Université de Limoges

Développement Logiciel Cryptographique

TD n° 1 : Cryptanalyse SQUARE de l'AES

1 Cryptanalyse SQUARE de l'AES à 4 tours

Commençons par quelques définitions simples permettant d'introduire la notion centrale de λ-set :

Définition 1 (État) *On appelle* état $(s)_{i,j}$ *une matrice* $4 \text{x} 4$ *d'octets représentant n'importe quel résultat intermédiaire dans le processus de calcul de l'AES.*

Définition 2 (Cellule) *Pour* $i, j \in \{0, 1, 2, 3\}$*, la* cellule (i, j) *est simplement définie comme l'élément à la ligne* i *et à la colonne* j *d'un état.*

Définition 3 (Cellule active) *Une cellule* (i, j) *est dite* active *à travers un ensemble* $(s^{(t)})_t$ *de 256 états si :*

$$\{s_{i,j}^{(t)} : t = 0, \dots, 255\} = \{0, \dots, 255\}$$

Définition 4 (Cellule inactive) *Une cellule* (i, j) *est dite* inactive *à travers un ensemble* $(s^{(t)})_t$ *de 256 états si il existe une valeur d'octet* c *telle que :*

$$\{s_{i,j}^{(t)} : t = 0, \dots, 255\} = \{c\}$$

Définition 5 (λ-set) *On dit qu'un ensemble de 256 états* $(s^{(t)})_{t=0\dots255}$ *est un* λ-set *si chacune de ses cellules* (i, j) *est soit active, soit inactive à travers l'ensemble de ces états.*

- Donnez un exemple de 256 états formant un λ-set.
- Donnez un exemple de 256 états ne formant pas un λ-set.

Notations : On s'autorisera l'utilisation des abréviations `SB`, `SR`, `MC` et `ARK` pour désigner respectivement les fonctions `SubBytes`, `ShiftRows`, `MixColumns` et `AddRoundKey` de l'AES.

Dans le processus de chiffrement AES d'un clair P avec une clé K, on note P_r l'état correspondant à l'entrée du tour r, et K_r la clé de tour utilisée dans ce même tour. On a notamment :

$$\begin{aligned} P_1 &= \mathtt{ARK}(P, K) \\ P_{r+1} &= \mathtt{ARK}(\mathtt{MC}(\mathtt{SR}(\mathtt{SB}(P_r))), K_r) \qquad (r = 1, \ldots, 3) \\ C &= \mathtt{ARK}(\mathtt{SR}(\mathtt{SB}(P_4)), K_4) \end{aligned}$$

Dans ce qui suit, nous nous intéressons à la conservation (ou non) du caractère actif ou inactif d'une cellule (à travers un ensemble de 256 états) durant le calcul des différentes fonctions de l'AES.

- Que peut-on dire de la cellule de sortie de la fonction `SB` lorsque la cellule d'entrée correspondante est active (resp. inactive) ?

- Que peut-on dire de la cellule de sortie de la fonction `ARK` lorsque la cellule d'entrée correspondante est active (resp. inactive) ?

- Comment se propage une cellule active (resp. inactive) à travers la fonction `SR` ?

- Supposons une colonne composée uniquement de cellules inactives. Que peut-on dire de la colonne correspondante en sortie de la fonction `MC` ?

- Supposons une colonne composée d'une cellule active et de trois cellules inactives. Que peut-on dire de la colonne correspondante en sortie de la fonction `MC` ?

- Qu'en est-il des cas où la colonne est composée de 2, 3 ou 4 cellules actives et de respectivement 2, 1 ou 0 cellules inactives ?

Nous nous plaçons dans le cadre d'une attaque à clairs choisis, et supposons que l'attaquant a pu obtenir les 256 chiffrés d'un AES à 4 tours correspondant à 256 clairs $(P^{(t)})_t$ formant un λ-set composé d'une cellule active et de 15 cellules inactives.

- Décrivez la propagation de ce λ-set à travers le processus de chiffrement. Exprimez une propriété $\mathcal{P}_1$ que vérifie l'ensemble des 256 états $(\mathtt{SR}(\mathtt{SB}(P_3^{(t)})))_t$ en entrée du `MixColumns` du troisième tour.

- Qu'en est-il de $(\mathtt{MC}(\mathtt{SR}(\mathtt{SB}(P_3^{(t)}))))_t$?

Définition 6 (Cellule équilibrée) *Une cellule* (i, j) *est dite* équilibrée *à travers un ensemble* $(s^{(t)})_t$ *de 256 états si :*

$$\bigoplus_{t=0}^{255} s_{i,j}^{(t)} = 0$$

- Montrez qu'une cellule active (resp. inactive) est a fortiori équilibrée.

- Que peut-on dire d'une colonne de sortie de la fonction `MC` lorsque la colonne d'entrée correspondante n'est composée que de cellules équilibrées ?

- Montrez qu'une cellule équilibrée le demeure à travers la fonction `ARK`.

- Qu'en est-il pour la fonction `SB` ?

- Exprimez une propriété $\mathcal{P}_2$ que vérifie l'ensemble des 256 états $(P_4^{(t)})_t$ en entrée du dernier tour.

Nous allons maintenant profiter de la propriété $\mathcal{P}_2$ pour concevoir une attaque permettant de retrouver un à un les octets de la clé K_4 du dernier tour.

- Quelle partie de K_4 doit-on connaître pour pouvoir calculer la cellule (i, j) de $(P_4^{(t)})_t$ à partir des chiffrés $(C^{(t)})_t$?
Application numérique : $(i, j) = (1, 1)$.

- Soit r_1 et r_2 deux valeurs d'octets aléatoires indépendantes. Quelle est la probabilité de l'événement $r_1 \oplus r_2 = 0$?

- Soit $(r^{(t)})_t$ 256 valeurs d'octets aléatoires indépendantes. Quelle est la probabilité de l'événement $\bigoplus_{t=0}^{255} r^{(t)} = 0$?

- Proposez une manière d'invalider un grand nombre de candidats concernant la valeur d'un octet arbitraire de K_4. Quelle est l'espérance du nombre de candidats restant valides pour cet octet ?

- Expliquez comment calculer un candidat K à partir d'un candidat K_4.

- Décrivez intégralement l'attaque SQUARE sur l'AES à 4 tour. Évaluez sa complexité – en temps (équivalent nombre de chiffrements) et en données (nombre de clairs choisis).

2 Extension à l'AES à 5 tours

Dans l'adaptation de l'attaque SQUARE à l'AES à 5 tours nous continuons à considérer un attaquant disposant des chiffrés (sur 5 tours) d'un λ-set n'ayant qu'une seule cellule active. La propriété $\mathcal{P}_2$ continue à être vérifiée en entrée du tour 4.

• Quelles parties de K_4 et de K_5 doit-on connaître pour pouvoir calculer la cellule (i, j) de $(P_4^{(t)})_t$ à partir des chiffrés $(C^{(t)})_t$?
Application numérique : $(i, j) = (2, 3)$.

• Dans quelle proportion est réduit le nombre de candidats valides lorsqu'on exploite par crible un λ-set chiffré ?

• Combien de λ-sets chiffrés sont nécessaires pour réduire le nombre de candidats suggérés à une toute petite valeur (resp. pour être quasi certain de n'avoir qu'une seule clé suggérée) ?

• Évaluez la complexité (en temps et données) de la version la plus simple de l'attaque SQUARE sur l'AES à 5 tours.

Une astuce permet de modifier l'attaque de base pour en réduire considérablement la complexité temporelle.

• En utilisant la linéarité de la fonction `MixColumns`, montrez que tout octet de P_4 peut être ré-écrit comme somme $\alpha \oplus \beta$ d'une donnée α ne dépendant que d'une colonne de P_5 et d'une combinaison linéaire β d'octets de K_4.

• En remarquant que l'inconnu sur β ne concerne que 8 bits, montrez que l'on peut profiter de cette ré-écriture pour diviser par un facteur 2^{24} le nombre de candidats clés à considérer.

• Évaluez la complexité (en temps et données) de cette version moins naïve de l'attaque SQUARE sur l'AES à 5 tours.