CRYPTOGRAPHY

DURATION : 1’30

Version : 0.0.3 Update : 02/02/2022

Summary

1. History of cryptography
2. Introduction to cryptograpgy
    1. Stream cipher
    2. Block cipher
    3. Symmetric encryption
    4. Asymmetric encryption
3. Hash functions
4. Steganography

HISTORY OF CRYPTOGRAPHY

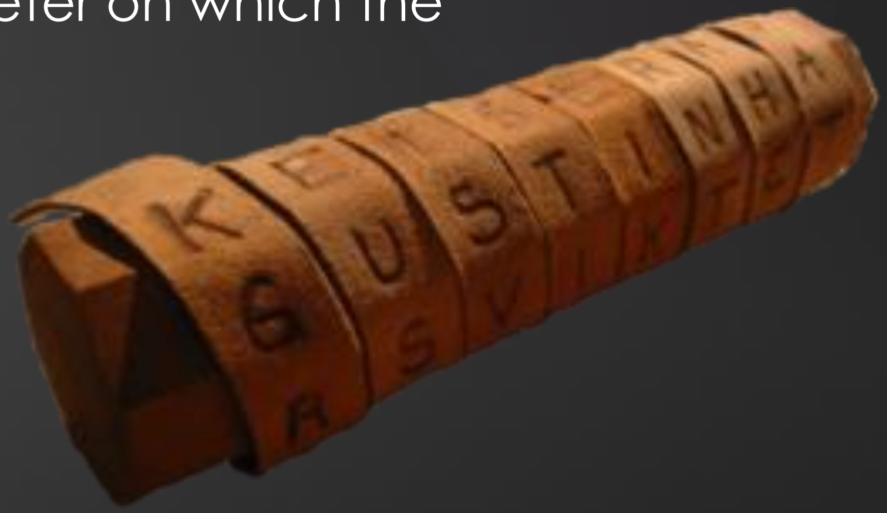
History

- Cryptography appeared in antiquity with **Scytale** (transposition).
- **Caesar cipher** with letters shift (monoalphabetic substitution).
- **Vigenère cipher** with letters shift (polyalphabetic substitution).

- **Enigma machine** created by the Germans (substitution and transposition).

Scytale

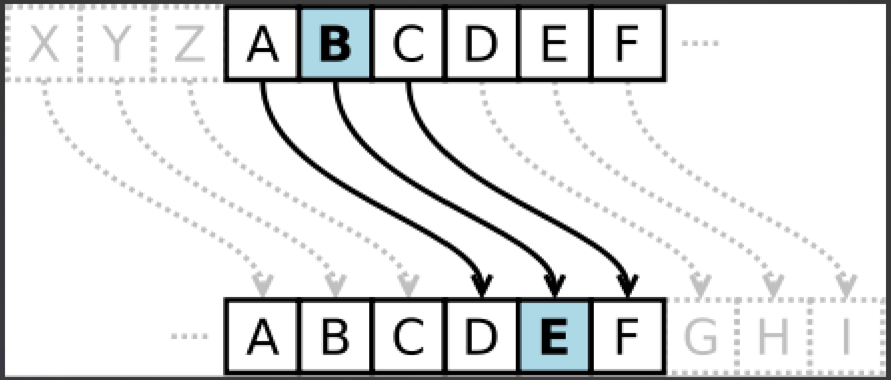
- It’s a baton used to perform a transposition cipher.
- It consist of a cylinder with a strip of parchment wound around it on which is written a message.
- The recipient uses a baton of the same diameter on which the parchment is wrapped to read the message.



- Exercise:
  - Decrypt “AQPETUASTERUAZLD”.

Caesar cipher

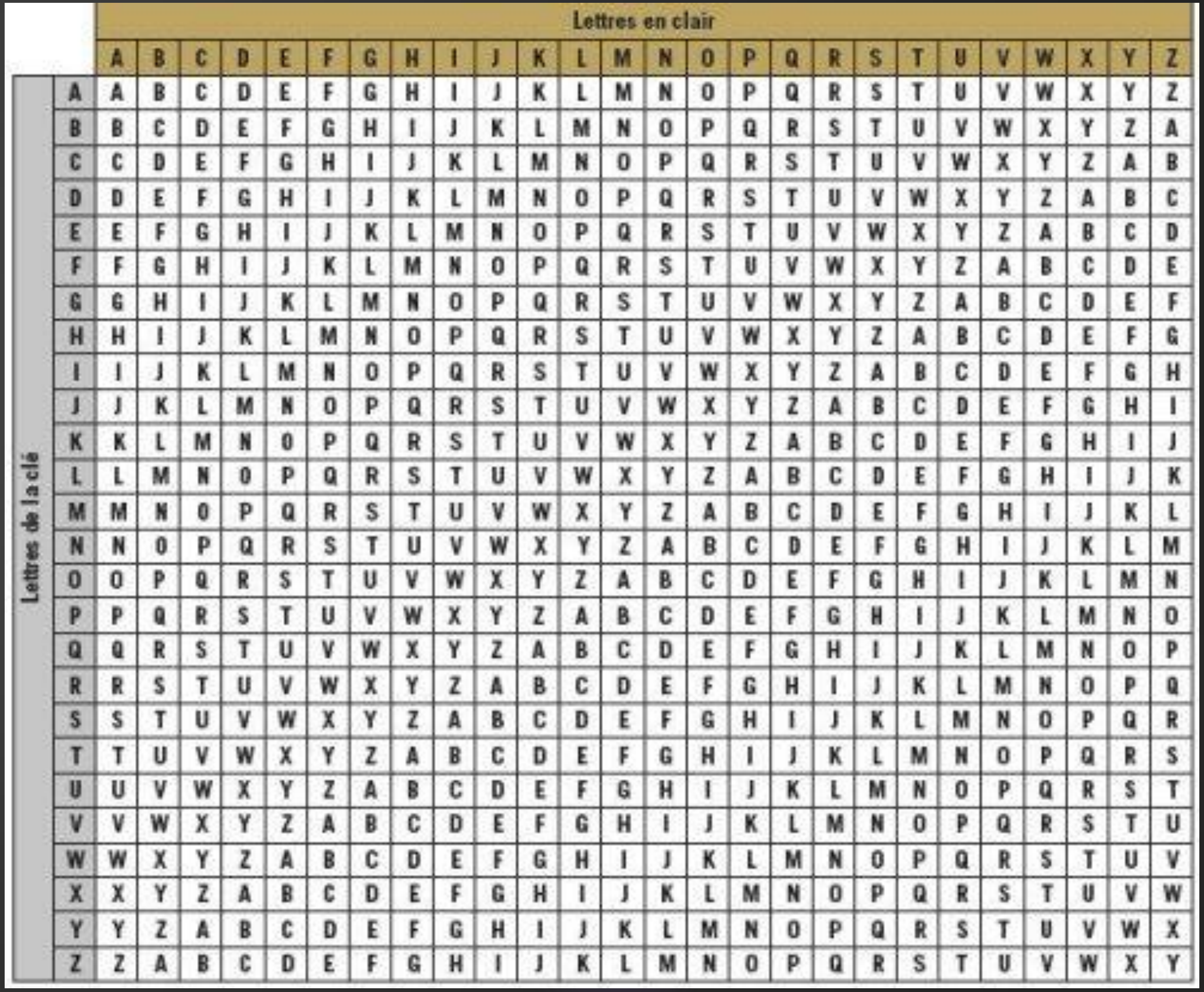
- It’s a type of substitution cipher in which each letter in the plaintext is replaced by a letter some fixed number of positions down the alphabet.



- Exercises:
  - Decrypt “Oh vruw hq hvw mhwh”.
  - Encrypt “Veni , vidi , vici”.

Vigenère cipher

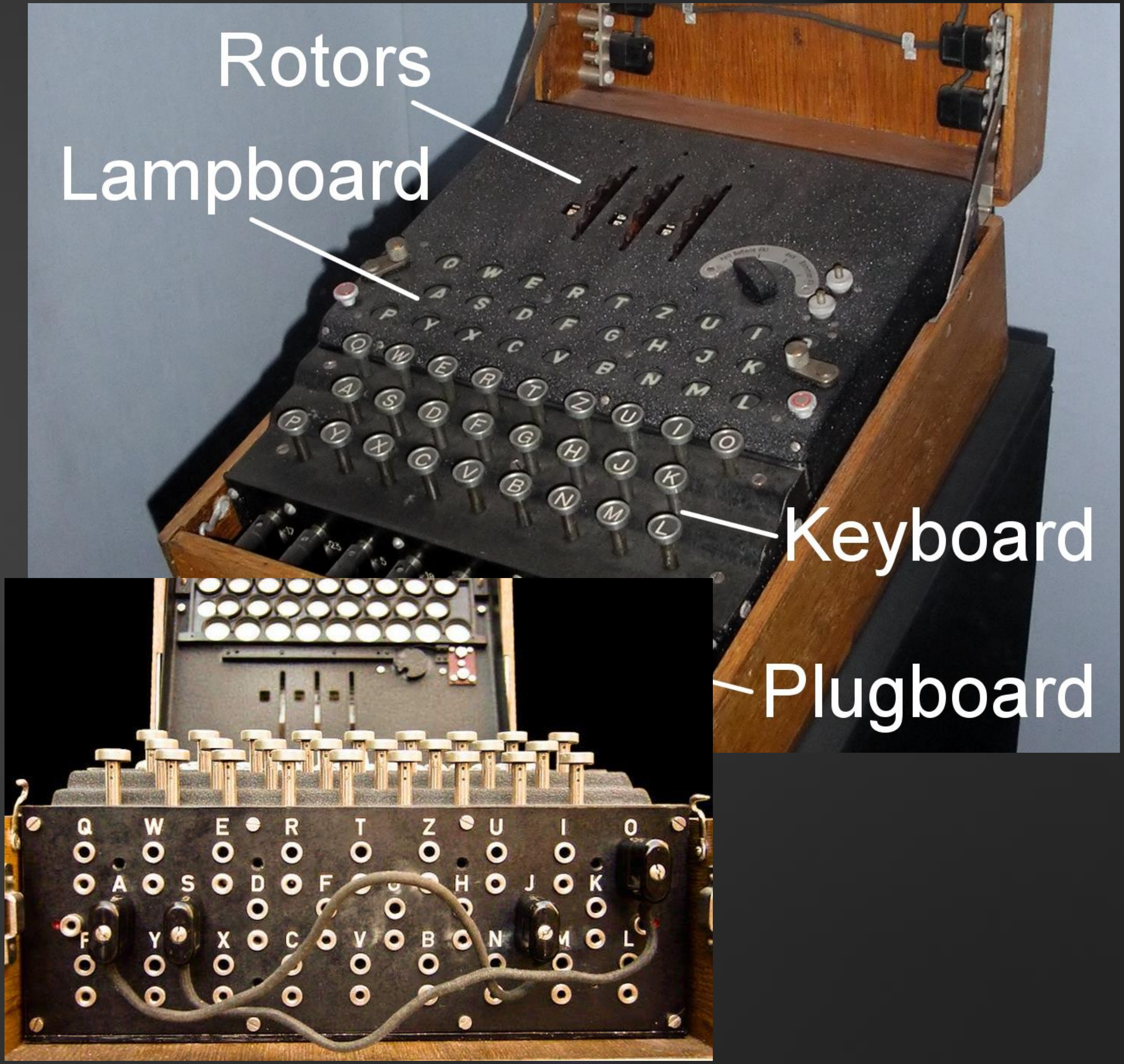
- It has several Caesar ciphers in sequence with different shift values.
- To encrypt and decrypt a table of alphabets must be used.
- The alphabet is written out 26 times in different rows, each alphabet shifted cyclically to the left compared to the previous alphabet, corresponding to the 26 possible Caesar ciphers.
- Exercises with "CRYPTO" key:
  - Encrypt "CHIFFRE".
  - Decrypt "XZETGSTV".



		Lettres en clair																									
		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
Lettres de la clé	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

Enigma machine

- Used by Nazi during World War 2.
- It was made up of:
  - **Plugboard**: one (or more) letter could be exchanged with another. When a key is pressed, the electric current first passes through the exchanged letter cable, before passing through the rotors.
  - 3 **rotors** (with 26 letters) must be chosen (each rotor moves the next like a clock).



INTRODUCTION TO CRYPTOGRAPHY

Concepts

- Hidden a secret for data confidentiality like authentication.
- Non repudiation (author cannot dispute).
- Data integrity : verify that the information received is exactly the same as the sent. Also called “tamper proof” or “tamper resistant”.

Terms

- **Cleartext / Plaintext** is an unencrypted message.
- **Cyphertext** is an encrypted message.
- **Cipher** is the algorithm used to encrypt and/or decrypt a message.
- **Cryptanalysis** is the art of cracking a cyphertext / an encryption (like mathematiciens who finding weaknesses in cyphers).

Terms : warning in French

- **Cryptologie** c'est la science de la cryptographie.
- **Chiffrement** c'est le procédé de la cryptographie.
- **Déchiffrer** c'est retrouver le message d'origine avec la clé de déchiffrement.
- **Décrypter** c'est retrouver le message d'origine sans la clé de déchiffrement.
- **/!\ Crypter /!** cela signifierait chiffrer un message sans la clé…

Keys

- A **key** is used by the cypher to encrypt message.
- Generally **larger key** are more secure against bruteforce.
- **Encryption can use one or more key**

Story of Alice, Bob and others

- **Alice and Bob:** Generic characters who want to exchange a message or a key.
- **Eve:** An eavesdropper / a passive attacker.

About time

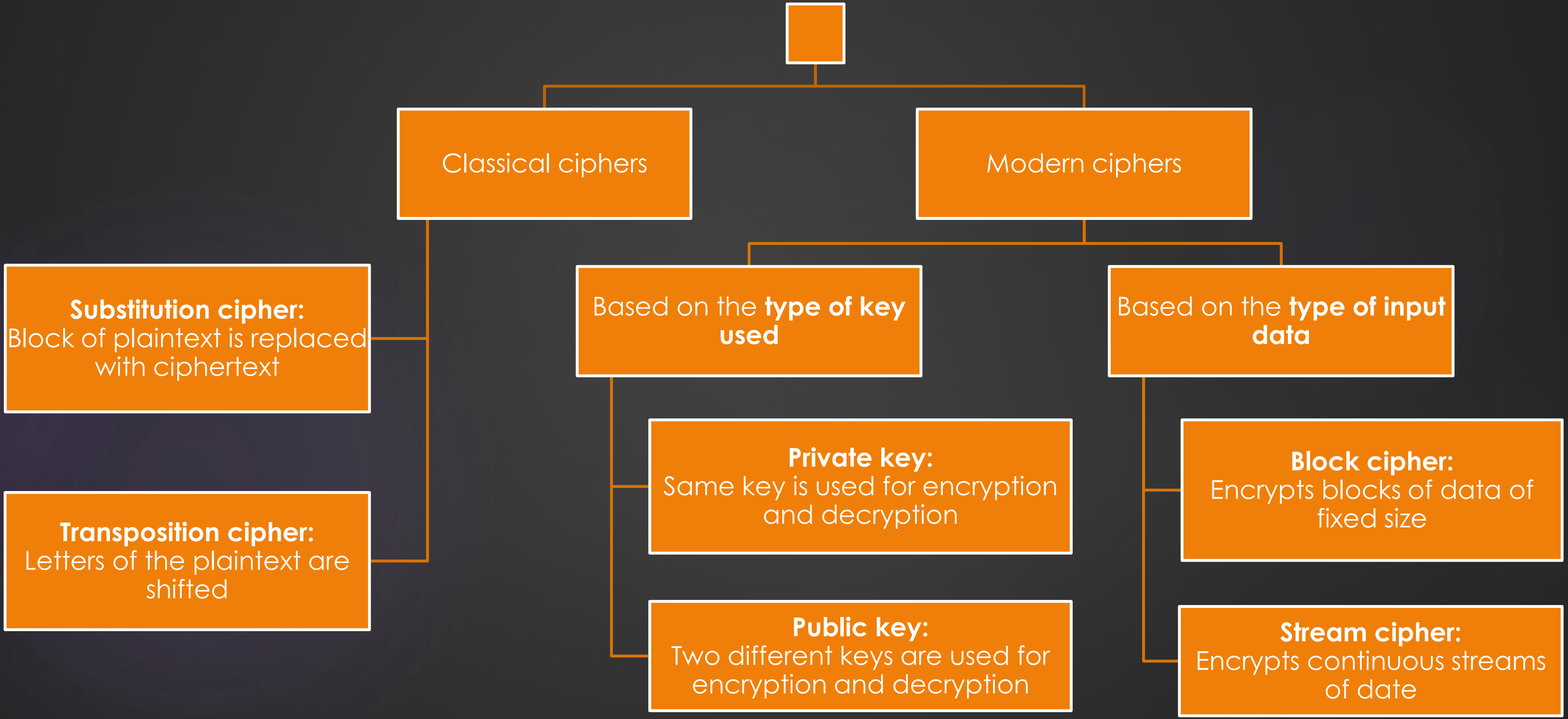
- Cryptographic methods use very large numbers. Table with values allowing to have an element of comparison with other values.

Title	Values
Number of seconds in a day	86 400 seconds
Number of seconds in a year	31,536,000 seconds 3 * 10 ^ 7 seconds
Number of seconds since the creation of the universe (13.7 billion years)	432 * 10 ^ 15 seconds
Number of atoms in the universe	10 ^ 80 atoms

- Now we can say that a message encrypted with the AES-256 algorithm and a machine that can make 100 billion attempts per second, it would take 10 ^ 58 seconds to test all the keys. This necessary computing time is much longer than the age of the universe.

Types of ciphers

Types of ciphers



Classical ciphers

Substitution cipher: Block of plaintext is replaced with ciphertext

Transposition cipher: Letters of the plaintext are shifted

Modern ciphers

Based on the **type of key used**

Private key: Same key is used for encryption and decryption

Public key: Two different keys are used for encryption and decryption

Based on the **type of input data**

Block cipher: Encrypts blocks of data of fixed size

Stream cipher: Encrypts continuous streams of date

STREAM CIPHER

Stream ciphers

- A stream cipher is a symmetric key cipher where plaintext digits are combined with a pseudorandom cipher digit stream (keystream).

- Stream ciphers are often used for their speed and simplicity of implementation in hardware, and in applications where plaintext comes in quantities of unknowable length like a secure wireless connection.

- Examples:
  - CryptMT,
  - RC4,
  - SEAL,
  - Etc.

BLOCK CIPHER

Introduction

- Blocks size must be fix.
- Initialization Vector (IV) is a unique random binary sequence used for each encryption operation.
- The IV is used to ensure distinct ciphertexts are produced even when the same plaintext is encrypted multiple times independently with the same key.
- The block cipher modes ECB, CBC, OFB, CFB, CTR, and XTS provide confidentiality, but they don't protect against accidental modification or malicious tampering.
- Modification or tampering can be detected with a separate message authentication code such as CBC-MAC, or a digital signature. The cryptographic community recognized the need for dedicated integrity assurances and NIST responded with HMAC, CMAC, and GMAC.

Common modes

- Authenticated encryption with additional data (AEAD) modes:
  - Galois/Counter Mode (GCM) means Mode compteur.

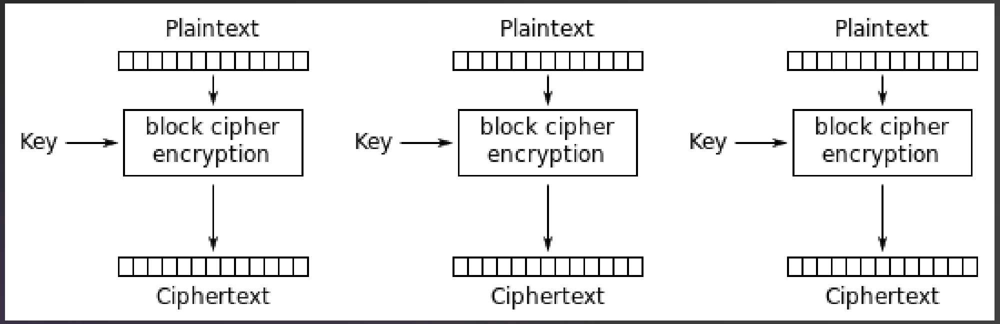
- Common modes:
  - Electronic Code Book, ECB means Dictionnaire de codes
  - Cipher Block Chaining, CBC means Enchaînement des blocs
  - Cipher Feedback, CFB means Chiffrement à rétroaction
  - Output Feedback, OFB means Chiffrement à rétroaction de sortie

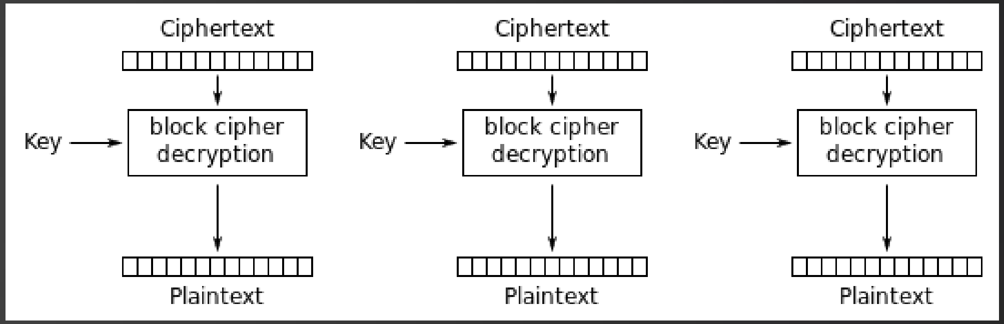
Galois/Counter Mode (GCM)

- GCM is an authenticated encryption algorithm designed to provide both data integrity, non-repudiation, and confidentiality. The encryption used is based on a counter mode in which the multiplication between the counter and the 128-bit key is performed in the Galois body.
- Galois Message Authentication Code (GMAC) is a variant limited to GCM authentication.
- The IEEE 802.1AE standard uses this algorithm, coupled with the Advanced Encryption Standard (AES).
- It’s a secure cipher mode (not a weak) unlike ECB, CBC, OFB, CFB, etc.

Electronic codebook (ECB)

- This mode is the simplest: the same block is always coded in the same way. There is no input or output feedback on the encryption function.





- Advantages
  - Encryption or decryption can be parallelized.
  - Machines or CPUs can work simultaneously on different parts of the message.
  - It allows random access in ciphertext. A one-bit transmission error only affects the decoding of the current block.
- Disadvantages
  - Plain text repeats are not masked and occur as cipher text repeats.
  - Complete portions of the message can be edited, repeated or replaced without difficulty.
  - Loss or addition of a bit is irrecoverable.

Electronic codebook (ECB)

Example with a salary list, the following two messages are encrypted with an ECB mode and a block cipher algorithm which works with a block of two characters.

JOHN__105000

JACK__500000

The encryption on the first message looks like this:

JO | HN | __ | 10 | 50 | 00

Q9 | 2D | FP | VX | C9 | IO

And on the second message, we get:

JA | CK | __ | 50 | 00 | 00

LD | AS | FP | C9 | IO | IO

We see that pairs of characters appear in the two encrypted messages, the same goes for the plain messages:

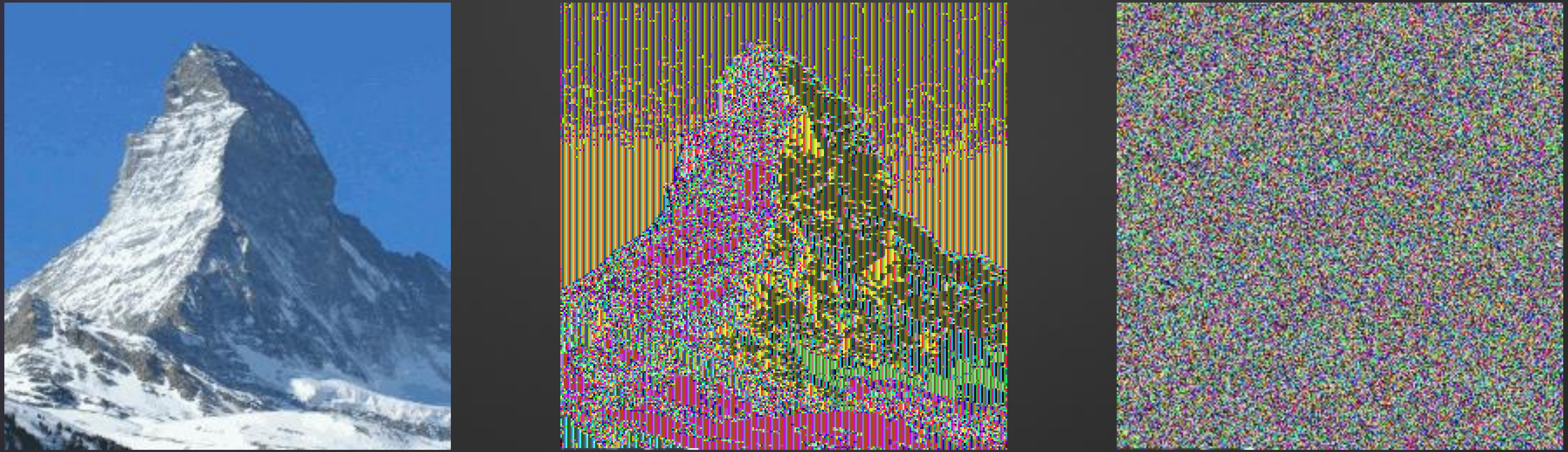
Q9 | 2D | FP | VX | C9 | IO

LD | AS | FP | C9 | IO | IO

Assuming that John knows his salary, he could guess Jack's salary because the sequence "C9" is "50" and "IO" is "00". John deduces that Jack's salary, quantified in "C9IOIO" corresponds to "500,000".

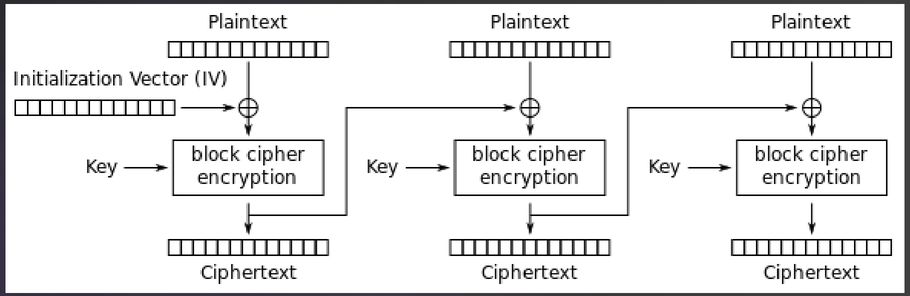
Electronic codebook (ECB)

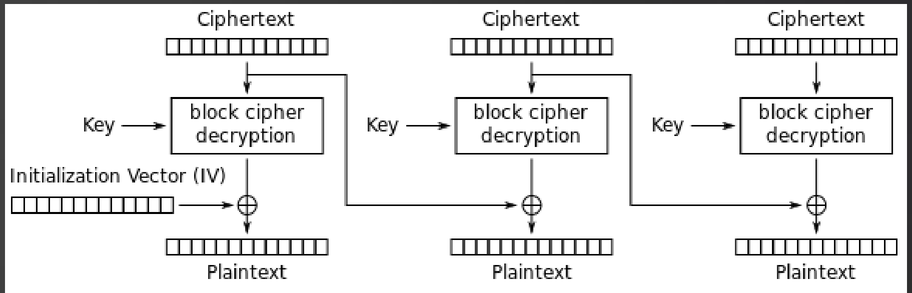
- The vulnerability is even more glaring in an image.
- Images consist of numerous redundancies which cause the blocks to be encrypted in the same way in ECB mode.
- In the example below, the ECB encryption is performed on blocks of 4 pixels. Last image is more secure (exemple with CBC).



Cipher Block Chaining (CBC)

- In this encryption mode, each block of plaintext is first combined by an **exclusive or** with the last block of the ciphertext. The output of this **exclusive or** is then applied to the encryption function.
- This encryption mode also has an **Initialization Vector** which allows the process to be initialized when no block has yet been encrypted.

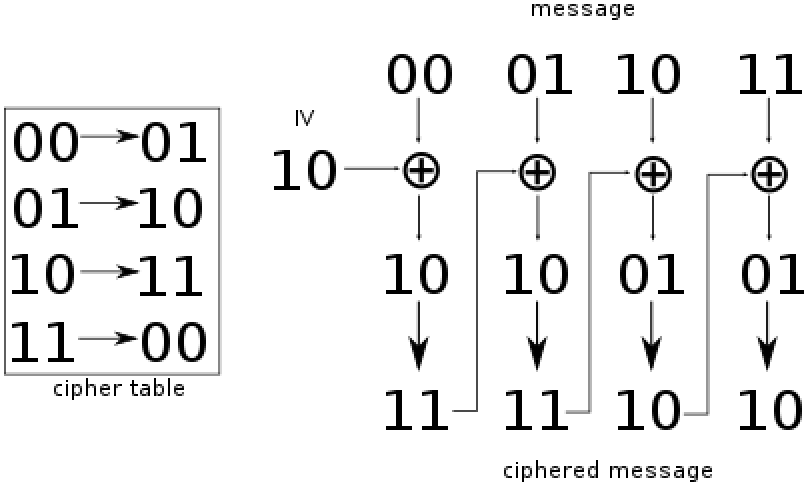


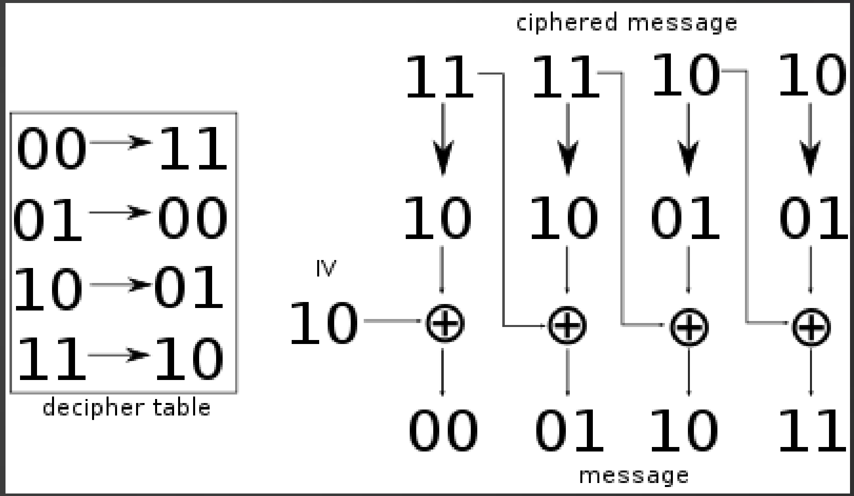


- Advantages
  - Clear text repetitions are hidden in the cipher text.
  - The value of the IV need not be secret.
- Disadvantages
  - Two identical plaintext will have the same beginning of cipher text.
  - A one-bit transmission error only affects the decoding of the current block as well as the decoding of the same bit in the following block;
  - The loss of synchronization (loss or addition of a bit) is irrecoverable.

Cipher Block Chaining (CBC)

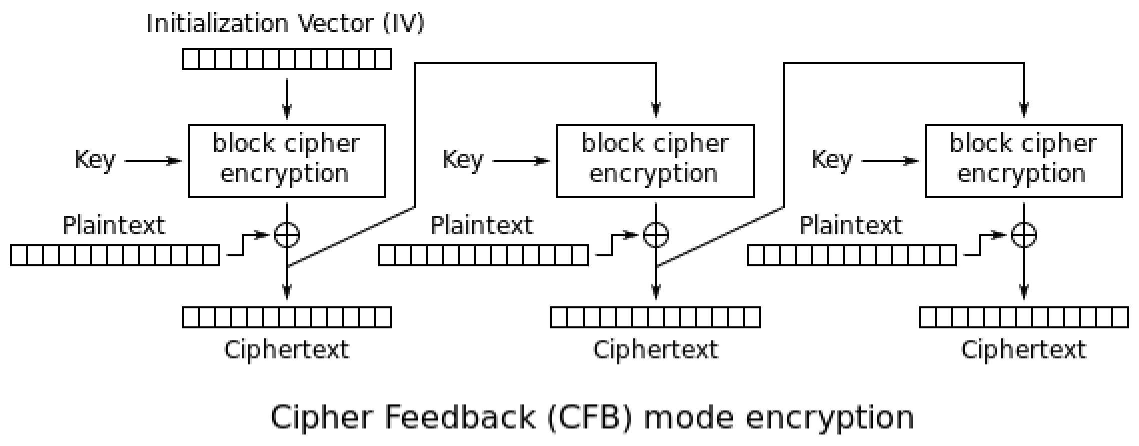
- Example :



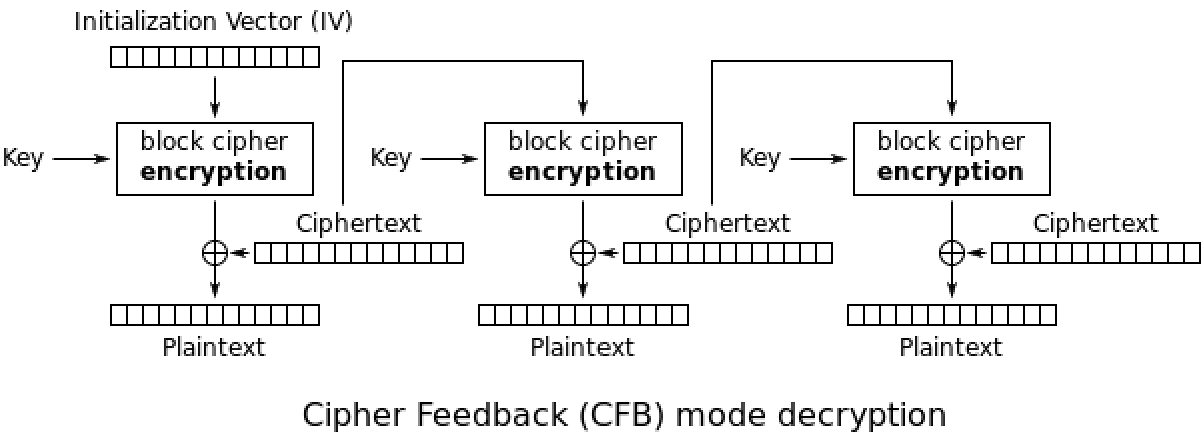


Cipher feedback (CFB)

- The cipher feedback (CFB) mode, in its simplest variation, is using the entire output of the block cipher. In this variation, it is very similar to CBC, makes a block cipher into a self-synchronizing stream cipher.



Cipher Feedback (CFB) mode encryption

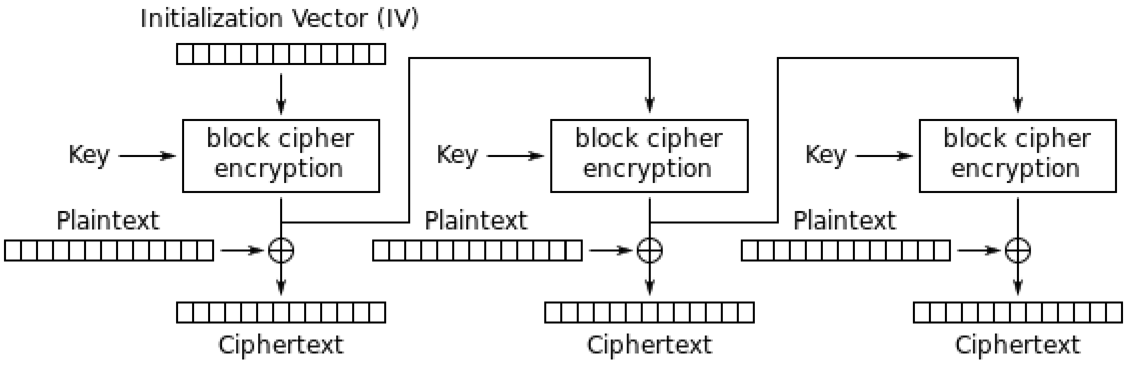


Cipher Feedback (CFB) mode decryption

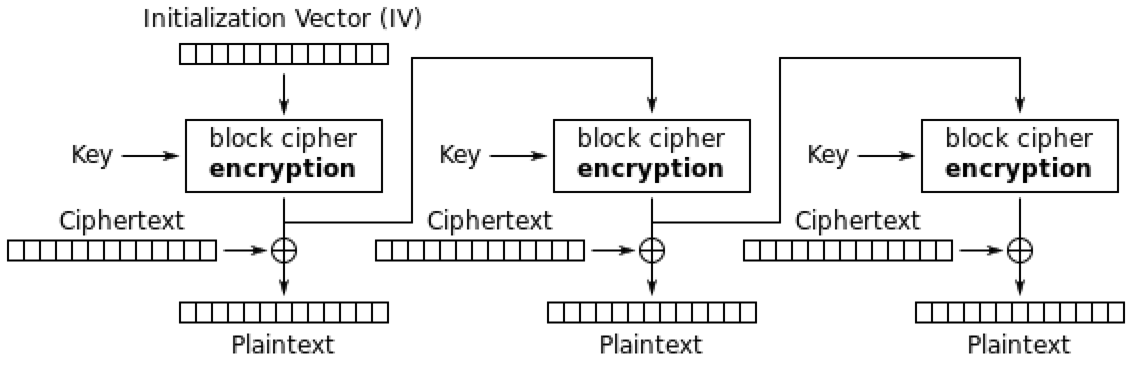
- Advantages
  - Clear text repetitions are hidden in the cipher text.
  - The value of the initialization vector IV need not be secret.
  - The loss of synchronization (loss or addition of a bit) is recoverable.
- Disadvantages
  - A one-bit transmission error only affects the decoding of the current block as well as the decoding of the same bit in the following block.

Output feedback (OFB)

- OFB mode resembles CFB mode. The only difference is that the byte injected into the shift register is the least significant byte of the ciphertext.



Output Feedback (OFB) mode encryption



Output Feedback (OFB) mode decryption

- Advantages
  - Clear text repetitions are hidden in the cipher text.
  - The value of the initialization vector IV need not be secret.
  - This mode does not amplify errors. A one-bit transmission error only affects that bit during decoding.
- Disadvantages
  - The loss of synchronization (loss or addition of a bit) is irrecoverable.

Conclusion

- After exploring the different encryption modes that can be used, we will now take a look at the different encryption algorithms.
- There are in all four major families of algorithms used in cryptography.
- These families are:
  - Hash functions,
  - Symmetric algorithms,
  - Asymmetric algorithms,
  - Methods of generating random numbers.

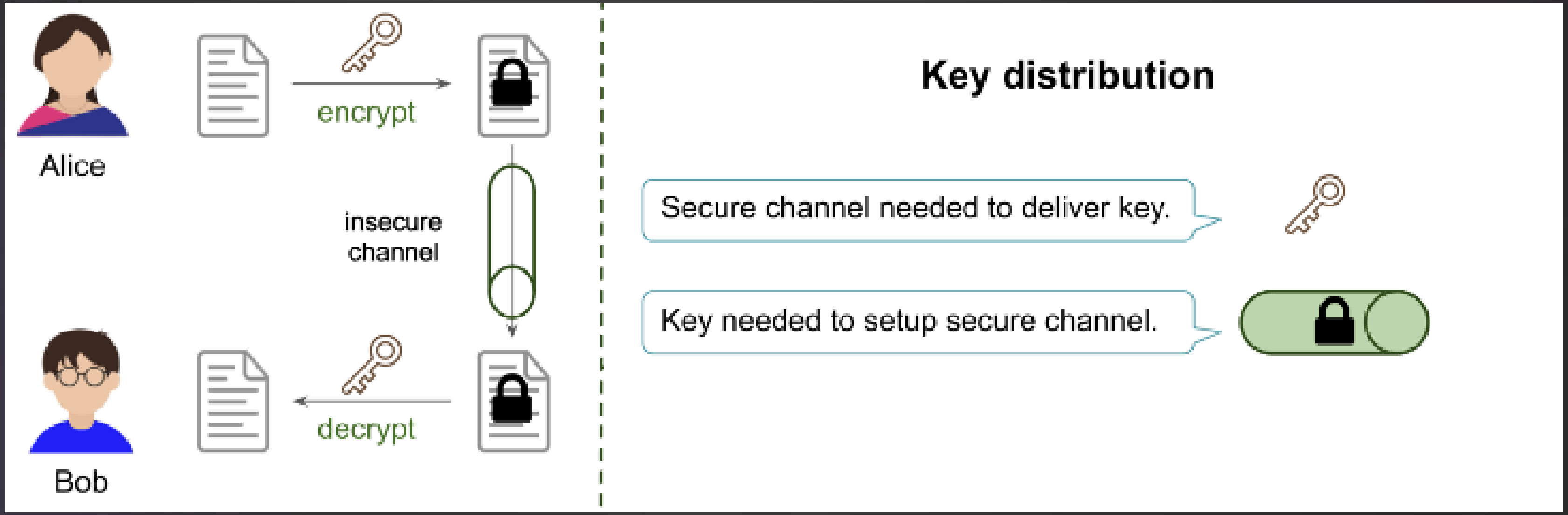
SYMMETRIC ENCRYPTION

Symmetric encryption

- **Single shared key** for encryption and decryption process.
- If **third party gains access** to this key you will need to throw that key away and you will **distribute a new key** both the sender and the recipient.
- Use a **shared key** algorithm (also called shared secret) only for people who need to either encrypt or decrypt this information.
- It's **difficult to distribute** these keys to everyone who might need it.
- **Very fast to encrypt and decrypt data** (compared to asymmetric encryption).
- You can **combine symmetric and asymmetric encryption** (example by encrypting a symmetric key using asymmetric encryption).

Symmetric encryption

- Sender and receiver share the same key for both encryption and decryption.


Alice
encrypt
insecure channel
Bob
decrypt
Key distribution
Secure channel needed to deliver key.
Key needed to setup secure channel.

SYMMETRIC ALGORITHMS

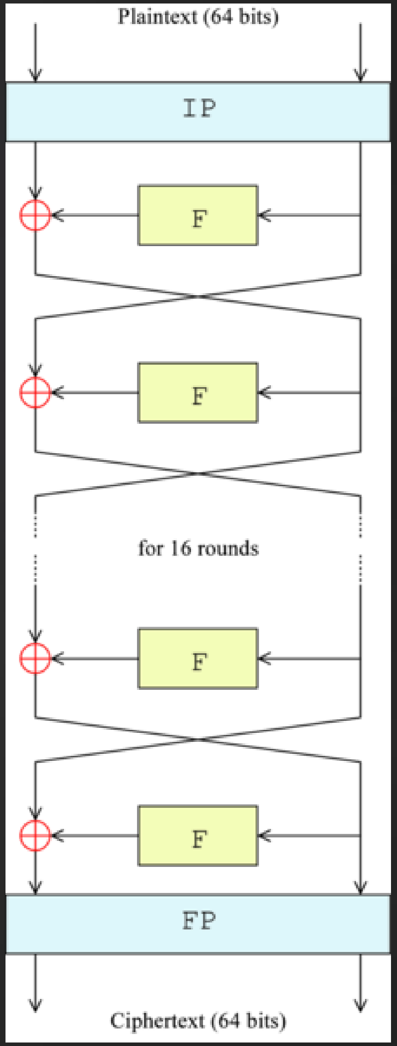
RC2, RC4, RC5 and RC6 Algorithms

- RCx for Rivest Cipher.
- All of these are vulnerable (**RC2** is vulnerable to a related-key attack, **RC4** have led to very insecure protocols such as WEP, etc.).

Algorithms	Category	Block sizes	Key sizes	Rounds
RC2	Block cipher	64 bits	8–1024 bits	16 of type MIXING, 2 of type MASHING
RC4	Stream cipher	NA	40-2048 bits	1-255
RC5	Block cipher	32, 64 or 128 bits	0-2040 bits	1-255
RC6	Block cipher	128 bits	128, 192, or 256 bits	20

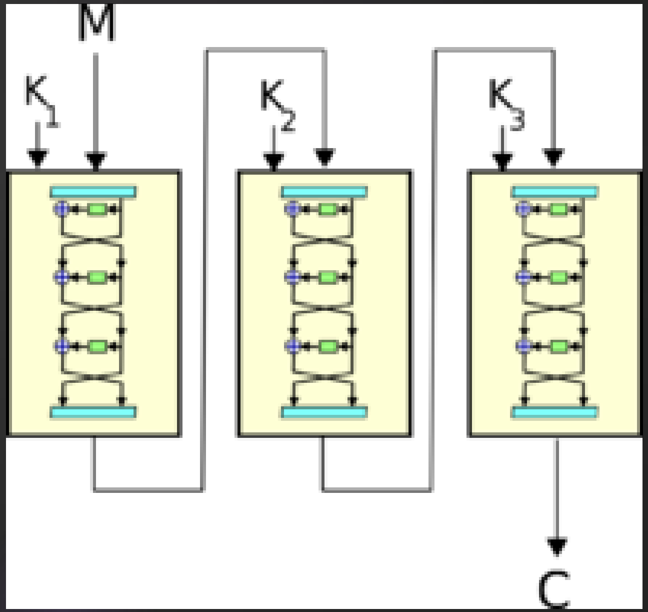
Data Encryprion Standard (DES)

- **DES** is designed to encrypt and decrypt blocks of data consisting of 64 bits under control of a 56 bit key.
  - Block sizes : 64 bits
  - Key sizes : 56 bits (+8 parity bits)
  - Rounds : 16 (Initial Permutation to Final Permutation with function F)
- **DES** has been considered **insecure** right from the start because short block size of 64 bits makes DES vulnerable to block collision attacks if it is used to encrypt large amounts of data with the same key.



Triple DES (3DES)

- **3DES** applies the DES cipher algorithm three times to each data block.
  - Block sizes : 64 bits
  - Key sizes : 168, 112 or 56 bits
  - Rounds : 3x16
- **3DES** is also considered **insecure** for same reason as DES. Feasibility of brute-force / collision attacks.
- **Vulnerability**
  Sweet32: Birthday attacks on 64-bit block ciphers



Blowfish

- **Blowfish** it is about 5 times faster than 3DES.
  - Block sizes : 64 bits
  - Key sizes : 32 at 448 bits
  - Rounds : 16
- **Blowfish** is currently considered **insecure**.
- **Blowfish's** use of a 64-bit block size (as opposed to e.g. AES's 128-bit block size) makes it **vulnerable to birthday attacks**. In 2016, the SWEET32 attack demonstrated how to leverage birthday attacks to perform plaintext recovery / decrypting ciphertext against ciphers with a 64-bit block size.

Twofish

- **Twofish** is the successor of Blowfish
  - Block sizes : 128 bits
  - Key sizes : 128, 192 or 256 bits
  - Rounds : 16

- **Twofish** is currently considered **secure** (finalist of the AES competition).

- **Twofish** is slightly slower than AES.
- Attractive alternative to the current AES if it becomes vulnerable.

Advanced Encryption Standard (AES)

- **AES** is a standardization process initiated by NIST to ask cryptologists to design a new block cipher algorithm for the US government.
- **Winner of the competition is Rijndael algorithm.**

- **AES** supersedes DES / 3DES. It use a method called Square.
  - Block sizes : 126 bits
  - Key sizes : 128, 192 or 256 bits
  - Rounds : 10, 12 or 14 (depending on key size)

- **Vulnerability** for all key sizes
  Biclique attack (based on MITM).

ASYMMETRIC ENCRYPTION

Asymmetric encryption

- Asymmetric encryption **involves both** a public key and a private key.
- **Public key** is publicly available and **Private key** must be kept secret.
- It'**s not the same key** that is used for encryption and decryption.
- They are intimately linked by a complex mathematical function (**one-way function**) is a function that is easy to compute on every input, but hard to invert. But a person with a piece of information (**private key**) can easily decode the message.
- Asymmetric algorithms have two modes of operation (**encryption mode** and **signature mode**).

Asymmetric encryption

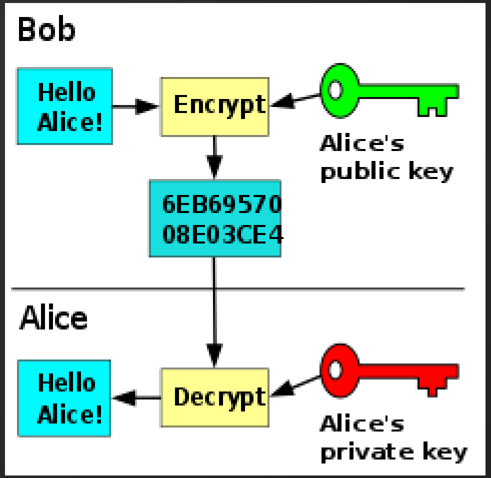
- Example :

Suppose Alice wants to receive a secret message from Bob on a channel that could be listened to by a passive attacker Eve.

1. Alice transmits to Bob a one-way function for which she alone knows the private key.
2. Bob uses the function sent by Alice to encrypt his secret message.
3. Alice receives the encrypted message then decodes it using the private key.
4. If Eve also receives the message while it is circulating on the public channel, she cannot decode it, because she has no knowledge of the private key.

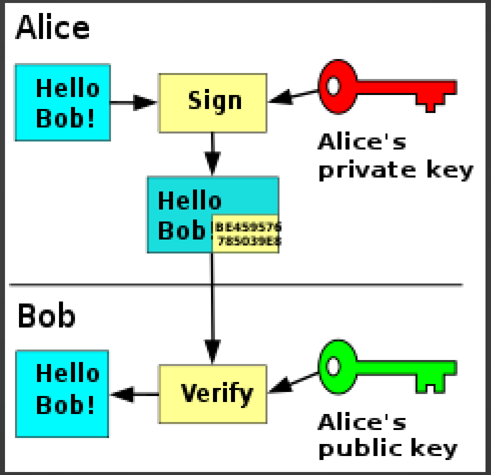
Asymmetric encryption scheme

- The sender encrypts with the recipient's public key, the recipient decrypts with her private key.
- In an asymmetric key encryption scheme, anyone can encrypt messages using the public key, but only the holder of the paired private key can decrypt.
- Security depends on the secrecy of the private key.



Digital Signatures (not encrypted)

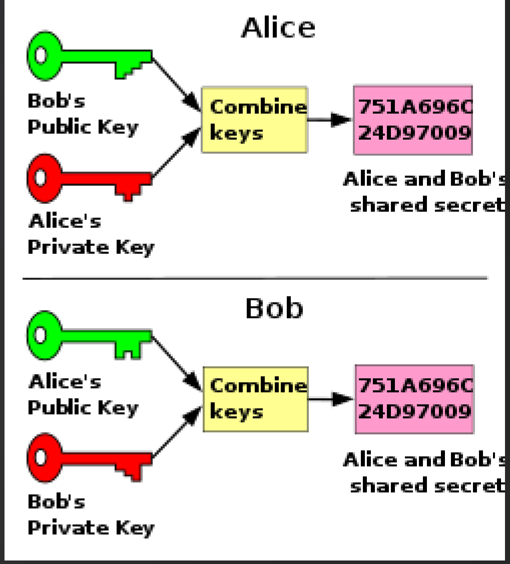
- The sender signs with her private key, the recipient verifies the signature with the sender's public key.



ASYMMETRIC ALGORITHMS

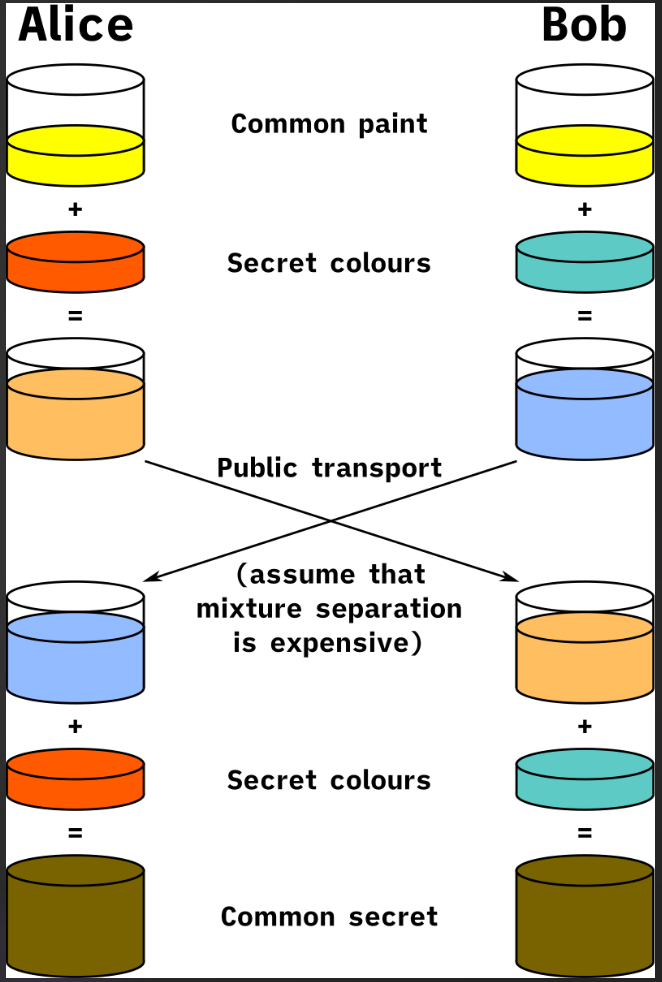
Diffie–Hellman (DH)

- **DH is a key exchange method** for securely exchanging cryptographic keys over a public channel.
- Traditionally, secure encrypted communication between two parties required that they first exchange keys by some secure physical means, such as paper key lists transported by a trusted courier.
- This method allows two parties that have no prior knowledge of each other to jointly establish a shared secret key over an insecure channel. This key can then be used to encrypt subsequent communications using a symmetric key cipher.



Diffie–Hellman (DH)

- Example with colors rather than numbers:
  - Alice and Bob, publicly agree on an arbitrary starting color that doesn't need to be kept secret.
  - Each person selects a secret color that they keep to themselves.
  - Each mix their own secret color with their mutually shared color.
  - Exchange publicly the two mixed colors.
  - Each of them mixes the color they received from the partner with their own private color.
  - If a third party listened to the exchange, it would only know the common color and the first mixed colors, but it would be difficult for this party to determine the final secret.
- To determine the color by calculation or with brute force is very time consuming.



Diffie–Hellman (DH)

- Function used:
  - DH turned their research towards non-reversible functions like functions of the form y = f (x) where when we know x, it is very easy to calculate y, but when we know y, it is impossible to recalculate x.
    - For example, the function y = x * x is a reversible function. Knowing x, it is easy to calculate y, and knowing y, it is easy to calculate x. The reversible function of y = x * x is x = $\sqrt{y}$.
  - DH used modulo function, because they are not reversible.
    - For example y = x (modulo 7) is not a reversible function. Indeed, for x = 11, we get y = 4 but for y = 4, we get x = 11 or x = 18 or even x = 25 and an infinity of other solutions.
  - The function used is of the type **y = Y ^ x (modulo P)** which allows this secure exchange.

Diffie–Hellman (DH)

- Function used with numbers. Alice wants to communicate with Bob and Eve can listen to the exchanges.
  - Alice says she will use the function **y = 11 ^ x modulo 13**
    - Bob receives the message and Eve heard / read this information.
  - Alice chooses a secret number **A = 5** and Bob **B = 8**.
    - Eve knows neither A nor B, Alice doesn't know B and Bob doesn't know A.
  - Alice calculates the number **A‘ = 11 ^ A modulo 13**. **A' = 7** and sends this number **A'= 7** to Bob and at the same time Bob does the same with B‘. **B‘=11 ^ B modulo 13 = 9** and send this number to Alice.
    - Eve heard / read these informations **A’=7** and **B’=9**.
  - Alice takes Bob's number and calculates the result is **B' ^ A modulo 13=3** Idem for Bob and Alice's number A'=7 the result is **A' ^ B modulo 13=3**
    - **Eve cannot calculate these numbers without A or B.**

Rivest Shamir Adleman (RSA)

- The algorithm they have developed is also based on a one-way function or rather a function that is very difficult to reverse.
- This difficult to reverse function is the **factorization of a number as a product of prime factors**.
- It is very easy to choose two numbers (48 and 52 for example) and to calculate their product (2496).
- It is long and painful to extract the prime factors of the number 2397 which are 47 and 51.
- The principle of the RSA algorithm is therefore based on the difficulty of factoring a number, especially when this number is very large and when it is the product of two very large prime numbers as well.

Rivest Shamir Adleman (RSA)

- Function used with numbers (lcm mean lowest common multiple):
  - Alice chose two distinct prime numbers, such as p=61 and q=53.
    - She compute **n = pq** = 61*53 = 3233
    - **λ(n) = lcm(p − 1, q − 1)** = lcm(60,52) = 780
    - She chose any number 1 < e < 780 that is coprime to 780 (E must not divide 780). Take **e = 17**
    - Alice calculates the modular multiplicative inverse **d = e (mod λ(n))**.
      So **d * e = 1 (mod λ(n)) ➔ d * 17 = 1 (mod 780) ➔ 413 * 17 = 1 (mod 780)**
      Because **413*17 = 7021 ≡ 1 (mod 780)**
      You can verify this by doing the following calculation : **7021 mod 780 = 1**
    - Alice calculates the number d given by **e * d = 1 modulo ((p-1) * (q-1))** (77 in our example). This number represents Alice's private key.
    - Public key is (n = 3233, e = 17). For a plaintext message m, the encryption function is : **c(m) = m^{17} mod 3233**
      Private key is (n = 3233, d = 413). For an encrypted ciphertext c, the decryption function is : **m(c) = c^{17} mod 3233**
  - Exercice: Convert the word “Hello” to ASCII and replace “m” with the value to have the encrypted message “c” then decrypt it.

HASH FUNCTIONS

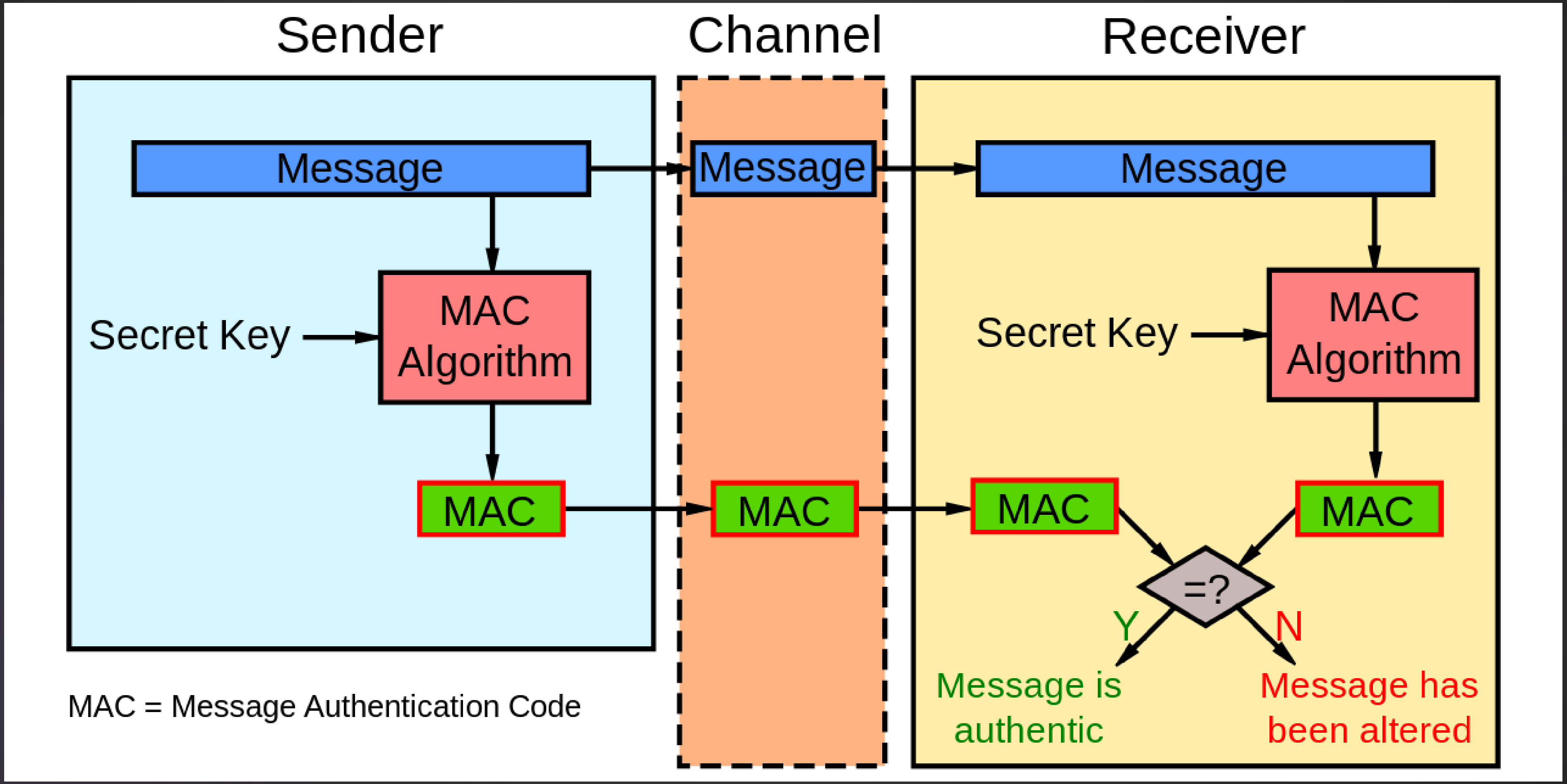
Introduction

- Hash function is a special function which, from data supplied as input, calculates a digital fingerprint used to quickly identify the initial data, just like a signature to identify a person. Hash functions are used in computing and in cryptography, in particular to quickly recognize files or passwords.
- This type of algorithm provides as a result a number whose size is fixed, whatever the size of the plain text input.
- A plaintext collision is when two different plaintext results in the same fingerprint.
- Checksum and CRC algorithms are not strong algorithms.
  - These two algorithms are not strong enough for cryptography, because the number of possible fingerprints is really too small (2 ^ 16 for the CHECKSUM algorithm or 2 ^ 32 for the CRC-32 algorithm to compare to 2 ^ 128 of the MD5 or even 2 ^ 160 of SHA-1).

Terms

- Message Authentication Code (MAC), is a short piece of information used to authenticate a message. In other words, to confirm that the message came from the stated sender (its authenticity) and hasn't been changed. The **MAC value protects** a message's **data integrity**, as well as its **authenticity**, by allowing verifiers (who also possess the secret key) to detect any changes to the message content.
- Hash-based Message Authentication Code (HMAC) is a specific type of message authentication code (MAC) involving a cryptographic hash function and a secret cryptographic key. HMAC can provide digital signatures using a shared secret instead of public key encryption.

Terms in image



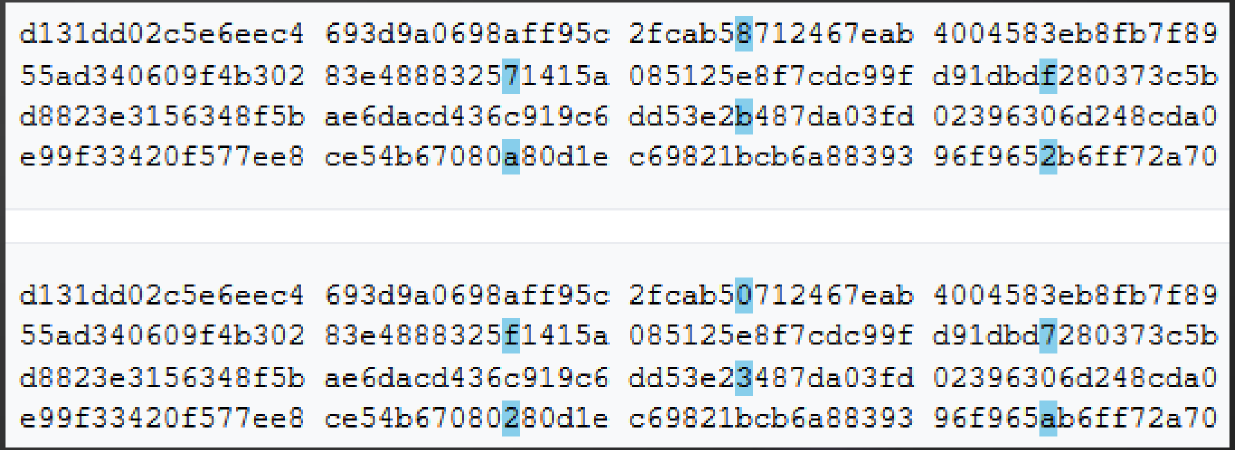
MAC = Message Authentication Code

Collision

- A hash function's “collision” is a pair of data distinct from its starting set, the checksums of which are identical.
- Collisions are generally considered undesirable but are generally impossible to avoid due to the size difference between the starting and ending sets of the function.
- This situation is considered rare, even impossible, depending on the quality level of the hash function. This is what makes it possible to consider that a file (or a password) corresponds to a unique signature. And therefore a given signature can only come from a single starting file (or password).
- A birthday attack is a type of cryptographic attack that exploits the mathematics behind the birthday problem in probability theory.

Collision

- Example:
  MD5 Collision Demo with
  these two blocks.
- Each of these blocks has
  MD5 hash :
  79054025255fb1a26e4bc422aef54eb4.



```
d131dd02c5e6eec4 693d9a0698aff95c 2fcab58712467eab 4004583eb8fb7f89
55ad340609f4b302 83e488832571415a 085125e8f7cdc99f d91dbdf280373c5b
d8823e3156348f5b ae6dacd436c919c6 dd53e2b487da03fd 02396306d248cda0
e99f33420f577ee8 ce54b67080a80d1e c69821bcb6a88393 96f9652b6ff72a70

d131dd02c5e6eec4 693d9a0698aff95c 2fcab50712467eab 4004583eb8fb7f89
55ad340609f4b302 83e4888325f1415a 085125e8f7cdc99f d91dbd7280373c5b
d8823e3156348f5b ae6dacd436c919c6 dd53e23487da03fd 02396306d248cda0
e99f33420f577ee8 ce54b67080280d1e c69821bcb6a88393 96f965ab6ff72a70
```

Length extension attack

- A length extension attack is a type of attack where an attacker can use Hash (message1) and the length of message1 to calculate Hash(message1 concatenate with message2) for an attacker-controlled message2, without needing to know the content of message1.
- Algorithms like MD5, SHA-1 and most of SHA-2 that are based on the Merkle–Damgård construction are susceptible to this kind of attack.
- When a Merkle–Damgård based hash is misused as a message authentication code with construction H(secret concatenate with message), and message and the length of secret is known, a length extension attack allows anyone to include extra information at the end of the message and produce a valid hash without knowing the secret.
- Since HMAC doesn’t use this construction, HMAC hashes aren’t prone to length extension attacks.

Message Digest

- The MD2, MD4, MD5 and MD6 algorithms are digital fingerprint calculation algorithms.
- If the MD5 algorithm is of significant historical interest, it is today considered to be outdated and absolutely unsuitable for any use in cryptography or in security.
- MD6 was developed to participate in the 2008 NIST hash function competition but was not selected.

Message Digest algorithm

- MD5 processes a variable-length message into a fixed-length output of 128 bits.
- The input message is broken up into chunks of 512-bit blocks (sixteen 32-bit words).
- The message is padded first a single bit “1”, is appended to the end of the message. This is followed by as many zeros as are required to bring the length.

Message Digest attacks

- John the ripper allows you to crack / reverse function MD5 by brute force.
- Rainbow tables can often crack in less than a second.
  - These tables use dictionaries established after several days, months or years of calculation.
  - These don’t contain all of the possible MD5 keys, nor are they intended for brute force breaking (one hash has 128 bits, which represents about 4.10^38 of combinations).
  - These allow by examining the footprint to eliminate very large classes of combinations not to be tested.
  - The efficiency of rainbow tables decreases if the footprint is calculated with a "salt".

Secure Hash Algorithms

- The Secure Hash Algorithms are a family of cryptographic hash functions published by the National Institute of Standards and Technology (NIST) as a U.S. Federal Information Processing Standard (FIPS), including:
  - SHA-0: It was withdrawn shortly after publication due to an undisclosed "significant flaw" and replaced by the slightly revised version SHA-1.
  - SHA-1: Designed by the NSA to be part of the Digital Signature Algorithm. Cryptographic weaknesses were discovered in SHA-1, and the standard was no longer approved for most cryptographic uses after 2010.
  - SHA-2: Two similar hash functions with different block sizes SHA-256 and SHA-512. They differ in the word size; SHA-256 uses 32-byte words where SHA-512 uses 64-byte words. There are also truncated versions of each standard, known as SHA-224, SHA-384, SHA-512/224 and SHA-512/256.
  - SHA-3: A hash function formerly called Keccak, chosen in 2012 after a public competition among non-NSA designers. It supports the same hash lengths as SHA-2, and its internal structure differs significantly from the rest of the SHA family.

Secure Hash Algorithms

Algorithm and variant	Output size (bits)	Internal state size (bits)	Block size (bits)	Rounds	Security (in bits) against collision attacks	Capacity against length extension attacks	First published
MD5	128	128	512	64	**≤18 (collisions found)**	**0**	1992
SHA-0	160	160	512	80	**<34 (collisions found)**	**0**	1993
SHA-1	160	160	512	80	**<63 (collisions found)**	**0**	1995
SHA-2: - SHA-224 - SHA-256	 224 256	 256 256	 512 512	 64 64	 **112** **128**	 **32** **0**	 2004 2001
SHA-2 : - SHA-384 - SHA-512	 384 512	 512 512	 1024 1024	 80 80	 **192** **256**	 **128** **0**	 2001 2001
SHA-2 : - SHA-512/224 - SHA-512/256	 224 256	 512 512	 1024 1024	 80 80	 **112** **128**	 **288** **256**	 2012 2012
SHA-3: - SHA3-224 - SHA3-256 - SHA3-384 - SHA3-512	 224 256 384 512	 1600 1600 1600 1600	 1152 1088 832 576	 24 24 24 24	 **112** **128** **192** **256**	 **448** **512** **768** **1024**	 2015 2015 2015 2015

STEGANOGRAPHY

Introduction to steganography

- It's an obfuscation technique. Security through obscurity is the belief that a system of any sort can be secure so long as nobody outside of its implementation group is allowed to find out anything about its internal mechanisms.
- It's the practice of concealing a file, message, image, or video within another file, message, image, or video.

Steganography techniques

- **Network:**
  Example using ICMP packets. You will need to fill in the data field with your data.
- **Image:**
  Example embed a message or a file in the image.
- **Machine Identification Code / Watermarks:**
  Example with printers “yellow dots”. Is a digital watermark which certain color laser printers and copiers leave on every single printed page. This technique allows identification of the device with which a document was printed and giving clues to the originator.

Steganography labs

1. Found the secret message in “Happy-dog_embed.jpg” file.
   Stegosuite is recommended.

2. Create a file with “Steganography_1.jpg” and the secret message “Joconde” file. Give the SHA1.
   Stegosuite is recommended.