Soutenance de stage de L3

Cryptanalyse linéaire expérimentale de DES

Lucas Pesenti [1]

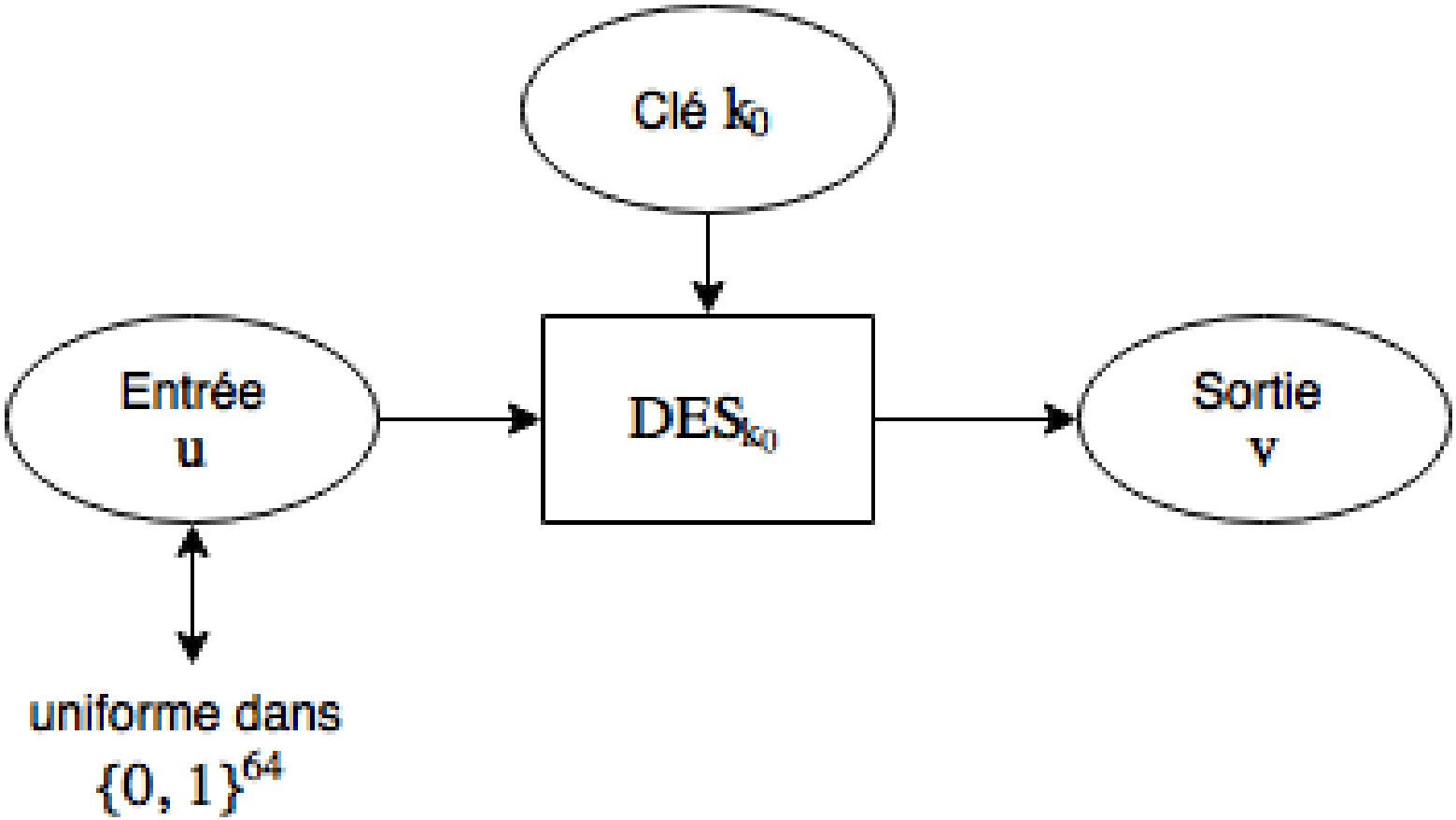
juin – juillet 2017

1. Sous la direction de François-Xavier Standaert dans l'équipe UCL Crypto de Louvain-la-Neuve (Belgique).

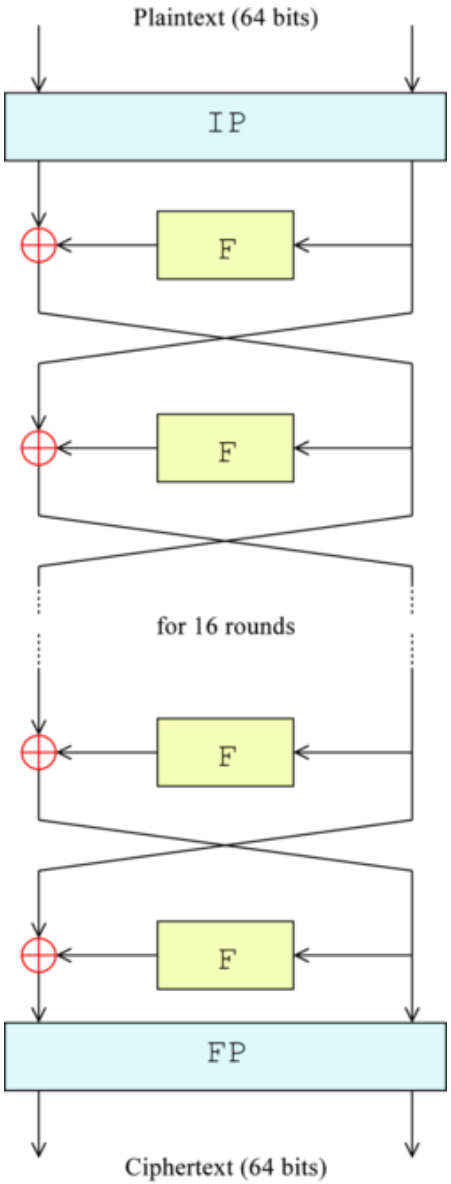
1. Vue d'ensemble de l'attaque
   - Présentation de DES
   - Fonctionnement général de l'attaque
2. Génération des approximations linéaires
   - Généralités
   - Combinaison d'approximations
   - Génération des approximations sur 1 round
   - Résultats
3. Phase d'attaque
   - Cryptanalyse multilinéaire (modèle d'Hermelin)
   - Optimisation temporelle du déchiffrement partiel
   - Estimation de rang
   - Résultats de l'attaque

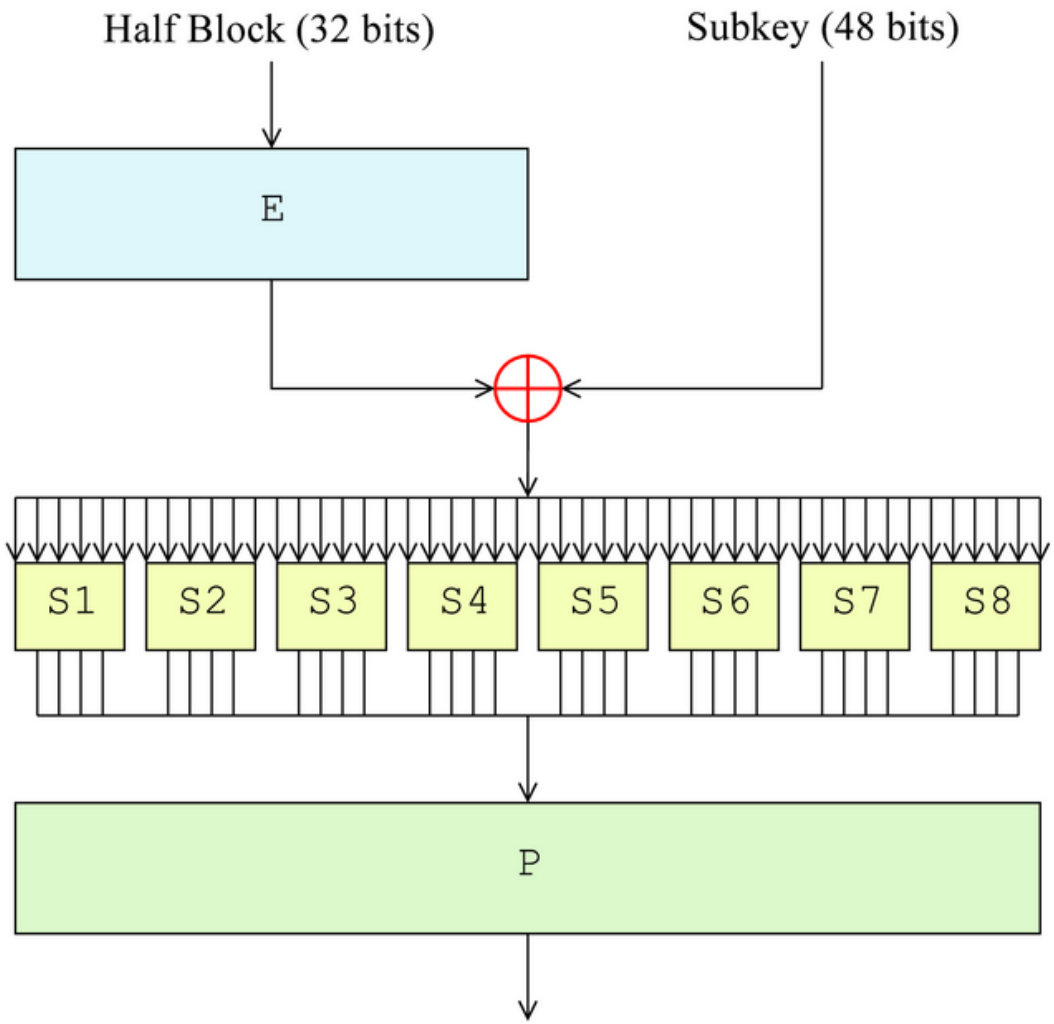
Présentation de DES

- Entrée sur 64 bits.
- Clé sur 56 bits.



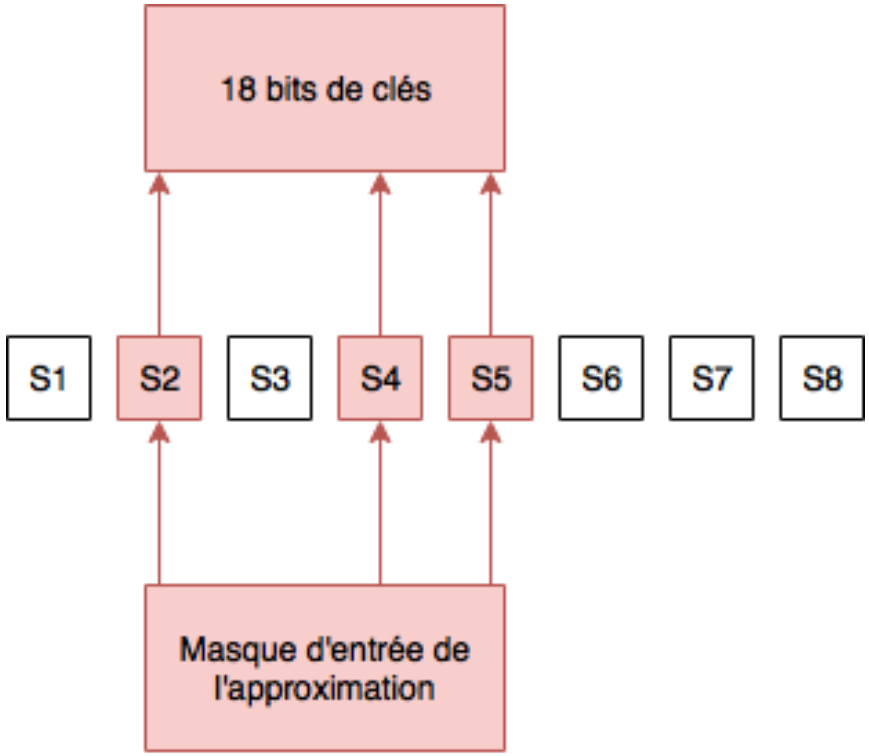
Présentation de DES (suite)





Fonctionnement général de l'algorithme 2 de Matsui

- N couples clair/chiffré.
- Approximations linéaires biaisées entre l'entrée et la sortie.
- Déchiffrement partiel.
- **Rang** de la clé : position dans le parcours exhaustif final.
- **Avantage** a si le rang est inférieur à 2^{56-a}.



Phase de cryptanalyse linéaire : associer à chacune des 2^{18} classes de clés un score.

Déchiffrement partiel

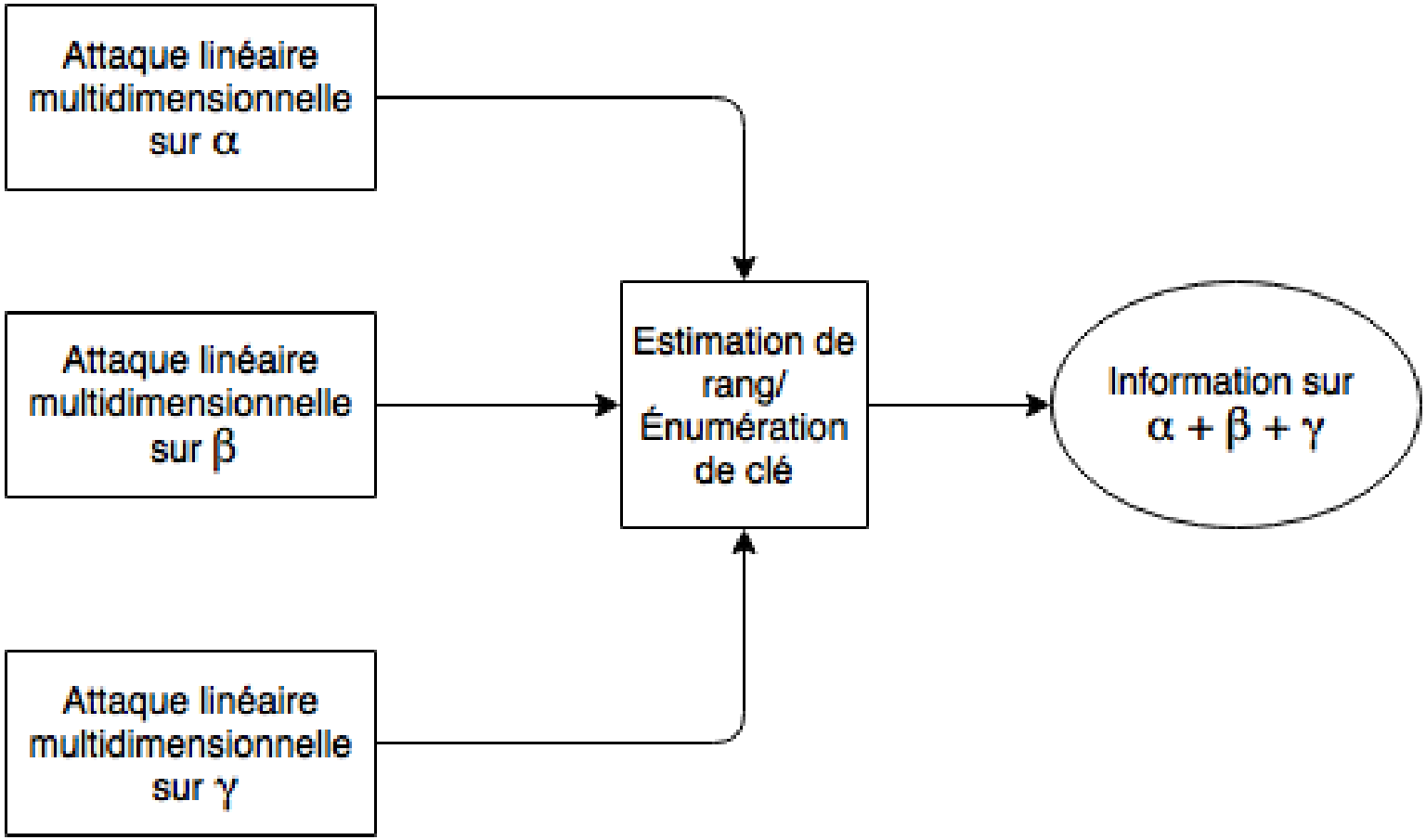
Nombre de bits de clé de dépendance :

Identifiant de la S-box j	1	2	3	4	5	6	7	8
Niveau $i = 1$	6	6	6	6	6	6	6	6
Niveau $i = 2$	38	40	39	36	37	36	38	39
Niveau $i = 3$	53	55	52	56	54	55	54	55
Niveau $i = 4$	56	56	56	56	56	56	56	56

1 round déchiffré partiellement en entrée, 1 en sortie.

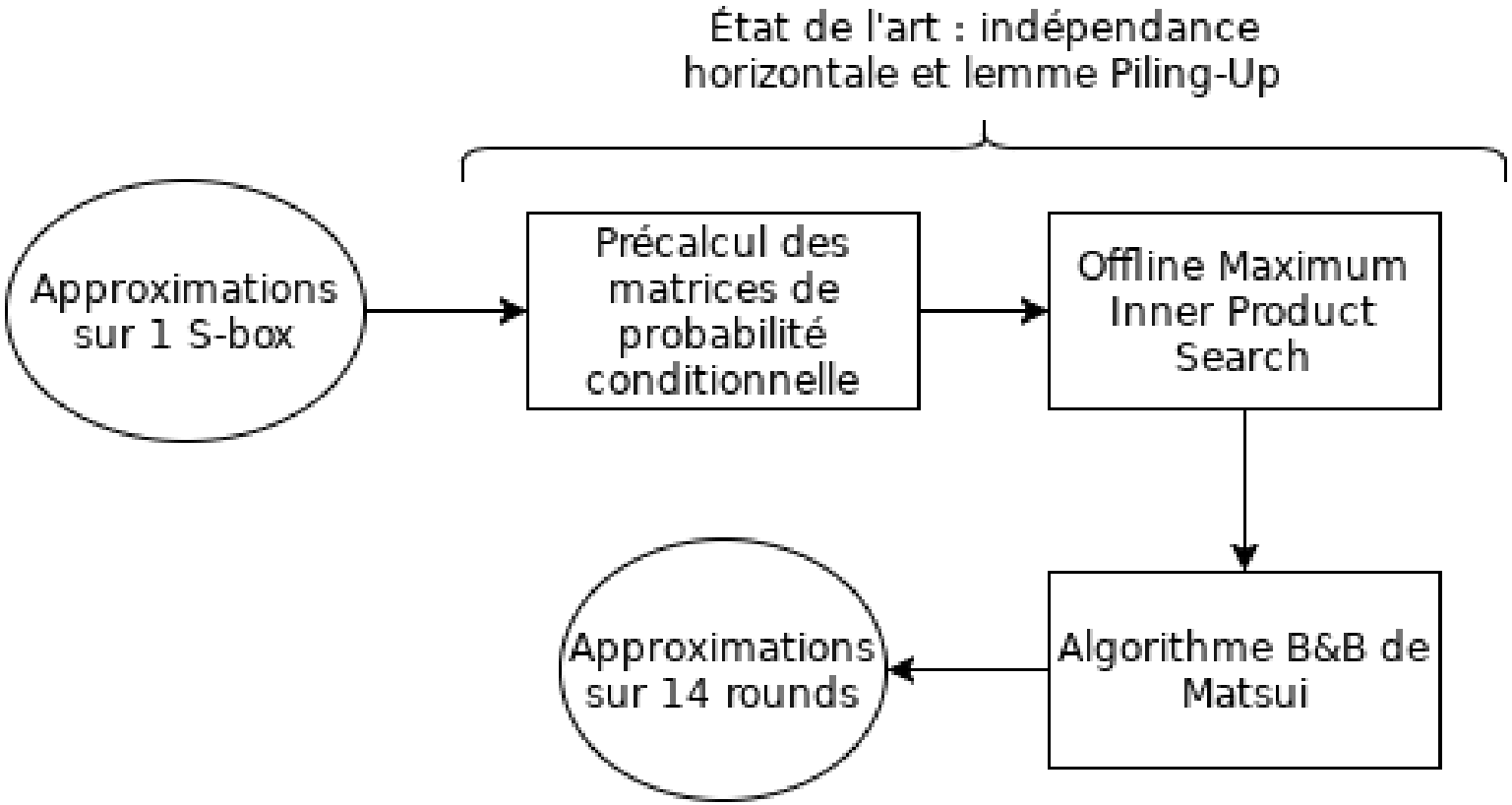
Fonctionnement général de l'attaque

- Attaque linéaire multidimensionnelle.
- Combinaison de plusieurs attaques indépendantes.



α, β, γ : masques de la clé

Schéma de la génération des approximations linéaires



Définitions

Définition (approximation linéaire)

- *Approximation linéaire* $\alpha \cdot u + \beta \cdot v = \gamma \cdot k_0$ *de biais :*

$$\epsilon = \left| P(\alpha \cdot u + \beta \cdot v = \gamma \cdot k_0) - \frac{1}{2} \right|$$

- *Système de m approximations linéaires* $Au + Bv = \Gamma k_0$ *de capacité :*

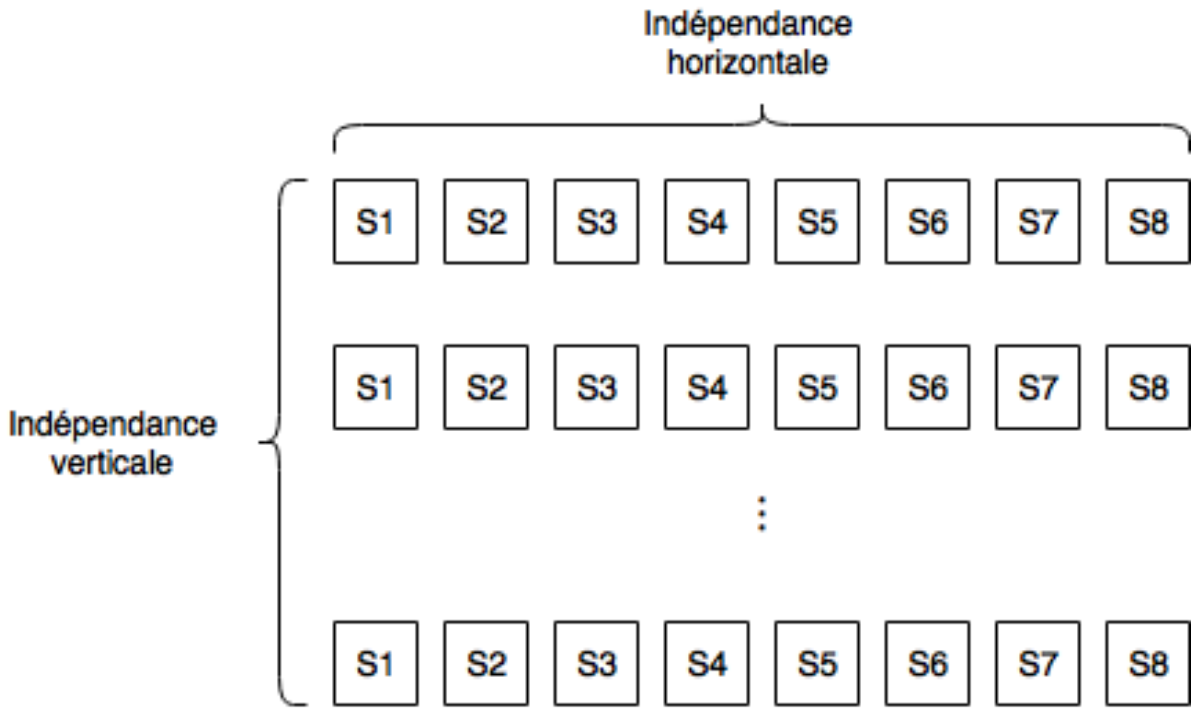
$$C = \sum_{\eta=0}^{2^m-1} \left(P(Au + Bv = \eta) - \frac{1}{2^m} \right)^2$$

Combinaison d'approximations

Lemma (Piling-Up)

Si $X_1, \ldots X_n$ *sont des variables aléatoires indépendantes à valeurs dans* $\{0, 1\}$, *alors :*

$$P(X_1 + \ldots + X_n = 0) = \frac{1}{2} + 2^{n-1} \prod_{i=1}^{n} \left(P(X_i = 0) - \frac{1}{2} \right)$$



Hypothèse d'indépendance horizontale

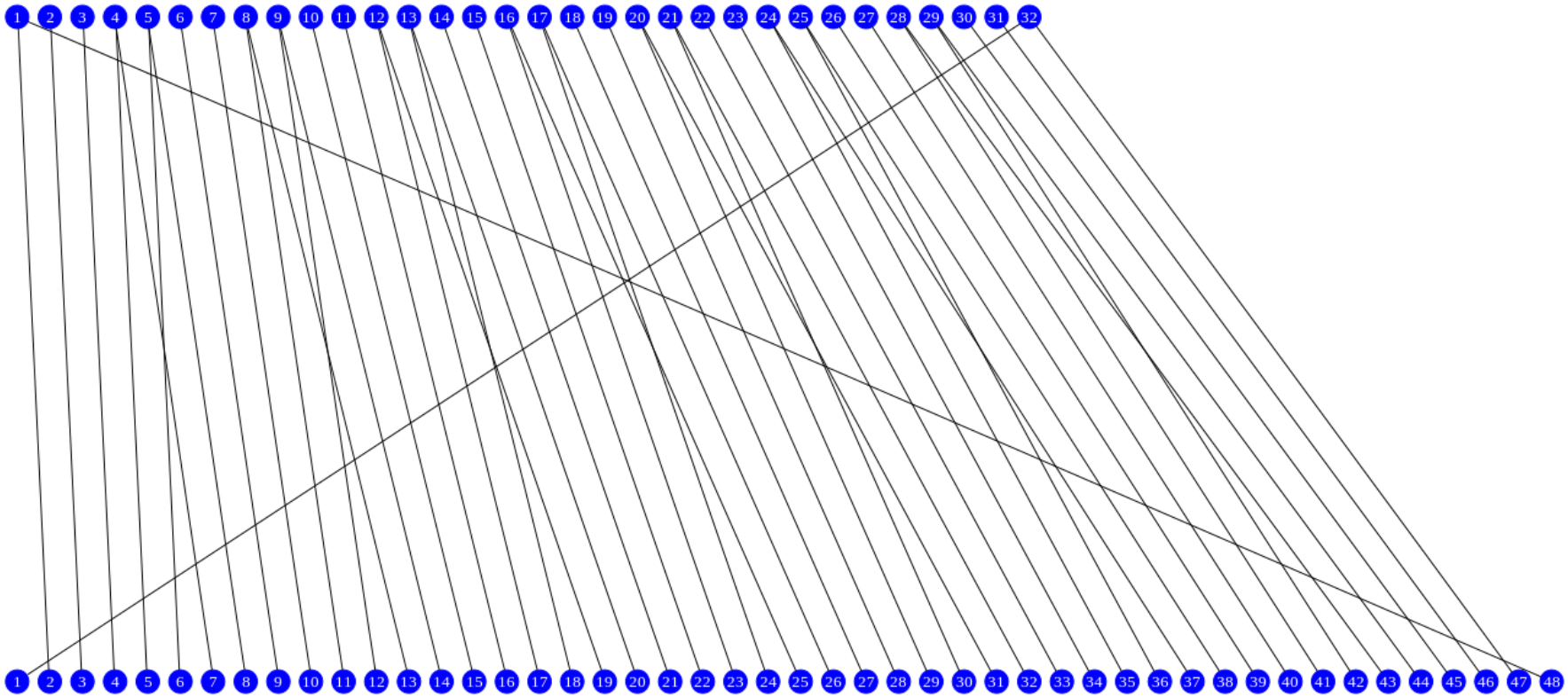


Schéma de la fonction d'expansion (Wikipedia)

Génération des approximations sur 1 round

Problème (OFF-MIPS)

Entrée : $(l_i)_{1\leq i\leq n}$ *et* $(r_j)_{1\leq j\leq n}$ *vecteurs de* $\mathbf{R}^d$.
Sortie :

$$\underset{\substack{1\leq i\leq n\\ 1\leq j\leq n}}{\operatorname{argmax}} \langle l_i, r_j \rangle$$

- Réduction à OFF-MIPS : précalcul de matrices 4×4 des biais sur 1 S-box.
- Résolution de OFF-MIPS : algorithme probabiliste.

Résultats

Algorithme B&B de Matsui :

- Plus court chemin dans un DAG, arêtes de poids positif, 2^{100} nœuds, 2^{132} arêtes.
- Matsui : élagage intelligent.
- En pratique : plus efficace que la génération sur 1 round.

Résultats :

- Simulation de 6h sur un bon serveur.
- Pas de changement de biais pour les approximations principales.
- Multithreading envisagé.
- $\sim$ 1400 masques de clé distincts.

1. Vue d'ensemble de l'attaque
   - Présentation de DES
   - Fonctionnement général de l'attaque
2. Génération des approximations linéaires
   - Généralités
   - Combinaison d'approximations
   - Génération des approximations sur 1 round
   - Résultats
3. Phase d'attaque
   - Cryptanalyse multilinéaire (modèle d'Hermelin)
   - Optimisation temporelle du déchiffrement partiel
   - Estimation de rang
   - Résultats de l'attaque

Cryptanalyse multilinéaire (modèle d'Hermelin)

$$Au^{k_0} + Bv^{k_0} = \Gamma k_0$$

Hypothèse (de la mauvaise clé)

Si $k \neq k_0$ et si u^k et v^k sont déchiffrés partiellement avec la clé k, alors $Au^k + Bv^k$ suit une distribution uniforme.

Test statistique :

- si $k = k_0$: distribution jointe (biaisée).
- si $k \neq k_0$: distribution uniforme.

Définition (distribution empirique)

$$P(k, \eta) = Card\{i \in \{1, \ldots, N\} | Ax_i^k + By_i^k = \eta\}$$

Test du χ^2 : $S(k) = \sum_{\eta=0}^{2^m-1} \left(\frac{P(k,\eta)}{N} - \frac{1}{2^m} \right)^2$

Des scores aux probabilités

Modèle bayésien :

- $D = \chi^2_{2^m-1}(NC)$: distribution limite quand $N \to \infty$ de $2^m NS(k)$ si $k = k_0$.
- $D' = \chi^2_{2^m-1}(0)$: idem si $k \neq k_0$.

$$P(k = k_0 | S = s) = \text{cte} \times \frac{f_D(S(k))}{f_{D'}(S(k))}$$

Optimisation temporelle du déchiffrement partiel

Rappel des notations :

- N : nombre de couples clair/chiffré.
- m : nombre d'approximations.
- t : cardinal du masque de clé.

Complexités temporelles :

- Complexité naïve : $O(Nm2^t)$.
- Optimisation de Matsui avec précalcul : $O(N + m2^{2t+m})$.
- (non implémenté entièrement) Optimisation de Nguyen avec transformée de Fourier/Walsh-Hadamard : $O(N + (t + m)2^{t+m})$.

Estimation de rang

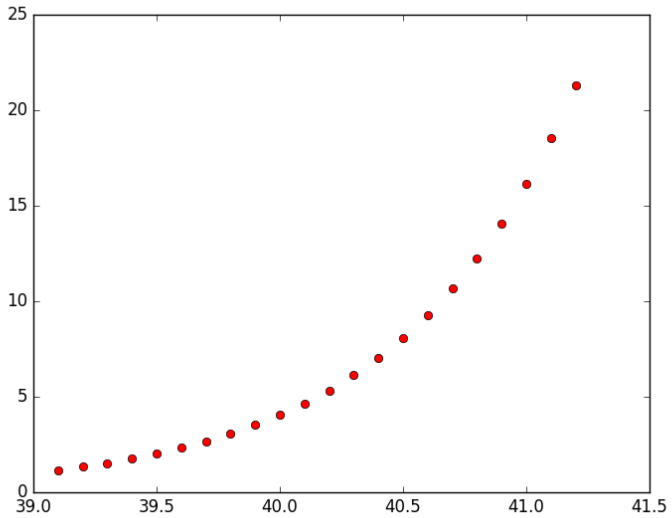
- Outils développés dans mon équipe d'accueil.
- Combine des listes de probabilité pour obtenir une estimation du rang de la clé si on les parcourt par probabilité décroissante.
- Contrainte : morceaux de clé disjoints.

Combinaison d'attaques multilinéaires

- Avantage fournie par chaque liste connu (Hermelin).
- Combinaison de deux listes de masques de clé disjoints : avantage sommé.
- Compression d'une liste en un masque plus petit : avantage réduit du nombre de bits compressés.

Problème

Maximiser l'avantage d'une liste pouvant être obtenue par combinaisons et compressions des listes de départ.



Résultats avec une file à priorité sur 14 rounds : meilleurs biais théoriques obtenus avec les approximations de Matsui.

Résultats de l'attaque sur 8 rounds

Biais 1	Biais 2
1.4648×2^{-11}	1.1250×2^{-12}
1.4648×2^{-12}	1.1250×2^{-13}
1.6875×2^{-15}	1.1250×2^{-13}

Figure – Biais des approximations choisies

N	2^{20}	2^{22}	2^{23}	2^{24}	2^{25}	$2^{25.5}$	2^{26}
R_{min}	$2^{54.205}$	$2^{54.652}$	$2^{40.340}$	2^{34}	$2^{41.175}$	2^{36}	2^{32}
R_{round}	$2^{54.211}$	$2^{54.658}$	$2^{40.384}$	$2^{34.322}$	$2^{41.224}$	$2^{36.170}$	2^{33}
R_{max}	$2^{54.211}$	$2^{54.658}$	$2^{40.384}$	$2^{34.322}$	$2^{41.224}$	$2^{36.170}$	2^{33}
a_{round}	1.789	1.342	15.616	21.618	14.776	19.830	23

Figure – Résultats des simulations

Conclusion

- Attaque complète du DES avec des outils nouveaux.
- Nécessité d'un outil pour faire de l'estimation de rang avec dépendance.
- Résultats difficiles à interpréter.
- Introduction à la cryptanalyse et à la recherche.

Bibliographie

Miia Hermelin.
Multidimensional linear cryptanalysis.
PhD thesis, Aalto University, Espoo, Helsinki, Finland, 2010.

Mitsuru Matsui.
Linear cryptanalysis method for des cipher.
In *Workshop on the Theory and Application of Cryptographic Techniques on Advances in Cryptology*, EUROCRYPT '93, pages 386–397, Secaucus, NJ, USA, 1994. Springer-Verlag New York, Inc.

Romain Poussier, François-Xavier Standaert, and Vincent Grosso.
Simple key enumeration (and rank estimation) using histograms : An integrated approach.
In *CHES*, pages 61–81. Springer, 2016.