Lightweight Construction of S-Boxes

Sébastien Duval

Inria Paris

8 nov. 2016

Table of Contents

Encryption

Send a secret message...

My secret Diary

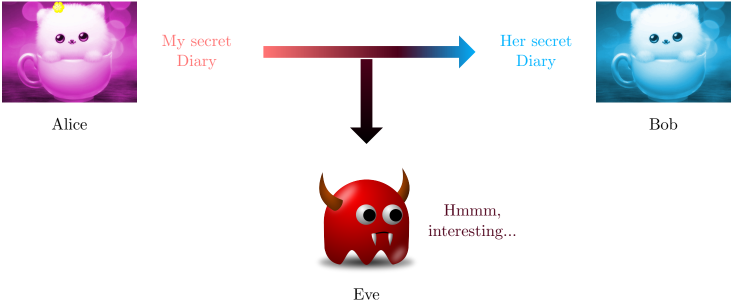
Her secret Diary

Alice

Bob

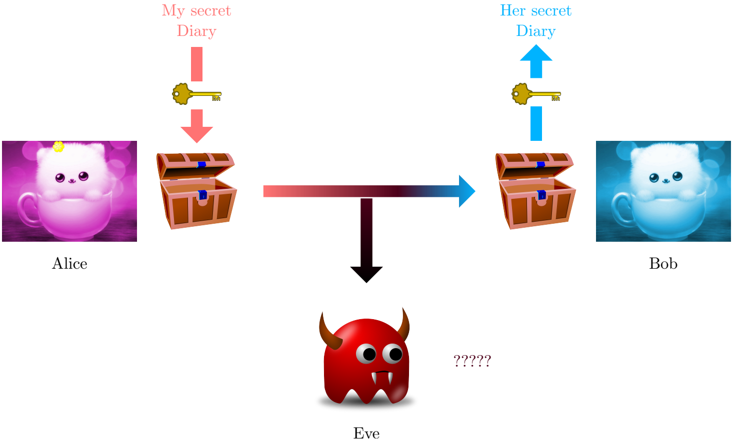
Encryption

But mind the enemy!



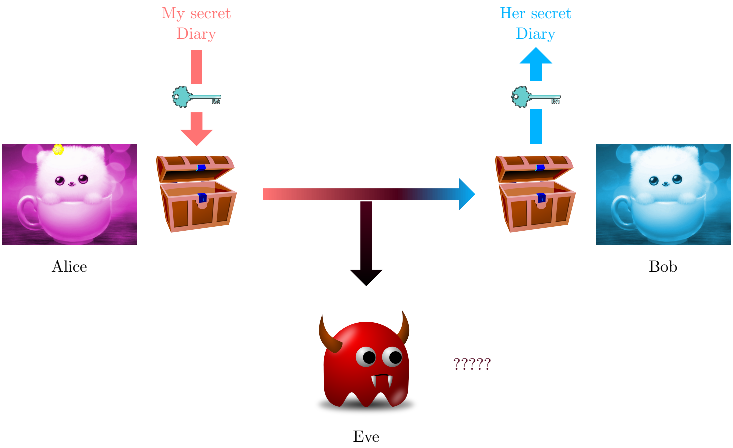
Encryption

Use encryption



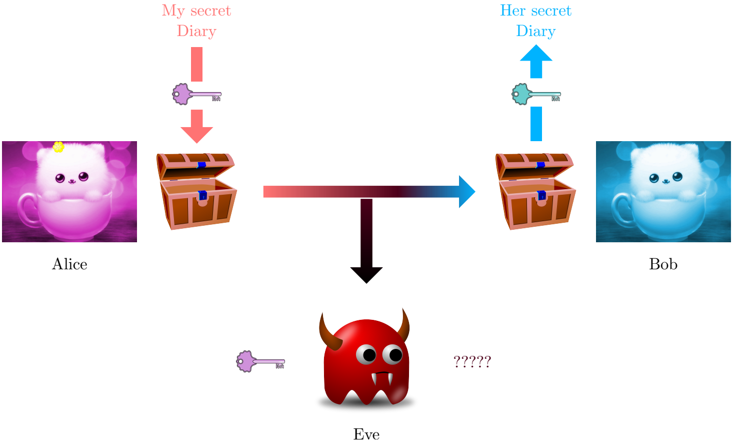
Encryption

Private key encryption



Encryption

Public key encryption

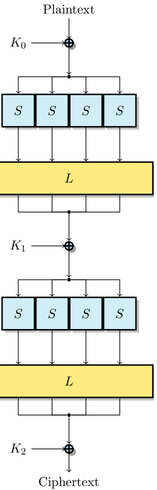


Symmetric Encryption: Security Criteria

Shannon's Criteria

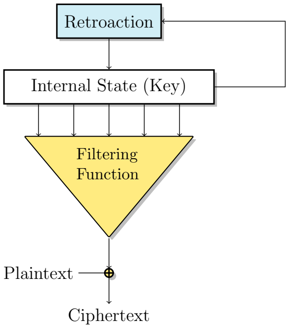
$$p \in \mathbb{F}_2^n \Rightarrow c \in \mathbb{F}_2^n$$

1. Diffusion
   - $\forall i, j$, p_i affects c_j.
   - Can be achieved using linear functions.
2. Confusion
   - Relation between p and c must be complex.
   - Requires non-linear functions.
   - Implemented as tables: S-Boxes.

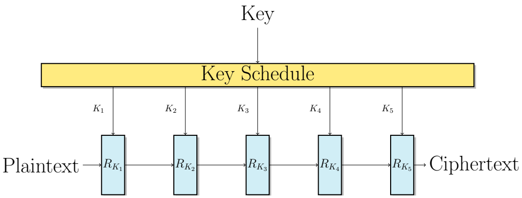


SPN Encryption

Stream & Block Ciphers

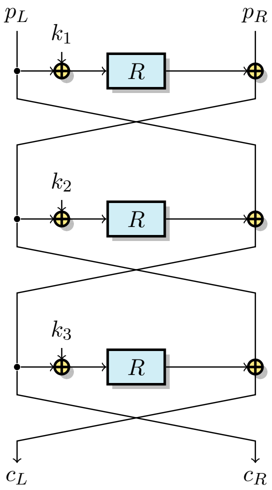


Stream Cipher



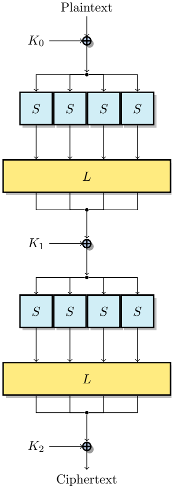
Block Cipher

Feistel Ciphers



- Lucifer/DES (H. Feistel, IBM, 1974)
- R built recursively
- Involution up to key ordering

SPN Ciphers



- Rijndael/AES (J. Daemen, V. Rijmen, 1988)
- Succession of confusion/diffusion layers
- Good for parallelism and easy to implement

Proven Security?

No NP-Complete Problem

- No reduction to an NP-complete problem
- No proven security
- Hypothesis: distinguishing from a random permutation is hard

Hard to Formalise

- Formal definition:
  - Chosen Plaintext Attack.
  - Cipher indistinguishable from a PRP $\Rightarrow$ secure against CPA.
  - i.e.: No Turing machine gives a different answer if given the cipher or a PRP.
- In practice:
  - How to define a "random" permutation ?
  - New property of random permutations $\Rightarrow$ new attack
  - We need cryptanalysis

Statistical Attacks

- Distinguish from random ⇒ attack
- Lots of properties:
  - Differential attacks
  - Linear attacks
  - Algebraic attacks
  - Subset attacks
  - ...
- Most efficient: differential and linear
- Very similar

Differential Attacks

Definition: Differential Uniformity

Let F be a function over $\mathbb{F}_2^n$. The table of differences of F is:

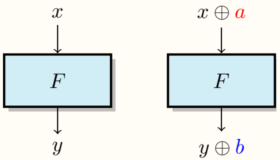
$$\delta_F(a \to b) = \#\{x \in \mathbb{F}_2^n | F(x \oplus a) = F(x) \oplus b\}.$$

Moreover, the differential uniformity of F is

$$\delta(F) = \max_{a \neq 0, b} \delta_F(a \to b).$$

We will also consider:

$$\delta_{min}(F) = \min_{a \neq 0} \max_{b} \delta_F(a \to b).$$



- F is resistant against differential attacks if $\delta(F)$ is small
- $\delta_F(a \to b)$ is even
- $\delta(F) = 2$ for APN functions

Table of Differences

a\b	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	2	0	0	0	0	2	2	0	4	0	4	2	0	0
2	0	0	2	4	2	0	0	0	2	4	0	0	0	0	2	0
3	0	2	2	0	4	0	0	0	2	0	0	2	0	0	4	0
4	0	2	0	0	0	0	0	2	0	4	0	2	4	2	0	0
5	0	4	0	0	2	0	0	2	0	0	0	4	0	2	2	0
6	0	0	0	4	0	4	0	0	0	4	4	0	0	0	0	0
7	0	0	2	0	0	4	0	2	2	4	0	0	0	2	0	0
8	0	2	2	2	0	2	0	0	2	0	0	0	2	0	2	2
9	0	0	2	2	2	0	2	0	0	0	0	0	0	2	2	4
10	0	4	2	0	0	0	4	2	0	0	2	0	0	0	0	2
11	0	0	0	0	0	2	2	0	0	0	2	2	2	4	2	0
12	0	0	2	2	2	2	0	0	2	0	0	2	2	0	0	2
13	0	0	0	2	2	0	2	2	2	0	0	0	0	0	2	4
14	0	0	0	0	0	0	4	0	2	0	2	4	0	2	0	2
15	0	2	0	0	2	2	2	4	0	0	2	0	2	0	0	0

All values are even:

$$S(x) \oplus S(x \oplus a) = b \Longleftrightarrow S((x \oplus a) \oplus a) \oplus S(x \oplus a) = b$$

Linear Attacks

Definition: Linearity

Let F be a function over $\mathbb{F}_2^n$. The table of linear biases of F is:

$$\lambda_F(a, b) = \sum_{x \in \mathbb{F}_2^n} (-1)^{a \cdot x \oplus b \cdot F(x)}.$$

Moreover, the linearity of F is

$$\boldsymbol{\mathcal{L}}(F) = \max_{a,b \neq 0} |\lambda_F(a, b)|.$$

- F is resistant to linear attacks if $\mathcal{L}(S)$ is small

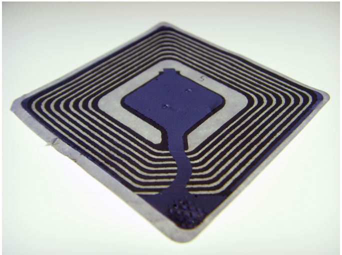
What Now?

We have good ciphers, considered secure and well studied with a powerful background theory: What now ?

- Still a lot of theory
- Cryptanalysis: Find new attacks
- ...
- Fit constrained specifications:
  - FHE
  - Side-channel attacks
  - Lightweight
  - ...

Lightweight Cryptography

- Secure and fast ciphers
- But too costly for dedicated environments...
- Useful for connected devices



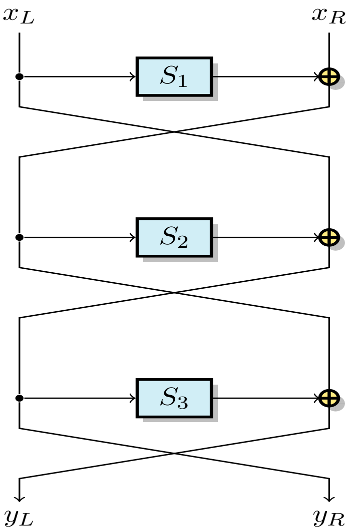
- Size of an RFID chip:$< 10\,000\,\mathrm{GE}$
- Smallest implementation of AES: $\sim 10\,000\,\mathrm{GE}$

Directions for Building S-Boxes

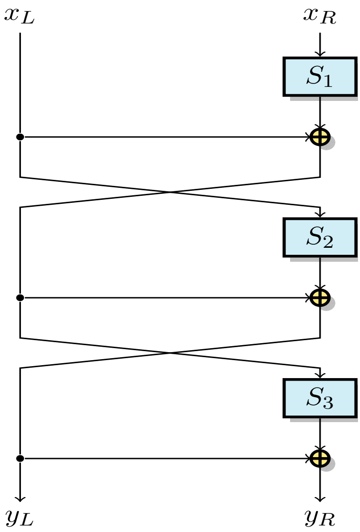
Problem: S-Box implementations are expansive

- Standard S-Box size: 8 bits (operations on bytes)
  - Implementation remains costly
- Smaller S-Boxes for a lesser cost:
  - Software implementation (table): Smaller table
  - Hardware implementation: Less logic gates
- But requires more rounds for same security
- Can we find a trade-off ?

Building Bigger S-Boxes From Small Ones



Feistel



MISTY

Objective of this Work

- Construction of S-Boxes using Feistel and MISTY networks
  - Construction of 8-bit S-Boxes from 4-bit ones
  - Trade-off between implementation cost and security

Results

- Determine the best properties reachable using MISTY and Feistel
  - Applied to 8-bit S-Boxes
- From theory to practice: Contruction of lightweight S-Boxes

Feistel and MISTY to Build Ciphers

- Initially used to define block ciphers (keyed networks)
- Well studied, many known results:

$$\mathrm{MEDP}(F_K) = \max_{a \neq 0, b} \frac{1}{2^k} \sum_{K \in \mathbb{F}_2^k} \frac{\delta_{F_K}(a \to b)}{2^n}$$

$$\mathrm{MELP}(F_K) = \max_{a, b \neq 0} \frac{1}{2^k} \sum_{K \in \mathbb{F}_2^k} \left(\frac{\lambda_{F_K}(a, b)}{2^n} \right)^2$$

- For MISTY and Feistel:

$$\mathrm{MEDP}(S_i) \leq p \Rightarrow \mathrm{MEDP}(F) \leq p^2$$

$$\mathrm{MELP}(S_i) \leq q \Rightarrow \mathrm{MELP}(F) \leq q^2$$

- Caution! Doesn't work when the key is fixed !

Feistel and MISTY with Fixed Key: Limits of MEDP

Example

- 3-round MISTY network.
- $S_1 = S_2 = S_3 = [A, 7, 9, 6, 0, 1, 5, B, 3, E, 8, 2, C, D, 4, F]$.
- $\delta(S_i) = 4$, $\text{MEDP}(S_i) = 2^{-2}$.
- $\text{MEDP}(F) \leq 2^{-4}$.
- For every key, there exists a differential with probability 2^{-3}.

- A bound on MEDP means:
  1. Choose an input and an output difference.
  2. For any chosen key, differential probability is low.
- No bound when the key is chosen before the differences!
- When building S-Boxes, there is no key, i.e. $K = 0$.

Feistel: Prior Results

Theorem (Li et Wang, CHES 2014)

Let F be a 3-round Feistel network with internal functions S_1, S_2 et S_3, then

- $\delta(F) \geq 2\delta(S_2)$
- $\delta(F) \geq 2^{n+1}$ if S_2 is not a permutation
- Pour $n = 4$, $\delta(F) \geq 8$, and if $\delta(F) = 8$, then $\mathcal{L}(F) \geq 64$
- $\delta(F) = 8$ and $\mathcal{L}(F) = 64$ is reachable

Feistel: New Results

Theorem

- $\delta(F) \geq \delta(S_2) \max(\delta_{\min}(S_1), \delta_{\min}(S_3))$
- $\delta(F) \geq 2^{n+1}$ if S_2 is not a permutation
- $\delta(F) \geq \max_{i \neq 2, j \neq i,2}(\delta(S_i)\delta_{\min}(S_j), \delta(S_i)\delta_{\min}(S_2^{-1}))$ if S_2 is a permutation with $\delta_{\min}(S) = \min_{a \neq 0} \max_b \delta_S(a \to b)$
- This bounds depend on all 3 S-Boxes

Pour $n = 4$

- $\delta(F) \geq 8$, tight
- $\mathcal{L}(F) \geq 48$, $\mathcal{L}(F) \geq 64$ if $\delta(F) < 32$

MISTY: New Results

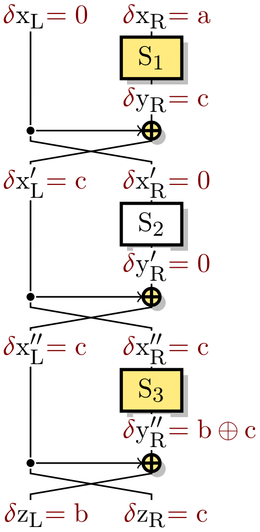
Theorem

- $\delta(F) \geq \delta(S_1) \max(\delta_{\min}(S_2), \delta_{\min}(S_3))$
- $\delta(F) \geq 2^{n+1}$ if S_1 is not a permutation
- $\delta(F) \geq \max_{i \neq 1, j \neq i, 1}(\delta(S_i)\delta_{\min}(S_j), \delta(S_i)\delta_{\min}(S_1^{-1}))$
  if S_1 is a permutation
  with $\delta_{\min}(S) = \min_{a \neq 0} \max_b \delta_S(a \to b)$
- There was no prior result for MISTY with fixed key

Pour $n = 4$

- $\delta(F) \geq 8$, tight
- $\mathcal{L}(F) \geq 48$, $\mathcal{L}(F) \geq 64$ if $\delta(F) < 32$

Sketch of Proof



Proposition

$$\delta_F(0 \,\|\, a \to b \,\|\, c) = \delta_{S_1}(a \to c) \times \delta_{S_3}(c \to b \oplus c)$$

Proof

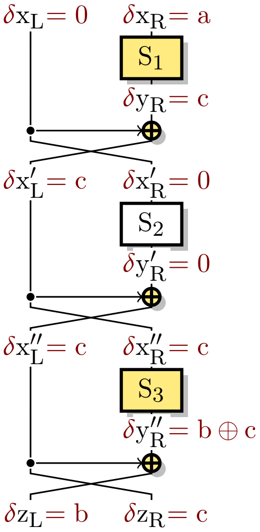
$$F(x_L \| x_R) \oplus F(x_L \| (x_R \oplus a)) = b \| c$$

$$\Leftrightarrow \begin{cases} S_3(S_1(x_R) \oplus x_L) \oplus S_3(S_1(x_R \oplus a) \oplus x_L) = b \oplus c, \\ S_2(x_L) \oplus S_1(x_R) \oplus x_L \oplus S_2(x_L) \oplus S_1(x_R \oplus a) \oplus x_L = c \end{cases}$$

$$\Leftrightarrow \begin{cases} S_3(S_1(x_R) \oplus x_L) \oplus S_3(S_1(x_R \oplus a) \oplus x_L) = b \oplus c, \\ S_1(x_R) \oplus S_1(x_R \oplus a) = c \end{cases}$$

$$\Leftrightarrow \begin{cases} x_R \in D_{S_1}(a \to c) \\ x_L \in S_1(x_R) \oplus D_{S_3}(c \to b \oplus c) \end{cases}$$

Sketch of Proof



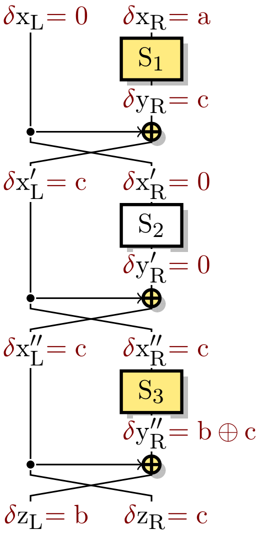
Proposition

$\delta_F(0 \| a \to b \| c) = \delta_{S_1}(a \to c) \times \delta_{S_3}(c \to b \oplus c)$

Application: if S_1 is not bijective

- Fix $b = c = 0$, $\delta_{S_3}(0 \to 0) = 2^n$
- Choose a such that $\delta_{S_1}(a \to 0) \geq 2$
- $\delta(F) \geq \delta_F(0 \| a \to 0 \| 0) \geq 2^{n+1}$

Sketch of Proof



Proposition

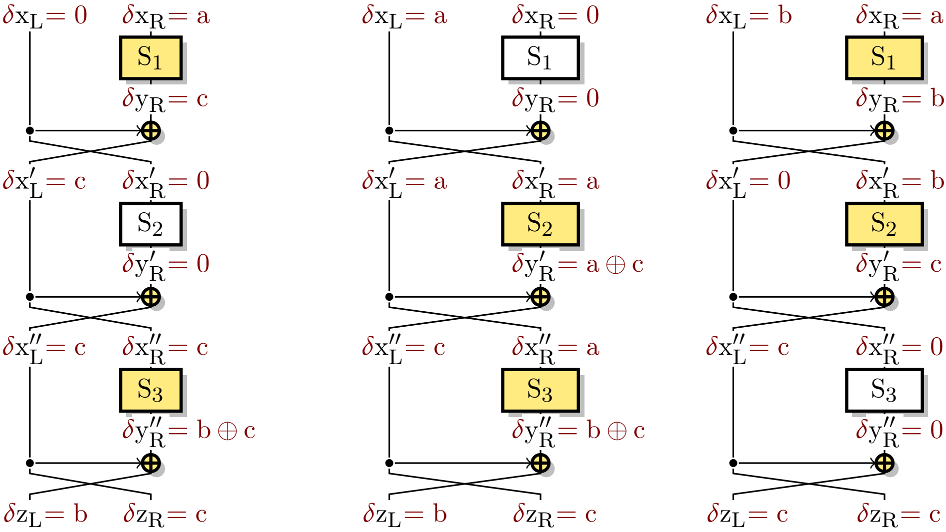
$$\delta_F(0 \,\|\, a \to b \,\|\, c) = \delta_{S_1}(a \to c) \times \delta_{S_3}(c \to b \oplus c)$$

Application: if S_1 is bijective

- Choose a, c such that $\delta_{S_1}(a, c) = \delta(S_1)$
- Choose b with $\delta_{S_3}(c, b \oplus c) \geq \delta_{min}(S_3)$
- $\delta(F) \geq \delta_F(0 \,\|\, a, b \,\|\, c) \geq \delta(S_1) \times \delta_{min}(S_3)$

- Choose b, c such that $\delta_{S_3}(c, b \oplus c) = \delta(S_3)$
- Choose a with $\delta_{S_1}(a, c) \geq \delta_{min}(S_1^{-1})$
- $\delta(F) \geq \delta_F(0 \,\|\, a, b \,\|\, c) \geq \delta(S_3) \times \delta_{min}(S_1^{-1})$

Sketch of Proof



Application to n = 4: Properties of 4-bit Functions

Properties of 4-bit S-Boxes

- Full classification of 4-bit permutations
  - 302 affine equivalence classes

    [De Cannière; Leander & Poschmann '07]
- Full classification of 4-bit APN functions
  - 2 extended affine equivalence classes [Brinkmann & Leander '08]
- There are 4-bit APN functions
  - $\delta(S_i) = 2$, $\delta_{min}(S_i) = 2$
- There are no 4-bit APN permutations
  - If S_i is a permutation, $\delta(S_i) \geq 4$, $\delta_{min}(S_i) \geq 2$

Refined bounds for n = 4 (MISTY and Feistel)

- If S_i are all non-bijective, then $\delta(F) \geq 32$
- If S_i bijective, $\delta(F) \geq \delta_{min}(S_j) \times \delta(S_i) \geq 8$

MISTY: Necessary Conditions for $\delta = 8$, $\mathcal{L} = 64$

Necessary Conditions for $\delta(F) = 8$

- S_1 permutation with $\delta(S_1) = 4$
- S_2, S_3 APN

Proof.

- Suppose $\delta(S_3) \geq 4$
  - $\delta(S_3) \geq 4$, therefore there exist c_1, b_1 with $\delta_{S_3}(c_1 \to b_1) \geq 4$
  - There are two pairs $(x, x \oplus c_1)$, $(y, y \oplus c_1)$ in $D_{S_3}(c_1 \to b_1)$
  - With $c_2 = x \oplus y$, $b_2 = S_3(x) \oplus S_3(y)$, there are two pairs (x, y), $(x \oplus c_1, y \oplus c_1)$ with $D_{S_3}(c_2 \to b_2)$
  - Similarly, there are two pairs $(x, y \oplus c_1)$, $(x \oplus c_1, y)$ with $D_{S_3}(c_1 \oplus c_2 \to b_1 \oplus b_2)$
  - At least 3 lines c_i with S_3 with a value ≥ 4

□

MISTY: Necessary Conditions for $\delta = 8$, $\mathcal{L} = 64$

Necessary Conditions for $\delta(F) = 8$

- S_1 permutation with $\delta(S_1) = 4$
- S_2, S_3 APN

Proof.

- Suppose $\delta(S_3) \geq 4$
  - At least 3 lines c_i with S_3 with a value ≥ 4
  - $\delta_F(0||a \to b||c) = \delta_{S_1}(a \to c) \times \delta_{S_3}(c \to b \oplus c)$
  - To get $c \leftarrow c_i$, we also need: c_i column of differences of S_1 with value $= 4$
  - If such a c_i does not exist $\Rightarrow L = \{c_1, c_2, c_3 = c_1 \oplus c_2\} \subseteq C$, C = columns of S_1 without value $= 4$
  - C for the representatives of affine equivalence classes does not contain any subset stable under XOR

□

Constructing Strong 8-bit S-Boxes with Feistel and MISTY

Feistel

- $\delta(F) \geq 8$, reached bound
  - S_1, S_3 must be APN, S_2 a permutation with $\delta(S_2) = 4$
- $\mathcal{L}(F) \geq 48$
  - $\mathcal{L}(F) \geq 64$ if $\delta(F) < 32$

MISTY

- $\delta(F) \geq 8$, reached bound
  - S_2, S_3 must be APN, S_1 a permutation with $\delta(S_1) = 4$
  - F is not a permutation
- $\mathcal{L}(F) \geq 48$
  - $\mathcal{L}(F) \geq 64$ if $\delta(F) < 32$
- F permutation: $\delta(F) \geq 16$, reached bound

Getting the Components

- From these results, Feistel is more adapted
- We need S_1, S_3 APN, S_2 permutation with $\delta(S_2) = 4$
  - Can we choose S_i with low implementation cost?
- Exhaustive search over small implementations until good properties are reached (Üllrich & al. 2011)
  - Search sequences of instructions for a bit-sliced implementation
  - We use equivalence classes to cut branches
  - Minimise the number of non-linear operations

Exhaustive Search Results

Permutation with $\delta = 4$

- Easy search
  Reuse results from Üllrich & al.
- 9 instructions
  - 4 non-linear
  - 4 XOR
  - 1 copy

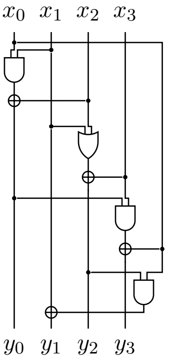
APN Function

- Costly search
  - No filtering permutations
  - 6k core-hours
- 10 instructions
  - But 6 non-linear
- 11 instructions
  - 4 non-linear
  - 5 XOR
  - 2 copies

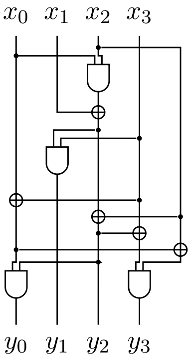
- 4 non-linear gates is optimal

- 4 non-linear gates is optimal

Concrete Example



Permutation with $\delta = 4$ (S_2)



APN Function (S_1, S_3)

A Feistel network using these functions is an 8-bit permutation with $\delta = 8$ and $\mathcal{L} = 64$.

Results: Better than Before

		Implem.		Properties		
S-Box	Construction	$\wedge, \vee$	$\oplus$	$\mathcal{L}$	δ	Cost
AES	Inversion	32	83	32	4	1
Whirlpool	Lai-Massey	36	58	56	8	1.35
CRYPTON	3-r. Feistel	49	12	64	8	1.83
Robin	3-r. Feistel	12	24	64	16	0.56
Fantomas	3-r. MISTY (3/5 bits)	11	25	64	16	0.51
LS (unnamed)	Whirlpool-like	16	41	64	10	0.64
New	3-r. Feistel	12	26	64	8	0.45

Conclusion

1. Bounds on the security of Feistel and MISTY networks with fixed key
2. Applied to 8-bit S-Boxes
   - Necessary conditions
   - Detailed bounds for permutations
   - Feistel is better for invertible 8-bit S-Boxes
3. Concrete construction of strong light S-Boxes
   - 8-bit S-Box from 3-round Feistel
   - Better than previously used S-Boxes

Questions ?