TP : Détection et cryptanalyse des messages (cryptogrammes) parallèles

On appelle **messages ou cryptogrammes parallèles** [Fil10] des chiffrés résultants de l'application d'une même masque jetable.
Ce masque est typiquement produit par un algorithme de chiffrement par flot - ou de manière équivalente par un algorithme de chiffrement par bloc en mode OFB (*output feedback mode*).
Observons que le parallèlisme est une relation d'équivalence. On peut donc considérer les classes de chiffrés résultants de l'utilisation d'un même masque jetable, que l'on appelera **classes de messages parallèles**. Deux problème se posent :

1. Identifier les classes de messages parallèles (détection) ;
2. Pour une classe de messages parallèles, retrouver la clé et le message en clair.

1 Détection des cryptogrammes parallèles

Exercice 1 : considérons le test statistique suivant :
- Hypothèse nulle : C_1 et C_2 ne sont pas des messages parallèles.
- Hypothèse alternative : C_1 et C_2 sont des messages parallèles.

Montrer que sous l'hypothèse nulle, $Z = wt(C_1 \oplus C_2 \oplus 1)$ (poids du vecteur binaire de taille n, $C_1 \oplus C_2 \oplus 1$ où l'addition est faite modulo 2) suit asymptotiquement une loi normale de moyenne $n/2$ et de variance $n/4$.
Que représente Z ?

Exercice 2 : montrer que sous l'hypothèse alternative, Z suit asymptotiquement une loi normale de moyenne np et de variance $np(1-p)$, où $p > 1/2$.
Que vaut p ? De quoi dépend la probabilité p ?

L'algorithme de détection consiste, pour tous les couples (C_1, C_2), à calculer $Z(C_1, C_2) = wt(C_1 \oplus C_2 \oplus 1)$ et à trier la liste des $Z(C_1, C_2)$, C_1, C_2 selon $Z(C_1, C_2)$.
Il est ensuite possible d'extraire les classes parallèles.

Exercice 3 : Modifier cet algorithme pour extraire les classes parallèles si des décalages existent entre les (C_1, C_2).

2 Cryptanalyse

Exercice 4 : L'attaquant dispose de 6 messages chiffrés avec le même masque. Il dispose d'un corpus des tetragrammes de la langue anglaise, qui donne pour chaque tetragramme sa fréquence d'apparition. L'objectif est de retrouver la clé et les messages en clair.

Proposer (et implanter en C) un algorithme permettant de retrouver
- les messages en clair ;
- le masque utilisé.

Références

[Fil10] E. Filiol. How to operationaly detect misuse or flawed implementation of weak stream ciphers (and even block ciphers sometimes) and break them - Application to the Office Encryption Cryptanalysis. In *Proceedings of Black Hat EU 2010*, 2010.