Master CSI 2 2017-1018

Cryptologie Avancée — 4TCY903U

Responsables : G. Castagnos – G. Zémor

Examen — 18 décembre 2017

Durée 3h — Documents non autorisés

Partie G. Castagnos

Exercice 1. On considère la variante suivante du chiffrement Elgamal.

— Soit k un paramètre de sécurité. Soit GenDH un algorithme polynomial qui prend en entrée 1^k et retourne la description d'un groupe cyclique G son ordre q premier de k bits et deux générateurs distincts g et h.

— L'algorithme KeyGen appelle GenDH puis choisit s, t aléatoires avec probabilité uniforme dans $\mathbf{Z}/q\mathbf{Z}$ et calcule $f = g^s h^t$. KeyGen retourne $pk = (G, q, g, h, f)$ et $sk = (s, t)$.

— L'algorithme $\mathsf{Encrypt}$ sur l'entrée (pk, m) avec $m \in G$ choisit r uniformément dans $\mathbf{Z}/q\mathbf{Z}$ et retourne $c = (g^r, h^r, f^r m)$.

(a) Donner un algorithme de déchiffrement.

(b) Rappeler les définitions de la notion de sécurité $\mathsf{IND-CPA}$ et de l'hypothèse DDH.

(c) Soit (X, Y, Z) un triplet Diffie-Hellman et m un message clair. Montrer qu'il est possible de construire une clef publique pk et un chiffré c de m bien distribués tels que X joue le rôle de h dans la clef publique et Y et Z constituent les deux premiers éléments de c.

(d) Soit (X, Y, Z) un triplet d'éléments aléatoires indépendants de G. Montrer que pk et c, construits à partir de (X, Y, Z) et de m comme à la question précédente, ne fournissent aucune information sur m même à un adversaire tout puissant (Indication : on pourra raisonner sur les valeurs des logarithmes discrets en base g des éléments auxquels l'attaquant a accès).

(e) Conclure sur la sécurité $\mathsf{IND-CPA}$ de ce chiffrement.

Exercice 2. On définit un protocole d'échange de clef $\mathcal{P}$ entre Alice et Bob à deux passes et sa sécurité. Soit k un paramètre de sécurité, on suppose connu par tous un groupe cyclique G, son ordre q avec $|q| = k$, et un générateur g. On suppose que toutes les quantités échangées et la clef secrète établie sont des éléments de G. Le protocole se déroule ainsi :

1. Bob à partir de G, q, g produit un état s_B et un élément $X \in G$ qu'il envoie à Alice ;
2. Alice à partir de G, q, g produit un état s_A et un élément $Y \in G$ qu'elle envoie à Bob ;
3. Alice calcule à partir de s_A et X une clef $K_A \in G$. De même Bob calcule à partir de s_B et Y une clef $K_B \in G$.

Le protocole $\mathcal{P}$ est correct si $K_A = K_B =: K$. De plus, il est sûr si un adversaire $\mathcal{A}$ observant les données échangées par Alice et Bob ne peut distinguer la clef K établie d'un élément de G aléatoire. Plus formellement on définit $\mathbf{Exp}_{\mathcal{P},k}(\mathcal{A})$:

1. Sous l'entrée 1^k Alice et Bob exécute le protocole $\mathcal{P}$. Ceci produit les quantités échangées X et Y et la clef K éléments de G d'ordre q avec $|q| = k$;
2. on choisit un bit aléatoire $b^\star \xleftarrow{\$} \{0,1\}$. Si $b^\star = 1$ alors $Z := K$ sinon Z est tiré uniformément dans G ;
3. on donne (G, q, g, X, Y, Z) à $\mathcal{A}$ qui sort un bit b ;
4. la sortie de l'expérience est 1 si $b = b^\star$ et 0 sinon.

Le protocole $\mathcal{P}$ est sûr si pour tout algorithme polynomial probabiliste $\mathcal{A}$, l'avantage de $\mathcal{A}$, défini par $\mathbf{Adv}_{\mathcal{P},k}(\mathcal{A}) = \left|\Pr(\mathbf{Exp}_{\mathcal{P},k}(\mathcal{A}) = 1) - \frac{1}{2}\right|$, est négligeable.

(a) Que sont $s_B, s_A, X, Y, K_A, K_B, K$ dans le cas du protocole de Diffie-Hellman ? Quelle est l'hypothèse qui assure que le protocole est sûr ?

(b) Montrer qu'à partir de n'importe quel protocole d'échange de clef $\mathcal{P}$ à deux passes correct, on peut construire un schéma de chiffrement à clef publique Π. Montrer que si $\mathcal{P}$ est sûr alors Π est sémantiquement sûr pour des attaques à clairs choisis.

Exercice 3. Soit $\Pi = (\mathsf{KeyGen}, \mathsf{Sign}, \mathsf{Verify})$ un schéma de signature : sous l'entrée 1^k, KeyGen retourne un couple (pk, sk) clef publique, clef privée. Sous l'entrée sk, m, Sign retourne une signature σ du message m. Sous l'entrée pk, m, σ, Verify retourne un bit b. Si $b = 1$ (resp. $b = 0$), on dit que la signature est valide (resp. invalide). On a de plus que $\mathsf{Verify}(pk, m, \mathsf{Sign}(sk, m)) = 1$.

On note $\mathcal{O}$ un oracle de signature pour Π : sous la requête m, $\mathcal{O}$ retourne une signature valide σ de m. On dit que Π est sûr si tout $\mathcal{A}$ polynomial probabiliste ne peut retourner un couple $(m^\star, \sigma^\star)$ avec $\sigma^\star$ signature valide de $m^\star$ qu'avec probabilité négligeable, en ayant accès à la clef publique et à $\mathcal{O}$ pour des requêtes $m \neq m^\star$.

(a) Donner une définition formelle de cette notion de sécurité en utilisant une expérience.

On considère dans les deux questions suivantes, le schéma de signature RSA basique : KeyGen retourne $pk = (N, e)$ et $sk = d$ avec N un module RSA de k bits et $ed \equiv 1 \pmod{\varphi(N)}$. Étant donné un message $m \in (\mathbf{Z}/N\mathbf{Z})^\times$, $\mathsf{Sign}(sk, m)$ retourne $m^d \mod N$ et $\mathsf{Verify}(pk, m, \sigma)$ vérifie si $m \equiv \sigma^e \pmod{N}$.

(b) Décrire un attaquant contre la sécurité de ce schéma qui n'utilise aucune requête à $\mathcal{O}$ et qui retourne un couple $(m^\star, \sigma^\star)$ valide avec probabilité 1.

(c) Soit $m^\star \in (\mathbf{Z}/N\mathbf{Z})^\times$ un message fixé. Décrire un attaquant qui utilise deux requêtes à $\mathcal{O}$ et qui retourne un couple $(m^\star, \sigma^\star)$ valide avec probabilité 1.

Dans toute la suite de l'exercice, on considère le schéma de signature suivant. Soit $\mathcal{H}$ un oracle aléatoire $\mathcal{H} : \{0,1\}^* \to (\mathbf{Z}/N\mathbf{Z})^\times$. L'algorithme KeyGen est le même que précédemment. Étant donné un message $m \in \{0,1\}^*$, $\mathsf{Sign}(sk, m)$ retourne $\mathcal{H}(m)^d \mod N$ et $\mathsf{Verify}(pk, m, \sigma)$ vérifie si $\mathcal{H}(m) \equiv \sigma^e \pmod{N}$.

(d) Les attaques de **(b)** et **(c)** fonctionnent elles sur ce schéma ?

(e) Dans cette question on considère un attaquant $\mathcal{A}$ contre ce schéma qui n'utilise aucune requête à $\mathcal{O}$ et qui retourne un couple $(m^\star, \sigma^\star)$ valide avec probabilité ϵ. Si ϵ est non négligeable, pourquoi peut on supposer que $\mathcal{A}$ a demandé la valeur $\mathcal{H}(m^\star)$ à l'oracle aléatoire ? On suppose que $\mathcal{A}$ fait q requêtes distinctes à l'oracle aléatoire. Construire à partir de $\mathcal{A}$ un algorithme résolvant le problème RSA avec probabilité ϵ/q. Que peut on en conclure ?

(f) Supposons que $\mathcal{A}$ fasse une requête à l'oracle de signature. Montrer que sans connaître d, il est possible de simuler conjointement l'oracle aléatoire et l'oracle de signature de manière à fournir une réponse σ parfaite à $\mathcal{A}$.

(g) En déduire que ce schéma de signature est sûr sous l'hypothèse RSA.

Partie G. Zémor

Exercice 4. Soit un code binaire C de longueur n,de dimension $n/2$ et de distance minimale d.

Décrire un algorithme probabiliste qui trouve un mot de code de poids d en adaptant la méthode de décodage par ensemble d'information. Estimer le temps de calcul nécessaire pour trouver ce mot sans tenir compte du temps requis pour mettre une matrice sous forme diagonale.

Exercice 5. Tous les vecteurs sont binaires. Soit $\mathbf{A}$ une matrice aléatoire uniforme à k_A lignes et n colonnes. Soit $\mathbf{E}$ une matrice aléatoire à n colonnes et k_E lignes, et dont toutes les lignes sont choisies indépendamment et uniformément parmi les lignes de poids t. On considère la matrice

$$\mathbf{H} = \begin{bmatrix} \mathbf{A} \\ \mathbf{Y} = \mathbf{SA} + \mathbf{E} \end{bmatrix}$$

où $\mathbf{S}$ est une matrice $k_E \times k_A$. La matrice $\mathbf{H}$ est donc $(k_A + k_E) \times n$. Soit C le code de matrice de parité $\mathbf{H}$, et soit $\mathbf{G}$ une matrice génératrice de C. On supposera C de dimension k, la matrice $\mathbf{G}$ sera donc $k \times n$ avec $k = n - k_A - k_E$. Par exemple on peut prendre $k = k_A = k_E = n/3$. Soit $\mathcal{M} = \{0,1\}^k$. On considère un système de chiffrement à clé publique défini sur l'ensemble des clairs $\mathcal{M}$, dont la clé publique est $\mathbf{G}$ et dont la clé secrète est $\mathbf{E}$ (ou $\mathbf{S}$). Le chiffrement d'un message $\mathbf{m} \in \mathcal{M}$ consiste en la transformation

$$\mathbf{m} \mapsto \mathbf{mG} + \mathbf{e}$$

où $\mathbf{e}$ est un vecteur de petit poids t aléatoire.

(a) Montrer que $\mathbf{E}(\mathbf{mG} + \mathbf{e})^T = \mathbf{Ee}^T$. En déduire un algorithme de déchiffrement fondé sur le décodage des codes MDPC (Moderate Density Parity-Check). Quelle condition sur t doit être réalisée pour que le déchiffrement fonctionne ?

(b) Pourquoi le système n'est-il pas sémantiquement sûr (IND-CPA) ?

(c) Pour le rendre sémantiquement sûr, on réduit l'ensemble des clairs de $\mathcal{M} = \{0,1\}k$ à $\{0,1\}^\ell$, $\ell < k$. Le chiffrement devient maintenant

$$\mathbf{m} \mapsto [\mathbf{m}||\mathbf{r}]\mathbf{G} + \mathbf{e}$$

où $||$ désigne la concaténation et où $\mathbf{r}$ est un vecteur uniforme de longueur $k - \ell$. Quelle justification informelle peut-on donner de la sécurité sémantique ?

(d) Démontrer rigoureusement la sécurité sémantique lorsque l'espace des clairs est réduit à deux messages $\mathbf{m}_0$ et $\mathbf{m}_1$. On pourra argumenter que
- L'attaque doit continuer à fonctionner lorsque la matrice $\mathbf{H}$ est remplacée par une matrice aléatoire uniforme, et donc que la matrice $\mathbf{G}$ peut être prise aléatoire uniforme.
- Puis que l'attaque peut permettre de distinguer un vecteur uniforme de $\{0,1\}^n$ d'un vecteur de la forme $\mathbf{x}+\mathbf{e}$ où $\mathbf{e}$ est de poids t et où $\mathbf{x}$ est un mot uniformément choisi dans un code linéaire aléatoire de longueur n et dont vous donnerez la dimension.

Exercice 6. Soit $\mathbf{H}$ une matrice de parité $r \times n$ d'un code binaire de longueur n et de dimension $k = n - r$. On définit la fonction syndrome associée

$$\begin{aligned} \mathbb{F}_2^n &\to \mathbb{F}_2^r \\ \mathbf{x} &\mapsto \sigma(\mathbf{x}) = \mathbf{H}\mathbf{x}^T. \end{aligned}$$

On souhaite établir un protocole qui prend en entrée $\mathbf{H}$, $\mathbf{s} \in \mathbb{F}_2^r$, et un entier positif w, et qui démontre, sans divulgation, la connaissance d'un vecteur $\mathbf{x} \in \mathbb{F}_2^n$ de poids w, et tel que $\sigma(\mathbf{x}) = \mathbf{s}^T$.

Dans une première étape, le prouveur permute aléatoirement les colonnes de la matrice $\mathbf{H}$ pour former une matrice $\mathbf{K}$. Puis il forme La matrice $\mathbf{M} = \mathbf{A}\mathbf{K}$ où $\mathbf{M}$ est une matrice $r \times r$ inversible aléatoire. En d'autres termes, le prouveur choisit r combinaisons linéaires arbitraires et linéairement indépendantes des lignes de $\mathbf{K}$, pour former les lignes de $\mathbf{M}$.

(a) Montrer qu'il existe un vecteur $\mathbf{x} \in \mathbb{F}_2^n$ de poids w tel que $\sigma(\mathbf{x}) = \mathbf{s}^T$ si et seulement s'il existe $\mathbf{x} \in \mathbb{F}_2^n$ de poids w tel que $\mathbf{M}\mathbf{x}^T = \mathbf{A}\mathbf{s}^T$.

(b) Le prouveur s'engage sur les colonnes de $\mathbf{M}$, c'est-à-dire qu'il donne au vérificateur $h(\mathbf{M}_j), j = 1 \dots n$, où $\mathbf{M}_j$ désigne la j-ième colonne de $\mathbf{M}$ et où h est une fonction à sens unique. Alternativement il donne au vérificateur des enveloppes contenant les $\mathbf{M}_j$. Il donne également $h(\mathbf{A}\mathbf{s}^T)$ ainsi que $h(\pi)$ où π est la permutation des colonnes qui transforme $\mathbf{H}$ en $\mathbf{K}$.

Le vérificateur renvoie un défi ε qui vaut 0 ou 1. Si $\varepsilon = 0$, le prouveur révèle π, $\mathbf{M}$, et $\mathbf{A}$ (ouvre les enveloppes). Que doit-il révéler lorsque $\varepsilon = 1$?

(c) Démontrer que le protocole est complet, valide, et sans divulgation. S'il n'est pas vrai qu'il existe un vecteur $\mathbf{x}$ de poids w tel que $\sigma(\mathbf{x}) = \mathbf{s}^T$, quelle est la plus petite probabilité d'échec du protocole ?