Cryptanalyse

Cours 5 - Cryptanalyse Différentielle

Maxime Bombar

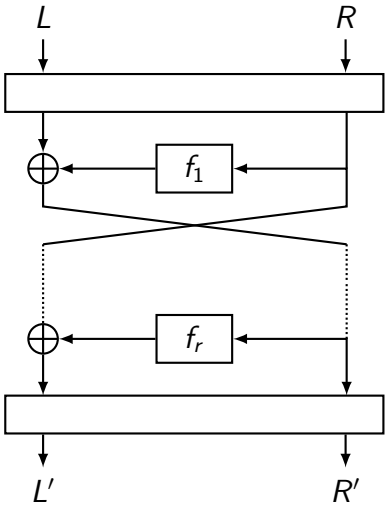
Mardi 01 Septembre 2024

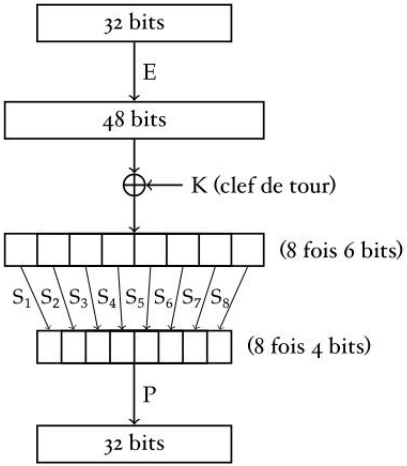
Rappels de la Semaine Dernière

Chiffrement par blocs

- Modes d'opération pour chiffrer tout un message.
- Attaque de Vaudeney sur le mode CBC : Attention au padding.
- Modes modernes assurent aussi l'intégrité.

Schémas de Feistel (*e.g.*, DES)





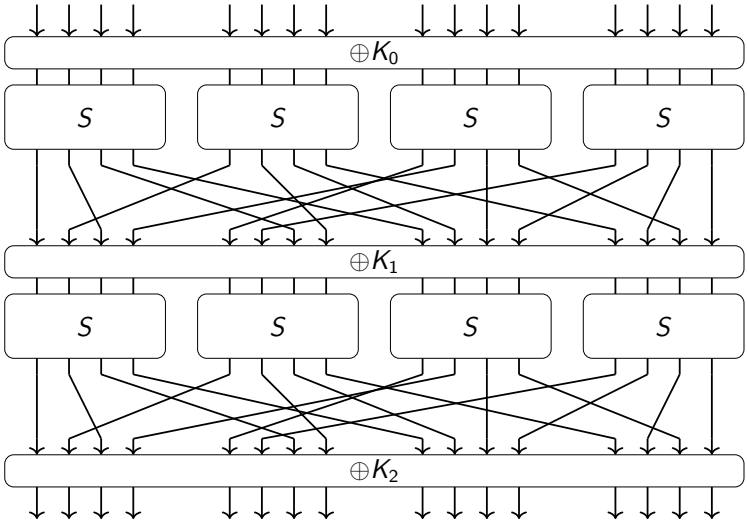
Faiblesse des fonctions de tour

V. Rijmen, B. Preneel (1997)

Utiliser des fonctions de tour non surjectives peut-être exploité pour monter des attaques.

Idée : Utiliser des fonctions booléennes **bijectives** $\mathbb{F}_2^n \to \mathbb{F}_2^n$.

Réseaux de Substitutions-Permutations (SPN)



Fonctions de tour **bijectives** :

Confusion : Chaque S-box est une bijection **non linéaire**.

Diffusion : Les sorties des S-box sont permutées.

Exemple : AES (Daemen, Rijmen)

Programme des prochaines séances

Aujourd'hui :
Cryptanalyse différentielle

08 Octobre :
Cryptanalyse linéaire

Attaques statistiques
Des points communs

15 Octobre :
Présentation des stages (13h-14h30)

25 et 23 Octobre :
Double Séance
Cryptanalyse Algébrique

Programme des prochaines séances

Aujourd'hui :
Cryptanalyse différentielle

08 Octobre :
Cryptanalyse linéaire

15 Octobre :
Présentation des stages (13h-14h30)

25 et 23 Octobre :
Double Séance
Cryptanalyse Algébrique

Séance pour digérer les deux précédentes
Probablement **fonctions de hachages**.

Décaler la séance ? TD uniquement ?

Programme des prochaines séances

Aujourd'hui :
Cryptanalyse différentielle

08 Octobre :
Cryptanalyse linéaire

15 Octobre :
Présentation des stages (13h-14h30)

25 et 23 Octobre :
Double Séance
Cryptanalyse Algébrique

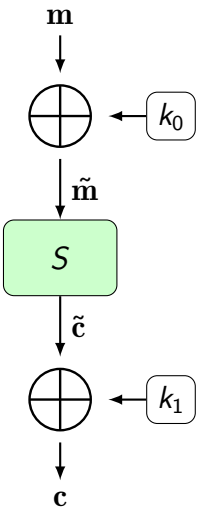
Rattrappe séance Carte à Puces du 5 Novembre.

Attaques Algébriques :

- Exploite la structure du chiffrement
- Résolution de systèmes polynomiaux
- Bases de Gröbner

Différentielles

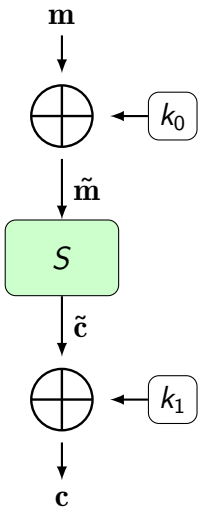
Introduction



Hypothèse : Attaque par clairs connus.
On connaît des paires $(\mathbf{m}, \mathbf{c})$.

Imaginez que l'on connaisse en plus $\tilde{\mathbf{c}}$.

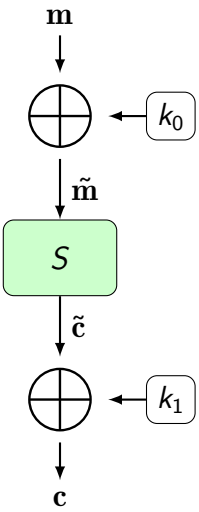
Introduction



Hypothèse : Attaque par clairs connus.
On connaît des paires $(\mathbf{m}, \mathbf{c})$.

Imaginez que l'on connaisse en plus $\tilde{\mathbf{c}}$.
On pourrait en déduire k_1, puis tout le reste.

Introduction

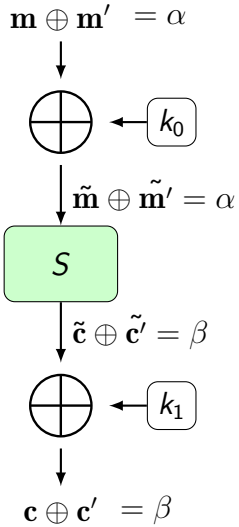


Hypothèse : Attaque par clairs connus.
On connaît des paires $(\mathbf{m}, \mathbf{c})$.

Imaginez que l'on connaisse en plus $\tilde{\mathbf{c}}$.
On pourrait en déduire k_1, puis tout le reste.

Problème : Avec k_0 clé uniforme, et S bijective, $\tilde{\mathbf{c}}$ est **uniformément distribué**.

Introduction



Hypothèse : Attaque par clairs connus.
On connaît des paires $(\mathbf{m}, \mathbf{c})$.

Imaginez que l'on connaisse en plus $\tilde{\mathbf{c}}$.
On pourrait en déduire k_1, puis tout le reste.

Problème : Avec k_0 clé uniforme, et S bijective, $\tilde{\mathbf{c}}$ est **uniformément distribué**.

Idée : $(\mathbf{m} \oplus k_0) \oplus (\mathbf{m}' \oplus k_0) = \mathbf{m} \oplus \mathbf{m}'$

Cryptanalyse Différentielle

- Attaque par **clairs choisis**.
- Formalisée par Eli Biham et Adi Shamir en 1990 pour la cryptanalyse de DES.
- En partie connue par IBM et NSA dès les années 1970 (Coppersmith 1994).
- Aujourd'hui, l'un des grands principes cryptanalytiques pour guider design et attaques.

Objectif

(1) Prendre deux messages $\mathbf{m}$ et $\mathbf{m}$ avec $\Delta\mathbf{m} \stackrel{\text{def}}{=} \mathbf{m} \oplus \mathbf{m}'$ donnée.
(2) Prédire la valeur $\Delta\mathbf{c} = \mathbf{c} \oplus \mathbf{c}'$.
(3) En déduire des informations sur la clé secrète.

Quelques formalités

Différence

La **différence** entre deux éléments $\mathbf{x}, \mathbf{x}'$ d'un groupe $(G, \otimes)$ est $\Delta\mathbf{x} \stackrel{\text{def}}{=} \mathbf{x} \otimes (\mathbf{x}')^{-1}$.

En pratique

On se contentera souvent de $G = \mathbb{F}_2^n$, et $\otimes$ sera alors l'addition (*i.e.,* le XOR bit à bit) : $\Delta\mathbf{x} = \mathbf{m} + \mathbf{m}'$, mais cette technique pourrait s'appliquer plus généralement (par exemple $G = (\mathbb{F}_q^n, +)$).

Dérivée d'une fonction booléenne

Soit $\alpha \in \mathbb{F}_2^n$ et $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$. La (fonction booléenne) **dérivée** de f en direction α est

$$\Delta_\alpha f(\cdot) \stackrel{\text{def}}{=} f(\cdot + \alpha) + f(\cdot).$$

Principes de la cryptanalyse différentielle

- On choisit des couples de clairs $\mathbf{x}, \mathbf{x}'$ de différence $\alpha \stackrel{\text{def}}{=} \Delta\mathbf{x}$ fixée.
- On considère leurs images $\mathbf{y} = f(\mathbf{x})$ et $\mathbf{y}' = f(\mathbf{x}')$ par une fonction booléenne $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$.
- On veut estimer $\beta \stackrel{\text{def}}{=} \Delta y = f(\mathbf{x} + \alpha) + f(\mathbf{x}) = \Delta_\alpha(f)(\mathbf{x})$.

- Un tel couple (α, β) est souvent noté $(\alpha \mapsto \beta)$ et est appellé une **différentielle** (possible) de f.
- La cryptanalyse différentielle cherche à exploiter l'existence de différentielles $(\alpha \mapsto \beta)$ qui apparaissent avec **grosse probabilité**.

Remarques importantes

Pour une fonction booléenne f **linéaire**, alors

$$(\Delta_\alpha f)(\mathbf{x}) = f(x+\alpha) + f(x) = f(\alpha)$$

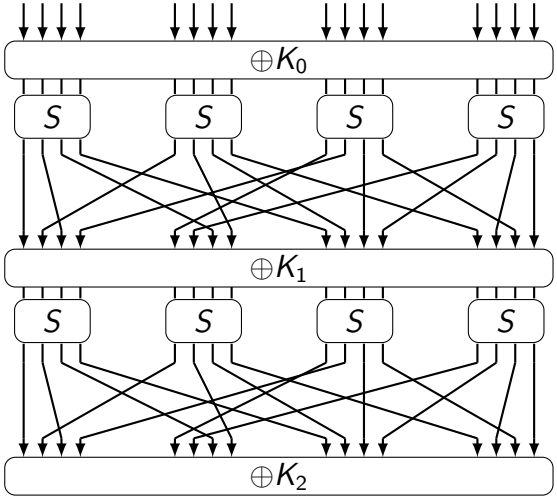
Pour une fonction linéaire, les seules différentielles possibles sont les $(\alpha \mapsto f(\alpha))$.

Pour une fonction booléenne **affine** $f = \ell + K_0$ où ℓ est **linéaire**, alors

$$(\Delta_\alpha f)(\mathbf{x}) = \ell(\alpha)$$

Une addition de clé **ne change pas** les différentielles.

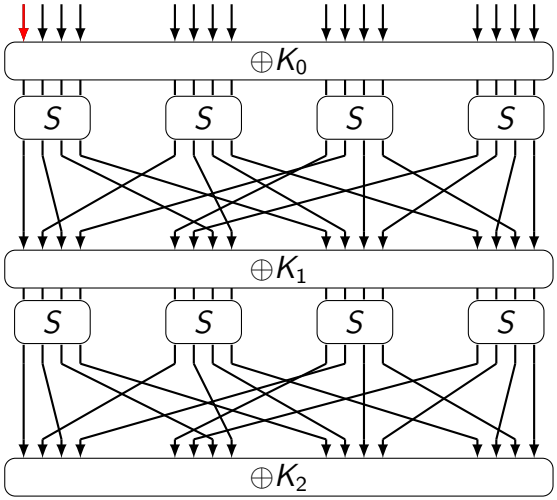
Un exemple



	00	01	10	11
00	2	0	4	3
01	9	5	6	7
10	1	d	e	f
11	a	8	c	b

Exemple : $S(9) = S(1001) = d$

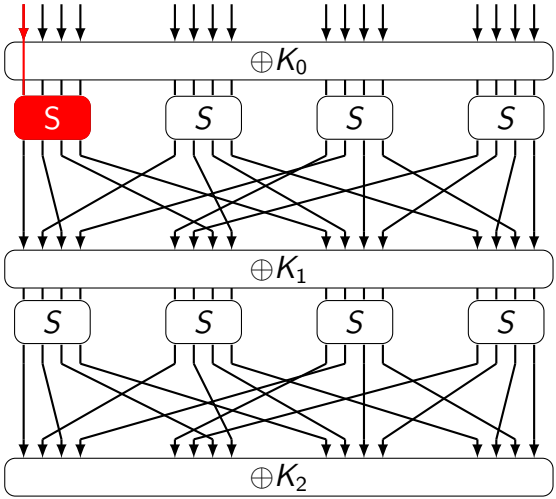
Une différence sur un bit



	00	01	10	11
00	2	0	4	3
01	9	5	6	7
10	1	d	e	f
11	a	8	c	b

On active la différentielle $\alpha = (8, 0, 0, 0)$

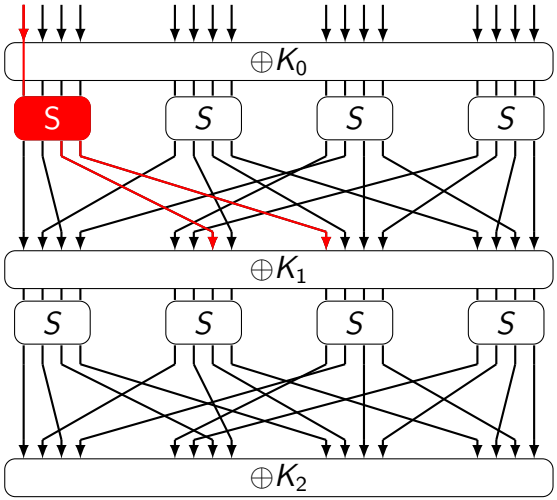
Une différence sur un bit



	00	01	10	11
00	2	0	4	3
01	9	5	6	7
10	1	d	e	f
11	a	8	c	b

Addition de clé ne change pas la différentielle.
On active la première boîte S.

Une différence sur un bit

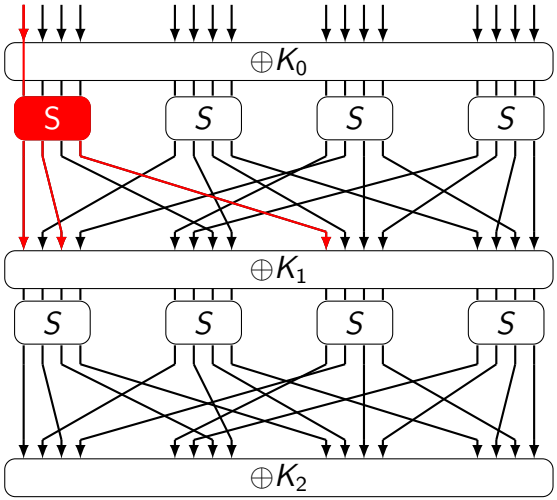


	00	01	10	11
00	2	0	4	3
01	9	5	6	7
10	1	d	e	f
11	a	8	c	b

$2 \oplus 1 = 3 = (0, 0, 1, 1)$

$(\Delta_{(1,0,0,0)} S)(\mathbf{x}) \in \{3, \}$

Une différence sur un bit

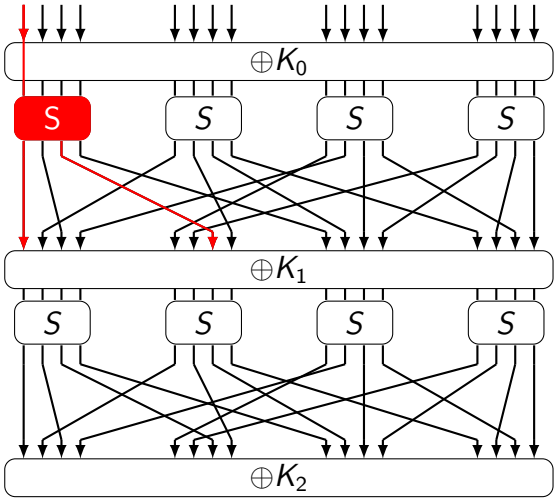


	00	01	10	11
00	2	0	4	3
01	9	5	6	7
10	1	d	e	f
11	a	8	c	b

$0 \oplus \mathrm{d} = \mathrm{d} = (1, 1, 0, 1)$

$(\Delta_{(1,0,0,0)} S)(\mathbf{x}) \in \{3, \mathrm{d}, \}$

Une différence sur un bit

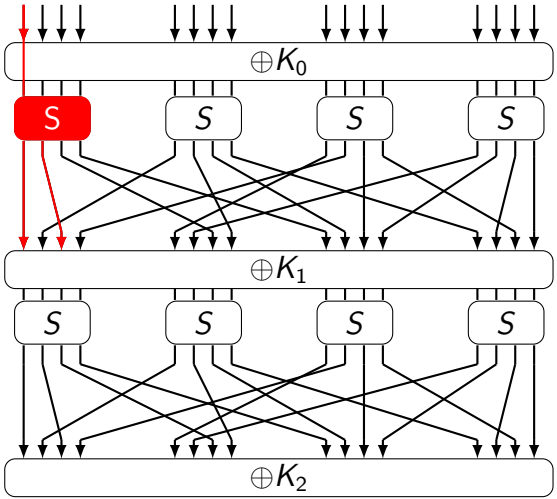


	00	01	10	11
00	2	0	4	3
01	9	5	6	7
10	1	d	e	f
11	a	8	c	b

$$4 \oplus \mathrm{e} = \mathrm{a} = (1,0,1,0)$$

$$(\Delta_{(1,0,0,0)} S)(\mathbf{x}) \in \{3, \mathrm{d}, \mathrm{a}, \}$$

Une différence sur un bit



	00	01	10	11
00	2	0	4	3
01	9	5	6	7
10	1	d	e	f
11	a	8	c	b

$$3 \oplus \mathrm{f} = \mathrm{c} = (1,1,0,0)$$

$$(\Delta_{(1,0,0,0)} S)(\mathbf{x}) \in \{3, \mathrm{d}, \mathrm{a}, \mathrm{c}\}$$

Nombre de solutions et Uniformité

Soit $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ une fonction booléenne, et $(\alpha \mapsto \beta)$ une différentielle. On note

$$\delta_f(\alpha, \beta) \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbb{F}_2^n \mid (\Delta_\alpha f)(\mathbf{x}) = \beta\}.$$

- Le tableau représentant $\#\delta_f(\alpha, \beta)$ pour toute $(\alpha \mapsto \beta)$ est appelé ***Difference Distribution Table*** (DDT).
- C'est une table de taille $2^n \times 2^n$.
- La valeur $\delta_f \stackrel{\text{def}}{=} \max_{\alpha \neq 0, \beta} \#\delta_f(\alpha, \beta)$ est appelée **Uniformité différentielle** de f.

Remarque : On a toujours $\delta_f(0, 0) = \mathbb{F}_2^n$. La différentielle $(0 \mapsto 0)$ est appelée différentielle triviale.

Probabilité d'une différentielle

La probabilité $\pi_f(\alpha, \beta)$ d'une différentielle $(\alpha \mapsto \beta)$ est la probabilité qu'elle apparaisse sous une entrée uniforme $\mathbf{x}$:

$$\pi_f(\alpha, \beta) = \mathbb{P}_{\mathbf{x}}(f(\mathbf{x} + \alpha) + f(\mathbf{x}) = \beta) = \frac{\#\delta_f(\alpha, \beta)}{2^n} \leqslant \frac{\delta_f}{2^n} \text{ pour } \alpha \neq 0.$$

- En cryptanalyse, on va chercher $(\alpha \mapsto \beta)$ offrant un **gros biais**.
- La résistance d'une S-box à la cryptanalyse différentielle est d'autant meilleure que son uniformité différentielle est **faible**.

DDT de la Sbox précédente

$\alpha \setminus \beta$	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	-	4	4	-	-	-	-	4	-	-	-	-	4	-	-	-
2	-	-	4	4	-	-	4	-	-	-	-	-	-	-	-	4
3	-	4	-	4	4	-	-	-	-	-	-	-	-	-	4	-
4	-	-	4	-	4	4	-	-	-	-	-	4	-	-	-	-
5	-	-	-	4	-	4	-	4	-	4	-	-	-	-	-	-
6	-	-	-	-	4	-	4	4	-	-	-	-	-	4	-	-
7	-	4	-	-	-	4	4	-	-	-	4	-	-	-	-	-
8	-	-	-	4	-	-	-	-	-	-	4	-	4	4	-	-
9	-	4	-	-	-	-	-	-	-	-	-	4	-	4	-	4
a	-	-	-	-	-	4	-	-	-	-	-	-	4	-	4	4
b	-	-	4	-	-	-	-	-	-	4	-	-	-	4	4	-
c	-	-	-	-	-	-	-	-	16	-	-	-	-	-	-	-
d	-	-	-	-	4	-	-	-	-	4	4	-	-	-	-	4
e	-	-	-	-	-	-	-	4	-	-	4	4	-	-	4	-
f	-	-	-	-	-	-	4	-	-	4	-	4	4	-	-	-

- La différentielle $(\mathrm{c} \mapsto 8)$ arrive avec probabilité 1.
- $(8 \mapsto \mathrm{a})$ et $(\mathrm{a} \mapsto \mathrm{c})$ arrivent toutes deux avec probabilité $1/4$.

Bonnes propriétés de confusion : Fonctions APN

(Nyberg, Knudsen 1993)

Soit $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ une fonction booléenne en n variables. Alors son uniformité différentielle vérifie $\delta(f) \geqslant 2$. Une fonction qui atteint cette borne est appelée *Almost Perfectly Nonlinear* (APN).

Preuve : $\delta(\alpha, \beta)$ est nécessairement pair, puisque si $\mathbf{x}$ vérifie $f(\mathbf{x}+\alpha) + f(\mathbf{x}) = \beta$, alors

$$f((\mathbf{x}+\alpha)+\alpha) + f(\mathbf{x}+\alpha) = f(\mathbf{x}) + f(\mathbf{x}+\alpha) = \beta.$$

Ainsi, pour toute solution x, alors $x + \alpha$ est aussi solution.

La conjecture APN

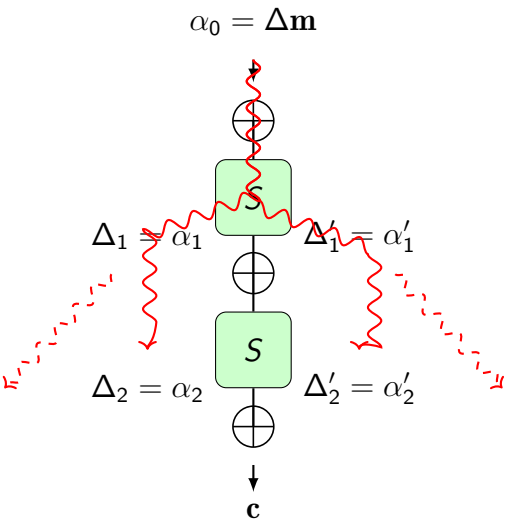
On ne connaît pas de **bijections** APN pour n pair, sauf pour $n = 6$.

Il était conjecturé pendant longtemps qu'il n'existait pas de bijections APN à un nombre pair de variables. La solution pour $n = 6$ date de 2009 (Dillon).

Les S-box dans l'AES sont des bijections $\mathbb{F}_2^8 \to \mathbb{F}_2^8$ qui vérifient $\delta(S) = 4$.

Gérer plusieurs tours

Multiplicité des différentielles



- Si trouver la ou les différentielles optimales pour les petites Sbox est facile, lorsque l'on agit sur tout un bloc cela devient très vite impossible.

- Il faut aussi tenir compte des couches assurant la **diffusion**.

- On va alors plutôt chercher une **borne inférieure** sur la probabilités d'une différentielle globale, en suivant un/des chemins particuliers.

Trace Différentielle

Définition : Soit E un chiffrement à r tours. Une **trace différentielle** (*Differential characteristic*) est un $(r+1)$-uplet $(\alpha_0, \ldots, \alpha_r)$ tel que $(\alpha_i \mapsto \alpha_{i+1})$ est une différentielle possible pour le tour i, et α_0 est la différence entre deux messages initiaux.

Si la fonction de tour était **linéaire**, alors il n'existerait qu'une seule trace différentielle issue d'une différence initiale α_0 :

$$\alpha_0 \mapsto F(\alpha_0) \mapsto F(F(\alpha_0)) \mapsto \cdots \mapsto F^r(\alpha_0).$$

Pour la sécurité, on veut **maximiser** le nombre de traces différentielles.

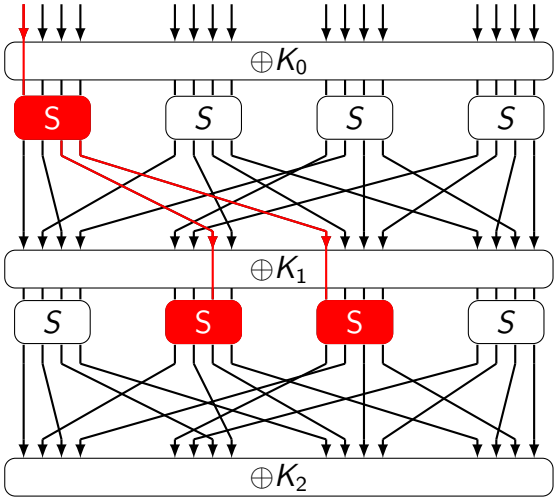
Probabilité d'une trace

Sous l'hypothèse raisonnable que les tours sont indépendants, la probabilité d'une trace est simplement le produit des probabilités de chaque différentielle intermédiaire :

$$\pi(\alpha_0, \ldots, \alpha_r) = \prod_{i=0}^{r-1} \pi_f(\alpha_i, \alpha_{i+1})$$

Remarque : La probabilité des traces ne dépend **que** de l'algorithme de chiffrement, et pas des données secrètes $\Longrightarrow$ précalculable.

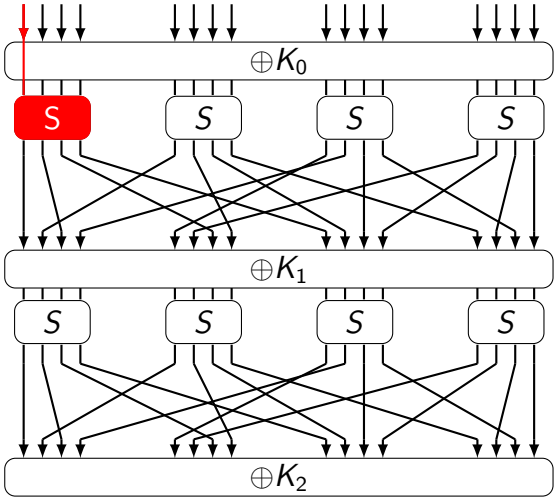
Retour de notre exemple



	00	01	10	11
00	2	0	4	3
01	9	5	6	7
10	1	d	e	f
11	a	8	c	b

La différentielle $(8,0,0,0) \mapsto (3,0,0,0)$ active 2 Sbox au tour 2.

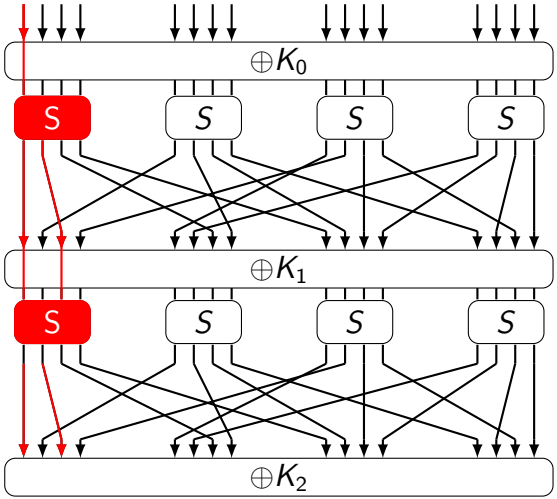
Retour de notre exemple



	00	01	10	11
00	2	0	4	3
01	9	5	6	7
10	1	d	e	f
11	a	8	c	b

Quel est le nombre **minimal** de Sbox activées au tour 2 ?

Retour de notre exemple



$(8,0,0,0) \mapsto (\mathrm{c},0,0,0) \mapsto (\mathrm{a},0,0,0) \mapsto \cdots$
n'active qu'1 seule Sbox

$(8,0,0,0) \mapsto (\mathrm{c},0,0,0) \mapsto (\mathrm{a},0,0,0)$
$\mapsto (\mathrm{c},0,0,0) \mapsto (\mathrm{a},0,0,0)$ est un
cycle possible, sur 2 tours
avec probabilité $\geqslant 2^{-2} \times 2^{-2} = 2^{-4}$.

Mauvaise propriété de **diffusion**.

Assurer une bonne diffusion

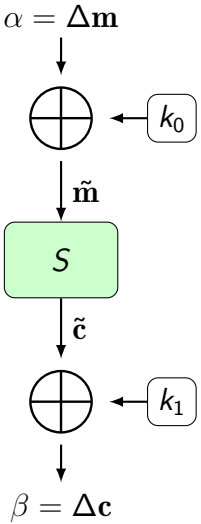
Nombre de branchement (*Branch Number*), Daemen 1995

- **Rappel** : La diffusion est assurée par les couches **linéaires**.
- **Définition :** Le **nombre de branchement** $\mathcal{B}$ d'un chiffrement par blocs est le nombre minimal de Sbox actives sur deux tours consécutifs.
- Se généralise à plus de tours.

- Dans l'exemple, on a $\mathcal{B} = 2$.
- Il existe une borne supérieure simple : $\mathcal{B} \leqslant 1+$ nombre de Sbox par tour.
- Dans l'AES on a 4 Sbox sur 32 bits (composition de $\mathrm{SubBytes}$ et $\mathrm{ShiftRows}$).
- La diffusion est assurée par $\mathrm{MixColumns}$.
- On peut vérifier que $\mathcal{B} = 5$, et après 4 tours $\mathcal{B} = 25$ qui est optimal.

Exploiter les Différentielles

Exemple avec un seul tour



(1) On a trouvé une bonne différentielle $(\alpha \to \beta)$ avec une probabilité p.
(2) On va générer N clairs $\mathbf{m}$ uniformes et calculer $\mathbf{m}' = \mathbf{m} \oplus \alpha$.
(3) On appelle notre **oracle de chiffrement** pour obtenir des couples de chiffrés $(\mathbf{c}, \mathbf{c}')$.
(4) Une paire $((\mathbf{m}, \mathbf{m}'), (\mathbf{c}, \mathbf{c}'))$ est **bonne** si $\Delta\mathbf{c} = \beta$. Il y en a $\approx pN$ pour N grand.
(5) On connaît les $\delta(\alpha, \beta) = p2^n$ paires $((\tilde{\mathbf{m}}, \tilde{\mathbf{m}}'), (\tilde{\mathbf{c}}, \tilde{\mathbf{c}}'))$ intermédiaires ayant la différentielle $(\alpha \to \beta)$.
(6) k_0 doit être de la forme $\mathbf{m} \oplus \tilde{\mathbf{m}}$ pour l'une de ces paires intermédiaires.

Comparaison avec la recherche exhaustive

Recherche exhaustive

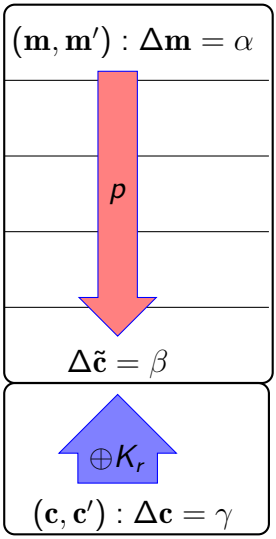
- 1 clair connu $(\mathbf{m}, \mathbf{c})$.
- Clé $= (k_0, k_1)$, espace de taille 2^{2n}.
- En réalité, k_0 et k_1 sont liés $\rightarrow 2^n$.

Cryptanalyse différentielle

- N clairs **choisis** pour trouver **une** bonne paire.
- Beaucoup de mauvaises paires ($\Delta\mathbf{m} = \alpha$ mais $\Delta\mathbf{c} \neq \beta$).
- On recherche k_0 dans un espace de taille $p2^n \rightarrow$ on a **décimé** l'espace des clés possibles.

- Si p est grand, il est **facile** de trouver une différentielle ($\alpha \mapsto \beta$) mais espace de recherche plus gros.
- $N = \Theta(\frac{1}{p})$. En pratique $N \approx 3 \times \frac{1}{p}$.

Attaque sur le dernier tour



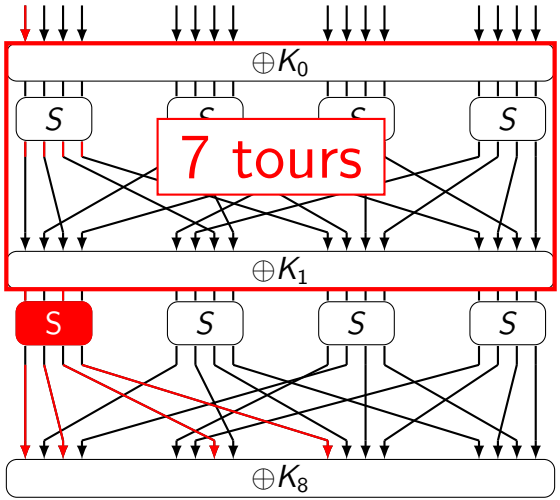
- On suppose avoir une trace différentielle ($\alpha \mapsto \beta$) avec probabilité $p >> 2^{-|\text{taille de block}|}$.
- On appelle l'**oracle de chiffrement** sur $N = \Theta(1/p)$ clairs de différence $\Delta\mathbf{m} = \alpha$.
- Pour chaque clé K_r possible (**recherche exhaustive**) on déchiffre 1 tour et on augmente un compteur si on observe $\Delta\tilde{\mathbf{c}} = \beta$.
- Vote majoritaire pour deviner la bonne clé K_r.
- Là encore, beaucoup de faux positifs (on peut observer $\Delta\tilde{\mathbf{c}} = \beta$ par chance).

On continue avec notre exemple : Encore la DDT

$\alpha \setminus \beta$	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	-	4	4	-	-	-	-	4	-	-	-	-	4	-	-	-
2	-	-	4	4	-	-	4	-	-	-	-	-	-	-	-	4
3	-	4	-	4	4	-	-	-	-	-	-	-	-	-	4	-
4	-	-	4	-	4	4	-	-	-	-	-	4	-	-	-	-
5	-	-	-	4	-	4	-	4	-	4	-	-	-	-	-	-
6	-	-	-	-	4	-	4	4	-	-	-	-	-	4	-	-
7	-	4	-	-	-	4	4	-	-	-	4	-	-	-	-	-
8	-	-	-	4	-	-	-	-	-	-	4	-	4	4	-	-
9	-	4	-	-	-	-	-	-	-	-	-	4	-	4	-	4
a	-	-	-	-	-	4	-	-	-	-	-	-	4	-	4	4
b	-	-	4	-	-	-	-	-	-	4	-	-	-	4	4	-
c	-	-	-	-	-	-	-	-	16	-	-	-	-	-	-	-
d	-	-	-	-	4	-	-	-	-	4	4	-	-	-	-	4
e	-	-	-	-	-	-	-	4	-	-	4	4	-	-	4	-
f	-	-	-	-	-	-	4	-	-	4	-	4	4	-	-	-

- La différentielle $(8 \mapsto c)$ a probabilité $1/4$.
- $(c, 0, 0, 0) \mapsto (a, 0, 0, 0)$ par la permutation, et n'active qu'une seule Sbox.
- Après 2 Sbox, la trace $(8 \mapsto c)$ a probabilité $\geqslant 2^{-4}$.
- Après 7 tours complets, la trace $(8 \mapsto a)$ a probabilité $\geqslant 2^{-14}$.

On continue l'exemple avec 8 tours

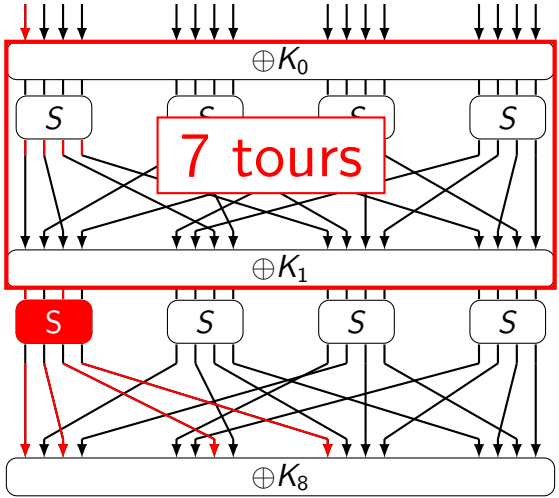


$\Delta = (8, 0, 0, 0)$

$p \geqslant 2^{-14}$

$\Delta = (a, 0, 0, 0)$

Complexité



- On a besoin de $\approx 3 \times \frac{1}{p} = 75\%$ de l'espace total !

- **Remarque :** On peut en réalité filtrer les paires de chiffrés qui ont une différence non nulle sur les bits non actifs.

- On peut retrouver toutes les clés K_8 qui ont les bon bits activés : Beaucoup moins cher et beaucoup moins de données.

- Brute-force les 12 bits restants ou nouvelle trace.

Pour aller plus loin

- Il existe des outils pour aider à automatiser ces attaques statistiques. Particulièrement utile pour les designers, mais aussi pour les cryptanalystes. Par exemple **Mixed Integer Linear Programming** (MILP).
- La cryptanalyse différentielle admet de nombreuses variantes : Différentielles impossibles, cryptanalyse boomerang, différentielles d'ordre supérieur.

Acknowledgement

Ces transparents sont très largement inspirés
du cours donné par Maria Eichlseder en 2021
https://www.youtube.com/watch?v=GQX8W8zKf2Q

Les exemples étant extrêmement bien choisis.