TP 5 - Cryptanalyse 101

9 décembre 2011

Pour toutes questions, suggestions, remarques ou autres, n'hésitez pas à m'envoyer un mail à edbonnet@hotmail.com en mettant en objet [TP Caml].

1 Introduction

La cryptanalyse est la science qui consiste à décrypter un message codé sans connaître a priori la clé du chiffrement. On va s'intéresser ici à des méthodes de cryptanalyse élémentaires, essentiellement basées sur l'étude fréquentielle.

2 Code de César

Selon Suétone, César codait les messages qu'il envoyait en réalisant un décalage constant sur toutes les lettres (par exemple A devient D, B devient E etc.). Pour l'anecdote, César faisait tondre ses messagers, écrivait sur leur crâne, attendait la repousse des cheveux, et les envoyait en mission. Ce qui fait de l'empereur romain un pionnier de la cryptographie asymétrique, où la clé publique est le messager et la clé privée est l'idée de le tondre.

Dans toute la suite, on peut supposer qu'un texte est une suite de minuscules sans ponctuation ni accent.

Question 1 *Écrire une fonction de type* `string -> int -> string` *qui prend un texte et un décalage et qui renvoie le texte codé 'à la César' correspondant.*

Question 2 *Proposer une cryptanalyse assistée pour le code de César et implémenter la.*

Question 3 *Décoder "cesar.txt".*

3 Code de permutation

On généralise maintenant l'idée de décalage à un code où la clé est une permutation des lettres $\sigma \in \Sigma_{26}$. Ainsi, si le texte clair est $a_1 a_2 \dots$ le texte codé est $\sigma(a_1)\sigma(a_2)\dots$

Question 4 *Écrire une fonction de type* `int -> int vect` *qui génère une permutation aléatoire.*

Question 5 *Écrire un programme qui code un texte selon le code de permutation.*

Question 6 *Écrire un programme qui prend en entrée un texte et qui calcule la fréquence d'apparition de ses lettres.*

L'appel de votre fonction de fréquence sur le texte contenu dans le fichier "clair.txt", qu'on jugera suffisamment long, vous donne donc une fréquence théorique pour un texte en français.

Question 7 *Proposer une cryptanalyse pour le code de permutation et implémenter la.*

Question 8 *Décoder "permutation.txt".*

4 Code de Vigénère

Le code de Vigénère est une autre façon de généraliser le code de César basée sur le décalage. Maintenant le décalage n'est pas constant mais est périodique et est donnée par un mot appelé clé.

Si le texte clair est $a_1a_2\dots a_ka_{k+1}\dots$ et la clé est le mot $b_1b_2\dots b_k$ le texte codé est $(a_1+b_1)(a_2+b_2)\dots(a_k+b_k)(a_{k+1}+b_1)\dots$ où la somme de deux lettres est leur somme modulo la taille de l'alphabet avec 'a'=0 et z='25'.

Question 9 *Écrire une fonction de type* `string -> string -> string` *qui prend un texte et une clé et qui renvoie le texte codé 'à la Vigénère'.*

Question 10 *Définir une quantité sur un texte, invariante par permutation et d'autant plus grande que le texte privilégie certaines lettres.*

Question 11 *Proposer une cryptanalyse pour le code de Vigénère et implémenter la.*

Question 12 *Décoder "vigenere.txt".*

5 Extensions

Question 13 *Proposer un code mélangeant permutation et Vigénère et implémenter le.*

Question 14 *Réfléchir à une cryptanalyse du code proposé et implémenter la.*