Université de Versailles – Saint-Quentin-en-Yvelines
Université Paris Saclay

le 09 / 11 / 2023

Jules Baudrin (Inria Paris)

Master 1 Informatique 2023–2024
Cryptographie

Correction exercice 4 TD 2 : permutations et substitutions

On considère un chiffrement sur n lettres, sur un alphabet de taille k. Le chiffrement consiste en la composition d'une permutation des positions des lettres du mot et d'une substitution mono-alphabétique.

1. Combien y a-t-il de clefs possibles ? Y a-t-il des clefs équivalentes ?

La clé est constitué d'une permutation des positions des lettres, ainsi que d'une permutation de l'alphabet. Les permutations d'un ensemble à t éléments sont au nombre de $t!$ (*c.f* compléments de maths par exemple). Il existe donc $n! \times k!$ clés pour ce chiffrement.

Soient $K \neq K'$ deux clés distinctes. Montrons qu'il existe un message x tel que $E_K(x) \neq E_{K'}(x)$ (autrement dit, qu'il n'existe pas de clés équivalentes).

Puisque $K \neq K'$, soit les permutations des positions $\sigma_K, \sigma_{K'} \in S_n$ sont distinctes, soit les permutations de l'alphabet sont distinctes $\tau_K, \tau_{K'} \in S_k$ (pas forcément de manière exclusive).

Supposons que $\sigma_K, \sigma_{K'} \in S_n$ sont distinctes. Dans ce cas, il existe un indice i tel que $\sigma_K(i) \neq \sigma_{K'}(i)$. Soient $a, b \in \mathcal{A}$ de lettres de l'alphabet. On considère le message $X = (a, a, \cdots, \underset{\text{pos } i}{b}, a, \cdots, a)$ Par construction, $E_K(x)$ ne contiendra que deux caractères distincts, l'un présent $n-1$ fois, l'autre présent 1 fois à la position $\sigma_K(i)$. $E_{K'}(x)$ ne contiendra que deux caractères distincts (potentiellement différents), l'un présent $n-1$ fois, l'autre présent 1 fois à la position $\sigma_{K'}(i)$. Puisque $\sigma_K(i) \neq \sigma_{K'}(i)$, les chiffrés $E_K(x)$ et $E_{K'}(x)$ sont nécessairement différents.

Supposons que $\tau_K, \tau_{K'} \in S_k$ sont distinctes. Dans ce cas, il existe une lettre $a \in \mathcal{A}$ tel que $\tau_K(a) \neq \tau_{K'}(a)$. Notons $\tau_K(a) = b$ et $\tau_{K'}(a) = c$ (on a donc $b \neq c$). On considère le message $X = (a, \cdots, a)$ uniquement constitué de a. Par construction, $E_K(x)$ ne contiendra que le caractère b présent n fois. De même, $E_{K'}(x)$ ne contiendra que le caractère c présent n fois. Puisque $b \neq c$, les chiffrés $E_K(x)$ et $E_{K'}(x)$ sont différents.

Ainsi il n'existe pas de clés équivalentes.

Etant donné les couples clairs-chiffrés suivants, donner le message correspondant au 5ème chiffré intercepté : BCGMMHN

clairs	chiffrés
MGNQSCO	JOANZGY
EFMJGOT	ZGMATFB
DBIHRWU	CKHLIUR
VKULDAP	HXWUSPE

Le chiffrement utilisé chiffre toujours les mêmes lettres de la même façon (substitution mono-alphabétique), en revanche en fonction de la position initiale dans le message, la position de la lettre chiffré diffère.

On commence par observer que dans les deux premiers clairs, les seules lettres identiques qui sont utilisées sont M, G, O. Ainsi {M, G, O} est nécessairement envoyé sur l'ensemble des lettres qui apparaissent dans les deux chiffrés, à savoir {A, G, Z}. On en déduit aussi que les positions de M, G, O dans les clairs sont envoyées sur les positions de A, G, Z dans les chiffrés. D'où $\{1, 2, 7\} \to \{3, 5, 6\}$ et $\{3, 5, 6\} \to \{4, 1, 2\}$.

De même avec les deux derniers clairs, les seules lettres identiques qui sont utilisées sont D, U. On en déduit d'une part une correspondance sur les lettres $\{D, U\} \to \{H, U\}$ et d'autre part deux correspondances sur les positions $\{1, 7\} \to \{3, 6\}$ et $\{3, 5\} \to \{1, 4\}$.

Ainsi puisque $\{1,\ 2,\ 7\} \to \{3,\ 5,\ 6\}$ et $\{1,\ 7\} \to \{3,\ 6\}$, nécessairement $2 \to 5$. Ainsi on peut associer le 2ème caractère des clairs au 5ème des chiffrés : $G \to Z$, $F \to T$, $B \to I$, $K \to S$.

De même, puisque $\{3,\ 5,\ 6\} \to \{4,\ 1,\ 2\}$ et $\{3,\ 5\} \to \{1,\ 4\}$, nécessairement $6 \to 2$. Ainsi on peut associer le 6ème caractère des clairs au 2ème des chiffrés : $C \to O$, $O \to G$, $W \to K$, $A \to X$.

Puisque $\{\mathrm{M},\ \mathrm{G},\ \mathrm{O}\} \to \{\mathrm{A},\ \mathrm{G},\ \mathrm{Z}\}$, $G \to Z$, $O \to G$, nécessairement $\mathrm{M} \to \mathrm{A}$. Pour le premier couple, M est la 1ère lettre du clair et A la 3ème du chiffré d'où $1 \to 3$.

$O \to G$ donne avec le premier couple $7 \to 6$.

$G \to Z$ donne avec le second couple $5 \to 1$.

$\{3,\ 5\} \to \{1,\ 4\}$ et $5 \to 1$ donc nécessairement $3 \to 4$.

Il ne reste pour 4 plus qu'un seul choix : $4 \to 7$.

On en déduit finalement que la permutation utilisée est :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 4 & 7 & 1 & 2 & 6 \end{pmatrix}$$

Le 5 ème chiffré utilise comme lettres B,C,G,M,H et N qui apparaissent toutes dans au moins un chiffré de la liste. On peut donc, grâce à la permutation désormais connue, retrouver les correspondances des lettres manquantes. Le déchiffrement du 5ème message est laissé au lecteur.

Quelle peut être la complexité en donnée dans une attaque à clair choisi quand le nombre de lettres est égal à la taille de l'alphabet.

Une attaque avec deux couples clairs-chiffrés existe. Par exemple si on demande le chiffré de $\mathrm{ABCDEFGHIJ}$ et $\mathrm{AABCDEFGHI}$. On en déduit le chiffré de A grâce à l'unique lettre, disons α, qui apparait en double dans le deuxième chiffré. Puis on en déduit à quelle position le premier caractère est envoyé en cherchant la position i de α dans le premier chiffré. On peut alors en déduire là où est envoyé le deuxième caractère, il s'agit de la position j du deuxième α dans le deuxième chiffré. Ainsi on en déduit le chiffré de B, β, en cherchant le caractère à la position j dans le premier chiffré. On continue successivement de trouver en alternance, image de la permutation des positions et image de la substitution mono-alphabétique.