T.P. Théorie de l'information
Application à la cryptanalyse par force brute

ESIEA Laval - 5ème Anné

25 septembre 2014

Le but de ce projet est d'expérimenter autour de l'*Advanced Equipartition Property* (AEP) d'une source d'information de Markov dans le cas de la cryptanalyse par force brute.

Le contexte est le suivant : vous disposez uniquement d'un texte chiffré et de la description (et de l'implémentation) de l'algorithme de chiffrement ayant servi à produire ce texte chiffré.

Toute technique de cryptanalyse passe par une phase (souvent finale) d'essais exhaustifs (force brute) sur tout ou partie de la clef. Le problème consiste à trouver la clef (unique, par définition) ayant servi à la production du texte chiffré à partir d'un texte clair inconnu. Ne disposant que du cryptogramme (situation opérationnelle), le problème est d'analyser et de valider le texte clair produit selon son entropie.

L'AEP cependant fait que plusieurs textes clairs (en nombre souvent réduits) sont " admissibles " et donc autant de clefs. Une analyse manuelle résiduelle (analyse sémantique manuelle par un linguiste ou automatique) permettra au sein de l'ensemble des séquences typiques de décider quelle clef unique a été utilisée et de là, accéder au texte clair.

Dans un but de gestion du temps, vous passerez par une recherche exhaustive partielle sur la clef (23 bits).

1 Attaque du système SC par force brute partielle

Le système SC est un système de chiffrement par flot conçu dans un but pédagogique. Sa clef secrète est de 59 bits. C'est un système à combinaison de registres (voir explication détaillée en cours). Le code source est fourni (utilisation en ligne de commande pour le chiffrement/déchiffrement). Dans un but de gestion du temps, vous passerez par une recherche exhaustive partielle sur la clef (registre 3, 23 bits). Les valeurs des registres 1 (17 bits) et 2 (19 bits) vous sont données.

Pour identifier et conserver les textes clair éligibles (séquences typiques selon l'AEP), vous devrez concevoir/imaginer une fonction filtre F qui décide si le clair correspond à une séquence typique ou non.

Pour chaque valeur possible I de l'état initial du registre 3 :

1. vous produirez le texte clair résultant P_I,
2. si $F(P_I) = 1$ la séquence est typique sinon ($F(P_I) = 0$) elle ne l'est pas et la valeur I est écartée.

Vous ferez cette expérience en utilisant respectivement
- un quart du cryptogramme,
- la moitié du cryptogramme,
- les trois-quarts du cryptogramme,
- la totalité du cryptogramme.

2 Analyse des résultats selon l'AEP

Le but est de d'analyser si la suite de ces quatre cryptanalyses produit des rédultats conformes ou non à l'AEP.

Dans un mini-rapport, vous devrez
- formaliser le problème : nature et modèle de la source, définir la fonction F que vous avez utilisée ;
- présenter les résultats des quatre cryptanalyses ;
- expliquer ces résultats et conclure relativement à l'AEP ;
- mettre le texte clair et la clef en annexe, ainsi que le code source de votre programme d'analyse.

3 Description des données

Vous disposez des données suivantes :

- du code source de l'algorithme (fichiers *scex.c* et *include.h*) ;
- de textes chiffrés (cryptogrammes), un par étudiant et appelé "`<nom étudiant>.cry`". Ces textes chiffrés ont été produits à partir de textes clairs en langue étrangère.
- d'un fichier contenant les valeurs initiales (en hexadécimal) des registres 1 et 2, en fonction du cryptogramme.
- du présent document.

Le rapport est à renvoyer avant le 30 septembre 2014 minuit à l'adresse *filiol@esiea.fr*.