COMPUTER SECURITY

3rd Year Computer Science

Chapiter 2 :

Introduction to Cryptography

Introduction

Cryptology is a very ancient science, humans have always needed to conceal information and transmit messages in complete confidentiality. The term **"cryptology"** comes from the Greek **"kruptos"** meaning secret or hidden, and **"logos"** meaning discourse.

Cryptology is therefore the science of secrecy. It encompasses two branches:

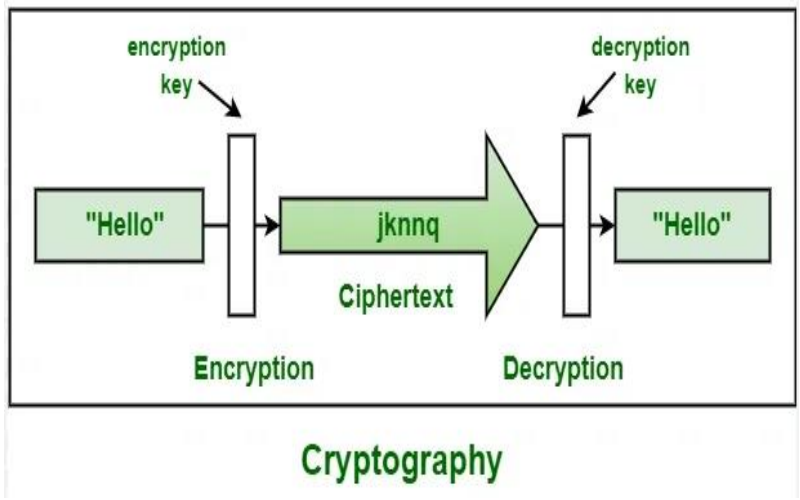
- **Cryptography.**
- **Cryptanalysis.**

1. Definitions and terminology

1. What is Cryptography?

Cryptography is the science, as well as art which focuses on encryption of communication and information by using cipher codes. Its essential function is to provide data protection through encryption, which involves converting text into unreadable cipher text and so access by the wrong people can be prevented. Cryptographic techniques, by way of algorithms and keys, are used to encode and decode messages, being capable of discerning the original content only by authorized parties.

1. What is Cryptography?



1. Definitions and terminology

2. What is Cryptanalysis?

Cryptanalysis is the science of breaking encrypted systems by reversing the encryption process to recover original keys and plaintext messages. Cryptanalysts use various methods including mathematical algorithms, statistical analysis, brute-force attacks, and frequency analysis. Its ultimate goal is to continuously improve cryptographic algorithms in response to new attack techniques.

3. Cryptography Terminologies

- **Plaintext:** The medium, the non-encrypted source data, and the clear message.
- **Ciphertext:** The ciphertext is the form of incomprehensible plaintext. Only the encryption key can open the cipher and let the message be read.
- **Encryption:** The process of applying cryptographic algorithms and keys to convert the plaintext into ciphertext.
- **Decryption:** The process that is equivalent to encryption does decryption by way of converting ciphertext to plaintext with the aid of decryption keys.
- **Key:** A cryptographic key is a piece of information used to encrypt and decrypt data.

4. Cryptanalysis Terminologies

- **Brute Force Attack:** A cryptanalysis technique whereby you keep on trying all the keys until you find the correct one.
- **Frequency Analysis:** A method involves untangling encrypted messages by studying the frequency of the letters or symbols in the ciphertext.

2. History of Cryptography

- The first encryption methods date back to Antiquity. As early as the 5th century BC, the Greeks concealed the contents of their communications by writing messages on scytales (wooden sticks) wrapped with leather strips, wax-covered wooden tablets, or directly on the shaved heads of their messengers. In the 1st century BC, Caesar's cipher appeared. In the 16th century, the Vigenère cipher was invented.
- From the 20th century onward, cryptography experienced a new surge, playing a key role in various wars. The Enigma machine was created and used by the Germans in World War II.

2. History of Cryptography

- With the advent of computing, its use became widespread. In 1977, the symmetric encryption standard DES was proposed by NIST. In 1976, asymmetric encryption was born with the Diffie-Hellman cipher, and in 1977, RSA was born and became globally used. The AES cipher is the current standard in symmetric cryptography, proposed in 2000.
- Finally, post-quantum cryptography allows us to surpass the limits of mathematical cryptography.

3. Classical Cryptography

Classical cryptography has been used since antiquity. It is based on the use of language letters for text encryption. The same key is used for both encryption and decryption. This category continued until the end of World War II, and these cryptosystems were applied to protect physical documents in military and diplomatic domains.

3. Classical Cryptography

3.1. Code Books Dictionaries:

used to replace certain words with different words. They were widely used until the early 20th century and were strongly criticized by Kerckhoffs.

3.2. Transposition Ciphers

Transposition methods consist of rearranging the data to be encrypted in order to make it incomprehensible for example, geometrically reordering the data to render it visually unusable.

3.2.1. Scytale cipher

This technique is likely the first evidence of encryption used in Greece as early as 600 BC to conceal messages written on strips of papyrus. It consisted of:

- Wrapping a strip of papyrus around a cylinder called a scytale.
- Writing the text longitudinally on the wrapped strip.

Once unrolled, the message is no longer comprehensible. The recipient simply needs a cylinder of the same radius to decrypt the message.

3.2.1. Scytale cipher

3.2.1. Scytale cipher

Example:

- To send a secret message between two parties.
- Both the sender and the receiver must first agree on a grid of a predetermined fixed width.
- The message is then written into this grid, with spaces between words replaced by the symbol □.
- Once the grid is filled, the ciphertext is obtained by reading the content column by column.

3.2.1. Scytale cipher

Improvement of the Scytale cipher:

- To allow the code to be changed quickly without altering its principle, and thereby increase security, both parties can decide to add a key.
- The goal is to easily change the encryption of a message while keeping the same encoding algorithm.
- To achieve this, a secret key is added specifying the column reading order.

3.3. Substitution Ciphers

Substitution ciphers consist of replacing one or more entities (generally letters) in a message with one or more other entities. There are four types:

- **Monoalphabetic substitution:** Each letter of the message is replaced by another letter of the alphabet.
- **Polyalphabetic substitution:** Uses a series of monoalphabetic ciphers reused periodically.
- **Homophonic substitution:** Maps each plaintext letter to a set of possible other characters.
- **Polygram substitution:** Substitutes a group of characters (polygram) with another group.

3.3.1. Caesar Cipher (50 BC)

One of the simplest and most well-known classical ciphers, based on shifting the letters of the alphabet. Julius Caesar used the following substitution during the Gallic Wars:

encoded letter = plain letter + 3 mod 26

- More generally, the family of codes is:

encoded letter = plain letter + n mod 26

where n (0–25) is the key.

- Decoding uses: plain letter = encoded letter − n mod 26.
- This is a symmetric stream cipher (secret key cipher).

3.3.2. Vigenère Cipher

A decisive improvement over Caesar's cipher. Its strength lies in using not one, but 26 shifted alphabets to encrypt a message

The Vigenère square works by using a key to determine how much each letter in the message gets shifted:

- The key letter is in the leftmost column.
- The plaintext letter is in the top row.
- The encrypted letter is at the intersection of the two.

The great strength of Vigenère is that the same letter will be encrypted in different ways, which makes classical frequency analysis ineffective.

19

Vigenère square

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

20

3.3.3. Vernam Cipher (One Time Pad - 1917)

The one-time pad, also known as the Vernam Cipher, is defined as a Vigenère cipher with the characteristic that the encryption key has the same length as the plaintext message.

To use this encryption, several properties must be respected:

- Choose a key as long as the text to be encrypted,
- Use a key made up of a sequence of random characters,
- Protect the key,
- Never reuse a key.

3.3.4. Affine Cipher

The idea is to use an affine function as the encryption function:

$$\mathbf{y = (ax + b) \bmod 26}$$

Where a and b are constants, and x and y are numbers corresponding to the letters of the alphabet (A = 0, B = 1, ...).

If a = 1, we recover the Caesar cipher where b is the shift.

Neutrality property: if b = 0, then "a" is always encrypted as "A" since it undergoes no shift. The original alphabet maps to itself.

For the affine cipher, the key consists of $\mathbf{(k_1, k_2)}$ where $k_1, k_2 \in [0, 25]$ such that: $\gcd(k_1, 26) = 1$.

3.3.4. Affine Cipher

Encryption and decryption formulas:

$ci = f(mi) = k_1 \times mi + k_2 \bmod 26$

$mi = f{-}1(ci) = k_1$

$-1 \times (ci - k_2) \bmod 26$

With the affine cipher there are 312 possible keys (12 choices for k_1 × 26 choices for k_2)

Example: Key = (k_1, k_2) = (3, 11)

- Encryption: $c_i = 3 \times m_i + 11 \bmod 26$
- Decryption: $k_1^{-1} = 3^{-1} \bmod 26 = 9$ [since $3 \times 9 \bmod 26 = 1$]

  $m_i = 9 \times (c_i - 11) \bmod 26$
- Thus: 'NSA' → 13 18 0 → 24 13 11 → 'YNL'

3.3.5. Polygraphic Encryption(Polygram substitution)

This involves encrypting a group of n letters with another group of n symbols. Notable examples include the Playfair cipher and the Hill cipher.

1. Playfair Cipher (1854)

- Two letters are encrypted as two others. The 25 letters of the alphabet (W excluded; V used instead) are arranged in a 5×5 grid built around a key.
- The English variant keeps W and merges I and J.
- The grid is filled with the letters of the keyword (ignoring duplicates), row by row, then the remaining alphabet letters fill the rest.

1. Playfair Cipher (1854)

Example: Key = example playfair

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

There are 4 rules to apply depending on the two letters being encrypted:

1. If the letters are at corners (different rows and columns), the encrypted letters are the other two corners of the rectangle. Ex: LA → ER, CA → BY

2. If the letters are on the same row, take the two letters immediately to their right. Ex: EM → XP, CH → DB

1. Playfair Cipher (1854)

3. If the letters are in the same column, take the two letters immediately below them. Ex: YK → CT, MU → FM

4. If they are identical (or only one remains), insert a null character (usually X) between them.

Example:

2. Hill Cipher (1929)

Simple substitution using polygrams. The algorithm replaces m successive plaintext letters with m ciphertext letters, using m linear equations.

Encryption and decryption in the Hill system:

$C = E_k(P) = K \cdot P \bmod 26$

$P = D_k(C) = K^{-1} \cdot C \bmod 26$

Where P and C are column vectors of size m (plaintext and ciphertext characters), and K is an invertible m×m matrix (the key), defined over the alphabet set.

Example: Encrypt ' TEXTEACRYPTER ' with m = 2 (blocks of 2 letters), using key matrix K = [[3,5],[6,17]] (det = 21, inverse mod 26 = 5).

27

2. Hill Cipher (1929)

• If we set $\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} \bmod 26$

• The ciphertext letters are computed using these linear equations:

$c_1 = (k_{11}P_1 + k_{12}P_2 + k_{13}P_3) \bmod 26$

$c_2 = (k_{21}P_1 + k_{22}P_2 + k_{23}P_3) \bmod 26$

$c_3 = (k_{31}P_1 + k_{32}P_2 + k_{33}P_3) \bmod 26$

28

2. Hill Cipher (1929)

Example:

- We perform encryption by blocks of 2 letters (m = 2).
- We want to encrypt the message 'TEXTEACRYPTER' using, as a key, a matrix K whose determinant is coprime with 26.
- For example, we use the matrix: $K = \begin{pmatrix} 3 & 5 \\ 6 & 17 \end{pmatrix}$
- whose determinant is 21. Since $5 \times 21 = 105 \equiv 1 \pmod{26}$, 5 is an inverse of det(K) modulo 26.

TEXTEACRYPTER → 19 ; 4 ; 23 ; 19 ; 4 ; 0 ; 2 ; 17 ; 24 ; 15 ; 19 ; 4 ; 17

- The letters are grouped into pairs, creating 7 two-dimensional vectors, the last pair being completed arbitrarily:

$X_1 = (19;4); X_2 = (23;19); X_3 = (4;0); X_4 = (2;17); X_5 = (24;15); X_6 = (19;4); X_7 = (17;6).$

2. Hill Cipher (1929)

- We then multiply (modulo 26) these vectors by the matrix K, for example for the first vector:

$$Y_1 = \begin{pmatrix} 3 & 5 \\ 6 & 17 \end{pmatrix} \begin{pmatrix} 19 \\ 4 \end{pmatrix} \bmod 26 = \begin{pmatrix} 25 \\ 0 \end{pmatrix}$$

- We then obtain 7 vectors, i.e., 14 letters:

(25 ; 0) ; (8 ; 19) ; (12 ; 24) ; (13 ; 15) ; (17 ; 9) ; (25 ; 0) ; (3 ; 22)

→ ZAITMYNPRJZADW

- To decrypt the cryptogram, the matrix K must be inverted:

$$K^{-1} = \begin{pmatrix} 17 & -5 \\ -6 & 3 \end{pmatrix}$$

- and multiplied (modulo 26) by the inverse of the determinant of K, that is, by 5:

$$B = \begin{pmatrix} 7 & 1 \\ 22 & 15 \end{pmatrix}$$

2. Hill Cipher (1929)

- Knowing the Y pairs, simply multiply them (modulo 26) by matrix B to recover the X pairs and successfully decrypt the message.
- For example, for the first vector:

$$X_1 = \begin{pmatrix} 7 & 1 \\ 22 & 15 \end{pmatrix} \begin{pmatrix} 25 \\ 0 \end{pmatrix} \bmod 26 = \begin{pmatrix} 19 \\ 4 \end{pmatrix}$$

3.4. Rotor Machines

quickly after the First World War, it became clear that if one wished to transmit large numbers of encrypted documents rapidly, and to be able to change encryption keys easily, it would be necessary to build machines capable of both encrypting and decrypting. The machines used for this purpose are rotor machines (the most famous of which was the ENIGMA machine with 3 rotors, invented in the 1930).This electric machine consists of an alphabetic keyboard, a light-up display screen, and three rotors. The system is straightforward: the user types a letter on the keyboard and the encrypted text then appears on the screen. With each keystroke, the first rotor would advance by one position, and once it had completed a full rotation, it would shift the second rotor by one position, and so on.

4. Rotor Machines

The rotors were initially set to any desired position, which defined the encryption key.To encrypt a message, once the key was set, all one had to do was type it on the machine; and to decrypt it, one simply had to place the rotors back in the same initial position and type in the encrypted message.