La simulation numérique
Cryptographie

Anis SAIED
Institut Préparatoire aux Etudes d'Ingénieurs de Nabeul (IPEIN),
Université de Carthage,
Tunisie
anis_saied@hotmail.com

2025/26

1 Introduction

Le principe du codage est de transformer/coder des données de manière à pouvoir enuite les retrouver/décoder. Le but du codage est la confidentialité d'un message de manière à ce qu'il ne soit compréhensible/décodable que par son destinataire.
L'objet de ce TP est de s'intéresser à deux méthodes pour coder/crypter un message sous forme de texte.

2 Le codage de César

Principe :

Pour coder un message, on décale ses lettres de 3 dans l'ordre de l'alphabet, par exemple A devient D, B devient E, ..., V devient Z, X devient A, Y devient B et Z devient C.
Soit le message suivant :
`M = "LE TRESOR EST CACHE DANS LE CHATEAU DE: SUITE AU PROCHAIN MESSAGE!"`
alors le codage de César de M est
`MC = 'OH WUHVRU HVW FDFKH GDQV OH FKDWHDX GH: VXLWH DX SURFKDLQ PHVVDJH!'`

Décodage

Pour décoder, il suffit de décaler de -3 : C devient A, etc.
De manière générale un codage de César décale les lettres d'un entier d, qui est appelé la *clé du codage*.
Si on connaît la clé, on peut facilement décoder.

Attaque du code de César

Une attaque consiste, à partir d'un message codé, à déterminer la clé et le message décodé. On dit aussi *casser le code*, et on parle de **cryptanalyse**.
Un codage de César est facile à casser : en français la lettre la plus fréquente est le *E*, il suffit donc de déterminer la lettre la plus fréquente dans le message codé pour connaître avec une forte probabilité la clé du code.
Exemple :
Dans le message codé `'SL AYLZVY LZA JHJOL KHUZ SL JOHALHB KL: ZBPAL HB WYVJOHPU TLZZHNL!'`
La lettre la plus fréquente est le `L`, qui a pour rang 12 dans l'alphabet, et `E` a pour rang 5, donc la clé est 12 - 5 = 7.
On vérifie que le décodage donne le message `'LE TRESOR ...'`.

Exercice 1.

Q1. Ecrire une fonction **cesar(m,d)** qui décale d'un entier **d** les caractères d'un mot **m** écrit avec l'alphabet `"ABCDEFGHIJKLMNOPQRSTUVWXYZ"`. Laisser inchangés les caractères hors de l'alphabet (espaces, ponctuation, minuscules etc).

Q2. Ecrire une fonction **decesar(m,d)** qui décode le mot **m** avec la cle **d**.

Q3. Ecrire une fonction **cassecesar(m)** qui décode un mot **m** codé avec un code de César dont on ignore la clé. Tester avec ce mot : `OVRA, NIRP IBHF, IREPVATRGBEVK NHENVG CRHG-RGER TNTAR!`

3 Le code de Vigenère

Principe

Toujours décaler les lettres, mais de manière variable et périodique. La clé est une chaîne de caractères, pour coder on décale le $i^{ème}$ caractère du message du rang dans l'alphabet de la $i^{ème}$ lettre de la clé. Quand on arrive au bout de la clé on repart du début.

Exemple : la clé est `CAE`, le message est `TRESOR`,
alors le `T` est décalé du rang de `C` qui est 3, donc devient `W`,
puis le `R` est décalé du rang de `A` qui est 1, donc devient `S`,
puis le `E` est décalé du rang de `E` qui est 5, donc devient `J`,
puis le `S` est décalé du rang de `C` qui est 3, donc devient `V`,
etc.
Ce qui donne le message codé `WSJVPW`

Décodage

Pour décoder, le principe est le même avec des décalages négatifs.

Attaque du code de Vigenère

Plus difficile que le code de César, on peut pourtant le casser aussi avec une analyse fréquentielle.
L'idée est de repérer dans le message codé des séquences de lettres identiques : elles auront alors toutes les chances d'être espacées d'un multiple de la longueur de la clé. Ce qui permet, en prenant le pgcd de ces décalages, de formuler une hypothèse sur la longeur de la clé.
Ensuite on se concentre sur les lettres espacées de la longeur de la clé, qui sont codées par un code de César, que l'on détermine par analyse des fréquences.

Exercice 2.

Dans la suite, l'alphabet utilisé est `"ABCDEFGHIJKLMNOPQRSTUVWXYZ"`.

Q1. Ecrire une fonction **rang(c)** qui rend le rang ≥ 0 d'un caractère c dans l'alphabet. Ecrire la fonction réciproque **lettre(r)**.

Q2. Ecrire une fonction **vigenere(m,cle)** qui code un mot **m** avec la clé **cle** selon la méthode de Vigenère.

Q3. Ecrire une fonction **devigenere(m,cle)** qui décode le mot **m** avec la clé **cle**.

Q4. Ecrire une fonction **cassevigenere(m)** qui décode le message **m** codé avec la méthode de Vigenère, en découvrant la clé avec une analyse fréquentielle.

4 Le codage de Hill

Comment encoder avec Hill? (Principe de chiffrement)

Le chiffre de `Hill` nécessite une matrice de chiffrement `M` et un `alphabet`.
Soient le texte `DCODE` a chiffrer avec l'alphabet `ABCDEFGHIJKLMNOPQRSTUVWXYZ` et la matrice `M` d'ordre 2 :

$$\mathbf{M} = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}$$

On découpe le texte en **n-grammes** avec n la taille de la matrice, on complète avec des lettres si besoin.

La matrice `M` est une matrice 2x2, `DCODE` devient `DC,OD,EZ` (on a rajouté un `Z` pour compléter le bigramme).

On substitue les lettres du message clair par leur rang dans l'alphabet à partir de 0.

Avec `ABCDEFGHIJKLMNOPQRSTUVWXYZ` on a `A=0, B=1, ..., Z=25`. On peut aussi utiliser `ZABCDEFGHIJKLMNOPQRS`
pour avoir `A=1, B=2, ... Y=25, Z=0`.
Les groupes de lettres `DC, OD, EZ` deviennent les groupes de valeurs (3,2), (14,3), (4,25)