Sécurité et systèmes embarqués

Pablo Rauzy
pr@up8.edu
pablo.rauzy.name/teaching/sese

UNIVERSITÉ PARIS 8 VINCENNES-SAINT-DENIS
UFR MITSIC / M1 informatique

Séance 3
La cryptanalyse

La cryptanalyse

Introduction

- La *cryptanalyse* est la science qui consiste à *décrypter* un message *chiffré*.
- C'est-à-dire tenter de *déchiffrer* ce message sans posséder la clef de chiffrement.

- Le processus par lequel on tente de décrypter un message est appelé une *attaque*.
- Une attaque est souvent caractérisée par les données qu'elle nécessite :
  - attaque sur texte chiffré seul (*ciphertext-only*, COA),
  - attaque à texte clair connu (*known-plaintext attack*, KPA),
  - attaque à texte clair choisi (*chosen-plaintext attack*, CPA),
  - attaque à texte chiffré choisi (*chosen-ciphertext attack*, CCA).

Niveau de résistance

- On défini la résistance d'un cryptosystème par rapport à la puissance d'un adversaire.
- Un cryptosystème est considéré comme sécurisé si un adversaire connaissant tout du système sauf la clef de chiffrement (*principe de Kerckhoffs*) n'est pas capable de *distinguer* deux chiffrés (IND).
- On parle d'*indistinguabilité* quand étant donné deux messages et le chiffré d'un de ces deux messages, l'adversaire ne peut différencier lequel des deux messages a été chiffré.
- Suivant la puissance de l'adversaire (par ordre croissant) on dit que le cryptosystème est :
  - IND-COA,
  - IND-KPA,
  - IND-CPA,
  - IND-CCA.

- Un cryptosystème est IND-COA si l'indistinguabilité de ses chiffrés résiste à un adversaire qui dispose du texte chiffré de plusieurs messages.
- Le but de l'adversaire est de déchiffrer les messages voire de retrouver la clef utilisée.
- En pratique, on a souvent une connaissance partielle des messages, comme par exemple leur langue.

- Un cryptosystème est IND-KPA si l'indistinguabilité de ses chiffrés résiste à un adversaire qui dispose de plusieurs paires de messages et chiffrés correspondants.
- Le but de l'adversaire est de retrouver la clef ou de déchiffrer d'autres messages chiffrés avec la même clef.

- Un cryptosystème est IND-CPA si l'indistinguabilité de ses chiffrés résiste à un adversaire qui dispose de plusieurs paires de messages et chiffrés correspondants, mais qui peut en plus choisir les messages en clair.
- Le but de l'adversaire est de retrouver la clef ou de déchiffrer d'autres messages chiffrés avec la même clef.

- Un cryptosystème est IND-CCA si l'indistinguabilité de ses chiffrés résiste à un adversaire qui dispose de plusieurs paires de messages et chiffrés correspondants, mais qui peut en plus choisir les chiffrés.
- Le but de l'adversaire est de retrouver la clef de chiffrement, ou de déchiffrer un message pour lequel il n'a pas le droit de demander le déchiffrement.
- On distingue deux sous-catégories d'attaque :
  - CCA "pause déjeuner",
  - CCA adaptative.

- Le principe des attaques CCA1 est que l'adversaire ne dispose que d'une quantité limité de paires de messages et chiffrés correspondants, car il ne peut plus en demander de nouveau à partir du moment où on lui fourni son challenge.

- Le principe des attaques CCA2 est que l'adversaire peut adapter ses demandes de déchiffrement au challenge, tant qu'il ne demande pas à déchiffrer le challenge lui-même.
- Ce genre d'attaque a peu de sens en pratique, mais il est utile pour l'étude formelle et la preuve de sécurité.

Techniques de cryptanalyse

- Il existe de nombreuses familles d'attaques cryptanalytiques.
- On les range en deux groupes :
  - les attaques cryptanalytiques génériques, et
  - les attaques cryptanalytiques modernes.

Les attaques cryptanalytiques génériques

- Ces attaques sont anciennes pour la plupart et ne concernent pas vraiment la cryptographie moderne.
- Elles sont toutefois intéressantes car plus simple à appréhender et “universelles”, au sens où elles s’utilisent directement sur tous les cryptosystèmes qu’elles permettent d’attaquer.

L'analyse fréquentielle

- L'*analyse fréquentielle*, découverte au 9ème siècle par Al-Kindi, examine les répétitions des lettres du message chiffré afin de trouver la clef.
- Elle est principalement utilisée contre les chiffrements mono-alphabétiques qui substituent chaque lettre par une autre et qui présentent un biais statistique.

L'indice de coïncidence

- L'*indice de coïncidence*, inventé en 1920 par William F. Friedman, permet de calculer la probabilité de répétitions des lettres du message chiffré.
- Il permet de savoir le type de chiffrement d'un message (chiffrement mono-alphabétique ou poly-alphabétique) ainsi que la longueur probable de la clef.
- Il permet de se ramener à l'analyse fréquentielle dans le cas poly-alphabétique.

L'attaque par mot probable

- L'*attaque par mot probable* consiste à supposer l'existence d'un mot probable dans le message chiffré.
- Il est donc possible d'en déduire (une partie de) la clef si le mot choisi est correct (toujours sur les chiffrements par substitution).
- Ce type d'attaque a été mené contre la machine Enigma durant la Seconde Guerre mondiale.

L'attaque par dictionnaire

- L'*attaque par dictionnaire* consiste à tester tous les mots d'une liste comme mot clef.
- Elle est souvent couplée à l'attaque par force brute.

L'attaque par force brute

- L'*attaque par force brute* consiste à tester toutes les solutions possibles de mots de passe ou de clefs.
- Elle est peu utilisée pour les clef un peu longue car le temps nécessaire devient vite trop important.

Attaque par paradoxe des anniversaires

- Le *paradoxe des anniversaires* est un résultat probabiliste qui est utilisé dans les attaques contre les fonctions de hachage.
- Il permet de donner une borne supérieure de résistance aux collisions (de l'ordre de $\sqrt{n}$ où n est la taille de la sortie).

Les attaques cryptanalytiques modernes

- Dès les années 70 apparaissent les méthodes de chiffrement modernes par blocs.
- Les tentatives d'attaques de ces algorithmes donnent naissance à de nombreuses nouvelles méthodes de cryptanalyse puissantes, faisant appel à des outils mathématiques avancés.
- Ces méthodes ne sont pas vraiment génériques et des modifications sont nécessaires pour attaquer un type de chiffrement donné.
- Souvent, on ne s'attaque pas à une version complète de l'algorithme de chiffrement mais une variante avec moins de tours.

Cryptanalyse linéaire

- La *cryptanalyse linéaire*, due à Mitsuru Matsui, consiste à faire une approximation linéaire de la structure interne de la méthode de chiffrement.
- Elle remonte à 1993 et s'avère être l'attaque la plus efficace contre DES.
- Les algorithmes plus récents sont insensibles à cette attaque.

Cryptanalyse différentielle

- La *cryptanalyse différentielle* est une analyse statistique des changements dans la structure de la méthode de chiffrement après avoir légèrement modifié les entrées.
- Avec un très grand nombre de perturbations, il est possible d'extraire la clef.
- Cette attaque a été présentée par Eli Biham et Adi Shamir en 1990.
- Toutefois, on sait maintenant que les concepteurs de DES connaissaient une variante de cette attaque nommée "attaque-T".
- Les algorithmes récents (AES, IDEA, etc.) sont conçus pour résister à ce type d'attaque.

Cryptanalyse χ^2

- La *cryptanalyse* χ^2, concept dû à Serge Vaudenay, permet d'obtenir des résultats similaires à des attaques linéaires ou différentielles.
- L'analyse statistique associée permet de s'affranchir des défauts de ces dernières en évitant d'avoir à connaître le fonctionnement exact du chiffrement.

Cryptanalyse modulo n

- La *cryptanalyse modulo n* a été introduite en 1999 par Bruce Schneier, David Wagner, et John Kelsey.
- Elle consiste à exploiter les différences de fonctionnement (selon une congruence variable) des algorithmes qui utilisent des rotations binaires.

Attaques par canaux auxiliaires

- Les *attaques par canaux auxiliaires* font partie d'une vaste famille de techniques cryptanalytiques qui exploitent des propriétés inattendues d'un algorithme de cryptographie lors de son implémentation logicielle ou matérielle.
- En effet, une sécurité *mathématique* ne garantit pas forcément une sécurité lors de l'utilisation *en pratique*.
- Les attaques portent sur différents paramètres : le temps, le bruit, la consommation électrique, etc.

Compromis temps/mémoire

- Le concept de *compromis temps/mémoire* a été introduit en cryptanalyse par Martin Hellman en 1980.
- Il a été amélioré en 1993 par Philippe Oechslin avec le concept de table arc-en-ciel.
- Il s'agit d'un compromis entre une attaque par force brute et l'utilisation de dictionnaires.
- C'est ce qui a permis par exemple d'attaquer les mots de passe de sessions Windows.

Attaques sur les systèmes asymétriques

- Casser un chiffrement assuré par de la cryptographie asymétrique nécessite d'autres approches.
  - Dans le cas de RSA, c'est la difficulté de la factorisation qui assure la résistance du chiffrement.
  - Pour ElGamal, c'est le problème du logarithme discret qui est employé.
- Toutefois, certaines failles peuvent apparaître selon l'utilisation que l'on fait de ces algorithmes.
  - RSA est vulnérable si des exposants de faible magnitude sont utilisés.
  - Sous des conditions particulières, un surchiffrement avec RSA peut être attaqué.

Autres propriétés analysées

- Clefs faibles.
- Biais statistiques.

Cryptanalyse pratique

- Les trois principales méthodes de cryptanalyse sont :
  - l'analyse fréquentielle,
  - la cryptanalyse linéaire,
  - la cryptanalyse différentielle.
- On a déjà mise en pratique l'analyse fréquentielle.
- Aujourd'hui, faisons une cryptanalyse différentielle.

Cryptanalyse différentielle de TAKToy

- Nous allons cryptanalyser un cryptosystème jouet, TAKToy.
- TAKToy est suffisamment petit pour comprendre tout ce qui se passe et faire la cryptanalyse entièrement.
- TAKToy est un RSP à deux tours avec une clef de 8 bits et un bloc de 4 bits.
- La clef est coupée en deux, une partie pour chaque tour qui est ajoutée (avec un `xor`) à l'état avant que celui ci soit passé dans une boîte S.
- La boîte S est la suivante :

n	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(n)$	3	e	1	a	4	9	5	6	8	b	f	2	d	c	0	7

→ C'est parti...

Crédits

- Merci à Jon King et aux contributeurices Wikipédia.