Notion de Sécurité CCA

Soit $\mathsf{KeyGen}, \mathsf{Enc}, \mathsf{Dec}$ un schéma de chiffrement à clé publique CCA-sûr défini sur $\mathcal{M}, \mathcal{C}$ où $\mathcal{C} = \{0,1\}^{\ell}$. Soit $(\mathsf{KeyGen}, \mathsf{Enc}', \mathsf{Dec}')$ un schéma défini sur $\mathcal{M}, \mathcal{C}' =$ où $\mathcal{C}' = \{0,1\}^{\ell+1}$ de la manière suivante :

$$\mathsf{Enc}'(\mathsf{pk}, m) := \mathsf{Enc}(\mathsf{pk}, m) \| 0$$

et

$$\mathsf{Dec}'(\mathsf{sk}, c) := \mathsf{Dec}(\mathsf{sk}, c[0 \dots \ell - 1])$$

autrement dit le dernier bit de chiffré peut être 0 ou 1, mais le déchiffrement l'ignore.

Question 1. *Le chiffrement est-il IND-CPA ?*

Solution. *Oui (pas la peine de détailler). Si on sait distinguer deux chiffrés, alors on peut enlever le 0.*

Question 2. *Montrer que* $(\mathsf{KeyGen}, \mathsf{Enc}', \mathsf{Dec}')$ *n'est pas IND-CCA.*

Solution. *Se souvenir du jeu IND-CCA.*

On a le droit de chiffrer, et de demander des requêtes de déchiffrement. Puis il faut distinguer.

L'astuce ici consiste à modifier le chiffré challenge en changeant le dernier bit, puis à demander le déchiffrement. C'est autorisé par le jeu IND-CCA car non trivial. Mais on récupère donc le message challenge et on casse le jeu.

(Ici on a cassé la sécurité IND-CCA2 car la requête de déchiffrement a lieu après réception du challenge, je pense qu'on pourrait prouver la sécurité IND-CCA1).

Combinaison de CBC et CBC-MAC

En cours, nous avons vu que la combinaison "encrypt-then-MAC" d'un chiffrement IND-CPA avec un MAC SUF-CMA nous donnait un chiffrement authentifié IND-CCA, et par ailleurs impossible à contrefaire (il est donc impossible pour un adversaire de créer une nouvelle paire "message, tag" valide). Cependant, de mauvaises combinaisons peuvent aboutir à des attaques. Nous donnons ici l'exemple de CBC combiné avec ECBC-MAC en mode "encrypt-and-MAC", dans lequel :

- On chiffre le message avec CBC ;
- On appelle le MAC *sur le message clair* ;
- On renvoie les deux résultats.

Suppsons qu'Alice chiffre un message à deux blocs (P_1, P_2) avec la clé K et une IV aléatoire et l'envoie à Bob.

Question 3. *Donner l'expression du chiffré* (C_1, C_2) *et du tag* T. *Montrer que si* $IV = 0$, $T = E_{K'}(C_2)$.

Solution.

$$\begin{cases} C_1 = E_K(IV + P_1) \\ C_2 = E_K(P_2 + C_1) \\ T = E_{K'}\left(E_K(P_2 + E_K(P_1))\right) \end{cases} \tag{1}$$

(trivial)

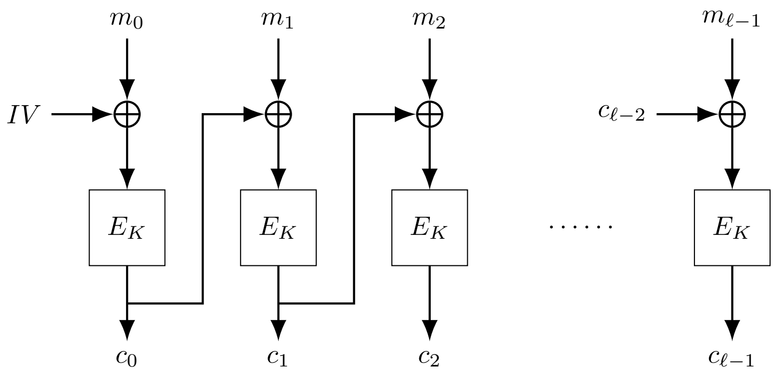

Figure 1 – Le mode CBC.

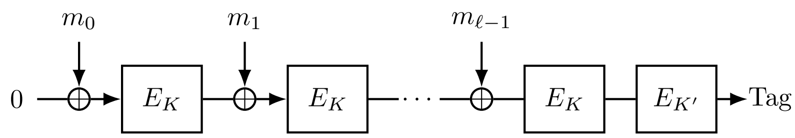

Figure 2 – ECBC-MAC.

Question 4. *En déduire qu'un attaquant Eve peut créer un nouveau triplet* (C'_1, C'_2, T') *valide.*

Solution. $T' = T$ *et on a juste à changer* C'_1*.*

Question 5. *On suppose maintenant qu'on utilise encrypt-then-MAC. Donner la nouvelle expression du tag. Montrer qu'il y a toujours un problème avec cette construction ; en particulier qu'on peut casser la sécurité IND-CPA.*

Solution.

$$\begin{aligned} T &= E_{K'}(E_K(C_2 + E_K(C_1))) \\ &= E_{K'} \circ E_K(E_K(P_2 + C_1) + E_K(C_1)) \end{aligned}$$

Si $P_2 = 0$ *le tag devient constant :* $T = E_{K'}E_K(0)$*. C'est vrai quelle que soit l'IV.*

Dans le jeu IND-CPA l'attaquant peut donc d'abord demander le tag d'un message avec $P_2 = 0$*, puis distinguer entre deux messages pour lequel* $P_2 = 0$ *sur l'un des deux et* $\neq 0$ *sur l'autre.*

Autour de ElGamal

Rappelons le principe du chiffrement ElGamal.

Soit $\mathbf{G}$ un groupe d'ordre q premier et g un générateur.

KeyGen. Tirer $\mathsf{sk} \hookleftarrow U([1, q-1])$ et calculer $\mathsf{pk} = g^{\mathsf{sk}}$.

$\mathsf{Enc}(m, \mathsf{pk})$. Tirer $r \hookleftarrow U(\mathbb{Z}_q)$ et renvoyer $(g^r, m \cdot \mathsf{pk}^r)$.

$\mathsf{Dec}((c_1, c_2), \mathsf{sk})$. Renvoyer $c_2 \cdot (c_1)^{-\mathsf{sk}}$.

Question 6. *Montrer que ElGamal est homomorphe pour la multiplication : si* $(c_1, c_2) = \mathsf{Enc}(m, \mathsf{pk})$ *et* $(c_1', c_2') = \mathsf{Enc}(m, \mathsf{pk})$ *alors* $(c_1 c_1', c_2 c_2') = \mathsf{Enc}(mm', \mathsf{pk})$.

Solution. *Trivial.*

Question 7. *ElGamal est-il IND-CCA ?*

Solution. *La réponse est non et cela vient de l'homomorphisme. Durant le jeu IND, on considère le chiffré challenge* (c_1^*, c_2^*). *On chiffre un message quelconque*(c_1^*, c_2^*) m *en* (c_1, c_2). *On demande de déchiffrer* $(c_1^* c_1, c_2^* c_2)$ *ce qui donne* mm^*. *On peut donc en déduire* m^*.

La variante qui a été cassée est IND-CCA2 (quand on fait du déchiffrement après avoir reçu le challenge).

Question 8. *La valeur aléatoire* r *(« sel ») peut-elle être réutilisée pour un autre message ?*

Solution. *Non.*

Question 9. *Soit* p, q *de grands premiers tels que* q *divise* $p - 1$. *Soit* $\mathbf{G}$ *le sous-groupe de* $\mathbb{Z}_p^*$ *d'ordre* q *engendré par* g *et supposons que DDH est difficile dans* $\mathbf{G}$.

Supposons que nous instancions le système ElGamal dans le groupe $\mathbf{G}$ *; cependant les messages sont pris dans le groupe* $\mathbb{Z}_p^*$ *entier. Montrer que le système qui résulte de ce choix n'est pas IND-CPA.*

Solution. *D'abord il faut bien comprendre ce qu'on est en train de faire ; la différence avec le ElGamal classique n'est pas très grande. On va travailler modulo* p *quand on chiffre et on déchiffre. Mais la puissance de* g *est toujours dans le groupe* $\mathbf{G}$.

L'astuce pour distinguer deux messages consiste à prendre l'un dans le groupe $\mathbf{G}$ *(par exemple 1) et l'autre quelconque dans* $\mathbb{Z}_p^*$. *Ensuite on regarde si le chiffré renvoyé* $m_b \cdot \mathsf{pk}^r$ *est dans* $\mathbf{G}$ *ou non.*

Pour faire ce test on calcule l'ordre de l'élément. En général un élément de $\mathbb{Z}_p^*$ *n'aura pas d'ordre* q. *Donc il suffit de calculer une puissance, et c'est efficace.*

Pendant des décennies, la question de savoir si ElGamal était IND-CCA1 est restée ouverte. On sait aujourd'hui qu'on ne peut pas prouver sa sécurité IND-CCA1 sous des hypothèses standard ("New limits of provable security and applications to ElGamal encryption", Schäge, EUROCRYPT 2024).

Une variante de RSA (*)

Le problème RSA est le suivant.

Problem 1. *Soit* $N = pq$ *où* p, q *sont premiers,* e *premier avec* $\phi(N)$ *et* y*, trouver* x *tel que* $x^e = y \pmod N$.

Considérons le schéma ci-dessous, où H est une fonction de hachage « idéale ».

KeyGen. Le même que textbook RSA : $\mathsf{pk} := (e, N)$ et $\mathsf{sk} := (d, N)$ où $ed = 1 \pmod{\phi(N)}$.

$\mathsf{Enc}(m, \mathsf{pk})$. Tirer $r \hookleftarrow U(\mathbb{Z}_N)$ et renvoyer $(r^e, H(r)m) \in \mathbb{Z}_N^2$.

$\mathsf{Dec}((c_1, c_2), \mathsf{sk})$. Renvoyer $c_2(H(c_1^d))^{-1} \in \mathbb{Z}_N$.

Question 10. *Montrer que le schéma est correct.*

Solution. *Cela suit de $x^{ed} = x \pmod N$.*

Question 11. *Montrer qu'on peut utiliser un adversaire contre le problème RSA pour monter un adversaire contre le jeu IND-CPA.*

Solution. *Considérons un adversaire $\mathcal{B}$ pour le problème RSA. Il reçoit y et renvoie x tel que $x^e = y \pmod N$.*

Nous allons utiliser $\mathcal{B}$ pour implémenter un adversaire $\mathcal{A}$ dans le jeu IND-CPA.

L'adversaire $\mathcal{A}$ choisit m_0, m_1, reçoit $(r^e, H(r)m_b)$ et doit renvoyer un bit b.

Quand il reçoit $(r^e, H(r)m_b)$ il envoie r^e à $\mathcal{B}$, qui trouve r et le renvoie. Ensuite $\mathcal{A}$ multiplie la deuxième partie du chiffré par $H(r)^{-1}$, trouve le message, et détermine b.

L'avantage de $\mathcal{A}$ dans le jeu IND-CPA est donc égal à la probabilité de succès de $\mathcal{B}$ pour résoudre le problème RSA.

La fonction H est considérée se comporter comme un *oracle aléatoire*. Ainsi, dans les preuves de sécurité suivantes, on supposera que chaque fois qu'un calcul de H est effectué sur une entrée x :

- si x a déjà été vu, on renvoie la même valeur $H(x)$ que précédemment ;
- sinon, $H(x)$ est sélectionné uniformément au hasard.

Question 12. *Soit G le jeu IND-CPA joué entre le challenger $\mathcal{C}$ et l'attaquant $\mathcal{A}$. Soit G' une modification de ce jeu dans lequel le chiffré challenge $c_1, c_2 = (r^e, H(r)m_b)$ est modifié comme suit : $H(r)$ est remplacé par une valeur aléatoire indépendante de r.*

Soit E l'évènement dans le jeu G' :

« $\mathcal{A}$ appelle H sur l'entrée r »

Justifier que tant que l'évènement E ne se produit pas, la vue de l'adversaire dans le jeu G et dans le jeu G' sont identiques.

Solution. *En effet si E se produit on a déjà déterminé la valeur de $H(r)$, elle ne peut pas être re-tirée au hasard. Sinon tout est aléatoire.*

Question 13. *Justifier que :*

$$\Pr[\mathcal{A} \textit{ gagne } G'] = \frac{1}{2}$$

Solution. *Tout est aléatoire.*

Question 14. *Déduire des questions précédentes que :*

$$\left| \Pr[\mathcal{A} \textit{ gagne } G] - \frac{1}{2} \right| \leq \Pr[E] \quad .$$

Solution. *Attention : E est bien un évènement du jeu G', il n'a pas vraiment de sens dans le jeu G.*

$$\Pr[\mathcal{A}\ gagne\ G'] = \Pr[E]\Pr[\mathcal{A}\ gagne\ G'|E] + (1-\Pr[E])\underbrace{\Pr[\mathcal{A}\ gagne\ G'|\neg E]}_{\Pr[\mathcal{A}\ gagne\ G]}$$

$$\frac{1}{2} \leq \Pr[E] + \Pr[\mathcal{A}\ gagne\ G]$$

$$\frac{1}{2} - \Pr[\mathcal{A}\ gagne\ G] \leq \Pr[E] \quad .$$

Au passage on peut faire le même raisonnement avec "$\mathcal{A}$ ne gagne pas G" à la place ce qui donne :

$$\frac{1}{2} - (1 - \Pr[\mathcal{A}\ gagne\ G]) \leq \Pr[E]$$

et donc :

$$|\Pr[\mathcal{A}\ gagne\ G] - \frac{1}{2}| \leq \Pr[E] \quad .$$

Question 15. *Soit $\mathcal{A}$ un adversaire IND-CPA contre le schéma proposé. On construit un adversaire $\mathcal{B}$ contre le problème RSA comme suit.*

$\mathcal{B}$ exécute localement $\mathcal{A}$. Pendant cette exécution, $\mathcal{A}$ "croit" se trouver dans le jeu G'.

- *Chaque fois que $\mathcal{A}$ fait un appel à H, on enregistre la valeur d'appel.*
- *$\mathcal{A}$ choisit une paire de messages m_0, m_1*
- *Lorsque $\mathcal{B}$ reçoit la valeur y, il envoie $(y, *m_b)$ à $\mathcal{A}$, où $*$ est une nouvelle valeur aléatoire.*
- *Enfin, quand $\mathcal{A}$ termine, $\mathcal{B}$ renvoie une des valeurs d'appel prise au hasard.*

On suppose qu'au maximum t requêtes à H sont faites. Soit E l'évènement défini plus haut, qui concerne donc $\mathcal{A}$. Montrer que :

$$\Pr[\mathcal{B}\ gagne] \geq \frac{1}{t}\Pr[E] \quad .$$

Solution. *Si E arrive, alors cela veut dire qu'une des valeurs d'appel de H contient z où $z^e = y$. Donc une solution au problème.*

Question 16. *Montrer que le schéma proposé est IND-CPA sous l'hypothèse que le problème RSA est difficile.*

Solution. *Sous l'hypothèse RSA, pour tout adversaire $\mathcal{B}$ on a $\Pr[\mathcal{B}\ gagne] = \text{negl}$ ce qui implique $\Pr[\mathcal{A}\ gagne\ G] = \text{negl}$.*