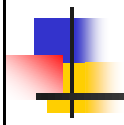
Réseaux Avancés
L3 Informatique



Cours 2 et 3: Cryptographie

Osman SALEM
Maître de conférences - HDR
osman.salem@parisdescartes.fr

UNIVERSITÉ
PARIS DESCARTES
UFR DE MATHÉMATIQUES ET INFORMATIQUE

2. Cryptographie

Introduction

La cryptographie

- Elle est caractérisée par
  - Le type des opérations de chiffrement utilisées
    - substitution / transposition / produit
  - Le nombre des clés utilisées
    - Clé unique partagée / paire de clés ou clés publiques
  - La façon dont le message en clair est traité
    - Par bloc / en flux

La sécurité apportée par la cryptographie

- Sécurité inconditionnelle
  - Quelles que soient les ressources de calcul disponibles, le message chiffré ne peut pas être cassé car il contient des informations insuffisantes pour reconstruire de manière unique le message en clair correspondant
- Sécurité ″liée au coût des calculs″
  - Avec des ressources de calcul limitées (par exemple, le temps des calculs est égal à l'âge de l'univers), le message chiffré ne peut pas être cassé

Chiffrement symétrique

- Emetteur et récepteur ***partagent*** une même clé secrète
- Tous les algorithmes de chiffrement classiques appartiennent à cette famille
- C'était l'unique méthode de chiffrement avant les années 1970 et l'invention des systèmes à clés publiques

Cryptographie: définition

- C'est la science qui étudie les principes, méthodes et techniques mathématiques pour réaliser la sécurité de l'information
- Pour une utilisation sûre, il faut
  - Un algorithme de chiffrement solide "**E**" (Encryption ou chiffrement)
  - Une clé secrète "**K**" connue seulement de l'émetteur et du récepteur

    $C = E_K(M)$

    $M = D_K(C) = D_K(E_K(M))$

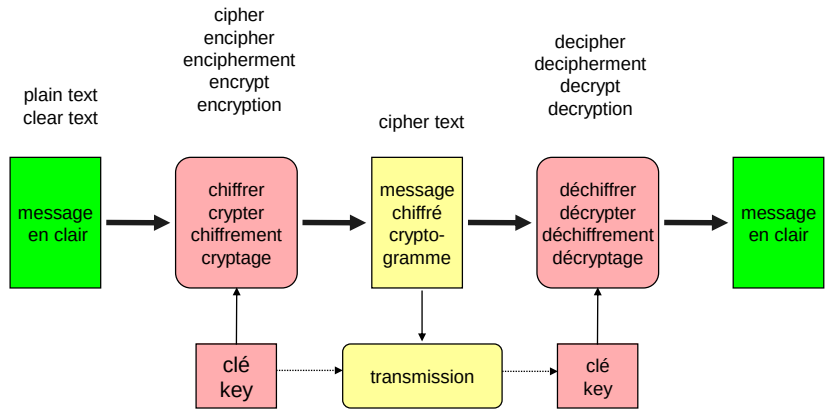
    Ex: message en clair: *Bob, je t'aime. Alice*

    message chiffré: *CPC, KF U'BJNF. BMJDF*

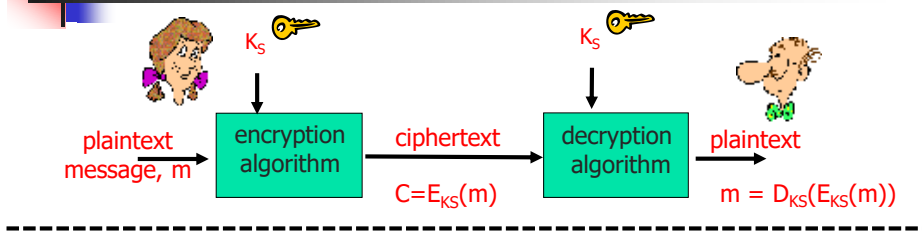
Cryptographie : synonymes

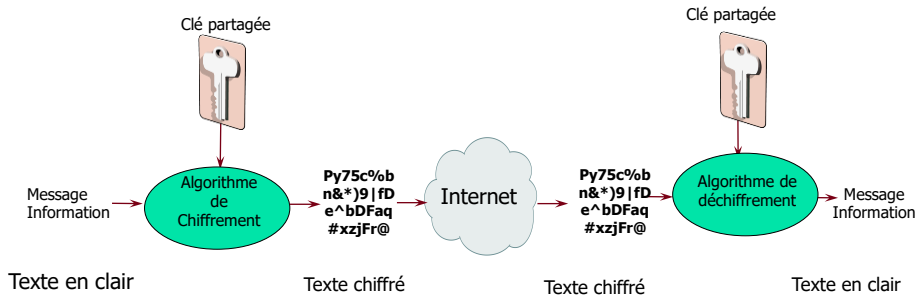
- Message en clair = message original = *plaintext*
- Message chiffré = ciphertext – *message original*
- (algorithme de) Chiffrement = cipher - algorithme de transformation du message en clair en message chiffré
- Clé = key – information utilisée dans l'algorithme de chiffrement et connue des seuls émetteur et récepteur
- Chiffrer, crypter = encipher, encrypt - convertir le message en clair en message chiffré
- Déchiffrer, décrypter = decipher, decrypt – restaurer le message en clair à partir du message chiffré
- Cryptographie = étude des principes et méthodes de chiffrement
- Cryptanalyse = ***codebreaking*** - étude des principes et méthodes de déchiffrement sans connaissance des clés
- Cryptologie = domaine complet comprenant la cryptographie et la cryptanalyse

Cryptographie: terminologie



Chiffrement symétrique



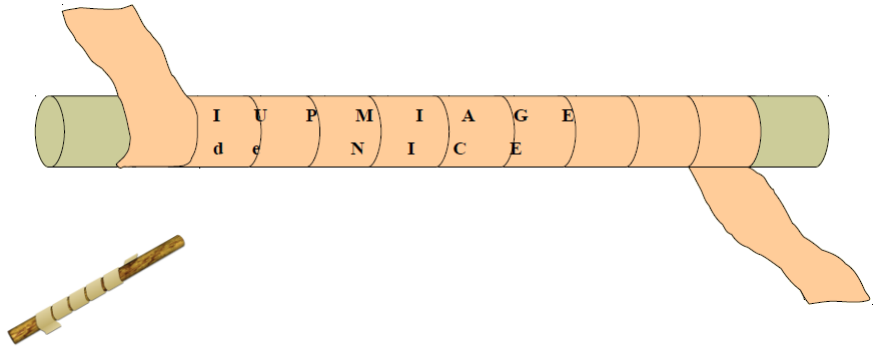


Cryptanalyse

- La sécurité d'un système de chiffrement doit reposer sur
  - le secret de la clé de chiffrement
  - et non pas sur celui de l'algorithme
- Le ***principe de Kerkhoff*** suppose en effet que l'attaquant connaît l'algorithme utilisé
  - L'algorithme est connu
  - La clé est **secrète**
- Il faut donc une solution sûre pour transmettre la clé de l'émetteur au récepteur => rôle des protocoles d'échange de clés

Chiffrement symétrique

- Grèce antique: la ***scytale*** utilisée à Sparte (**Spartiates**)
  - Algorithme : texte écrit sur un ruban enroulé autour d'un bâton
  - Clé: diamètre du bâton

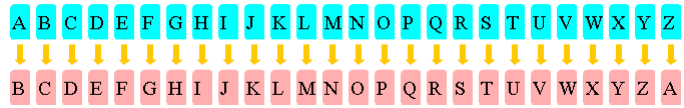


Chiffrement symétrique

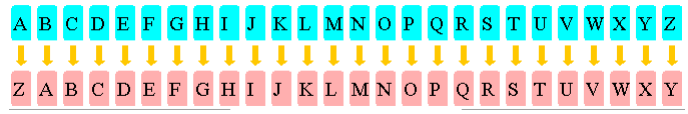
- Les algorithmes de chiffrement symétrique se fondent sur une clé unique pour chiffrer et déchiffrer un message
  - Clé unique partagée
- **Basée sur 2 mécanismes**
  - Substitution
  - Permutation

Substitution

- Substitution: Ex1: chiffrement par décalage



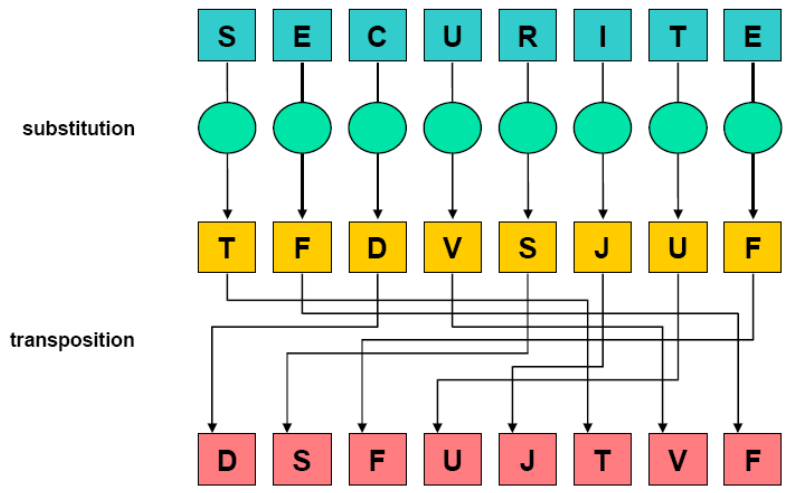
- Substitution alphabétique inversée



- Substitution, Ex3: utilisation de tables d'association

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

Substitution et transposition

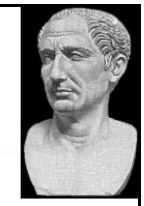


Méthode classique de substitution

- Les lettres du message en clair sont remplacées par d'autres lettres, des chiffres ou d'autres symboles
- Si le message en clair est vu comme une suite de données binaires, ce sont des séquences de bits qui seront remplacées par d'autres séquences de bits

input	output	input	output
000	110	100	011
001	111	101	010
010	101	110	000
011	100	111	001

Le code de César



- Plus ancien code de substitution connu
- Inventé par Jules César
- Utilisé pour des affaires militaires
  - Chaque lettre est remplacée par celle qui est située trois (par exemple) lettres après dans l'alphabet
- Exemple

Rendezvousapreslecours

UHQGHCYRXVDSUHVOHFRXUV

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

http://www.bibmath.net/crypto/substi/cryptcesar.php3

Mécanisme du code de César

- Une transformation

```
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
```

- En associant une valeur numérique à chaque lettre

```
a  b  c  d  e  f  g  h  i  j  k   l   m
0  1  2  3  4  5  6  7  8  9  10  11  12
n  o  p  q  r  s  t  u  v  w  x  y  Z
13 14 15 16 17 18 19 20 21 22 23 24 25
```

- $C = E_3(m) = (m + 3) \bmod (26)$

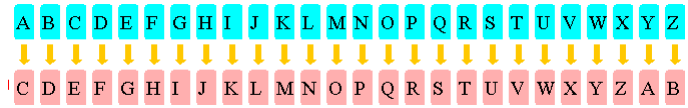
  $m = D_3(C) = (C - 3) \bmod (26)$

Cryptanalyse du code de César

- 26 possibilités de codage
  - ″A″ est transformé en ″A″, ″B″, ... ou ″Z″
- Ce code est facilement cassé par attaque brute
  - Il suffit d'essayer les 26 solutions !
- Pour cela, il faut pouvoir reconnaître le message en clair
- Essayez "HCEKNG FG FGEJKHHTGT"

Cryptanalyse du code de César

- 26 possibilités de codage
  - "A" est transformé en "A", "B", ...ou "Z"
- Ce code est facilement cassé par attaque brute
  - Il suffit d'essayer les 26 solutions !
- Pour cela, il faut pouvoir reconnaître le message en clair
- Essayez "HCEKNG FG FGEJKHHTGT"
  - A ➔ C
  - FACILE DE DECHIFFRER



http://www.bibmath.net/crypto/substi/cryptcesar.php3

Codage monoalphabétique

- Meilleur qu'un simple décalage des lettres
- Les lettres sont codées "au hasard"
- Cela revient à un code dont la clé serait longue de 26 lettres
- Exemple

```
clair:    abcdefghijklmnopqrstuvwxyz
chiffré: DKVQFIBJWPESCXHTMYAUOLRGZN
Message en clair:   etsinousreplacionsleslettres
Message chiffré:    FUAWXHOAYFTSDWHXASFYSFUUYFA
```

Sécurité apportée par un code monoalphabétique

- Nous avons $26! \geq 4 \times 10^{26}$ clés
- Avec un tel nombre de clés, on pourrait croire que le système est sûr
- Il y a malheureusement un problème : les langues naturelles ont des caractéristiques bien connues !

Cryptanalyse

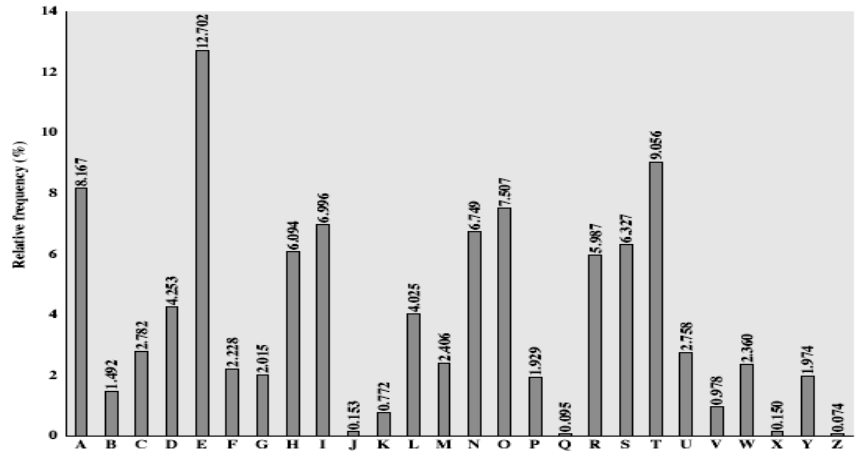
- Les méthodes de chiffrement par substitution ne changent pas la fréquence des lettres
- Découverte des scientifiques arabes au IX^{e} siècle
  - Al-Kindi
- Il suffit de calculer la fréquence des lettres dans le message chiffré et de comparer avec les fréquences connues
- Idée: examiner la fréquence des lettres d'un message chiffré.

Statistiques de la langue et Cryptanalyse

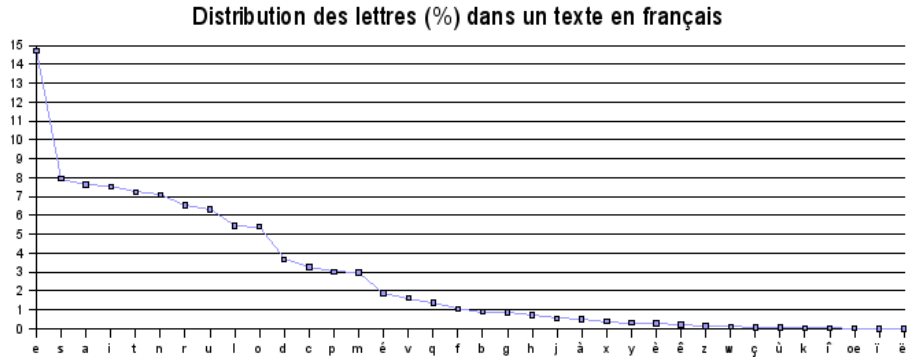
- Les lettres ont des fréquences d'apparition différentes :
  - En anglais, "E" est la lettre la plus fréquente, suivie par "T", "A", "O", "I", "N", "S", "H", "R"
  - Certaines lettres "Z", "J", "K", "Q", "X" sont rares
  - Il existe des tables de fréquences des digrammes (deux lettres consécutives), trigrammes (trois lettres consécutives),…

  - En **français**, "**E**" est la lettre la plus fréquente, suivie par "S", "A", "I", "T", "N", "R" , "U", "L", "O",
  - Certaines lettres "J", "X", "Y", "W", "K" sont rares
  - Il existe des tables de fréquences des digrammes (deux lettres consécutives), trigrammes (trois lettres consécutives),…

Fréquence des lettres en anglais



Fréquence des lettres en français



Source:
http://fr.wikipedia.org/wiki/Fr%C3%A9quence_d%27apparition_des_lettres_en_fran%C3%A7ais

Cryptanalyse

- Les méthodes de chiffrement par substitution ne changent pas la fréquence des lettres
- Il suffit de calculer la fréquence des lettres (des digrammes, des trigrammes) dans le message chiffré et de comparer avec les fréquences connues
- Pour le code de César (en français)
  - E, S, A les lettres les plus fréquentes, ES, EN, LE et DE sont les digrammes les plus fréquents, ENT, LES et ION les trigrammes les plus fréquents
  - Les raretés : y w k

Chiffrement "playfair"

- La substitution monoalphabétique ne suffit pas à apporter une sécurité suffisante
- Une approche "polyalphabétique" permet d'améliorer la sécurité
- Exemple : playfair

Matrice de lettres

- Une matrice 5X5 basée sur un mot-clé
- On place (sans espaces et sans duplication) les lettres du mot-clé
- On remplit le reste de la matrice avec les lettres restantes
- Par exemple avec le mot MONARCHIE

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

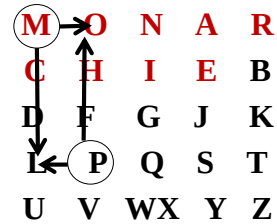
http://www.apprendre-en-ligne.net/crypto/subst/playfair.html

Chiffrement

- Le message en clair est chiffré en prenant les lettres deux par deux
  - Message en clair = **loup blanc**
  - Chaque groupe de 2 lettres est codé par la lettre à l'intersection de la ligne de la première et la colonne de la seconde puis à l'intersection de la ligne de la seconde et de la colonne de la première
  - Si les deux lettres tombent sur la même ligne, on remplace chacune par celle de droite ; Si les deux lettres tombent sur la même colonne, on remplace chacune par celle de dessous (avec rotation circulaire)
  - En cas de lettre double, et en cas de lettre unique (nombre total de lettres impair), une lettre "parasite" est insérée (w)
  - Exemple : LO UP BL AN CW
  - Chiffré en : PM VL CT RA IU

Déchiffrement

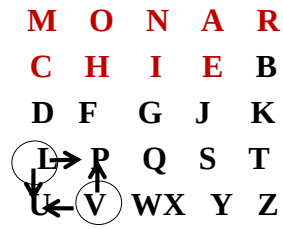
- Le message chiffré est lu en prenant les lettres deux
- par deux
  - Message chiffré = PM VL CT RA IU



- PM➔ ligne de P et colonne de M : intersection = L ; colonne de P et ligne de M : intersection = 0

Déchiffrement

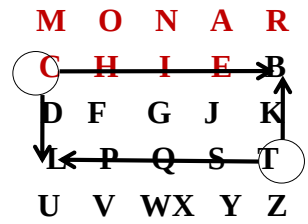
- Le message chiffré est lu en prenant les lettres deux
- par deux
  - Message chiffré = PM VL CT RA IU



- PM➔ ligne de P et colonne de M : intersection = L ; colonne de P et ligne de M : intersection = O
- VL ➔ ligne de V et colonne de L : intersection = U; colonne de V et ligne de L : intersection = P

Déchiffrement

- Message chiffré = PM VL CT RA IU



- PM➔ ligne de P et colonne de M : intersection = L ; colonne de P et ligne de M : intersection = 0
- VL ➔ ligne de V et colonne de L : intersection = U; colonne de V et ligne de L : intersection = P
- CT ➔ ligne de C et colonne de T : intersection = B ; colonne de c et ligne de T: intersection = L

Déchiffrement

- Message chiffré = PM VL CT RA IU

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

- PM➔ ligne de P et colonne de M : intersection = L ; colonne de P et ligne de M : intersection = 0
- VL ➔ ligne de V et colonne de L : intersection = U; colonne de V et ligne de L : intersection = P
- CT ➔ ligne de C et colonne de T : intersection = B ; colonne de c et ligne de T: intersection = L
- RA ➔ sont sur la même ligne donc décalage vers la gauche ==AN

Puis IU donne C et W ou X (lettre parasite)

Sécurité apportée par ce code

- Nettement meilleure qu'un code monoalphabétique
- Possède 26x26 combinaisons au lieu de 26 avec un code monoalphabétique
- Nécessite la connaissance de plusieurs échantillons de messages chiffrés… mais peut tout de même être cassé !

Code de vigenère

- Le premier et le plus simple, utilise plusieurs codes de César
- La clé possède d lettres $K = k_1\ k_2 \ldots K_d$
- $i^{ème}$ lettre détermine le $i^{ème}$ alphabet à utiliser
- Exemple avec le mot-clé MONARCHIE
- key: monarchiemonarchiemonarchiem
- plaintext: ausecoursnousommesdecouverts
- ciphertext: MIFETQBZWZCHSFOTMWPSPOLXLZXE

Code 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

Sécurité du code de Vigenère

- Une lettre du message en clair est transformée en plusieurs lettres différentes dans le message chiffré : cela brouille les fréquences
- Recherche des fréquences d'occurrence
  - Recherche de la longueur du mot clé en découpant le message chiffré en blocs de toutes les tailles possibles
- Les répétitions dans le message chiffré peuvent donner des indications

Auto-Chiffrement

- Pour éliminer la nature périodique du code précédent, la clé n'est utilisée qu'une seule fois, en préfixe du message
- Le message lui-même devient la suite de la clé
- Exemple avec MONARCHIE

  plaintext: ausecoursnoussommesdecouverts

  key: monarchieausecoursnoussommesd

  ciphertext: ...

Les chiffrements par tranposition

- Permet de mieux cacher le message en changeant l'ordre des lettres
- Peut même être utilisé sans changer les lettres elles-mêmes ce que l'on pourra reconnaître puisque toutes les fréquences d'apparition seront conservées…

Chiffrement "Rail Fence"

- Ecrire le message en posant les lettres en diagonale sur un certain nombre de lignes
- Puis lire le message ligne par ligne
- exemple

```
r   n   e   v   u   a   r   s   e   o   r
  e   d   z   o   s   p   e   l   c   u   s
```

le message chiffré est : **RNEVUARSEOREDZOSPELCUS**

http://www.apprendre-en-ligne.net/crypto/transpo/railfence.html

Transposition de colonnes

- Un schéma plus complexe
- Ecrire le message en clair en ligne sur un nombre spécifié de colonnes
- Réordonner les colonnes conformément à la clé

```
Key:        4 3 1 2 5 6 7
            1 2 3 4 5 6 7
Plaintext:  r e n d e z v
            o u s a p r e
            s l e c o u r
            s
```

le message chiffré est : **DACNSEROSSEULEPOZRUVER**

Chiffrements hybrides

- Les chiffrements utilisant substitutions et transpositions ne sont pas totalement sûrs à cause des caractéristiques statistiques des langues
- On peut imaginer d'en utiliser plusieurs en cascade pour rendre la tâche de l'adversaire plus difficile
  - deux substitutions sont plus complexes qu'une seule, mais cela reste une substitution
  - deux transpositions sont plus complexes qu'une seule, mais cela reste une transposition
  - Mais une substitution suivie par une transposition est beaucoup plus difficile
- Nous arrivons dans le monde du chiffrement moderne…

Machines à rotor

- Les machines à rotor ont été les machines les plus répandues, surtout pendant la seconde guerre mondiale
  - Enigma
- Elles implémentaient une méthode de substitution vraiment complexe et variable
- Avec une série de cylindres, chacun définissant une substitution, qui fait tourner et change après chaque lettre chiffrée
- Avec 3 cylindres, 26^3=17576 alphabets

Machines à rotor

- Les travaux commencés par les polonais

https://fr.wikipedia.org/wiki/Enigma_(machine)

Dès 1931, le Service français de renseignement (surnommé le « 2e Bureau ») était parvenu à recruter une source (Hans-Thilo Schmidt) au sein même du bureau du chiffre du ministère de la Reichswehr. Il obtient de lui de premières copies de la documentation ; il les proposa à l'Intelligence Service britannique, qui se montra sceptique, et au service polonais, qui fut très intéressé. Une coopération s'instaura, qui allait durer jusqu'en 1939. Les Français continuèrent de fournir de nouveaux renseignements obtenus de la même source, et les Polonais montèrent une équipe qui parvint à reproduire la machine à partir de la documentation de plus en plus précise qui leur parvenait.

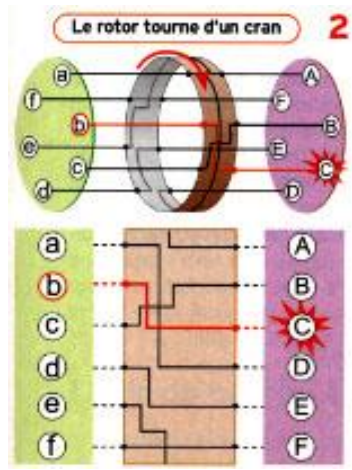
- Transmis à l'ambassade de Grande-Bretagne
  - deux jours avant l'invasion par l'Allemagne
- Les informations obtenues donnaient un net avantage dans la poursuite de la guerre
- Le conflit en Europe s'est considérablement écourté grâce à la cryptanalyse du code allemand

Enigma Machine

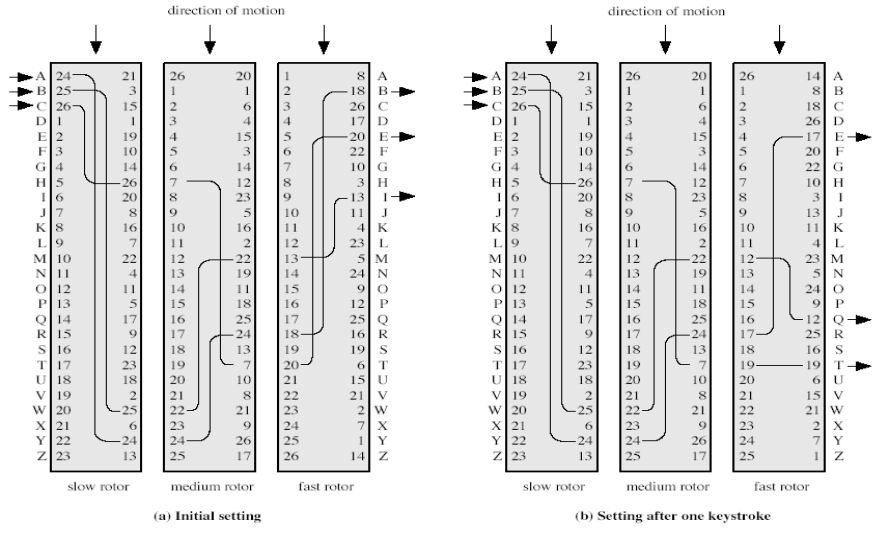




Un exemple de machine à rotor

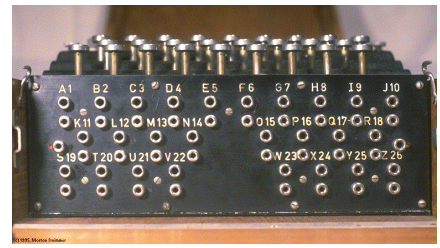


Un exemple de machine à rotor

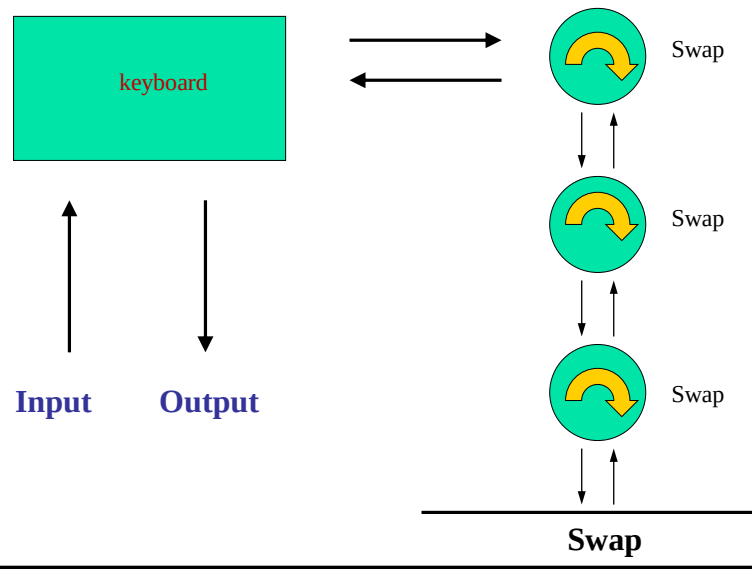


Plug Board

- Changement de lettres avant et après le passage par les cylindres



Procédure



Stéganographie

- Une alternative au chiffrement ; (du grec steganos, couvert et graphein, écriture)
- Cacher l'existence même des messages
- Utiliser un sous-ensemble de lettres ou de mots dans un message plus long
- Utiliser un pixel précis dans une séquence d'images vidéo…
- Inconvénients
  - Énorme overhead pour cacher peu d'informations

Intérêts de la stéganographie

- Communiquer en toute liberté même dans des conditions de censure et de surveillance
- Protéger ses communications privées là où l'utilisation de la cryptographie n'est normalement pas permise ou soulèverait des suspicions
- Contrebalancer toutes les législations ou barrières possibles empêchant l'usage de la cryptographie
- Publier ouvertement (mais à l'insu de tous) des informations qui pourront ensuite être révélées et dont l'antériorité sera incontestable et vérifiable par tous

Stéganographie (exemple 1)

- **Alfred de Musset écrit à George Sand :**

Quand je vous jure, hélàs, un éternel hommage
Voulez-vous qu'un instant je change de langage
Que ne puis-je, avec vous, goûter le vrai bonheur
Je vous aime, ô ma belle, et ma plume en délire
Couche sur le papier ce que je n'ose dire
Avec soin, de mes vers, lisez le premier mot
Vous saurez quel remède apporter à mes maux.

- **George Sand a répondu :**

Cette grande faveur que votre ardeur réclame
Nuit peut-être à l'honneur mais répond à ma flamme.

*George Sand est le pseudonyme d' Amandine Aurore Lucile Dupin

Stéganographie (exemple 2)

- http://lwh.free.fr/pages/algo/crypto/steganographie.htm

Stéganographie (exemple 3)

- Cacher un fichier text dans un autre
  - echo how are you doing > file.txt
  - echo password > file.txt:hidden.txt
  - dir file.txt
  - notepad file.txt:hiddent.txt
- Cacher une image dans un fichier text
  - type image.jpg>file.txt:image.jpg
  - Mspaint file.txt:image.jpg

Fonction de hachage : intégrité

- Fonction mathématique qui, à un ensemble de nombres en entrée, fait correspondre un ensemble de nombres de cardinal plus petit en sortie ;
  - La modification d'un élément en entrée engendre une modification de sa fonction de hachage en sortie.

M	h(M)
`remarquez la fin de cette ligne,`	`4b:2c:65:c0:e8:ee:95:5f:eb:05:9f:3c:6d:2f:2f :0f:9a:26:00:b7`
`remarquez la fin de cette ligne!`	`9e:e9:22:99:11:7f:41:23:7c:ce:38:3a:d8:05:18 :0c:4a:fc:ab:4c`
`??`	`75:cc:4b:e0:a9:7c:76:34:78:58:bf:04:db:3b:90 :2b:45:6a:2b:c0`

- Propriétés
  - Deux messages différents ont deux empreintes différentes
  - Connaissant h(M), il est impossible de trouver M

Fonction de hachage

- Fonctions de hachage :
  - si *H(x)* est une telle fonction, pour tout y donné, il doit être **quasi-impossible** de trouver un x tel que *H(x)=y* ;
  - si *y=H(x)*, et *x' ≈ x* à un tout petit détail près (ex : changer 1 bit), on doit avoir *y'=H(x')* très ≠ de y
  - Collision : si *H(y)=H(x)* sachant *x ≠ y*
  - Problème: hachage doit avoir un petit nb de collisions
  - Ex de fonctions de hachage: MD5, SHA-1, MD4, RIPEMD-160

Hash Function Algorithms

- **MD5**
  - Digest sur 128-bit
- **SHA-1**
  - Digest sur 160-bit