Error-Correcting Code

An error-correcting code is an algorithm for expressing a sequence of numbers such that any errors which are introduced can be detected and corrected (within certain limitations) based on the remaining numbers. The study of error-correcting codes and the associated mathematics is known as coding theory.

Error detection is much simpler than error correction, and one or more "check" digits are commonly embedded in credit card numbers in order to detect mistakes. Early space probes like Mariner used a type of error-correcting code called a block code, and more recent space probes use convolution codes. Error-correcting codes are also used in CD players, high speed modems, and cellular phones. Modems use error detection when they compute checksums, which are sums of the digits in a given transmission modulo some number. The ISBN used to identify books also incorporates a check digit.

A powerful check for 13 digit numbers consists of the following. Write the number as a string of digits . Take and double. Now add the number of digits in odd positions which are to this number. Now add . The check number is then the number required to bring the last digit to 0. This scheme detects all single digit errors and all transpositions of adjacent digits except 0 and 9.

Let denote the maximal number of (0,1)-vectors having the property that any two of the set differ in at least places. The corresponding vectors can correct errors. is the number of s with precisely 1s (Sloane and Plouffe 1995). Since it is not possible for -vectors to differ in places and since -vectors which differ in all places partition into disparate sets of two,

(1)

Values of can be found by labeling the (0,1)- -vectors, finding all unordered pairs of -vectors which differ from each other in at least places, forming a graph from these unordered pairs, and then finding the clique number of this graph. Unfortunately, finding the size of a clique for a given graph is an NP-complete problem.

Caesar Cipher Technique

The Caesar cipher is the simplest and oldest method of cryptography. The Caesar cipher method is based on a mono-alphabetic cipher and is also called a shift cipher or additive cipher. Julius Caesar used the shift cipher (additive cipher) technique to communicate with his officers. For this reason, the shift cipher technique is called the Caesar cipher. The Caesar cipher is a kind of replacement (substitution) cipher, where all letter of plain text is replaced by another letter.

Let's take an example to understand the Caesar cipher, suppose we are shifting with 1, then A will be replaced by B, B will be replaced by C, C will be replaced by D, D will be replaced by C, and this process continues until the entire plain text is finished.

Caesar ciphers is a weak method of cryptography. It can be easily hacked. It means the message encrypted by this method can be easily decrypted.

Plaintext: It is a simple message written by the user.

Ciphertext: It is an encrypted message after applying some technique.

The formula of encryption is:

$E_n(x) = (x + n) \bmod 26$

The formula of decryption is:

$D_n(x) = (xi - n) \bmod 26$

If any case (Dn) value becomes negative (-ve), in this case, we will add 26 in the negative value.

Where,

E denotes the encryption
D denotes the decryption
x denotes the letters value
n denotes the key value (shift value)

Error-detecting codes are a sequence of numbers generated by specific procedures for detecting errors in data that has been transmitted over computer networks.

When bits are transmitted over the computer network, they are subject to get corrupted due to interference and network problems. The corrupted bits leads to spurious data being received by the receiver and are called errors.

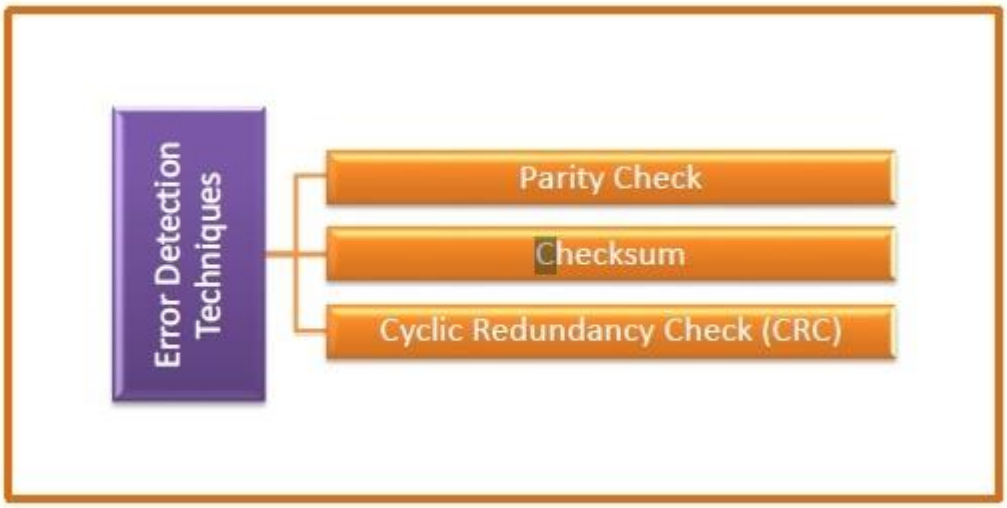
Error – detecting codes ensures messages to be encoded before they are sent over noisy channels. The encoding is done in a manner so that the decoder at the receiving end can detect whether there are errors in the incoming signal with high probability of success.

Features of Error Detecting Codes

- Error detecting codes are adopted when backward error correction techniques are used for reliable data transmission. In this method, the receiver sends a feedback message to the sender to inform whether an error-free message has been received or not. If there are errors, then the sender retransmits the message.
- Error-detecting codes are usually block codes, where the message is divided into fixed-sized blocks of bits, to which redundant bits are added for error detection.
- Error detection involves checking whether any error has occurred or not. The number of error bits and the type of error does not matter.

Error Detection Techniques

There are three main techniques for detecting errors



Parity Check

Parity check is done by adding an extra bit, called parity bit to the data to make number of 1s either even in case of even parity, or odd in case of odd parity.

While creating a frame, the sender counts the number of 1s in it and adds the parity bit in following way

- In case of even parity: If number of 1s is even then parity bit value is 0. If number of 1s is odd then parity bit value is 1.
- In case of odd parity: If number of 1s is odd then parity bit value is 0. If number of 1s is even then parity bit value is 1.

On receiving a frame, the receiver counts the number of 1s in it. In case of even parity check, if the count of 1s is even, the frame is accepted, otherwise it is rejected. Similar rule is adopted for odd parity check.

Parity check is suitable for single bit error detection only.

Checksum

In this error detection scheme, the following procedure is applied

- Data is divided into fixed sized frames or segments.
- The sender adds the segments using 1's complement arithmetic to get the sum. It then complements the sum to get the checksum and sends it along with the data frames.
- The receiver adds the incoming segments along with the checksum using 1's complement arithmetic to get the sum and then complements it.

- If the result is zero, the received frames are accepted; otherwise they are discarded.

Cyclic Redundancy Check (CRC)

Cyclic Redundancy Check (CRC) involves binary division of the data bits being sent by a predetermined divisor agreed upon by the communicating system. The divisor is generated using polynomials.

- Here, the sender performs binary division of the data segment by the divisor. It then appends the remainder called CRC bits to the end of data segment. This makes the resulting data unit exactly divisible by the divisor.
- The receiver divides the incoming data unit by the divisor. If there is no remainder, the data unit is assumed to be correct and is accepted. Otherwise, it is understood that the data is corrupted and is therefore rejected.

In cryptography, a **substitution cipher** is a method of encrypting in which units of plaintext are replaced with the ciphertext, in a defined manner, with the help of a key; the "units" may be single letters (the most common), pairs of letters, triplets of letters, mixtures of the above, and so forth. The receiver deciphers the text by performing the inverse substitution process to extract the original message.

Substitution ciphers can be compared with transposition ciphers. In a transposition cipher, the units of the plaintext are rearranged in a different and usually quite complex order, but the units themselves are left unchanged. By contrast, in a substitution cipher, the units of the plaintext are retained in the same sequence in the ciphertext, but the units themselves are altered.

There are a number of different types of substitution cipher. If the cipher operates on single letters, it is termed a **simple substitution cipher**; a cipher that operates on larger groups of letters is termed **polygraphic**. A **monoalphabetic cipher** uses fixed substitution over the entire message, whereas a **polyalphabetic cipher** uses a number of substitutions at different positions in the message, where a unit from the plaintext is mapped to one of several possibilities in the ciphertext and vice versa.

Drawbacks

The first ever published description of how to crack simple substitution ciphers was given by Al-Kindi in *A Manuscript on Deciphering Cryptographic Messages* written around 850 CE. The method he described is now known as frequency analysis.

What is a Playfair Cipher

The Playfair cipher encryption technique can be used to encrypt or encode a message. It operates exactly like typical encryption. The only difference is that it encrypts a digraph, or a pair of two letters, as opposed to a single letter.

An initial 5×5 matrix key table is created. The plaintext encryption key is made out of the matrix's alphabetic characters. Be mindful that you shouldn't repeat the letters. There are 26 alphabets, however, there are only 25 spaces in which we can place a letter. The matrix will delete the extra letter because there is an excess of one letter (typically J). Despite this, J is there in the plaintext before being changed to I.

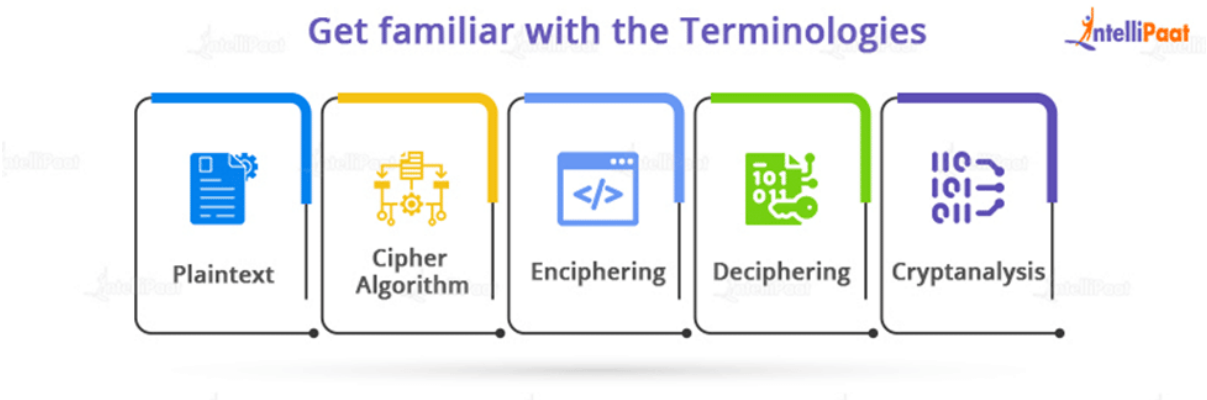
History of Playfair Cipher

The Playfair cipher is the earliest and best-known digraph substitution cipher to use symmetry encryption. Charles Wheatstone developed the cipher in 1854, and Lord Playfair, who supported its use, gave it its name. In contrast to a simple substitution cipher, which simply encrypts single letters, the Playfair cipher technique encrypts digraphs or parts of letters.

The Playfair cipher is rapid and doesn't require any additional tools to use. British forces employed it tactically in World War I, the Second Boer War, and World War II, as did Australian forces. The major goal of the encryption was to safeguard non-critical yet important information during actual combat. By the time the opposition's cryptanalysts were able to decrypt it, the data was useless.

Want to learn more and make a career in the field of Cyber Security, make sure to check out Cyber Security Course.

Get familiar with the Terminologies

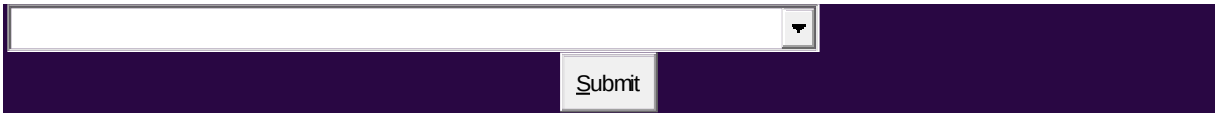


- Plaintext: The entire conversation must be encrypted. It's also referred to as a message. It is an encrypted or ciphertext message.
- The cipher algorithm converts plaintext into ciphertext, it serves as the decryption or encryption key for text. Only the sender and the recipient are aware of it. The key-table or key-matrix is filled in character by character.
- Enciphering is the process of converting plaintext into ciphertext.
- Deciphering is the process of separating plaintext from the ciphertext.
- Cryptanalysis is the study of non-key decoding algorithms and concepts for ciphertext.

Get a deep understanding of the Access Control List through this blog!

Playfair Cipher's Advantages:

- If we carefully study the algorithm, we can see that each stage of the process results in a distinct ciphertext, which makes it more challenging for the cryptanalyst.
- Brute force attacks do not affect it.
- Cryptanalysis is not possible (decode a cipher without knowing the key).
- eliminates the flaw in the simple Playfair square cipher.
- Making the substitution is easy.



Playfair Cipher Restrictions:

The Playfair cipher is constrained by the following:

- Only 25 alphabets are supported.
- It is incompatible with characters that number.
- Only capital or lowercase letters are accepted.
- Special characters like spaces, newlines, punctuation, etc. are not allowed.

Importance of Playfair Cipher

Due to its relative complexity to other ciphers at the time during World Wars I and II, the Playfair cipher became very well-known. Furthermore, neither the encryption nor decryption of the data required any specialized tools or methods.

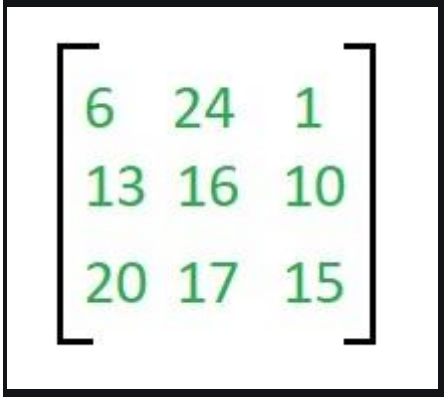
However, with the advent of computers, the Playfair cipher was no longer used since computers can easily apply break codes to decrypt Playfair ciphers.

As a result, with the advancement of digital encryption and the passage of time, the Playfair cipher was no longer a feasible form of message encoding due to the possibility of data getting into the wrong hands. Therefore, using the Playfair cipher for corporate enterprises is not recommended.

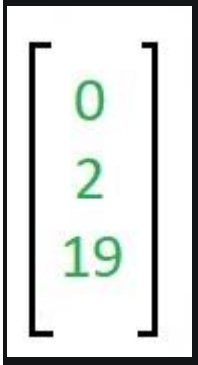
Hill cipher is a polygraphic substitution cipher based on linear algebra.Each letter is represented by a number modulo 26. Often the simple scheme A = 0, B = 1, ..., Z = 25 is used, but this is not an essential feature of the cipher. To encrypt a message, each block of n letters (considered as an n-component 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:

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:

$$\begin{bmatrix} 6 & 24 & 1 \\ 13 & 16 & 10 \\ 20 & 17 & 15 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 19 \end{bmatrix} = \begin{bmatrix} 67 \\ 222 \\ 319 \end{bmatrix} \equiv \begin{bmatrix} 15 \\ 14 \\ 7 \end{bmatrix} \pmod{26}$$

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:

$$\begin{bmatrix} 6 & 24 & 1 \\ 13 & 16 & 10 \\ 20 & 17 & 15 \end{bmatrix}^{-1} \equiv \begin{bmatrix} 8 & 5 & 10 \\ 21 & 8 & 21 \\ 21 & 12 & 8 \end{bmatrix} \pmod{26}$$

For the previous Ciphertext 'POH':

$$\begin{bmatrix} 8 & 5 & 10 \\ 21 & 8 & 21 \\ 21 & 12 & 8 \end{bmatrix} \begin{bmatrix} 15 \\ 14 \\ 7 \end{bmatrix} \equiv \begin{bmatrix} 260 \\ 574 \\ 539 \end{bmatrix} \equiv \begin{bmatrix} 0 \\ 2 \\ 19 \end{bmatrix} \pmod{26}$$

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.

Given a plain-text message and a numeric key, cipher/de-cipher the given text using Rail Fence algorithm.
The rail fence cipher (also called a zigzag cipher) is a form of transposition cipher. It derives its name from the way in which it is encoded.
Examples:

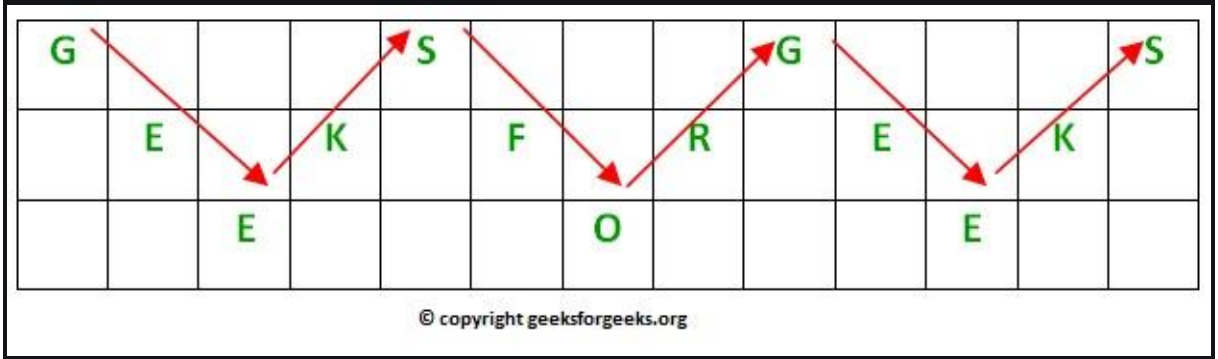
Recommended: Please try your approach on ***{IDE}*** first, before moving on to the solution.

Encryption
In a transposition cipher, the order of the alphabets is re-arranged to obtain the cipher-text.

- In the rail fence cipher, the plain-text is written downwards and diagonally on successive rails of an imaginary fence.

- When we reach the bottom rail, we traverse upwards moving diagonally, after reaching the top rail, the direction is changed again. Thus the alphabets of the message are written in a zig-zag manner.
- After each alphabet has been written, the individual rows are combined to obtain the cipher-text.

For example, if the message is “GeeksforGeeks” and the number of rails = 3 then cipher is prepared as:



.’.Its encryption will be done row wise i.e. GSGSEKFREKEOE

Decryption

As we’ve seen earlier, the number of columns in rail fence cipher remains equal to the length of plain-text message. And the key corresponds to the number of rails.

- Hence, rail matrix can be constructed accordingly. Once we’ve got the matrix we can figure-out the spots where texts should be placed (using the same way of moving diagonally up and down alternatively).
- Then, we fill the cipher-text row wise. After filling it, we traverse the matrix in zig-zag manner to obtain the original text.

Implementation:

Let cipher-text = “GsGsekfrek eoe” , and Key = 3

- Number of columns in matrix = len(cipher-text) = 13
- Number of rows = key = 3

Hence original matrix will be of 3*13 , now marking places with text as '*' we get

Below is a program to encrypt/decrypt the message using the above algorithm.

Given a plain-text message and a numeric key, cipher/de-cipher the given text using Columnar Transposition Cipher

The Columnar Transposition Cipher is a form of transposition cipher just like Rail Fence Cipher. Columnar Transposition involves writing the plaintext out in rows, and then reading the ciphertext off in columns one by one.

Examples:

Recommended: Please try your approach on ***{IDE}*** first, before moving on to the solution.

Encryption

In a transposition cipher, the order of the alphabets is re-arranged to obtain the cipher-text.

1. 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.
2. Width of the rows and the permutation of the columns are usually defined by a keyword.
3. For example, the word HACK is of length 4 (so the rows are of length 4), and the permutation is defined by the alphabetical order of the letters in the keyword. In this case, the order would be “3 1 2 4”.
4. Any spare spaces are filled with nulls or left blank or placed by a character (Example: _).
5. Finally, the message is read off in columns, in the order specified by the keyword.

Encryption

Given text = Geeks for Geeks
Keyword = HACK **Length of Keyword** = 4 (no of rows) **Order of Alphabets in HACK** = 3124

H	A	C	K
3	1	2	4
G	e	e	k
s	_	f	o
r	_	G	e
e	k	s	_

Print Characters of column 1,2,3,4
Encrypted Text = e kefGsGsrekoe_

Decryption

1. To decipher it, the recipient has to work out the column lengths by dividing the message length by the key length.
2. Then, write the message out in columns again, then re-order the columns by reforming the key word.

product cipher, data encryption scheme in which the ciphertext produced by encrypting a plaintext document is subjected to further encryption.
By combining two or more simple transposition ciphers or substitution ciphers, a more secure encryption may result.

In the days of manual cryptography, product ciphers were a useful device for cryptographers, and in fact double transposition or product ciphers on key word-based rectangular matrices were widely used. There was also some use of a class of product ciphers known as fractionation systems, wherein a substitution was first made from symbols in the plaintext to multiple symbols (usually pairs, in which case the cipher is called a biliteral cipher) in the ciphertext, which was then encrypted by a final transposition, known as superencryption. One of the most famous field ciphers of all time was a fractionation system, the ADFGVX cipher employed by the German army during World War I. This system used a 6 × 6 matrix to substitution-encrypt the 26 letters and 10 digits into pairs of the symbols A, D, F, G, V, and X. The resulting biliteral cipher was then written into a rectangular array and route encrypted by reading the columns in the order indicated by a key word, as illustrated in the figure.

cryptography, a **ciphertext-only attack** (**COA**) or **known ciphertext attack** is an attack model for cryptanalysis where the attacker is assumed to have access only to a set of ciphertexts. While the attacker has no channel providing access to the plaintext prior to encryption, in all practical ciphertext-only attacks, the attacker still has some knowledge of the plaintext. For instance, the attacker might know the language in which the plaintext is written or the expected statistical distribution of characters in the plaintext. Standard protocol data and messages are commonly part of the plaintext in many deployed systems and can usually be guessed or known efficiently as part of a ciphertext-only attack on these systems.

Attack[edit]

The attack is completely successful if the corresponding plaintexts can be deduced, or even better, the key. The ability to obtain any information at all about the underlying plaintext beyond what was pre-known to the attacker is still considered a success. For example, if an adversary is sending ciphertext continuously to maintain traffic-flow security, it would be very useful to be able to distinguish real messages from nulls. Even making an informed guess of the existence of real messages would facilitate traffic analysis.

In the history of cryptography, early ciphers, implemented using pen-and-paper, were routinely broken using ciphertexts alone. Cryptographers developed statistical techniques for attacking ciphertext, such as frequency analysis. Mechanical encryption devices such as Enigma made these attacks much more difficult (although, historically, Polish cryptographers were able to mount a successful ciphertext-only cryptanalysis of the Enigma by exploiting an insecure protocol for indicating the message settings). More advanced ciphertext-only attacks on the Enigma were mounted in Bletchley Park during World War II, by intelligently guessing plaintexts corresponding to intercepted ciphertexts.

Modern[edit]

Every modern cipher attempts to provide protection against ciphertext-only attacks. The vetting process for a new cipher design standard usually takes many years and includes exhaustive testing of large quantities of ciphertext for any statistical departure from random noise. *See:* Advanced Encryption Standard process. Also, the field of steganography evolved, in part, to develop methods like mimic functions that allow one piece of data to adopt the statistical profile of another. Nonetheless poor cipher usage or reliance on home-grown proprietary algorithms that have not been subject to thorough scrutiny has resulted in many computer-age encryption systems that are still subject to ciphertext-only attack. Examples include:

stream cipher, one byte is encrypted at a time while in block cipher ~128 bits are encrypted at a time.

Initially, a key(k) will be supplied as input to pseudorandom bit generator and then it produces a random 8-bit output which is treated as keystream.

The resulted keystream will be of size 1 byte, i.e., 8 bits.

1. Stream Cipher follows the sequence of pseudorandom number stream.
2. One of the benefits of following stream cipher is to make cryptanalysis more difficult, so the number of bits chosen in the Keystream must be long in order to make cryptanalysis more difficult.
3. By making the key more longer it is also safe against brute force attacks.
4. The longer the key the stronger security is achieved, preventing any attack.
5. Keystream can be designed more efficiently by including more number of 1s and 0s, for making cryptanalysis more difficult.
6. Considerable benefit of a stream cipher is, it requires few lines of code compared to block cipher.

Encryption :
For Encryption,

- Plain Text and Keystream produces Cipher Text (Same keystream will be used for decryption.).
- The Plaintext will undergo XOR operation with keystream bit-by-bit and produces the Cipher Text.

Example –

```
Plain Text : 10011001

Keystream  : 11000011
`````````````````````

Cipher Text : 01011010
```

**Decryption :**
For Decryption,

- Cipher Text and Keystream gives the original Plain Text (Same keystream will be used for encryption.).
- The Ciphertext will undergo XOR operation with keystream bit-by-bit and produces the actual Plain Text.

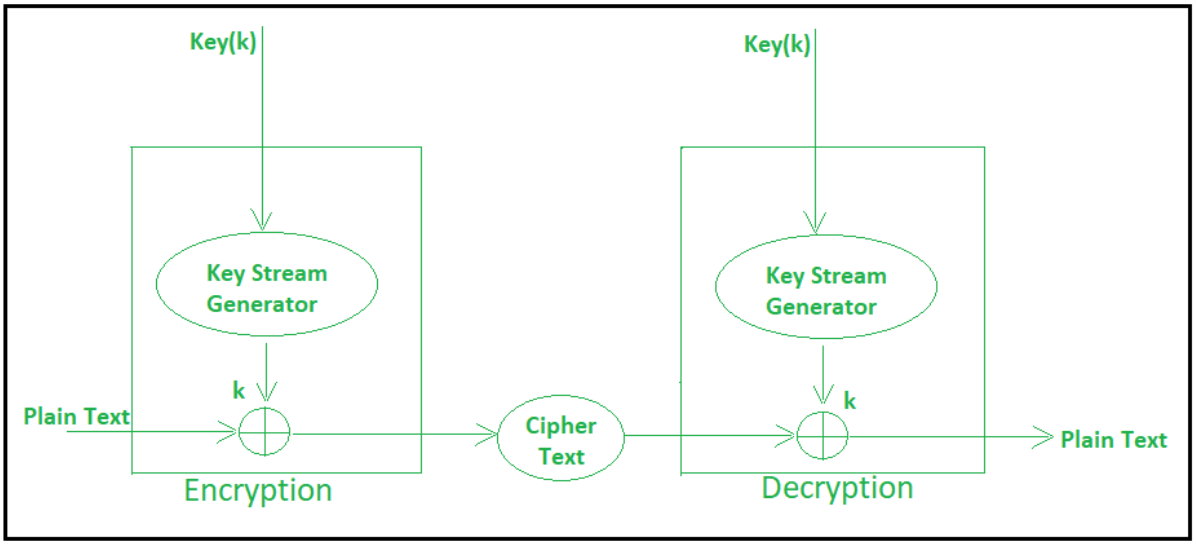
**Example –**

```
Cipher Text : 01011010

Keystream   : 11000011
`````````````````````

Plain Text  : 10011001
```

Decryption is just the reverse process of Encryption i.e. performing XOR with Cipher Text.


Key(k)
Key Stream Generator
k
Plain Text
Encryption
Cipher Text
Key(k)
Key Stream Generator
k
Plain Text
Decryption