Histoire des codes secrets

Razvan Barbulescu

parimaths, avril 2016

La cryptologie tente de résoudre plusieurs problèmes dont le plus ancien est la confidentialité : on chiffre les messages de façon que seul le destinataire puisse les lire. La première solution est que toute paire de personnes utilise une langue différente ou une méthode de chiffrement différente, mais cela demande un effort trop important. Une alternative, connue sous le nom de principe de Kerckhoffs, est la création d'une méthode de chiffrement connue par tout le monde et d'utiliser une clé différente pour chaque paire de personnes qui communiquent. On a de nombreux exemples de telles méthodes de chiffrement : le scytale des Spartiates (404 av. J.-C.), César (Ier siècle av. J.-C.), substitution (jusqu'au XII siècle ap. J.-C.), Enigma (XXe siècle) et masque jetable (XXe siècle).

En 1976 Withfield Diffie et Martin Hellman ont proposé un changement majeur : créer des méthodes de chiffrement qui rassemble à des cadenas : tout le monde peut chiffrer mais personne ne peut déchiffrer à moins d'avoir la clé. C'est la cryptologie asymétrique, alors que les méthodes précédentes s'appellent symétriques. En 2016 Diffie et Hellman ont reçu le prix Turing pour cette idée, qui est utilisée sur internet. En effet, beaucoup de sites internet utilisent le protocole https où les messages sont chiffrés : après un échange de clé à la Diffie et Hellman les communications sont chiffrés par une méthode symétrique.

La cryptologie asymétrique requiert des fonctions mathématiques particulières. Plusieurs telles fonctions ont été proposées :

1. Merkle-Hellman (méthode du sac à dos);
2. McEliece (apparenté aux codes correcteurs);
3. Rivest Shamir Adleman (RSA).

La méthode du sac a dos a été cassée et ne peut plus être utilisée, mais elle constitue un exemple intéressant, que nous traitons dans un exercice.

Exercice : Alice utilise le système de chiffrement de Merkle-Hellman qui permet à tout le monde de lui envoyer des messages secrets tout en garantissant qu'elle est la seule à pouvoir les lire.

Pour cela Alice prépare des données comme suit :

1. elle choisit un ensemble d'entiers $a_0, a_1, \ldots, a_n$ tels que chaque terme domine la somme des précédents, ou plus formellement pour tout $k \leq n$, $a_0 + \cdots + a_{k-1} < a_k$.
2. elle choisit un entier N supérieur à $a_0 + \cdots + a_n$.
3. elle prend un entier B au hasard dans l'intervalle $[1, N-1]$ tel que $\text{pgcd}(B, N) = 1$ et calcule $b_0 := Ba_0 \bmod N$, $b_1 = Ba_1 \bmod N$, ..., $b_n = Ba_n \bmod N$, où $\ll \mod N \gg$ désigne le reste à la division par N.

Pour chiffrer un message $m = m_0, \ldots, m_n$ de n chiffres binaires, 0 ou 1, Bob calcule $e = (m_0b_0 + m_1 + \cdots + m_nb_n) \bmod N$.

1. Voyons comment Alice fait pour déchiffrer. Prenons l'exemple où Alice choisit $n = 8$ entiers : $a_0 = 1$, $a_1 = 3$, $a_2 = 6$, $a_3 = 13$, $a_4 = 24$, $a_5 = 50$, $a_6 = 98$ et $a_7 = 200$. Elle choisit $N = 400 > 395 = a_0 + \cdots + a_7$ et $B = 267$, puis calcule $b_0 = 267$, $b_1 = 1$, $b_2 = 202$, $b_3 = 471$, $b_4 = 8$, $b_5 = 350$, $b_6 = 166$ et $b_7 = 200$. Bob a chiffré un message $m = m_0m_1m_2m_3m_4m_5m_6m_7$ et a obtenu $e = (267m_0 + m_1 + 2m_2 + 271m_3 + 8m_4 + 150m_5 + 166m_6 + 200m_7) \bmod 400 = 233$.

   Calculer deux entiers u et v tels que $267u + 400v = 1$ et remarquer que pour tout entier $a < 400$, $(u \cdot 267 \cdot a) \bmod 400 = a$. Calculer $e' = ue \bmod 400$. Utiliser e' pour déchiffrer le message de Bob.

2. Voyons aussi le point de vu du cryptanalyste. Prenons un nouveaux exemple où Alice choisit autres $n = 8$ entiers a_0, a_1, ..., a_7 et elle les garde secrets. Elle choisit un entier N supérieur à $a_0 + \cdots + a_7$ et annonce $N = 1000$. Ensuite elle choisit B au hasard et le garde secret, puis elle calcule $b_0 = 893$, $b_1 = 155$, $b_2 = 679$, $b_3 = 251$, $b_4 = 181$, $b_5 = 41$, $b_6 = 901$ et $b_7 = 481$ et les rend publiques.

   Bob chiffre le message $m = m_0m_1m_2m_3m_4m_5m_6m_7$ et obtient $e = (893m_0 + 155m_1 + 679m_2 + 251m_3 + 181m_4 + 41m_5 + 901m_6 + 481m_7) \bmod 1000 = 990$. Éve intercepte le message chiffré e et veut retrouver m sans la clé secrète B. Elle a calculé $s = 893m_0 + 155m_1 + 679m_2 + 251m_3$ pour toutes les possibilités des chiffres binaires m_0, m_1, m_2 and m_3 et les a rangé dans le tableau ci-dessous. Aider-la à retrouver m.

m_1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
m_2	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
m_3	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
m_4	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
s	0	893	155	48	679	572	834	727	251	144	406	299	930	823	85	978

Solution : i) Pour calculer u et v on calcule :

$$\begin{aligned} 0 \cdot 267 + 1 \cdot 400 &= 400 \quad (1) \\ 1 \cdot 267 + 0 \cdot 400 &= 267 \quad (2) \end{aligned}$$

On fait la division : $400 = 267 \cdot 1 + 133$ et déduit l'équation

$$-1 \cdot 267 + 1 \cdot 400 \quad = \quad 133 \quad (3) := (1) - (2)$$

On fait de nouveau la division des deux dernier membres droits : $267 = 2 \cdot 133 + 1$.

$$3 \cdot 267 - 2 \cdot 400 \quad = \quad 1 \quad (4) := (2) - 2 \cdot (3)$$

Ainsi on choisit $u = 3$ et $v = -2$.

On suit les indications de l'exercice : $e' = u \cdot e \bmod N = 3 \cdot 233 \bmod 400 = 299$.

Comme $e \equiv Ba_0m_0 + \cdots Ba_7m_7 \mod N$ et $uB \equiv 1 \mod N$ on a

$$\begin{aligned} e' &\equiv uBa_0m_0 + \cdots uBa_7m_7 \mod N \\ e' &\equiv a_0m_0 + \cdots a_7m_7 \mod N. \end{aligned}$$

D'après le choix de N on a $N > a_0 + \cdots a_7$ donc $N > a_0m_0 + \cdots a_7m_7$. En connaissant le reste modulo N d'un entier inférieur à N on connait cet entier, donc

$$a_0m_0 + \cdots a_7m_7 = 299.$$

Pour retrouver m_0, ..., m_7 on doit retrouver un sous-ensemble S de a_0, ..., a_7 dont la somme est 299. Comme $a_0 + a_1 + \cdots a_6 < 299$, S doit contenir a_7. Il reste à retrouver un sous-ensemble de $\{a_0, \ldots, a_6\}$ de somme $299 - 200 = 99$. Comme $a_0 + \cdots + a_5 < 99$ le sous-ensemble S doit contenir a_6. Il reste à trouver un sous-ensemble de somme $99 - 98 = 1$. On finit en rajoutant $a_0 = 1$. Donc $S = \{a_0, a_6, a_7\}$ et alors

$$m = 10000011.$$

ii) Éve doit trouver $m_0, m_1, \ldots, m_7$ tels que $m_0b_0 + \cdots m_7b_7 = 990$. Cela revient à trouver s et t tels que s s'écrit comme $s = (m_0b_0 + m_1b_1 + m_2b_2 + m_3b_3) \bmod 1000$ et t s'écrit comme $t = (m_4b_4 + m_5b_5 + m_6b_6 + m_7b_7) \bmod 1000$ et on a

$$s = t.$$

On dresse un deuxième tableau :

m_4	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
m_5	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
m_6	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
m_7	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
t	990	809	949	768	89	908	48	967	509	328	468	287	608	467	567	386

Parmi les valeurs possible pour s et t on retrouve 48 qui est commune. On lit $m_0m_1m_2m_3$ dans le premier tableau et $m_4m_5m_6m_7$ dans le deuxième, donc

$$m = 11000110.$$