Master CSI 2 2020-2021

Cryptologie Avancée — 4TCY903U

Responsables : G. Castagnos – G. Zémor

Examen — 14 décembre 2020

Durée 3h — Documents non autorisés
Répondre sur deux copies séparées

Partie G. Castagnos

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

— Soit k un paramètre de sécurité. Soit GenDL 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 GenDL 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 dans G 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. Soit k un entier et $\Pi = (\mathsf{KeyGen}, \mathsf{Encrypt}, \mathsf{Decrypt})$ un schéma de chiffrement asymétrique. On suppose que l'espace des messages clairs est $\mathcal{M} := \{0,1\}^{2k}$. On suppose de plus que le chiffrement d'un message $m \in \mathcal{M}$ avec la clef publique pk consiste à prendre $r \in \{0,1\}^k$

avec distribution uniforme puis poser $c = E_{pk}(m, r)$ où E_{pk} est une fonction de $\{0,1\}^{2k} \times \{0,1\}^k$ à valeurs dans l'espace des chiffrés.

Soit $\mathcal{H} : \{0,1\}^{2k} \to \{0,1\}^k$ un oracle aléatoire. À partir de Π, on construit un nouveau schéma de chiffrement $\Pi' = (\mathsf{KeyGen}', \mathsf{Encrypt}', \mathsf{Decrypt}')$ dans le modèle de l'oracle aléatoire. L'algorithme de génération de clefs est inchangé : $\mathsf{KeyGen}' := \mathsf{KeyGen}$. L'algorithme de chiffrement $\mathsf{Encrypt}'$ est défini comme suit. L'espace des messages clairs est $\mathcal{M}' := \{0,1\}^k$. Soit $m \in \{0,1\}^k$ à chiffrer avec la clef publique pk. On tire $t \in \{0,1\}^k$ avec probabilité uniforme et on pose $c = E_{pk}(m||t, \mathcal{H}(m||t))$ où $||$ désigne la concaténation des chaînes de bits.

(a) Donner la description d'un algorithme de déchiffrement $\mathsf{Decrypt}'$ pour Π' qui vérifie que le chiffré est bien formé avant de retourner le message clair.

Dans la suite on note $\mathcal{A}' = (\mathcal{A}'_1, \mathcal{A}'_2)$ un attaquant polynomial probabiliste contre la notion de sécurité $\mathsf{IND-CPA}$ du schéma Π' dans le modèle de l'oracle aléatoire. À partir de $\mathcal{A}'$, on construit $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ un attaquant contre la notion de sécurité $\mathsf{IND-CPA}$ du schéma Π comme suit.

$\underline{\mathcal{A}_1(pk)}$

1. $(m_0, m_1, s) \leftarrow \mathcal{A}'_1(pk)$
2. $t_0, t_1 \xleftarrow{\$} \{0,1\}^k$
3. Retourne $(m_0||t_0, m_1||t_1, s)$

$\underline{\mathcal{A}_2(c^\star, s)}$

1. $b' \leftarrow \mathcal{A}'_2(c^\star, s)$
2. Retourne b'

On note $b^\star$ le bit choisi lors de l'expérience $\mathsf{IND-CPA}$ que joue $\mathcal{A}$. Soit E l'événement « $\mathcal{A}'_2$ demande $m_{b^\star}||t_{b^\star}$ à son oracle aléatoire » et F l'événement « $\mathcal{A}'_2$ demande $m_{\bar{b}^\star}||t_{\bar{b}^\star}$ à son oracle aléatoire ». Ici $\bar{b}^\star$ désigne le complémentaire du bit $b^\star$.

(b) Compléter la description de $\mathcal{A}$ pour répondre aux requêtes faites par $\mathcal{A}'$ à son oracle aléatoire. De plus comment $\mathcal{A}$ peut-il utiliser ces requêtes pour résoudre l'expérience $\mathsf{IND-CPA}$ avec un meilleur avantage ?

(c) Que valent les probabilités suivantes $\Pr(\mathbf{Exp}^{\mathsf{IND-CPA}}_{\Pi,k}(\mathcal{A}) = 1|E)$, $\Pr(\mathbf{Exp}^{\mathsf{IND-CPA}}_{\Pi,k}(\mathcal{A}) = 1|(\bar{E} \text{ et } F))$ et $\Pr(\mathbf{Exp}^{\mathsf{IND-CPA}}_{\Pi,k}(\mathcal{A}) = 1|(\bar{E} \text{ et } \bar{F}))$?

(d) En déduire que

$$\Pr(\mathbf{Exp}^{\mathsf{IND-CPA}}_{\Pi,k}(\mathcal{A}) = 1) - \Pr(\mathbf{Exp}^{\mathsf{IND-CPA}}_{\Pi',k}(\mathcal{A}') = 1) \geqslant -\Pr(\bar{E} \text{ et } F).$$

Conclure.

(e) Adapter ce qui précède pour montrer que Π' est $\mathsf{IND-CCA2}$ dans le modèle de l'oracle aléatoire si Π est $\mathsf{IND-CPA}$.

Partie G. Zémor

Exercice 3. La loi $C(t)$ produit un couple $(\mathbf{A},\mathbf{y})$ où $\mathbf{A}$ est une matrice binaire $k\times n$ aléatoire uniforme, et $\mathbf{y}=\mathbf{sA}+\boldsymbol{\varepsilon}$ où $\mathbf{s}$ est choisi uniformément dans $\mathbb{F}_2^k$, $\boldsymbol{\varepsilon}$ est choisi uniformément dans l'ensemble des vecteurs de $\mathbb{F}_2^n$ de poids t, et $\mathbf{A},\mathbf{s},\boldsymbol{\varepsilon}$ sont indépendants.

La loi $H(t)$ produit un couple $(\mathbf{H},\boldsymbol{\sigma})$ où $\mathbf{H}$ est une matrice binaire $(k+1)\times n$ aléatoire uniforme et $\boldsymbol{\sigma}=\mathbf{He}^{\intercal}$ où $\mathbf{e}$ est choisi indépendamment de $\mathbf{H}$ et uniformément dans l'ensemble des vecteurs de $\mathbb{F}_2^n$ de poids t.

La loi $U_{k\times n,n}$ (uniforme) produit un couple $(\mathbf{A},\mathbf{y})$ où $\mathbf{A}$ et $\mathbf{y}$ sont choisi uniformément et indépendamment dans $\mathbb{F}_2^{k\times n}$ et $\mathbb{F}_2^n$ respectivement.

les entiers n,k,t sont des paramètres pour lesquels on fait l'hypothèse qu'aucun algorithme de complexité raisonnable n'est capable de distinguer si le couple $\mathbf{A},\mathbf{y}$ est produit suivant la loi $C(t)$ ou la loi U. De même on suppose que les lois $H(t)$ et $U_{(k+1)\times n,k+1}$ sont indistinguables.

(a) On définit la loi DH qui produit des quadruplets $(\mathbf{A},\mathbf{y},\mathbf{Ae}^{\intercal},\mathbf{ye}^{\intercal})$ où $(\mathbf{A},\mathbf{y})$ est produit suivant la loi $C(t)$ et $\mathbf{e}\in\mathbb{F}_2^n$ est indépendant de $(\mathbf{A},\mathbf{y})$ et uniforme dans l'ensemble des vecteurs de poids t.

Démontrer que la loi DH est indistinguable d'un quadruplet $(\mathbf{A},\mathbf{y},\mathbf{x},b)$ uniforme, c'est-à-dire où $\mathbf{A},\mathbf{y},\mathbf{x},b$ sont indépendants, uniformes dans leurs espaces respectifs, soit $\mathbb{F}_2^{k\times n},\mathbb{F}_2^n,\mathbb{F}_2^k,\mathbb{F}_2$.

(b) En déduire un système de chiffrement dont la clé publique est donnée par le couple $\mathbf{A},\mathbf{y}=\mathbf{sA}+\boldsymbol{\varepsilon}$ et dont l'espace des clairs est $\mathcal{M}=\{0,1\}$. Expliquer le fonctionnement du chiffrement et du déchiffrement. Démontrer la sécurité du système.

Exercice 4. On considère un code C de longueur n, donné par une matrice de parité $\mathbf{H}$ à $n/2$ lignes. On rappelle que mettre $\mathbf{H}$ sous forme systématique consiste à trouver une matrice de parité $\mathbf{H}'$ du même code C, dont une sous-matrice $(\mathbf{H}'_{ij})_{\substack{1\leq i\leq n/2\\ j\in J}}$ où $|J|=n/2$, est la matrice identité $n/2\times n/2$.

On suppose que l'on a mis $\mathbf{H}$ sous forme systématique suivant une partition $[1,n]=J\cup\overline{J}$ aléatoire des coordonnées. Soit $\mathbf{x}$ un mot de poids d du code C.

(a) Quel est la probabilité que $|\mathrm{supp}(\mathbf{x})\cap J|=d-1$ et $|\mathrm{supp}(\mathbf{x})\cap\overline{J}|=1$? Comment peut-on reconnaître si on est dans un tel cas de figure ?

(b) Quel est approximativement le coût de chercher un mot de poids d de C de cette manière ?

(c) Sachant que systématiser la matrice $\mathbf{H}$ coûte de l'ordre de n^2 additions de vecteurs (n-uples), est-ce plus ou moins avantageux de chercher des partitions $(J,\overline{J})$ telles que $|\mathrm{supp}(\mathbf{x})\cap J|=d-2$? Telles que $|\mathrm{supp}(\mathbf{x})\cap J|=d-3$?

Exercice 5. Tous les vecteurs sont binaires. Soit $\mathbf{A}$ une matrice aléatoire uniforme à n colonnes et $n/3$ lignes. Soit $\mathbf{E}$ une matrice aléatoire à n colonnes et $n/3$ lignes, et dont toutes les lignes sont choisies indépendamment et uniformément parmi les lignes de poids t, et indépendamment de $\mathbf{A}$. On considère la matrice

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

où $\mathbf{S}$ est une matrice aléatoire uniforme $n/3 \times n/3$. La matrice $\mathbf{H}$ est donc $2n/3 \times n$. Soit C le code de matrice de parité $\mathbf{H}$, et soit $\mathbf{G}$ une matrice génératrice de C, typiquement de dimension $k = n/3$. 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 $\mathbf{E}(\mathbf{mG} + \mathbf{e})^\intercal = \mathbf{E}\mathbf{e}^\intercal$. 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) Montrer que le système n'est pas sémantiquement sûr (IND-CPA).

(c) Sous l'hypothèse où le message clair $\mathbf{m}$ est choisi uniformément dans $\mathcal{M}$, justifier la sécurité du système de chiffrement.