Master CSI 2 2019-2020

Cryptologie Avancée — 4TCY903U

Responsables : G. Castagnos – G. Zémor

Examen — 16 décembre 2019

Durée 3h — Documents non autorisés

Partie G. Castagnos

Exercice 1. Soit $\Pi = (\mathsf{KeyGen}, \mathsf{Encrypt}, \mathsf{Decrypt})$ un schéma de chiffrement asymétrique. On note $\mathcal{M}$ l'espace des messages clairs pour Π. Soit $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$ un algorithme attaquant Π et k un paramètre de sécurité. On définit l'expérience *real of random* CPA, $\mathbf{Exp}^{\mathsf{RR-CPA}}_{\Pi,k}(\mathcal{B})$ comme suit :

1. On lance l'algorithme $\mathsf{KeyGen}(1^k)$ pour obtenir les clefs (pk, sk)
2. $\mathcal{B}_1$ reçoit pk et retourne (m_0, s) un message clair de $\mathcal{M}$ et un état s
3. On choisit un bit aléatoire avec équiprobabilité $b^\star \xleftarrow{\$} \{0,1\}$ et un message aléatoire $m_1 \xleftarrow{\$} \mathcal{M}$
4. On calcule $c^\star$ un chiffré de $m_{b^\star}$: $c^\star \leftarrow \mathsf{Encrypt}(pk, m_{b^\star})$
5. On donne $(s, c^\star)$ à $\mathcal{B}_2$ qui sort un bit b
6. La sortie de l'expérience est 1 si $b = b^\star$ et 0 sinon.

L'avantage de l'attaquant $\mathcal{B}$ est défini par

$$\mathbf{Adv}^{\mathsf{RR-CPA}}_{\Pi,k}(\mathcal{B}) = \left| \Pr\left(\mathbf{Exp}^{\mathsf{RR-CPA}}_{\Pi,k}(\mathcal{B}) = 1 \right) - \frac{1}{2} \right|.$$

Le schéma Π est dit sûr au sens $\mathsf{RR-CPA}$ si pour tout algorithme polynomial probabiliste $\mathcal{B}$ cet avantage est négligeable.

(a) Rappeler la définition de la notion de sécurité $\mathsf{IND-CPA}$ pour le schéma Π, en particulier décrire l'expérience $\mathsf{IND-CPA}$.

(b) Soit $\mathcal{B}$ un attaquant polynomial probabiliste contre la notion $\mathsf{RR-CPA}$ pour le schéma Π. Construire à partir de $\mathcal{B}$ un algorithme polynomial probabiliste $\mathcal{A}$ attaquant la notion $\mathsf{IND-CPA}$ ayant le même avantage que $\mathcal{B}$. Conclure.

(c) Réciproquement, soit $\mathcal{A}$ un attaquant polynomial probabiliste contre la notion $\mathsf{IND-CPA}$ pour le schéma Π avec avantage ϵ. Construire à partir de $\mathcal{A}$ un algorithme $\mathcal{B}$ attaquant la notion $\mathsf{RR-CPA}$ ayant un avantage $\epsilon/2$ (bien détailler le calcul de cet avantage). Conclure.

(d) Soit $\mathcal{C}$ l'espace des chiffrés de Π. On suppose dans cette question que prendre un message aléatoire $m \xleftarrow{\$} \mathcal{M}$ puis un chiffré $c \leftarrow \mathsf{Encrypt}(pk, m)$ de m est équivalent à prendre directement $c \xleftarrow{\$} \mathcal{C}$. On suppose de plus que $(\mathcal{M}, +)$ est un groupe et que Π est homomorphe additif. Montrer que Π est $\mathsf{IND} - \mathsf{CPA}$ si et seulement si il n'existe pas d'algorithme polynomial distinguant un chiffré aléatoire d'un chiffré de 0 avec avantage non négligeable.

Exercice 2. Une variante du chiffrement de Paillier $\mathsf{IND} - \mathsf{CCA2}$

On considère la variante suivante du chiffrement de Paillier, définie dans le modèle de l'oracle aléatoire. On désigne par $||$ la concaténation de deux chaînes de bits.

Soit un algorithme polynomial probabiliste Gen, qui prend en entrée k et qui retourne (n, p, q) où n est le produit de deux nombres premiers distincts de $k+1$ bits p et q, tel que $\gcd(n, \varphi(n)) = 1$.

Soit $\mathcal{H} : \{0,1\}^{2k} \mapsto (\mathbf{Z}/n\mathbf{Z})^\times$ un oracle aléatoire.

L'algorithme de génération de clefs KeyGen appelle Gen, retourne la clef publique $pk = n$ et la clef privée $sk = \varphi(n)$.

Soit $m \in \{0,1\}^k$. L'algorithme de chiffrement Encrypt sur l'entrée (pk, m) choisit $r \xleftarrow{\$} \{0,1\}^k$. Puis il pose M égal à l'entier dont la représentation binaire est la concaténation $m||r$ et $R = \mathcal{H}(m||r)$. Enfin il ressort $c \equiv (1+n)^M R^n \pmod{n^2}$.

L'algorithme de déchiffrement Decrypt sur l'entrée (sk, c) effectue les mêmes calculs que le déchiffrement du système de Paillier standard vu en td pour retrouver $M' \mod n$ tel que $c \equiv (1+n)^{M'} R'^n \pmod{n^2}$ avec $R' \in (\mathbf{Z}/n\mathbf{Z})^\times$. Il écrit ensuite s, la représentation binaire de M' sur exactement $2k$ bits, et pose m' égal aux k premiers bits de s et r' égal aux k derniers bits de s, de telle sorte que $s = m'||r'$. L'algorithme de déchiffrement vérifie ensuite que le chiffrement par Encrypt de m' avec l'aléa r' donne bien c. Si c'est le cas, il ressort m', sinon il ressort une erreur, notée $\perp$.

On rappelle l'hypothèse de la résidualité composite. Soit $\mathcal{D}$ un attaquant, on définit l'expérience $\mathbf{Exp}^{RC}_{\mathsf{Gen},k}(\mathcal{D})$:

1. Lancer Gen avec entrée k pour obtenir n
2. Prendre $b^\star \xleftarrow{\$} \{0,1\}$
3. Si $b^\star = 0$ alors $x \xleftarrow{\$} (\mathbf{Z}/n^2\mathbf{Z})^\times$, sinon choisir $r \xleftarrow{\$} (\mathbf{Z}/n\mathbf{Z})^\times$, et calculer $x = r^n$ dans $(\mathbf{Z}/n^2\mathbf{Z})^\times$.
4. $\mathcal{D}$ prend n et x en entrée et renvoie un bit b
5. La sortie de l'expérience est 1 si $b = b^\star$ et 0 sinon

On définit l'avantage de $\mathcal{D}$:

$$\mathbf{Adv}^{RC}_{\mathsf{Gen},k}(\mathcal{D}) = \left| \Pr\left(\mathbf{Exp}^{RC}_{\mathsf{Gen},k}(\mathcal{D}) = 1 \right) - \frac{1}{2} \right|.$$

L'hypothèse de la résidualité composite (RC) est qu'il existe un générateur Gen tel que tout algorithme probabiliste polynomial $\mathcal{D}$ ait un avantage négligeable pour cette expérience.

Le but de l'exercice est de montrer que ce schéma de chiffrement est IND – CCA2 sous l'hypothèse RC, dans le modèle de l'oracle aléatoire.

(a) Dans l'algorithme Decrypt, quels sont les calculs effectués par l'algorithme de déchiffrement du système de Paillier standard vu en td pour retrouver M' mod n tel que $c \equiv (1+n)^{M'} R'^n \pmod{n^2}$ avec $R' \in (\mathbf{Z}/n\mathbf{Z})^\times$?

Dans le reste de l'exercice, on note $\mathcal{A}$ un attaquant polynomial probabiliste IND – CCA2 contre ce schéma de chiffrement et on suppose que l'avantage ϵ de $\mathcal{A}$ lors de l'expérience IND – CCA2 est non négligeable.

(b) Donner l'expérience IND – CCA2 que joue $\mathcal{A}$. Comment est défini l'avantage de $\mathcal{A}$? Quels sont le ou les oracles que $\mathcal{A}$ peut interroger ?

À partir de $\mathcal{A}$, on veut construire un attaquant $\mathcal{D}$ avec un avantage non négligeable pour résoudre le problème de la résidualité composite, RC.

(c) Donner un algorithme qui permet à $\mathcal{D}$ d'interagir avec $\mathcal{A}$ pour simuler l'expérience IND – CCA2 et utiliser la sortie de $\mathcal{A}$ pour trouver la solution du problème RC. Dans cette question on ne simulera pas les oracles que $\mathcal{A}$ peut interroger.

(d) Comment $\mathcal{D}$ peut-il simuler l'oracle aléatoire $\mathcal{H}$ auquel $\mathcal{A}$ a accès ? Cette simulation est telle parfaite ?

(e) Montrer comment $\mathcal{D}$ peut utiliser la liste des requêtes de $\mathcal{A}$ à l'oracle aléatoire pour simuler l'oracle de déchiffrement. Préciser en particulier quels sont les chiffrés que $\mathcal{D}$ ne déchiffre pas correctement.

(f) On suppose qu'au cours de son exécution $\mathcal{A}$ fait q_D requêtes de déchiffrement. On note ces requêtes c_i pour $i = 1, \ldots, q_D$. Quelle est la probabilité que $\mathcal{D}$ ne sache pas correctement déchiffrer c_i ? En déduire une majoration de la probabilité qu'au moins une requête de $\mathcal{A}$ ne soit pas correctement déchiffrée par $\mathcal{D}$.

(g) Donner une minoration de l'avantage, en fonction de ϵ, de l'algorithme $\mathcal{D}$ que vous avez construit. Conclure.

Partie G. Zémor

Exercice 3. Soit n un multiple de 4. Soit $\mathbf{E}$ une matrice binaire $n/4 \times n$, où chaque ligne de $\mathbf{E}$ est de poids t, petit devant n. La matrice $\mathbf{E}$ est choisie aléatoire uniforme sous ces contraintes. On définit le code linéaire binaire $C = \{\mathbf{x} \in \mathbb{F}_2^n, \mathbf{E}\mathbf{x}^{\mathrm{T}} = 0\}$.

(a) Soit $\mathbf{G}$ une matrice génératrice aléatoire fixée de C. Soit $k = 3n/4$. On chiffre un message $\mathbf{m} \in \mathbb{F}_2^k$ par la correspondance

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

où $\mathbf{e}$ est un vecteur de petit poids w. Comment peut-on déchiffrer grâce à la clé secrète $\mathbf{E}$? Comment faut-il choisir les paramètres t et w pour que cela fonctionne ?

(b) Soit $\mathbf{A}$ une matrice aléatoire uniforme d'ordre $n/4 \times n$. Soit C maintenant le code $C = \{\mathbf{x} \in \mathbb{F}_2^n, \mathbf{E}\mathbf{x}^{\mathrm{T}} = 0 \text{ et } \mathbf{A}\mathbf{x}^{\mathrm{T}} = 0\}$, et $\mathbf{G}$ de nouveau une matrice génératrice aléatoire de C. Que devient la dimension de l'espace des messages si on continue à appliquer le chiffrement précédent ? Il y a-t-il d'autres changements ?

(c) On supposera :
- qu'étant donné un code binaire aléatoire de longueur n et de dimension $n/2$, il n'est pas réaliste de trouver un mot de code $\mathbf{c}$ à partir de $\mathbf{c} + \mathbf{e}$ où $\mathbf{c}$ est choisi uniformément dans le code et où $\mathbf{e}$ est aléatoire uniforme de poids w.
- qu'étant donné une matrice $\mathbf{A}$ aléatoire uniforme $n/4 \times n$ il est algorithmiquement difficile de distinguer si une matrice $\mathbf{B}$ d'ordre $n/4/ \times n$ est
  - aléatoire uniforme indépendante de $\mathbf{A}$,
  - ou de la forme $\mathbf{E} = \mathbf{SA} + \mathbf{E}$, où $\mathbf{S}$ est aléatoire d'ordre $n/4 \times n/4$.

Dans ces conditions, démontrer que le système de chiffrement introduit en (b) est sûr, si on suppose que le chiffrement

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

est appliqué à des messages clairs $\mathbf{m}$ uniformes dans $\mathbb{F}_2^k$.

(d) Soit $\mathbf{A}$ une matrice $n/4 \times n$ comme précédemment. On définit maintenant le code $C = \{\mathbf{x} \in \mathbb{F}_2^n,\ \mathbf{A}\mathbf{x}^{\mathrm{T}} = 0 \text{ et } \mathbf{E}\mathbf{x}^{\mathrm{T}} \in C_0\}$, où C_0 est un code de longueur $n/4$, et venant avec un algorithme de décodage efficace corrigeant un nombre suffisant d'erreurs que l'on précisera. De nouveau la clé publique du système est une matrice génératrice $\mathbf{G}$ aléatoire de C, et le chiffrement se fait par la correspondance

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

où l'on supposera que le clair $\mathbf{m}$ est réparti uniformément dans $\mathbb{F}_2^k$ où k est la dimension de C. Expliquer comment on déchiffre.

(e) Démontrer la sécurité du nouveau système sous des hypothèses similaires à celles du (c).

Exercice 4. Alice et Bob partagent une matrice aléatoire uniforme $\mathbf{A}$ d'ordre $n/4 \times n$. Ils s'échangent les données suivantes.
- Alice envoie à Bob $\mathbf{SA} + \mathbf{E}$ où $\mathbf{S}$ est une matrice $n/4 \times n/4$ aléatoire uniforme et où $\mathbf{E}$ est une matrice $n/4 \times n$ aléatoire dont les lignes sont toutes d'un petit poids t.
- Bob envoie à Alice $\mathbf{AE}'^{\mathrm{T}}$ où $\mathbf{E}'$ est une matrice $n/4 \times n$ aléatoire dont les lignes sont toutes d'un petit poids t.

(a) Montrer que Alice peut calculer une matrice $\mathbf{M}_A$, Bob peut calculer une matrice $\mathbf{M}_B$ telles que $\mathbf{M}_A$ et $\mathbf{M}_B$ diffèrent sur peu de coordonnées.

(b) En déduire un système de chiffrement à clé publique où la clé publique est constituée de la donnée de $\mathbf{A}$ et de $\mathbf{SA} + \mathbf{E}$. On expliquera le chiffrement et le déchiffrement.

(c) En faisant l'hypothèse, de type Diffie-Hellman décisionnel, que la distribution $(\mathbf{SA} + \mathbf{E}, \mathbf{AE}'^{\mathrm{T}}, \mathbf{M}_B)$ est indistinguable de la donnée de trois matrices aléatoires uniformes indépendantes de dimensions appropriées, démontrer la sécurité du système de chiffrement ci-dessus.

(d) On formule les deux hypothèses suivantes :

(i) la distribution $\mathbf{SA} + \mathbf{E}$ est indistinguable d'une matrice uniforme de mêmes dimensions,

(ii) la distribution de $\mathbf{UE}'^{\mathrm{T}}$, où $\mathbf{U}$ est une matrice $n/2 \times n$ aléatoire uniforme, est indistinguable d'une matrice aléatoire uniforme de même dimension que $\mathbf{UE}'^{\mathrm{T}}$.

Démontrer que dans ces conditions la distribution $(\mathbf{SA} + \mathbf{E}, \mathbf{AE}'^{\mathrm{T}}, \mathbf{M}_B)$ est indistinguable de la donnée de trois matrices aléatoires uniformes indépendantes. On pourra procéder en deux étapes.

(e) Montrer que la distribution $(\mathbf{SA} + \mathbf{E}, \mathbf{AE}'^{\mathrm{T}}, \mathbf{M}_A)$ n'est pas indistinguable de la distribution uniforme.