

# Introduction à la cryptographie

# Introduction

- ❑ La cryptologie est la technique qui traite de la communication en présence d'adversaires
- ❑ Le but d'un système cryptographique est de chiffrer un texte clair  $P$  en un cryptogramme  $C$  au moyen d'une clé  $K$ .
- ❑ Ce cryptogramme est ensuite transmis à un destinataire sur le canal.
- ❑ Le destinataire légitime doit pouvoir déchiffrer le cryptogramme  $C$  à l'aide de la clé  $K$ .
- ❑ *Stéganographie* : Branche particulière de la cryptographie qui consiste non pas à rendre le message inintelligible, mais à le camoufler dans un support (texte, image, etc.) de manière à masquer sa présence.

# Domaines d'applications

- ☐ Secret militaire, diplomatique
- ☐ Secret industriel
- ☐ Transactions commerciales
- ☐ Droits d'auteurs
- ☐ Protection de la vie privée
- ☐ Communications numériques

Le secret est nécessaire, partout, tout le temps...

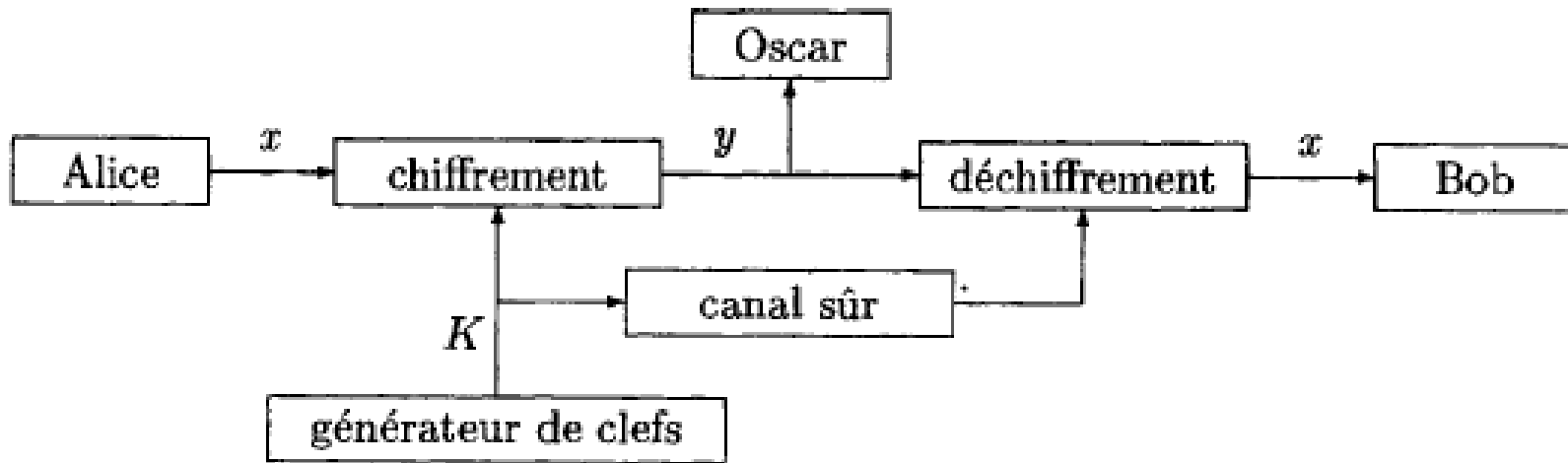
# Objectifs de la cryptographie

- ❑ Objectifs de la cryptologie
  - Confidentialité
  - Authenticité
  - Intégrité
- ❑ Ne couvre qu'une partie du problème de la sécurité informatique qui compte également :
  - *Non-répudiation*
  - *Disponibilité*

## Suite

- ❑ **Confidentialité** : s'assurer que l'information *n'est seulement accessible qu'à ceux dont l'accès est autorisé*
- ❑ **Authenticité** : L'authentification est la *procédure qui consiste, pour un système informatique, à vérifier l'identité d'une entité (personne, ordinateur...), afin d'autoriser l'accès de cette entité à des ressources (systèmes, réseaux, applications...)*.
  - Ne pas confondre *authentification* avec *identification*
    - Authentification : **vérifier l'identité**
    - Identification : **connaître l'identité**
- ❑ **Intégrité des données** : s'assurer que les données ont, ou non, été modifiées

# Modèle de Shannon pour le secret



- Un ennemi (cryptanalyste) ne doit pas être en mesure de réaliser une opération de cryptanalyse (décrypter) qui lui permet de retrouver le clair sans connaître  $K$ .
- Le cryptanalyste peut tenter selon différentes techniques d'attaques.


## Autres notions

- ❑ **Cryptanalyse** : Art d'analyser un message chiffré afin de le décrypter. On parle aussi de décryptement.
- ❑ **Déchiffrement** : Opération inverse du chiffrement, i.e. obtenir la version originale d'un message qui a été précédemment chiffré en connaissant la méthode de chiffrement et les clefs.
- ❑ **Décryptement** : Restauration des données qui avaient été chiffrées à leur état premier ("en clair"), sans disposer des clefs théoriquement nécessaires.

# Concepts de la cryptographie

- ❑ Il est possible de résumer la philosophie de la cryptographie moderne selon le principe de Kerchoff (1883):
  - " la sécurité d'un système de chiffrement ne doit pas dépendre du secret de l'algorithme mais seulement du secret de la clé "
- ❑ La cryptographie repose sur les systèmes de chiffrement à clé secrète:
  - Un espace de messages  $M$ : ensemble de mots sur l'alphabet des messages en clair;
  - Un espace de cryptogrammes  $C$ : ensemble des cryptogrammes
  - Un espace de clés  $K$ : ensemble des clés sur un alphabet



- 
- ❑ On pourrait envisager de fonder la confidentialité sur le fait que seules les personnes autorisées connaissent les algorithmes E et D:
    - il faut changer les algorithmes chaque fois qu'un initié, quitte le groupe.
  - ❑ On préfère rendre publics les algorithmes utilisés.
  - ❑ On fonde alors la confidentialité sur le fait que seules les personnes autorisées connaissent la clé de déchiffrement  $K_2$ .

## Suite

- Un algorithme de chiffrement qui est une application


$$\mathcal{E}: \mathcal{K} \times \mathcal{M} \longrightarrow \mathcal{C}$$

- Un algorithme de déchiffrement qui est une application


$$\mathcal{D}: \mathcal{K} \times \mathcal{C} \longrightarrow \mathcal{M}$$

- Les deux algorithmes  $\mathcal{E}$  et  $\mathcal{D}$  doivent vérifier:

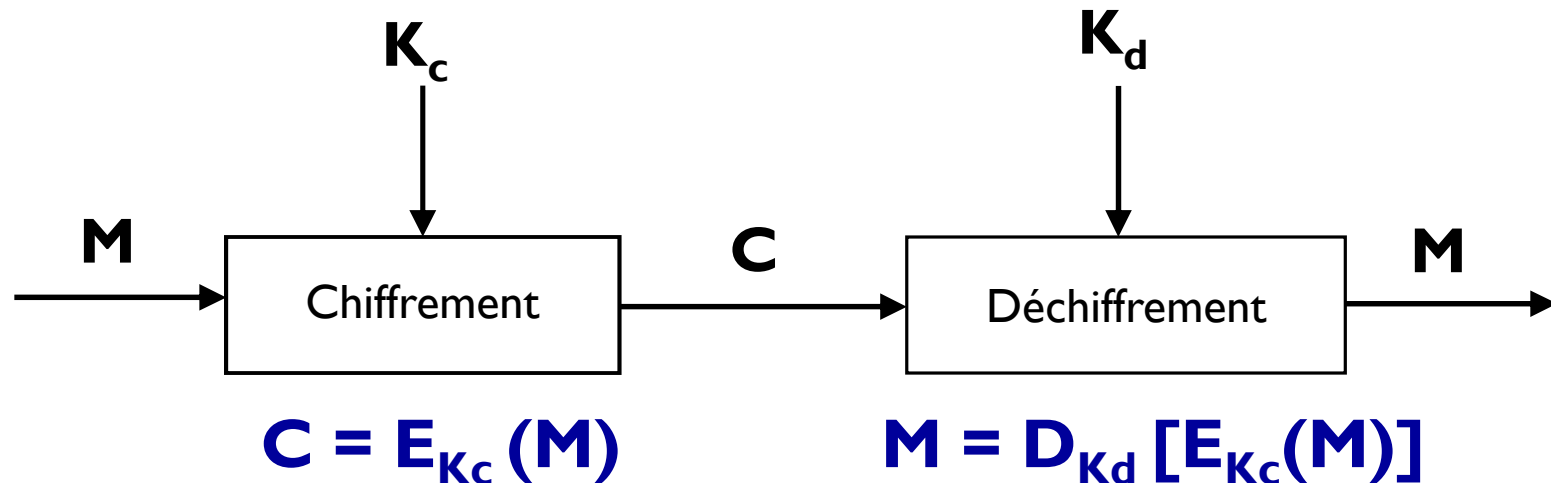
$$\mathcal{D}(\mathcal{K}_2, \mathcal{E}(\mathcal{K}_1, M)) = M, \forall \mathcal{K} \in \mathcal{K}, \forall M \in \mathcal{M}$$



**Cryptographie = Science de l'écriture secrète**  
Issu du grec **cryptos** (caché ou secret) et **graphie** (écriture)

- ❑ Service de base fourni par la **cryptographie**:  
capacité de transmettre une information entre deux correspondants sans que celle-ci ne soit accessible à des tierces personnes.
  - ❑ Informations de nature critique et nécessitant plus de protection
-  Recours à la **cryptographie** ou **chiffrement** pour renforcer la confidentialité des données et le contrôle d'accès à ces données.

# Système cryptographique



**M** : Message en clair

**C** : Message **chiffré** ou **Cryptogramme**

- **Chiffrement** (Emetteur): Algorithme **E**, Clé **K<sub>c</sub>**
- **Déchiffrement** (Récepteur): Algorithme **D**, Clé **K<sub>d</sub>**

# Caractéristiques d'un algorithme cryptographique

## ❑ Réversibilité

- **Réversible**: il est possible de retrouver M à partir de C en appliquant la transformation inverse.
- **Irréversible**: aucun moyen de retrouver M à partir de C.
  - Sert à vérifier le contenu de M



Destiné aux contrôles d'intégrité et/ou origine.

## ❑ Symétrie


- **Symétrique**:  $K_c$  et  $K_d$  sont les mêmes, ou peuvent se déduire **facilement** l'une de l'autre.
- **Asymétrique**:  $K_c$  et  $K_d$  ne peuvent se déduire **facilement** l'une de l'autre.



## ❑ Publication

- Publié: spécifications disponibles et décrites dans la littérature.
- Non publié: spécifications gardées secrètes par les concepteurs et/ou les commanditaires de l'algorithme.

## ❑ Implémentation

- Logicielle (programme): portabilité.
- Matérielle (circuit intégré): rapidité et protection,  
 Indispensable pour les algorithmes non publiés.



## ❑ Mode de chiffrement

- Chiffrement par blocs de longueur fixe de  $M$ .
- Chiffrement par flux ou en continu des éléments binaires de  $M$ .


## ❑ Autorisation de mise en œuvre

- L'usage des algorithmes de cryptographie en France est réglementé et soumis au régime des matériels de guerre de seconde catégorie.
- Leur mise en œuvre est régie par des décrets sur les démarches à effectuer pour la fabrication, la commercialisation, l'acquisition, la détention et l'utilisation des moyens cryptographiques.

## ❑ Résistance

- Mesure la capacité d'un algorithme à résister à la **cryptanalyse** exprimée en termes de temps et de moyens nécessaires.
- Cryptanalyse: méthodes mises en œuvre par un intrus afin de retrouver des informations secrètes (clés, message en clair) à partir d'informations publiques (cryptogrammes, algorithmes).



- 
- Attaque à texte chiffré connu: l'opposant ne connaît que le message chiffré.
  - Attaque à texte clair connu: l'opposant dispose d'un texte clair  $x$  et du message chiffré correspondant  $y$
  - Attaque à texte clair choisi : l'opposant a accès à une machine chiffrente. Il peut choisir un texte clair et obtenir le texte chiffré correspondant  $y$ , mais il ne connaît pas la clef de chiffrement.
  - Attaque à texte chiffré choisi : l'opposant a accès à une machine chiffrente. Il peut choisir un texte chiffré,  $y$  et obtenir le texte clair correspondant  $x$ , mais il ne connaît pas la clef de déchiffrement.

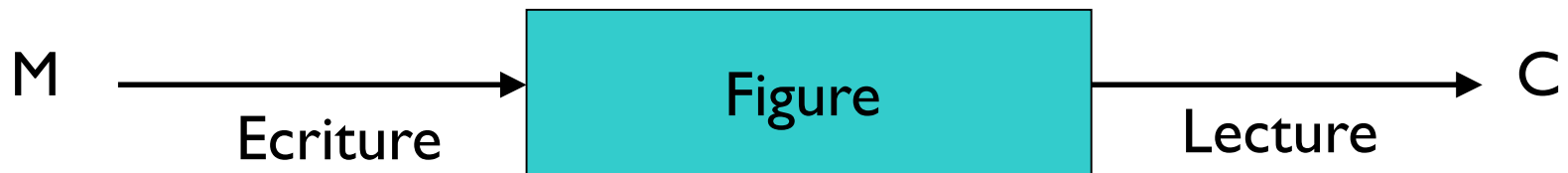
## Résistance: Méthodes de cryptanalyse

- **Systématiques:** Essais successifs sur la liste exhaustive des clés.
  - Possible si moyens illimités (temps de calcul, espace mémoire).
  - Protection: choisir des clés longues et des algorithmes longs.
- **Analytiques:** Relations entre K, M et C.
  - Protection: usage d'algorithmes complexes exprimés par un système d'équations
- **Statistiques:** Relations statistiques K, M et C.
  - Protection: brouiller les statistiques et réduire la redondance



# **Quelques techniques de chiffrement**

## ❑ Réarrangement des caractères selon une figure donnée



- Figure géométrique à 2-dimensions (Matrice)
- Clé: Figure + Modes d'Ecriture / Lecture.

## □ Transposition par colonnes

- Le message en clair M est transcrit dans une matrice (n,m).
- Le message chiffré C est obtenu en prenant les colonnes dans un certain ordre.

M : CONFIDENTIEL

Ordre: 

2	4	1	3
---	---	---	---

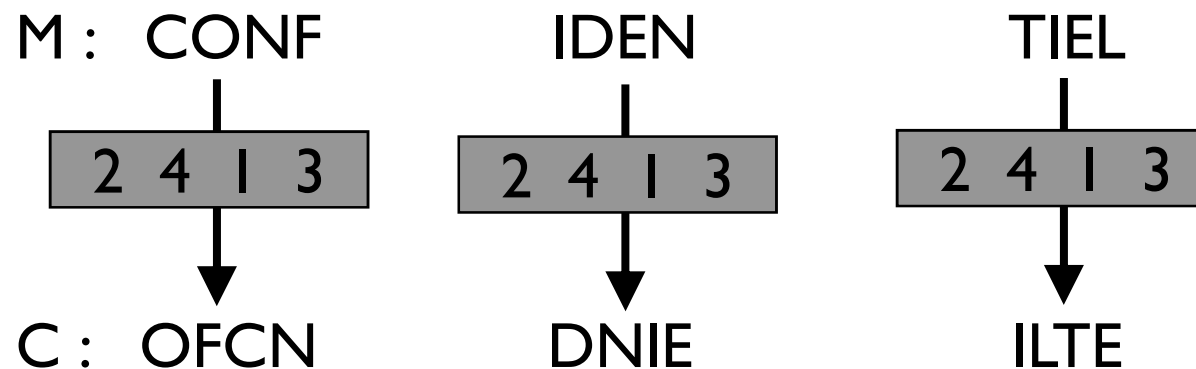
C : ODIFNLCITNEE

1	2	3	4
C	O	N	F
I	D	E	N
T	I	E	L

## ❑ Transpositions périodiques

- Le message en clair **M** est décomposé en blocs de taille fixe.
- Le message chiffré **C** est obtenu en prenant les caractères de chaque bloc selon un ordre donné.

Exemple:  $d=4$



# Algorithmes de substitution

## □ Substitutions simples

- Chaque caractère de M provenant de l'alphabet A est remplacé par le caractère correspondant dans un alphabet de substitution S.

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

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

M: C O N F I D E N T I E L



C: S M L B G A R L T G R J

# Cryptanalyse des chiffres par substitution monoalphabétique

- ❑ Le nombre de substitutions est phénoménal : 26! ( $\sim 4 \cdot 10^{26}$ )

- L'age de l'univers :  $\sim 4 \cdot 10^{25}$  ms

- ❑ La méthode est-elle robuste ?

- Non, pas du tout...



- ❑ Le message clair est rédigé dans une langue dont les propriétés statistiques sont connues

- *Attaque par analyse fréquentielle*

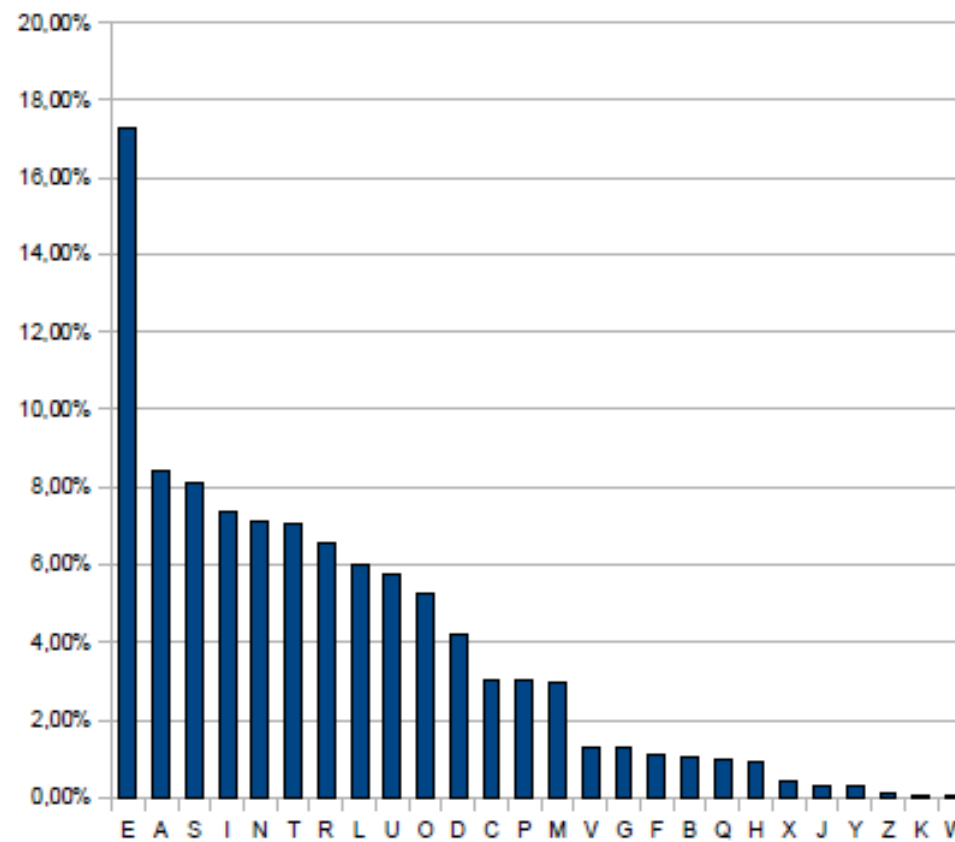
- *Possible si le texte est suffisamment long*



- ❑ Exemple : structure des textes français (étude sur 100 000 caractères)

<i>Lettre</i>	<i>Fréquence</i>	<i>Lettre</i>	<i>Fréquence</i>
<i>A</i>	8.40 %	<i>N</i>	7.13 %
<i>B</i>	1.06 %	<i>O</i>	5.26 %
<i>C</i>	3.03 %	<i>P</i>	3.01 %
<i>D</i>	4.18 %	<i>Q</i>	0.99 %
<i>E</i>	17.26 %	<i>R</i>	6.55 %
<i>F</i>	1.12 %	<i>S</i>	8.08 %
<i>G</i>	1.27 %	<i>T</i>	7.07 %
<i>H</i>	0.92 %	<i>U</i>	5.74 %
<i>I</i>	7.34 %	<i>V</i>	1.32 %
<i>J</i>	0.31 %	<i>W</i>	0.04 %
<i>K</i>	0.05 %	<i>X</i>	0.45 %
<i>L</i>	6.01 %	<i>Y</i>	0.30 %
<i>M</i>	2.96 %	<i>Z</i>	0.12 %

## Représentation par Histogramme



## Les 20 digrammes les plus fréquents

Bigrammes	ES	DE	LE	EN	RE	NT	ON	ER	TE	EL	AN	SE	ET	LA	AI	IT	ME	OU	EM	IE
Fréquence	3318	2409	2366	2121	1885	1694	1646	1514	1484	1382	1378	1377	1307	1270	1255	1243	1099	1086	1056	1030

## Les 20 trigrammes les plus fréquents

Trigrammes	ENT	LES	EDE	DES	QUE	AIT	LLE	SDE	ION	EME	ELA	RES	MEN	ESE	DEL	ANT	TIO	PAR	ESD	TDE
Fréquence	900	801	630	609	607	542	509	508	477	472	437	432	425	416	404	397	383	360	351	350

## Exemple à décrypter

1	J	A	F	R	U	J	B	T	F	U	G	D	U	W	B	T	A	F	F	U
2	V	X	E	C	O	Q	E	V	U	X	A	F	R	U	C	A	X	E	T	F
3	U	B	V	I	T	J	B	A	T	C	U	Q	F	T	M	U	C	J	U	V
4	V	U	E	X	H	T	B	T	U	Q	G	U	B	H	C	T	V	V	E	F
5	B	T	V	J	E	W	W	Q	K	E	J	Q	C	V	U	D	A	Q	C	E
6	F	B	C	U	S	A	C	X	E	B	U	Q	C	U	B	R	U	X	E	P
7	A	P	Q	U	W	A	Q	C	J	A	F	E	J	D	U	F	J	T	A	F
8	W	A	V	T	B	T	O	Q	U	J	B	C	E	B	U	P	U	U	B	B
9	E	D	B	T	D	T	U	F	I	E	H	T	V	U	T	V	C	U	W	A
10	Q	J	J	E	V	U	J	S	C	A	F	B	T	U	C	U	J	C	A	X
11	E	T	F	U	J	N	Q	J	O	Q	E	Q	C	I	T	F	U	B	E	V
12	A	D	U	E	F	E	B	V	E	F	B	T	O	Q	U	U	F	D	A	F
13	O	Q	U	C	E	F	B	V	E	P	E	Q	V	U	W	Q	T	J	Q	B
14	T	V	T	J	E	J	U	J	V	U	P	T	A	F	J	W	A	Q	C	J
15	U	X	W	E	C	U	C	R	Q	W	A	Q	M	A	T	C				



## Fréquences

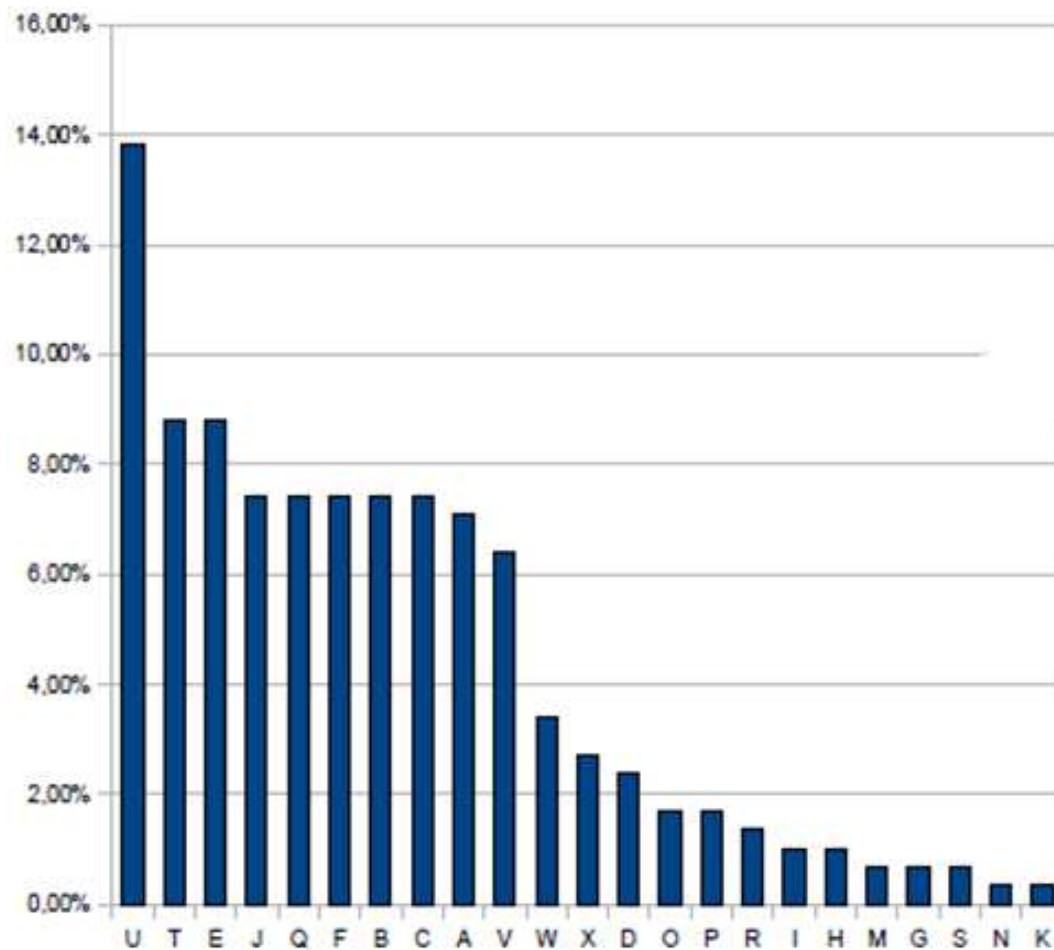
Fréquences des lettres											
U	T	E	J	Q	F	B	C	A	V	W	X
41	26	26	22	22	22	22	22	21	19	10	8
14,75%	9,35%	9,35%	7,91%	7,91%	7,91%	7,91%	7,91%	7,55%	6,83%	3,60%	2,88%

Fréquences des digrammes											
UJ	VU	CU	AQ	TF	UL	RU	ET	SE	UF	FJ	JE
21	13	10	7	7	6	6	5	5	5	5	5
16,41%	10,16%	7,81%	5,47%	5,47%	4,69%	4,69%	3,91%	3,91%	3,91%	3,91%	3,91%

Fréquences des trigrammes											
TFU	WAQ	EFB	OQU	AQC	FBT	TAF	OQE	UJB	CAX	ETF	FUB
4	4	4	3	3	3	3	2	2	2	2	2
8,00%	8,00%	8,00%	6,00%	6,00%	6,00%	6,00%	4,00%	4,00%	4,00%	4,00%	4,00%



Rapprochement avec  
les statistiques réelles

## ❑ Substitutions homophoniques

- A chaque caractère de l'alphabet A, on associe un ensemble de caractères de substitution appelés **homophones**.

<u>Lettre</u>	<u>Homophones</u>
E	17 19 84 41 56 60 67 83
I	08 22 53 65 88 90
C	03 44 76
N	02 09 15 27 32 40 59
O	01 11 23 28 42 54 70 80
D	33 91 45 58 64 78
F	05 10 20 29

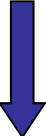
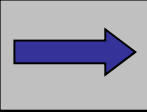
M: C O N F I D E N T I E L



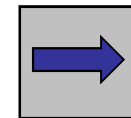
C: 76 11 59 05 65 91 17 27 03 41

- Substitutions homophoniques (second ordre)

- Le message M est associé à un faux message X pour générer le cryptogramme C grâce à une matrice d'homophones

	E	F	H	O	M
E	10	22	18	02	11
F	12	01	25	05	20
H	19	06	23	13	07
O	03	16	08	04	15
M	17	09	21	14	24



M: H O M M E

X: F E M M E



C: 06 03 24 24 10



# Chiffrement par décalage

- ❑ Le chiffrement par décalage se définit dans  $\mathbb{Z}_{26}$  car on utilise 26 lettres dans l'alphabet mais on pourrait le définir sur n'importe quel  $\mathbb{Z}_m$ .
- ❑ Il est facile de voir que le chiffrement par décalage forme un système cryptographique c'est-à-dire:

$$d_K(e_K(x)) = x \quad \forall x \in \mathbb{Z}_{26}$$

Soient  $M$ ,  $C$  et  $K \in \mathbb{Z}_{26}^*$ . Pour  $0 \leq K \leq 25$ , on définit

$$e_K(x) = (x + K) \bmod 26$$

et

$$d_K(y) = (y - K) \bmod 26$$

où  $x, y \in \mathbb{Z}_{26}$

## ❑ Algorithme

- On numérote de 0 à 25 les lettres de l'alphabet.
- On choisit une clé  $K$  comprise entre 1 et 25 et on chiffre le caractère  $X$  par:

$$e_K(x) = (x + K) \bmod 26$$

## ❑ Exemple :

- Chiffrement du message ZORRO avec  $K = 3$ 
  - $Z \rightarrow 25, \quad 25 + 3 = 28, \quad 28 \bmod 26 = 2 \rightarrow C$
  - $O \rightarrow 14, \quad 14 + 3 = 17, \quad 17 \bmod 26 = 17 \rightarrow R$
  - $R \rightarrow 17, \quad 17 + 3 = 20, \quad 20 \bmod 26 = 20 \rightarrow U$
  - ZORRO devient CRUUR

## Exemple: chiffre alphabétique par décalage

- ❑ Avec la clé  $K=3$  (chiffre de César), le texte ABC est chiffré par DEF
- ❑ La clé  $K=13$  donne le chiffre ROT13, encore parfois utilisé dans certains systèmes UNIX

## Chiffrement affine

- On choisit un alphabet à  $M \geq 2$  lettres, on associe à chaque lettre de l'alphabet un entier entre 0 et  $M - 1$ . Un code affine sur cet alphabet est un code dont la fonction de codage est :

$$E: \mathbb{Z}/M\mathbb{Z} \longrightarrow \mathbb{Z}/M\mathbb{Z}$$
$$x \mapsto ax + b$$

- On suppose qu'on utilise l'alphabet latin avec 26 lettres. On considère le code affine avec  $M = 26$ :

## Suite

□ Soient  $P, C \in \mathbb{Z}_{26}^*$  et soit

$$\mathbb{K} = \{(a, b) \in \mathbb{Z}_{26} \times \mathbb{Z}_{26} : \text{pgcd}(a, 26) = 1\}$$

Pour tout  $K = (a, b) \in \mathbb{K}$ , on définit

$$e_K(x) = (ax + b) \bmod 26$$

et

$$d_K(y) = a^{-1}(y - b) \bmod 26$$

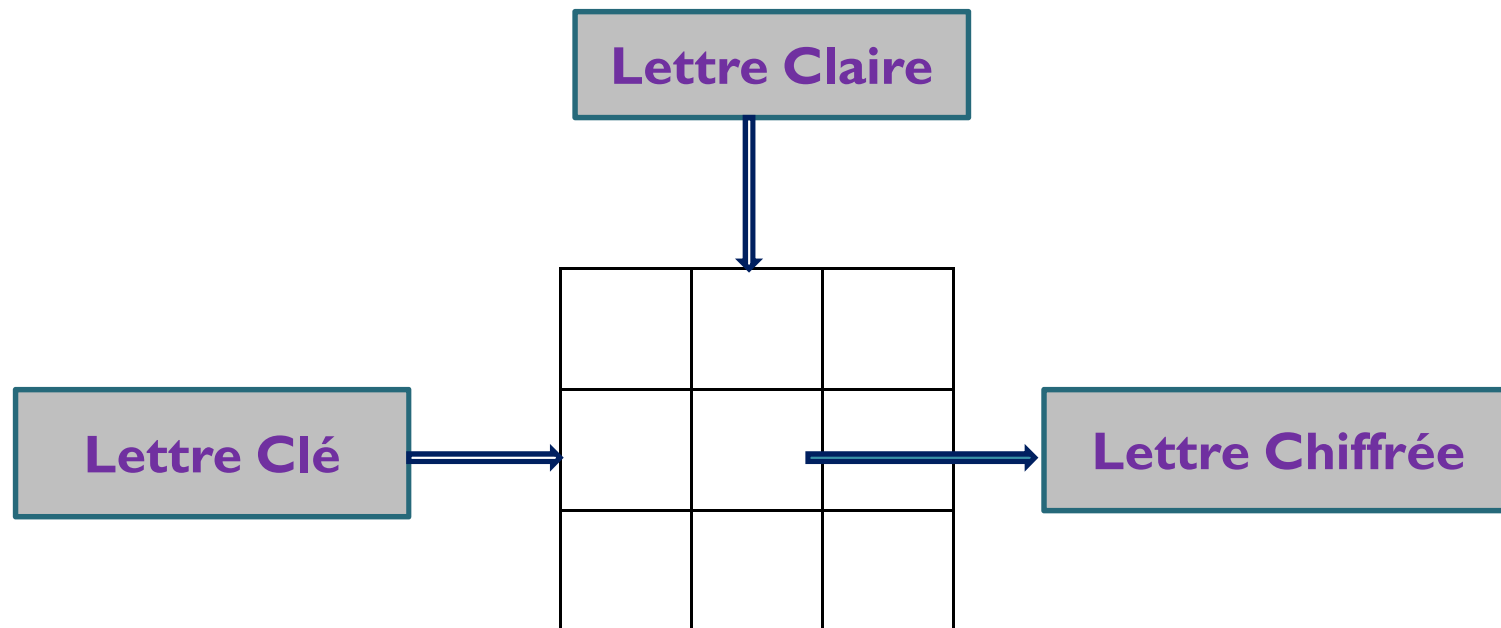
# Chiffrement par substitution

## ❑ Substitutions polyalphabétiques

- Chiffrement périodique à plusieurs alphabets de substitutions Vigenère [16 siècle]: table de 26 alphabets de substitution.
- Caractère de la clé K: nombre de décalages dans le  $i^{\text{ème}}$  alphabet.

M:	C O N F	I D E N	T I E L
	■	■	■
K:	P I L E	P I L E	P I L E
	↓	↓	↓
C:	R W Y J	X L P R	I Q P P

- ❑ S'appuie sur une clé
- ❑ Principe



## Carré de Vigenère

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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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



## ❑ Exemple

Texte clair	r	e	n	d	e	z	v	o	u	s
Clé	B	O	N	B	O	N	B	O	N	B
Texte chiffré	S	S	A	E	S	M	W	C	H	T

- Lettre 1 : la clé est B donc substitution dans le 2e alphabet (r->S)
- Lettre 2 : la clé est O donc substitution dans le 15e alphabet (e->S)
- etc.

# Le chiffre de Vigenère

- ❑ Le chiffre de Vigenère utilise des substitutions alphabétiques multiples par décalage.
  - On choisit un mot comme clé.
  - Le rang de chaque lettre de la clé définit un décalage à appliquer.
- ❑ Exemple: avec la clé **DECEPTION**, on chiffrera le texte clair **NOUSSOMMESDECOUVERTS**

# Le carré de Vigenère

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	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	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	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	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
...																									
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
...																									
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	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	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
...																									
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

Texte clair	N	O	U	S	S	O	M	M	E	S	D	E	C	O	U	V	E	R	T	S
Clé répétée	D	E	C	E	P	T	I	O	N	D	E	C	E	P	T	I	O	N	D	E

# Le carré de Vigenère

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	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	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	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	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
...																									
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
...																									
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	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	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
...																									
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

Texte clair	N	O	U	S	S	O	M	M	E	S	D	E	C	O	U	V	E	R	T	S
Clé répétée	D	E	C	E	P	T	I	O	N	D	E	C	E	P	T	I	O	N	D	E
Texte chiffré	Q	S	W	W	H	H	U	A	R	V	H	G	G	D	N	D	S	E	W	W

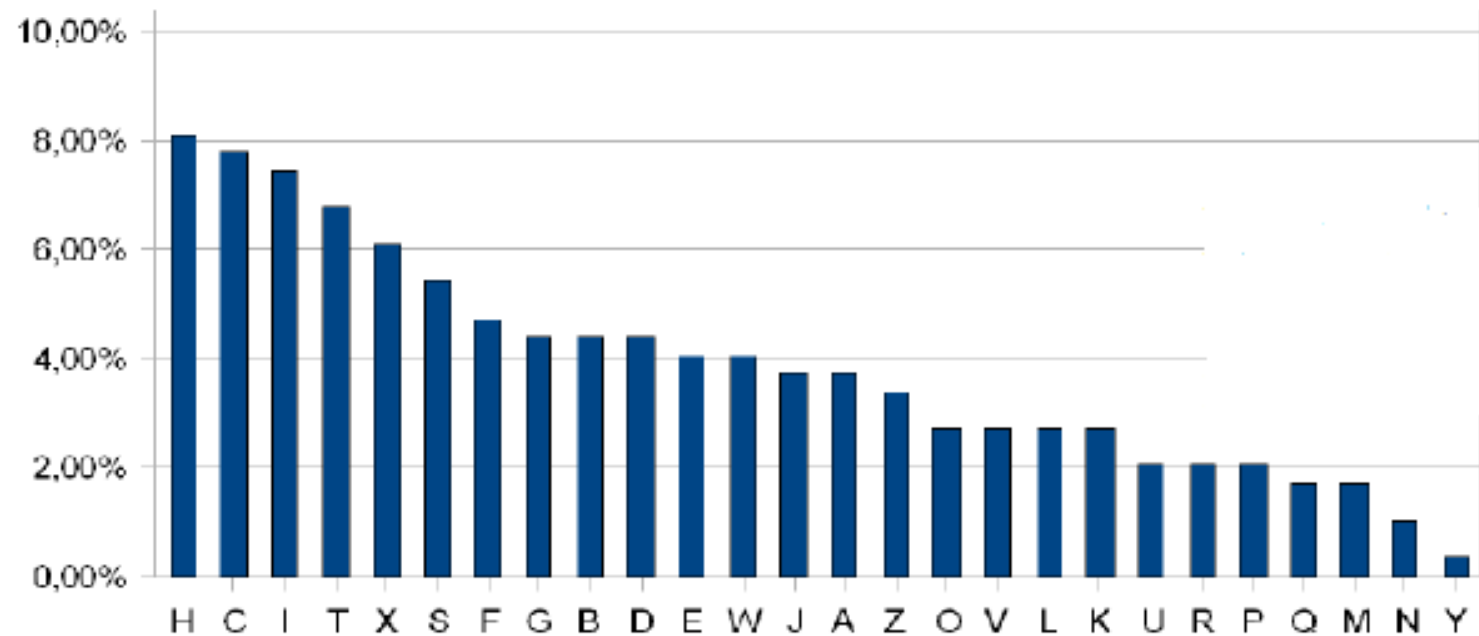
## Exemple de texte chiffré

1	H	C	E	W	S	H	H	Z	G	S	M	Q	V	I	H	X	C	E	G	S
2	A	A	R	K	E	J	O	C	X	A	D	B	U	X	F	D	A	R	B	B
3	T	H	C	A	W	H	H	F	B	F	T	I	E	B	J	T	F	J	X	Z
4	A	S	R	F	P	X	H	Z	X	I	M	S	K	U	F	X	Z	C	T	B
5	I	W	C	L	O	E	D	L	R	O	H	I	I	E	S	R	C	L	K	O
6	C	H	I	X	T	D	F	D	T	H	T	I	I	X	H	S	S	D	T	U
7	D	U	L	X	D	D	I	I	L	C	C	O	J	V	S	C	G	Z	H	B
8	E	C	C	B	H	X	E	L	X	G	I	F	R	M	S	V	S	V	M	H
9	P	Q	K	B	Q	X	S	E	A	O	Q	W	C	X	W	A	F	V	I	C
10	J	G	J	T	Z	T	G	W	K	C	C	H	Z	X	F	T	G	I	H	A
11	P	W	E	X	G	Y	I	J	J	I	P	I	I	A	W	C	S	K	T	Z
12	D	Q	V	T	B	P	H	C	T	B	I	W	H	N	S	T	B	T	H	B
13	F	I	V	K	O	C	H	C	T	U	P	I	C	X	D	J	W	J	N	H
14	X	Z	Z	L	O	H	S	J	E	S	V	W	F	G	G	E	C	L	K	G
15	T	A	G	T	F	T	F	U	N	D	D	I	M	H	W	G				

## Fréquences

Fréquence des lettres

H	C	I	T	X	S	F	G	B	D	E	W	J	A	Z	O	V	L	K	U	R	P	Q	M	N	Y
24	23	22	20	18	16	14	13	13	13	12	12	11	11	10	8	8	8	8	6	6	6	5	5	3	1



# Cryptanalyse des chiffres par substitution polyalphabétique

- ❑ Principe général : trouver la longueur de la clé
  - Test de Kasiski (1863)/Babbage(1854)
    - Recherche de répétitions de lettres (causées par l'emploi cyclique de la même clé)
    - L'écart entre les différentes occurrences donne des indices
    - Les diviseurs communs des différents écarts permettent de trouver la longueur de la clé



- ❑ Exemple : on localise les répétitions de BIW, OCH et DDI

1	H	C	E	W	S	H	H	Z	G	S	M	Q	V	I	H	X	C	E	G	S
2	A	A	R	K	E	J	O	C	X	A	D	B	U	X	F	D	A	R	B	B
3	T	H	C	A	W	H	H	F	B	F	T	I	E	B	J	T	F	J	X	Z
4	A	S	R	F	P	X	H	Z	X	I	M	S	K	U	F	X	Z	C	T	B
5	I	W	C	L	O	E	D	L	R	O	H	I	I	E	S	R	C	L	K	O
6	C	H	I	X	T	D	F	D	T	H	T	I	I	X	H	S	S	D	T	U
7	D	U	L	X	D	D	I	I	L	C	C	O	J	V	S	C	G	Z	H	B
8	E	C	C	B	H	X	E	L	X	G	I	F	R	M	S	V	S	V	M	H
9	P	Q	K	B	Q	X	S	E	A	O	Q	W	C	X	W	A	F	V	I	C
10	J	G	J	T	Z	T	G	W	K	C	C	H	Z	X	F	T	G	I	H	A
11	P	W	E	X	G	Y	I	J	J	I	P	I	I	A	W	C	S	K	T	Z
12	D	Q	V	T	B	P	H	C	T	B	I	W	H	N	S	T	B	T	H	B
13	F	I	V	K	O	C	H	C	T	U	P	I	C	X	D	J	W	J	N	H
14	X	Z	Z	L	O	H	S	J	E	S	V	W	F	G	G	E	C	L	K	G
15	T	A	G	T	F	T	F	U	N	D	D	I	M	H	W	G				



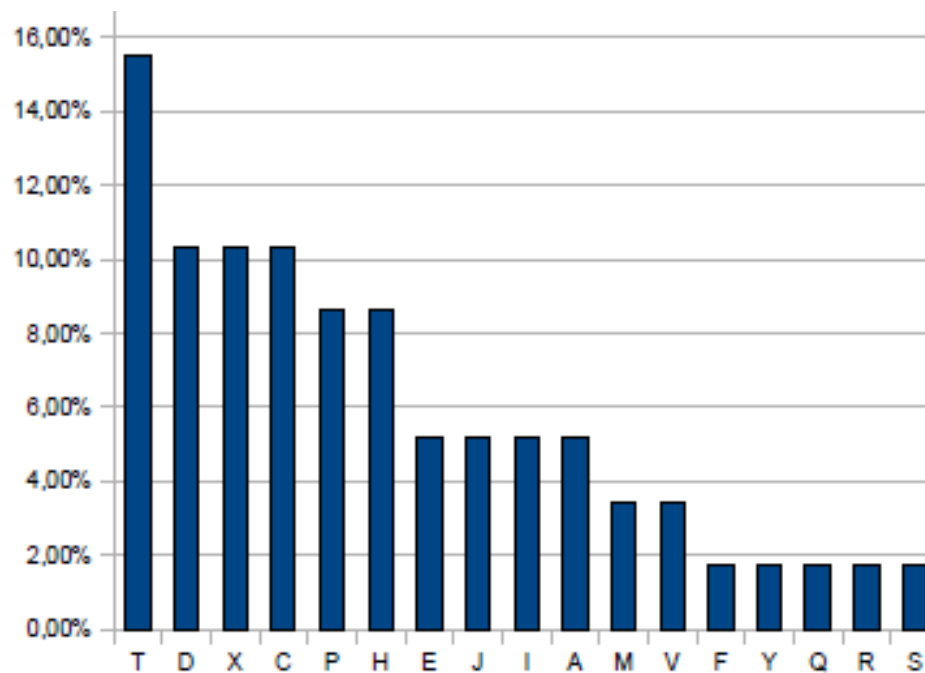
- L'analyse de ces répétitions donne :

	<i>Position 1</i>	<i>Position 2</i>	<i>Écart</i>	<i>Décomposition</i>			
<i>OCH</i>	99	244	145			5	29
<i>BIW</i>	79	229	150	2	3	5	
<i>DDI</i>	124	289	165		3	5	

- 5 étant le seul diviseur commun, on peut conjecturer que c'est la longueur de la clé

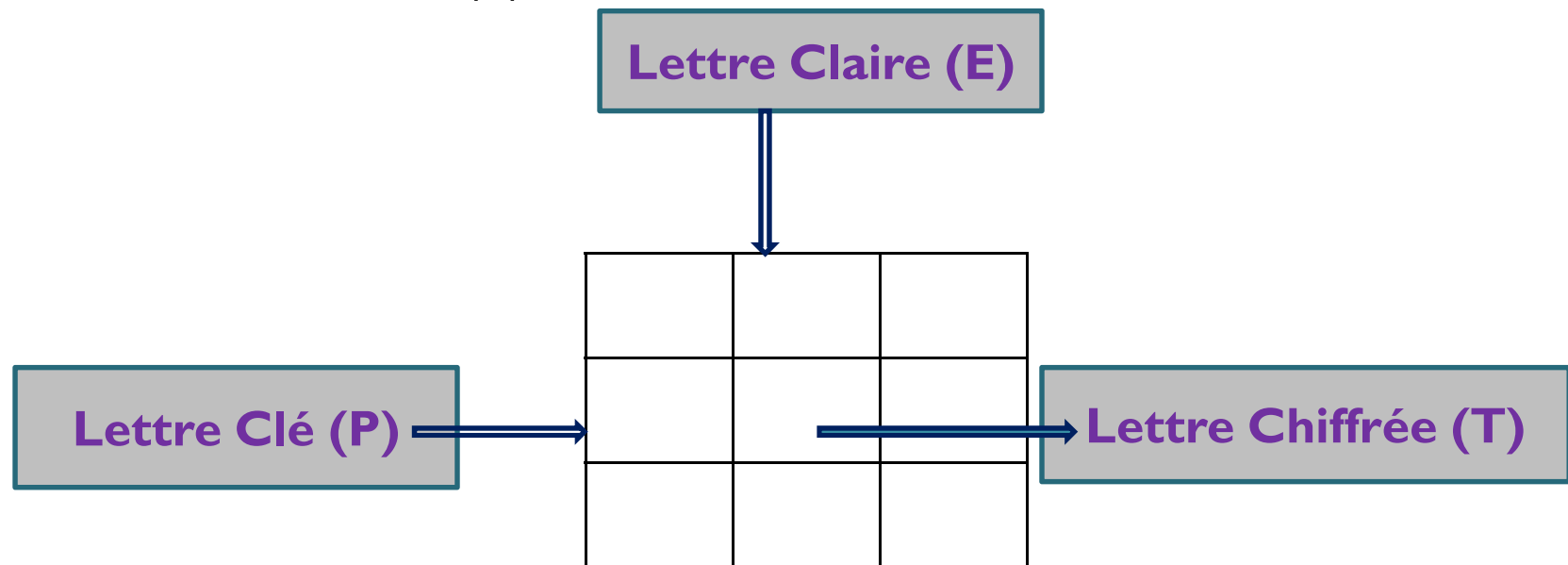
❑ Etude du premier alphabet (1 lettre sur 5 en partant de la première) :

H	H	M	X	A	J	D	D	T	H	T	T	A	X	M	X	I	E	H	R
C	D	T	S	D	D	C	C	E	X	I	V	P	X	Q	A	J	T	C	T
P	Y	P	C	D	P	I	T	F	C	P	J	X	H	V	E	T	T		



L'histogramme des fréquences est proche de celui d'une substitution monoalphabétique

- ❑ On connaît la lettre chiffrée (T) et l'on a une bonne hypothèse pour la lettre claire (E), d'où l'on peut déduire la lettre de la clé (P)



- ❑ En appliquant la même analyse sur les 4 autres alphabets, on retrouve toute la clé (PORTO)



## ❑ Substitutions à clé automatique:

- Si la clé est aussi longue que le message, l'algorithme est théoriquement incassable.
- Clé: séquence aléatoire de caractères, sans répétition;
- Chiffres "On-Time Pad" ou "Clé aléatoire une fois "
- ~~Caractères: 5 positions (marque: 1 et espace: 0)~~
- Chaque bit de M est additionné, modulo 2, au bit de la clé

## ❑ Substitutions à clé automatique:

- $M = m_1 m_2 m_3 \dots\dots K = k_1 k_2 k_3 \dots\dots$

- $C = E_k(M) = c_1 c_2 c_3 \dots\dots$

- Chiffrement:  $c_i = (m_i + k_i) \bmod 2$


- Déchiffrement:  $(c_i + k_i) \bmod 2 = (m_i + k_i + k_i) \bmod 2 = m_i$

❑ **Exemple:** l'addition du caractère "A" (codé 11000) de M et du caractère "D" (codé 10010) de la clé K.

$$M = 1 \ 1 \ 0 \ 0 \ 0$$

$$K = 1 \ 0 \ 0 \ 1 \ 0$$

$$E_k(M) = 0 \ 1 \ 0 \ 1 \ 0$$

- 
- ❑ Substitutions polygrammiques: Chiffre de Playfair[1854]
    - Utilisé par les anglais durant la première guerre mondiale
    - Clé: Matrice carré de 25 lettres
    - Chiffre de Playfair: Substitution de bigrammes
    - Chiffrement de chaque paire de caractères ( $m_1m_2$ ) de  $M$  selon 3 règles basées sur la position de  $m_1$  et  $m_2$  dans la matrice.

# Exemple: Chiffre de Playfair

m1	c1	m2	c2

**Règle 1**

m1			c1
c2			m2

**Règle 2**

	m1	
	c1	
	m2	
	c2	

**Règle 3**

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

M: CO NF ID EN TI EL

↓ ↓ ↓ ↓ ↓ ↓

C: BV LH KB RK OM SK

# Chiffre de Hill

## ■ Substitutions polygrammiques: Chiffre de Hill

- Transformation linéaire sur  $d$  caractères de  $M$  pour générer  $d$  caractères de  $C$ .
- Si  $d = 2$ , le bloc  $(m_1 \ m_2)$  de  $M$  est chiffré par le bloc  $(c_1 c_2)$  de  $C = E_k(M)$  par un système d'équations de dimension  $d$ :

$$c_1 = (k_{11}m_1 + k_{12}m_2) \bmod n$$

$$c_2 = (k_{21}m_1 + k_{22}m_2) \bmod n$$



□ Chiffrement:

$$\mathbf{C} = \mathbf{E}_k(\mathbf{M}) = \mathbf{K} \cdot \mathbf{M} \bmod n$$

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \mathbf{K} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} \quad \text{avec } \mathbf{K} = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix}$$

□ Déchiffrement:

$$\mathbf{D}_k(\mathbf{M}) = \mathbf{K}^{-1} \mathbf{C} \bmod n = \mathbf{K}^{-1} \mathbf{K} \bmod n = \mathbf{M} \bmod n$$

■ Chiffrement du bigramme EG = [4,6] par le bigramme YQ = [24,16]

$$\begin{pmatrix} 3 & 2 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \end{pmatrix} \bmod 26 = \begin{pmatrix} 24 \\ 16 \end{pmatrix} = \text{YQ} \quad K = \begin{pmatrix} 3 & 2 \\ 3 & 5 \end{pmatrix}$$

$$\begin{pmatrix} 15 & 20 \\ 17 & 9 \end{pmatrix} \begin{pmatrix} 24 \\ 16 \end{pmatrix} \bmod 26 = \begin{pmatrix} 4 \\ 6 \end{pmatrix} = \text{EG} \quad K^{-1} = \begin{pmatrix} 15 & 20 \\ 17 & 9 \end{pmatrix}$$