Aalto University
School of Science

Multiple differential cryptanalysis using $\mathtt{LLR}$ and χ^2 Statistics

Céline Blondeau

joint work with Benoît Gérard and Kaisa Nyberg

October 8, 2012

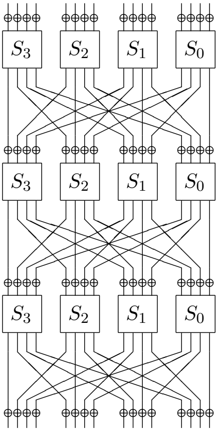
Outline

Outline

Block ciphers

$$E_K : \mathbb{F}_2^m \to \mathbb{F}_2^m$$

- K: Master key
- F: Round function
- K_i: Round key



SMALLPRESENT-[4]

Statistical Attacks

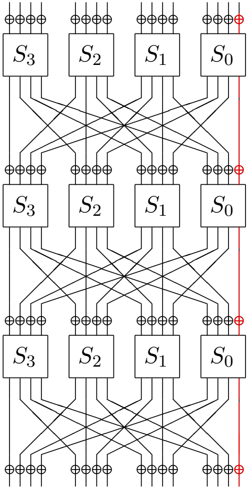
Statistical attacks:

- Take advantage of a non-uniform behavior of the cipher
- Two families: Linear and Differential cryptanalysis

Improvement of differential cryptanalysis

- Differential cryptanalysis [Biham Shamir 91]
- Truncated differential cryptanalysis [Knudsen 95]
- Impossible differential cryptanalysis [Biham Biryukov Shamir 99]
- Higher order differential cryptanalysis [Lai 94] [Knudsen 95]
- Multiple differential cryptanalysis (First approach) [BG 11]

Linear cryptanalysis



[Tardy-Gilbert91], [Matsui93]

Linear relation using

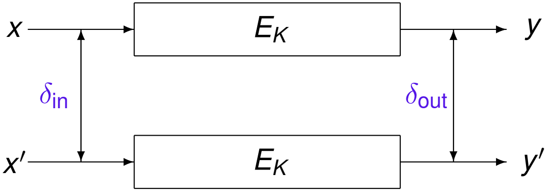
- plaintext bits,
- key bits,
- ciphertext bits.

$$\pi \cdot x \oplus \kappa \cdot K \oplus \gamma \cdot y = 0$$

with probability $p = \frac{1}{2} + \varepsilon$

Differential Cryptanalysis

Given an input difference between two plaintexts, some output differences occur more often than others.

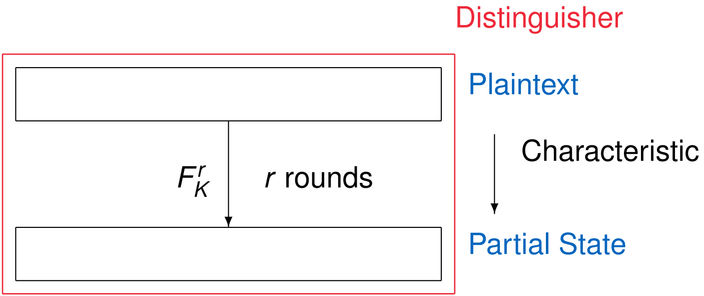


Differential: pair of input and output difference $(\delta_{\text{in}}, \delta_{\text{out}})$

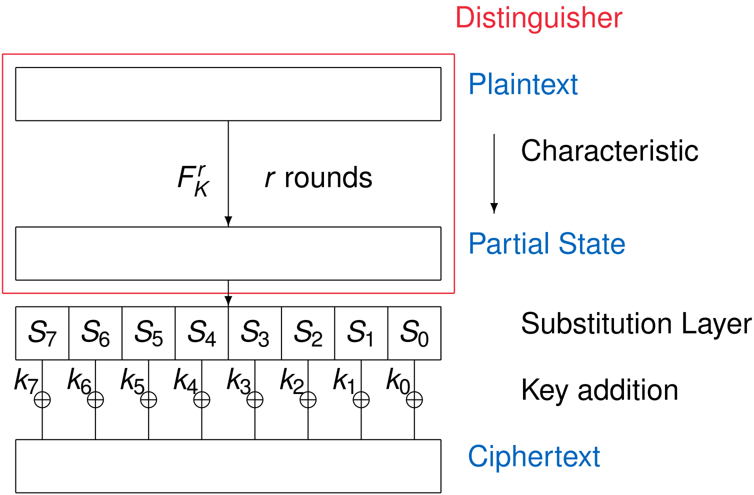
Differential probability: $p = P_{X,K}[\, E_K(X) \oplus E_K(X \oplus \delta_{\text{in}}) = \delta_{\text{out}} \,]$

Uniform probability: $\theta = 2^{-m}$

Last Round Attack



Last Round Attack



Related Work

Linear Cryptanalysis:

- Multiple linear cryptanalysis [Baignères, Junod, Vaudenay 04]
- Multidimensional linear cryptanalysis [Hermelin, Cho, Nyberg 08]

Both use $\mathtt{LLR}$ and/or χ^2 statistical tests.

Differential Cryptanalysis:

- [Blondeau, Gérard 11]:
  The frequencies are sum up
- Here:
  We study the $\mathtt{LLR}$ and/or χ^2 statistical tests.

Multiple differential cryptanalysis (First Approach)

- Set of differences $\delta_{\text{in}}{}^{(v)}, \delta_{\text{out}}{}^{(v)}$
- With probabilities $p_v = P_{X,K}[\ E_K(X) \oplus E_K(X \oplus \delta_{\text{in}}^{(v)}) = \delta_{\text{out}}^{(v)}\ ]$.
- Set of input differences $\delta_{\text{in}}{}^{(v)} \in \Delta_{in}$.
- $p = \frac{1}{\Delta_{in}} \sum_v p_v$ expected probability.
- $\theta = \frac{1}{\Delta_{in}} \sum_v \frac{1}{2^m}$ uniform probability.

Outline

Multiple Differential Cryptanalysis

- Fix input difference δ_{in} (To simplify the analysis)
- Vector of "difference": $V = [\delta_{\text{out}}^{(i)}]$ after r rounds,
- $p = [p_v]_{v \in V}$ vector of expected probabilities.
- $\theta = [\theta_v]_{v \in V}$ vector of uniform probabilities.

Discussion

Parallel Work for small ciphers: [Albrecht Leander 2012]

Whole distribution taken for SMALLPRESENT-[4] (16-bit cipher)
Whole distribution taken for KATAN-32 (32-bit cipher)

Limits:

For actual ciphers the output size is too large (2^{64} or 2^{128})

Application to real cipher:

Introduction of partitioning functions.

Partitioning function

We analyze two “orthogonal” cases

- Unbalanced partitioning
  - Take a subset of simple differences
- Balanced partitioning
  - Group the differences in order to be able to use information of the whole output space.

Unbalanced Partitioning

Idea: Subset of simple differences

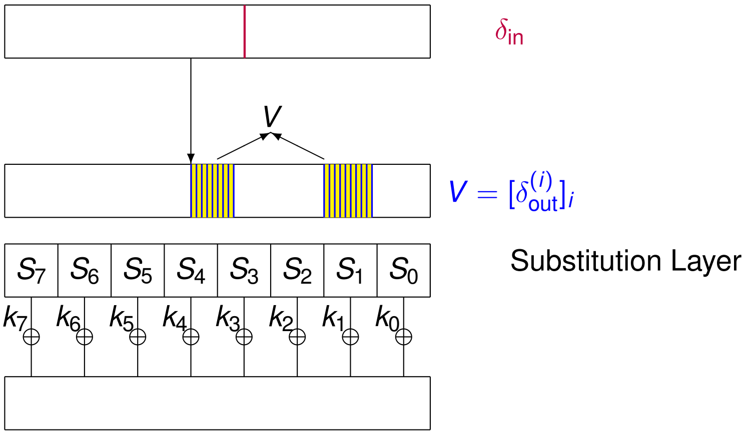
- Output differences $(\delta_{out}^{(i)})_{1 \leq i \leq A}$,
- Counter for each of these differentials q_i^k.

- As $\sum_{i=1}^{A} q_i^k \neq 1$
- We have a "trash" counter q_0^k which gather all other output differences.

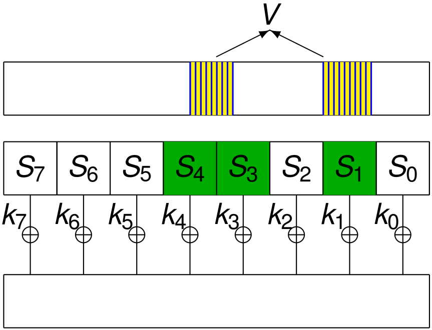
We increment the counter q_i^k

if the difference $\delta_{out}^{(i)}$ is obtained after partial deciphering.

Unbalanced Partitioning: Last Round Attack



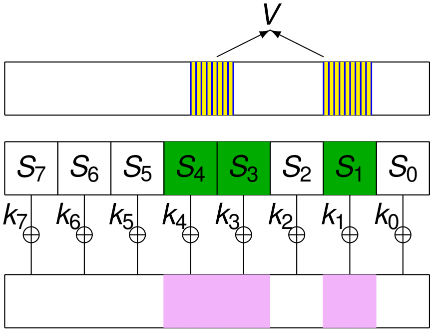
Unbalanced Partitioning: Last Round Attack



$V = [\delta_{\text{out}}^{(i)}]_i$

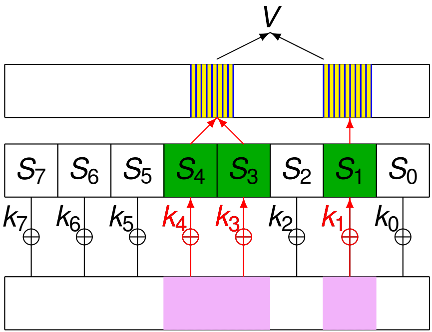
Active Sboxes

Unbalanced Partitioning: Last Round Attack



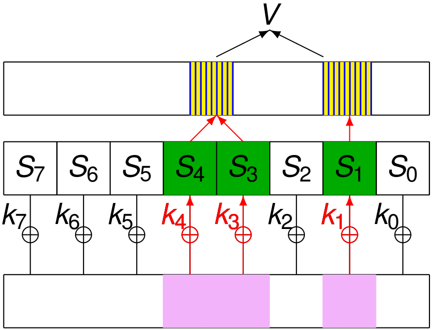
Sieving process
Discard some ciphertext pairs

Unbalanced Partitioning: Last Round Attack



For all key candidates, partially decipher

Unbalanced Partitioning: Last Round Attack



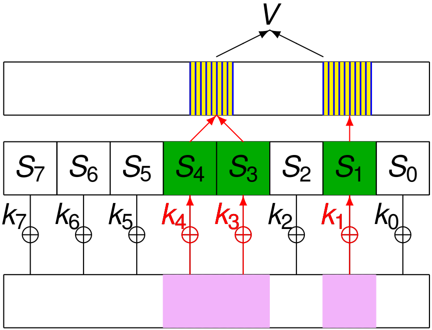
If $\delta = \delta_{\text{out}}^{(i)}$

Increment q_i^k

Otherwise

Increment q_0^k

Unbalanced Partitioning: Last Round Attack



If $\delta = \delta_{\text{out}}^{(i)}$

Increment q_i^k

Otherwise

Increment q_0^k

Analyse the vectors q^k for each key

Scoring function

Unbalanced Partionning: Remarks

Corresponding known/former attacks:

- Differential cryptanalysis.

Advantage:

- A sieving process $\Rightarrow$ “smaller” time complexity

Disadvantage:

- Subset of output space $\Rightarrow$ not all information
- Small Probabilities $\Rightarrow$ Non-tightness of the information

Balanced Partitioning

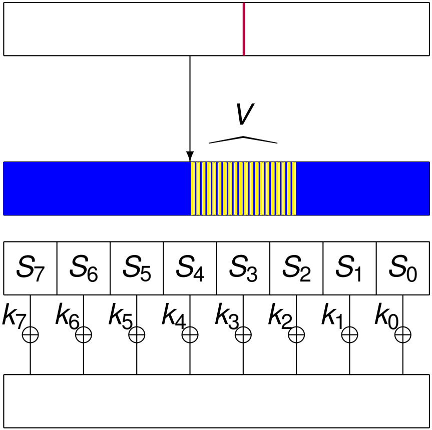
Idea: Using information from all output differences by grouping them.

Let $V = [\delta_{\text{out}}^{(i)}]_i$ a subspace of $\mathbb{F}_2^m$

A group of differences $\Delta_{\text{out}}^{(i)} = \delta_{\text{out}}^{(i)} \oplus \bar{V} \qquad (\bar{V} \oplus V = \mathbb{F}_2^m)$

A counter q_i^k for each group of differences.
We increment the counter q_i^k
if the difference $\delta \in \Delta_{out}^{(i)}$ is obtained partial deciphering.

Balanced Partitioning: Last Round Attack

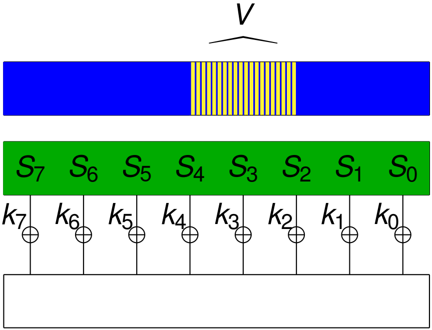


δ_{in}

$\Delta_{\text{out}} = \delta_{\text{out}} \oplus \bar{V}$

Substitution Layer

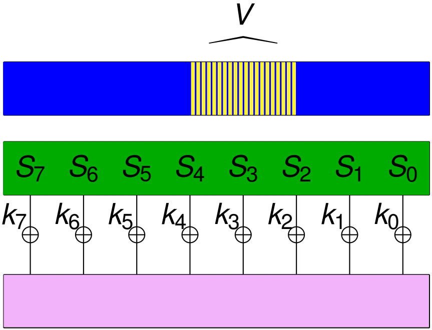
Balanced Partitioning: Last Round Attack



$$\Delta_{\text{out}} = \delta_{\text{out}} \oplus \bar{V}$$

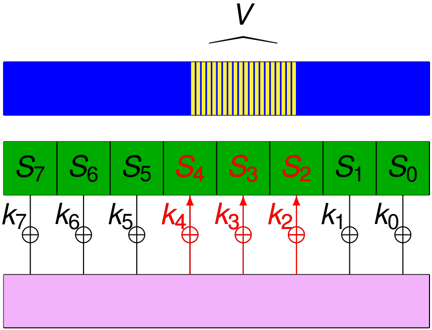
Active Sboxes

Balanced Partitioning: Last Round Attack



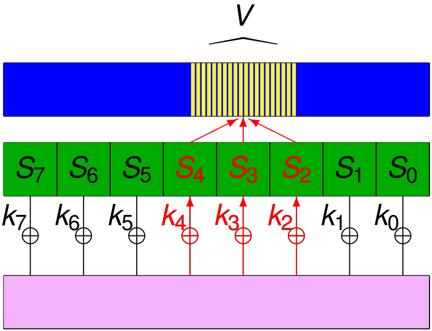
No Sieving process
Partially decipher for all pairs

Balanced Partitioning: Last Round Attack



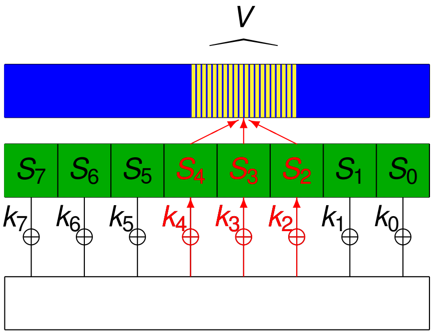
For all key candidates, partially decipher

Balanced Partitioning: Last Round Attack



If $\delta \in \delta_{\text{out}}^{(i)} \oplus \bar{V}$
Increment q_i^k

Balanced Partitioning: Last Round Attack



If $\delta \in \delta_{\text{out}}^{(i)} \oplus \bar{V}$
Increment q_i^k

Analyse the vectors q^k for each key
Scoring function

Balanced Partitioning: Remarks

Corresponding known/former attacks:

- Truncated Differential cryptanalysis.

Advantage:

- Whole output space $\Rightarrow$ More information
- Bigger Probabilities $\Rightarrow$ Tightness of the information

Disadvantage:

- No sieving process $\Rightarrow$ More time complexity

Statistical Tests

Probability distribution vectors

- Expected: $p = [p_v]_{v \in V}$
- Uniform: θ
- Observed: q^k (for a given key candidate)

$\mathtt{LLR}$ test: requires the knowledge of the theoretical probability p.

$$S_k = \mathtt{LLR}_k(q^k, p, \theta) \stackrel{\text{def}}{=} N_s \sum_{v \in V} q_v^k \log\left(\frac{p_v}{\theta_v}\right).$$

χ^2 test: Does not require the knowledge of p for the attack

$$S_k = \chi_k^2(q^k, \theta) = N_s \sum_{v \in V} \frac{(q_v^k - \theta_v)^2}{\theta_v}.$$

Complexities

Let $S(k)$ be the statistic obtained for a key candidate k.

$$S(k) = \mathtt{LLR}_k(q^k, p, \theta) \text{ or } = \chi^2_k(q^k, \theta)$$

Then,

$$S(k) \sim \begin{cases} \mathcal{N}(\mu_R, \sigma_R^2) & \text{if } k = K_r, \\ \mathcal{N}(\mu_W, \sigma_W^2) & \text{otherwise.} \end{cases}$$

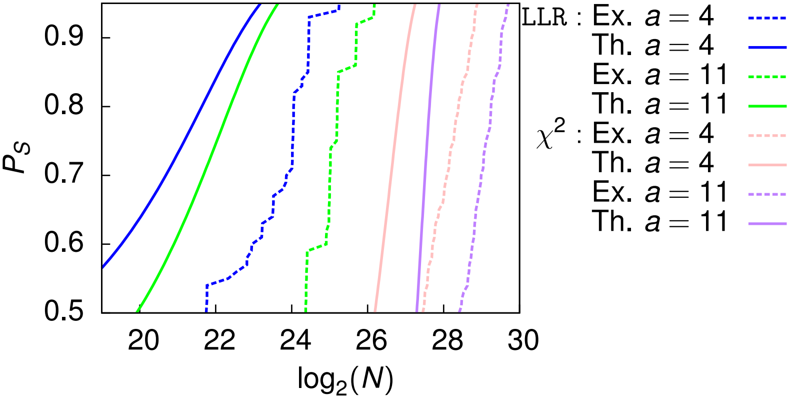
In the paper:

- Estimates of the value of $\mu_R, \mu_W, \sigma_R, \sigma_w$ for both $\mathtt{LLR}$ and χ^2 statistical tests.
- Estimates of the Data Complexity

Outline

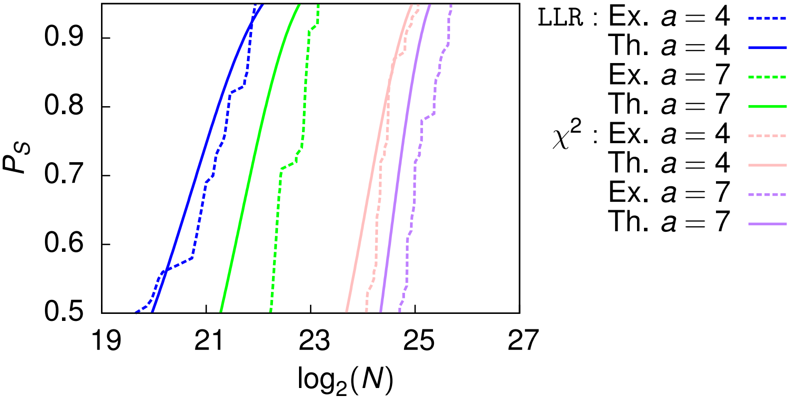
Using unbalanced partitioning

Subset of output differences



Using balanced partitioning

Set of groups of output differences



Conclusions

Balanced or Unbalanced partitioning ?

- Time Complexity: unbalanced $\Rightarrow$ faster attack.
- Data Complexity: depends on the cipher.

$\mathtt{LLR}$ or χ^2?

- If we have a good estimate of the expected probabilities $\Rightarrow$ $\mathtt{LLR}$ provides better Data and Memory complexities
- Otherwise $\mathtt{LLR}$ is not effective

Work in Progress

Estimation of the Differential Probabilities

In Theory

- Estimation of truncated differential probabilities can be done correlations.

In Practice

- Estimation of the correlations are “easy” on PRESENT CHO
- We use them to compute the distribution vector.
- We provide a multiple differential attack on PRESENT