ENS Cachan, antenne de Ker Lann
Agrégation de Mathématiques 2008
Préparation à l'oral de modélisation
Option C : Calcul Formel

Rudiments de Cryptologie

Richard Leroy

1 Introduction

La cryptologie, étymologiquement la science du secret, ne peut être vraiment considérée comme une science que depuis peu de temps. Cette science englobe la cryptographie et la cryptanalyse.

La cryptographie s'attache à protéger des messages (assurant confidentialité et/ou authenticité) en s'aidant souvent de secrets ou clés. Elle est utilisée depuis l'Antiquité, mais certaines de ses méthodes les plus importantes, comme la cryptographie asymétrique, n'ont que quelques dizaines d'années d'existence.

La cryptanalyse s'oppose, en quelque sorte, à la cryptographie. En effet, si déchiffrer consiste à retrouver le message clair au moyen d'une clé, cryptanalyser c'est tenter de se passer de cette dernière.

1.1 Cryptographie

La cryptographie se scinde en deux parties nettement différenciées :

◦ d'une part la cryptographie à clef secrète, encore appelée symétrique ou bien classique,

◦ d'autre part la cryptographie à clef publique, dite également asymétrique ou moderne.

La première est la plus ancienne, on peut la faire remonter à l'Égypte de l'an 2000 av. J.-C. en passant par Jules César ; la seconde remonte à l'article de W. Diffie et M. Hellman, *New directions in cryptography,* daté de 1976.

Toutes deux visent à assurer la confidentialité de l'information, mais la cryptographie à clef secrète nécessite au préalable la mise en commun entre les destinataires d'une certaine information : la clef (symétrique), nécessaire au chiffrement ainsi qu'au déchiffrement des messages. Dans le cadre de la cryptographie à clef publique, ce n'est plus nécessaire. En effet, les clefs sont alors différentes, ne peuvent se déduire l'une de l'autre, et servent à faire des opérations opposées, d'où l'asymétrie entre les opérations de chiffrement et de déchiffrement.

Bien que beaucoup plus récente et malgré d'énormes avantages — signature numérique, échange de clefs... — la cryptographie à clef publique ne remplace

pas totalement celle à clef secrète, qui pour des raisons de vitesse de chiffrement et parfois de simplicité reste présente. À ce titre, signalons la date du dernier standard américain en la matière, l'AES : décembre 2001, ce qui prouve la vitalité encore actuelle de la cryptographie symétrique.

1.2 Cryptanalyse

Le pendant de cette confidentialité se trouve dans la cryptanalyse. Évidemment, depuis l'existence de ces codes secrets, on a cherché à les casser, à comprendre les messages chiffrés bien que l'on n'en soit pas le destinataire légitime, autrement dit décrypter.

2 Cryptographie à clef publique

La cryptographie à clef publique (ou asymétrique) s'est peu à peu immiscée dans la vie quotidienne : sécurité des cartes à puces, du commerce électronique...

Le principe de la cryptographie à cléf publique est d'utiliser les limites de la technologie. En bref, la fonction de chiffrement est connue de tous, facile à programmer, mais l'inversion de cette fonction est impossible en temps raisonnable avec les technologies informatiques actuelles, sauf pour le destinataire qui a des informations privées supplémentaires sur cette fonction. La construction d'un tel système peut être basée sur l'existence de fonctions mathématiques dites à sens unique et à trappe.

◦ *Avantages* :

- le fait que le code soit connu n'est pas un danger (intérêt militaire ou lorsqu'il y a des millions d'utilisateurs)
- il est possible de "signer" un message de telle sorte que l'on soit sûr de sa provenance

◦ *Inconvénients* :

- lenteur des algorithmes

Le paragraphe suivant est consacré au cryptosystème à clef publique le plus répandu à l'heure actuelle : le protocole RSA.

2.1 Protocole RSA (Rivest, Shamir, Adleman)

2.1.1 Principe de la méthode

Algorithme 1 ***[Génération des clefs]***
- Générer deux (grands !) nombres premiers p et q et calculer $n = pq$ et $\varphi = \varphi(n) = (p-1)(q-1)$
- Trouver un entier e tel que $1 < e < \varphi$ et $\gcd(e, \varphi) = 1$

- *Calculer* $d \equiv e^{-1} \bmod \varphi$
La clef publique diffusée est (n, e) *et la clef privée est* d.

Algorithme 2 ***[Cryptage]***
Soit $m \in [0, n-1]$ *le message clair à transmettre.*
Le message crypté transmis est $c \equiv m^e \bmod n$.

Algorithme 3 ***[Décryptage]***
Soit $c \in [0, n-1]$ *le message crypté.*
Alors $m = c^d \bmod n$.

2.1.2 Correction de l'algorithme

Proposition 4 *Soient* p *et* q *deux nombres premiers distincts, et* $n = pq$. *Soient* d *et* e *deux entiers tels que*

$$de \equiv 1 \bmod (p-1)(q-1).$$

Alors on a :

$$\forall m \in \mathbb{Z}, m^{de} \equiv m \bmod n.$$

Preuve. Soit $m \in \mathbb{Z}$, montrons que $m^{de} \equiv m \bmod n$.
∘ Premier cas : on suppose m premier avec p.
Par conséquent, $m^{de} = m^{1+k\varphi(n)} = m^{1+k(p-1)(q-1)} = m\left(m^{p-1}\right)^{k(q-1)} \equiv m \bmod p$.
∘ Le cas où m n'est pas premier avec p est trivial (mais souvent oublié lors d'une présentation à l'oral). ■

2.1.3 Mise en oeuvre du protocole RSA

∘ **Génération de grands nombres premiers** :

Pour construire un nombre premier de 100 chiffres, on génère au hasard des entiers impairs de 100 chiffres et on leur applique un test de primalité. La pratique actuelle est d'utiliser un test probabiliste (Miller-Rabin). A noter qu'un algorithme de test de primalité en temps polynomial a été très récemment proposé (2002). Si l'entier impair n'est pas premier, on l'incrémente de 2 et on reteste.

∘ **Génération de** e **:**

On tire au hasard jusqu'à ce que $\gcd(e, \varphi) = 1$.

∘ **Calcul de** d **:**

On utilise ici l'algorithme d'Euclide étendu.

∘ **Algorithmes de cryptage et de décryptage :**

On utilise un algorithme dichotomique d'exponentiation modulaire.

2.1.4 Utilisations du protocole RSA

Outre le chiffrement que l'on vient d'étudier, le protocole RSA peut être utilisé pour résoudre des problèmes de signature et d'authentification (cartes bancaires).

2.1.5 Cryptanalyse et attaques

Lemme 5 *La connaissance de $\varphi(n)$ est équivalente à la connaissance de la factorisation de n.*

Preuve. "$\Leftarrow$" trivial
"$\Rightarrow$" En substituant n/p à q dans la relation $\varphi(n) = (p-1)(q-1)$, on obtient l'équation du second degré

$$p^2 - (n + 1 - \varphi(n))p + n = 0$$

dont la résolution est facile. ■

Remarque 6 *Cette question faisait l'objet d'une question d'un texte proposé à l'oral de modélisation en 2006.*

Lemme 7 *La connaissance de d est équivalente à la connaissance de la factorisation de n.*
Plus précisemment, il existe un algorithme probabiliste qui calcule la factorisation $n = pq$ à partir des valeurs n, e et d.

Preuve. • *On montre d'abord que la connaissance d'une racine carrée non triviale de 1 modulo n permet le calcul en temps polynomial de la factorisation de n.*
On considère l'équation $x^2 \equiv 1 \bmod n$. Elle est équivalente, d'après le théorème des restes chinois, au système :

$$\begin{cases} x^2 \equiv 1 \mod p \\ x^2 \equiv 1 \mod q \end{cases}$$

lui-même équivalent au système :

$$\begin{cases} x \equiv \pm 1 \mod p \\ x \equiv \pm 1 \mod q \end{cases}$$

L'équation initiale possède donc 4 solutions, dont les deux triviales $x = \pm 1 \bmod n$. Soit x est une solution non triviale. Alors $n|(x+1)(x-1)$ mais $n \nmid x+1$ et $n \nmid x-1$. Par conséquent, $\gcd(x+1, n) = p$ ou q et $\gcd(x+1, n) = q$ ou p.

• *Décrivons maintenant un algorithme probabiliste de calcul d'une racine carrée non triviale de 1 modulo n.*

On tire au hasard un entier $\omega \in]1, n-1]$.
Si $\gcd(\omega, n) > 1$, *on arrête : on a trouvé un facteur premier de* n.
Supposons donc que $\gcd(\omega, n) = 1$. *On peut écrire*

$$ed - 1 = 2^s r$$

où $s \geq 1$ *et* r *est impair. On a alors :*

$$\omega^{2^s r} = \omega^{ed-1} = \omega^{k\varphi(n)} \equiv 1 \bmod n.$$

Soit s_0 *le plus petit entier tel que* $\omega^{2^{s_0} r} \equiv 1 \bmod n$.
Dans le cas où $s_0 > 0$, *on pose* $v_0 = \omega^{2^{s_0-1} r}$.
Alors $v_0 \neq 1 \bmod n$ *et* $v_0^2 = 1 \bmod n$. *Si* $v_0 \neq -1 \bmod n$, *alors on a trouvé une racine carrée non triviale de l'unité modulo* n.
D'après le théorème de Rabin, la probabilité d'être dans le cas favorable où $s_0 > 0$ *et* $v_0 \neq 1 \bmod n$ *est supérieure à* $3/4$.
Si l'algorithme échoue, alors on recommence avec un autre ω. *L'espérance du nombre d'itérations est alors inférieure à* $4/3$. ■

Quelques faiblesses Si l'on ne prend pas garde à son utilisation, le protocole RSA peut présenter des faiblesses :

• Envoi d'un même message à deux utilisateurs en employant le même n. On utlise donc les deux clés publiques (n, e_1) et (n, e_2), et les clés secrètes d_1 et d_2. Si e_1 et e_2 sont premiers entre eux, l'algorithme d'Euclide étendu permet de trouver deux entiers u et v tels que $ue_1 + ve_2 = 1$. On peut dès lors retrouver le message m sans connaître les clés secrètes, car en effet :

$$(m^{e_1})^u (m^{e_2})^v \equiv m \bmod n.$$

• Envoi d'un même message à plusieurs utilisateurs en utilisant le même e. Les messages envoyés sont de la forme $m_i \equiv m^e \bmod n_i$, pour $i = 1, \cdots, k$. Si les n_i sont premiers entre eux, alors le théorème des restes chinois permet de calculer $m^e \bmod \prod n_i$. Si $e \leq k$, ce n'est autre que m^e lui-même, et on obtient alors m très facilement.

Attaque par factorisation La sécurité du cryptosystème RSA est dûe à la difficulté à factoriser un grand nombre entier n. On ne connaît à l'heure actuelle aucun algorithme de factorisation suffisamment efficace pour pouvoir casser en pratique le système RSA.

L'algorithme naïf des divisions successives a une complexité en $O(\sqrt{n})$. Il consiste à diviser successivement n par tous les nombres compris entre 2 et $\sqrt{n}$. C'est en fait une utilisation du crible d'Eratosthène, dont on pourra donner à l'oral une visualisation graphique.

D'autres algorithmes de factorisation existent. On peut notamment utiliser des algorithmes probabilistes, dont la complexité est moindre.

Attaque par fractions continues (Wiener 1989) On commence par quelques rappels sur les fractions continues.

On note :

$$[a_0, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots \quad \ddots \quad a_{n-2} + \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}}$$

Pour $x \in \mathbb{R}$, on pose $a_0 = \lfloor x \rfloor$, $x_0 = \dfrac{1}{x - a_0}$, puis $a_i = \lfloor x_{i-1} \rfloor$ et $x_i = \dfrac{1}{x_{i-1} - a_i}$, pour $i \in \mathbb{N}^*$, et tant que ces expressions ont un sens.

On a donc $\forall n \in \mathbb{N}, x = \big[a_0, \ldots, a_{n-1}, x_n\big]$. La suite (finie ou non) des a_i s'appelle le développement en fractions continues de x.

Les nombres rationnels $p_n/q_n = [a_0, \ldots, a_n]$ sont appelés les réduites de x.

On a alors le résultat suivant :

Théorème 8 *Soient $x \in \mathbb{R}$ et $p/q \in \mathbb{Q}$ tels que $|x - p/q| < \dfrac{1}{2q^2}$.*
Alors p/q est une réduite de x.

Ce théorème permet donc l'attaque suivante :

On suppose que $q < p < 2q$ (p et q ont le même nombre de bits).

On peut écrire $ed - 1 = k\varphi(n)$ pour un certain $k \in \mathbb{Z}$. Puisque $0 < e < \varphi(n)$, on a $0 < k < d$. De plus, l'égalité

$$\varphi(n) = (p-1)(q-1) = n - p - q + 1$$

permet d'écrire successivement :

$$ed = 1 + k(n - p - q + 1)$$

$$\frac{e}{n} = \frac{k}{d} + \frac{1 + k - k(p+q)}{dn}$$

$$\left|\frac{e}{n} - \frac{k}{d}\right| = \frac{k(p+q) - k - 1}{dn} < \frac{k(p+q)}{dn} < \frac{3kq}{dn} < \frac{3k}{d\sqrt{n}} < \frac{3}{\sqrt{n}}$$

Par conséquent, si $d < \dfrac{n^{1/4}}{\sqrt{6}}$, on obtient :

$$\left|\frac{e}{n} - \frac{k}{d}\right| < \frac{1}{2d^2}.$$

Le théorème précédent permet donc de conclure que k/d est une réduite de e/n.

L'attaquant procède donc ainsi : puisque e et n sont publics, il peut développer e/n en fraction continue. Il teste alors toutes les réduites successives. Il s'arrête dès que le dénominateur permet de factoriser n.

2.1.6 Attaque par le milieu

Il s'agit cette fois de s'attaquer à la communication des clés. Voici son principe :

- Alice demande la clé publique de Bob.
- Bob envoie e à Alice.
- Le message est intercepté par Charles, qui envoie sa clé e'.
- Alice crypte alors avec e'.
- Charles intercepte le message c' d'Alice, et décrypte avec sa clé secrète.
- Charles envoie le message d'Alice en le cryptant avec e.

Cette attaque montre bien qu'il est nécessaire pour les interlocuteurs de s'identifier. On peut à cet effet utiliser également le protocole RSA, comme vu plus haut. Mais une attaque par le milieu est encore possible sur l'identification. En pratique, des protocoles dits de preuve sans transferts de connaissance permettent de se prévenir de ce genre d'attaque.