Cryptographie et Cryptanalyse

Un peu d'histoire avant de commencer

De la stéganographie à la cryptographie

- Stéganographie : Étymologie grecque
  - *Steganos* : Couvert
  - *Graphein* : Ecriture

- Vers 600 avant J.C. : Nabuchodonosor utilisait des crânes rasés

- Vers 480 avant J.C. : Dematarus (Grec) isolé en Perse utilise des tablettes de cire pour prévenir son pays des projets de Xerxes (Perse)

- 1ier siècle : Apparition de l'encre invisible (fluides organiques riches en carbone)

  Si le message est intercepté...

De la stéganographie à la cryptographie

- Cryptographie : Étymologie grecque
  - *Kryptos* : Caché
  - *Graphein* : Ecriture

- 5 siècle avant J.C., Spartiates : Utilisation de scytales

- Chiffre de César : Décalage alphabétique
- 1580 : Chiffre par substitution de symboles - Marie Stuart
- 1586 : *Traité des chiffres* – Chiffre de Vigenère
- 1930 : Enigma
- Fin 20e : Apparition du chiffrement à clé publique - RSA

La cryptographie

- Étude des algorithmes et des protocoles utilisés pour préserver la confidentialité de l'information et garantir son intégrité.
- Elle constitue la base de la plupart des techniques de sécurité:
  - échanges sécurisés sur Internet,
  - confidentialité des transactions bancaires,
  - protection des secrets industriels et commerciaux,
  - protection du secret médical,
  - protection des systèmes informatiques contre les intrusions,
  - Par le biais de la signature électronique:
    - identification des correspondants,
    - non répudiation des transactions,
    - Garantie d'intégrité des documents.

Cryptanalyse

but de la cryptanalyse:
trouver des attaques permettant de *casser* les protections mises en place à l'aide de la cryptographie.

Codes et Chiffres

Les Codes et leurs Limites

- Une table à double entrée, *le code secret*, constituait la clé. Elle donnait la correspondance entre les unités du texte clair et les unités correspondantes du message secret.

Exemple.

Le sous marin est attendu à	Jean
10 heures	est là
20 heures	n'est pas là

- On voit que la taille de la clé était nécessairement supérieure ou égale à celle du message. La confidentialité de l'information était fondée sur le fait que seules les personnes autorisées connaissaient la clé.
- Le message secret pouvait masquer totalement la structure du texte clair.

- Dans l'exemple précédent, un attaquant ne disposant pas du code, ne peut pas déterminer si le message secret annonce l'arrivée d'un sous-marin ou le décollage d'un avion.

- Un code secret peut donc apporter une excellente garantie de confidentialité (tant que le code utilisé reste ignoré des attaquants !).
- le type d'information que l'on peut ainsi échanger est très limité :
  - une information qui n'a pas d'entrée dans la table ne peut pas être codée.

Quelques Principes du Chiffrement

Codes modernes

- A la différence des codes, les algorithmes cryptographiques actuels, appelés chiffres, fonctionnent au niveau des bits ou des groupes de bits, permettant ainsi la transmission de toute information : texte, fichier binaire, image,...

- Un algorithme cryptographique comprend deux fonctions, une fonction E (Encryption) pour chiffrer le message et sa fonction inverse D (Decryption) pour le déchiffrer.

Fonctions de bases

- fonction E (Encryption):

$$C=E(K1,M)$$

chiffre le message M avec la clé de chiffrement K1

- fonction D (Decryption):

$$M=D(K2,C)$$

déchiffre C avec la clé de déchiffrement K2

- 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.
  - Toute standardisation des logiciels de communication est impossible.

➔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 K2.

Chiffres par décalage

Exemple: chiffre alphabétique par décalage.

- **P**ermutation circulaire sur l'ensemble des lettres de l'alphabet

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
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

- 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

Exemple: chiffre alphabétique par décalage.

- 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:

$$E(K,X) = (X+K) \bmod 26$$

- Chiffrement du message ZORRO avec K = 3
  - Z = 25, 25 + 3 = 28, 28 mod 26 = 2 : C
  - O = 14; 14 + 3 = 17, 17 mod 26 = 17 : R
  - R = 17, 17 + 3 = 20, 20 mod 26 = 20 : U
- ZORRO devient CRUUR

Petite pause exercice...

- **Exercice.** Ecrire la fonction D de déchiffrement correspondant au chiffre par décalage. Quelle est la clé de déchiffrement?

Vraiment petite...

- **Exercice.** Ecrire la fonction D de déchiffrement correspondant au chiffre par décalage. Quelle est la clé de déchiffrement?
- Fonction= la même
- Clef= 26-k
- -k mod 26 = 26 -k mod 26

De manière générale,
les chiffres par substitutions...

Chiffres alphabétiques par substitution

- Un chiffre alphabétique par substitution généralise le chiffre par décalage.
- Le principe consiste à choisir comme clé une permutation arbitraire P des 26 lettres de l'alphabet et à chiffrer le caractère X par:

$$E(P,X)=P(X)$$

Comment construire une permutation sans trop de difficultés ?

- Etablir une clé qui peut être un mot ou une phrase
  - Exemple : prenons le mot « securite »
- Nettoyer la clé i.e. supprimer les lettres doubles
  - « securite » devient « securit »
- Faire correspondre les premières lettres de l'alphabet aux lettres de la clé et complèter

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

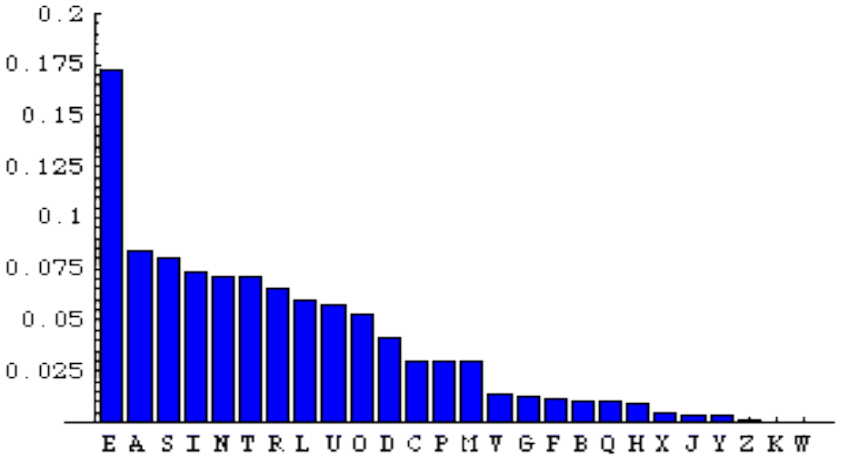
Les limites d'un Chiffre alphabétique par substitution

- Un chiffre alphabétique par substitution généralise le chiffre par décalage.
- Le principe consiste à choisir comme clé une permutation arbitraire P des 26 lettres de l'alphabet et à chiffrer le caractère X par:

$$E(P,X)=P(X)$$

➔ Les chiffres par substitution ne permettent pas de garantir la confidentialité de l'information.

- Voici un diagramme des fréquences des lettres en français:



- Le texte chiffré laisse apparaître la structure du texte clair.
  - Toutes les occurrences d'une même lettre ou d'un même groupe de lettres sont toujours chiffrées de la même manière.

➔ Une analyse statistique simple permet alors de reconstituer le texte clair.

Comment avoir un chiffrement résistant à l'analyse des fréquences ?

- Idée: essayer de masquer autant que possible la structure du texte clair (répétitions de caractères ou de petits groupes de caractères)

- Une approche possible consiste à appliquer un décalage fonction de la position du caractère dans le texte.

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

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

Version « *light* » du 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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

N O U S S O MMESDECOUVERTS
13 14 20 18 18 14...
D E C E P T IONDECEPTIONDECE
3 4 2 4 15 19....
16 18 22 22 7 7 ....
Q S W W H H ...

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

N O U S S O MMESDECOUVERTS
13 14 20 18 18 14...
D E C E P T IONDECEPTIONDECE
3 4 2 4 15 19....
16 18 22 22 7 7 ....
Q S W W H H ...

Des chiffrements de plus en plus pointus...

Chiffre de Hill

- Le chiffre de Hill substitue m caractères X1, ..., Xm du texte clair par m caractères C1,..,Cm, selon le système d'équations linéaires suivant:

$$C_1 = K_{11} X_1 + ... + K_{1m} X_m \bmod 26$$

$$C_2 = K_{21} X_1 + ... + K_{2m} X_m \bmod 26$$

.....

$$C_m = K_{m1} X_1 + ... + K_{mm} X_m \bmod 26$$

ou encore, sous forme matricielle:

$$C = KX$$

- La matrice K constitue la clé de chiffrement.
- Elle doit évidemment être inversible pour que le déchiffrement soit possible.
- L'algorithme de déchiffrement est alors le même que l'algorithme de chiffrement, en remplaçant simplement K par K^{-1}

On voit que, si p<m , alors deux occurrences d'une même chaîne X1...Xp de p caractères ne seront en général pas chiffrées de la même manière

Exemple. Avec la clé

$$K = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix}$$

le texte PAYMOREMONEY est chiffré par LNSHDLEWMTRW.

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Exemple. Avec la clé

$$K = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix}$$

le texte:

P	A	Y	M	O	R	E	M	O	N	E	Y
15	0	24	12	14	17	4	12	14	13	4	24

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Exemple. Avec la clé

$$K = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix}$$

le texte:

P	A	Y	M	O	R	E	M	O	N	E	Y
15	0	24	12	14	17	4	12	14	13	4	24

$$\begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix} * \begin{pmatrix} 15 \\ 0 \\ 24 \end{pmatrix}$$

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Exemple. Avec la clé

$$K = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix}$$

le texte:

P A Y M O R E M O N E Y

15 0 24 12 14 17 4 12 14 13 4 24

est chiffré par:

11 13 18 7 3 11 4 22 12 19 17 22

L N S H D L E W M T R W.

$$\begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix} * \begin{pmatrix} 15 \\ 0 \\ 24 \end{pmatrix}$$

17*15+5*24=255+120=375
375 mod 26 = 11

21*15+21*24=315+504=819
819 mod 26 = 13

Chiffre XOR

- Le chiffre XOR manipule les bits et non plus les caractères.
- Il est basé sur l'opération *ou exclusif bit à bit* (XOR, noté encore $\oplus$).
- On utilise comme clé une chaîne de bits K de longueur donnée L.
- On chiffre en effectuant le ou exclusif bit à bit de la clé K avec le texte clair, découpé en blocs M de longueur L.

Chiffre XOR

- **Exemple.** Avec la clé K=0101 le texte clair 11110000 sera chiffré ainsi:

11110000 $\oplus$ texte clair

01010101 $\oplus$ clef

10100101 $\oplus$ texte codé

- Rappelons que le ou exclusif est associatif, commutatif, qu'il possède un élément neutre 0, et que toute chaîne K est son propre inverse: $K \oplus K = 0$
- Ainsi, on voit que l'algorithme de déchiffrement est identique au chiffrement, avec la même clé :

$$M = (M \oplus K) \oplus K = C \oplus K$$