Substitution Cipher Technique in Cryptography

Substitution technique is a classical encryption technique where the characters present in the **original message** are **replaced** by the **other characters or numbers or by symbols.** If the plain text (original message) is considered as the string of bits, then the substitution technique would replace bit pattern of plain text with the bit pattern of cipher text.

We will discuss some of the substitution techniques which will help us to understand the procedure of converting plain text o cipher text. In this section, we will study the following substitution techniques:

Substitution CipherTechnique:

1. Caesar Cipher
2. Monoalphabetic Cipher
3. Playfair Cipher
4. Hill Cipher
5. Polyalphabetic Cipher
6. One-Time Pad

1. Caesar Cipher

This the simplest substitution ciphers by Julius Caesar. In this substitution technique, to encrypt the plain text, each alphabet of the plain text is replaced by the alphabet three places further it. And to decrypt the cipher text each alphabet of cipher text is replaced by the alphabet three places before it.

Let us take a simple example:

Plain Text: meet me tomorrow

Cipher Text: phhw ph wrpruurz

Look at the example above, we have replaced, ‘m’ with ‘p’ which occur three places after, ‘m’. Similarly, ‘e’ is replaced with ‘h’ which occurs in three places after ‘e’.

Note: If we have to replace the letter ‘z’ then the next three alphabets counted after ‘z’ will be ‘a’ ‘b’ ‘c’. So, while counting further three alphabets if ‘z’ occurs it circularly follows ‘a’.

There are also some drawbacks of this simple substitution technique. If the hacker knows that the Caesar cipher is used then to perform brute force cryptanalysis, he has only to try 25 possible keys to decrypt the plain text.
The hacker is also aware of the encryption and decryption algorithm.

2. Monoalphabetic Cipher

Monoalphabetic cipher is a substitution cipher, where the cipher alphabet for each plain text alphabet is fixed, for the entire encryption.

In simple words, if the alphabet ‘p’ in the plain text is replaced by the cipher alphabet ‘d’. Then in the entire plain text wherever alphabet ‘p’ is used, it will be replaced by the alphabet ‘d’ to form the ciphertext.

3. Playfair Cipher

Playfair cipher is a substitution cipher which involves a 5X5 matrix. Let us discuss the technique of this Playfair cipher with the help of an example:

Plain Text: meet me tomorrow

Key: KEYWORD

Now, we have to convert this plain text to ciphertext using the given key. We will discuss the further process in steps.

Step 1: Create a 5X5 matrix and place the key in that matrix row-wise from left to right. Then put the remaining alphabets in the blank space.

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

Note: If a key has duplicate alphabets, then fill those alphabets only once in the matrix, and I & J should be kept together in the matrix even though they occur in the given key.

Step 2: Now, you have to break the plain text into a pair of alphabets.

Plain Text: meet me tomorrow

Pair: me et me to mo rx ro wz

Note

- Pair of alphabets must not contain the same letter. In case, pair has the same letter then break it and add ‘x’ to the previous letter. Like in our example letter ‘rr’ occurs in pair so, we have broken that pair and added ‘x’ to the first ‘r’.
- In case while making pair, the last pair has only one alphabet left then we add ‘z’ to that alphabet to form a pair as in our above example, we have added ‘z’ to ‘w’ because ‘w’ was left alone at last.
- If a pair has ‘xx’ then we break it and add ‘z’ to the first ‘x’, i.e. ‘xz’ and ‘x_’.

Step 3: In this step, we will convert plain text into ciphertext. For that, take the first pair of plain text and check for cipher alphabets for the corresponding in the matrix. To find cipher alphabets follow the rules below.

Note

- If both the alphabets of the pair occur in the **same row** replace them with the alphabet to their **immediate right**. If an alphabet of the pair occurs at extreme right then replace it with the first element of that row, i.e. the last element of the row in the matrix circularly follows the first element of the same row.
- If the alphabets in the pair occur in the **same column**, then replace them with the alphabet **immediate below** them. Here also, the last element of the column circularly follows the first element of the same column.
- If the alphabets in the pair are **neither in the same column and nor in the same row,** then the alphabet is replaced by the element in its own row and the corresponding column of the other alphabet of the pair.

Pair: me et me to mo rx ro wz

Cipher Text: go op go zc tc au ck oy

So, this is how we can convert a plain text to ciphertext using Playfair cipher. When compared with monoalphabetic cipher Playfair cipher is much more advanced. But still, it is easy to break.

4. Hill Cipher

Hill cipher is a polyalphabetic cipher introduced by Lester Hill in 1929. Let us discuss the technique of hill cipher.

Plain text: Binary

Key: HILL

Choose the key in such a way that it always forms a **square matrix**. With HILL as the key, we can form a 2×2 matrix.

Now, of plain text, you have to form a column vector of length similar to the key matrix. In our case, the key matrix is 2×2 then the column vectors of plain text would be 2×1.

The general equation to find cipher text using hill cipher is as follow:

C = KP mod 26

$$(c_1\ c_2) = \begin{pmatrix} k_1 & k_2 \\ k_3 & k_4 \end{pmatrix} \begin{pmatrix} p_1 \\ p2 \end{pmatrix} mod\ 26$$

For our example, our key matrix would be:

$$\begin{pmatrix} H & I \\ L & L \end{pmatrix}$$

And our plain text matrices of 2×1 will be as follow:

$$\begin{pmatrix} B \\ I \end{pmatrix} \begin{pmatrix} N \\ A \end{pmatrix} \begin{pmatrix} R \\ Y \end{pmatrix}$$

Now, we have to convert the key matrix and plain text matrices into numeric matrices. For that number the alphabets such as A=0, B=1, C=2, ……, Z=25. So, considering the alphabet numbering:

Key matrix will be:

$$K = \begin{pmatrix} 7 & 8 \\ 11 & 11 \end{pmatrix}$$

Plain text matrices would be:

$$\begin{pmatrix} 1 \\ 8 \end{pmatrix} \begin{pmatrix} 13 \\ 0 \end{pmatrix} \begin{pmatrix} 17 \\ 24 \end{pmatrix}$$

In the first calculation, we would get two cipher alphabets for plain text alphabet ‘B’ & ‘I’.

$$(c_1\ c_2) = \begin{pmatrix} 7 & 8 \\ 11 & 11 \end{pmatrix} \begin{pmatrix} 1 \\ 8 \end{pmatrix} mod\ 26$$

$$(c_1\ c_2) = \begin{pmatrix} 71 \\ 99 \end{pmatrix} mod\ 26$$

$$(c_1\ c_2) = \begin{pmatrix} 71 \\ 99 \end{pmatrix} mod\ 26$$

$$(c_1\ c_2) = \begin{pmatrix} 19 \\ 21 \end{pmatrix}$$

$$(c_1\ c_2) = \begin{pmatrix} T \\ V \end{pmatrix}$$

So, the cipher alphabet for plain text alphabet ‘B’ & ‘I’ is ‘T’ & ‘V’. Similarly, we have to calculate ciphertext for remaining plain text. And then accumulate them to form the ciphertext.

The calculated **ciphertext** for ‘**Binary**’ using hill cipher is ‘**TVNNZJ**’.

5. Polyalphabetic Cipher

Polyalphabetic cipher is far more secure than a monoalphabetic cipher. As monoalphabetic cipher maps a plain text symbol or alphabet to a ciphertext symbol and uses the same ciphertext symbol wherever that plain text occurs in the message.
But polyalphabetic cipher, each time replaces the plain text with the different ciphertext.

6. One-Time Pad

The one-time pad cipher suggests that the **key length** should be **as long as the plain text** to prevent the repetition of key. Along with that, the **key** should be **used** only **once** to encrypt and decrypt the single message after that the key should be discarded.

Onetime pad suggests a new key for each new message and of the same length as a new message. Now, let us see the one-time pad technique to convert plain text into ciphertext. Assume our plain text and key be:

Plain text: Binary

Key: Cipher

Now again convert the plain text and key into the numeric form. For that number the alphabets such as A=0, B=1, C=2, ……, Z=25. So, our plain text and key in numeric form would be:

Plain text: 1 8 13 0 17 24

Key: 2 8 15 7 4 17

Now, you have to add the number of the plain text alphabet, to the number of its corresponding key alphabet. That means, for this example, we will add:

B+C = 1+2 = 2

I+I = 8+8 = 16

N+P = 13+15 = 28

A+H = 0+7 = 7

R+E = 17+4 = 21

Y+R = 24+17 = 41

The resultant ciphertext numbers we get are (2, 16, 28, 7, 21, and 41)

If the addition of any plain text number and the key number is >26, then subtract only that particular number from 26. We have the addition of two pair of plain text number and a key number, greater than 26, i.e. N+P=28 & Y+R=41.

Subtract them by 26.

N+P = 28 – 26 = 2

Y+R = 41 – 26 = 15

So, the final **ciphertext numbers are (2, 16, 2, 7, 21, and 1)**. Now convert this number to alphabets assuming A to be numbered 0 and B to be 1…..Z to 25.

Ciphertext: Cqchvb.

In this way, we can convert plain text to cipher text using a one-time pad.

So, this is all about the substitution cipher techniques. It has a monoalphabetic cipher and polyalphabetic cipher technique. Substitution technique is also called classical substitution technique.

Transposition Cipher Technique in Cryptography

In cryptography, a transposition cipher is methods of encryption by which the positions held by units of plaintext (which are commonly characters or groups of characters) are shifted according to a regular system, so that the ciphertext constitutes a permutation of the plaintext. That is, the order of the units is changed. Mathematically a bijective function is used on the characters' positions to encrypt and an inverse function to decrypt.

Following are some implementations.

Contents

1. Rail Fence Cipher

The Rail Fence cipher is a form of transposition cipher that gets its name from the way in which it is encoded. In the rail fence cipher, the plaintext is written downwards on successive "rails" of an imaginary fence, then moving up when we get to the bottom. The message is then read off in rows. For example, using three "rails" and a message of 'WE ARE DISCOVERED. FLEE AT ONCE', the cipherer writes out:

W . . . E . . . C . . . R . . . L . . . T . . . E

.E . R . D . S . O . E . E . F . E . A . O . C.

. . A . . . I . . . V . . . D . . . E . . . N . .

Then reads off:

WECRL TEERD SOEEF EAOCA IVDEN

(The cipherer has broken this ciphertext up into blocks of five to help avoid errors.)

2. Route Cipher

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

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

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

EJXCTEDECDAEWRIORFEONALEVSE

Route ciphers have many more keys than a rail fence. In fact, for messages of reasonable length, the number of possible keys is potentially too great to be enumerated even by modern machinery. However, not all keys are equally good. Badly chosen routes will leave excessive chunks of plaintext, or text simply reversed, and this will give cryptanalysts a clue as to the routes.

An interesting variation of the route cipher was the Union Route Cipher, used by Union forces during the American Civil War. This worked much like an ordinary route cipher, but transposed whole words instead of individual letters. Because this would leave certain highly sensitive words exposed, such words would first be concealed by code. The cipher clerk may also add entire null words, which were often chosen to make the ciphertext humorous.

3. 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. Both the width of the rows and the permutation of the columns are usually defined by a keyword. For example, the word ZEBRAS is of length 6 (so the rows are of length 6), and the permutation is defined by the alphabetical order of the letters in the keyword. In this case, the order would be "6 3 2 4 1 5".

In a regular columnar transposition cipher, any spare spaces are filled with nulls; in an irregular columnar transposition cipher, the spaces are left blank. Finally, the message is read off in columns, in the order specified by the keyword. For example, suppose we use the keyword ZEBRAS and the message WE ARE DISCOVERED. FLEE AT ONCE. In a regular columnar transposition, we write this into the grid as:

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	K	J	E	U

Providing five nulls (QKJEU) at the end. The ciphertext is then read off as:

EVLNE ACDTK ESEAQ ROFOJ DEECU WIREE

In the irregular case, the columns are not completed by nulls:

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					

This results in the following ciphertext:

EVLNA CDTES EAROF ODEEC WIREE

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

Columnar transposition continued to be used for serious purposes as a component of more complex ciphers at least into the 1950's.

4. Double Transposition

A single columnar transposition could be attacked by guessing possible column lengths, writing the message out in its columns (but in the wrong order, as the key is not yet known), and then looking for possible anagrams. Thus to make it stronger, a double transposition was often used. This is simply a columnar transposition applied twice. The same key can be used for both transpositions, or two different keys can be used.

As an example, we can take the result of the irregular columnar transposition in the previous section, and perform a second encryption with a different keyword, STRIPE, which gives the permutation "564231":

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

As before, this is read off columnwise to give the ciphertext:

CAEEN SOIAE DRLEF WEDRE EVTOC

If multiple messages of exactly the same length are encrypted using the same keys, they can be anagrammed simultaneously. This can lead to both recovery of the messages, and to recovery of the keys (so that every other message sent with those keys can be read).

During World War I, the German military used a double columnar transposition cipher, changing the keys infrequently. The system was regularly solved by the French, naming it Übchi, who were typically able to quickly find the keys once they'd 4 intercepted a number of messages of the same length, which generally took only a few days. However, the French success became widely-known and, after a publication in Le Matin, the Germans changed to a new system on 18 November 1914. [1]

During World War II, the double transposition cipher was used by Dutch Resistance groups, the French Maquis and the British Special Operations Executive (SOE), which was in charge of managing underground activities in Europe.[2] It was also used by agents of the American Office of Strategic Services[3] and as an emergency cipher for the German Army and Navy.

Until the invention of the VIC cipher, double transposition was generally regarded as the most complicated cipher that an agent could operate reliably under difficult field conditions.

5. Myszkowski Transposition

A variant form of columnar transposition, proposed by Émile Victor Théodore Myszkowski in 1902, requires a keyword with recurrent letters. In usual practice, subsequent occurrences of a keyword letter are treated as if the next letter in alphabetical order, e.g., the keyword TOMATO yields a numeric keystring of "532164."

In Myszkowski transposition, recurrent keyword letters are numbered identically, TOMATO yielding a keystring of "432143."

4	3	2	1	4	3
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					

Plaintext columns with unique numbers are transcribed downward; those with recurring numbers are transcribed left to right:

ROFOA CDTED SEEEA CWEIV RLENE

6. Disrupted Transposition

In a disrupted transposition, certain positions in a grid are blanked out, and not used when filling in the plaintext. This breaks up regular patterns and makes the cryptanalyst's job more difficult.

7. Grilles

Another form of transposition cipher uses grilles, or physical masks with cut-outs, rather than a mathematical algorithm. This can produce a highly irregular transposition over the period specified by the size of the grille, but requires the correspondents to keep a physical key secret. Grilles were first proposed in 1550, and were still in military use for the first few months of World War One.

8. Detection and Cryptanalysis

Since transposition does not affect the frequency of individual symbols, simple transposition can be easily detected by the cryptanalyst by doing a frequency count. If the ciphertext exhibits a frequency distribution very similar to plaintext, it is most likely a transposition. This can then often be attacked by anagramming - sliding pieces of ciphertext around, then looking for sections that look like anagrams of English words, and solving the anagrams. Once such anagrams have been found, they reveal information about the transposition pattern, and can consequently be extended.

Simpler transpositions also often suffer from the property that keys very close to the correct key will reveal long sections of legible plaintext interspersed by gibberish. Consequently such

ciphers may be vulnerable to optimum seeking algorithms such as genetic algorithms.[citation needed]

A detailed description of the cryptanalysis of a German transposition cipher can be found in chapter 7 of Herbert Yardley's "The American Black Chamber."

9. Combinations

Transposition is often combined with other techniques. For example, a simple substitution cipher combined with a columnar transposition avoids the weakness of both. Replacing high frequency ciphertext symbols with high frequency plaintext letters does not reveal chunks of plaintext because of the transposition. Anagramming the transposition does not work because of the substitution. The technique is particularly powerful if combined with fractionation (see below). A disadvantage is that such ciphers are considerably more laborious and error prone than simpler ciphers.

10. Fractionation

Transposition is particularly effective when employed with fractionation - that is, a preliminary stage that divides each plaintext symbol into several ciphertext symbols. For example, the plaintext alphabet could be written out in a grid, then every letter in the message replaced by its co-ordinates (see Polybius square and Straddling checkerboard). Another method of fractionation is to simply convert the message to Morse code, with a symbol for spaces as well as dots and dashes.

When such a fractionated message is transposed, the components of individual letters become widely separated in the message, thus achieving Claude E. Shannon's diffusion. Examples of ciphers that combine fractionation and transposition include the bifid cipher, the trifid cipher, the ADFGVX cipher and the VIC cipher.

Another choice would be to replace each letter with its binary representation, transpose that, and then convert the new binary string into the corresponding ASCII characters. Looping the scrambling process on the binary string multiple times before changing it into ASCII characters would likely make it harder to break. Many modern block ciphers use more complex forms of transposition related to this simple idea.