TD 1 chiffrement symétrique

Remarque

Les questions marquées d'une étoile ($\star$) sont optionnelles.

Cryptanalyse en boite noire

Vous êtes un espion de Cléopâtre et vous venez de recevoir un message de l'un de vos informateurs. Voici le contenu du message probablement chiffré:
H A Y W J W N Z A O P F W Q J A A P O A O Y W J A P K J O O K J P X H A Q O

Question 1

A partir d'une analyse fréquentielle, essayez de décrypter le message. On donne la fréquence des 10 lettres les plus utilisées dans les pages wikipedia :

E	:	12,10%	A	:	7,11%	I	:	6,59%	S	:	6,51%	N	:	6,39%
R	:	6,07%	T	:	5,92%	O	:	5,02%	L	:	4,96%	U	:	4,49%

Cryptanalyse du chiffrement de Vigenère

L'indice de coïncidence

L'indice de coïncidence est une technique de cryptanalyse inventée par William F. Friedman en 1920[a]. Il permet de savoir si le chiffrement est mono-alphabétique (i.e. Un seul alphabet utilisé) ou poly-alphabetique (i.e. plusieurs alphabets utilisés) et donne le nombre de caractères utilisés par alphabet (donc la taille de la clé). Dans le cas d'un chiffrement monoalphabétique, voici ça définition :

$$IC = \sum_{q=A}^{Z} \frac{n_q(n_q - 1)}{n(n-1)}$$

avec n_q l'occurence de la q^{eme} lettre de l'alphabet, et n le nombre total de caractères dans la phrase.
Pour la langue française, cet indice vaux environ 0.0746.

[a] William F. Friedman, *The Index of Coincidence and its Applications in Cryptography*, Technical paper, Washington War Department, 1920

Question 2

Calculez la valeur de l'indice de coïncidence dans le cas où les caractères sont uniformément distribués.

$\star$ *Question 3*

Calculez l'indice de coïncidence du chiffré de l'exercice précédent. En déduire si le chiffré était vulnérable au calcul de l'indice de coïncidence. On donne la fréquence des caractères du chiffré :

A	:	7	F	:	1	H	:	2	J	:	5	K	:	2	N	:	1	O	:	6	P	:	4
Q	:	2	W	:	4	X	:	1	Y	:	2	Z	:	1									

On considère maintenant le chiffré suivant:
LVLTISPJYXLDCJOJXJQLQPKFKMCGBSBRNTKQGIKDGQQUCHXQTVR

⋆ *Question 4*

Calculez l'indice de coïncidence de ce nouveau chiffré. On donne la fréquence des caractères du chiffré :

A	:	0	B	:	2	C	:	3	D	:	2	E	:	0	F	:	1	G	:	3	H	:	1
I	:	2	J	:	4	K	:	4	L	:	4	M	:	1	N	:	1	O	:	1	P	:	2
Q	:	6	R	:	2	S	:	2	T	:	3	U	:	1	V	:	2	W	:	0	X	:	3
Y	:	1	Z	:	0																		

Ici le problème vient du fait que l'on utilise un chiffrement poly-alphabetique, c'est-à-dire que l'alphabet utilisé est différent par segments dont la taille est donnée par la taille de la clé.

Question 5

Proposez une méthode pour déterminer la longueur de la clé dans un chiffrement de Vigenère en ce basant sur le calcul de l'indice de coïncidence.

Question 6

On suppose une attaque à clair choisi, c'est-à-dire que l'on peut choisir arbitrairement un clair et connaître le chiffré associé. Donnez un exemple de message permettant de récupérer la clé secrète.

Question 7

On suppose que l'on connaît un certain nombre de chiffrés (C_i). Donnez une méthode pour récupérer de l'information sur les messages en clair.

Le chiffrement de Hill (1929)

Le chiffrement de Hill, proposé en 1929 par Lester S. Hill, est un chiffrement polygraphique, c'est-à-dire que les caractères sont chiffrés par paquets de caractères. Dans le cas où l'on chiffre par paquets de deux caractères, l'algorithme de chiffrement est le suivant:

$$\begin{pmatrix} C_k \\ C_{k+1} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} P_k \\ P_{k+1} \end{pmatrix} [26]$$

Avec P_k le k^{eme} caractère du message, C_k le k^{eme} caractère du chiffré, et (a, b, c, d) la clé de chiffrement associé.

Question 8

Déterminez l'algorithme de déchiffrement et donnez la condition sur la clé pour que le déchiffrement soit possible.

Question 9

On suppose une attaque à clair choisi, c'est-à-dire que l'on peut choisir arbitrairement un clair et connaître le chiffré associé. Donnez un exemple de message permettant de récupérer la clé secrète.

⋆ *Question 10*

On suppose une attaque à clair connu, c'est-à-dire que l'on connaît un certain nombre de paires (P_i,C_i). Donnez une méthode pour récupérer la clé secrète. Indice : on pourra remarquer qu'il faut au moins 4 caractères chiffrés pour mener une attaque.

⋆ *Question 11*

Utilisez la méthode précédente pour trouver la clé de chiffrement pour la paire clair/chiffré suivante :

Clair	:	CRYPTO
Chiffré	:	GFOBDR

Le cryptosystème AES

Question 12

A partir de la définition de l'opération SubBytes, vérifiez si cet étage est sensible au calcul de l'indice de coïncidence.

Question 13

A partir de la définition des opérations AddRoundKey, SubBytes, ShiftRows et MixColumns, essayez de déterminer si certaines opérations sont issues d'un chiffrement de Cesar, de Vigenère ou de Hill. Ces opérations sont-ils sensibles au calcul de l'indice de coïndidence ?