throughput of several GB/s)to improve the speed and security of applications that use AES for encryption and decryption. Even though its been 20 years since its introduction we have failed to break the AES algorithm as it is infeasible even with the current technology. Till date the only vulnerability remains in the implementation of the algorithm.

**Differential and Linear Cryptanalysis**

Cryptanalysis is the process of transforming or decoding communications from non-readable to readable format without having access to the real key. OR we may say it is the technique of retrieving the plain text of the communication without having access to the key. Cryptoanalysis is the art, science, or practice of decrypting encrypted messages. The secret key used for encryption and decoding is considered to be unknown to the cryptologists, mathematicians, and other scientists participating in the process. In contrast to a brute force attack, this form of analysis seeks vulnerabilities in a cryptosystem.

Cryptanalysis frequently comprises a direct evaluation of the cryptosystem in use, which is essentially an advanced concentrated mathematical attempt at decryption utilizing knowledge about the encryption scheme that is already available. They can employ intercepted encrypted messages (ciphertext), intercepted complete, partial, likely, or similar original messages (plaintext), or information (encrypted or original) that is known to be used adaptively in subsequent trials**.**



Cryptanalysis is used to break cryptographic security systems and gain access to the contents of the encrypted messages, even if the cryptographic key is unknown.

**Types of Cryptanalytic Attacks:**

**1. Ciphertext only attack:**

In this type of cryptanalytic attack, the attacker has the knowledge of only the ciphertext.

* The attacker has to detect the plain text using the ciphertext only.
* This type of attack is not very easy to be implemented.

**2. Known plain text only attack:**

In this type of cryptanalytic attack, the attacker has the knowledge of some plain text as well as ciphertext.

* The attacker tries to decrypt the messages using these two.
* This type of attack is somewhat easy to implement.
* Different Forms of Cryptanalysis:
* Cryptanalysis basically has two forms:

**1. Linear Cryptanalysis:**

Linear cryptanalysis is a general type of cryptanalysis based on discovering affine approximations to a cipher’s action in cryptography. Block and stream ciphers have both been subjected to attacks. Linear cryptanalysis is one of the two most common attacks against block ciphers, with differential cryptanalysis being the other.

**2. Differential Cryptanalysis:**

Differential cryptanalysis is a sort of cryptanalysis that may be used to decrypt both block and stream ciphers, as well as cryptographic hash functions. In the widest sense, it is the study of how alterations in information intake might impact the following difference at the output. In the context of a block cipher, it refers to a collection of strategies for tracking differences across a network of transformations, finding where the cipher displays non-random behavior, and using such attributes to recover the secret key (cryptography key).

**Difference between Linear Cryptanalysis and Differential Cryptanalysis**

|  |  |
| --- | --- |
| **Linear Cryptanalysis** | **Differential Cryptanalysis** |
| Linear cryptanalysis was basically invented by Matsui and Yamagishi in the year 1992. | Differential cryptanalysis was first defined in the year 1990 by Eli Biham and Adi Shamir. |
| Linear cryptanalysis always works on a single bit (one bit at a time). | Differential cryptanalysis can work on multiple bits at a time. |
| In the case of Linear cryptanalysis, ciphertext attack is a very big disadvantage. | In the case of differential cryptanalysis plain text attack is a very big disadvantage. |
| The use of linear cryptanalysis is to figure out what is the linear relationship present between some plaintext bits, ciphertext bits, and unknown key bits very easily. | The use of differential cryptanalysis is to get clues about some critical bits, reducing the need for an extensive search. |
| Subsets of input attributes refer to the internal structures of a single input. | The underlying structure of each individual input is unimportant in this case since the input attributes are differential. |
| The cryptanalyst decrypts each ciphertext using all available subkeys and analyses the resultant intermediate ciphertext to determine the random outcome for one encryption cycle. | After several encryption rounds, Cryptanalyst analyses the changes in the intermediate ciphertext obtained. The practice of combining assaults is known as differential linear cryptanalysis. |

**S-box design**

1. The substitution function of a set of 8 S- boxes, each of which accepts 6 bits as input and produces 4 bits as output.
2. The main characteristics of S box is its size.
3. The n × m S box has n inputs and m outputs bits. e.g DES has 6 × 4 S boxes.
4. Larger the S box more is it resistant to the crypto analysis. However if n is large, the look up table is large, the look up table is large. Hence a limit of n equals to about 8 to 10 is usually imposed.
5. Also if S box is large, it is more difficult to design it properly.

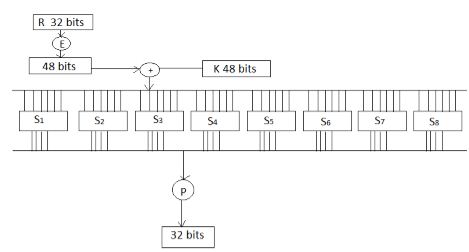


Figure 4.9 S Box DesignFigure 4.9 S Box Design

**The S- Box Design Criteria**

1. No output bits of any S box should be close to a linear function of input bits.
2. Each row of a S box should include all possible bit combinations.
3. If 2 inputs to an S box differs in exactly, the output must in exactly in at least 2 bits.
4. If 2 inputs to a S box differ in two middle bits, the output must differ in at least 2 bits.
5. If 2 inputs to a S box differ in 2 first bits are identical in their last 2 bits, the 2 outputs must not be same.
6. The above criteria make the algorithm robust and can’t be easily broken. If it also gives good confusion properties.

**The S-Box Design Approaches**

1. Random: It was some Pseudo random number generation or some table of random digits to generate the entries in ‘S’ box.
2. Random with Testing: Choose S box entries randomly then let the results against various criteria and throw away those that do not pass.
3. Human Made: This is a manual approach with only simple maths to support it ( Not Suitable for large S box)
4. Math Made: Generate S box according to mathematical principles.

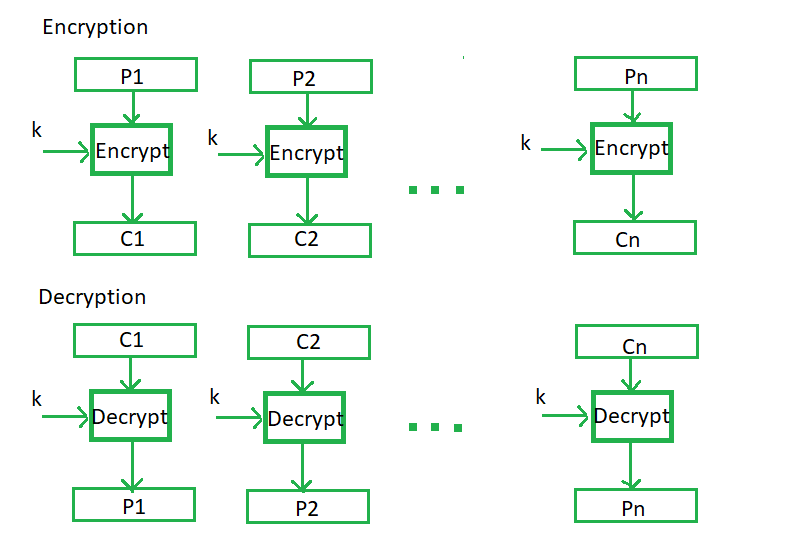
**Block Cipher modes of Operation**

Encryption algorithms are divided into two categories based on the input type, as a block cipher and stream cipher. Block cipher is an encryption algorithm that takes a fixed size of input say b bits and produces a ciphertext of b bits again. If the input is larger than b bits it can be divided further. For different applications and uses, there are several modes of operations for a block cipher.

**Electronic Code Book (ECB) –**

Electronic code book is the easiest block cipher mode of functioning. It is easier because of direct encryption of each block of input plaintext and output is in form of blocks of encrypted ciphertext. Generally, if a message is larger than b bits in size, it can be broken down into a bunch of blocks and the procedure is repeated.

**Procedure of ECB is illustrated below:**



**Advantages of using ECB –**

* Parallel encryption of blocks of bits is possible, thus it is a faster way of encryption.
* Simple way of the block cipher.

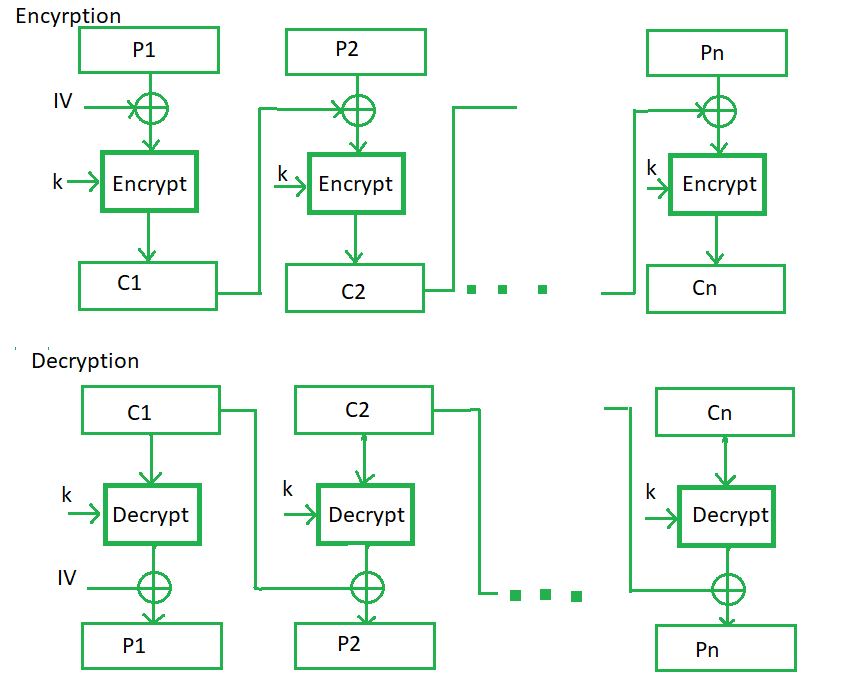
**Disadvantages of using ECB –**

* Prone to cryptanalysis since there is a direct relationship between plaintext and ciphertext.

**Cipher Block Chaining –**

Cipher block chaining or CBC is an advancement made on ECB since ECB compromises some security requirements. In CBC, the previous cipher block is given as input to the next encryption algorithm after XOR with the original plaintext block. In a nutshell here, a cipher block is produced by encrypting an XOR output of the previous cipher block and present plaintext block.

**The process is illustrated here:**



**Advantages of CBC –**

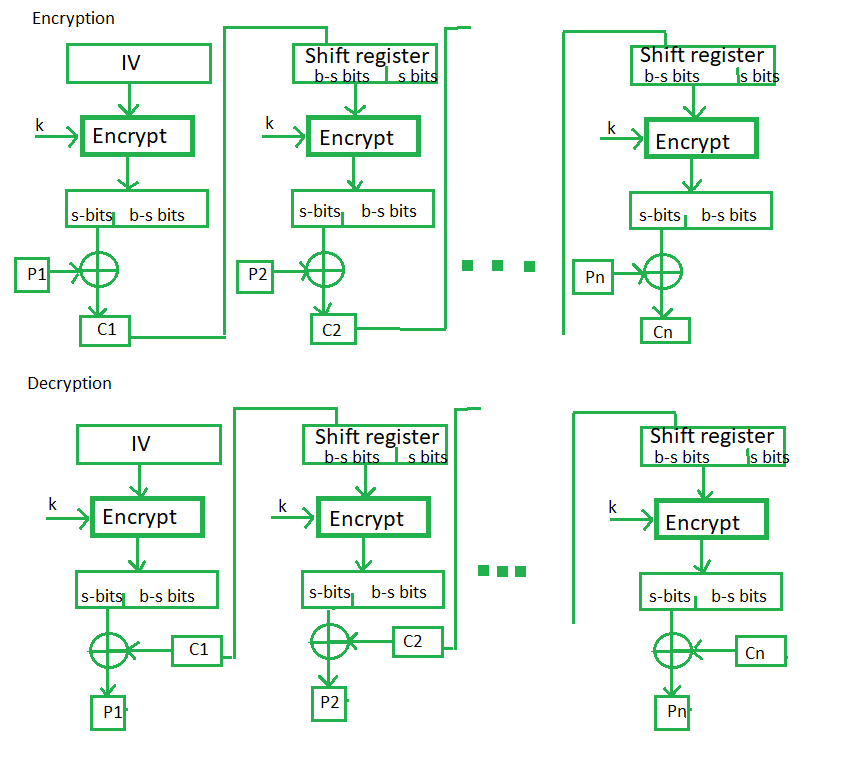
* CBC works well for input greater than b bits.
* CBC is a good authentication mechanism.
* Better resistive nature towards cryptanalysis than ECB.

**Disadvantages of CBC –**

* Parallel encryption is not possible since every encryption requires a previous cipher.

**Cipher Feedback Mode (CFB) –**

In this mode the cipher is given as feedback to the next block of encryption with some new specifications: first, an initial vector IV is used for first encryption and output bits are divided as a set of s and b-s bits.The left-hand side s bits are selected along with plaintext bits to which an XOR operation is applied. The result is given as input to a shift register having b-s bits to lhs,s bits to rhs and the process continues. The encryption and decryption process for the same is shown below, both of them use encryption algorithms.

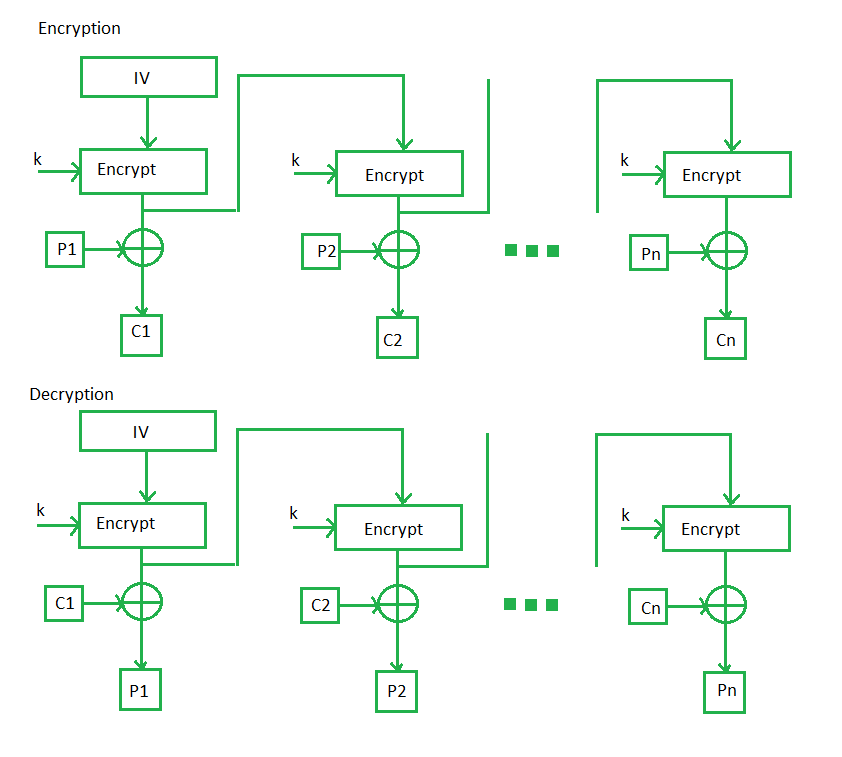


**Advantages of CFB –**

* Since, there is some data loss due to the use of shift register, thus it is difficult for applying cryptanalysis.

**Output Feedback Mode –**

The output feedback mode follows nearly the same process as the Cipher Feedback mode except that it sends the encrypted output as feedback instead of the actual cipher which is XOR output. In this output feedback mode, all bits of the block are sent instead of sending selected s bits. The Output Feedback mode of block cipher holds great resistance towards bit transmission errors. It also decreases the dependency or relationship of the cipher on the plaintext.



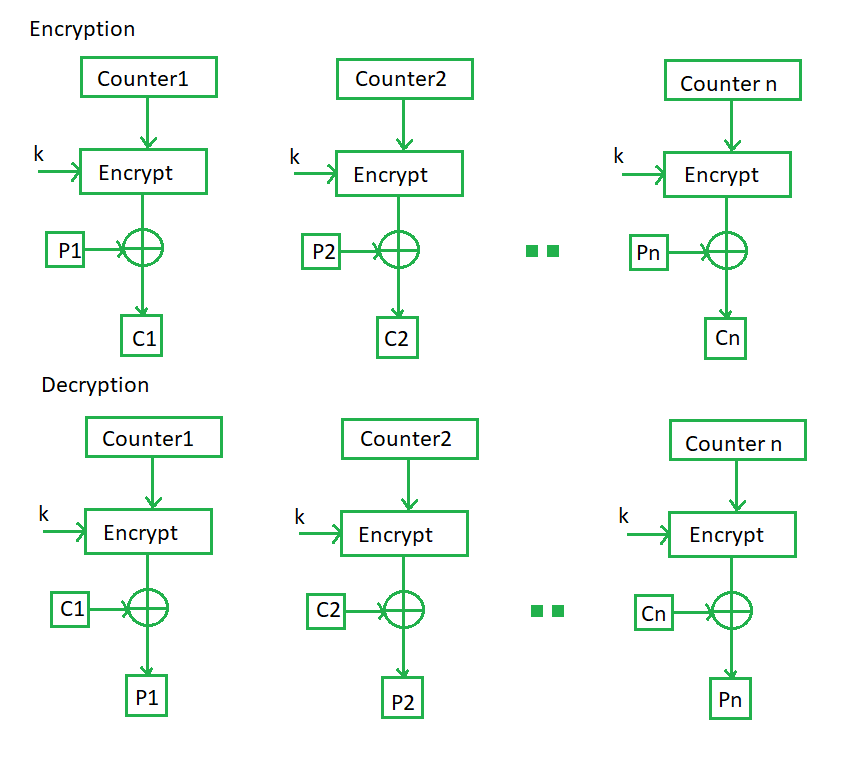
**Advantages of OFB –**

In the case of CFB, a single bit error in a block is propagated to all subsequent blocks. This problem is solved by OFB as it is free from bit errors in the plaintext block.

**Counter Mode –**

The Counter Mode or CTR is a simple counter-based block cipher implementation. Every time a counter-initiated value is encrypted and given as input to XOR with plaintext which results in ciphertext block. The CTR mode is independent of feedback use and thus can be implemented in parallel.

**Its simple implementation is shown below:**



**Advantages of Counter –**

* Since there is a different counter value for each block, the direct plaintext and ciphertext relationship is avoided. This means that the same plain text can map to different ciphertext.
* Parallel execution of encryption is possible as outputs from previous stages are not chained as in the case of CBC.