ASI331 - TP1
Cryptanalyse d'un schéma de chiffrement de Vigenère

Françoise Levy-dit-Vehel

10 septembre 2020

Il s'agit ici de cryptanalyser la substitution polyalphabétique de Vigenère. Celle-ci est définie par un mot-clef, qui constitue la clef secrète (cf. cours).

Méthode

Étape 1 - Détermination de la longueur de la clef : indice de coïncidence (Friedman, 1920)
On note $c_0c_1c_2\ldots$ les lettres du chiffré. Pour $m = 3,\ldots$ on calcule les quantités

$$IC(\mathbf{y}_i) = \frac{\sum_{k=0}^{25} f_k^{(i)}(f_k^{(i)}-1)}{|\mathbf{y_i}|\cdot(|\mathbf{y_i}|-1)}$$

où, pour $0 \leq i < m$, $|\mathbf{y}_i|$ est la longueur de la sous-suite chiffrée $\mathbf{y}_i = c_ic_{m+i}c_{2m+i}\ldots$ et $f_k^{(i)}$ est le nombre d'apparitions de la k-ème lettre de l'alphabet dans la sous-suite $\mathbf{y}_i$. On calcule $IC(\mathbf{y}_i)$ pour toutes les sous-suites $\mathbf{y}_i$, $0 \leq i < m$. Si m est égal à la longueur de la clef, alors les $I_c(\mathbf{y}_i)$ sont tous proches de $I_c = 0.071$ qui est l'indice de coïncidence du français. Sinon les $I_c(\mathbf{y}_i)$ sont tous proches de la valeur $1/26 \approx 0.038$, correspondant à une répartition aléatoire des lettres.

Étape 2 - Détermination de la clef
On suppose que la longueur m de la clef est connue.
1-Recouvrement de la clef à décalage près : indice de coïncidence mutuelle
L'indice de coïncidence mutuelle permet de retrouver les positions relatives des lettres de la clef dans l'alphabet.
Soit $\mathbf{y} = c_0c_1c_2\ldots$ le chiffré et soit $\mathbf{y}_i$, $0 \leq i < m$, les sous-suites de $\mathbf{y}$ définies comme précédemment.

Fixer $i \in \{1,\ldots,m-1\}$
Pour $g = 0...25$, calculer les indices de coïncidence mutuelle

$$ICM_{0,i,g} = ICM(\mathbf{y}_0, \mathbf{y}_i^g) = \frac{1}{|\mathbf{y}_0|\cdot|\mathbf{y}_i|}\cdot\sum_{k=0}^{25} f_k^{(0)} f_{k-g}^{(i)}$$

où $k-g$ s'entend $\bmod 26$ et $f_k^{(i)}$ est la fréquence de la $k$ème lettre de l'alphabet dans la sous-suite $\mathbf{y}_i$.

Garder le $(0, i, g)$ dont l'indice de coïncidence mutuelle $ICM_{0,i,g}$ est le plus élevé (idéalement, proche de $0,071$).
Répéter ce calcul pour tous les $i \in \{1, \ldots, m-1\}$.

Déterminer la clef utilisée à décalage k_0 près en résolvant un système linéaire issu des triplets d'indices conservés précédemment.

2-Recouvrement de la clef : déterminer k_0 et en déduire la clef.

Étape 3 - Recouvrement du clair
En inversant le procédé de chiffrement, retrouver le texte clair.

Mise en oeuvre informatique

Chiffrement et Cryptanalyse

1. Chiffrer le message se trouvant dans le fichier `Clair.txt` à l'aide de la clef `BEAU` ; déchiffrer ensuite le message chiffré obtenu et vérifier que l'on retrouve bien le clair.
2. En appliquant la méthode ci-dessus, retrouver la clef et le texte clair correspondant au chiffré présent dans le fichier `Chiffre.txt`.

Challenge 4 de Simon Singh
Dans son livre "Histoire des codes secrets", Simon Singh propose un prix de 10.000 Livres Sterling à la personne capable de résoudre les 10 challenges proposés dans l'appendice.
Le quatrième challenge est un texte chiffré par un cryptosystème de Vigenère, dont le texte se trouve dans le fichier *Vigenere_ Singh.txt*.
Retrouver la clef et le texte original.