Aalto University
School of Science

Advanced Attacks on Block Ciphers

Kaisa Nyberg

Email: kaisa dot nyberg at aalto dot fi
Department of Computer Science
Aalto University School of Science

Teheran February 2020

Outline

- Part I: Truncated differential attack on PRESENT
- Part II: Affine multidimensional linear distinguishers for Simon
- Part III: Links between differential and linear type properties (for general vectorial Boolean functions)

Part I: Attacks on PRESENT

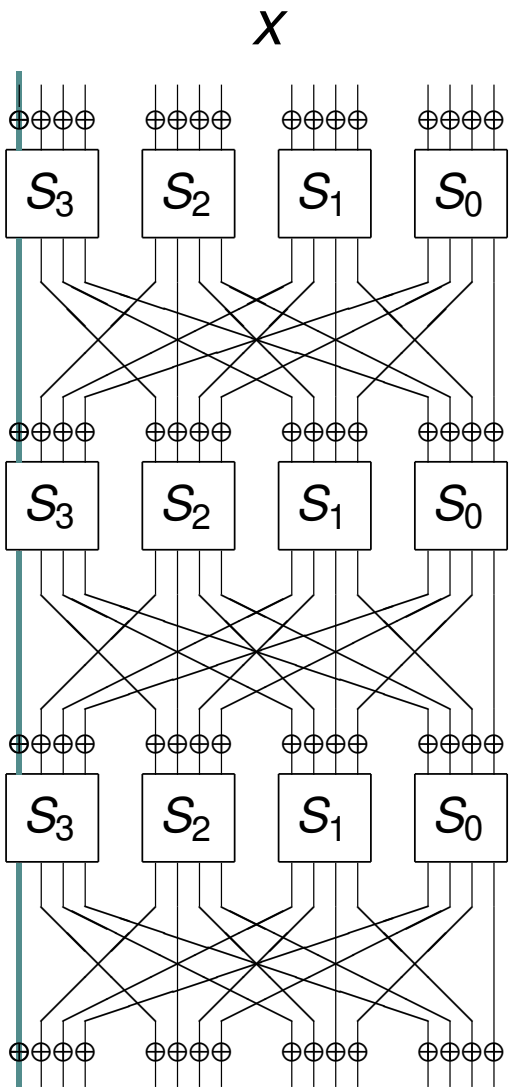
Block cipher

- Arguably the most important cryptographic primitive
- Symmetric key (not public key)
- Used as building block for other type of cryptographic primitives and protocols: stream ciphers, message authentication codes, authentication protocols.
- Several standards: standalone block ciphers for general purpose, special purpose (dedicated) block ciphers intended for particular application.

Block cipher – Is it secure?

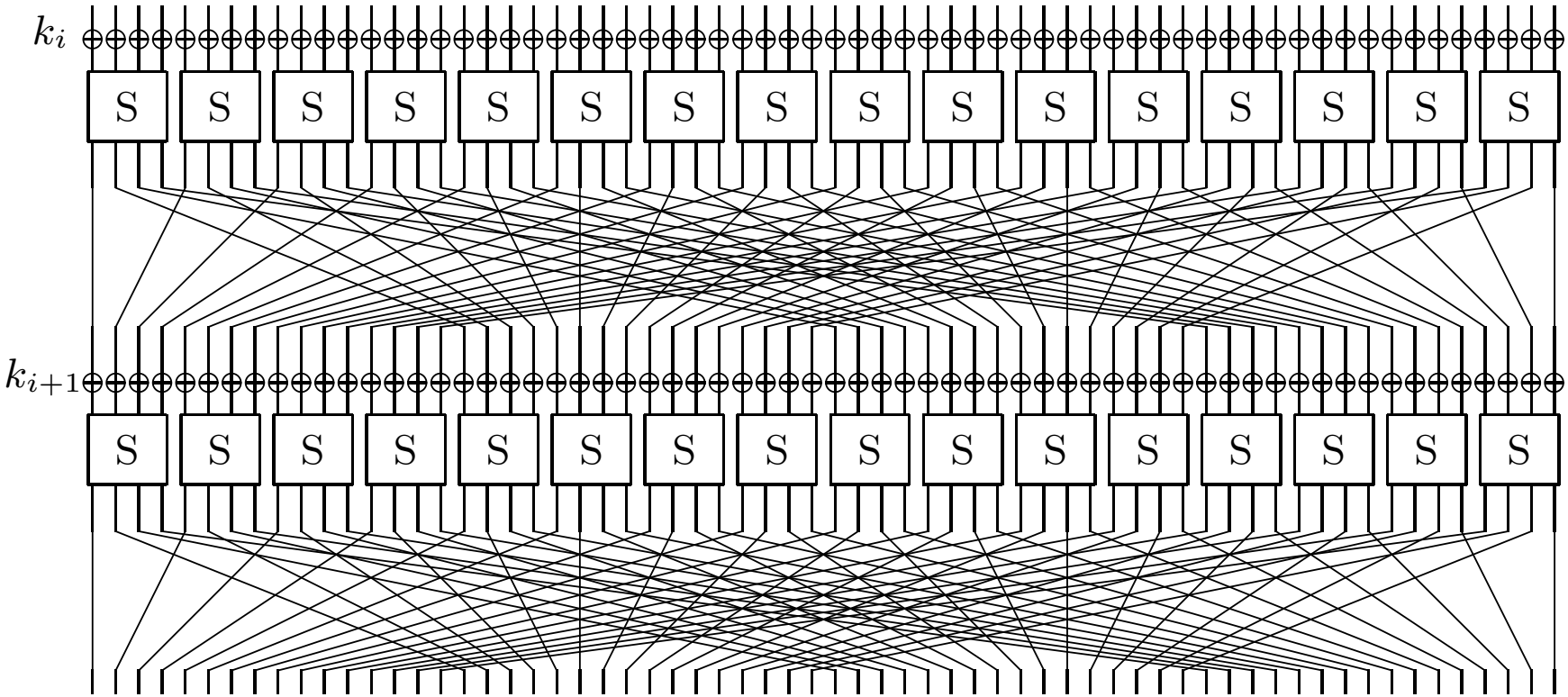
- Smart and well-founded design by experienced cryptographers
- Cannot be proven secure as such
- Security tested against known methods
- Need strong, generic, adequate cryptanalysis methods
- Design of *lightweight* block ciphers needs accurate cryptanalysis
- Lightweight ciphers: tight trade-off between security and efficiency
- Reducing security margins – not security!

Example SPN



$$y = E_k(x)$$

Two rounds of PRESENT



PRESENT in Wikipedia – Design

- Lightweight block cipher
- Orange Labs (France), Ruhr University Bochum (Germany) and the Technical University of Denmark in 2007
- Notable for its compact size (about 2.5 times smaller than AES).
- The block size is 64 bits and the key size can be 80 bit or 128 bit
- The non-linear layer is based on a single 4-bit S-box which was designed with hardware optimizations in mind.
- Intended to be used in situations where low-power consumption and high chip efficiency is desired
- Standardized by ISO/IEC

PRESENT in Wikipedia - Cryptanalysis

- A truncated differential key-recovery attack on 26 out of 31 rounds of PRESENT was suggested in 2014 by Blondeau and Nyberg (Eurocrypt 2014)
- Several full-round attacks using biclique cryptanalysis have been introduced on PRESENT.
- By design all block ciphers with a block size of 64 bit can have problems with block collisions if they are used with large amounts of data.

Outline

Linear cryptanalysis

- Presented by M. Matsui in 1993
- Makes use of xor sums of bits: xor of some input bits may predict xor of some output bits with high probability independently of the key.
- Input mask $u = u_1, u_2, \ldots, u_n$ is a bit string of the same length n as the input $x = x_1, x_2, \ldots, x_n$. Then xor sum of input bits determined by mask u is

$$u \cdot x = u_1 x_1 \oplus u_2 x_2 \oplus \cdots \oplus u_n x_n$$

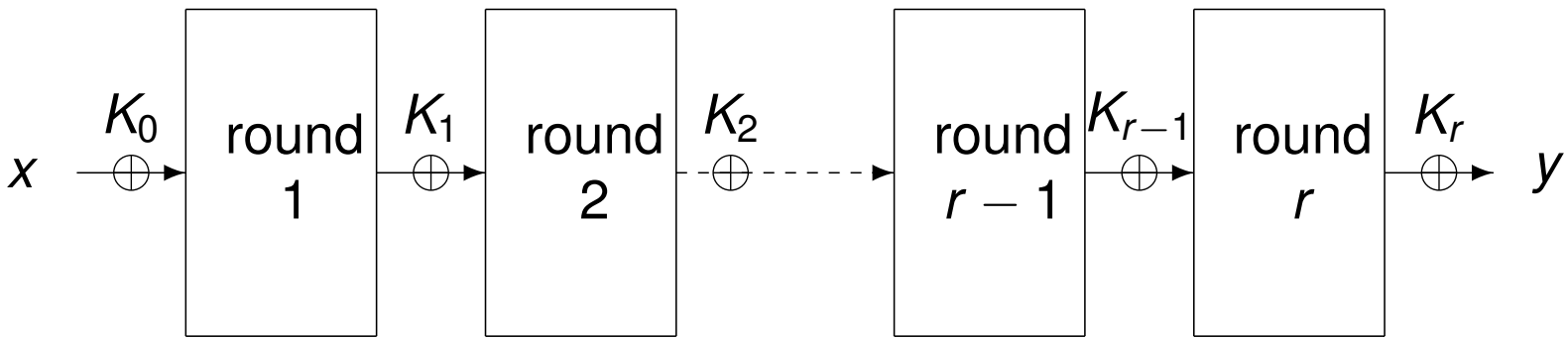
  It means that the mask u picks certain bits to be included in the xor sum.
- Similarly, output mask $v = v_1, v_2, \ldots, v_m$ is a bit string of the same length m as the output $y = y_1, y_2, \ldots, y_m$. Then xor sum of output bits is

$$b \cdot y = v_1 y_1 \oplus v_2 y_2 \oplus \cdots \oplus v_m y_m.$$

Iterated block cipher

SPN and Feistel ciphers are *iterated* block ciphers.

Iterated block cipher processes data and key by iterating a round function:



Correlation

F vectorial Boolean function

$$\mathrm{cor}_F(u, v) = \#\{x \mid u \cdot x = v \cdot F(x)\}$$

E_K is iterated block cipher with round keys K_j, $j = 0, \ldots, r$, and round functions F_i, $i = 1, \ldots, r$

$$\mathrm{cor}_{E_K}(u, v) = \sum_{\gamma} (-1)^{\sum_{j=0}^{r} \gamma_j \cdot K_j} \prod_{i=1}^{r} \mathrm{cor}_{F_i}(\gamma_{i-1}, \gamma_i)$$

where the sum is taken over all $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_r)$, $\gamma_0 = u$, $\gamma_r = v$.

The term, independent of the keys,

$$\prod_{i=1}^{r} \mathrm{cor}_{F_i}(\gamma_{i-1}, \gamma_i)$$

is called the trail correlation of the trail γ.

Trails and hulls

- Given u and v, all trails with non-zero trail correlations contribute to the correlation $\mathrm{cor}_{E_K}(u, v)$. Such family of trails is also called the linear hull.
- Sometimes only part of the hull gives good estimate of the correlation:
  - "dominant" trails for DES
  - single-bit trails for PRESENT
- For n-bit block cipher, one trail correlation less than $2^{-n/2}$ is not a useful estimate for the correlation of the related linear approximation.

Outline

Linear cryptanalysis

Linear and Differential attacks on PRESENT

Other linear cryptanalysis

Link between multidimensional linear and truncated differential attacks

PRESENT: Single-bit masks of Sbox

Sbox

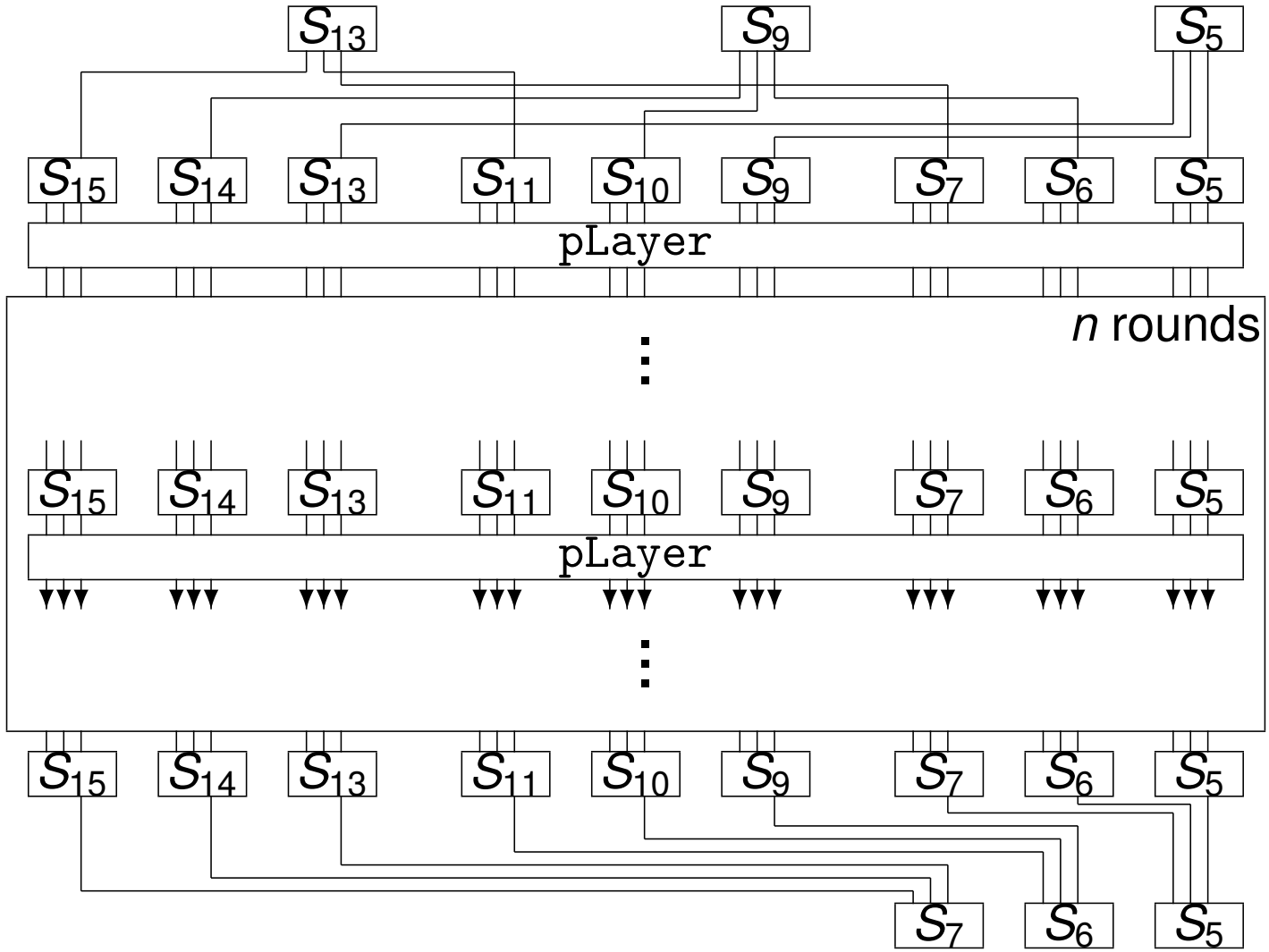
x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	c	5	6	b	9	0	a	d	3	e	f	8	4	7	1	2

Denoting masks u and v as integers 2^i, $i = 0, 1, 2, 3$.

Table gives correlations $\mathrm{cor}_S(u, v)$

$u \backslash v$	1	2	4	8
1	0	0	0	0
2	0	0.25	-0.25	0.25
4	0	-0.25	-0.25	-0.25
8	0	0.25	0	-0.25

Partial hull of single-bit trails



Capacity

Let the set $\Lambda = \{(u, v)\}$ be a linear space of linear approximations

$$\mathrm{Cap}(\Lambda) = \sum_{\substack{(u,v) \in \Lambda \\ (u,v) \neq 0}} \mathrm{cor}(u, v)^2$$

$2^n \mathrm{Cap}(\Lambda)$ follows χ^2 distribution with degrees of freedom $\leq \#\Lambda - 1$
[IACR ePrint 2019/934]

PRESENT – multidimensional linear distinguisher

- input to Sboxes S_5 (mask $u^{(5)}$), S_9 (mask $u^{(9)}$), S_{13} (mask $u^{(13)}$)
- output from Sboxes S_7 (mask $v^{(7)}$), S_6 (mask $v^{(6)}$), S_5 (mask $v^{(5)}$),
- The entire twelve bits to twelve bits multidimensional linear approximation not used
- Instead, we take 9 multidimensional linear approximations

$$u^{(i)} \cdot x + v^{(j)} \cdot E_K(x), i = 5, 9, \text{ or } 13 \text{ and } j = 7, 6, \text{ or } 5$$

  where E_K is the encryption over r rounds, and all the 4-bit masks are of the form $1, 1, 1, 0$.
- This is justified by the assumption: if the mask has nonzero bits on more than one Sbox, then the bits are independent.
- By clustering all single-bit trails Cho obtained the capacity $C = 2^{-55.38}$ over 24 rounds.

Key recovery

First round: guess 16 bits

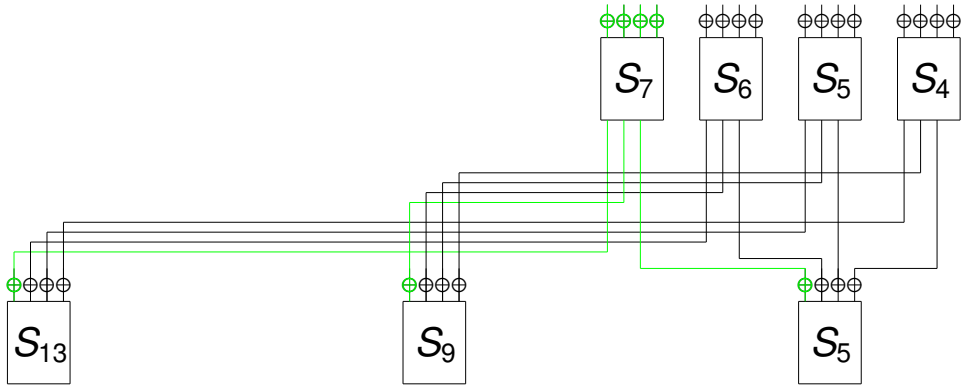


Figure: Key recovery on the first round of PRESENT

Last round similarly.

- χ^2 distinguisher with $9 \cdot (2^8 - 1)$ degrees of freedom (24 rounds)
- Data requirements are

$$N \approx 2C^{-1}\sqrt{a \cdot 9(2^8 - 1)} \approx 2^8 C^{-1}$$

with advantage $a = 8$ bits and success probability 0.95.

Outline

Linear cryptanalysis

Linear and Differential attacks on PRESENT

Other linear cryptanalysis

Link between multidimensional linear and truncated differential attacks

Distinguisher using individual linear approximations

Bogdanov, et al. 2018
Set of $9 \cdot (3 \cdot 6 + 3)$ of individual single-bit mask pairs

input masks $\quad u^{(k)} = 2^{4k+3},\ k = 5, 6, 7, 9, 10, 11, 13, 14, 15$

output masks $\quad v^{(ij)} = \begin{cases} j \cdot 2^{4i+2}, & i = 5, 6, 7, 9, 10, 11,\ j = 1, 2, 3 \\ 2^{4i+3}, & i = 13, 14, 15,\ j = 1 \end{cases}$

26-round key-recovery: distinguisher over 22 rounds

27-round key-recovery: distinguisher over 23 rounds

The attack uses χ^2 distinguisher with 189 degrees of freedom.

Assumption: Linear approximations statistically independent.

Estimating capacity

- Partial clustering of the hull of each approximation.
- An average of 103 482 624 trails were enumerated per hull.
- All trails with Hamming weight 1 masks were enumerated, as well as some trails with Hamming weight 2 and 3 masks.
- The distribution of the “capacity" was then estimated using 2000 random master keys.

A new multiple linear attack Eurocrypt 2020

- María Naya-Plasencia et al.
- Essentially the same distinguisher over 24 rounds
- Using FFT and relations between round-key bits to speed up the key search
- Adds 2 + 2 rounds, to get an attack over 28 rounds.

Outline

Linear cryptanalysis

Linear and Differential attacks on PRESENT

Other linear cryptanalysis

Link between multidimensional linear and truncated differential attacks

Linear Approximation Table LAT

Let F be a function which maps n-bit strings to m-bit strings.
The linear approximation table of F has 2^n rows and 2^m columns.
An entry in position (u, v) is denoted by $\mathrm{LAT}(u, v)$ and defined as follows:

$$\mathrm{LAT}(u, v) =$$
$$\#\{x \in \{0,1\}^n \mid u \cdot x = v \cdot F(x)\} - \#\{x \in \{0,1\}^n \mid u \cdot x \neq v \cdot F(x)\}$$

Then

$$\mathrm{cor}(u, v) = 2^{-n}\mathrm{LAT}(u, v)$$

is the correlation of the linear approximation.

Differential Distribution Table DDT

Let F be a function which maps n-bit strings to m-bit strings.
The DDT of F has 2^n rows and 2^m columns.
An entry in position (a, b) is denoted by $\mathrm{DDT}(a, b)$ and it is defined as follows:

$$\mathrm{DDT}(a, b) = \#\{x \in \{0,1\}^n \mid F(x + a) + F(x) = b\}$$

Then

$$\Pr(a, b) = 2^{-n}\mathrm{DDT}(a, b)$$

is the probability of the differential $F(x + a) + F(x) = b$

The link

$$2^{m+n}\,\mathrm{DDT}(a,b) = \sum_u \sum_v (-1)^{(u,v)\cdot(a,b)} \mathrm{LAT}(u,v)^2$$

$$\mathrm{LAT}(u,v)^2 = \sum_a \sum_b (-1)^{(u,v)\cdot(a,b)}\,\mathrm{DDT}(a,b)$$

A single $\mathrm{DDT}(a,b)$ value can be large even if all $\mathrm{LAT}(u,v)$ values are small, and *vice versa*.

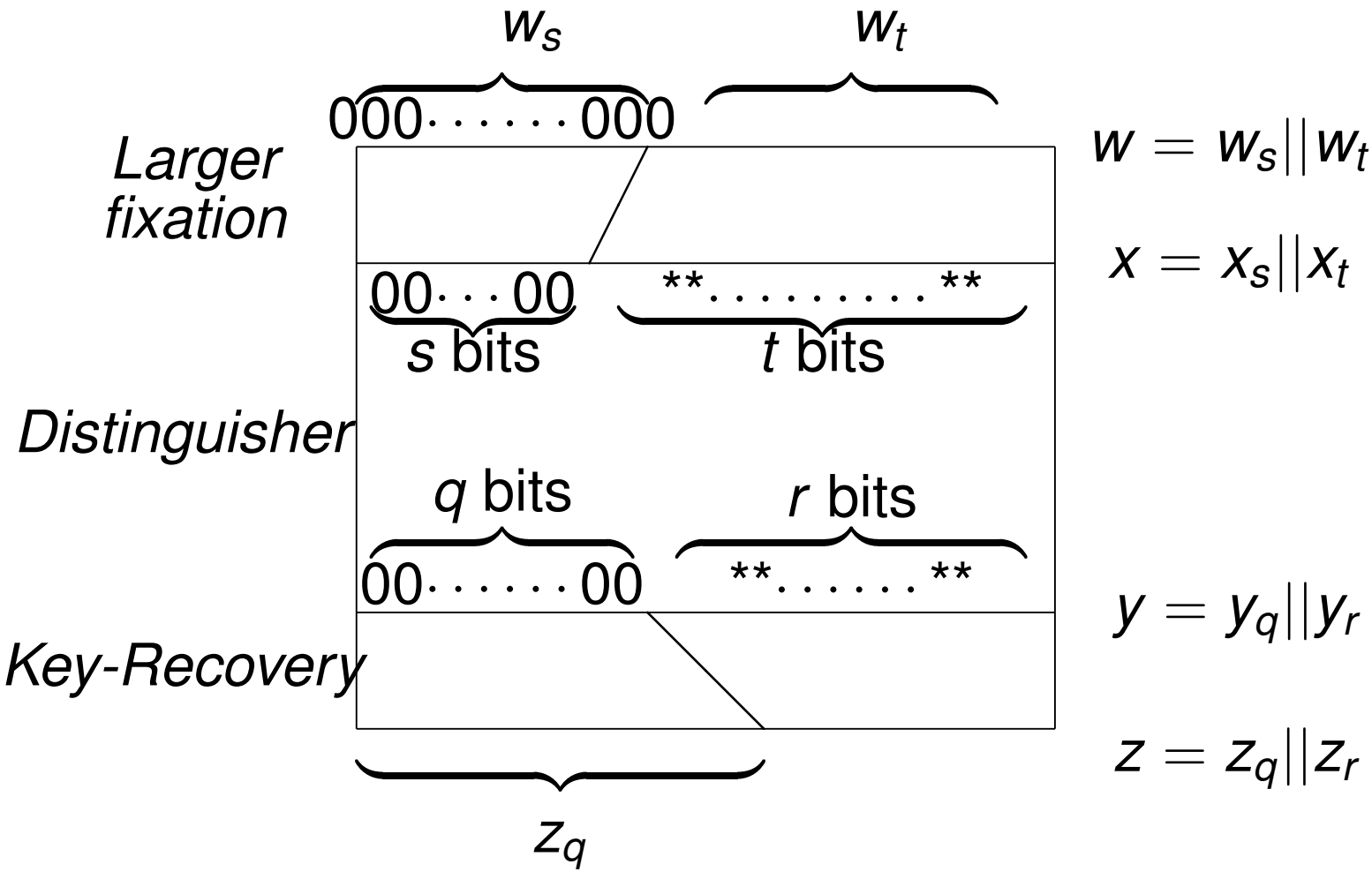
BUT: Taking a linear subspace of differentials gives a property whic is equivalent in distinguishing strength with a property given by a linear subspace of linear approximations [Blondeau-Nyberg 2013].

Link between multidimensional linear approximation and truncated differential

$$2^{-t} \sum_{a_t \in \mathbb{F}_2^t, b_r \in \mathbb{F}_2^r} \Pr_F(a_s||a_t,\, b_q||b_r) =$$

$$2^{-q} \sum_{u_s \in \mathbb{F}_2^s, v_q \in \mathbb{F}_2^q} (-1)^{u_s \cdot a_s \oplus v_q \cdot b_q} \operatorname{cor}_F\left((u_s||0_t) \cdot x \oplus (v_q||0_r) \cdot F(x)\right)^2$$

Subspaces of differentials and linear approximations



Mathematical link

Multidimensional linear approximation:

$$\Lambda = \{(u_s||0_t, v_q||0_r) \mid u_s \in \mathbb{F}_2^s, v_q \in \mathbb{F}_2^q\}$$

Truncated differential Δ with

input differences $\{0_s||a_t \mid a_t \in \mathbb{F}_2^t\}$,
output differences $\{0_q||b_r \mid b_r \in \mathbb{F}_2^r\}$

Probability of Δ is defined as the average probability of all differentials taken over the input differences

$$\Pr(\Delta) = 2^{-t} \sum_{a_t \in \mathbb{F}_2^t, b_r \in \mathbb{F}_2^r} \Pr\left(0_s||a_t, 0_q||b_r\right)$$

The Link:

$$\Pr(\Delta) = 2^{-q}\left(\mathrm{Cap}(\Lambda) + 1\right)$$

Truncated differential attack on PRESENT

- Chosen plaintext pairs with zero difference in certain s bits
- Observing probability of getting output difference zero in certain r bits
- Data requirements are about the same as for the corresponding multidimensional attack; time requirements can possibly be reduced.

It means than such a property can be used either with

- known plaintext (linear cryptanalysis), or
- chosen plaintext (differential cryptanalysis).

Conclusions

- We discussed the best linear attacks on PRESENT.
- The equivalence between multidimensional linear and truncated differential properties.
- The linear attack on PRESENT can be turned to a truncated differential attack.
- The attack of Bogdanov et al. uses less approximations but is more efficient. It relies, however, on artificial assumtions of independence.
- The new attack further improves by extending over 28 rounds.
- Transferring the new attacks to truncated differential ones remains to explored.

Part II: Affine Multidimensional Distiguisher

Outline

Affine multidimensional distinguisher

Affine distinguisher of SIMON

Drawbacks of multidimensional linear approximation

- Multidimensional linear approximations Λ usually involve many linear approximations with correlation zero. Their impact to the capacity is zero!
- This problem was observed already in cryptanalysis of SERPENT [Hermelin et al. 2008) and an ad hoc method of just ignoring them were used in practical computations. Still the statistical χ model required to use $\#\Lambda - 1$ degrees of freedom.
- Recently, two possibilities to avoid involving trivial approximations were proposed [Kahn-Nyberg, IACR ePrint 2019/934]:
  - Davies-Mayer type approximation
  - Affine multidimensional approximation

Structure of multidimensional linear approximation

$$\Lambda = U \oplus V \oplus W$$

Λ contains three disjoint subspaces (only (0,0) in common)

- U composed of masks of the form $(a, 0)$
- V composed of masks of the form $(0, b)$
- W composed of masks of the form (a, b), where $a = 0$ iff $b = 0$.

Let $\dim\Lambda = t$, $\dim U = u$, $\dim V = v$.

Previously, people used χ^2 with $2^t - 1$ degrees of freedom, while the correct value is

$$2^t - 2^u - 2^v + 1.$$

Davies-Mayer structure

So in general, degree of freedom is (at most)

$$2^t - 2^u - 2^v + 1.$$

It is equal to $2^t - 1$ if and only if Λ has $U = \{0\}$ and $V = \{0\}$. In this case we call it Davies-Mayer approximation, that is, all linear approximations of P in Λ are of the form

$$a \cdot (x + A \circ P(x)),$$

where A is a linear permutation.

A single linear approximation is Davies-Mayer. A linear space of approximations is Davies-Mayer if and only if it has a basis (a_i, b_i), $i = 1, \ldots, t$, where $\{a_i,\ i = 1, \ldots, t\}$ and $\{b_i,\ i = 1, \ldots, t\}$ are linearly independent.

Case $\dim(W) = 1$ used in practice. Not known if $\dim(W) > 1$ useful in practice.

Affine multidimensional approximation

Affine multidimensional linear approximation consists of linear approximations of the form

$$(a_0, b_0) + (a, b), \quad \text{where } (a, b) \in H,$$

H is a linear space, and $(a_0, b_0) \notin H$.

We only consider affine sets which do not contain mask pairs of the form $(a, 0)$ where $a \neq 0$, or $(0, b)$, $b \neq 0$.

Let $\dim H = s$. Then χ^2 test with 2^s degrees of freedom can be used.

That is, all linear approximations are meaningful, no trivial approximations included.

Affine sets of linear approximations arise naturally in practical ciphers.

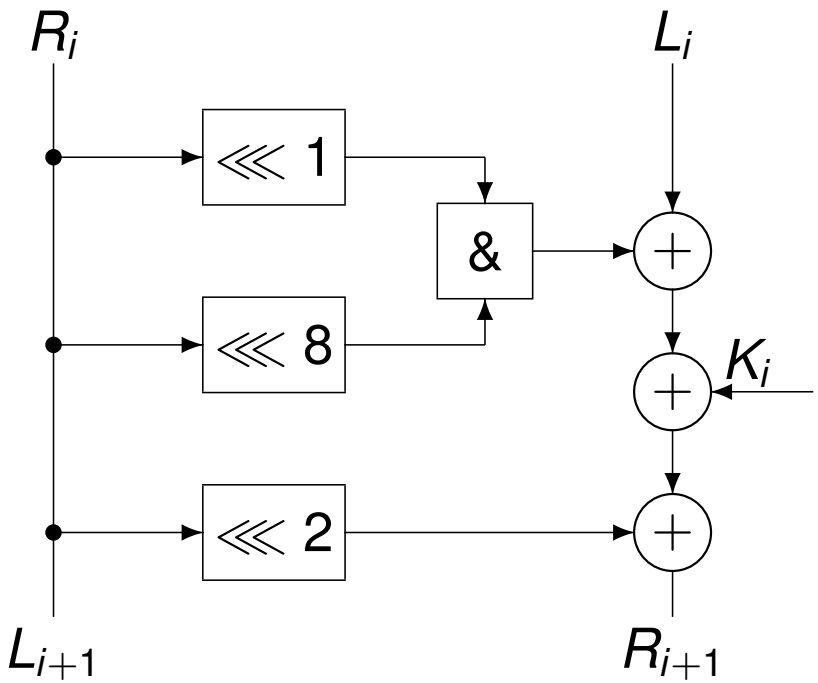
Outline

Affine multidimensional distinguisher

Affine distinguisher of SIMON

SIMON

SIMON is a family of lightweight block ciphers designed by the US National Security Agency (NSA) and published in 2013.
It has 10 members differing in their block and key sizes, and numbers of rounds.
All members of the family are Feistel ciphers. One round:



Linear approximations of AND

- Bitwise AND maps two bits x and y to the single bit $x\&y$.
- All four linear combinations of the bits x and y have nonzero correlation with $x\&y$, all with the same absolute value 2^{-1}.
- The 2-bit to 1-bit AND-function is a bent function and its four linear approximations with nonzero output mask form an affine set with capacity equal to 1.

Linear approximations over one round of SIMON

a the mask on left input half X_i
b the mask on right input half Y_i
Then must have b as left output mask on X_{i+1}.

If b has a single 1-bit, only a single AND operation is activated.
All masks on the right half Y_{i+1} form two-dimensional affine space

$$a \oplus b_{\ggg 2} \oplus \mathrm{sp}\left\{b_{\ggg 1}, b_{\ggg 8}\right\}.$$

Backwords: Mask $b||c$ on the output data $X_{i+1}||Y_{i+1}$, where b has Hamming weight one.
Then the four input masks on the data $X_i||Y_i$ are of the form $a||b$, where

$$a \in c \oplus b_{\ggg 2} \oplus \mathrm{sp}\left\{b_{\ggg 1}, b_{\ggg 8}\right\}.$$

Statistical Test

Null Hypothesis
Permutation under test is a truly random permutation.

Alternative Hypothesis
The permutation is not a truly random permutation.

A affine set of linear approximations of a permutation
Test statistic

$$T = 2^n \mathrm{Cap}(A) = 2^n \sum_{(a,b)\in A} \mathrm{cor}(a,b)^2$$

If permutation is truly random then T follows χ^2 distribution with $\#A$ degrees of freedom.

Distinguishing Advantage

Given threshold τ, the null hypothesis is accepted if $T \leq \tau$, else the alternative hypothesis is accepted.
Set the threshold τ_α in such a way that

$$\Pr(\text{ Alternative Hypothesis is accepted } | \text{ Permutation is random }) = \alpha.$$

Note. Here we need to know the distribution of T in case the permutation is random.

Then the success probability $P_S(\alpha)$ is defined as

$$\Pr(\text{ Alternative Hypothesis is accepted } | \text{ Permutation is cipher })$$

Distinguishing advantage is given as

$$|P_S(\alpha) - \alpha|.$$

Single linear approximations of SIMON

Table: Experimental success probabilities and average squared correlations of single linear approximations derived from the core trail with input mask $4000_x||0001_x$. Experiments used 2^{13} keys.

No. rounds	output mask	av. squared correlation	success probability $P_S(\alpha)$		
			$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0.05$
13	$0100_x\|\|0000_x$	$2^{-29.873}$	0.499	0.391	0.308
14	$0000_x\|\|0100_x$	$2^{-29.873}$	0.499	0.391	0.308
15	$0100_x\|\|0040_x$	$2^{-31.131}$	0.326	0.212	0.143
16	$0040_x\|\|0110_x$	$2^{-31.739}$	0.236	0.131	0.076
17	$0110_x\|\|0004_x$	$2^{-31.975}$	0.204	0.103	0.054
18	$0004_x\|\|0111_x$	$2^{-32.024}$	0.202	0.1	0.047

Affine linear approximations on SIMON

Table: Experimental success probabilities and average capacities for a 2-dimensional affine subspace A of linear approximations with four input masks:
$4000_x||0001_x$, $C000_x||0001_x$, $4100_x||0001_x$, $C100_x||0001_x$ and one output mask. Experiments used 2^{13} keys.

No. rounds	output mask	average capacity	success probability $P_S(\alpha)$		
			$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0.05$
13	$0100_x\|\|0000_x$	$2^{-27.887}$	0.628	0.533	0.463
14	$0000_x\|\|0100_x$	$2^{-27.887}$	0.628	0.533	0.463
15	$0100_x\|\|0040_x$	$2^{-29.133}$	0.422	0.312	0.232
16	$0040_x\|\|0110_x$	$2^{-29.736}$	0.276	0.165	0.099
17	$0110_x\|\|0004_x$	$2^{-29.968}$	0.208	0.105	0.055
18	$0004_x\|\|0111_x$	$2^{-30.000}$	0.198	0.101	0.049

Using multiple linear approximations

- Multiple ($\approx$ 6000) linear approximations over 16 rounds with squared trail correlations 2^{-42}, 2^{-44}, and 2^{-46} collected to sum up to 2^{-32} [Alizadeh et al. 2014].
- Then claimed that this set can be used to distinguish from random.
- Two main problems:
  - Trail correlations less than 2^{-16} do not give any meaningful lower bound estimate to absolute values of correlations. All correlations of linear approximations of any permutation are expected to be around 2^{-16}.
  - Even if the correlations would have been around 2^{-16}, taking a large number of them does not help distinguishing from random.
- But, if some linear approximations have absolute correlations slightly above 2^{-16} then they can be combined to get a better distinguisher.

Conclusions

- Affine sets of linear approximations admit a statistical model (for the random case) without additional assumptions.
- Approaches how to model a cipher given by Blondeau and Nyberg DCC 2017 and ToSC 2016. They do not handle the structure of multidimensional approximation properly.Not even for the random case.
- Combinations of linear approximations are useful only if they can distinguish the cipher from randomm.
- We established a model for the random behavior and showed experiments with a cipher.

Part III: More Links

Outline

Boomerang and Differential-Linear Properties

Extended Tables and Links

Linear Approximation Table LAT

- F a function which maps n-bit strings to m-bit strings
- LAT of F has 2^n rows and 2^m columns.
- An entry in position (u, v) is denoted by $\mathrm{LAT}(u, v)$ and it is defined as follows:

$$\mathrm{LAT}(u, v) =$$
$$\#\{x \in \{0,1\}^n \mid u \cdot x = v \cdot F(x)\} - \#\{x \in \{0,1\}^n \mid u \cdot x \neq v \cdot F(x)\}$$

- Then

$$\mathrm{cor}(u, v) = 2^{-n}\mathrm{LAT}(u, v)$$

  is the correlation of the linear approximation.
- $2^{-n}LAT$ is called the correlation matrix of F [Daemen 1994]

Correlation matrices and nonlinear invariants

[Beyne, Asiacrypt 2018]
$f : \mathbb{F}_2^n \to \mathbb{F}_2$ is non-linear invariant of E_K if

$$f(x) + f(E_K(x)) = c \text{ where } c \text{ is constant.}$$

Link between eigenvectors of the correlation matrix and nonlinear invariants:

Let $\mathcal{C}$ be the correlation matrix[1] of E_K and $(e_0\ v)^\top$ be the correlation matrix of f. Then f is a nonlinear invariant for E_K with constant c if and only if v is an eigenvector of $\mathcal{C}$ with eigenvalue $(-1)^c$.

Here we denoted $e_0 = (1, 0, \ldots, 0)^\top$ (length n).

New weak key classes for Midori and Mantis admitting nonlinear invariants.

[1] It seems that Beyne's correlation matrix is the transpose of the one of Daemen

Differential Distribution Table DDT

- F a function which maps n-bit strings to m-bit strings.
- DDT of F has 2^n rows and 2^m columns.

- An entry in position (a, b) is denoted by $\mathrm{DAT}(a, b)$ and it is defined as follows:

$$\mathrm{DDT}(a, b) = \#\{x \in \{0,1\}^n \mid F(x + a) + F(x) = b\}$$

- Then

$$\Pr(a, b) = 2^{-n}\,\mathrm{DDT}(a, b)$$

  is the probability of the differential $F(x + a) + F(x) = b$.

Link between differential and linear cryptanalysis

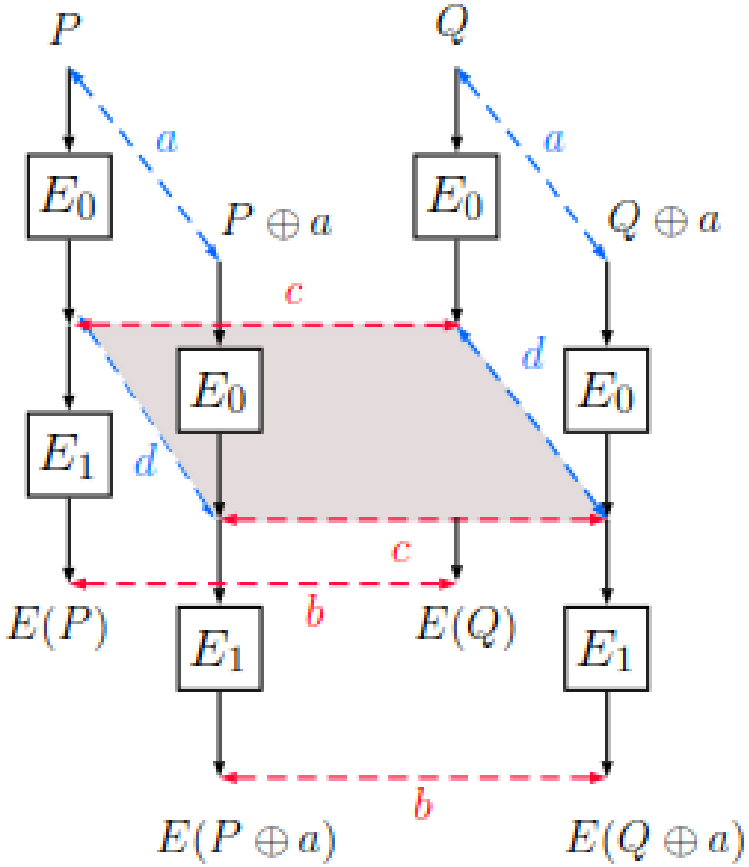
$$2^{m+n}\mathrm{DDT}(a,b) = \sum_u \sum_v (-1)^{(u,v)\cdot(a,b)}\,\mathrm{LAT}((u,v)^2$$

$$\mathrm{LAT}((u,v)^2 = \sum_a \sum_b (-1)^{(u,v)\cdot(a,b)}\,\mathrm{DDT}((a,b)$$

A single $\mathrm{DDT}(a,b)$ value can be large even if all $\mathrm{LAT}(u,v)$ values are small, and *vice versa*.

A linear subspace of differentials is equivalent to a linear subspace of linear approximations [Blondeau-Nyberg2013], see also Part I.

Boomerang attack [Wagner1999]



Boomerang relation $E^{-1}(E(P) \oplus b) \oplus E^{-1}(E(P \oplus a) \oplus b) = a$

Probability estimate: $2^{-4n}\,\mathrm{DDT}_{E_0}(a,d)^2\,\mathrm{DDT}_{E_1}(c,b)^2$

Independence of data between rounds

The differential trails assume independence of data between rounds for one differential, that is, for one pair of data.

The boomerang trails assume independence of data between rounds for a quartet of data. This has been shown problematic [Murphy 2011].

Solution [Cid at al. 2018]

Solution: Handle the boomerang quartet within one layer (in the middle).

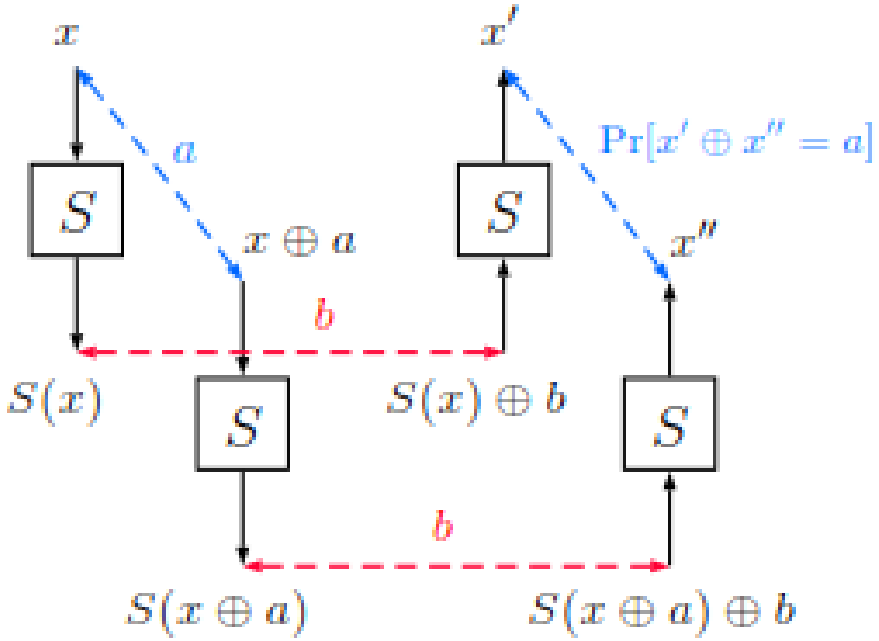


Figure 2: A quartet at the Sbox level

Boomerang Connectivity Table

$F : \{0,1\}^n \to \{0,1\}^m$

Given $a \in \{0,1\}^n$, $b \in \{0,1\}^m$, we set:

$$\mathrm{BCT}_F(a,b) =$$
$$\#\{(x,y) \in \{0,1\}^n \mid F(x) \oplus F(y) = b \text{ and } F(x \oplus a) \oplus F(y \oplus a) = b\}$$

[Cid et al.2018] (bijective F), [Li et al.2019] (general F)

Cipher is decomposed as $E = E_1 \circ F \circ E_0$

$$\mathrm{BCT}_E(a',b') \approx 2^{-4n}\,\mathrm{DDT}_{E_0}(a',a)^2\,\mathrm{BCT}_F(a,b)\,\mathrm{DDT}_{E_1}(b,b')^2$$

BCT table of x^{-1}

Table 2: BCT of the permutation $x \mapsto x^{2^n-2}$ for $n = 4$

	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
1	16	4	0	0	0	0	6	6	0	2	0	2	0	2	2	0
2	16	0	0	6	0	0	0	2	0	4	6	0	2	2	0	2
3	16	0	6	0	0	0	0	2	2	0	2	2	6	0	4	0
4	16	0	0	0	0	6	2	0	6	0	2	0	0	4	2	2
5	16	0	0	0	6	0	2	0	2	2	0	4	2	0	0	6
6	16	6	0	0	2	2	6	4	0	0	2	0	2	0	0	0
7	16	6	2	2	0	0	4	6	2	0	0	0	0	0	0	2
8	16	0	0	2	6	2	0	2	0	0	0	6	0	0	2	4
9	16	2	4	0	0	2	0	0	0	2	0	0	6	0	6	2
a	16	0	6	2	2	0	2	0	0	0	0	0	4	2	6	0
b	16	2	0	2	0	4	0	0	6	0	0	2	2	6	0	0
c	16	0	2	6	0	2	2	0	0	6	4	2	0	0	0	0
d	16	2	2	0	4	0	0	0	0	0	2	6	0	2	0	6
e	16	2	0	4	2	0	0	0	2	6	6	0	0	0	2	0
f	16	0	2	0	2	6	0	2	4	2	0	0	0	6	0	0

BCT table of APN function with $m = n$

$\sum_b \mathrm{BCT}(a,b) = \sum_b \mathrm{DDT}(a,b)^2$, for all a

$\mathrm{BCT}(a,b) \geq \mathrm{DDT}(a,b)$

Then for all $a \neq 0$

$\max_{b \neq 0} \mathrm{BCT}(a,b) = 2$ if and only if $\max_b \mathrm{DDT}(a,b) = 2$.

max BCT values of optimal Sboxes

Bijective 4×4 Sboxes with optimal nonlinearity (8) and differential uniformity (4)

	Representative	$\mathcal{L}(S)$	[DeC07]	[LP07]	n_0	n_2	n_4	n_6	n_8	n_{10}	n_{16}	β_S
1	[8, 0, 1, 12, 15, 5, 6, 7, 4, 3, 10, 11, 9, 13, 14, 2]	8	3	G_3	120	60	15	30	0	0	0	**6**
2	[2, 0, 1, 8, 3, 11, 6, 7, 4, 9, 10, 15, 12, 13, 14, 5]	8	6	G_5	108	72	27	18	0	0	0	**6**
3	[8, 0, 1, 12, 2, 5, 6, 9, 4, 3, 10, 11, 7, 13, 14, 15]	8	2	G_6	104	80	27	10	4	0	0	**8**
4	[8, 0, 1, 9, 2, 5, 13, 7, 4, 6, 10, 11, 12, 3, 14, 15]	8	8	G_{11}	100	85	30	5	5	0	0	**8**
5	[4, 0, 1, 15, 2, 11, 6, 7, 3, 9, 10, 5, 12, 13, 14, 8]	8	1	G_{13}	105	78	28	11	2	1	0	**10**
6	[2, 0, 1, 8, 3, 13, 6, 7, 4, 9, 10, 5, 12, 11, 14, 15]	8	4	G_4	112	72	23	14	0	4	0	**10**
7	[2, 0, 1, 8, 3, 15, 6, 7, 4, 9, 5, 11, 12, 13, 14, 10]	8	5	G_7	105	80	30	5	0	5	0	**10**
8	[4, 8, 1, 2, 3, 11, 6, 7, 0, 9, 10, 14, 12, 13, 5, 15]	8	7	G_{12}	110	75	25	10	0	5	0	**10**
9	[8, 14, 1, 2, 3, 5, 6, 7, 4, 12, 10, 11, 9, 13, 0, 15]	8	9	G_9	108	69	28	14	5	1	0	**10**
10	[8, 14, 1, 2, 3, 5, 6, 7, 4, 9, 15, 11, 12, 13, 0, 10]	8	10	G_{14}	108	70	27	13	6	1	0	**10**
11	[8, 15, 1, 2, 3, 5, 12, 7, 4, 9, 10, 11, 6, 13, 14, 0]	8	11	G_{15}	108	70	27	13	6	1	0	**10**
12	[8, 15, 1, 2, 3, 5, 6, 13, 4, 9, 10, 11, 12, 7, 14, 0]	8	12	G_{10}	108	69	30	12	3	3	0	**10**
13	[12, 0, 1, 9, 3, 5, 4, 7, 6, 2, 10, 11, 8, 13, 14, 15]	8	13	G_2	107	64	32	8	12	0	2	**16**
14	[12, 11, 1, 2, 3, 5, 4, 7, 6, 9, 10, 0, 8, 13, 14, 15]	8	14	G_1	107	60	36	12	8	0	2	**16**
15	[12, 9, 1, 2, 3, 5, 4, 7, 6, 0, 10, 11, 8, 13, 14, 15]	8	15	G_8	103	72	32	0	16	0	2	**16**
16	[8, 14, 1, 2, 3, 5, 4, 7, 6, 9, 10, 0, 12, 13, 11, 15]	8	16	G_0	107	64	32	8	12	0	2	**16**

Differential-linear cryptanalysis

$F : \{0,1\}^n \to \{0,1\}^m$

$a \in \{0,1\}^n$, $v \in \{0,1\}^m$

Differential-linear relation: $v \cdot (F(x) \oplus F(x \oplus a))$

$$\begin{aligned} \mathrm{DLT}_F(a,v) \quad = & \\ & \#\{x \in \{0,1\}^n \mid v \cdot (F(x) \oplus F(x \oplus a)) = 0\} \\ - & \#\{x \in \{0,1\}^n \mid v \cdot (F(x) \oplus F(x \oplus a)) \neq 0\} \end{aligned}$$

New approach [Bar-On et al. 2019]

Traditionally [Langford-Hellman1995], [Biham et al. 2002] $E = E_1 \circ E_0$

$$\mathrm{DLT}_E(a, v) \approx 2^{-2n}\mathrm{DDT}_{E_0}(a, b)\mathrm{LAT}_{E_1}(u, v)^2$$

where $b \in \mathrm{sp}\{u\}^{\perp}$.

New $E = E_1 \circ F \circ E_0$

$$\mathrm{DLT}_E(a, v) \approx 2^{-3n}\,\mathrm{DDT}_{E_0}(a, b)\,\mathrm{DLT}_F(b, u)\,\mathrm{LAT}_{E_1}(u, v)^2$$

[Li et al. 2019] DLT is the same as previously known autocorrelation table ACT [Zhang et al. 2000]

DLT values of optimal Sboxes [Li et al. 2019]

TABLE I
REPRESENTATIVES FOR ALL 16 CLASSES OF OPTIMAL 4 BIT SBOXES

F_0	$0, 1, 2, 13, 4, 7, 15, 6, 8, 11, 12, 9, 3, 14, 10, 5$
F_1	$0, 1, 2, 13, 4, 7, 15, 6, 8, 11, 14, 3, 5, 9, 10, 12$
F_2	$0, 1, 2, 13, 4, 7, 15, 6, 8, 11, 14, 3, 10, 12, 5, 9$
F_3	$0, 1, 2, 13, 4, 7, 15, 6, 8, 12, 5, 3, 10, 14, 11, 9$
F_4	$0, 1, 2, 13, 4, 7, 15, 6, 8, 12, 9, 11, 10, 14, 5, 3$
F_5	$0, 1, 2, 13, 4, 7, 15, 6, 8, 12, 11, 9, 10, 14, 3, 5$
F_6	$0, 1, 2, 13, 4, 7, 15, 6, 8, 12, 11, 9, 10, 14, 5, 3$
F_7	$0, 1, 2, 13, 4, 7, 15, 6, 8, 12, 14, 11, 10, 9, 3, 5$
F_8	$0, 1, 2, 13, 4, 7, 15, 6, 8, 14, 9, 5, 10, 11, 3, 12$
F_9	$0, 1, 2, 13, 4, 7, 15, 6, 8, 14, 11, 3, 5, 9, 10, 12$
F_{10}	$0, 1, 2, 13, 4, 7, 15, 6, 8, 14, 11, 5, 10, 9, 3, 12$
F_{11}	$0, 1, 2, 13, 4, 7, 15, 6, 8, 14, 11, 10, 5, 9, 12, 3$
F_{12}	$0, 1, 2, 13, 4, 7, 15, 6, 8, 14, 11, 10, 9, 3, 12, 5$
F_{13}	$0, 1, 2, 13, 4, 7, 15, 6, 8, 14, 12, 9, 5, 11, 10, 3$
F_{14}	$0, 1, 2, 13, 4, 7, 15, 6, 8, 14, 12, 11, 3, 9, 5, 10$
F_{15}	$0, 1, 2, 13, 4, 7, 15, 6, 8, 14, 12, 11, 9, 3, 10, 5$

TABLE II
AUTOCORRELATION SPECTRA AND DLU OF F_i FOR $0 \leq i \leq 15$

F_i	$i = 3 \sim 7, 11 \sim 13$	$i = 0 \sim 2, 8 \sim 10, 14, 15$
DL-Walsh spectrum	$\{-8, 0, 8\}$	$\{-16, -8, 0, 8, 16\}$
DLCT spectrum	$\{-4, 0, 4\}$	$\{-8, -4, 0, 4, 8\}$
DLU	4	8

Known Links Summary

$$\mathrm{DDT}(a,b) = 2^{-m}2^{-n}\sum_u(-1)^{u\cdot a}\sum_v(-1)^{v\cdot b}\,\mathrm{LAT}((u,v)^2$$

$$\mathrm{DLT}(a,v) = 2^{-n}\sum_u(-1)^{u\cdot a}\,\mathrm{LAT}(u,v)^2$$

$$\mathrm{DLT}(a,v) = \sum_b(-1)^{v\cdot b}\,\mathrm{DDT}(a,b)$$

$$\sum_b \mathrm{BCT}(a,b) = \sum_b \mathrm{DDT}(a,b)^2$$

$$\sum_b \mathrm{BCT}(\mathrm{a},\mathrm{b}) = 2^{-m}\sum_v \mathrm{DLT}(a,v)^2$$

Discussion

- We discussed old and more recently discovered Sbox criteria
- DDT, LAT and DLT are linked: given one of them the other two can be computed. BCT looks like an outlier.
- Links imply equivalences between the corresponding attacks, but with different parameters. E.g. a differential type of attack can be feasible, while the corresponding linear type of attack is not.
- On the other hand, e.g. multidimensional differential-linear attack is equivalent to a truncated differential attack.

Outline

Boomerang and Differential-Linear Properties

Extended Tables and Links

Extended Autocorrelation Table EACT

$$\mathrm{EACT}_F(a; v, w) = 2\#\{\, x \in \mathbb{F}_2^n \mid v \cdot F(x) \oplus w \cdot F(x \oplus a) = 0 \,\} - 2^n.$$

Then

$$\mathrm{EACT}(a; v, w) = \mathrm{EACT}(a; w, v) \text{ for all } v, w \in \mathbb{F}_2^m.$$

In particular,

$$\mathrm{EACT}(a; v, v) = \mathrm{ACT}(a, v) \text{ for all } a \in \mathbb{F}_2^n,\ v \in \mathbb{F}_2^m.$$

EACT relates to a differential-linear attack where the masks on $F(x)$ and $F(x \oplus a)$ can be different.

Example

EACT of the following 4×4 Sbox S

x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
$S(x)$	E	4	D	1	2	F	B	8	3	A	6	C	5	9	0	7

Input difference $a = 1011$
Output masks $v = 0100$ and $w = 0001$:

$$\mathrm{EACT}(a; v, w) = -16$$
$$\mathrm{EACT}(a; w, v) = -16$$
$$\mathrm{EACT}(a; v \oplus w, v \oplus w) = \mathrm{ACT}(a, v \oplus w) = 16$$

where $v \oplus w$ is a two-bit mask.

$$\max_{Hwt(v')=1} |\mathrm{ACT}(a, v')| = 8$$

obtained for $v' = 1000$.

Link between boomerangs and EACT

$$\mathrm{BCT}(a,b) = 2^{-2m} \sum_{v,w} (-1)^{(u+w)\cdot b} \, \mathrm{EACT}(a; v, w)^2.$$

Inverse direction? We must extend BCT.

Extended BCT

$$\mathrm{EBCT}_F(a; b, c) =$$
$$\#\{\,(x, y) \in \mathbb{F}_2^n \times \mathbb{F}_2^n \mid F(x) \oplus F(y) = b \text{ and } F(x \oplus a) \oplus F(y \oplus a) = c\,\}.$$

Then

$$\mathrm{EACT}(a; v, w)^2 = \sum_{b,c} (-1)^{v \cdot b \oplus w \cdot c}\, \mathrm{EBCT}(a; b, c).$$

We can show, certain multidimensional differential-linear properties are equivalent to ceratain truncated boomerangs.

Conclusions

- We considered tools for evaluating known non-linearity properties of Sboxes and super Sboxes.
- We gave unified summary of links between these properties for general vectorial Boolean functions (some literature only considers bijective functions)
- Autocorrelation and boomerang connectivity tables are not directly linked.
- We established the link via extended autocorrelation and boomerang tables.
- Using the approach of [Blondeay-Nyberg 2013] one can establish equivalence between truncated boomerangs and multidimensional diffrential-linear properties.