La sécurité des données

Valérie Ménissier-Morain
Laboratoire d'Informatique de Paris 6, UPMC

Systèmes de chiffrement symétriques (suite)

Transposition

- La scytale (cf. dispositif physique).
- Le plus souvent en combinaison avec d'autres systèmes de chiffrement.
- On écrit le message sur des lignes de longueur fixe, on permute les colonnes obtenues puis on relit le texte colonne par colonne.
- La clef est évidemment la permutation de colonnes.
- Elle est souvent décrite par un mot-clef. À chaque caractère du mot-clef est associée sa position dans ce mot puis on réordonne alphabétiquement les lettres du mot, ce qui induit une permutation des colonnes.
- Par exemple pour chiffrer la phrase :
  `Mais dis donc, on n'est quand même pas venus pour beurrer les sandwichs!`
  avec le mot-clef MONTAUBAN, on numérote les colonnes puis on les réordonne

MONTAUBAN	AABMNNOTU
123456789	587139246

  on écrit le message en majuscules sans accents ni ponctuation
  MAISDISDONCONNESTQUANDMEMEPASVENUSPOURBEURRERLESSANDWICHS
  sur 9 colonnes (la longueur de la permutation), en complètant la dernière ligne par des X (du *padding*)

123456789
MAISDISDO
NCONNESTQ
UANDMEMEP
ASVENUSPO
URBEURRER
LESSANDWI
CHSXXXXXX

  puis on relit le texte en commençant par la colonne 5 puis la colonne 8, etc.
  DNMNUAXDTEPEWXSSMSRDXMNUAULCIONVBSSOQPORIXACASREHSNDEESXIEEURNX
- Cryptanalyse : retrouver la longueur de la permutation (c'est un diviseur de la longueur du texte et des distances entre les paddings), s'aider des fréquences des bigrammes pour replacer les colonnes l'une par rapport à l'autre.

Techniques pour brouiller l'analyse de fréquences

Substitution homophonique

- On substitue aux caractères les plus fréquents de la langue plusieurs symboles. Lorsqu'on doit chiffrer un tel caractère on choisit au hasard parmi les symboles qui représentent le caractère.
- Pour brouiller encore davantage l'esprit du cryptanalyste on introduit sovuent des lettres nulles, des symboles spéciaux pour masquer les bigrammes les plus fréquents, et un petit nomenclateur pour représenter des mots de liaison courants par un seul symbole.
- Le premier document qui utilise cette technique est dû au doge de Venise Steno en 1411

- Cette technique sera énormément utilisée au XVe et XVI siècle en Europe.

Surchiffrement

Principe chiffrer avec un autre système de chiffrement un message déjà chiffré
Deux exemples :

Le chiffre du Che substitution mono-alphabétique + substitution poly-alphabétique : un premier chiffrement associe au hasard à chaque lettre de l'alphabet un nombre de 1 ou 2 chiffres, un nombre de 1 chiffre n'étant pas le chiffre des dizaines d'un nombre de 2 chiffres pour pouvoir déchiffrer sans ambiguïté, le second chiffrement consiste à ajouter un multi-décalage représenté par un nombre.

ADFGVX substitution par carré de Polybe dont les lignes et les colonnes sont indexées par les lettres A, D, F, G. V et X, puis transposition. Inventé par Nebel (mars 1918) ...et cryptanalysé par Painvin (juin 1918).

Cryptosystème réputé cryptanalysé mais toujours utilisé ?!

- Playfair, double Playfair, ou double transposition, bien que cryptanalysés pendant la première guerre mondiale, ont été utilisés pendant la seconde guerre mondiale pour des communications importantes mais non-critiques sur les champs de bataille.
- Résistance : temps de cryptanalyse / durée de vie de l'information
  On peut utiliser un cryptosystème pour chiffrer une information tant que le temps pour effectuer la cryptanalyse du système dépasse la durée de la pertinence de l'information.
  Évidemment on ne sait pas quand la cryptanalyse devient assez rapide pour fournir le clair avant la péremption de l'information ! D'où la question : existe-t-il un chiffrement inattaquable ? Est-ce une panacée ?

Chiffrement parfait

- En voulant contrer la faiblesse du système de chiffrement de Vigenère, Vernam a inventé un système de chiffrement parfait :
  la connaissance d'un message chiffré ne révèle aucune fuite d'information sur la clef ou le message clair correspondant et aucune information non plus sur les textes chiffrés futurs.
- Principe : un Vigenère avec une clef aléatoire (au moins) aussi longue que le message à chiffrer, utilisée une seule fois.
- Utilisé dans la cryptographie Top Secrète (téléphone rouge, valise diplomatique, militaire (Atomique)).
- C'est le seul cryptosystème à chiffrement parfait mais c'est très peu pratique !
  - Il faut transmettre au préalable une clef assez longue pour anticiper la longueur possible du message
  - Cet usage unique d'une clef impose le casse-tête de la transmission sécurisée de la nouvelle clef à chaque transmission de message
  - L'usage unique d'une clef impose d'être sûr même après des années de ne pas réutiliser une clef
  - Il est difficile de produire des clefs vraiment aléatoires
  - Donc non ce n'est pas la panacée et il est très peu usité

Substitution poly-alphabétique : Vigenère

- D"après une idée de Alberti en 1466, et après quelques autres mathématiciens européens, Vigenère propose en 1586 un système de chiffrement poly-alphabétique qui ne sera cassé qu'en 1863.
- La clef est un mot ou une phrase qui va servir à induire des décalages différents selon la position des caractères dans le message.
- Chaque caractère est décalé d'autant de positions dans l'alphabet que le caractère correspondant de la clef l'est de A. Visuellement sur la table ci-dessous, on suit la ligne correspondant au caractère du message, la colonne qui correspond au caractère de la clef et à l'intersection on trouve le caractère chiffré.

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

Remarque : Vous trouvez que cela ressemble à une table d'addition ? c'est normal puisque en associant à un caractère sa position dans l'alphabet, le caractère c décalé de k positions peut s'écrire $(c + k) \bmod 26$.

Par exemple si le message est QUININE et la clef est MONTAUBAN :
  - Q est chiffré par le caractère à l'intersection de la ligne commençant par Q et la colonne commençant par M, ce qui donne C
  - U est chiffré par le caractère à l'intersection de la ligne commençant par U et la colonne commençant par O, ce qui donne I
  - I est chiffré par le caractère à l'intersection de la ligne commençant par I et la colonne commençant par N, ce qui donne V
  - etc.
  - finalement QUININE est chiffré par CIVGIHF.
- Si le message est plus long que la clef, ce qui est généralement le cas, on découpe le message en ligne de la longueur de la clef, on chiffre chacune de ces lignes puis on recolle les chiffrés des lignes pour obtenir le chiffré du message. Cela revient à écrire le texte en colonnes et à appliquer le même décalage à tous les caractères de la colonne.
  Par exemple pour la phrase :
  `Et pourquoi pas de la quinine et un passe-montagne?`
  réécrite tout en majuscules sans espaces ni ponctutaiton et en lignes de 9 colonnes :

ETPOURQUO
IPASDELAQ
UININEETU
NPASSEMON
TAGNE

  le résultat écrit par paquets de 5 caractères comme pour les communications usuelles :
  QHCHU LRUBU DNLDY MADGW ABNYF THZDN LSYNO AFOTG E
- Utilisations très variées : dans des livres, dans des films et à titre militaire, pendant la guerre de Sécession, mais avec en toutt et pour tout seulement 3 clefs différentes donc une utilsation peu sûre.
- Cryptanalyse détection de la longueur de la clef puis/et analyse des décalages (analyse de fréquences plus poussée)

Chiffrement polygrammique : le chiffrement de Playfair

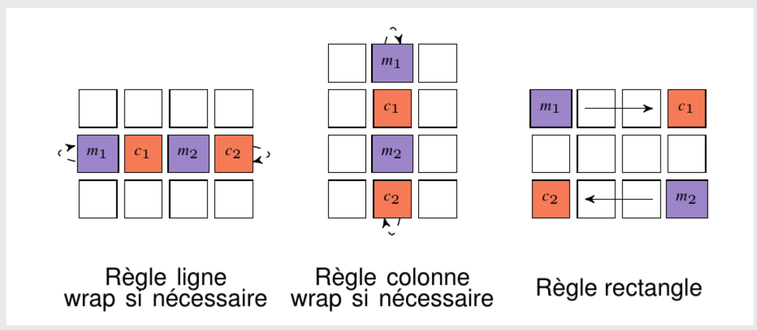
Polygrammique chiffrer non plus des caractères individuels mais des blocs de caractères (des polygrammes). On travaille sur un alphabet à 26^2 caractères donc le nombre de clef possibles possibles augmente énormément et l'analyse de fréquences sur les caractères individuels ne sert à rien et l'analyse sur les polygrammes est très ardue.

Playfair Exemple de substitution polygrammique : le chiffrement de Playfair
Inventé par Wheatstone en 1854 et popularisé par Lord Playfair, utilisé pendant la Première Guerre Mondiale par les anglais et la Seconde Guerre Mondiale par les australiens, bien que ce chiffrement ait été analysé en 1914 par Mauborgne.

La clef secrète se présente sous la forme d'un carré 5x5 qui contient toutes les lettres de l'alphabet, en en confondant 2, telles que I et J ou V et W.

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

Principe : on analyse le texte par groupe de 2 caractères et on applique selon ces caractères l'une des 3 règles suivantes :



S'il ne reste qu'un seul caractère à chiffrer ou si les deux caractères sont identiques : on ajoute une lettre neutre telle que X après le premier caractère et on applique à nouveau les règles à partir du premier caractère.

ATTAQUE PARIS DEMAIN AVEC PETIT TRAIN BLEU
AT TA QU EP AR IS DE MA IN AV EC PE TI TT RA IN BL EU
AT TA QU EP AR IS DE MA IN AV EC PE TI TXTR AI NB LE UX
RK KR ZD RG EA RI YI KM EF MC AN GR WR XL XP ER BF NR DV
RKKRZ DRGEA RIYIK MEFMC ANGRW RXLXP ERBFN RDV