Master CSI 2 2016-2017

Cryptanalyse — 4TCY902U

Responsable : G. Castagnos

Examen — mardi 13 décembre 2016

Durée 3h
Documents non autorisés
Les 3 exercices sont indépendants

1 Le chiffrement à flot *SG*

On considère un chiffrement à flot synchrone additif noté *SG* dans la suite. Ce système utilise deux LFSR : $\mathrm{LFSR_A}$ et $\mathrm{LFSR_S}$ de longueurs respectives ℓ_A et ℓ_S. Les états initiaux des deux LFSR notés K_A et K_S constituent la clef secrète. Les rétroactions de ces deux LFSR sont publiques, on choisit comme polynômes de rétroaction des polynômes primitifs. La production de la suite de bits chiffrante se fait ainsi :

On répète les trois points suivants jusqu'à produire assez de suite chiffrante :

1. Les $\mathrm{LFSR_A}$ et $\mathrm{LFSR_S}$ sont mis à jour, produisant deux bits a et s ;
2. Si le bit de sortie s du $\mathrm{LFSR_S}$ vaut 0 alors on ne fait rien ;
3. Sinon si $s = 1$, alors le bit de sortie de *SG* est le bit de sortie du $\mathrm{LFSR_A}$, a.

Par exemple, le $\mathrm{LFSR_A}$, de longueur $\ell_A = 3$, de polynôme de rétroaction $P_A = 1 + X + X^3$ initialisé par $K_A = [1, 0, 1]$ produit la suite de période $1, 0, 1, 0, 0, 1, 1$. Le $\mathrm{LFSR_S}$ de longueur $\ell_S = 2$ de polynôme de rétroaction $P_S = 1 + X + X^2$ initialisé par $K_S = [0, 1]$ produit la suite de période $0, 1, 1$. Avec ces choix, la suite produite par *SG* est $0, 1, 0, 1, 1, 0, 0, 0, \ldots$

Dans la suite on notera $(a_t)_{t\in\mathbf{N}}$, $(s_t)_{t\in\mathbf{N}}$ et $(z_t)_{t\in\mathbf{N}}$ les suites de sorties respectives du $\mathrm{LFSR_A}$, du $\mathrm{LFSR_S}$ et de *SG*. Soit $i \in \mathbf{N}$, on note k_i la position du $(i+1)$-ième « 1 » dans la suite $(s_t)_{t\in\mathbf{N}}$. Ainsi, on a $z_i = a_{k_i}$.

(a) On note T_A et T_S les périodes respectives des suites $(a_t)_{t\in\mathbf{N}}$ et $(s_t)_{t\in\mathbf{N}}$. Que valent, d'après le cours, T_A et T_S en fonction de ℓ_A et ℓ_S ?

(b) On note W_S le nombre de 1 dans une période de la suite produite par le $\mathrm{LFSR_S}$. Que vaut W_S en fonction de ℓ_S ? Démontrer ce résultat.

(c) Montrer que pour $i, j \in \mathbf{N}$, $z_{i+jW_S} = a_{k_i+jT_S}$. En déduire que la période de la suite $(z_t)_{t\in\mathbf{N}}$ divise $T_A W_S$.

(d) On suppose que $pgcd(T_A, T_S) = 1$. On admet que pour tout $k \in \mathbf{N}$, la suite $(a_{k+jT_S})_{j\in\mathbf{N}}$ est une m-suite produite par un LFSR de longueur ℓ_A. En déduire un polynôme de rétroaction pour la suite $(z_t)_{t\in\mathbf{N}}$ et une majoration de sa complexité linéaire.

(e) Question indépendante des questions précédentes. On suppose que l'on a accès aux bits de sortie de *SG*. Proposer une attaque détaillée sur ce générateur visant à retrouver la clef secrète. On précisera en particulier la complexité de l'attaque et le nombre de bits de sortie nécessaires. Que doit on prendre comme taille de paramètres pour se mettre à l'abri de cette attaque ?

2 Attaques sur un chiffrement par bloc

On note $\mathsf{DES}_K(X)$ la fonction de chiffrement du DES, qui prend en entrée un bloc de message X de 64 bits, une clef K de 56 bits et ressort un bloc de 64 bits. On construit un nouveau chiffrement par bloc en faisant 6 tours de schéma de Feistel avec 6 clefs de 56 bits indépendantes notées $K_1, K_2, \dots, K_6$, en utilisant la fonction DES comme fonction de tour. Plus précisément, soit M un bloc de 128 bits à chiffrer. On pose $M = L_0 \| R_0$ avec L_0 et R_0 de 64 bits, puis pour $i \in \{1, \dots, 6\}$,

$$L_i = R_{i-1}, \qquad R_i = L_{i-1} \oplus \mathsf{DES}_{K_i}(R_{i-1})$$

Le chiffré est $C = R_6 \| L_6$.

(a) On suppose connu un couple clair chiffré (M, C) par ce nouveau schéma de chiffrement. Donner une attaque utilisant $2^{169} = 2^{3\times 56+1}$ chiffrements DES visant à retrouver les clefs $K_1, \dots, K_6$. À l'issue de votre attaque combien reste-t-il en moyenne de candidats possibles pour ces clefs (en faisant l'hypothèse habituelle que les sorties des tours sont uniformément distribuées quand on fait varier les clefs) ? Combien faudrait il de couples clairs chiffrés pour n'avoir plus qu'un candidat en moyenne ? Décrire cette attaque utilisant plusieurs couples clairs chiffrés.

(b) Soit $\Delta \neq 0$ un bloc de 64 bits non nul. On considère deux messages clairs $M = L_0 \| R_0$ et $M^\star = L_0^\star \| R_0^\star$ tel que $L_0 \oplus L_0^\star = \Delta$ et $R_0 = R_0^\star$. On note $L_5 \| R_5$ (resp. $L_5^\star \| R_5^\star$) l'entrée du dernier tour lors du chiffrement de M (resp. de $M^\star$). Montrer qu'il est impossible d'avoir simultanément $L_5 = L_5^\star$ **et** $R_5 \oplus R_5^\star = \Delta$.

(c) Soit X un bloc fixé de 64 bits. On considère N messages clairs de la forme $M^{(i)} = L_0^{(i)} \| X$ pour $i = 1, \dots, N$ où $L_0^{(i)}$ est choisi au hasard de manière uniforme parmi les blocs de 64 bits. On note $C^{(i)} = R_6^{(i)} \| L_6^{(i)}$ le chiffré de $M^{(i)}$ pour $i = 1, \dots, N$. En supposant les $C^{(i)}$ uniformément distribués et indépendant des messages clairs, combien y aura-t-il en moyenne de paires d'éléments distincts i, j tels que $L_6^{(i)} \oplus L_6^{(j)} = L_0^{(i)} \oplus L_0^{(j)}$?

(d) Déduire des deux questions précédentes une attaque à messages clairs choisis visant à retrouver la clef K_6 de dernier tour. Combien de messages clairs choisis et combien de chiffrements DES sont nécessaires pour éliminer environ la moitié des clefs possibles pour K_6 ?

3 Attaques sur la construction de Merkle-Damgård

On rappelle la construction de Merkle-Damgård pour construire une fonction de hachage. On note f une fonction de compression de $\{0,1\}^{k+n}$ dans $\{0,1\}^n$, avec n et k deux entiers strictement positifs.

Soit IV un élément fixé de $\{0,1\}^n$. Soit m message à hacher. On commence par découper m en ℓ blocs de k bits, $m_0, \dots, m_{\ell-1}$ en « paddant » $m_{\ell-1}$ par 10000 ... pour obtenir un bloc de k bits. On rajoute un bloc m_ℓ dans lequel on code sur exactement k bits le nombre de bits du message m.

On note $z_0 = \text{IV}$ un bloc de n bits avec IV une valeur fixe, et pour $i = 0, \dots, \ell$, $z_{i+1} = f(m_i \| z_i)$. Le haché $h(m)$ de m est alors $h(m) = z_{\ell+1}$.

(a) Si f est résistante aux collisions, que peut on dire de h ? Rappeler la démonstration de ce résultat.

(b) Soit E un système de chiffrement par blocs qui utilise des clefs de k bits et des blocs de messages clairs de n bits et ressort des blocs de messages chiffrés de n bits. On note $\mathrm{E}_\mathrm{K}(\mathrm{X})$ le chiffré d'un message X avec la clef K. On construit la fonction de compression suivante : à une chaîne de k bits m et une chaîne de n bits z, on associe $f_\mathrm{E}(m\|z) := \mathrm{E}_m(z)$. Montrer que f_E n'est pas à sens unique.

(c) On utilise cette fonction de compression f_E pour définir une fonction de hachage h_E par la construction de Merkle-Damgård **sans utiliser le bloc supplémentaire contenant la longueur du message haché**. Montrer que si le vecteur d'initialisation z_0 peut être choisi librement alors h_E n'est pas à sens-unique.

(d) On suppose maintenant que le vecteur d'initialisation a une valeur fixée : $z_0 := \text{IV}$. On veut montrer que h_E n'est toujours pas à sens-unique. On prend N_1 blocs de messages aléatoires de k bits, notés m_i pour $i = 0, \dots, \mathrm{N}_1 - 1$ et on construit à partir de z_0, une liste L_1 contenant les valeurs $z_{i+1} = f_\mathrm{E}(m_i \| z_i)$ pour $i = 0, \dots, \mathrm{N}_1 - 1$. Soit z un haché par h_E. Montrer comment construire une pré-image de z en utilisant la liste L_1. Quel est la complexité en temps et en mémoire de cette attaque afin d'avoir une bonne probabilité de réussite ? (on peut ajuster N_1).

(e) On suppose dans toute la suite que l'on utilise une fonction de compression quelconque, f de $\{0,1\}^{k+n}$ dans $\{0,1\}^n$ pour définir une fonction de hachage h par la construction de Merkle-Damgård mais toujours **sans utiliser le bloc supplémentaire contenant la longueur du message haché**. Soit un message m constitué d'exactement N_1 blocs de longueur k. On note $m = m_0 \| m_1 \| \dots \| m_{\mathrm{N}_1 - 1}$. Montrer qu'à partir de m et de $z = h(m)$, il est possible, avec une bonne probabilité, de construire une seconde pré-image $m' \neq m$ tel que $h(m') = h(m)$ en $\mathscr{O}(2^n/\mathrm{N}_1)$ évaluations de f.

(f) Montrer comment construire avec bonne probabilité et $\mathscr{O}(2^{n/2})$ évaluations de f, des blocs de k bits $m_0, \dots m_6$ tels que $h(m_0\|m_1) = h(m_0\|m_4\|m_5\|m_6) = h(m_2\|m_3\|m_1) = h(m_2\|m_3\|m_4\|m_5\|m_6)$.

(g) Soit $t > 0$ un entier. Généraliser la méthode de la question précédente pour construire des messages $m^{(i)}$ constitués de $t + i$ blocs de k bits, pour $i = 0, \dots, 2^t - 1$ et tels que $h(m^{(0)}) = h(m^{(1)}) = \dots = h(m^{(2^t-1)})$. Quel est le coût de cette attaque ?

(h) En déduire une attaque pour adapter celle de la question (e) à une fonction de hachage construite par Merkle-Damgård avec le bloc supplémentaire contenant la longueur du message haché.