Cryptologie : Éléments de cryptanalyse

Matthieu Amiguet

2006 – 2007

haute école arc ingénierie
neuchâtel berne jura — saint-imier le locle porrentruy

Principe de Kerckhoffs 2

Auguste Kerckhoffs

Jean Guillaume Hubert Victor Françoise Alexandre Auguste Kerckhoffs von Niuewenhof (1835–1903)

- Dans “la cryptographie militaire”, énonce 6 principes dont le deuxième reste célèbre :

Le principe de Kerckhoffs

“Il faut qu’il [le système cryptographique] n’exige pas le secret, et qu’il puisse sans inconvénient tomber entre les mains de l’ennemi”.

Attaques actives, attaques passives 3

Attaque passive

- L’attaquant se contente d’observer la communication
- Menace la confidentialité

Attaque active

- L’attaquant tente de modifier, ajouter ou détruire une partie de l’information transmise
- Menace l’intégrité et l’authentification.

Texte crypté seulement 4

- Se contente de la connaissance d’un message crypté
- Se base sur une connaissance minimale sur le texte en clair (p.ex. sa langue de rédaction)
- Un système de cryptage sensible à une telle attaque n’est pas considéré comme sûr
- Aucune attaque de ce type connue sur les cryptages modernes.

Texte clair connu 5

- Se base sur la connaissance d'une partie du texte en clair pour déduire le reste du message
- Il est souvent difficile d'empêcher toute connaissance de ce type. Le scénario est donc réaliste.

Texte clair choisi 6

- Se base sur la possibilité de choisir un texte clair et d'obtenir son encryptage
- Scénarios possibles :
  - On possède une "boîte noire" qui fait le cryptage
  - On peut convaincre Alice d'encrypter le texte pour nous
  - Encryptage à clé publique...

Texte crypté choisi 7

- Se base sur la possibilité d'obtenir le texte en clair pour certains messages cryptés choisis
- Bien entendu, on n'a pas accès au décryptage du message lui-même !
- Scénario plus théorique (ev. plausible avec des systèmes embarqués ?).

Texte clair/crypté choisi adaptatif 8

- Similaires aux cas précédents, mais avec possibilité d'adapter les textes clairs/cryptés à traduire
- Peut être vu comme un problème d'apprentissage
  - Dans ce contexte, un "bon" cryptage est une fonction difficile à apprendre...

Voies détournées 9

- Attaques basées sur
  - le temps de calcul
  - l'énergie consommée
  - la chaleur dégagée
  - les radiations émises
  - ...
- Utilisé p.ex. contre les cartes à puces...
- ... et très récemment contre RSA (à confirmer !)

Attaques distinctives 10

- Le but est de distinguer un texte encrypté d'une chaîne aléatoire
- Souvent utilisé pour identifier le type de cryptage utilisé
- Première étape vers une attaque plus "méchante".

Connaissance sur le texte clair 11

- Toute attaque qui permet d'obtenir des connaissances *sur* le texte en clair
- Ne nécessite pas forcément le décryptage
- Exemples : longueur, parité, régularités diverses, ...

Connaissance du texte clair (décryptage) 12

- L'attaque la plus "évidente" : récupérer le texte clair
- Ne nécessite pas forcément de trouver la clé.

Encryptage 13

- Vise à encrypter un message choisi selon la même méthode/même clé qu'un message donné
- Les cryptages sensibles à une telle attaque ne se prêtent pas à l'authentification/l'identification.

Récupération de la clé 14

- Vise à récupérer la clé utilisée
- L'attaque la plus puissante : elle permet toutes les autres.

Recherche exhaustive de la clé 15

- Parcours de toutes les clés pour trouver la bonne
- Attaque à texte crypté seulement
- Permet la récupération de la clé
- Applicable en principe à tous les cryptages moins forts que le masque jetable
- Il "suffit" d'avoir un espace des clés assez grand pour éviter cette attaque
- Un cryptage est considéré comme (partiellement) cassé si on trouve une attaque plus efficace que celle-ci. . .
  - . . . ce qui ne veut pas forcément dire que le cryptage n'est plus utilisable !

Recherche exhaustive des messages 16

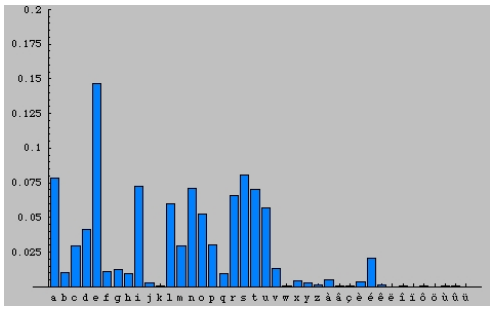
- Attaque à texte clair choisi
- Vise le décryptage (sans connaissance de la clé !)
- Il "suffit" d'avoir un espace des messages assez grand pour éviter cette attaque
  - si nécessaire, "saler" le message (i.e. ajouter de l'aléatoire).

Analyse fréquentielle 17

- Chaque langue possède un profil particulier en ce qui concerne la fréquence des lettres, des digrammes, trigrammes, etc.
- Ceci peut être utilisé pour monter une attaque à texte crypté seulement (attaque distinctive, décryptage, récupération de la clé)
  - contre les substitutions monoalphabétiques ou de polygrammes
- On peut contrer partiellement cette attaque en réduisant la redondance du message (p.ex. compression).

Exemple 18

- Féquences relatives des lettres dans un texte en français



- Voir aussi `http://mantis.free.fr/articles/freq.htm`

Méthode de Kasiski 19

- Contre les substitutions polyalphabétiques
- Attaque à texte crypté seulement
- Basé sur l'observation suivante :

  Deux portions de texte clair identiques encryptées avec la même partie du mot de passe donnent deux portions de texte crypté identiques
- Le nombre de caractères entre deux portions de texte identiques a donc des chances d'être un multiple de la longueur du mot de passe.

Méthode de Kasiski – suite 20

- Idéalement, il suffirait de calculer le pgdc de toutes les distances entre textes répétés
- Mais des répétitions peuvent arriver par coïncidence
- Il faut donc parfois chercher un peu. . .
- Une fois la longueur n de la clé déterminée, on se retrouve avec n substitutions monoalphabétiques
  - ⇝ Il ne reste "plus qu'à" faire n analyses fréquentielles.

Indice de coïncidence 21

Indice de coïncidence

Probabilité que deux lettres tirées au hasard dans un texte soient les mêmes

- L'indice de coïncidence est plus ou moins constant pour une langue donnée

Langue	Français	Anglais	Allemand	Espagnol	Italien
IC	0.074	0.065	0.072	0.074	0.075

- Un texte aléatoire (uniforme) construit sur 26 lettres a une IC d'environ 0,038.

Indice de coïncidence – utilisation 22

- Contre les substitutions polyalphabétiques
- Attaque à texte crypté seulement
- Pour $n = 1, 2, 3, \ldots$
  - On extrait les n sous-chaînes de c : $c_0 c_n c_{2n} \ldots$, $c_1 c_{n+1} c_{2n+1} \ldots$, $\ldots$
  - On calcule leur IC
  - Si n est (un multiple de) la longueur de la clé, les IC vont s'approcher de la valeur caractéristique de la langue
    - Sinon, on a de fortes chances de tourner autour de la valeur "uniforme" (p.ex. 0,038)
- Une fois la longueur n de la clé déterminée, on se retrouve avec n substitutions monoalphabétiques.

Comment calculer l'IC ? 23

- Soient
  - N la longueur du texte
  - n_A le nombre de "A" dans le texte
  - n_B le nombre de "B" dans le texte
  - $\ldots$
  - n_Z le nombre de "Z" dans le texte
- Le nombre de tirages possibles donnant "AA" est $\frac{n_A(n_A-1)}{2}$
- Le nombre total de tirages possibles est $\frac{N(N-1)}{2}$
- La probabilité de tirer "AA" est donc $\frac{n_A(n_A-1)}{N(N-1)} \ldots$
- $\ldots$ et donc $IC = \sum_{i=A}^{Z} \frac{n_i(n_i-1)}{N(N-1)}$.

Cryptanalyse différentielle 24

- Attaque contre les cryptages par bloc principalement
  - Aussi contre les cryptages par flux et les fonctions de hachage
- Publié par Biham&Shamir à la fin des années 80
- Destiné à casser DES
- DES s'est révélé particulièrement résistant à cette attaque (les concepteurs y avaient-ils pensé ?)
- D'autres cryptages ont été cassés ainsi (FEAL, ...).

Cryptanalyse différentielle – 2 25

- Attaque à texte clair choisi
- L'attaquant choisi deux textes clairs et en calcule la *différence* (souvent, ou exclusif)
- Il crypte ces deux textes et calcule la différence des textes cryptés
- En répétant cette opération un certain nombre de fois, les statistiques obtenues permettent de monter
  - une attaque distinctive
  - parfois une récupération de la clé.

Cryptanalyse différentielle – 3 26

- Cette méthode requiert une analyse fine du fonctionnement de la méthode de cryptage
  - Elle nécessite donc une parfaite connaissance de celle-ci
- Les cryptages récents (AES, . . .) sont généralement "blindés" contre la cryptanalyse différentielle.

Cryptanalyse linéaire 27

- Similaire à la cryptanalyse différentielle
  - vise cryptages par blocs
  - attaque à clairs choisis
- L'idée est de trouver des approximations linéaires de l'action du cryptage
- Peu praticable sur DES : nécessite 2^{43} paires texte clair/texte crypté (!)

Pour aller plus loin. . . 28

- `http://www.apprendre-en-ligne.net/crypto/`
- `http://fr.wikipedia.org/wiki/Cryptanalyse`
- `http://www.brics.dk/RS/95/9/BRICS-RS-95-9.pdf`
- `http://ditwww.epfl.ch/SIC/SA/publications/FI00/fi-sp-00/sp-00-page8.html`
- Alfred J. Menezes, Paul C. van Oorschot and Scott A. Vanstone, *Handbook of Applied Cryptography*, 1996, téléchargeable sur `http://www.cacr.math.uwaterloo.ca/hac/`