Feuille TD 2 - RSA

1 Codage et décodage RSA.

On considère la clef publique RSA (11, 319), c'est-à-dire pour $n = 319$ et $e = 11$.
Note : on pourra utiliser les résultats suivants :

- $319 = 11 \times 29$; $10^{11} = 263 \pmod{319}$; $263^2 = 216 \times 319 + 265$;
- $133^3 = 12 \pmod{319}$; $133^{25} = 133 \pmod{319}$;
- $11^2 = 121 \pmod{280}$; $11^4 = 81 \pmod{280}$; $11^8 = 121 \pmod{280}$; $11^{16} = 81 \pmod{280}$;
- $95 = 64 + 31$; $81.11 = 51 \pmod{280}$; $81.121 = 1 \pmod{280}$.

1. Quel est le message correspondant au codage avec cette clé du message $M = 100$?

   Correction: $M' = 100^{11} \pmod{319} = 265$

2. Calculer d la clé privée correspondant à la clé publique e.

   Correction: *On doit résoudre* $11 * d = 1 \pmod{280}$. *On trouve* $d = 51$
   - *soit en utilisant l'algorithme d'Euclide étendu;*
   - *soit en essayant "à la main" car* $51 = (280/11) * 2$ *donc il suffit de deux essais pour trouver;*
   - *soit en utilisant Euler,* $d = 11^{-1} \pmod{280}$ *d'où* $d = 11^{\phi(280)-1} \pmod{280} = 11^{\phi(7.5.8)-1} \pmod{280} = 11^{(6.4.4)-1} \pmod{280} = 11^{95} \pmod{280} = 11^{64+16+8+4+2+1} \pmod{280} = 81.81.121.81.121.11 = 81.11 = \pmod{280} = 51 \pmod{280}$.

3. Décoder le message $M' = 133$.

   Correction: *on doit calculer* $133^{51} \pmod{319}$. *Dans les notes on donne* $133^{25} = 133 \pmod{319}$. *Le résultat est* $133 * 133 * 133 \pmod{319} = 12$.

4. Le message codé 625 peut-il résulter d'un codage avec la clé publique ? Même question avec la clé privée.

   Correction: *évidemment non pour les deux car le résultat doit être inférieur à 319.*

2 Cryptographie RSA et authentification

Un professeur envoie ses notes au secrétariat de l'École par mail. La clef publique du professeur est (3,55); celle du secrétariat est (3,33).

1. Vérifier que la clef privée du professeur (supposée connue de lui seul) est 27; et que celle du secrétariat est 7.

2. Pour assurer la confidentialité de ses messages, le professeur chiffre les notes avec la clef RSA du secrétariat. Quel message chiffré correspond à la note 12 ?

3. Pour assurer l'authenticité des messages contenant les notes, le professeur signe ses messages pour le secrétariat aprés les avoir chiffrés. Le secrétariat reçoit le message 23. Quelle est la note correspondante ?

Correction: *1. Pour le professeur:* $\phi(55) = 40$ *et* $27.3 = 81 = 1 \mod 40$.
Pour le secrétariat: $\phi(33) = 20$ *et* $7.3 = 21 = 1 \mod 20$.

2. Le professeur envoit $m = 12^3 \mod 33$. *Or* $12^2 = 12[33]$*; donc* $m = 12 \mod 33$.

3. Le professeur a calculé $(x^{e_S} \mod n_S)^{d_P} \mod n_P$.
Le sécrariat reçoit y et calcule $(y^{e_P} \mod n_P)^{d_S} \mod n_S$.
D'où la note $(23^3 mod 55)^7 \mod 33 = (12^7 \mod 33) = 12$.

3 Quelques attaques de RSA.

1. Montrer que casser un code RSA de clef publique (n, e) avec $e = 3$ est polynomialement aussi difficile que factoriser n.
Ceci fonde la robustesse de l'algorithme RSA; mais cela ne justifie pas pour autant une confiance aveugle en n'importe quelle implémentation de RSA car de nombreuses attaques sont possibles liées à des mauvais choix d'implémentation ou de protocoles. Les questions suivantes présentent quelques unes de ces attaques.

Correction: *cf cours*

2. Attaque par diffusion de messages sur un même exposant e petit. William, Jack et Averell ont respectivement les clefs RSA publiques $(n_W, 3)$, $(n_J, 3)$ et $(n_A, 3)$. Joe envoie en secret à chacun d'eux le même message x avec $0 \leq x < \text{Min}(n_W, n_J, n_A)$.
Montrer que Lucky Luke, qui voit passer sur le réseau $x^3 \mod n_W$, $x^3 \mod n_J$ et $x^3 \mod n_A$ peut facilement calculer x.
Indication. On rappelle (ou on admettra !) que pour a et k entier, la méthode de Newton permet de calculer très rapidement $\lfloor a^{1/k} \rfloor$, en temps $O(\log^2 a)$.

Correction: *Grâce au théorème des restes chinois, Lucky Luke peut calculer facilement – en temps polynomial $O(\log^2(n_A + n_J + n_W))$ – l'entier $C = x^3$ mod $n_W.n_J.n_A$ i.e. $C = x^3$ puisque $x < Min(n_W, n_J, n_A)$.*
Comme x est entier, il suffit ensuite de calculer dans $\mathbb{R}$ $x = C^{1/3}$ par Newton pour obtenir x en clair.

3. Attaque par texte chiffré bien choisi. Eve intercepte le message c chiffré envoyé par Bob à Alice : $c = m^{e_A} \mod n_A$. Pour déchiffrer c, Eve procède comme suit :

1. Eve choisit un entier $0 < r < n_A$ au hasard et calcule $x{:=}r^{e_A} \mod n_A$;
2. Eve calcule $y{:=}x.c \mod n_A$;
3. Eve demande à Alice de signer y avec sa clef privée; Alice renvoit à Eve $u = y^{d_A} \mod n_A$.

Montrer que Eve peut alors facilement découvrir le message m émis par Bob (on calculera $u.r^{-1} \mod n_A$). Moralité ?

Correction: *En effet: $u.r^{-1} \mod n_A = y^{d_A}.r^{-1} \mod n_A = c^{d_A}x^{d_A}r^{-1} \mod n_A = c^{d_A}(x^{d_A} \mod n_A)r^{-1} \mod n_A = c^{d_A}rr^{-1} \mod n_A = c^{d_A} \mod n_A = m$.*
Le calcul de $u.r^{-1} \mod n_A$ peut être très facilement fait par Eve en temps $O(\log^2 n_A)$ et donne le message m.
Moralité: ne jamais signer un message illisible (aléatoire ou chiffré); ou alors ne chiffrer que le résumé du message.

4. Attaque par modulo commun. Une implémentation de RSA donne à deux personnes (Alice et Bob) le même nombre n (produit de deux nombres premiers) mais des clefs (e_A, d_A) et (e_B, d_B) différentes. On suppose de plus que e_A et e_B sont premiers entre eux (ce qui est le plus général).
Supposons alors que Alice et Bob chiffrent un même message m et que Eve intercepte les deux messages $c_A = m^{e_A} \mod n_A$ et $c_B = m^{e_B} \mod n_B$ qu'elle sait être deux chiffrements du même message m.
Montrer que Eve peut alors très facilement découvrir le message m (on pourra utiliser le calcul des coefficients de Bezout associés à e_A et e_B). Moralité ?

Correction: *Eve connaît (e_A, e_B, n) et (c_A, c_B). Par l'algorithme d'Euclide étendu, elle calcule facilement (en temps -presque- linéaire du nombre de bits de n) les coefficients de Bezout r et s tels que $re_A + se_B = 1$.*
Eve peut alors calculer (par exponentiation rapide) $c_A^r.c_B^s \mod n = m^{r.e_A+s.e_B} \mod n = m$.
Moralité: ne jamais utiliser le même n pour un groupe d'utilisateurs.

4 Cryptographie à clef publique par résidu quadratique

Soient a et b deux entiers; on dit que $a \neq 0$ est un *carré* (ou *résidu quadratique*) modulo b ssi il existe x tel que $x^2 \equiv a \mod b$.
On dit alors que x est une *racine carrée* de a modulo b.
Dans tout l'exercice, p et q désignent deux nombres premiers différents de 2 et $n = p.q$.

1. Dénombrement des carrés dans $\mathbb{Z}/n\mathbb{Z}^\star$

a. Vérifier que si $x^2 \equiv a \mod b$, alors $(b-x)^2 \equiv a \mod b$.

b. Montrer que si a est un carré modulo n, alors a est aussi un carré modulo p et modulo q.

c. Montrer que tout carré $a \neq 0$ modulo p a exactement 2 racines : x et $y = p - x$.

d. En déduire que tout carré a dans $\mathbb{Z}/n\mathbb{Z}$, tel que a est premier avec p et q, admet exactement quatre racines carrées distinctes x_1, $n - x_1$, x_2 et $n - x_2$. **Indication :** utiliser le théorème chinois des restes.

e. En utilisant que $(\mathbb{Z}/p\mathbb{Z}^\star, \times)$ est cyclique, montrer qu'il y a $\frac{p-1}{2}$ carrés non nuls modulo p.

f. En déduire le nombre de carrés dans $\mathbb{Z}/n\mathbb{Z}^\star$.

Correction: *a.* $(b-x)^2 = b^2 - 2bx + x^2 = x^2 = a \mod b$.

b. $a = x^2 + kpq$; *donc* $a \equiv x^2 \mod p$ *est un carré modulo* p *(de même pour* q*).*

c. Soit x *et* y *distincts tels que* $a = x^2 = y^2 \mod p$. *D'où* $x^2 - y^2 = (x-y)(x+y) = 0 \mod p$. *Comme* $\mathbb{Z}/p\mathbb{Z}$ *est intègre (car* p *est premier) et* $x - y \neq 0 \mod p$, *on en déduit que* $x + y = 0 \mod p$; *d'où* $y = p - x$.

d. D'après b., tout carré $a \mod n$ *est un carré* $\mod p$ *et* $\mod q$. *Comme* $a \neq 0 \mod p$, a *admet exactement 2 racines* $u_1 = u$ *et* $u_2 = p - u$ *modulo* p *(resp.* $v_1 = v$ *et* $v_2 = q - v$ *modulo* q*). On en déduit, par le théorème chinois des restes, l'existence de exactement 4 racines distinctes pour* a *dans* n *:* $u_i.q.q^{-1[p]} + v_j.p.p^{-1[q]} \mod n$ *avec* $1 \leq i, j \leq 2$.
D'après a., on déduit que ces racines s'écrivent nécessairement x_1, $n - x_1$, x_2 *et* $n - x_2$.

e. Soit g *un générateur de* $(\mathbb{Z}/p\mathbb{Z}^\star, .)$ *qui est cyclique. Supposons qu'il existe* x *tel que* $g = x^2$. *D'après Fermat,* $g^{p-1} = 1 \mod p$; *comme* g *est générateur et* p *impair, on en déduit que* $g^{\frac{p-1}{2}} = -1$. *D'où* $x^{p-1} = -1 \neq 1$ *car* $p \neq 2$. *Donc, d'après le théorème de Fermat,* x *n'appartient pas à* $\mathbb{Z}/p\mathbb{Z}$. *Ainsi,* g *n'est pas un carré modulo* p.
On en déduit que les seuls carrés non nuls sont les éléments g^{2i}; *il y en a* $\frac{p-1}{2}$.
NB: g^{2i} *a deux uniques racines carrées:* $x = g^i$ *et* $g^{i+\frac{p-1}{2}} = -x = p - x$.

f. *Soit g un générateur de $(\mathbb{Z}/p\mathbb{Z}^\star, .)$ qui est cyclique. Par le théorème chinois des restes, à tout couple de carré $(u, v) \in \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$, on associe un et un seul carré dans $\mathbb{Z}/n\mathbb{Z}$. En comptant 0, il y a $\frac{p+1}{2}$ carrés modulo p. D'où $\frac{(p+1)(q+1)}{4}$ carrés modulo n, soit $\frac{n+p+q-3}{4}$ carrés non nuls modulo n.*

2. Difficulté du calcul des racines carrées. Soit $a < n$; le but de cette question est de montrer que calculer les racines carrées x de $a \neq 0$ modulo n est (polynomialement) plus difficile que factoriser n.
On suppose donc dans toute cette question que l'on connaît les 4 racines distinctes x_1, x_2, $(n - x_1)$ et $(n - x_2)$ de a modulo n; on veut montrer qu'il est alors possible de factoriser rapidement n.

a. Soit $u = x_1 - x_2 \mod n$ et $v = x_1 + x_2 \mod n$. Montrer que $u.v \equiv 0 \mod n$.

b. En justifiant que $1 \leq u, v < n$, expliquer comment calculer alors les deux facteurs p et q de n à partir de u et v.

c. Donner une majoration du coût de ce calcul en fonction du nombre de bits de n.

d. En déduire que la fonction *carré* de $\mathbb{Z}/n\mathbb{Z}$ définie par *carré*$(x) = x^2 \mod n$ peut être considérée comme une fonction à sens unique.

Correction: *a.* $u.v = x_1^2 - x_2^2 = a^2 - a^2 = 0 \mod n$.

b. On peut supposer $1 \leq x_1, x_2 < n$. Comme $x_1 \neq x_2$, $u = x_1 - x_2 \neq 0$. Comme $x_1 \neq n - x_2$, $v = x_1 + x_2 = x_1 - (n - x_2) \neq 0$. Donc $1 \leq u, v < n$.
On a $u.v = k.n$; donc $n = p.q$ divise $u.v$. Comme $u < p.q$, et p et q premiers, on en déduit que p divise u ou q divise u mais $p.q$ ne divise pas u. Donc $pgcd(n, u)$ fournit un des 2 facteurs de n et $n/pgcd(n,u)$ l'autre.

c. Il suffit de faire une addition, un pgcd et une division. Le coût dominant est le pgcd; par Euclide, cela donne un coût en $O(t^2)$ (ou $O(t \log^2 t \log\log t)$ par Schonhagge).

d. Calculer $x^2 \mod n$ se calcule efficacement en $O(t^{1+\epsilon})$. Par contre le calcul de sa réciproque est plus difficile que la factorisation; en effet, si on sait calculer les racines carrées de $a \mod p$, on sait factoriser a. Comme la factorisation est un problème réputé difficile, on en déduit que carré *peut être considérée comme une fonction à sens unique.*

3. Protocole d'identification quadratique. Soit $n = pq$ un nombre de 512 bits, produit de deux nombres premiers; p et q ne sont connus que d'un tiers de confiance TTP, mais pas d'Alice et de Bob.
Pour s'identifier, Alice choisit l'entier $x_A < n$ comme clef secrète unique. Soit $a = x_A^2 \mod n$; TTP délivre alors à Alice un passeport sur lequel figure les entiers publics n et a.

a. On suppose que seule Alice (et peut-être TTP) connaît x_A et que personne en dehors de TTP ne sait calculer les racines carrées modulo n; est-ce raisonnable ?

b. Pour identifier Alice, le douanier Bob qui consulte le passeport d'Alice utilise le protocole suivant (qu'il répète 2 ou 3 fois) :

1. Alice choisit un nombre r au hasard qu'elle garde secret;
2. Alice calcule $y = r^2 \mod n$ et $z = x_A.r \mod n$;
3. Alice envoie y et z à Bob;
4. Bob teste l'identité d'Alice en vérifiant que $a.y - z^2 = 0 \mod n$.

Montrer que si un espion, qui ne sait pas calculer des racines carrées, a pu calculer r, c'est nécessairement qu'il connaît la clef secrète x_A de Alice. Qu'en déduisez-vous ?

c. Cependant, avec le protocole précédent, un espion peut se faire passer pour Alice : à la place des étapes 1 et 2, l'espion tire au hasard un nombre z et calcule $y = z^2/a \mod n$.
Pour éviter cela, le protocole suivant (dit protocole à *zéro-connaissance*) est utilisé :

1. Alice choisit r au hasard, calcule $y = r^2 \mod n$ et envoie y à Bob;
2. Bob tire au hasard $b \in \{0, 1\}$; il envoie b à Alice;
3. Si Alice reçoit 0, elle envoye $z = r$ à Bob (i.e. une racine de y modulo n);
   si elle reçoit 1, elle envoye à Bob $z = x_A.r \mod n$ (i.e. une racine de $y.m \mod n$).
4. Bob teste l'identité d'Alice en vérifiant que $y.a^b - z^2 = 0 \mod n$.

Majorer alors la probabilité que l'espion a de répondre correctement à Bob après k passages dans ce protocole.

Correction: *a. On ne connaît pas d'algorithme rapide (inférieur à 5 ans disons, la durée d'un passeport) pour factoriser un entier de 512 bits. Donc, on peut supposer que personne ne connaît p et q sauf TTP.*
De plus, on peut supposer que personne ne connaît x_A sauf Alice et éventuellement TTP.
La seule solution pour calculer x_A est alors d'extraire la racine carrée de $a \mod n$; d'après 2., ce calcul des racines carrées x_a de a est plus difficile que factoriser n. Donc on peut supposer que personne ne connaît x_A. L'hypothèse est donc raisonable.
Hors-question: *en fait, même si cela n'est pas demandé, on peut supposer que TTP, qui connaît p et q ne connaît pas x_A. En effet, connaitre x_A à partir de a demande de savoir calculer des racines carrées modulo p. Or, montrons qu'alors TTP saurait calculer le log discret, qui est supposé difficile. En effet, soit g un générateur de $\mathbb{Z}/p\mathbb{Z}^\star$. Soit $y < p$; en calculant une racine carrée de $y \mod p$ (ou de y/g si y n'a pas de racine carrée), TTP peut trouver y_1 tel que $y_1^2 = y$. Si $y_1 = g$, il en déduit l'indice de y: 2 (ou 1). Sinon, en réitérant ce calcul de racine carré à partir de y_1 jusqu'à trouver $y_k = g$, il peut en déduire l'indice i de y_1; et donc l'indice $2.i$ (ou $2.i+1$) de y. TTP saurait donc calculer le log discret modulo p.*

b. *Si l'on ne sait pas calculer les racines carrées, r étant pris au hasard, la connaissance de y n'est d'aucune utilité. La seule solution pour calculer r est alors d'utiliser z; mais calculer r à partir de z est équivalent à calculer x_A. La seule solution pour calculer r est donc de connaître x_A.*
On en déduit que Bob peut ainsi identifier Alice.

c. *D'après 2., sans connaître la clef de Alice et sans savoir calculer les recines carrées, la seule solution pour l'espion est de tricher. Il doit parier sur ce que Bob va lui envoyer (0 ou 1) pour envoyer y. Mais sa probabilité de parier juste à 1 tirage est 1/2. Donc sa probabilité de tromper Bob après k tirages est inférieure à 2^{-k}.*