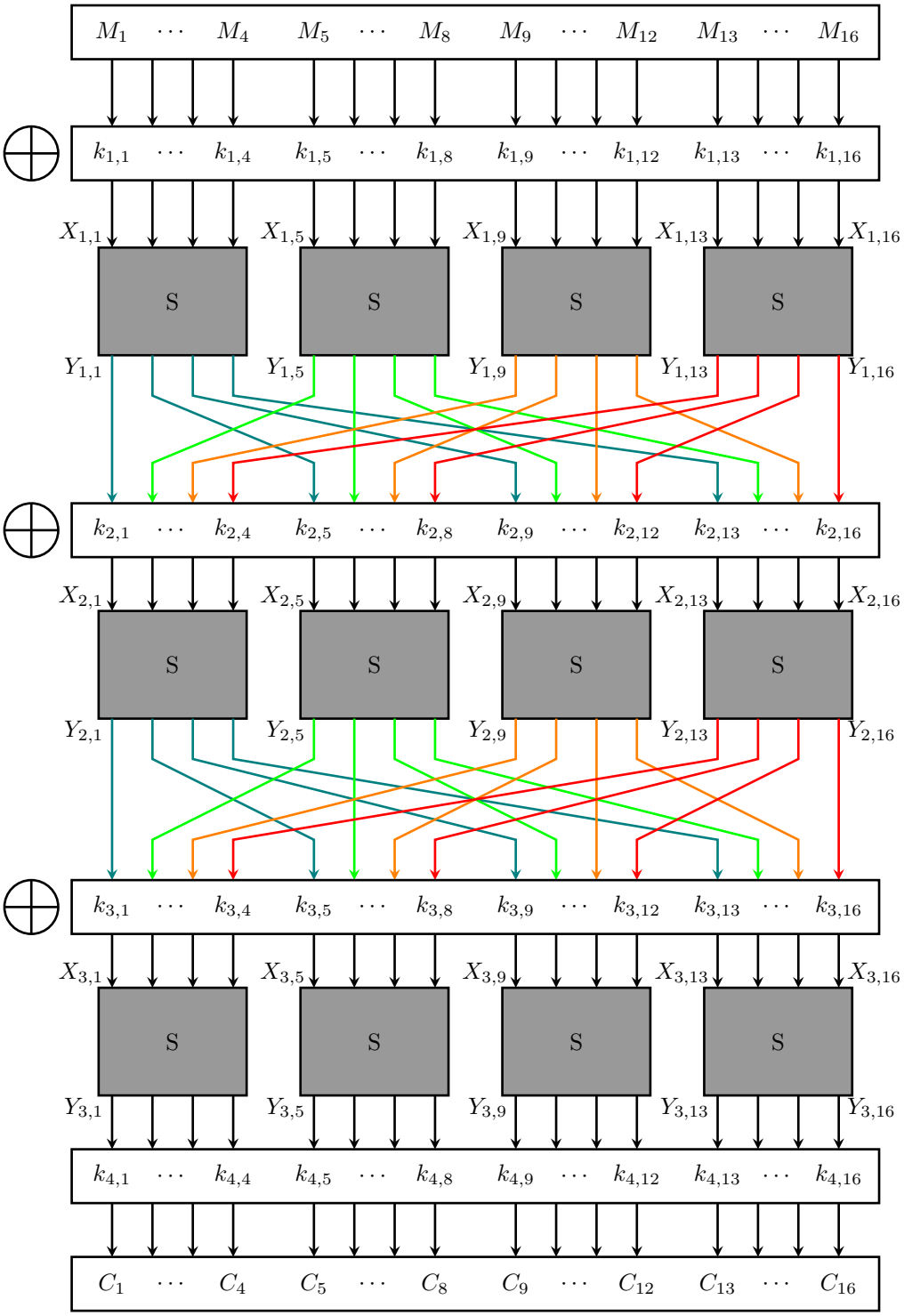
CY Cergy-Paris Université — **Enseignant : Kévin Carrier**

Introduction à la Cryptographie — kevin.carrier@cyu.fr

TD 5 : Cryptanalyse Linéaire et Différentielle

On propose le schéma de chiffrement suivant :



où S est une S-box (boite de substitution) définie sur $\mathbb{F}_2^4$ comme suit :

x	$0x0$	$0x1$	$0x2$	$0x3$	$0x4$	$0x5$	$0x6$	$0x7$	$0x8$	$0x9$	$0xA$	$0xB$	$0xC$	$0xD$	$0xE$	$0xF$
$S(x)$	$0x7$	$0xC$	$0x2$	$0xF$	$0xD$	$0xE$	$0x5$	$0x6$	$0xA$	$0x0$	$0x1$	$0x9$	$0x4$	$0x8$	$0xB$	$0x3$

Remarque. *Dans ce TD, on représente les variables aléatoires par des lettres majuscules scripts comme M_j, $X_{i,j}$, $Y_{i,j}$ et C_j tandis que les lettres minuscules représentent des valeurs fixées comme $k_{i,j}$.*

Exercice 1. Cryptanalyse Linéaire.

Trouver des relations linéaires biaisées dans la S-box. On note $\overrightarrow{X} := (X_1, X_2, X_3, X_4)$ l'entrée de la boite S et $\overrightarrow{Y} := S(\overrightarrow{X}) := (Y_1, Y_2, Y_3, Y_4)$ sa sortie. Soit f une combinaison linéaire de $X_1, X_2, X_3, X_4, Y_1, Y_2, Y_3$ et Y_4 :

$$f(\overrightarrow{X}, \overrightarrow{Y}) := c_1X_1 \oplus c_2X_2 \oplus c_3X_3 \oplus c_4X_4 \oplus c_5Y_1 \oplus c_6Y_2 \oplus c_7Y_3 \oplus c_8Y_4$$

avec $c_1, \cdots, c_8 \in \mathbb{F}_2$ des coefficients qui définissent f.

Quelque soit f, si $(\overrightarrow{X}, \overrightarrow{Y})$ suit une distribution uniforme sur $\mathbb{F}_2^8$, alors on devrait avoir idéalement

$$\mathbb{P}[f(\overrightarrow{X}, \overrightarrow{Y}) = 0] = \tfrac{1}{2}$$

Cependant, en pratique, les boites de substitution ne sont pas parfaites et on a des biais :

$$\mathbb{P}[f(\overrightarrow{X}, \overrightarrow{Y}) = 0] = \tfrac{1}{2} + \varepsilon$$

Dans la suite, on identifie une combinaison linaire f par un octet dont les bits correspondent aux coefficients $c_1, \cdots, c_8$. Par exemple, la combinaison linéaire $0x5A$ correspond à $c_1 = 0, c_2 = 1, c_3 = 0, c_4 = 1, c_5 = 1, c_6 = 0, c_7 = 1, c_1 = 0$.

1. Pour chaque combinaison linéaire des entrées et sortie de la S-box, calculer (à l'aide d'un programme Python) le biais ε. Complétez le tableau suivant :

f	$\mathbb{P}[f(\overrightarrow{X}, \overrightarrow{Y}) = 0]$	$\lvert\varepsilon\rvert$
$0x00$		
$0x01$		
$0x02$		
$\vdots$	$\vdots$	$\vdots$
$0xFD$		
$0xFE$		
$0xFF$		

2. Justifiez que les combinaisons linéaires $0x00, \cdots, 0x0F$ sont inutiles.

3. Déterminez les probabilités des événements suivants :

   i. $M_1 \oplus k_{1,1} \oplus Y_{1,2} = 0$

   ii. $Y_{1,2} \oplus k_{2,5} \oplus Y_{2,6} = 0$

4. On admet le *Pilling-up Lemma* : Soient n variables aléatoires de Bernoulli $B_1, \cdots B_n$. On note $p_i := \mathbb{P}(B_i = 0)$ pour tout $i \in \{1, \cdots, n\}$. On a alors

$$\mathbb{P}(B_1 \oplus \cdots \oplus B_n = 0) = \frac{1 + \prod_{i=1}^{n}(2p_i - 1)}{2}$$

   Montrez que

$$\mathbb{P}[M_1 \oplus k_{1,1} \oplus k_{2,5} \oplus k_{3,6} \oplus X_{3,6} = 0] = 0.78125.$$

5. En déduire que

$$\mathbb{P}[M_1 \oplus X_{3,6} = 0] = 0.21875 \text{ ou } 0.78125$$

Exploiter les biais pour retrouver une partie de la clé. Dans la question 5, nous avons estimé la probabilité $\mathbb{P}[M_1 \oplus X_{3,6} = 0]$ et avons constaté qu'elle était biaisée par rapport à $\frac{1}{2}$. Nous allons à présent estimer heuristiquement cette même probabilité grâce à un ensemble de couples clair/chiffré mis à votre disposition et en faisant certaines hypothèses sur la clé. L'hypothèse permettant de nous rapprocher le plus du résultat de la question 5 sera celle que nous validerons.

On note $\mathcal{S}$ l'ensemble des couples clair/chiffré à notre disposition :

$$\mathcal{S} \subseteq \{(\mathbf{m}, \mathbf{c}) : \mathbf{c} := Enc(\mathbf{k}, \mathbf{m})\}$$

avec $\sharp\mathcal{S} := N$.

6. Exprimez $X_{3,6}$ en fonction de $k_{4,5}, k_{4,6}, k_{4,7}, k_{4,8}$ et C_5, C_6, C_7, C_8.
7. Pour chaque valeur possible du 4-uplet $(k_{4,5}, k_{4,6}, k_{4,7}, k_{4,8}) \in \mathbb{F}_2^4$, estimez numériquement la probabilité $\mathbb{P}[M_1 \oplus X_{3,6} = 0]$ à l'aide de $\mathcal{S}$. En déduire cette partie de la clé secrète.
8. Pour distinguer statistiquement une probabilité $\frac{1}{2} + \varepsilon$ par rapport à une probabilité $\frac{1}{2}$, on estime qu'il faut au moins $\frac{1}{\varepsilon^2}$ échantillons statistiques. Quelle est donc la complexité de notre attaque?
9. En suivant une stratégie similaire, retrouvez une autre partie de la clé.
10. La cryptanalyse linéaire nécessite-t-elle un modèle d'attaquant de type COA, KPA, CPA ou bien CCA?
11. Comment peut-on se protéger des attaques linéaires? Justifier.

Exercice 2. Cryptanalyse Différentielle.

Le template d'une cryptanalyse différentielle ressemble beaucoup à celui d'une cryptanalyse linéaire. Au lieu d'étudier les biais dans les relations linéaires de la boite S, nous allons ici étudier les biais dans les différentielles.

Trouver des différentielles biaisées dans la S-box. On note X et X' deux variables aléatoires sur $\mathbb{F}_2^4$ représentant deux entrées de la boite S. On pose $Y := S(X)$ et $Y' := S(X')$ les sorties correspondantes.

Pour tout $(\Delta_X, \Delta_Y) \in \mathbb{F}_2^4 \times \mathbb{F}_2^4$, on aimerait idéalement avoir

$$\mathbb{P}\left[Y \oplus Y' = \Delta_Y | X \oplus X' = \Delta_X\right] = \frac{1}{16}$$

Cependant, en pratique, les boites de substitution ne sont pas parfaites et on a des biais :

$$\mathbb{P}\left[Y \oplus Y' = \Delta_Y | X \oplus X' = \Delta_X\right] = \frac{1}{16} + \varepsilon$$

Dans la suite, on identifie une différentielle sur l'entrée et la sortie de la boite S par un octet dont le nibble (4 bits) de gauche représente Δ_X et le nibble de droite Δ_Y. Par exemple, $0x5A$ signifie que la différentielle Δ_X sur l'entrée est 0101 et la différentielle Δ_Y sur la sortie est 1010.

1. Pour chaque différence d'entrées et de sorties de la S-box, calculer (à l'aide d'un programme Python) le biais ε correspondant. Complétez le tableau suivant :

$\Delta_X\Delta_Y$	$\mathbb{P}\left[Y \oplus Y' = \Delta_Y \mid X \oplus X' = \Delta_X\right]$	$\lvert\varepsilon\rvert$
$0x00$		
$0x01$		
$0x02$		
$\vdots$	$\vdots$	$\vdots$
$0xFD$		
$0xFE$		
$0xFF$		

2. Justifiez que les différentielles $0x00, \cdots, 0x0F$ sont inutiles.

3. On note $A_{i,j..k} := (A_{i,j}, A_{i,j+1}, \cdots, A_{i,k})$. De plus, on duplique les variables aléatoires décrites dans le schéma de la page 1 en ajoutant la notation "prime" pour représenter deux entrées distinctes. Déterminez les probabilités des événements suivants :

   i. $Y_{1,1..16} - Y'_{1,1..16} = 0x2000 \,|\, M_{1,1..16} - M'_{1,1..16} = 0x5000$

   ii. $Y_{2,1..16} - Y'_{2,1..16} = 0x0060 \,|\, Y_{1,1..16} - Y'_{1,1..16} = 0x2000$

4. En déduire
$$\mathbb{P}\left[X_{3,5..12} - X'_{3,5..12} = 0x22 \,|\, M_{1,1..16} - M'_{1,1..16} = 0x5000\right]$$

Exploiter les biais pour retrouver une partie de la clé. Dans la question 4, nous avons estimé la probabilité $\mathbb{P}\left[X_{3,5..12} - X'_{3,5..12} = 0x22 \,|\, M_{1,1..16} - M'_{1,1..16} = 0x5000\right]$ et avons constaté qu'elle était biaisée par rapport à $\frac{1}{256}$. Nous allons à présent estimer heuristiquement cette même probabilité grâce à un ensemble de couples clair/chiffré mis à votre disposition et en faisant certaines hypothèses sur la clé. L'hypothèse permettant de nous rapprocher le plus du résultat de la question 4 sera celle que nous validerons.

On note $\mathcal{S}$ l'ensemble des couples clair/chiffré à notre disposition :

$$\mathcal{S} \subseteq \{(\mathbf{m}, \mathbf{m}', \mathbf{c}, \mathbf{c}') \,:\, \mathbf{c} := Enc(\mathbf{k}, \mathbf{m}) \text{ et } \mathbf{c}' := Enc(\mathbf{k}, \mathbf{m}') \text{ et } \mathbf{m} \oplus \mathbf{m}' = 0x5000\}$$

avec $\sharp\mathcal{S} := N$.

5. Exprimez $X_{3,5..12}$ (resp. $X'_{3,5..12}$) en fonction de $k_{4,5..12}$ et $C_{5..12}$ (resp. $C'_{5..12}$).

6. Pour chaque valeur possible du 8-uplet $k_{4,5..12} \in \mathbb{F}_2^8$, estimez numériquement et à l'aide de $\mathcal{S}$ la probabilité
$$\mathbb{P}\left[X_{3,5..12} - X'_{3,5..12} = 0x22 \,|\, M_{1,1..16} - M'_{1,1..16} = 0x5000\right].$$
   En déduire la valeur de $k_{4,5..12}$.

7. En suivant une stratégie similaire, retrouvez une autre partie de la clé.

8. La cryptanalyse différentielle nécessite-t-elle un modèle d'attaquant de type COA, KPA, CPA ou bien CCA?

9. Comment peut-on se protéger des attaques différentielle? Justifier.