Cryptographie

Christophe Bourguignat Thomas Fajour Rémi Zara

10 février 1999

Aujourd'hui, la théorie des nombres, longtemps considérée comme la branche la plus abstraite et la moins utile des mathématiques, connaît un vaste essor à cause notamment de ses applications fondamentales dans le domaine de la cryptographie. La recherche actuelle s'appuie sur deux axes principaux :

- comment savoir en un temps raisonnable si un nombre donné est premier ou non.
- comment décomposer en facteurs premiers un nombre donné.

En ce qui concerne le premier problème, de nombreux progrès ont été accomplis depuis quelques années puisqu'il faut aujourd'hui 45 secondes à l'aide d'un ordinateur puissant pour savoir si un nombre de 100 chiffres est premier, 6 minutes pour un nombre de 200 chiffres et une semaine pour un nombre de 1000 chiffres.

Les méthodes utilisées, qui permettent de conclure quant à la question de la primalité d'un nombre, sont en revanche inefficaces en ce qui concerne la décomposition en facteurs premiers de ce même nombre. La solution du second problème, qui apporte évidemment une solution au premier problème, n'a donc pas encore été trouvée. Autrement dit, il est aujourd'hui pratiquement impossible de trouver la décomposition primaire d'un nombre non premier de 100 chiffres et plus.

Cet obstacle est à la base d'un système de cryptographie très performant (le système R.S.A. dont nous exposerons le principe plus bas) qui donne à la théorie des nombres une dimension pratique et concrète encore mésestimée il y a peu de temps.

La cryptographie

Depuis toujours, l'homme a eu besoin de transmettre à ses alliés des messages secrets censés ne pas pouvoir être compris par ses ennemis. Il a fait appel pour cela à une science, la cryptographie, qui lui permettait de coder des messages qui devenaient ainsi indéchiffrables par des personnes non averties. Même s'il est vrai que ces procédés répondaient au début avant tout à des besoins militaires, diplomatiques ou d'espionnage, la cryptographie est aujourd'hui largement utilisée dans d'autres domaine : sécurité des transactions bancaires, sécurité des fichiers ...

Donnons quelques exemples de procédés de cryptage.

Un des premiers procédés fut utilisé par Jules César. Il consistait à remplacer chaque lettre du message à coder par la lettre qui la suivait dans l'ordre naturel de l'alphabet ($BONJOUR$ devenait $CPOKPVS$). Mais il était peu fiable car il était facile pour l'ennemi de découvrir l'algorithme de chiffrement et de décrypter ainsi des messages entiers. Ainsi voit-on sur cet exemple que ce genre de système de cryptage nécessite de garder secret l'algorithme même de codage.

Puis est apparue en 1975 la cryptographie à «algorithme public». Si tout le monde connaît le principe de codage, seules les personnes concernées possèdent la clé de codage et de décodage nécessaire pour coder et décoder le messages. Cette idée fut à la base du système DES. Ce procédé de cryptage utilise le langage binaire et la table d'addition suivante :

+	0	1
0	0	1
1	1	0

Une clé binaire (prenons-la égale à 101 dans notre exemple) est maintenue secrète. Pour coder un message, disons 1100010, on opère de la façon suivante :

- on écrit en dessous l'un de l'autre le message en clair et la clé répétée :

$$\begin{array}{ccc} \text{message en clair} & \rightarrow & 110\ 001\ 0 \\ \text{clé répétée} & \rightarrow & 101\ 101\ 1 \end{array}$$

- On obtient le message crypté (ou cryptogramme) en effectuant l'addition : cela donne ici 0111001
- Le déchiffrage se fait toujours à l'aide de la clé secrète par opération inverse et sans risque d'ambiguïté (car $0, 1$ muni de l'addition définie ci-dessus est un groupe donc chaque élément a un unique opposé).

L'efficacité de ce système provient du fait que si la clé secrète est de très grande longueur, il est impossible pour l'ennemi de décrypter (même s'il connaît l'algorithme) en essayant toutes les clés tant les combinaisons sont nombreuses. En revanche, son gros défaut est l'extrême difficulté de garder secrète cette fameuse clé qui, si elle est découverte par l'ennemi, rend totalement inefficace le procédé. En effet, les transmissions de clés, souvent longues, sont délicates surtout au sein de réseaux étendus dans lesquels de nombreuses personnes doivent connaître la clé pour à la fois crypter et décrypter.

C'est justement pour parer cet inconvénient majeur qu'a été développé un nouveau procédé : la cryptographie à clés publiques. En voici le principe :

Soit f le procédé de codage à clés publiques et M le message à coder. $f(M)$ représente le message codé. Pour le décoder, il faut lui appliquer une fonction inverse f^{-1} telle que $f^{-1}(f(M)) = M$. Supposons que **A** veuille envoyer un message à **B**. **A** commence par se procurer la clé de chiffrement rendue publique par **B** qui l'a par exemple diffusée dans un annuaire. Muni de cette clé de chiffrement, **A** peut expliciter la fonction f qui est de complexité très réduite et ainsi coder facilement le message M en $f(M)$. Mais il est très difficile pour **A** (ou toute autre personne que **B**) d'expliciter f^{-1} car la clé de chiffrement ne suffit pas au déchiffrement ; autrement dit, même l'émetteur **A** ne peut pas déchiffrer les messages émis par d'autres correspondants avec la même clé de chiffrement publique.

Seul **B** possède une clé complémentaire ultra-secrète qui elle seule permet d'expliciter f^{-1} et donc de déchiffrer. Cette clé complémentaire est bien sûr très difficile à trouver par l'ennemi même en utilisant de gros moyens comme des ordinateurs.

Ainsi comprend-on l'intérêt de la cryptographie à clé publique : dans un vaste réseau, chaque utilisateur possède une clé publique personnelle qu'il publie et une clé complémentaire ultra-secrète qu'il n'a à aucun moment à communiquer à quiconque. Le risque de «fuite» est ainsi considérablement réduit car chacun veille à ce que sa propre clé complémentaire reste secrète.

Nous allons exposer un exemple de système de cryptographie à clé publique : le système *R.S.A*.

1 Le système R.S.A

On appellera **Marie** la personne qui désire recevoir un message codé, et **Patrick** la personne qui envoie le message.

1.1 Choix de la clé

Marie choisit deux grands entiers naturels premiers p et q (d'environ 100 chiffres chacun ou plus) et fait leur produit $n = pq$. Puis elle choisit un entier E tel que E soit premier avec $(p-1)(q-1)$. Enfin, elle publie dans un annuaire par exemple l'information $(R.S.A,\ n,\ E)$.

1.2 Codage

Patrick veut donc envoyer un message à **Marie**. Il cherche dans l'annuaire le système de codage et la clé de chiffrement qu'elle a publiés. Il sait maintenant qu'il doit utiliser le système *R.S.A* avec les deux entiers n et E (prenons par exemple $n = 1073 = 29 \times 37$ et $E = 5$, premier avec 28×36. Il transforme en chiffre son message en remplaçant par exemple chaque lettre par son rang dans l'alphabet.

«Au tableau» devient : «01 21 20 01 02 12 05 01 21»

Puis il le découpe en tronçons de même longueur représentant chacun un nombre plus petit que n.

Son message devient : «012 120 010 212 050 121»

Il élève ensuite chaque tronçon B à la puissance E et apelle C le reste de la division euclidienne de B^E par n. C constitue un tronçon du message codé que **Patrick** envoie tel quel à **Marie**.

Le message codé est : «969 589 211 788 553 544»

1.3 Décodage

Marie calcule à partir de p et q qu'elle a gardés secrêts la clé D de déchiffrage. Celle-ci est donnée par la formule $DE = 1 \bmod (p-1)(q-1)$. Ici, $D = 605$ par exemple. Puis elle élève chacun des tronçons C du message codé à la puissance D et retrouve B en prenant le reste de la division euclidienne de C^D par n.

Elle retrouve : «012 120 010 212 050 121»

Puis en groupant les chiffres deux par deux et en remplaçant les nombres ainsi obtenus par les lettres correspondante, elle sait enfin qu'elle doit aller au tableau, sans que personne d'autre ne soit au courant. Malheureusement, le temps qu'elle décode le message envoyé par **Patrick**, le cours est déjà terminé

1.4 Intérêt

L'efficacité du système $R.S.A$ repose sur le fait qu'à l'heure actuelle il est pratiquement impossible de retrouver p et q à partir de n si celui-ci est très grand. **Marie** est donc la seule à pouvoir calculer D. De plus, comme elle n'a jamais à transmettre les entiers p et q, le risque que **Paul** ou **Pierre** les interceptent est nul.

2 Démonstrations

2.1 Petit théorème de Fermat

Soient p un nombre premier et a dans $\mathbf{N}^\star$. Alors p divise $a^p - a$.

2.2 Preuve

Cela se démontre par récurrence sur a dans $\mathbf{N}^\star$.

Pour $a = 1$, on a effectivement $1^p - 1 = 0 \times p$.

Supposons maintenant que p divise $a^p - a$ et montrons que p divise $(a+1)^p - (a+1)$.

$$(a+1)^p = a^p + \sum_{k=1}^{p-1} C_p^k a^k + 1$$

Or

$$\forall k \in 1, 2, \dots, p-1, \quad p \mid C_{p-1}^k$$

car

$$k!C_{p-1}^k = \frac{k!p!}{k!(p-k)!} = p(p-1)...(p+k-1)$$

et

$$p \mid p(p-1)(p+k-1)$$

De plus p et k sont premiers entre eux car p est premier et $k \in 1, 2, \dots, p-1$. Donc par le théorème de Gauss, p divise C_{p-1}^k.

Ainsi p divise $\sum_{k=1}^{p-1} C_p^k a^k$, donc il existe q dans $\mathbf{N}$ tel que $\sum_{k=1}^{p-1} C_p^k a^k = qp$. Et par hypothèse de récurrence, il existe r dans $\mathbf{N}$ tel que $a^p = rp + a$. Donc on a : $(a+1)^p = (r+q)p + a + 1$ i.e. p divise $(a+1)^p - (a+1)$.$\square$

2.3 Corollaire

Soient p un nombre premier et a dans $\mathbf{N}^\star$ tels que p ne divise pas a. Alors comme p divise $a(a^{p-1} - 1)$, par le théorème de Gauss, p divise $a^{p-1} - 1$.

2.4 Démonstration du principe mathématique sur lequel repose le système R.S.A

Montrons que $C^D \equiv B \pmod n$.

Par hypothèse E est premier avec $(p-1)(q-1)$. Ainsi d'après l'identité de Bézout, il existe (D, F) dans $\mathbf{N} \times \mathbf{Z}$ tel que $DE + F(p-1)(q-1) = 1$ i.e. $F(p-1)(q-1) = 1 - DE$. Donc il existe D dans $\mathbf{N}$ tel que $DE-1$ soit divisible par $(p-1)(q-1)$.

Or $C \equiv B^E \pmod n$, donc $C - B^E$ est un multiple de n. De plus comme $C^D - B^{ED} = (C-B^E)(C^{D-1} + B^E C^{D-2} + \ldots + B^{ED-2}C + B^{ED-1})$, $C^D - B^{ED}$ est aussi un multiple de n.

Si B n'est divisible ni par p ni par q, alors, par le petit théorème de Fermat, p et q divisent respectivement $B^{p-1} - 1$ et $B^{q-1} - 1$. Donc p et q divisent $B^{(p-1)(q-1)} - 1$ car cet entier est multiple de $B^{p-1} - 1$ et $B^{q-1} - 1$.

Or comme $DE - 1$ est divisible par $(p-1)(q-1)$, il existe k dans $\mathbf{N}^\star$ tel que $DE - 1 = k(p-1)(q-1)$. Ainsi on a :

$$
\begin{aligned}
& B^{ED} - B \\
& \quad = \quad B^{1+k(p-1)(q-1)} - B \\
& \quad = \quad B(B^{k(p-1)(q-1)} - 1) \\
& \quad = \quad B(B^{(p-1)(q-1)} - 1)(B^{(k-1)(p-1)(q-1)} + \cdots + 1).
\end{aligned}
$$

Donc $B^{ED} - B$ est divisible par $n = pq$. Comme $C^D - B = (C^D - B^{ED}) + (B^{ED} - B)$, $C^D - B$ est divisible par n. Donc $C^D \equiv B \pmod n$.$\square$