L'algorithme RSA

Jean-Baptiste Campesato

10 avril 2010

Le but de ces quelques pages est de présenter l'algorithme RSA de façon rigoureuse tout en ayant un minimum de prérequis (juste quelques notions d'arithmétique). En effet toutes les notions nécessaires sont introduites et expliquées.

1 Introduction

RSA est un algorithme de cryptographie à clé publique développé par Ron **R**ivest, Adi **S**hamir et Len **A**dleman en 1977.
On parle d'algorithme de cryptographie à clé publique (ou d'algorithme de cryptographie asymétrique) lorsqu'une personne A dispose d'une fonction *à sens unique*, c'est à dire injective, très rapide à appliquer tout en étant très difficile à inverser si l'on ne connait pas une brèche que la personne A a en sa possession. Cette fonction est nommée clé publique et la brèche est quant à elle nommée clé privée.
A met à disposition de tout le monde la clé publique mais garde très secrètement la clé privée. Ainsi lorsqu'une autre autre personne, disons B, veut envoyer un message à A, elle crypte son message en lui appliquant la clé publique et envoie le message chiffré obtenu à A. Comme A est la seule personne connaissant la clé privée, B sait que même si le message est intercepté seule A pourra le déchiffrer.
La sécurité d'un tel procédé réside dans la difficulté à trouver la clé privée et surtout dans la difficulté à inverser la fonction donnée par la clé publique.

$f : A \to B$ est injective $\Leftrightarrow (\forall (x,y) \in A^2, x \neq y \Rightarrow f(x) \neq f(y))$ $\Leftrightarrow (\forall (x,y) \in A^2, f(x) = f(y) \Rightarrow x = y)$.

Un tel procédé permet d'identifier avec certitude le destinataire, un autre problème dans ce domaine est d'identifier l'expéditeur, ce que l'on peut faire avec RSA en *signant* le message.

L'algorithme RSA est très utilisé dans le domaine du numérique (pour les payements en ligne, lors des transactions avec des cartes bancaires...) et repose sur des propriétés arithmétique relativement simples et connues depuis un certain temps.
Nous ne traiterons pas des difficultés d'implémentation de RSA qui sont de toute façons traitées dans les documents de la bibliographie.

2 Les mathématiques de RSA

2.1 Le cadre d'étude

On va faire de l'arithmétique modulaire, c'est à dire travailler dans les anneaux $\mathbb{Z}/n\mathbb{Z}$ (ici et dans toute la suite on fixe la condition $n \geq 1$).

Pour les plus aguerris d'entre vous on peut définir $\mathbb{Z}/n\mathbb{Z}$ comme l'ensemble quotient que l'on obtient en considérant la relation d'équivalence $\forall (a,b) \in \mathbb{Z}^2$, $a\mathcal{R}b \Leftrightarrow (a \equiv b\,[n])$ (cela signifie que a et b sont en relation si et seulement s'ils ont le même reste par la division euclidienne par n si et seulement si $b-a$ est un multiple de n...). On peut encore voir $\mathbb{Z}/n\mathbb{Z}$ comme l'anneau quotient obtenu par la donnée de l'idéal $n\mathbb{Z}$. Il s'agit d'un anneau unitaire commutatif et même principal et de Bézout.
Les constructions mathématiques ci-dessus permettent d'obtenir directement des propriétés mathématiques intéressantes découlant de nos théories abstraites sur les anneaux quotient et de notre connaissance de $\mathbb{Z}$.

$n|(m-k)$ $\Leftrightarrow \exists l \in \mathbb{Z}$, $m-k = nl$ $\Leftrightarrow n$ divise $(m-k)$ $\Leftrightarrow (m-k)$ est un multiple de n.

Sinon de façon plus simple on peut voir $\mathbb{Z}/n\mathbb{Z}$ comme un ensemble de parties (ou sous-ensembles) de $\mathbb{Z}$: $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$ où $\overline{k}$ est l'ensemble des entiers ayant k pour reste lors de la division euclidienne par n, c'est à dire $\overline{k} = \{m \in \mathbb{Z},\ n|(m-k)\} = \{m \in \mathbb{Z},\ m \equiv k\,[n]\}$.
Ainsi dans $\mathbb{Z}/5\mathbb{Z}$: $\overline{0} = \overline{5} = \overline{15} = \overline{-5}$ et de même $\overline{2} = \overline{-3} = \overline{7}$.
Les propriétés des anneaux, des anneaux quotients et de $\mathbb{Z}/n\mathbb{Z}$ (ainsi que les différentes constructions) sont disponibles dans les documents de la bibliographie.
Dans la suite nous donnerons au fur et à mesure les propriétés nécessaires à l'explication de RSA (Où on admettra les notions qui ne posent pas de difficulté comme le fait que le produit et l'addition sont bien définis, d'ailleurs du fait de cela certains auteurs écrivent de façon abusive a au lieu de $\overline{a}$ sans nuire à la compréhension).

Soient $n \in \mathbb{N}^*$ et $x \in \mathbb{Z}/n\mathbb{Z}$, on dit que x est inversible (dans $\mathbb{Z}/n\mathbb{Z}$) s'il existe $y \in \mathbb{Z}/n\mathbb{Z}$ tel que $xy = \overline{1}$. On dit que y est l'inverse de x (en effet s'il existe, on peut montrer qu'il est unique).
On peut montrer que $\overline{a}$ est inversible dans $\mathbb{Z}/n\mathbb{Z}$ si et seulement si $\text{pgcd}(a,n) = 1$. (†)
On note $(\mathbb{Z}/n\mathbb{Z})^*$ l'ensemble des éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$.

Si un ensemble A admet un nombre fini d'éléments, alors son cardinal, noté $\text{card}(A)$, et le nombre d'éléments qu'il contient.

On définit désormais $\varphi : \begin{array}{ccc} \mathbb{N}^* & \longrightarrow & \mathbb{N}^* \\ n & \longmapsto & \text{card}\left((\mathbb{Z}/n\mathbb{Z})^*\right) \end{array}$ l'application qui à n associe le nombre d'éléments inversibles dans $\mathbb{Z}/n\mathbb{Z}$. La fonction φ est nommée *indicatrice d'Euler* et intervient dans de très nombreux théorèmes d'arithmétique et de la théorie des nombres.
On sait par ailleurs calculer $\varphi(n)$ de façon effective à partir de la décomposition en facteurs premiers de n : $\varphi\left(\prod_{i=1}^{m} p_i^{\alpha_i}\right) = \prod_{i=1}^{m} p_i^{\alpha_i} - p_i^{\alpha_i - 1} = \prod_{i=1}^{m} p_i^{\alpha_i - 1}(p_i - 1)$.

On va aussi utiliser le théorème d'Euler qui est une généralisation du (petit) théorème de Fermat (qui se restreint au cas où n est premier) :

Théorème: Théorème d'Euler

Soient $n \in \mathbb{N}$ et $a \in \mathbb{Z}$ tels que $\text{pgcd}(a,n) = 1$ (i.e. $\overline{a} \in (\mathbb{Z}/n\mathbb{Z})^*$).
Alors $a^{\varphi(n)} \equiv 1\,[n]$ (i.e. $\overline{a}^{\varphi(n)} = \overline{1}$).

On peut aussi démontrer le théorème d'Euler en utilisant le théorème de Lagrange.

Démonstration :

On peut vérifier que $f : \begin{array}{ccc} (\mathbb{Z}/n\mathbb{Z})^* & \longrightarrow & (\mathbb{Z}/n\mathbb{Z})^* \\ x & \longmapsto & \overline{a}x \end{array}$ est une bijection.

On a alors : $\prod_{x \in (\mathbb{Z}/n\mathbb{Z})^*} \overline{a}x = \prod_{x \in (\mathbb{Z}/n\mathbb{Z})^*} f(x) = \prod_{x \in (\mathbb{Z}/n\mathbb{Z})^*} x$

$\Rightarrow \left(\prod_{x \in (\mathbb{Z}/n\mathbb{Z})^*} \overline{a}\right)\left(\prod_{x \in (\mathbb{Z}/n\mathbb{Z})^*} x\right) = \prod_{x \in (\mathbb{Z}/n\mathbb{Z})^*} x.$

$\Rightarrow \prod_{x \in (\mathbb{Z}/n\mathbb{Z})^*} \overline{a} = \overline{1}.$

On a donc le résultat du fait que $\overline{a}^{\varphi(n)} = \prod_{x \in (\mathbb{Z}/n\mathbb{Z})^*} \overline{a}$.

■

Une application $f : A \to B$ est bijective si et seulement si elle est injective et surjective, c'est à dire si et seulement si pour tout $y \in B$ il existe un unique $x \in A$ tel que $y = f(x)$.

2.2 Les outils mathématiques de RSA

RSA repose sur le fait que l'on ne sait pas réaliser rapidement (du moins sans ordinateur quantique (algorithme de Shor) ou sans une certaine puissance de calcul encore inaccessible) :

1. Factoriser un grand entier en facteurs premiers (factorisation qui existe toujours et qui est unique, rappellons le).
2. Inverser $f_q : \begin{array}{ccc} \mathbb{Z}/n\mathbb{Z} & \longrightarrow & \mathbb{Z}/n\mathbb{Z} \\ x & \longmapsto & x^q \end{array}$ avec $q \geq 1$.

On va donc construire pas à pas notre fonction à sens unique, mais avant cela il nous faut quelques résultats mathématiques :

Lemme

Soient p et q deux nombres premiers distincts et soit $n = pq$.
Soit $t \in \mathbb{N}$ tel que $t \equiv 1\,[\varphi(n)]$.
Alors $\forall x \in \mathbb{Z}/n\mathbb{Z}$, $x^t = x$ (i.e. $\forall a \in \mathbb{Z}$, $a^t \equiv a\,[n]$).

Démonstration :

On sait déjà que $\varphi(n) = (p-1)(q-1)$, posons $t = 1 + k\varphi(n)$ avec $k \in \mathbb{N}$ et soit $a \in \mathbb{Z}$.

Il y a 4 cas possibles à traiter :
• ($\text{pgcd}(a,p)=1$ et $\text{pgcd}(a,q)=1$) $\Leftrightarrow \text{pgcd}(a,n)=1$:
Alors d'après le théorème d'Euler $a^{\varphi(n)} \equiv 1\,[n]$ et donc : $a^t = a(a^{\varphi(n)})^k \equiv a\,[n]$.
• $\text{pgcd}(a,p)=1$ et a est un multiple de q :
D'après le théorème d'Euler $a^{p-1} = a^{\varphi(p)} \equiv 1\,[p]$ d'où une première relation :
$a^t = a^{1+k(p-1)(q-1)} = a(a^{p-1})^{k(q-1)} \equiv a\,[p] \Rightarrow p|(a^t-a)$
Puis comme a est un multiple de q : $a^t \equiv a\,[q] (\equiv 0\,[q]) \Rightarrow q|(a^t-a)$.
On a donc que (a^t-a) est à la fois un multiple de p et de q et donc de $n=pq$ (comme $\text{pgcd}(p,q)=1$).
Et ainsi $a^t \equiv a\,[n] (\equiv 0\,[n])$.
• $\text{pgcd}(a,q)=1$ et a est un multiple de p :
Se traite de la même façon que le dernier point en inversant les rôles de p et de q.
• a est un multiple de $n=pq$:
On a $a \equiv 0\,[n]$ et donc $a^t \equiv a\,[n]$.

■

Théorème 1: Construction des clés

Soient p et q deux nombres premiers distincts et soit $n=pq$.
Soit $e \in \mathbb{N}$ premier avec $\varphi(n)$.
On note d l'inverse de e modulo $\varphi(n)$ (qui existe d'après la remarque (†) de 2).
Alors :

1. $f_e : \begin{array}{ccc} \mathbb{Z}/n\mathbb{Z} & \longrightarrow & \mathbb{Z}/n\mathbb{Z} \\ x & \longmapsto & x^e \end{array}$ est une bijection de réciproque
   $f_d : \begin{array}{ccc} \mathbb{Z}/n\mathbb{Z} & \longrightarrow & \mathbb{Z}/n\mathbb{Z} \\ x & \longmapsto & x^d \end{array}$.
2. On en déduit que :
   $\pi_e : \begin{array}{ccc} [\![2;n-1]\!] & \longrightarrow & [\![2;n-1]\!] \\ x & \longmapsto & \text{reste de la DE de } x^e \text{ par } n \end{array}$
   et $\pi_d : \begin{array}{ccc} [\![2;n-1]\!] & \longrightarrow & [\![2;n-1]\!] \\ x & \longmapsto & \text{reste de la DE de } x^d \text{ par } n \end{array}$
   sont des bijections réciproques de $[\![2;n-1]\!]$.

$[\![2;n-1]\!] = \{2,3,\dots,n-1\}$.

DE : division euclidienne.

π_e et π_d sont en fait respectivement nos fonctions de chiffrement et de déchiffrement.

Démonstration :

$\forall x \in \mathbb{Z}/n\mathbb{Z},\ f_e \circ f_d(x) = f_d \circ f_e(x) = x^{ed} = x$ d'après le lemme puisque $ed \equiv 1\,[\varphi(n)]$.
On passe au point 2 en remarquant que $f_e(\overline{0}) = f_d(\overline{0}) = \overline{0}$ et que $f_e(\overline{1}) = f_d(\overline{1}) = \overline{1}$.

■

3 La construction de RSA

Munis de tout cet arsenal mathématique, nous sommes en mesure de construire l'algorithme RSA :
Tout d'abord il faut se fixer les mots que l'on peut transmettre (par exemple des chiffres, des lettres, des octets ou des paquets d'une certaine taille...) et à chacun de ces mots on associe un unique entier de façon à obtenir un ensemble $[\![2;C]\!]$ ($C \in \mathbb{N}$, $C \geq 2$). Ce sont donc des éléments de $[\![2;C]\!]$ que l'on cryptera et que l'on transmettra. Le passage des mots aux éléments de $[\![2;C]\!]$ est l'encodage et il est connu de tous.
Notre personne A de l'introduction va donc construire ses clés de la façon suivante :

1. Elle se donne deux nombres premiers distincts p et q et calcule $n=pq$ et $\varphi(n) = (p-1)(q-1)$. Comme on travaille modulo n, il faut $n > C$.
2. Elle choisit un entier $d > 1$ premier avec $\varphi(n)$ et calcul son inverse e modulo $\varphi(n)$.
3. Le couple (e,n) définit la clé publique que A publie et le couple (d,n) définit la clé privée que A garde pour elle. Toutes les autres données peuvent être supprimées (notamment p, q et $\varphi(n)$).

3.1 Envoi d'un message crypté (certitude du bon destinataire)

Imaginons que B veuille envoyer le message $m \in [\![2; C]\!]$ à A. Alors B prend connaissance de la clé publique (e, n) de A et envoie le message crypté $M = \pi_e(m)$ à A.
Ensuite lorsque A reçoit M, elle regarde sa clé privée (d, n) et calcule $\pi_d(M)$, d'après le théorème on retrouve bien m.

Imaginons qu'une personne E intercepte M. Ayant de plus accès à la clé publique (e, n), E doit en déduire d qui est l'inverse de e modulo $\varphi(n) = (p-1)(q-1)$. E doit donc calculer p et q, c'est à dire factoriser n en facteurs premiers, ce qui est actuellement extrêmement long pour p et q suffisamment grands.

3.2 Signature d'un message (certitude du bon expéditeur)

Ici il ne s'agit plus d'envoyer un message secret mais de vérifier qu'il n'y a pas usurpation de l'identité de l'expéditeur.
Cette fois c'est A qui envoie un message à B. B a toujours accès à la clé publique (e, n) de A.
Si A veut envoyer le message $m \in [\![2; C]\!]$ à B, elle lui envoie d'abord m puis $s = \pi_d(m)$.
B reçoit m et s et calcule $\pi_e(s)$, d'après le théorème elle doit normalement retomber sur m.
Et comme seule A connait d, B est certain que c'est bien A qui lui a envoyé le message.

4 Les atouts et les faiblesses de RSA

Atouts	Faiblesses
Si p et q suffisamment grands, le temps de factorisations est actuellement très long	Algorithme de Shor dès que l'on aura un ordinateur quantique
...	Difficulté d'implémentation : comment choisir des premiers de grande taille ? calcul des $m^d[n]$ (pour faire la division euclidienne en fait on applique une exponentiation modulaire) ?
...	On peut tenter de s'attaquer non plus à la factorisation mais à la façon dont sont générés p et q
...	...

5 Bibliographie

5.1 Pour RSA (et les outils mathématiques nécessaires) :

http://www.edu.upmc.fr/maths/wassef_LM220/documents/arithmetique.pdf
http://www.math.jussieu.fr/~hindry/Cours-arith.pdf
http://perso.univ-rennes1.fr/antoine.chambert-loir/2005-06/a2/coursa2.pdf
http://people.math.jussieu.fr/~boyer/enseignement/L2/cours.pdf
Michel Demazure, « Cours d'algèbre », ISBN-13 : 9782842251277 chez Cassini.

5.2 Pour approfondir l'algèbre (plus abstraite) :

Gostiaux, Perrin, Godement, Lang... à rédiger...