Faculté des sciences exactes et naturelles
Licence Informatique / INFO0603 : Compression et cryptographie
Année universitaire 2014-2015
Pascal Mignot

Examen terminal du 12 mai 2015
Première session

Instructions :

- Le polycopié de cours est autorisé. Tout autre document (TDs, notes personnelles, livres, ...) est interdit.
- L'usage de la calculatrice est autorisé. Néanmoins, celle-ci ne devra pas avoir de fonctions communicantes (*i.e.* l'utilisation d'un téléphone portable comme calculatrice est interdit).
- Un résultat donné sans justification sera compté faux (*i.e.* joindre vos brouillons si nécessaire).

Notation : A_x signifie que le nombre A est écrit en base x. Par exemple, 101_2 est 101 en binaire (soit 5 en décimal), 12_{16} est le nombre hexadécimal 12 (soit 18 en décimal). Néanmoins, la base n'est pas précisée lorsqu'elle est évidente dans le contexte.

Exercice 1: cryptographie classique

Le message `OZSOXOQHOSGH` a été codé avec un chiffre mono-alphabétique par décalage. On rappelle les fréquences des lettres les plus fréquentes : e=17.5%, s=8,0%, a=7,6%, i=7,5%, t=7,2%, n=6,6%.

$\boxed{1}$ Donner deux méthodes permettant d'éviter de tester toutes les clefs.

$\boxed{2}$ Décoder le message.

$\boxed{3}$ Que peut-on en déduire sur la répartition des lettres dans le message ?

$\boxed{4}$ Quel est le principal inconvénient de ce cryptosystème ?

Exercice 2: chiffre de Vigenère

On souhaite effectuer la cryptanalyse du message suivant `SRSASTSNTRRSANTRRRTR` codé avec un chiffre de Vigenère qui utilise un alphabet de 5 symboles $\{A, N, R, S, T\}$. Ce message est constitué d'une suite de mots français utilisant les mêmes lettres (*i.e.* pour le code et le texte clair : $A = 0$, $N = 1$, ..., $T = 4$).

Les fréquences de lettres dans le texte clair sont $f(A) = 42,4\%$, $f(N) = 12,1\%$, $f(R) = 9,1\%$, $f(S) = 18,2\%$, $f(T) = 18,2\%$.

$\boxed{1}$ Expliquer pourquoi le calcul des fréquences de lettres n'est d'aucune aide dans ce cas.

$\boxed{2}$ Appliquer le test de Kasiski avec un motif de taille 3 afin de déterminer la taille de la clef.

$\boxed{3}$ En quoi la connaissance de la longueur de la clef va nous aider dans la cryptanalyse ?

$\boxed{4}$ Calculer l'histogramme des fréquences des lettres associées à la deuxième lettre de la clef.

$\boxed{5}$ En se basant uniquement sur la fréquence des lettres, quelle serait la seconde lettre de clef ?

$\boxed{6}$ Calculer les indices de coïncidence de Friedman manquants dans la table ci-dessous.

	A	N	R	S	T
k_0	0.3933	0.1127	0.0452	0.1938	0.2550
k_1	?	?	0.0524	0.3751	0.1666
k_2	0.2213	0.4099	0.0964	0.0918	0.1806

$\boxed{7}$ En déduire la clef.

$\boxed{8}$ Pourquoi le test de coïncidence de Friedman est-il en général plus fiable ?

$\boxed{9}$ Utiliser la clef afin de décoder le premier mot du message (ou au plus les 8 premiers caractères).

Exercice 3: cryptosystème produit d'un chiffre de Hill

On considère un chiffre de Hill S_1 dans $\mathbb{Z}_{26}$ représentant l'alphabet (a $= 0, \ldots,$ z $= 25$). On veut effectuer un codage par paquet de 2.

$\boxed{1}$ Quelle est la taille de l'espace des clefs ?

$\boxed{2}$ Quelle condition doit vérifier la clef K de manière à être inversible dans $\mathbb{Z}_{26}$?

$\boxed{3}$ Soit $K_1 = \left[\begin{smallmatrix} 11 & 19 \\ 9 & 20 \end{smallmatrix}\right]$. Calculer le chiffre de Hill du texte "zouave".

$\boxed{4}$ Calculer l'inverse de K_1 dans $\mathbb{Z}_{26}$.

$\boxed{5}$ Effectuer le déchiffrement du chiffre obtenu à la question 2 avec la matrice obtenue à la question précédente.

$\boxed{6}$ On considère maintenant un deuxième chiffre de Hill S_2 constitué du couple de matrices (K_2, K_2^{-1}). Donner les expressions des fonctions de chiffrement et de déchiffrement de $S_1 \times S_2$.

$\boxed{7}$ Le cryptosystème $S_1 \times S_2$ obtenu est-il commutatif ?

$\boxed{8}$ Le cryptosystème $S_1 \times S_2$ obtenu est-il idempotent ?

$\boxed{9}$ Soit $K_2 = \left[\begin{smallmatrix} 9 & 1 \\ 14 & 1 \end{smallmatrix}\right]$ et $K_2^{-1} = \left[\begin{smallmatrix} 5 & 21 \\ 8 & 19 \end{smallmatrix}\right]$ un autre couple de matrices de chiffrement/déchiffrement d'un chiffre de Hill. Donner le cryptosystème équivalent à $S_1 \times S_2$. On donnera les clefs de chiffrement et déchiffrement équivalentes.

Exercice 4: variation sur DES

On considère une méthode de chiffrement par bloc à deux étages, s'inspirant de DES, codant des blocs de 8 bits, et définie de la façon suivante :

- La fonction d'étage est un schéma de Feistel sur un bloc de 8 bits (découpé en deux blocs **L** et **R** de 4 bits chacuns).
- La fonction d'étage f est définie comme indiqué sur le schéma ci-dessous :
  - ◆ Application de la fonction d'expansion $E : \mathbb{B}_4 \to \mathbb{B}_6$ qui découpe l'entrée par bloc c_i de 2 bits et ajoute à bloc un bit de parité p_i. Par exemple, si $c_1c_2 = 00\,01$ est étendu comme $c_1p_1c_2p_2 = 00\underline{0}\,01\underline{1}$. On note $\mathbf{P} = E(\mathbf{L})$.
  - ◆ Incorporation de la sous-clef associée à l'étage qui effectue un ou-exclusif entre le résultat de l'expansion et la sous-clef ($\mathbf{B} = \mathbf{P} \oplus \mathbf{K}_1$).
  - ◆ Application d'une fonction de substitution $s : \mathbb{B}_3 \to \mathbb{B}_2$ qui substitue un bloc de 3 bits par un bloc de 2 bits. Si $b = x_2x_1x_0$ est un bloc de 3 bits x_i, s est définie par la table suivante :

$x_2x_0 =$	0	1	2	3
$x_1 = 0$	3	2	0	1
$x_1 = 1$	1	3	2	0

    Par exemple, si $x_2x_1x_0 = 011_2 = 3$, $x_1 = 1_2 = 1$ et $x_2x_0 = 01_2 = 1$, alors $s(x_2x_1x_0) = 3 = 11_2$.

    La substitution sur un bloc $\mathbf{B} = b_1b_2$ de 6 bits, on l'appliquera sur les paquets b_i de 3 bits. A savoir, $\mathbf{S} = s(\mathbf{B}) = s(b_1)s(b_2)$.
  - ◆ Une fonction de permutation $p : \mathbb{B}_4 \to \mathbb{B}_4$ par la table de permutation $\pi = (0, 2, 1, 3)$ où $\pi(i)$ indique la position du $i^{\text{ème}}$ bit après permutation. Par exemple, $p(0011_2) = 0101_2$. On a $\mathbf{P} = p(\mathbf{S})$.
- Pour la fonction de diversification de la clef $K = k_1k_2k_3k_4$ de 8 bits sous forme de 4 paquets k_i de 2 bits chacun. On construit les sous-clefs comme $K_1 = k_1k_2k_3$ et $K_2 = k_2k_3k_4$.

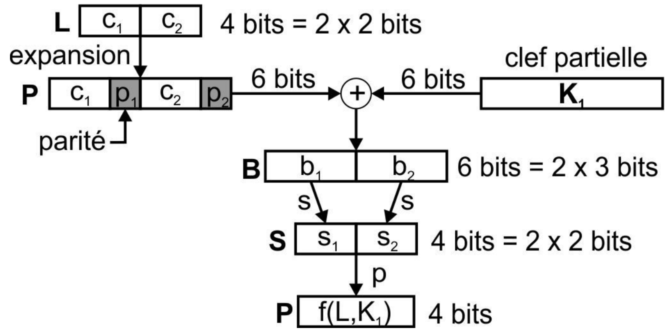

Figure 1 : Détail de la fonction d'étage

1. On souhaite coder des messages issus de l'alphabet $\{a, e, t, r\}$. Donner le plus petit codage de taille fixe possible associé à cet alphabet.
2. Donner l'ensemble des blocs de 8 bits produits avec le code de la question précédente pour le message suivant : "taratata". Les espaces seront supprimés et les accents ignorés. On complètera si nécessaire les blocs avec des 0.
3. Soit la clef $K = \mathrm{a}5_{16}$. Utiliser l'algorithme de diversification de la clef pour produire les deux sous-clefs.
4. Appliquer le chiffrement du premier étage sur les blocs du texte clair calculés à la question 2.
5. Appliquer le chiffrement du second étage sur les résultats du premier étage.
6. Donner alors le résultat du chiffrement des blocs du texte clair obtenu à la question 2 par un mode opérateur ECB.
7. Quel serait l'intérêt d'utiliser plutôt le mode CBC ?

Exercice 5: RSA

On souhaite construire la clef privée (n, b) et clef publique (p, q, a) du chiffre RSA, afin de coder un message avec des blocs B dont chaque bloc a une valeurs est strictement inférieure ou égale à 255.

1. Quelles sont les propriétés que doivent respecter n dans ce cas ?
2. Rappeler la définition d'un nombre premier, et montrer que si l'on veut tester la primalité d'un nombre a, il suffit qu'il ne soit divisible par aucun en entier entre 2 et $\sqrt{a}$.
3. On choisit $p = 7$ et $q = 37$. Vérifier de manière certaine que ces deux nombres sont premiers.
4. Parmi les valeurs suivantes $\{2, 13, 27, 108, 221\}$, choisir l'unique valeur de la clef publique a compatible avec les conditions requises pour son choix. Pour les autres valeurs, on indiquera pourquoi ce choix est inadéquat.
5. En déduire la valeur de clef privée.
6. Quelles sont les information minimales qui doivent être transmise à un correspondant afin de lui permettre d'effectuer un code RSA.
7. Les messages à transmettre sont codés avec l'alphabet $(\mathrm{a}, \mathrm{i}, \mathrm{n}, \mathrm{s})$. Combien de blocs ai-je besoin pour envoyer le message "assassin" ? Donner alors la valeur des blocs correspondant au texte clair.
8. Chiffrer les deux premiers bloc du message. Si $a > 10$, on utilisera l'algorithme d'exponentiation modulaire.
9. Pourquoi l'hasardisation du message transmis améliore la sécurité du codage ?
10. Déchiffrer les deux blocs du message 217 206. Si $b > 10$, on utilisera l'algorithme d'exponentiation modulaire.
11. Utiliser la méthode rapide de déchiffrement pour déchiffrer le premier bloc du message.
12. Un attaquant voudrait décoder le code RSA défini ci-dessus. Utiliser l'algorithme $\rho - 1$ de Pollard pour factoriser n.
13. Expliquer alors comment l'attaquant peut obtenir la clef de décodage b.

Exercice 6: Nombres premiers

1. Soit un nombre entier stocké sur n bits en supposant que le bit de poids le plus fort soit à 1. De combien de chiffres a-t-on besoin pour représenter ce nombre en base 10 ?
2. Appliquer la formule obtenue à la question précédente afin de déterminer le nombre p de chiffres a un nombre entier stocké sur 512 bits.
3. En déduire la borne inférieure du nombre de nombres premiers qui ont exactement 512 chiffres.
4. Tester la primalité de 37 en effectuant deux fois le test de Fermat (*i.e.* $k = 2$).
5. Tester la primalité de 37 avec le test de Miller-Rabin. On utilisera uniquement deux témoins.

6 Pour les tests de primalité effectué aux deux questions précédentes, peut-on dire quelque chose sur chance effective que 37 soit effectivement premier dans chacun des cas ?

7 La validité du test de primalité de Miller-Rabin dépend-il du nombre de chiffres nécessaires pour représenter le nombre ? Si oui, on précisera comment. Si non, on dira de quoi elle dépend.

Exercice 7: Sécurisation de RSA

On veut sécuriser un message RSA sur 8 bits. On donne dans ce but les fonctions suivantes :

- une fonction d'expansion $G : \mathbb{B}^2 \to \mathbb{B}^6$ définie comme $G(b_0b_1) = b_0b_1b_0b_1b_0b_1$.
- une fonction de condensation $H : \mathbb{B}^6 \to \mathbb{B}^2$ définie comme $H(b_0b_1b_2b_3b_4b_5) = b_0b_1 \oplus b_2b_3 \oplus b_4b_5$.

L'expéditeur veut envoyer le message $m = 011001$ en utilisant comme clef aléatoire $r = 10$.

1 Calculer le message hasardisé : $g = m \oplus G(r)$.

2 Calculer la signature du message hasardisé : $h = r \oplus H(g)$.

3 Calculer le message M à envoyer ($M = g \circ h$).

4 Après déchiffrement, le destinataire reçoit le message M. Vérifier qu'il peut bien reconstruire m à partir de M sans avoir la connaissance de r.

Exercice 8: Gain et information

A la suite d'un meurtre lors d'une réception dans un lieu clos, une liste de n suspects $(s_1, \ldots, s_n)$ est établie sur la base des personnes présentes. On note X la loi de probabilité associée ($\Pr[X = s_i]$ est la probabilité pour s_i d'être le meurtrier).

1 Dans un premier temps, nous n'avons aucune information sur qui peut être le meurtrier. Quelle est la probabilité pour chacun des suspects d'être le meurtrier ?

2 Calculer alors l'entropie de X.

3 A la suite d'une information, la police est presque sûre que le suspect s_1 est l'assassin, pendant que l'on a toujours aucune information sur les autres suspects. On a alors $\Pr[X = s_1] = 1 - \epsilon$.

4 Calculer alors l'entropie de X avec cette information supplémentaire.

5 A la suite d'une nouvelle information, s_1 est complètement écarté de la liste des suspects. Calculer alors les nouvelles probabilités de la loi X.

6 Calculer alors l'entropie de X avec cette information supplémentaire.

7 Que peut-on alors toujours en déduire sur l'ajout d'information et l'entropie ?

Exercice 9: Codage `LZW`

On veut effectuer un codage `LZW` avec un dictionnaire de taille 16 (à savoir, que l'on arrête de faire grossir le dictionnaire dès que sa taille atteint 16).

1 Donner le codage `LZW` de :

$$AAAABBCAAAABBCAAAABBCAAAABBCAAAABBCAAAABBCAAAAB$$

2 Donner le codage binaire minimum de la compression `LZW` obtenu à la question précédente en utilisant un codage binaire à taille fixe pour l'alphabet pour coder les symboles.

3 En déduire la taille du codage `LZW` sur le message M complet, ainsi que le nombre de bits par symboles atteinte par ce codage.