Cours : Cryptographie et Sécurité Informatique

Cryptographie classique II
Chiffres par substitution

I. Akharraz
Université Sidi Mohamed Ben Abdellah

ismail.akharraz@usmba.ac.ma

Sommaire.

1. Introduction.

- Chiffre par substitution : un caractère du texte clair est remplacée par un autre caractère du même alphabet ou d'un autre alphabet suivant un algorithme précis.

- Si le symbole de substitution est fixe : Chiffre mono-alphabétique.
  - Le chiffre de César,
  - Le chiffre affine,
  - Les chiffres désordonnés.

- Si le symbole de substitution change : Chiffre poly-alphabétique.
  - Le chiffre de Vigenère.
  - Le chiffre de Hill.
  - Le chiffre de Vernam.

2. Chiffre de César.

- Le code César est mono-alphabétique.
- Pour coder on décale les lettres :

A devient D, B devient E, ..., Z devient C: décalage par 3

- Formulation mathématique du code César :
- A=0, B=1, C=2,D=3, E=4, F=5, G=6, H=7, I=8, J=9, K=10, L=11, ,M=12, N=13, O=14, P=15, Q=16, R=17, S=18, T=19, U=20, V= 21, W=22, X=23, Y=24, Z=25
- $\{A, B, \ldots, Z\} \cong (\mathbb{Z}/26\mathbb{Z}, +, *)$.
- Décalage par trois $\Longleftrightarrow$ C(x) = x + 3 modulo 26

Exemple : Clair : RENDEZ VOUS DEMAIN MIDI VILLETANEUSE

17 4 13 3 4 25 21 16 20 18 3 4 12 0 8 13 12 8 3 8 21 8 11 11 4 19 0 13 4 20 4

+ 3 mod 26

20 7 17 6 7 2 24 19 23 21 6 7 3 11 17 15 11 6 11 24 11 14 14 7 22 3 16 7 23 7

Chiffré : UHQGHC YRXV GHPDLQ PLGL YLOOHWDQHXVH

- Le chiffre de césar se généralise par :

$$C(x) = x + N \text{ modulo } 26$$

où $N \in \{0, 1, \ldots, 25\}$ est la clef du code.

Exemple. N = 7 : le chiffre du texte codé précédement devient :

YLUKLG CVBZ KLTHPU TPKP CPSSLAHULBZLBZL

- Le déchiffrement se fait en utlisant la relation :

$$x = y - N \text{ mod } 26$$

- Si on sait que le code utilisé est Cesar, On peut facilement le casser par force brute : 25 clés sont possibles $\Longrightarrow$ 25 essais au maximum.

(Attaque par force brute : On essaye toutes les clés possibles.)

- Le problème c'est que le chiffre mono-alphabétique restait un mystère pendant des siecles.

2. Chiffre de César.

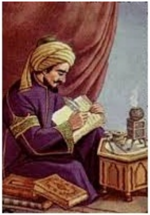
- Jusqu'à ce que Al-Kindi (801-873) : découvre l'analyse des fréquences des lettres.

- Al Kindi avait étudié la fréquence des lettres de l'arabe : أ et ل sont les plus courantes, à cause notamment de l'article ال alors que le ج apparaît dix fois moins souvent, ...

- Naissance de la science de la cryptanalyse

  - Langue française :

Lettre	%	Lettre	%	Lettre	%	Lettre	%
A	9,4	H	0,8	N	7,2	U	6,2
B	1,0	I	8,4	O	5.1	V	2,1
C	2,6	J	0,9	P	2,9	W	0
D	3,4	K	0	Q	1,1	X	0,3
E	15,9	L	5,3	R	6,5	Y	0,2
F	1	M	3,2	S	7,9	Z	0,3
G	1			T	7,3		



Al Kindi

3. Chiffre affine.

- Fonction affine : $x \longrightarrow a.x + b$ (un polynôme de degré 1).

- Utiliser les fonctions affines pour chiffrer :
Le chiffre d'un caractère x est : c(x) = a.x + b mod 26,

- La clé de ce chiffre est le couple: (a,b)

- Déchiffrement : x = a^{-1}(y-b).

- Pour déchiffrer il faut que a soit inversible pour la multiplication(a^{-1} existe dans $(\mathbb{Z}/26\mathbb{Z}, \times)$) :
Sinon, on ne pourra pas déchiffrer;

- Si a = 1, on retrouve le chiffre de César.

3. Chiffre affine.

- Arithmétique : a est inversible dans $(\mathbb{Z}/26\mathbb{Z}, \times)$ si et seulement si, pgcd(a , 26) = 1 .
- $(\mathbb{Z}/26\mathbb{Z}, \times)^*=\{1, 3, 5, 6, 7, 9, 11, 17, 19, 21, 23, 25\}$
- Nombre de clés possible pour un chiffre affine :
  - 12 choix possibiles pour a.
  - 26 choix possibiles pour b.

Donc : 12 $\times$ 26 =312 choix possibles pour (a,b).

312 clés possibles.

Exemple : Soient la clé = $(a, b) = (3, 11)$

Chiffrement : $c_i = f(m_i) = 3 * m_i + 11 mod 26.$

Déchiffrement :

$a^{-1} = 3^{-1} mod 26 = 9, \ m_i = f^{-1}(c_i) = 9 * (c_i - 11) mod 26$

'NSA' $\longrightarrow$ 13 18 0 $\longrightarrow$ 24 13 11 $\longrightarrow$ 'YNL'.

Cryptanalyse du chiffre affine.

- La cryptanalyse du chiffre affine se fait en deux étapes :
  (1) Analyse des fréquences des lettres.
  (2) Recherche et résolution des équations.

Exemple.

Soit HGAHY RAEFT GAGRH DGAGM OEHIY RAAOT ZGAGJ GKFDG AZGSB INNTG KGRHE NNIRG le chiffré d'un texte en français.

(1) G apparait 12 fois, A apparait 8 fois.
(2) G est un E : f(E) = G c.à.d f(4) = 6.
(2) A est un S : f(S) = A c.à.d f(18) = 0

- On résout les équations pour que retrouver la clé (a,b).

$$
\begin{gathered}
f(4) = 6,\ f(18) = 0 \\
\Downarrow \\
4 * a + b \equiv 6 (\mathrm{mod} 26) \\
18 * a + b \equiv 0 (\mathrm{mod} 26) \\
\Downarrow \\
14a \equiv -6 \bmod 26 \iff 14a \equiv 20 \bmod 26 \\
\Downarrow \\
a = 7 \Longrightarrow b = 4.
\end{gathered}
$$

- La fct de déchiffrement est : $m_i = 15 * (c_i - 4) mod 26.$

HGAHYRAEFTGAGRHDGAGMOEHIYRAAOTZGAGJGKFDGAZGSBINNTGKGRHENNIRG
devient
TESTONSAPRESENTLESEQUATIONSSURDESEXEMPLESDECHIFFREMENTAFFINE

4. Chiffre polyalphabétique.

- Les chiffres : par permutation, césar et affine sont monoalphabétique.

- La faiblesse de ces chiffre est qu'une lettre est remplacée par une même lettre le long du cryptage. Ceci permet une cryptanalyse facile par analyse de fréquences.

- Pour améliorer la sécurité on a pensé aux codes polyalphabétique :
  - On découpe le texte clair en blocs.
  - On applique à chaque bloc une permutation différente des autres blocs.

- Vigemère , Hill, ...

5. Chiffre de Vigenère.

- Chiffre de Vigenère (1568).
- Amélioration décisive du chiffre de César.
- Ce chiffre utilise un mot comme clef qui définit le décalage appliqué au caractères de texte clair.

Exemple :

- Texte claire : CHIFFRE DE VIGENERE
- Clef : BACHELIER , la clef est répétée plusieurs fois pour être aussi longue que le texte clair.

Clair	C	H	I	F	F	R	E	D	E	V	I	G	E	N	E	R	E
Clé	B	A	C	H	E	L	I	E	R	B	A	C	H	E	L	I	E
Décalage	1	0	2	7	4	11	8	4	17	1	0	2	7	4	11	8	4
Chiffré	D	H	K	M	J	C	M	H	V	W	I	I	L	R	P	Z	I

5. Chiffre de Vigenère.

- Le code de Vigenère est un code par blocs.
- En général :
  - On se fixe une longueur de bloc m.
  - On découpe le message en blocs de m lettres.
  - On chiffre par blocs de m lettres :

la première lettre d'un bloc de m est codée avec un César de clef k_1, la deuxième avec un César de clef k_2 et la m^{eme} par un César de clef k_m.

Formulation mathématique :

Soit $m > 0$ *et* $\mathcal{P} = \mathcal{C} = \mathcal{K} = (\mathbb{Z}_{26})^m$. *Pour la clé* $K = (k_1, k_2, \ldots, k_m)$. *Le chiffre de Vigenère :*

$e_K(x_1, x_2, \ldots, x_m) = (x_1 + k_1, x_2 + k_2, \ldots, x_m + k_m)$

$d_K(y_1, y_2, \ldots, y_m) = (y_1 - k_1, y_2 - k_2, \ldots, y_m - k_m)$

5. Chiffre de Vigenère.

Exemple.

```
  jadoree couterl aradiot outelaj ournee
+ MUSIQUE MUSIQUE MUSIQUE MUSIQUE MUSIQU
= VUVWHYI OIMBULP MLSLYIX AOLMBUN AOJVUY
```

- Force du chiffre de Vigenère : la même lettre chiffrée de différentes manières = perte de la fréquence des lettres.

- Très sûr pendant 4 siècles. Ils ont été cryptanalysés officieellement par Charles Babbage et Friedrich Wilhelm Kasiski au 19^{me} siècle.

6. Chiffre de Hill.

- Cryptosystème qui généralise celui de Vigénère.
- Il a été publié par L. S. Hill en 1929.

- On choisit un alphabet de n lettres (n = 26, dans nos exemples) et une taille m pour les blocs, par exemple m = 2. Alors $\mathcal{P} = \mathcal{E} = (\mathbb{Z}/26\mathbb{Z})^2$.

- La clef de codage est une matrice inversible $K_2 \in GL_m(\mathbb{Z}/26\mathbb{Z})$, si m = 2

$$K = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{Z}/26\mathbb{Z})$$

- Si $(x_1, x_2) \in (\mathbb{Z}/26\mathbb{Z})^2$ est le message clair alors le codé sera:

6. Chiffre de Hill.

$$(y_1, y_2) = e_K((x_1, x_2)) = (x_1, x_2)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (ax_1 + cx_2, bx_1 + dx_2)$$

- La clé de déchiffrage est la matrice inverse de K dans $GL_m(Z/26Z)$.
- Par exemple avec m = 2 et

$$K = \begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix} \quad ona \quad K^{-1} = \begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix}$$

- Le cryptosystème de Hill succombe facilement aux attaques à texte clair choisi.

7. Chiffre de Vernam.

- Claude Shannon : un chiiffre est à confidentialité parfaite si un message chiffré ne permet de retrouver ni la clé secrète ni le message clair en un temps raisonnable.
- De tels codes existent : code de Vernam ou masque jetable, (Gilbert Vernam, 1917).
- Les codes de Vernam :
  - La clé est de la taille du message à envoyer.
  - Les lettres de cette clé sont choisies de façon totalement aléatoire.
  - La clé ne doit servir qu'une seule et unique fois.
- Rarement utilisés, mais, parfaitement sûr : il a longtemps protégé le fameux "téléphone rouge", qui reliait la Maison Blanche au Kremlin.
- Problème : stockage et transmission des clés.

7. Chiffre de Vernam.

Principe :

- On tranforme le texte en une suite de chiffres en base b (souvent b = 2).
- On fabrique ensuite une suite aléatoire de chiffres de même longueur (la clé)
- On ajoute les deux suites ainsi obtenues.

sans retenue, c.à.d. que l'on fait une addition chiffre à chiffre modulo b.

- Un XOR si b =2. Exemple b = 26 .

lettre codée =lettre claire + lettre de la clé mod 26

- On code le message CHIENS par la clef KZUTEG :

CHIENS + KZUTEG = NHDYSZ

- Le même que le chiffre de CERISE par la clé KCNPZC :

CERISE + KCNPZC = NHDYSZ