Module 1

Lecture Two: Elementary Cryptography

What is Cryptography?

In very simple terms, cryptography refers to hidden or coded writing. It is increasingly used to protect information to ensure confidentiality, Integrity, Authenticity and non-repudiation too. Therefore, it plays an important role in computer and information security.

Cryptography Terms (Speak Like a Crypto Geek)

Cryptology (to be very precise) Cryptography --- code designing

Cryptanalysis: the science of breaking cryptographic algorithms.

Cryptanalyst: a person who breaks cryptographic codes; also referred to as "the attacker". A cryptanalyst studies encryption and encrypted message and tries to find the hidden meanings (to break an encryption).

Cryptologist: Cryptographer & cryptanalyst

Plaintext –A message in its natural format readable by an attacker. It is the message or data before it gets encrypted.

Ciphertext – Message altered to be unreadable by anyone except the intended recipients. The encrypted (scrambled) version of the message.

Cipher: The algorithm that does the encryption.

Encryption/encipherment: scrambling a message or data using a specialized cryptographic algorithm.

Key – Sequence that controls the operation and behavior of the cryptographic algorithm

Keyspace – Total number of possible values of keys in a crypto algorithm Initialization Vector – Random values used with ciphers to ensure no patterns are created during encryption

Cryptosystem – The combination of algorithm, key, and key management functions used to perform cryptographic operations

Decryption/decipherment: the process of converting ciphertext back to the original plaintext.

History of Cryptography

We shall look at the history of Cryptography based on three eras as below:-

The Manual Era

Dates back to at least 2000 B.C. and was a Pen and Paper Cryptography. Examples of ciphers in this era includes Scytale Atbash Caesar Vigenère

The Mechanical Era

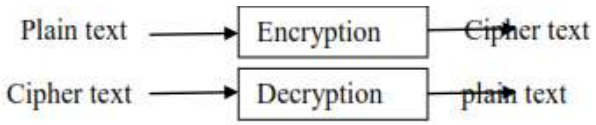
This era brought the Invention of cipher machines. Examples of ciphers in this era includes the Confederate Army's Cipher Disk, the Japanese Red and Purple Machines and the German Enigma

The Modern Era
This era is marked by the use of Computers and electronic devices. It is the modern era of cryptography. Examples of crypto algorithms in this era are Lucifer, Rijndael, RSA, ElGamal and many others.

Substitution Cipher

As seen above Encryption is the process of encoding a message so that its meaning is not obvious; decryption is the reverse process, transforming an encrypted message back into its normal, original form. Alternatively, the terms encode and decode or encipher and decipher are used instead of encrypt and decrypt. That is, we say that we encode, encrypt, or encipher the original message to hide its meaning. Then, we decode, decrypt, or decipher it to reveal the original message. A system for encryption and decryption is called a cryptosystem.

The original form of a message is known as plaintext, and the encrypted form is called cipher text. For convenience, we denote a plaintext message P as a sequence of individual characters $P = <p1, p2, \ldots, pn>$. Similarly, cipher text is written as $C = <c1, c2, \ldots, cm>$.



For instance, the plaintext message "I want cookies" can be denoted as the message string <*I, ,w,a,n,t,c,o,o,k,i,e,s*>. It can be transformed into cipher text<*c*1, *c*2, …,*c*14>, and the encryption algorithm tells us how the transformation is done.

We use this formal notation to describe the transformations between plaintext and cipher text. For example:

we write C = E (P) and P = D(C), where C represents the cipher text, E is the encryption rule, P is the plaintext, and D is the decryption rule.

P = D (E(P)).

In other words, we want to be able to convert the message to protect it from an intruder, but we also want to be able to get the original message back so that the receiver can read it properly.

The cryptosystem involves a set of rules for how to encrypt the plaintext and how to decrypt the cipher text. The encryption and decryption rules, called algorithms, often use a device called a key, denoted by *K*, so that the resulting cipher text depends on the original plaintext message, the algorithm, and the key value. We write this dependence as *C =E* (*K, P*). Essentially, *E* is a *set* of encryption algorithms, and the key *K* selects one specific algorithm from the set.

There are many types of encryption. In the next sections we look at two simple forms of encryption: substitutions in which one letter is exchanged for another and transpositions, in which the order of the letters is rearranged.

Cryptanalyst: cryptanalyst is a person who studies encryption and encrypted message and tries to find the hidden meanings (to break an encryption).

Confusion: it is a technique for ensuring that ciphertext has no clue about the original message.

Diffusion: it increases the redundancy of the plaintext by spreading it across rows and columns.

Substitutions Cipher: It basically consists of substituting every plaintext character for a different cipher text character.

It is of two types-

I. Mono alphabetic substitution cipher
II. Poly alphabetic substitution cipher

Mono alphabetic substitution cipher:

Relationship between cipher text symbol and plain text symbol is 1:1.

- Additive cipher:
  Key value is added to plain text and numeric value of key ranges from 0 – 25.

  Example:

  Plain text(P)- H E L LO (H=7,E=4,L=11,L=11,O=14)
  Key (K)=15
  Cipher text (C)= 7+15,4+15,11+15,11+15,14+15
  = 22,19, 26,26,(29%26)=3
  = W T A AD

Affine cipher:

It is the combination of additive and multiplicative cipher

C= (P+K) mod 26 P= (C-K) mod 26

Let K1 and K2 are two keys

C= [(P x K1) + K2] mod 26 P= [(C-K2) x $K1^{-1}$] mod 26

Polyalphabetic substitution cipher

In polyalphabetic cipher each occurrence of a character may have different substitution. The relationship between characters in plain text and cipher text is 1 to many.

- Auto key cipher
- Playfair cipher
- Vigegeire cipher
- Hill cipher

Auto key cipher:

- In this cipher, key is stream of subkeys in which subkey is used to encrypt the corresponding character in the plain text.
- Here 1st subkey is predefined and 2nd subkey is the value of the 1st character of the plain text 3rd subkey is the value of the 2nd plain text and so on.

Example: A T T A C K

0 19 19 0 2 10

Key=12 ↘ ↘ ↘ ↘ ↘

12 0 19 19 0 2

Cipher text(C)= (12,19,38 19,2 12)%26 ⟶ M T M T C M

Playfair cipher

In playfair cipher the secret key is made of 25 characters arranged in 5x5 matrix

Rules:-

- If 2 letters in a plaintext are located in the same row of the secret key then the corresponding encrypted character for each letter is next letter to the right.
- If 2 letters in a pair are in same column then the corresponding encrypted character is next below in the same column.
- If 2 letters are neither in same row or in same column then encrypted character is in its own row but in the same column as the other character.

Example:

$$K = \begin{vmatrix} L & G & D & B & A \\ Q & M & H & E & C \\ U & R & N & I/J & F \\ X & V & S & O & K \\ Z & Y & W & T & P \end{vmatrix}$$

Plain text= HELLO

It is then made as pair.

H E	L X	L O
H → E	L → Q	L → B
E → C	X → Z	O → X

Vigener cipher:

The key stream is the repetition of the initial secret key stream of length m. (1<=m<=26)

Example:

Plaintext- A B C D E F G H

Ks= 0, 5, 8

A	B	C	D	E	F	G	H	(B=1 =>1+5=6=>G)
0	5	8	0	5	8	0	5	
0	**6**	**10**	**3**	**9**	**13**	**6**	**12**	
A	**G**	**K**	**D**	**J**	**N**	**G**	**M**	<= ciphertext

Transposition cipher:

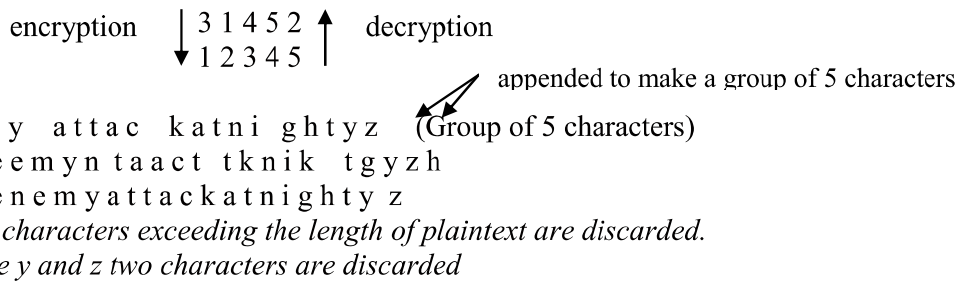
A transposition cipher is a method of encryption by which the positions held by units of plaintext (which are commonly characters or groups of characters) are shifted according to a regular system, so that the ciphertext constitutes a permutation of the plaintext. That is, the order of the units is changed.

The goal of substitution is confusion; *the transposition method is an attempt to make it difficult i.e diffusion.*

1. Keyless transposition cipher
   There are two methods for permutation of characters
   - Text is written into a table column by column and transmitted row by row
     Example: plaintext- meet me at the park
     m e m a t e a k
     e t e t h p r
     ciphertext- memateaketethpr
   - Text is written into the table row by row and then transmitted column by column.
     Example: **m e e t**
     m e a t
     t h e p
     a r k
     ciphertext- mmtaeehreaekttp

2. Keyed transposition cipher
   Plaintext is divided into groups and permutes the character in each group.
   Example: plaintext- “enemy attack at night”

   keys:



encryption: e e m y n t a a c t t k n i k t g y z h
decryption: e n e m y a t t a c k a t n i g h t y z
the characters exceeding the length of plaintext are discarded.
Like y and z two characters are discarded

3. Combining the two approaches:
   Encryption and decryption is done in three steps.
   - Text is written into a table row by row.
   - Permutation is done by reordering the column.
   - New table is read column by column

1.5 MAKING GOOD ENCRYPTION ALGORITHM

So far, the encryption algorithms we have seen are trivial, intended primarily to demonstrate the concepts of substitution and permutation. At the same time, we have examined several approaches cryptanalysts use to attack encryption algorithms. Now we examine algorithms that are widely used in the commercial world.
For each type of encryption we considered, has the advantages and disadvantages. But there is a broader question: What does it mean for a cipher to be "good"? The meaning of good depends on the intended use of the cipher. A cipher to be used by military personnel in the field has different requirements from one to be used in a secure installation with substantial computer support. In this section, we look more closely at the different characteristics of ciphers.

Shannon's Characteristics of "Good" Ciphers

In 1949, Claude Shannon [SHA49] proposed several characteristics that identify a good cipher.

1. The amount of secrecy needed should determine the amount of labor appropriate for the encryption and decryption.

2. The set of keys and the enciphering algorithm should be free from complexity.

   This principle implies that we should restrict neither the choice of keys nor the types of plaintext on which the algorithm can work. For instance, an algorithm that works only on plaintext having an equal number of A's and E's is useless. Similarly, it would be difficult to select keys such that the sum of the values of the letters of the key is a prime number. Restrictions such as these make the use of the encipherment prohibitively complex. If the process is too complex, it will not be used. Furthermore, the key must be transmitted, stored, and remembered, so it must be short.

3. The implementation of the process should be as simple as possible.
Principle 3 was formulated with hand implementation in mind: A complicated algorithm is prone to error or likely to be forgotten. With the development and popularity of digital computers, algorithms far too complex for hand implementation became feasible. Still, the issue of complexity is important. People will avoid an encryption algorithm whose implementation process severely hinders message transmission, thereby undermining security. And a complex algorithm is more likely to be programmed incorrectly.

4. Errors in ciphering should not propagate and cause corruption of further information in the message.
Principle 4 acknowledges that humans make errors in their use of enciphering algorithms. One error early in the process should not throw off the entire remaining ciphertext. For example, dropping one letter in a columnar transposition throws off the entire remaining encipherment. Unless the receiver can guess where the letter was dropped, the remainder of the message will be unintelligible. By contrast, reading the wrong row or column for a polyalphabetic substitution affects only one character and remaining characters are unaffected.
5. The size of the enciphered text should be no larger than the text of the original message.

The idea behind principle 5 is that a ciphertext that expands dramatically in size cannotpossibly carry more information than the plaintext, yet it gives the cryptanalyst more datafrom which to infer a pattern. Furthermore, a longer ciphertext implies more space for storage and more time to communicate.

Properties of "Trustworthy" Encryption Systems
Commercial users have several requirements that must be satisfied when they select an encryption algorithm. Thus, when we say that encryption is "commercial grade," or "trustworthy," we mean that it meets these constraints:

- It is based on sound mathematics. Good cryptographic algorithms are not just invented; they are derived from solid principles.
- It has been analyzed by competent experts and found to be sound. Even the best cryptographic experts can think of only so many possible attacks, and the developers may become too convinced of the strength of their own algorithm. Thus, a review by critical outside experts is essential.
- It has stood the atest of time.a As a new algorithm gains popularity, people continue to review both its mathematical foundations and the way it builds on those foundations. Although a long period of successful use and analysis is not a guarantee of a good algorithm, the flaws in many algorithms are discovered relatively soon after their release.

We can divide all the cryptography algorithms (ciphers) into two groups: symmetric key cryptography algorithms and asymmetric cryptography algorithms. Figure shows the taxonomy.

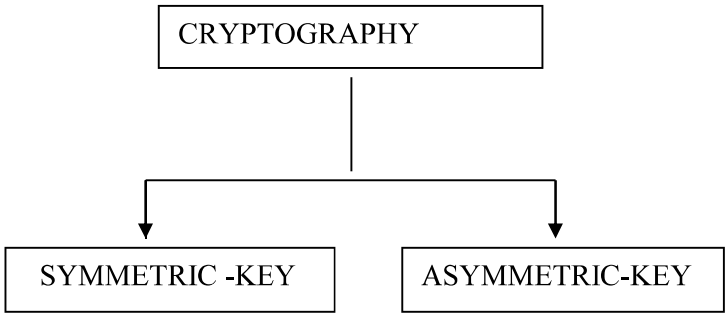


Fig :Categories of Cryptography

1. *Symmetric·Key Cryptography*

In symmetric-key cryptography, the same key is used by both parties. The sender uses this key and an encryption algorithm to encrypt data; the receiver uses the same key and the corresponding decryption algorithm to decrypt the data.

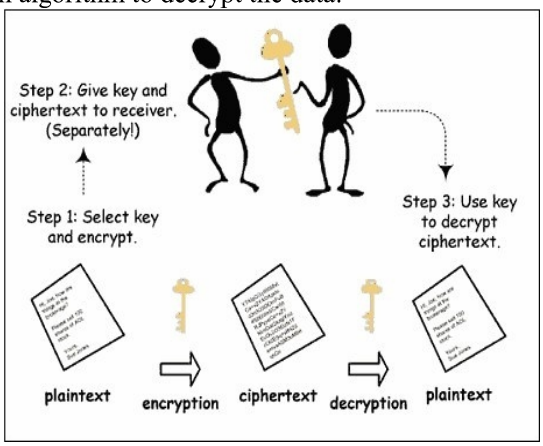


Fig :Symmetric-key Cryptography

2. *Asymmetric-Key Cryptography:*

In asymmetric or public-key cryptography, there are two keys: a private key and a public key. The private key is kept by the receiver. The public key is announced to the public.