UNIT-II – CLASSICAL ENCRYPTION TECHNIQUES

Basics of Information and Network Security

- In daily life we use information for various purposes and use network for communication and exchange information between different parties.
- In many cases these information are sensitive so we need to take care that only authorized party can get that information.
- For maintaining such privacy we require some mechanism or physical device which ensures that it is safe. Such mechanism or physical devices are known as **security system**.
- **Computer Security:** The protection afforded to an automated information system in order to attain the applicable objectives of preserving the **integrity**, **availability**, and **confidentiality** of information system resources.
- This definition of computer security introduces three key objectives that are at the heart of computer security:

1.Confidentiality: It covers two concepts

Data Confidentiality: Assures that private or confidential information is not made available or disclosed to unauthorized individuals.

Privacy: Assures that individuals control or influence what information related to them may be collected and stored and by whom and to whom that information may be disclosed.

2.Integrity: It covers two concepts

Data Integrity: Assures that information and programs are changed only in specified and authorize manner.

System Integrity: Assures that a system performs its intended function in an unimpaired manner, free from deliberate or inadvertent unauthorized manipulation of the system.

3.Availability: Assures that systems work promptly and service is not denied to authorize user.

•**Threat**: A potential for violation of security, which exists when there is a circumstance, capability,action, or event that could breach security and cause harm. That is, a threat is a possible danger that might exploit vulnerability.

Security services

- A security service is a processing or communicating service that can prevent or detect the various attacks. Various security services are:
  - **Authentication:** the recipient should be able to identify the sender, and verify that the sender, who claims to be the sender, actually did send the message.
  - **Data Confidentiality:** An attacker should not be able to read the transmitted data or extract data in case of encrypted data. In short, confidentiality is the protection of transmitted data from passive attacks.
  - **Data Integrity:** Make sure that the message received was exactly the message the sender sent.
  - **No repudiation**: The sender should not be able to deny sending the should not be able to deny receiving the message.message.The receiver should no be able to deny receiving the message.

Cryptography

- An original message is known as the **plaintext.**
- The Coded message is called the **ciphertext.**

UNIT-II – CLASSICAL ENCRYPTION TECHNIQUES

- The Process of converting from plaintext to ciphertext is known **as enciphering or encryption.**
- Restoring the plaintext from the ciphetext is **deciphering or decryption.**
- The many schemes used for encryption constitute the area of study known as **cryptography.**
- Techniques used for deciphering a message without any knowledge of the enciphering details is known as **cryptanalysis.** It also known as **"Breaking the Code".**
- The areas of cryptography and cryptanalysis together are called **cryptology.**
- A **cryptanalyst** develops mathematical methods and codes that protect data from computer hackers. This involves the decryption of a cipher text into plain text in order to transmit a message over insecure channels.

Symmetric cipher model

- Symmetric encryption is a form of cryptosystem in which encryption and decryption are performed using the same key.
- It is also known as conventional encryption.
- Symmetric encryption transforms plaintext into cipher text using a secret key and an encryption algorithm.
- Using the same key and a decryption algorithm, the plaintext is recovered from the cipher text.
- A symmetric encryption scheme has five ingredients
  - **Plaintext**: This is the original intelligible message or data that is fed into the algorithm as input.
  - **Encryption algorithm**: The encryption algorithm performs various substitutions and transformations on the plaintext.
  - **Secret key**: The secret key is also input to the encryption algorithm. The key is a value independent of the plaintext and of the algorithm.
  - **Ciphertext**: This is the scrambled message produced as output. It depends on the plaintext and the secret key. The ciphertext is an apparently random stream of data and, as it stands, is unintelligible.
  - **Decryption algorithm**: This is essentially the encryption algorithm run in reverse. It takes the ciphertext and the secret key and produces the original plaintext.

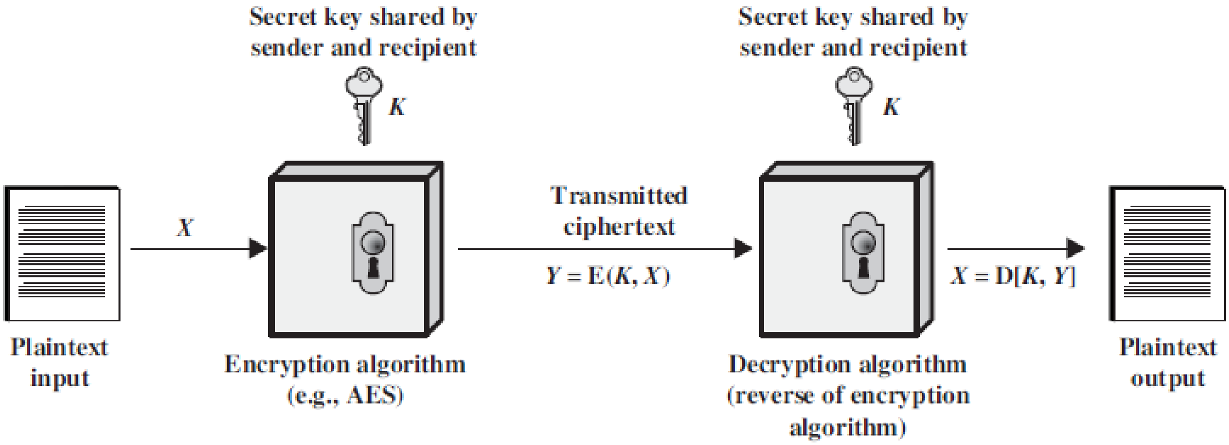
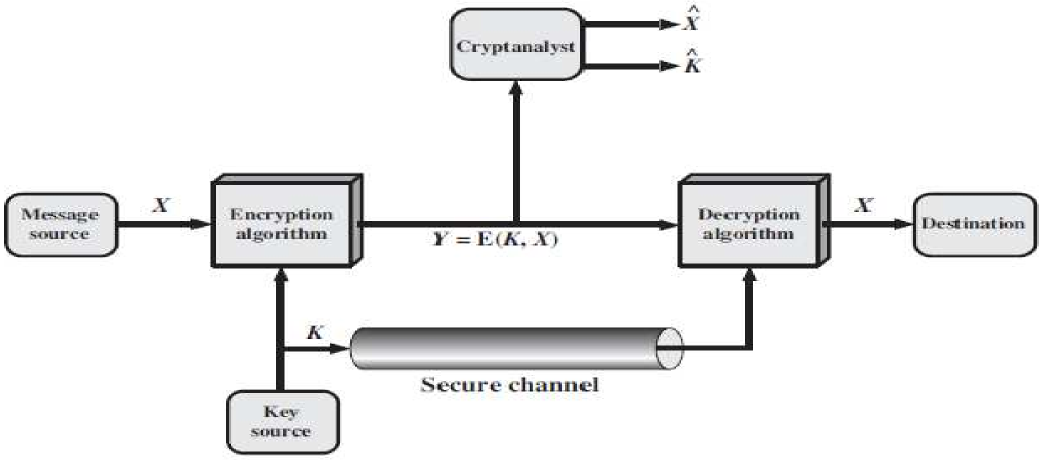

Secret key shared by sender and recipient
K
Secret key shared by sender and recipient
K
X
Transmitted ciphertext
Y = E(K, X)
X = D[K, Y]
Plaintext input
Encryption algorithm (e.g., AES)
Decryption algorithm (reverse of encryption algorithm)
Plaintext output

Fig: Simplified Model of Symmetric Encryption

UNIT-II – CLASSICAL ENCRYPTION TECHNIQUES

Symmetric Cipher Model



- A symmetric cipher model are broadly contains five parts.
- **Plaintext:** This is the original intelligible message.
- **Encryption algorithm:** The encryption algorithm performs various substitutions and transformations on the plaintext. It takes in plaintext and key and gives the cipher text.
- **Secret key:** The key is a value independent of the plaintext and of the algorithm. Different keys will yield different outputs.
- **Ciphertext:** This is the scrambled message produced as output. It depends on the plaintext and the secret key.
- **Decryption algorithm:** Runs on the cipher text and the key to produce the plaintext.This is essentially the encryption algorithm run in reverse.

- Two basic requirements of encryption are:
  1. Encryption algorithm should be strong. An attacker knowing the algorithm and having any number of cipher text should not be able to decrypt the cipher text or guess the key.
  2. The key shared by the sender and the receiver should be secret.
- Let the plaintext be $X = [X1, X2,\ldots, X_M]$, key be $K = [K1, K2,\ldots, K_J]$ and the cipher text produced be $Y = [Y1, Y2,\ldots, Y_N]$. Then, we can write

$$Y = E(K, \mathrm{X})$$

- Here E represents the encryption algorithm and is a function of plaintext X and key K.
- The receiver at the other ends decrypts the cipher text using the key.

$$\mathrm{X} = D(K, Y)$$

- Here D represents the decryption algorithm and it inverts the transformations of encryption algorithm.
- An opponent not having access to X or K may attempt to recover K or X or both.
- It is assumed that the opponent knows the encryption (E) and decryption (D) algorithms.
- If the opponent is interested in only this particular message, then the focus of the effort is to recover by generating a plaintext estimate $\hat{\mathrm{X}}$.
- If the opponent is interested in being able to read future messages as well then he will attempt to recover the key by making an estimate $\hat{K}$.

Encryption Animation https://youtu.be/1y1M2fZqIlQ

- Cryptographic systems are characterized along three independent dimensions.
  1. **The types of operations used for transforming plaintext to ciphertext.** All encryption algorithms are based on two general principles substitution, and transposition. Basic requirement is that no information be lost. Most systems referred to as product system, involves multiple stages of substitutions and transpositions.
  2. **The number of keys used.** If both sender and receiver use the same key, the system is referred to as **symmetric, single-key, secret-key, or conventional encryption**. If the sender and receiver use different keys the system is referred to as **asymmetric, two-key, or public-key encryption**.
  3. **The way in which the plaintext is processed.** A block cipher process a block at a time and produce an output block for each input block. A stream cipher process the input element continuously, producing output one element at a time, as it goes along.

Cryptanalysis and Brute-Force Attack

- **Cryptanalysis:** Cryptanalytic attacks rely on the nature of the algorithm plus perhaps some knowledge of the general characteristics of the plaintext or even some simple plaintext-ciphertext pairs. This type of attack finds characteristics of the algorithm to find a specific plaintext or to find key.
- **Brute-force attack:** The attacker tries every possible key on a piece of ciphertext until plaintext is obtained. On average, half of all possible keys must be tried to achieve success.
- Based on the amount of information known to the cryptanalyst cryptanalytic attacks can be categorized as:
  - **Cipher text Only Attack:** The attacker knows only cipher text only. It is easiest to defend.
  - **Known plaintext Attack:** In this type of attack, the opponent has some plaintext-cipher text pairs. Or the analyst may know that certain plaintext patterns will appear in a message. For example, there may be a standardized header or banner to an electronic funds transfer message and the attacker can use that for generating plaintext-cipher text pairs.
  - **Chosen plaintext:** If the analyst is able somehow to get the source system to insert into the system a message chosen by the analyst, then a *chosen-plaintext* attack is possible. In such a case, the analyst will pick patterns that can be expected to reveal the structure of the key.
  - **Chosen Cipher text:** In this attack, the analyst has cipher text and some plaintext-cipher text pairs where cipher text has been chosen by the analyst.
  - **Chosen Text:** Here, the attacker has got cipher text, chosen plaintext-cipher text pairs and chosen cipher text-plaintext pairs.
- Chosen cipher text and chosen text attacks are rarely used.
- It is assumed that the attacker knows the encryption and decryption algorithms.
- Generally, an encryption algorithm is designed to withstand a known-plaintext attack.

UNIT-II – CLASSICAL ENCRYPTION TECHNIQUES

Substitution Techniques

It is one in which the letters of plaintext are replaced by other letters or by numbers or symbols.

Caesar cipher

- The encryption rule is simple; replace each letter of the alphabet with the letter standing 3 places further down the alphabet.
- The alphabet is wrapped around so that Z follows A.
- Generally Plain text is in lower case and Cipher text is Upper Case.
- Example:
  Plaintext: meet me after the party
  Ciphertext: PHHW PH DIWHU WKH SDUWB

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12
n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

- Here, the key is 3. If different key is used, different substitution will be obtained.
- Mathematically, starting from a=0, b=1 and so on, Caesar cipher can be written as:

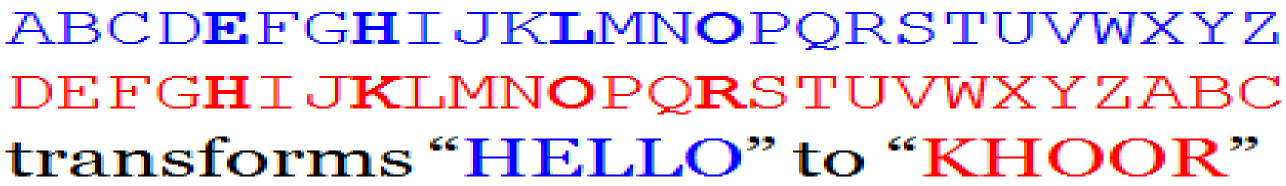
$$\mathrm{E}(p) = (p + k) \bmod (26)$$
$$\mathrm{D}(\mathrm{C}) = (\mathrm{C} - k) \bmod (26)$$

Encryption k=3 $\mathrm{E}(p) = (p + k) \bmod (26)$	**Result**	**Cipher Text** **D(C) = (C – k) mod (26)**	**Result**
M=E(M)=(12+3)mod26	15=P	D(P)=(15-3)mod26	12=m
E=E(E)=((4+3)mod26	7=H	D(H)=(7-3)mod26	4=e
E= E(E)=((4+3)mod26	7=H	D(H)=(7-3)mod26	4=e
T=E(T)=((19+3)mod26	21=V	D(V)=(21-3)mod26	19=t

- This cipher can be broken
  - If we know one plaintext-cipher text pair since the difference will be same.
  - By applying Brute Force attack as there are only 26 possible keys.

Animation Link :http://brianveitch.com/maze-runner/caesar/index.html

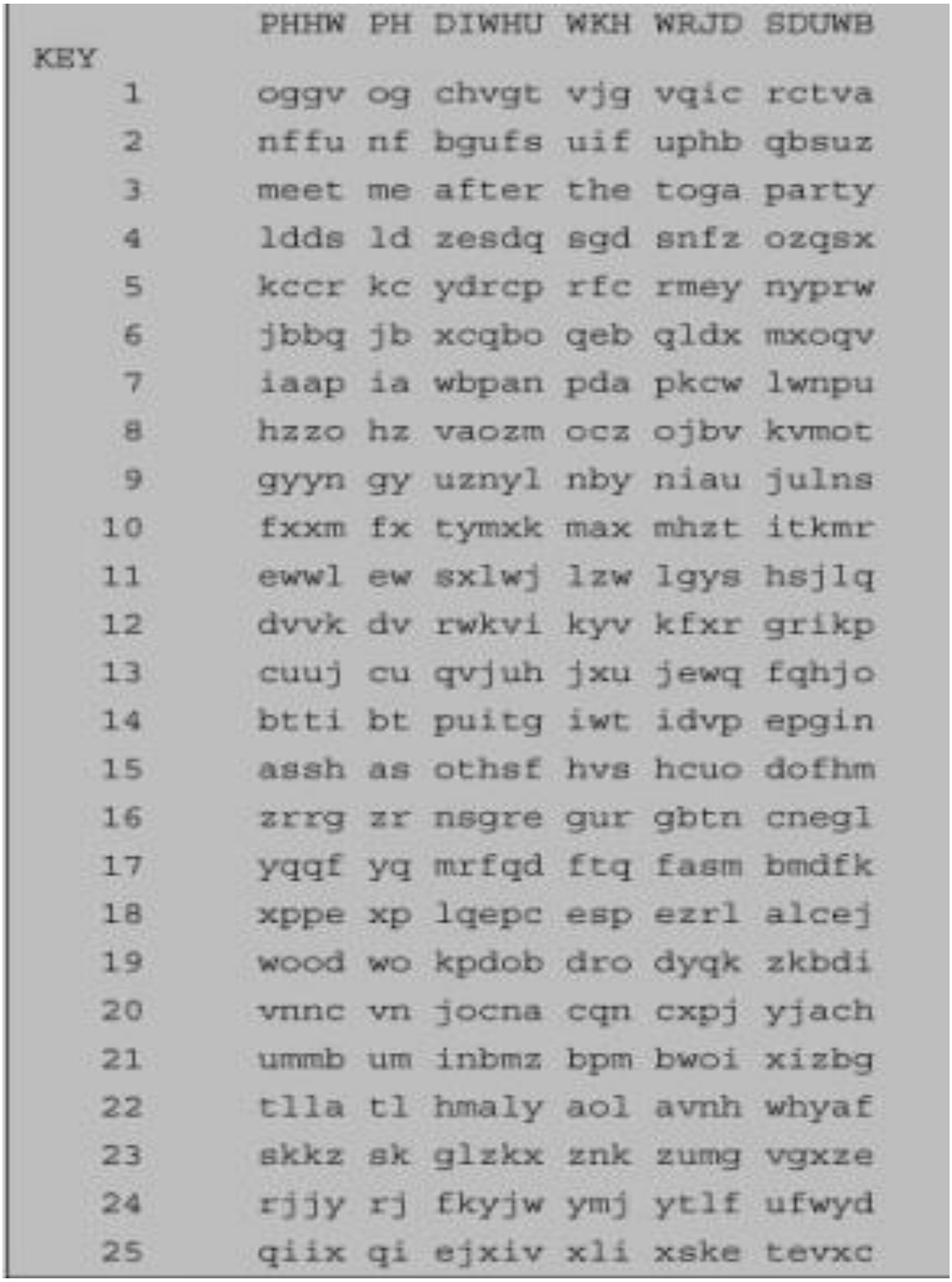
Example2:



ABCDEFGHIJKLMNOPQRSTUVWXYZ

DEFGHIJKLMNOPQRSTUVWXYZABC

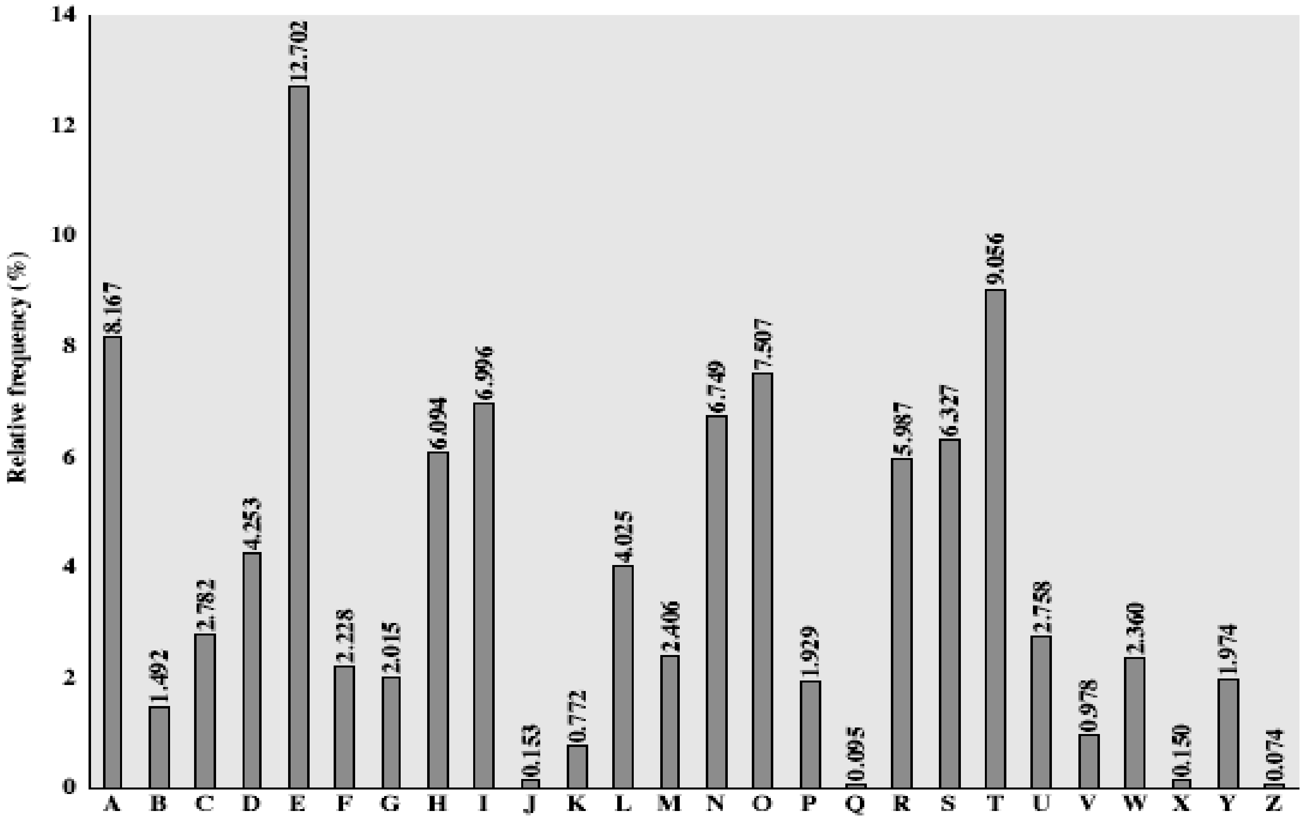
transforms “HELLO” to “KHOOR”



```
            PHHW PH DIWHU WKH WRJD SDUWB
KEY
   1        oggv og chvgt vjg vqic rctva
   2        nffu nf bgufs uif uphb qbsuz
   3        meet me after the toga party
   4        ldds ld zesdq sgd snfz ozqsx
   5        kccr kc ydrcp rfc rmey nyprw
   6        jbbq jb xcqbo qeb qldx mxoqv
   7        iaap ia wbpan pda pkcw lwnpu
   8        hzzo hz vaozm ocz ojbv kvmot
   9        gyyn gy uznyl nby niau julns
  10        fxxm fx tymxk max mhzt itkmr
  11        ewwl ew sxlwj lzw lgys hsjlq
  12        dvvk dv rwkvi kyv kfxr grikp
  13        cuuj cu qvjuh jxu jewq fqhjo
  14        btti bt puitg iwt idvp epgin
  15        assh as othsf hvs hcuo dofhm
  16        zrrg zr nsgre gur gbtn cnegl
  17        yqqf yq mrfqd ftq fasm bmdfk
  18        xppe xp lqepc esp ezrl alcej
  19        wood wo kpdob dro dyqk zkbdi
  20        vnnc vn jocna cqn cxpj yjach
  21        ummb um inbmz bpm bwoi xizbg
  22        tlla tl hmaly aol avnh whyaf
  23        skkz sk glzkx znk zumg vgxze
  24        rjjy rj fkyjw ymj ytlf ufwyd
  25        qiix qi ejxiv xli xske tevxc
```

Monoalphabetic Substitution Cipher

- Instead of shifting alphabets by fixed amount as in Caesar cipher, any random permutation is assigned to the alphabets. This type of encryption is called monoalphabetic substitution cipher.
- For example, A is replaced by Q, B by D, C by T etc. then it will be comparatively stronger than Caesar cipher.
- The number of alternative keys possible now becomes 26!.
- Thus, Brute Force attack is impractical in this case.
- However, another attack is possible. Human languages are redundant i.e. certain characters are used more frequently than others. This fact can be exploited.
- In English ‘e’ is the most common letter followed by ‘t’, ‘r’, ‘n’, ’o’, ‘a’ etc. Letters like ‘q’, ‘x’, ‘j’ are less frequently used.
- Moreover, digrams like ‘th’ and trigrams like ‘the’ are also more frequent.
- Tables of frequency of these letters exist. These can be used to guess the plaintext if the plaintext is in uncompressed English language.
- The most common two letter combinations are called as **digrams**. e.g. th, in, er, re and an.
- The most common three letter combinations are called as **trigrams.** e.g. the, ing, and, and ion



UNIT-II – CLASSICAL ENCRYPTION TECHNIQUES

Playfair Cipher

- In this technique multiple (2) letters are encrypted at a time.
- This technique uses a 5 X 5 matrix which is also called key matrix.

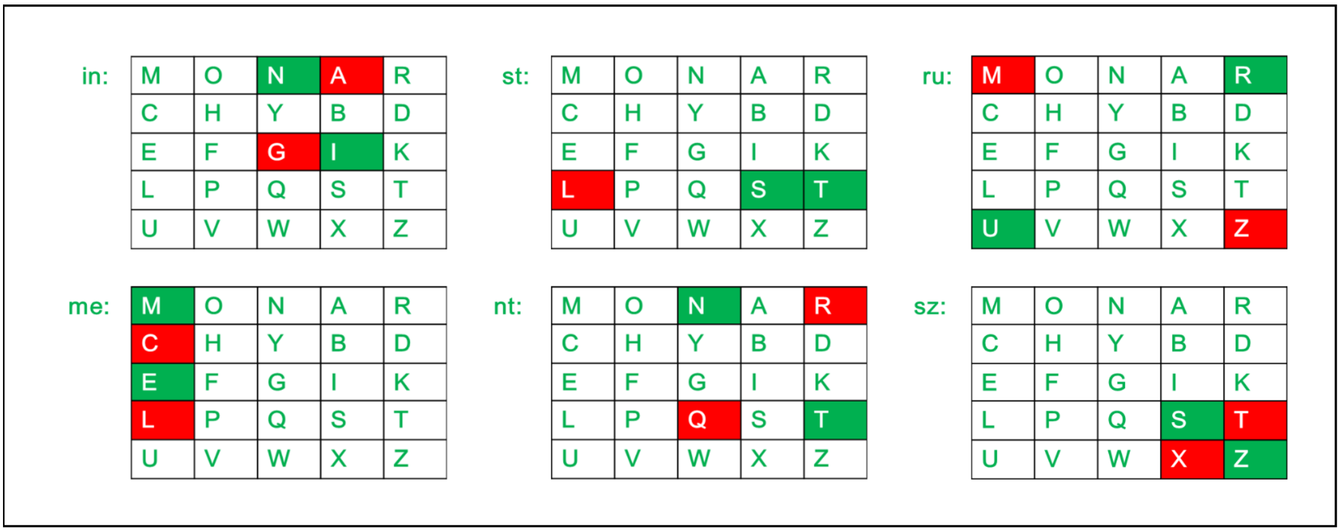
M	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

- The plaintext is encrypted **two letters at a time**:
  - Break the plaintext into pairs of two consecutive letters.
  - If a pair is a repeated letter, insert a filler like 'X'in the plaintext, eg. "Balloon" is treated as "ba lx lo on".
  - If both letters fall in the same row of the key matrix, replace each with the letter to its right (wrapping back to start from end), eg. “AR" encrypts as "RM".
  - If both letters fall in the same column, replace each with the letter below it (again wrapping to top from bottom), eg. “MU" encrypts to "CM".
  - Otherwise each letter is replaced by the one in its row in the column of the other letter of the pair, eg. “HS" encrypts to "BP", and “EA" to "IM" or "JM" (as desired)
- Security is much improved over monoalphabetic as here two letters are encrypted at a time and hence there are 26 X 26 =676 diagrams and hence it needs a 676 entry frequency table.

However, it can be broken even if a few hundred letters are known as much of plaintext structure isretained in cipher text.

```
Example 2:  PlainText: "instruments" keyword: monarchy
After Split: 'in' 'st' 'ru' 'me' 'nt' 'sz'
cipher text : ga tl mz cl rq tx
```

*For both **encryption** and **decryption**, the **same key** is to be used.*



in:

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

st:

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

ru:

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

me:

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

nt:

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

sz:

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

video link: https://www.youtube.com/watch?v=quKhvu2tPy8

Strength of playfair cipher Playfair cipher is a great advance over simple mono alphabetic ciphers. Since there are 26 letters, 26x26 = 676 diagrams are possible, so identification of individual diagram is more difficult.

Hill Cipher

source link

https://www.youtube.com/watch?v=sYdq6-kquKc

https://www.educative.io/edpresso/what-is-the-hill-cipher

- This cipher is based on linear algebra.
- Each letter is represented by numbers from 0 to 25 and calculations are done modulo 26.
- This encryption algorithm takes m successive plaintext letters and substitutes them with m cipher text letters.
- The substitution is determined by m linear equations. For *m* = 3, the system can be described as:

$$c_1 = (k_{11}p_1 + k_{12}p_2 + k_{13}p_3) \; mod \; 26$$
$$c_2 = (k_{21}p_1 + k_{22}p_2 + k_{23}p_3) \; mod \; 26$$
$$c_3 = (k_{31}p_1 + k_{32}p_2 + k_{33}p_3) \; mod \; 26$$

- This can also be expressed in terms of row vectors and matrices.

$$(c_1 \; c_2 \; c_3) = (p_1 \; p_2 \; p_3) \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} mod \; 26$$

Where **C** and **P** are row vectors of length 3 representing the plaintext and cipher text, and **K** is a 3 X 3 matrix representing the encryption key

- Key is an invertible matrix K modulo 26, of size m. For example:

$$K = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix} \qquad K^{-1} = \begin{pmatrix} 4 & 19 & 15 \\ 15 & 17 & 6 \\ 24 & 0 & 17 \end{pmatrix}$$

- Encryption and decryption can be given by the following formulae:

  Encryption: $C = PK \; mod \; 26$

  Decryption: $P = CK^{-1} \; mod \; 26$

- The strength of the Hill cipher is that it completely hides single-letter frequencies.
- Although the Hill cipher is strong against a cipher text-only attack, it is easily broken with a known plaintext attack.
  - Collect m pair of plaintext-cipher text, where m is the size of the key.
  - Write the m plaintexts as the rows of a square matrix P of size m.
  - Write the m cipher texts as the rows of a square matrix C of size m.
  - We have that C=PK mod 26.
  - If P is invertible, then K=P^{-1}C mod 26,
  - If P is not invertible, then collect more plaintext-cipher text pairs until an invertible P is obtained.

The Vigenère cipher

- This is a type of polyalphabetic substitution cipher (includes multiple substitutions depending on the key). In this type of cipher, the key determines which particular substitution is to be used.
- To encrypt a message, a key is needed that is as long as the message. Usually, the key is a repeating keyword.
- For example, if the keyword is *deceptive*, the message "we are discovered save yourself" is encrypted as follows:

  Key: *deceptivedecept*

  Plaintext: wearediscovered

  Ciphertext: ZICVTWQNGRZGVTW
- Encryption can be done by looking in the Vigenere Table where ciphertext is the letter key's row and plaintext's column or by the following formula:

$$C_{\mathrm{i}} = (P_{\mathrm{i}} + K_{\mathrm{i}\,mod\,m})\ mod\ 26$$

- Decryption is equally simple. The key letter again identifies the row. The position of the cipher text letter in that row determines the column, and the plaintext letter is at the top of that column.
- The strength of this cipher is that there are multiple ciphertext letters for each plaintext letter, one for each unique letter of the keyword.
- Thus, the letter frequency information is obscured however, not all knowledge of the plaintext structure is lost.

Vernam Cipher

- This system works on binary data (bits) rather than letters.
- The technique can be expressed as follows:

$$C_{\mathrm{i}} = P_{\mathrm{i}} \oplus K_{\mathrm{i}}$$

  Where

  P_{i} = ith binary digit of plaintext.

  K_{i} = ith binary digit of key.

  C_{i} = ith binary digit of ciphertext.

  $\oplus$ = exclusive-or (XOR) operation
- Thus, the ciphertext is generated by performing the bitwise XOR of the plaintext and the key.
- Decryption simply involves the same bitwise operation:

$$P_{\mathrm{i}} = C_{\mathrm{i}} \oplus K_{\mathrm{i}}$$

- The essence of this technique is the means of construction of the key.
- It was produced by the use of a running loop of tape that eventually repeated the key, so that in fact the system worked with a very long but repeating keyword.

- Although such a scheme has cryptanalytic difficulties, but it can be broken with a very long ciphertext or known plaintext as the key is repeated.

One-Time Pad

- In this scheme, a random key that is as long as the message is used.
- The key is used to encrypt and decrypt a single message, and then is discarded. Each new message requires a new key of the same length as the new message.
- This scheme is unbreakable.
- It produces random output that bears no statistical relationship to the plaintext.
- Because the ciphertext contains no information whatsoever about the plaintext, there is simply no way to break the code.
- For any plaintext of equal length to the ciphertext, there is a key that produces that plaintext.
- Therefore, if you did an exhaustive search of all possible keys, you would e d up with many legible plaintexts, with no way of knowing which the intended plaintext was.
- Therefore, the code is unbreakable.
- The security of the one-time pad is entirely due to the randomness of the key.
- The one-time pad offers complete security but, in practice, has two fundamental difficulties:
  - There is the practical problem of making large quantities of random keys. Any heavily used system might require millions of random characters on a regular basis. Supplying truly random characters in this volume is a significant task.
  - Another problem is that of key distribution and protection. For every message to be sent, a key of equal length is needed by both sender and receiver.
- Because of these difficulties, the one-time pad is used where very high security is required.
- The one-time pad is the only cryptosystem that exhibits **perfect secrecy**.

Transposition Techniques

- A very different kind of mapping is achieved by performing some sort of permutation on the plaintext letters. This technique is referred to as a transposition cipher.
- The simplest such cipher is the **rail fence** technique.

Rail Fence Technique

- Encryption involves writing plaintext letters diagonally over a number of rows, then read off cipher row by row.
- For example, the text “meet me after the party” can be written (in 2 rows) as:

```
m e m a t r h p r y
 e t e f e t e o a t
```

- Ciphertext is read from the above row-by-row:
  MEMATRHPRYETEFETEAT
- This scheme is very easy to cryptanalyze as no key is involved.
- Transposition cipher can be made significantly more secure by performing more than one stage of transposition. The result is a more complex permutation that is not easily reconstructed.

- d it much faster.