Cryptologie : Introduction

Matthieu Amiguet

2006 – 2007

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

Qu'est-ce que la cryptologie ?

- Crypto – logie, étymologiquement :
  - "science du secret"
- La cryptologie englobe deux disciplines :
  - La cryptographie
    - "art de l'écriture secrète"
  - La cryptanalyse
    - "analyse du secret"
    - En clair : étude des moyens permettant de casser les efforts des cryptographes !

Cryptographie

- Discipline s'intéressant aux moyens de coder un message sous une forme difficilement décodable sans posséder un *secret*
- Usages traditionnels :
  - La guerre. . .
  - . . . et l'amour !
    Dans le Kama-sutra, l'art de l'écriture secrète est un des 64 arts devant être maîtrisés par les femmes !
- Plus récemment
  - Commerce, finance, . . .
  - Sécurité informatique
  - Usage privé.

Cryptanalyse

- Discipline s'intéressant aux moyens de décoder de l'information cryptée *sans connaître le secret*
- Usages
  - Utilisation militaire
  - Surveillance (notamment étatique)
  - Utilisation frauduleuse (notamment dans le domaine de la finance)
  - Amélioration des techniques cryptographiques.

Terminologie

- À strictement parler, le français utilise les termes
  - *chiffrer* et *déchiffrer* pour les opérations de codeage/décodage de la cryptographie
  - *chiffre* pour le système utilisé.
- Pour éviter les ambiguïtés, utiliserons aussi les termes *crypter*/*décrypter*/*cryptage*
- En anglais :
  - *encryption*/*decryption*
  - *cipher*

Disciplines connexes

- Théorie des codes
  - Notions en commun, mais buts distincts !
- Théorie de l'information
- Algorithmique
- Théorie des nombres
- Stéganographie
  - "Art de cacher des messages"
  - La difficulté est de *réaliser* qu'un message est présent, pas de le lire !
  - La cryptologie, au contraire, ne cache pas le message, elle le rend illisible !

L'importance du contexte

- L'applicabilité et surtout la sécurité d'une technique cryptographique dépendent énormément de son contexte
  - Qui est l'adversaire ?
  - Quel est son niveau de formation, de connaissances ?
  - Quelle est sa "puissance de calcul" ?
  - De combien de temps dispose-t-il ?
  - . . .
- Dans ce contexte, l'évolution technique des ces dernières décennies a bouleversé le domaine !

Les débuts

- On pourrait dire que la cryptographie a commencé avec l'écriture : le simple fait d'écrire réservait le contenu aux initiés
  - Si nécessaire, on pouvait inventer ou modifier un alphabet pour réduire encore le nombre de personnes susceptibles de comprendre
- Plus tard (~500 av. J.-C.), on a commencé à mettre en place
  - des substitutions de lettres
  - des permutations de leur ordre
- Ces deux principes ont formé le coeur de la cryptographie pendant près de deux millénaires
  - Avec une mécanisation croissante

Le tournant

- Dans la première moitié du XXe siècle sont mises au point les machines à tambour
  - Parmi les machines de cryptologie mécanique les plus complexes
  - La plus célèbre sera ENIGMA, utilisée par l'armée allemande durant la seconde guerre mondiale
- Les alliés vont mettre au point des machines mécaniques, puis électroniques de plus en plus sophistiquées pour tenter de casser ce code
- L'avènement de l'informatique va définitivement changer le paysage de la cryptographie, rendant obsolète pratiquement tout ce qui a été fait jusque là. . .

Pendant ce temps. . .

- En 1918, Vernam invente un algorithme de cryptage
  - incassable (quelle que soit la puissance de calcul de l'adversaire)
  - inutilisable (sauf dans des circonstances très particulières)
- En 1976, Diffie et Hellman lancent l'idée d'un cryptage à clé publique
  - Le secret n'est plus nécessaire pour chiffrer le message, mais seulement pour le déchiffrer

Aujourd'hui. . .

- Presque tout ce qui a été fait avant l'apparition de l'informatique est dépassé
  - Les grands principes subsistent tout de même. . .
- La cryptologie moderne requiert systématiquement un ordinateur
  - Aussi bien pour le cryptage, de décryptage que pour la cryptanalyse. . .
- La cryptographie se divise en deux grands domaines
  - La cryptographie à clé privée
  - La cryptographie à clé publique

Aujourd'hui. . . (suite)

- Avec le développement des télécommunications, les domaines d'application de la cryptologie se sont beaucoup étendus
  - Besoins accrus de *confidentialité*
- Mais la perte d'importance du support papier au profit de supports électroniques a fait aparaître d'autre utilités au techniques cryptographiques :
  - Intégrité des données
  - Authentification
  - Non-répudiation

Et demain ?

- La technique continue d'avancer à grand pas, ce qui oblige la cryptologie à se renouveler constamment
- Depuis les années 1970, l'évolution se fait tout de même dans une certaine continuité...
- Prochaine grande cassure : le calcul quantique ?
  - S'il est mis au point, ce sera une remise en question aussi fondamentale que l'avènement de l'informatique "sur silicone"...
  - Mais des techniques de *cryptographie quantique* voient déjà le jour...

Fonctions

Définition

Une *fonction* f est définie par deux ensembles X et Y et une règle qui assigne à chaque élément x de X un élément $f(x)$ de Y.

- Exemples :
  - $X = Y = \mathbb{N}, f(x) = x + 1$
  - $X = Y = \mathbb{R}, f(x) = sin(x)$
  - $Y =$"fichiers informatiques", $X =$"groupes de fichiers informatiques", $f(x) = \text{zip}(x) =$"L'archive zip contenant le groupe de fichiers x"

Images

- $f(x)$ est appelé l'*image* de x. x est une *préimage* de $f(x)$
  - $X = Y = \mathbb{R}, f(x) = \cos(x)$
    - L'image de 0 est 1
    - 0 est *une* préimage de 1 (mais 2π aussi !)
- L'ensemble
$$f(X) = \{f(x) \in Y | x \in X\}$$
est appelé l'*image* de f.
  - $X = Y = \mathbb{N}, f(x) = x + 1 : f(X) = Y = \mathbb{N}$
  - $X = Y = \mathbb{R}, f(x) = sin(x) : f(X) = [0, 1]$
  - l'image de la fonction "zip" est l'ensemble de tous les fichiers zip possibles

Propriétés (mathématiques) des fonctions

Soit $f : X \longrightarrow Y$ une fonction.

- f est *injective* ssi $x \neq y \Rightarrow f(x) \neq f(y)$
  - Pas de perte d'information ("lossless")
- f est *surjective* ssi $f(X) = Y$
  - Tout point de Y est une "sortie" potentielle
- f est *bijective* ssi elle est injective et surjective
  - La fonction cos : $\mathbb{R} \longrightarrow \mathbb{R}$ n'est ni injective ni surjective
  - La fonction cos : $\mathbb{R} \longrightarrow [0, 1]$ n'est pas injective mais elle est surjective
  - La fonction zip (telle que définie ci-dessus) est injective mais elle n'est pas surjective
  - La fonction $f(x) = x + 1$ définie ci-dessus est bijective.

Inverse

- Si f est bijective, elle admet une *fonction inverse* $f^{-1} : Y \to X$ vérifiant
  - $f(f^{-1}(y)) = y$
  - $f^{-1}(f(x)) = x$
- Exemples :
  - $X = Y = \mathbb{N}, f(x) = x + 1 : f^{-1}(y) = y - 1$
  - La fonction zip possède un inverse… et pourtant nous avons vu qu'elle n'est pas surjective ! Comment est-ce possible ?

Permutation

- Une bijection $f : X \to X$ est une *permutation*.
- Exemples :
  - Mélanger des cartes
  - $X = \mathbb{N}, f(x) = x + 1$
  - $X = [0, \ldots, 26], f(x) = (x + 13) \bmod 26$
- Pour les ensembles finis, on peut représenter une permutation par une table :

$$\begin{pmatrix} A & B & C & D & \ldots & X & Y & Z \\ W & R & A & N & \ldots & Z & K & Y \end{pmatrix}$$

Propriétés (cryptographiques) des fonctions – 1

- $f : X \to Y$ est à *sens unique* si
  - $f(x)$ est "relativement facile" à calculer pour tous les x, mais
  - pour presque tous les $y \in f(X)$, il est "calculatoirement impossible" de trouver x avec $f(x) = y$.

Exemple

- $X = Y = \{1, 2, 3, \ldots, 16\}$
- $f(x) =$ le reste de 3^x divisé par 17

x	1	2	3	4	5	6	7	8	9	10	…
$f(x)$	3	9	10	13	5	15	11	16	14	8	

- Préimage de 7 ?

Propriétés (cryptographiques) des fonctions – 2

Soit f une fonction à sens unique,

- f est dite *à brèche secrète* (*trapdoor one-way function*) si la donnée d'une information supplémentaire (la *brèche*) rend la préimage calculable.

Exemple

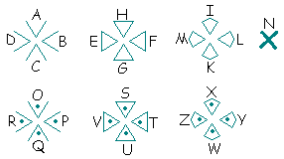
Plus tard !

Cryptage (ou chiffrement)

- Pour définir un cryptage, on a généralement besoin de
  - Un alphabet pour les messages en clair : $\mathcal{A}_M$
    qui définit un espace des messages clairs $M = \mathcal{A}_M^*$
  - Un alphabet pour les messages cryptés : $\mathcal{A}_C$
    qui définit un espace des messages cryptés $C = \mathcal{A}_C^*$
    - Parfois $\mathcal{A}_M = \mathcal{A}_C$, parfois pas
- Un cryptage est alors une bijection $E : M \to C$
  - C'est une bijection parce qu'on veut pouvoir décrypter !

Un premier exemple

- Le code :



- Exemple :

Le problème de la clé

- L'exemple ci-dessus marche bien tant que personne n'a compris comment il marche... puis on peut le jeter !
- Nous verrons que ce n'est pas une bonne idée de tabler sur le secret de l'algorithme... Ça ne marche pas !
- C'est ainsi que prend forme l'idée suivante : un cryptage doit être *paramétrable* par une valeur appelée *clé*.
- Seule cette valeur doit rester secrète.

Un exemple avec clé

- Alphabet $\mathcal{A} = \mathcal{A}_M = \mathcal{A}_C = \{0, 1, 2, ..., 255\}$
- Espace des messages $M = C = \mathcal{A}^*$
- Espace des clés $\mathcal{K} = \{0, 1, 2, ..., 255\}$
- Pour $k \in \mathcal{K}$, $E_k = D_k : \mathcal{A}^* \to \mathcal{A}^*$

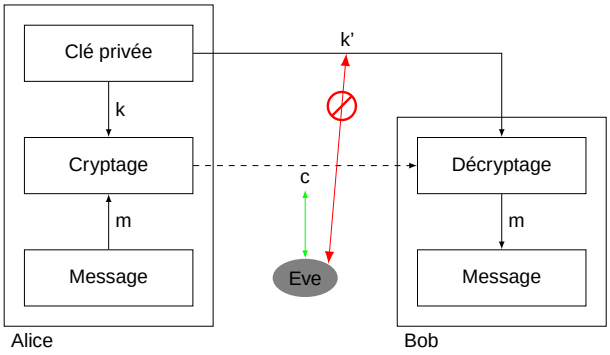
$$a_1 a_2 \dots a_n \mapsto (a_1 \oplus k)(a_2 \oplus k) \dots (a_n \oplus k)$$

  où $\oplus$ désigne le ou-exclusif bit par bit.
- Exercice : comment décrypter ?
- Exercice (bis) : comment casser ce cryptage ?

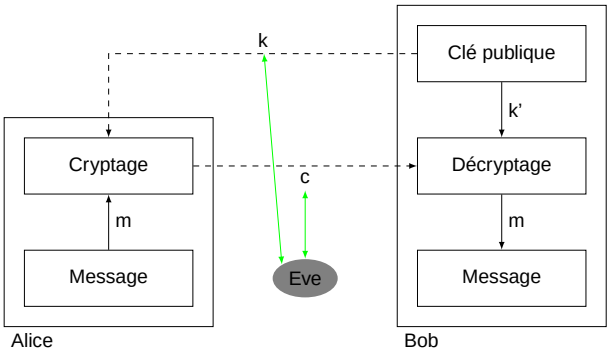
Clés secrète et publiques

- On peut reformuler la définition du cryptage en tenant compte du besoin de clé :
  - Un cryptage est une bijection $E_K : M \rightarrow C$ paramétrée par une clé K
  - Pour décrypter, il faudra connaître la clé de décryptage K' et la fonction paramétrée $D_{K'}$ telle que $D_{K'}(E_K(m)) = m, \forall m \in M$
- Parfois $K = K'$, parfois pas
- Si K' se retrouve facilement à partir de K, on parle de cryptage à *clé privée*
- S'il est "calculatoirement impossible" de retrouver K' à partir de K, le cryptage est dit à *clé publique*.

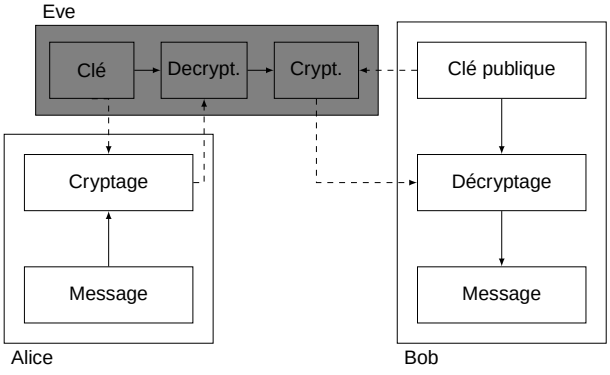
Clé privée – fonctionnement général



Clé publique – fonctionnement général



Mais si Eve est active. . .



Talons d'Achille

- L'attaque d'Eve est connue sous le nom de "Man in the Middle"
- Elle met en évidence le talon d'Achille de la cryptologie a clé publique : l'*authentification*
- La cryptologie à clé privée a aussi son talon d'Achille : la *transmission des clés*.

Taille des clés : quelle importance

- Vrai ou faux : "Plus la clé est grande, plus le cryptage sera difficile à casser"

⇝ Vrai, mais

  - La recherche exhaustive des clés n'est pas toujours la manière la plus efficace de casser un cryptage
  - Ce n'est pas non plus une méthode universelle
  - Entre une clé qui prendrait 12 millions d'années à casser et une qui prendrait 100 milliards d'années, peut-on parler d'amélioration ?

Qu'est-ce que "grand"

- Pour tenter de répondre à cette question, nous allons établir un tableau donnant, pour un chiffre codé sur n bits
  - Combien de chiffres sont nécessaires pour l'écrire en base 10
  - Le temps nécessaire pour parcourir toutes les valeurs sur un PC actuel avec un bon algorithme (2mio clés/sec)
  - Le temps nécessaire pour parcourir toutes les valeurs sur un cluster distribué de type distributed.net (100 mia clés/sec)
  - Un exemple (parfois très approximatif !) de l'ordre de grandeur.

Ordres de gandeur

#bits	#chiffres	PC	cluster	exemple
16	4	0,01s		
32	9	8,3 min	0,01s	Âge du système solaire (années)
56	16	158 ans	2,7h	Clés DES
64	19	158'400 ans	3,17 ans	
80	24	1 AU	316'880 ans	#grains de sable au monde
96	28	10'562 AU	0,21 AU	#atomes dans le corps humain
128	38		2,1 mia AU	#adresses IPv6
				#empreintes MD5

Ordres de gandeur (suite)

#bits	#chiffres	PC	cluster	exemple
240	72		2e43 AU	#nombres premiers de 75 chiffres
256	77		3e58 AU	clés AES (max)
				#électrons dans l'univers
512	154		2e125 AU	
832	250		2e221 AU	#parties d'échecs possibles
1024	308		2e279 AU	Clés RSA
			...	
2372	700			#parties de go possibles

Plan du cours

- Introduction
- Cryptographie historique
- Cryptographie moderne à clé secrète
- Cryptographie à clé publique
- Éléments de cryptanalyse
- Générateurs pseudo-aléatoires
- Signatures numériques