Dpt Telecommunications 1A INPG - Avril 2002 Théorie des codes
Durée : 2 heures. Tous documents et calculatrices autorisés.

Exercice 1. Identification par résidu quadratique (12 points)

On dit que $a \neq 0$ est un *carré* (ou *résidu quadratique*) modulo b ssi $\exists x : x^2 \equiv a \mod b$.
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. 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.

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$\star$. Hors-sujet: On rappelle que $\mathbb{Z}/p.\mathbb{Z}^\star$ est cyclique : il existe donc un générateur $g \in \mathbb{Z}/p.\mathbb{Z}^\star$ tel que $\mathbb{Z}/p.\mathbb{Z}^\star = \{g^i \mod p; i = 0, \ldots, p-2\}$.
Montrer qu'il y a $\frac{p-1}{2}$ carrés non nuls modulo p. En déduire le nombre de carrés dans $\mathbb{Z}/n.\mathbb{Z}^\star$.
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* $g^{i+\frac{p-1}{2}} = -x = p - x$.
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. 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. 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}.*

Exercice 2. Codes Correcteurs (8 points)

On veut comparer le rendement en terme de bits de différents codes 2-correcteurs.

1. Nous considérons tout d'abord un code de répétition sur $V = \{0, 1\}$, c'est à dire un code (mk, k) où chaque groupe de k bits est simplement répété m fois.

a. Quelle est la distance minimale entre deux mots de code ?

b. Proposer un code de répétition qui soit 2-correcteur

c. Quel est son rendement ?

Correction:

a. $\delta(C)$ *est le minimum des poids :* $Min(w(x); x \neq [0 \dots 0]) = m$.

b. Il faut $m = 2 * 2 + 1 = 5$. *Donc* $(5, 1)$ *convient.* $(5k, k)$ *aussi.*

c. $(5k, k)$, *est de rendement* $\frac{k}{5k} = 0.2$.

2. Nous considérons maintenant un code de Reed-Solomon sur $\mathbb{F}_8$. Nous étudions tout d'abord la construction de $\mathbb{F}_8$, puis un code de Reed-Solomon 2-correcteur et son rendement en terme de bits.

a. Enumérer tous les polynômes de degré 0, 1, 2 de $\mathbb{Z}/_{2\mathbb{Z}}[Y]$ et donner leur codage sur 3 bits $[b_2 b_1 b_0]$, le bit b_2 étant le coefficient du monôme de plus grand degré. Lesquels sont irréductibles ?

b. Montrer que $Q = 1 + Y + Y^3$ est irréductible sur $\mathbb{Z}/_{2\mathbb{Z}}[Y]$, en déduire une construction possible de $\mathbb{F}_8$.

c. Que valent $[010]^3$, $[010]^4$ dans cette construction ? On admettra dans la suite que Q est primitif sur $\mathbb{Z}/_{2\mathbb{Z}}[Y]$, c'est à dire que $\mathbb{F}_8^\star = \{[010]^i \mod Q; i = 0, \dots, 6\}$.

d. Nous construisons maintenant un code de Reed-Solomon sur $\mathbb{F}_8$. Quelle doit être la taille n des messages ?

e. Nous choisissons maintenant comme polynôme générateur de notre code $g = (X - [010]) * (X - [100]) * (X - [011]) * (X - [110])$. Montrer que g définit un code 4-détecteur et 2-correcteur.

f. Quel est le rendement de ce code en terme d'éléments de $\mathbb{F}_8$? en terme de bits totaux ?

Correction:

a. 0, 1, Y, $1+Y$, Y^2, $1+Y^2$, $Y+Y^2$, $1+Y+Y^2$ *correspondent à* $[000]$, $[001]$, $[010]$, $[011]$, $[100]$, $[101]$, $[110]$, $[111]$. *Seuls* 0, 1, Y, $1+Y$, $1+Y+Y^2$ *sont irréductibles.*

b. $(1+Y)(1+Y+Y^2) = 1+Y^3 \neq Q$, $Y(1+Y+Y^2) = Y+Y^2+Y^3 \neq Q$, $Y^3 \neq Q$, $(1+Y)^3 = 1+Y+Y^2+Y^3 \neq Q$. *On construit* $\mathbb{F}_8$ *comme étant* $\mathbb{Z}/_{2\mathbb{Z}}[Y]/Q$.

c. $[010]^3 = Y^3 = 1+Y = [011]$, $[010]^4 = Y^4 = Y(1+Y) = Y+Y^2 = [110]$. *Pour montrer qu'il est primitif il faudrait exhiber toutes les puissances :* $[010]^5 = Y(Y^4) = Y(Y+Y^2) = Y^2+1+Y = [111]$, $[010]^6 = Y(Y^5) = Y^3+Y+Y^2 = Y^2+1 = [101]$ *et* $[010]^7 = Y(Y^7) = Y^3+Y = 1 = [001]$. *Cela prouve que* Y *est générateur de* $\mathbb{F}_8^*$.

d. $n = q - 1 = 8 - 1 = 7$.

e. A l'aide de (c.), on remarque que $g = (X-[010])*(X-[010]^2)*(X-[010]^3)*(X-[010]^4)$. *Donc les racines de* g *sont une suite de* 4 *puissances successives d'une racine primitive. Donc* g *donne un code* 4*-détecteur et* 2*-correcteur.*

f. Le code est $(7, 3)$ *car* g *est de degré* 3. *Donc le rendement est* $\frac{3}{7} = \frac{9}{21} \approx 0,43$ *en terme de bits ou d'éléments.*

3. En admettant qu'il existe un polynôme primitif de degré 4 sur $\mathbb{Z}/_{2\mathbb{Z}}[Y]$, donner les caractéristiques (n, k, d) d'un code de Reed-Solomon 2-correcteur sur $\mathbb{F}_{16}$. Quel est le rendement de ce code en terme de bits ?

Correction:

Le code sera $(15, 11, 5)$, *d'où un rendement de* $\frac{11}{15} \approx 0.73$.