Master CSI 2 — 2006-2007

Crypto : DS du 5 mars 2007

Les exercices sont indépendants.

– EXERCICE 1. On considère le système de chiffrement donné par le tableau suivant, où l'espace des messages en clair est $\mathcal{M} = \{a, b, c, d\}$, l'espace des messages chiffrés $\mathcal{C} = \{1, 2, 3, 4\}$ et l'espace des clés est $\mathcal{K} = \{K_1, K_2, K_3\}$.

$\mathcal{K}^{\mathcal{M}}$	a	b	c	d
K_1	1	2	3	4
K_2	2	3	4	1
K_3	3	4	1	2

Les clés sont choisies, comme d'habitude, avec la loi de probabilité uniforme et indépendamment du message en clair. La loi de probabilité sur l'espace des messages en clair est donnée par $P(a) = 1/2$, $P(b) = P(c) = P(d) = 1/6$.

a) Calculer les probabilités conditionnelles $P(M = x \,|C = y)$ pour $x \in \mathcal{M}$ et $y \in \mathcal{C}$.

– **Solution.** On a

$$\begin{aligned} P(C=1) &= P(M=a, K=K_1) + P(M=c, K=K_3) \\ &\quad + P(M=d, K=K_2) \\ &= \frac{1}{2}\frac{1}{3} + \frac{1}{6}\frac{1}{3} + \frac{1}{6}\frac{1}{3} = \frac{1}{3}\frac{5}{6} \end{aligned}$$

D'où

$$P(M = a \,|\, C = 1) = \frac{\frac{1}{2}\frac{1}{3}}{\frac{1}{3}\frac{5}{6}} = \frac{3}{5}.$$

Et de manière analogue

$$P(M = b \,|\, C = 1) = 0, \quad P(M = c \,|\, C = 1) = (M = d \,|\, C = 1) = \frac{1}{5}.$$

$$P(M = a \,|\, C = 2) = \frac{3}{5}$$

$$P(M = c \,|\, C = 2) = 0, \quad P(M = b \,|\, C = 2) = (M = d \,|\, C = 2) = \frac{1}{5}.$$

$$P(M = a \,|\, C = 3) = \frac{3}{5}$$

$$P(M = d \,|\, C = 3) = 0, \quad P(M = b \,|\, C = 3) = (M = c \,|\, C = 3) = \frac{1}{5}.$$

$$P(M = a \,|\, C = 4) = 0$$

$$P(M = b \,|\, C = 4) = P(M = c \,|\, C = 4) = P(M = d \,|\, C = 4) = \frac{1}{3}.$$

b) Calculer $H(M)$, $H(C)$ et $H(M \mid C)$.

– **Solution.**

$$H(M) = \frac{1}{2}\log_2 2 + 3\frac{1}{6}\log_2 6 \approx 1,79.$$

$$H(C) = 3\frac{1}{3}\frac{5}{6}\log_2 \frac{3.6}{5} + \frac{1}{6}\log_2 6 \approx 1,97.$$

$$H(M \,|\, C) = 3(\frac{1}{6}\log_2 \frac{5}{3} + \frac{1}{9}\log_2 5) + \frac{1}{6}\log_2 3 \approx 1,41.$$

c) Qu'apprend-t-on sur la clé K après interception d'un cryptogramme ? En d'autres termes, comparez $\log_2 3$ et $H(K \mid C)$.

– **Solution.**

$$H(K \,|\, C) = H(K, C) - H(C) = H(K, M, C) - H(C) = H(K, M) - H(C)$$

Car K, C déterminent M et M, K déterminent C. Comme K et M sont indépendants, on en déduit

$$H(K \,|\, C) = H(K) + H(M) - H(C) = \log_2 3 + H(M) - H(C) \approx 1,41.$$

On constate que c'est une quantité inférieure à $\log_2 3 \approx 1,58$, ce qu'on obtiendrait si la connaissance de C n'apportait rigoureusement rien sur la clé.

On peut aussi constater que $H(K \,|\, C) = H(M \,|\, C)$. En effet, $H(M \,|\, C) = H(M, C) - H(C) = H(M, C, K) - H(C)$. Ceci parce que la connaissance conjointe de M et de C détermine K : attention, ce dernier point n'est pas vrai dans le cas général, c'est une particularité du présent tableau.

– Exercice 2. Vous devez chiffrer le résultat d'un sondage sur le deuxième tour de l'élection présidentielle. L'ensemble des messages en clair est donc un ensemble à deux éléments (candidat A ou candidat B). Vous devez mettre au point un système de chiffrement à $|\mathcal{K}| = 6$ clés secrètes et $|\mathcal{C}| = 3$ messages chiffrés possibles.

a) Proposez un système (vous pouvez le représenter par un tableau) à confidentialité parfaite, et de probabilité de substitution la plus faible que vous pouvez.

– **Solution.**

$\mathcal{K}^{\mathcal{M}}$	A	B
K_1	1	2
K_2	2	3
K_3	3	1
K_4	3	2
K_5	1	3
K_6	2	1

La confidentialité est parfaite. Par exemple,

$$P(M = A \,|\, C = 1) = \frac{P(M = A, C = 1)}{P(C = 1)} =$$

$$\frac{P(M = A, K \in \{K_1, K_5\})}{P(M = A, K \in \{K_1, K_5\}) + P(M = B, K \in \{K_3, K_6\})}$$
$$= P(M = A)\frac{2/6}{\frac{2}{6}(P(M = A) + P(M = B))}$$
$$= P(M = A).$$

b) Quelles sont les probabilités de substitution et d'imposture de votre système ?

– **Solution.** Chaque cryptogramme admet deux substitutions authentiques, elles sont équiprobables : $P_S = \frac{1}{2}$. Chaque cryptogramme a quatre chances sur six d'être authentique : $P_I = \frac{4}{6}$.

– Exercice 3. On considère la suite binaire $a = (a_i)$ qui commence ainsi :

$$1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1 \cdots$$

a) Trouver le plus petit générateur linéaire qui engendre cette séquence. Quelle est la période de la suite ainsi engendrée ?

– **Solution.**

$$a_i = a_{i-4} + a_{i-5}.$$

La période de la suite est 21.

b) Quel est le polynôme de rétroaction $h(X)$ de la suite a ? Le décomposer en facteurs irréductibles.

– **Solution.**

$$h(X) = X^5 + X + 1 = (X^2 + X + 1)(X^3 + X^2 + 1).$$

c) Combien y a-t-il de suites distinctes satisfaisant la récurrence linéaire trouvée en **a)** ? Trouver, parmi l'ensemble de suites satisfaisant cette récurrence, des suites dont les polynômes de rétroaction sont les facteurs irréductibles de $h(X)$.

– **Solution.**

L'ensemble des suites satisfaisant la récurrence est un espace vectoriel de dimension 5 sur $\mathbb{F}_2$: il comporte donc $2^5 = 32$ éléments.

Le polynôme de rétroaction $X^2 + X + 1$ est associé à la récurrence $a_i = a_{i-1} + a_{i-2}$ qui engendre la suite de période 3

$$011011011\cdots$$

ainsi que ses 2 décalées. On constate que ces 3 suites satisfont *aussi* la récurrence $a_i = a_{i-4} + a_{i-5}$ trouvée en **a)**.

De même, le polynôme de rétroaction $X^3 + X^2 + 1$ est associé à la récurrence $a_i = a_{i-1} + a_{i-3}$ qui engendre la suite de période 7

$$00111010011101001110100\cdots$$

ainsi que ses 6 décalées. Là encore, on constate que ces 7 suites satisfont aussi la récurrence de **a)**.

Remarque : on constate que l'ensemble des 32 suites solutions de **a)** se décompose en la suite nulle, les 3 suites de période 3 suscitées, les 7 suites de période 7 que l'on vient d'exhiber, et les 21 décalées de la suite que l'on est en train d'étudier.

d) Montrer que la suite a se décompose en la somme de suites, chacune desquelles a pour polynôme de rétroaction un facteur irréductible de $h(X)$.

– **Solution.** On constate que la suite a de l'exercice est la somme de la suite de période 3 et de la suite de période 7 exhibées à la question précédente. Pour trouver les bons décalages, il suffisait de les synchroniser sur les 5 premiers bits.

– Exercice 4. On considère un système de signature à clé secrète (MAC) de la forme $M \mapsto S = f_K(M)$ où la signature S prend ses valeurs dans $\{0,1\}^{64}$. Le signataire malveillant se prépare à signer un message de la forme :

«...début de message... Je reconnais devoir à M. Z la somme de 1000 Euros.»

Le signataire dispose juste d'une boite noire qui calcule f_K, sans avoir une connaissance explicite de la fonction. Le signataire a l'intention de répudier sa propre signature. Pour cela il prépare, en faisant des variantes sur le début du message, un premier ensemble A de m messages de ce type ainsi que leurs signatures. Puis il prépare un deuxième ensemble B de m messages de la forme

ainsi que les signatures correspondantes. Son but est de trouver un message M_A de l'ensemble A et un message M_B de l'ensemble B admettant la même signature, $f_K(M_A) = f_K(M_B)$. Il pourra ainsi signer M_A, puis plus tard prétendre qu'il n'a jamais signé M_A mais que c'est M_B qu'il a signé.

En faisant l'hypothèse que $f_K(x)$ se comporte comme une variable aléatoire à valeurs dans $\{0,1\}^{64}$, de loi uniforme, estimer la probabilité, en function de m, que le signataire frauduleux réalise son objectif.

– **Solution.**

L'ensemble des signatures des messages de l'ensemble A est un ensemble à approximativement m éléments. La probabilité qu'*aucune* signature d'un message de l'ensemble B n'appartienne à cet ensemble vaut donc

$$\left(1 - \frac{m}{2^{64}}\right)^m \approx e^{-m^2/2^{64}}.$$

La probabilité que le fraudeur réalise son objectif vaut donc approximativement

$$1 - e^{-m^2/2^{64}}.$$