

Cours 05 - Cryptanalyse Différentielle

Maxime Bombar

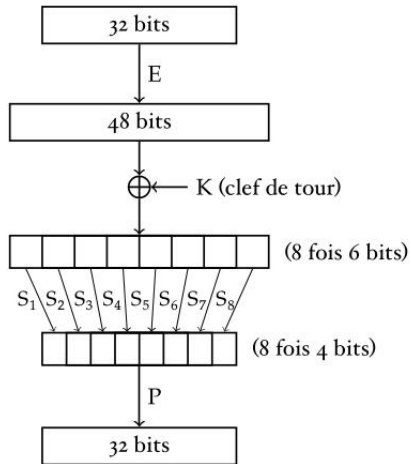
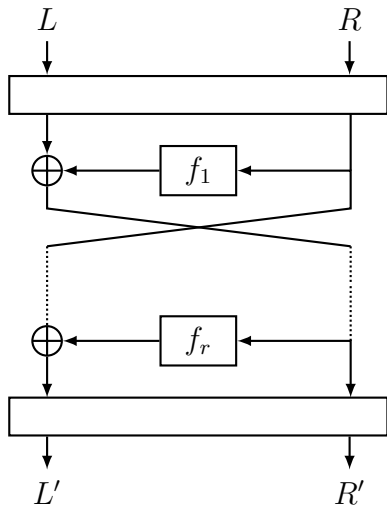
Introduction

Modes d'Opérations de Chiffrement par Blocs

- Permettent d'étendre une **primitive** opérant sur un bloc, à des messages de taille arbitraire.
- Sécurité **prouvée**, souvent si la primitive est une **permutation aléatoire**.
- Attention au choix du padding (cf Attaque de Vaudenay, Challenge 04)
- Certains modes permettent d'assurer aussi l'**intégrité**.

Objectif : Analyser les primitives de chiffrement par blocs.

Rappel : Schémas de Feistel (e.g. DES)



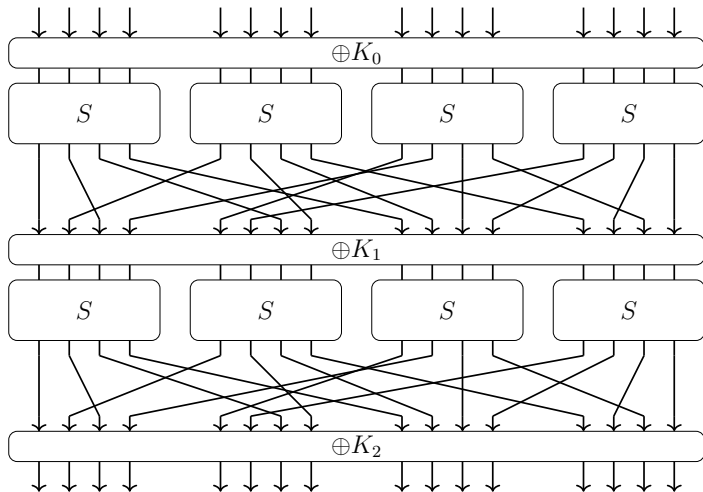
Faiblesse des fonctions de tour

V. Rijmen, B. Preneel (1997)

Attaques exploitant la non surjectivité des fonctions de tour pour monter une attaque.

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

Rappel : Réseaux de Substitutions-Permutations (SPN)



Fonctions de tour **bijectives** :

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

Diffusion : (ou Mixing) Permutation **linéaire** bit à bit.

Exemple : AES
(Daemen, Rijmen)

Cryptanalyse Différentielle

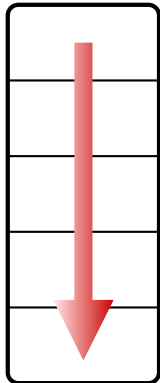
- Technique de cryptanalyse proposée par Biham et Shamir en 1990 pour DES.
- Certaines variantes étaient déjà connues des concepteurs de DES (IBM, NSA).

Idées principales

- On veut exploiter la **structure** de la primitive pour observer un **biais** statistique la **distinguant** d'une permutation aléatoire.
- *Chosen-Plaintext Attack*
- Pour deux messages ($\mathbf{m}, \mathbf{m}'$) différents, leurs images ($\mathbf{c}, \mathbf{c}'$) devraient être uniformément aléatoires, et idéalement indépendantes.
- **Idée** : prédire l'effet sur les chiffrés d'une légère différence $\Delta \mathbf{M} = \mathbf{m} \oplus \mathbf{m}'$.

Distingueur Différentiel

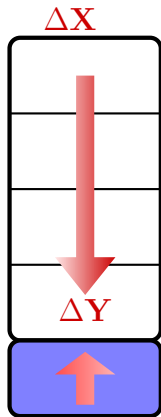
$$\Delta \mathbf{M} = \mathbf{m} \oplus \mathbf{m}'$$



$$\Delta \mathbf{C} = \mathbf{c} \oplus \mathbf{c}'$$

- Dans une permutation aléatoire, tous les $\Delta \mathbf{C}$ devraient être **équiprobables**.
- On estime théoriquement la probabilité $\mathbb{P}(\Delta \mathbf{C} = \mathbf{y} \mid \Delta \mathbf{M} = \mathbf{x})$.
- On évalue empiriquement si on observe (ou non) cette probabilité.
- Si le biais est **suffisamment** important, on a un **distingueur**.

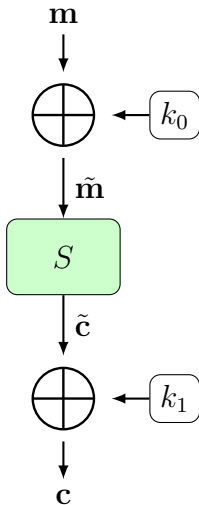
Exploiter le Distingueur



Recherche exhaustive intelligente sur la clé de dernier tour.

Caractéristiques Différentielles

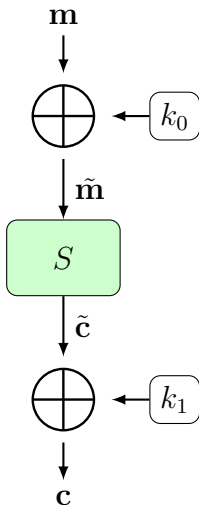
Mise en bouche



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

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

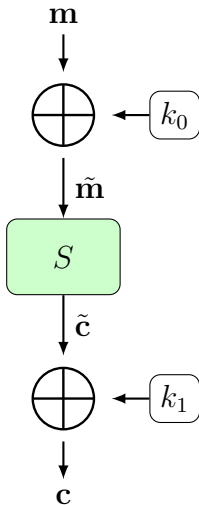
Mise en bouche



Hypothèse : Attaque par clairs connus.
On connaît des paires (m, c) .

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

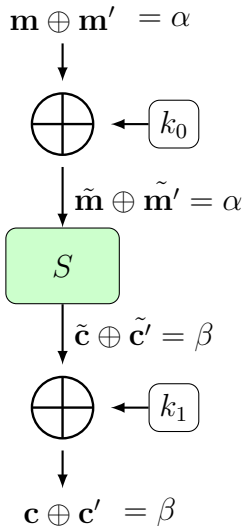
Mise en bouche



Hypothèse : Attaque par clairs connus.
On connaît des paires (m, c) .

- **Imaginez** que l'on connaisse en plus $\tilde{c}$.
- On pourrait en déduire k_1 , puis tout le reste.
- **Problème :** Avec k_0 clé uniforme, et S bijective, $\tilde{c}$ est **uniformément distribué**.

Mise en bouche



Hypothèse : Attaque par clairs connus.
On connaît des paires (m, c) .

- Imaginez que l'on connaisse en plus $\tilde{c}$.
- On pourrait en déduire k_1 , puis tout le reste.
- **Problème :** Avec k_0 clé uniforme, et S bijective, $\tilde{c}$ est **uniformément distribué**.
- **Idée :** $(m \oplus k_0) \oplus (m' \oplus k_0) = m \oplus m'$

Une remarque sur les différentielles

Différence

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

En pratique

On se contentera souvent de $G = \mathbb{F}_2^n$, et $\odot$ sera alors l'addition (*i.e.* le XOR bit à bit) :

$$\Delta x = m \oplus m',$$

mais cette technique pourrait s'appliquer plus généralement, comme par exemple avec des primitives définies sur des alphabets différents comme $\mathbb{F}_q$ (*e.g.* primitives dites « orientées vers l'arithmétisation ») ou $\mathbb{Z}/2^k\mathbb{Z}$ (*e.g.* chiffrement FEAL, cryptanalyse de MD5, SHA1, ...).

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 $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$.
 - On veut estimer $\beta \stackrel{\text{def}}{=} \Delta \mathbf{y} = f(\mathbf{x} + \alpha) + f(\mathbf{x}) = \Delta_\alpha(f)(\mathbf{x})$.
-
- Un tel couple (α, β) (parfois noté $(\alpha \mapsto \beta)$) est appelé une **différentielle** (possible) de f .
 - La cryptanalyse différentielle cherche à exploiter l'existence de différentielles $(\alpha \mapsto \beta)$ qui apparaissent avec **grosse probabilité**.

Le cas des chiffrements par blocs

- Un chiffrement est en général une fonction **très compliquée** (c'est l'objectif).
- Elle admet cependant une **structure** :
 - Elle est formée de plusieurs tours **nettement plus simples**
 - Les différentes boîtes S opèrent sur un **petit nombre** de bits (e.g. 8 bits pour l'AES, à comparer aux 128 bits d'un bloc).
- On va en général se concentrer sur quelques tours uniquement (**voire un seul**), et essayer de **propager les différentielles** qu'on trouve au reste du chiffrement.

Remarques importantes

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

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

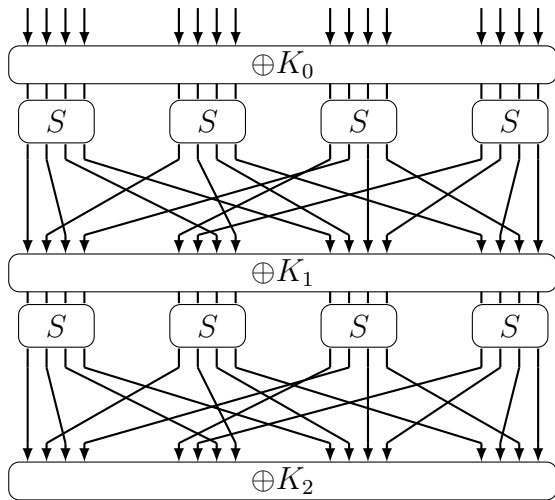
Pour une fonction linéaire (e.g. couches de diffusion), les seules différentielles possibles sont les $(\alpha \mapsto f(\alpha))$.

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.

Exemple sur un Chiffrement Jouet

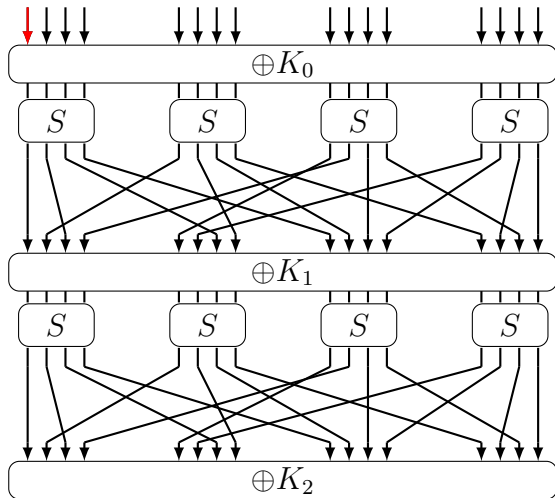


x	0	1	2	3	4	5	6	7
$S(x)$	2	0	4	3	9	5	6	7

x	8	9	a	b	c	d	e	f
$S(x)$	1	d	e	f	a	8	c	b

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

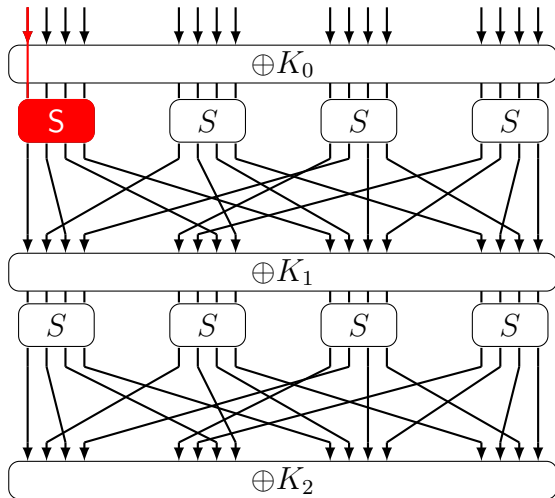
Une différence sur un bit



Correspond à la différentielle

$$\Delta \mathbf{X} = (1000, \mathbf{0}, \mathbf{0}, \mathbf{0}) = (8, 0, 0, 0)$$

Une différence sur un bit



- Addition de clé $\rightarrow$ **ne change pas** la différentielle.
- On **active** la première boîte S , avec la différentielle $\Delta_{\text{in}} = 8$.

Action de la Boîte S sur la Différence

$\Delta_{\text{in}} = 1000_2 = 8$. Peut-on prédire Δ_{out} en sortie de S ?

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	2	0	4	3	9	5	6	7	1	d	e	f	a	8	c	b

Action de la Boîte S sur la Différence

$$\Delta_{\text{in}} = 1000_2 = 8. \Delta_{\text{out}} \in \{3, \dots\}$$

$\Delta_{\text{in}} = 8$

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	2	0	4	3	9	5	6	7	1	d	e	f	a	8	c	b

$\Delta_{\text{out}} = 2 \oplus 1 = 0010_2 \oplus 0001_2 = 0011_2 = 3$

Action de la Boîte S sur la Différence

$$\Delta_{\text{in}} = 1000_2 = 8. \Delta_{\text{out}} \in \{3, d, \dots\}$$

$\Delta_{\text{in}} = 8$

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	2	0	4	3	9	5	6	7	1	d	e	f	a	8	c	b

$\Delta_{\text{out}} = 0 \oplus d = 0000_2 \oplus 1101_2 = 1101_2 = d$

Action de la Boîte S sur la Différence

$$\Delta_{in} = 1000_2 = 8. \Delta_{out} \in \{3, d, a, \dots\}$$

$\Delta_{in} = 8$

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	2	0	4	3	9	5	6	7	1	d	e	f	a	8	c	b

$\Delta_{out} = 4 \oplus e = 0100_2 \oplus 1110_2 = 1010_2 = a$

Action de la Boîte S sur la Différence

$$\Delta_{\text{in}} = 1000_2 = 8. \Delta_{\text{out}} \in \{3, \text{d}, \text{a}, \text{c}, \dots\}$$

$\Delta_{\text{in}} = 8$

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	2	0	4	3	9	5	6	7	1	d	e	f	a	8	c	b

$\Delta_{\text{out}} = 3 \oplus \text{f} = 0011_2 \oplus 1111_2 = 1100_2 = \text{c}$

Action de la Boîte S sur la Différence

$$\Delta_{\text{in}} = 1000_2 = 8. \Delta_{\text{out}} \in \{3, d, a, c\}$$

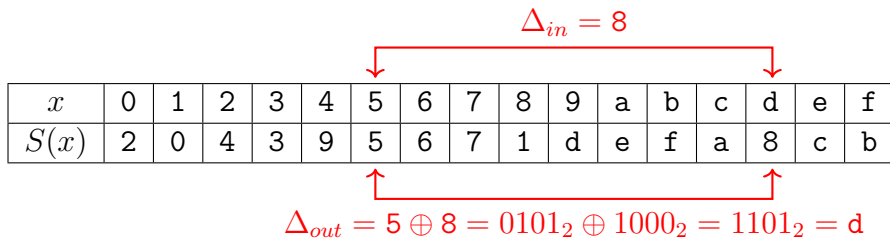
$\Delta_{\text{in}} = 8$

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	2	0	4	3	9	5	6	7	1	d	e	f	a	8	c	b

$\Delta_{\text{out}} = 9 \oplus a = 1001_2 \oplus 1010_2 = 0011_2 = 3$

Action de la Boîte S sur la Différence

$$\Delta_{\text{in}} = 1000_2 = 8. \Delta_{\text{out}} \in \{3, d, a, c\}$$



Tables des différences et Uniformité

Soit $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ 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 $\mathbf{x}$:

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

- 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 notre Sbox

$\alpha \backslash \beta$	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	-	-	-

On observe bien que

$$\Delta_{\text{in}} = 8 \Rightarrow \Delta_{\text{out}} \in \{3, a, c, d\}$$

DDT de notre Sbox

$\alpha \backslash \beta$	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 ($c \mapsto 8$) arrive avec probabilité 1 $\rightarrow$ **Problématique.**
- ($8 \mapsto a$) et ($a \mapsto c$) arrivent toutes deux avec probabilité $1/4$.

Bonnes propriétés de confusion : Fonctions APN

(Nyberg, Knudsen 1993)

Soit $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ une fonction booléenne en n variables. Alors son uniformité différentielle vérifie $\delta(f) \geq 2$. 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

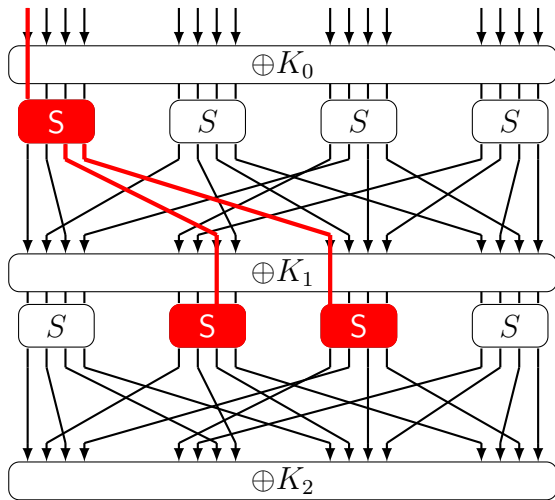
On ne connaît pas de **bijections** APN pour n pair, sauf pour $n = 6$.

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 $\mathbb{F}_2^8 \rightarrow \mathbb{F}_2^8$ qui vérifient $\delta(S) = 4$.

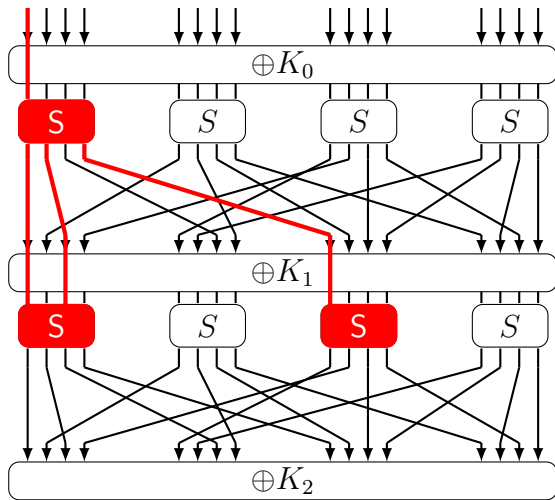
Gérer plusieurs tours

Propagation de la différence et Diffusion



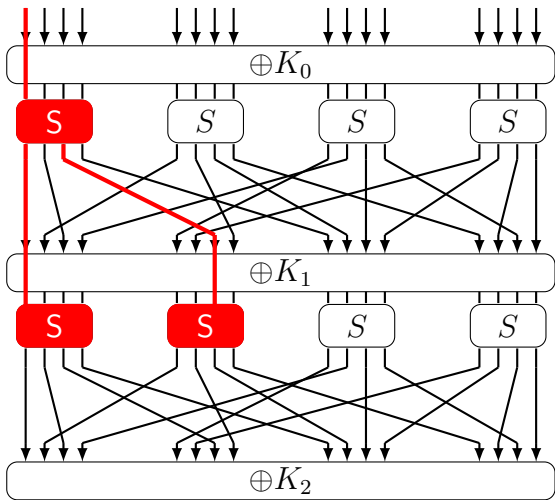
$$\Delta_{in} = 8 \Rightarrow \Delta_{out} \in \{3, d, a, c\}$$

Propagation de la différence et Diffusion



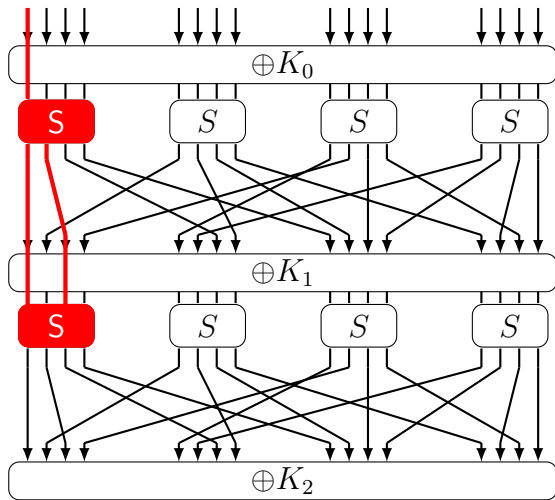
$$\Delta_{in} = 8 \Rightarrow \Delta_{out} \in \{3, \mathbf{d}, a, c\}$$

Propagation de la différence et Diffusion



$$\Delta_{in} = 8 \Rightarrow \Delta_{out} \in \{3, d, \mathbf{a}, c\}$$

Propagation de la différence et Diffusion



$$\Delta_{in} = 8 \Rightarrow \Delta_{out} \in \{3, d, a, \mathbf{c}\}$$

La couche linéaire ne **diffuse** la différentielle ($8 \mapsto c$) qu'à **une seule** boîte S de la couche suivante!

Chemin Différentiel

Définition : Soit E un chiffrement à r tours. Une **trace différentielle** ou **chemin différentiel** (*Differential characteristic*) est un $(r + 1)$ -uplet $(\alpha_0, \dots, \alpha_r)$ 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 :

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

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

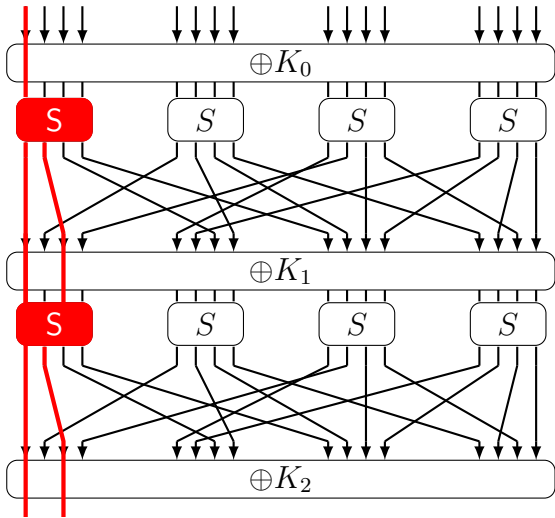
Probabilité d'une trace

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

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

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

Retour de notre exemple



- Couche 1 : $(\Delta_{\text{in}}^{(1)}, \Delta_{\text{out}}^{(1)}) = (8, c)$ avec probabilité $p_1 = 2^{-2}$.
- Couche linéaire L1 : $(c \mapsto a)$, une seule boîte S active.
- Couche 2 + L2 : (a, a) avec probabilité $p_2 = 2^{-2}$.
- Après 2 tours, $(\Delta_{\text{in}}^{(1)}, \Delta_{\text{out}}^{(2)}) = (8, a)$ avec probabilité $\geq 2^{-2} \times 2^{-2} = 2^{-4}$.
On peut itérer.
- Après r tours, $(8, a)$ se produit donc avec probabilité $\geq 2^{-2r}$.

Assurer une bonne diffusion

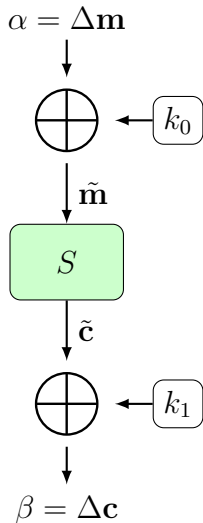
Nombre de branchement (*Branch Number*), Daemen 1995

Le **nombre de branchement** $\mathcal{B}$ d'un chiffrement par blocs est le nombre minimal de Sbox actives sur deux tours consécutifs.

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

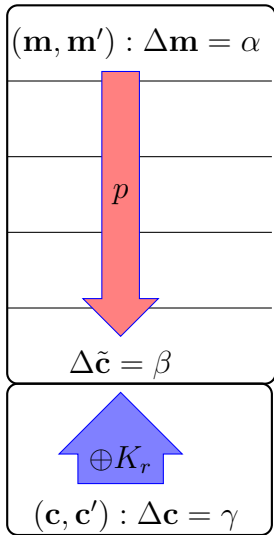
Exploiter les Différentielles

Exemple avec un seul tour



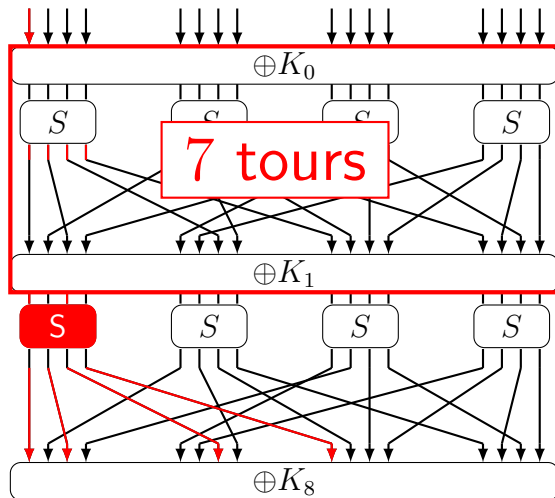
- (1) On a trouvé une bonne différentielle ($\alpha \rightarrow \beta$) avec une probabilité p .
- (2) On va générer N clairs $\mathbf{m}$ uniformes et calculer $\mathbf{m}' = \mathbf{m} \oplus \alpha$.
- (3) On appelle notre **oracle de chiffrement** pour obtenir des couples de chiffrés $(\mathbf{c}, \mathbf{c}')$.
- (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 ($\alpha \rightarrow \beta$).
- (6) k_0 doit être de la forme $\mathbf{m} \oplus \tilde{\mathbf{m}}$ pour l'une de ces paires intermédiaires.

Attaque sur le dernier tour



- On suppose avoir une trace différentielle $(\alpha \mapsto \beta)$ avec probabilité $p \gg 2^{-|\text{taille de bloc}|}$.
- On appelle l'**oracle de chiffrement** sur $N = \Theta(1/p)$ clairs de différence $\Delta \mathbf{m} = \alpha$. En pratique $N \approx 3 \times 1/p$.
- Pour chaque clé K_r possible (**recherche exhaustive**) on déchiffre 1 tour et on augmente un compteur si on observe $\Delta \tilde{\mathbf{c}} = \beta$. Ne regarder que les bits actifs!
- 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 l'exemple avec 8 tours

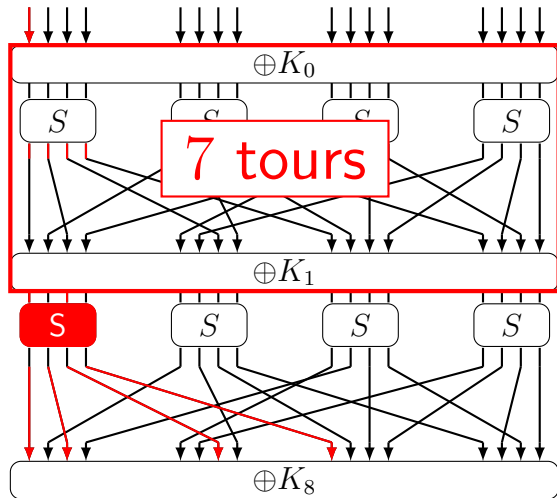


$$\Delta = (8, 0, 0, 0)$$

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

$$p \geq 2^{-14}$$

Complexité



- On a besoin d'environ $3 \times 2^{14} \approx 50000$ paires (clair, chiffré) pour exploiter la différentielle.
- On peut retrouver les bits des clés K_8 qui correspondent aux boîtes S actives (ici 4 bits).
- Brute-force les 12 bits restants ou nouvelle différentielle.

À retenir

- Pour une bonne confusion, il faut des boîtes S avec une petite uniformité différentielle.
- Pour une bonne diffusion, il faut un gros nombre de branchement.
- Dans le cas de l'AES, on est quasiment optimal !

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.