Aalto University
School of Science

Complexity of Statistical Attacks

On the Relation between Chosen and Known Plaintext Attacks

Céline Blondeau and Kaisa Nyberg

January 2014

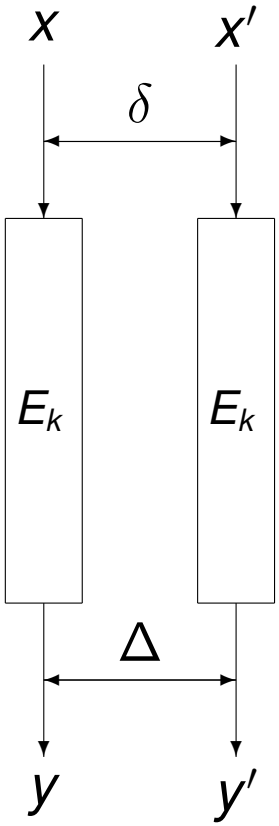
Dagstuhl

Outline

Outline

Differential Cryptanalysis [Biham Shamir 90]

Difference between plaintext and ciphertext pairs



Input difference : δ
Output Difference : Δ

Differential Probability :

$$\mathbf{P}[\delta \rightarrow \Delta] \quad = \quad P_{k,x}[\, E_k(x) \oplus E_k(x \ \oplus \ \delta) = \Delta]$$

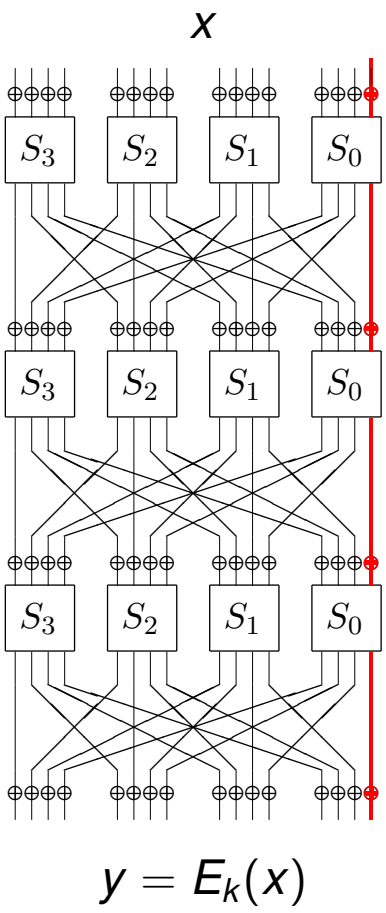
Truncated Differential (TD) [Knudsen 94] :

Set of input differences : $\delta \in A$
Set of output differences : $\Delta \in B$

$$\mathbf{P}[A \rightarrow B] = \frac{1}{|A|} \sum_{\delta \in A} \sum_{\Delta \in B} P[\delta \rightarrow \Delta]$$

Linear Cryptanalysis [Tardy Gilbert 91] [Matsui 93]

Linear relation involving plaintext, key and ciphertext bits



Input mask : u

Key mask : κ

Output mask : v

Bias :

$$\varepsilon = 2^{-n}\#\{x \in \mathbb{F}_2^n |\ u \cdot x \oplus \kappa \cdot k \oplus v \cdot y = 0\} - \frac{1}{2}$$

Correlation : $\mathbf{cor}_x(u, v) = 2\varepsilon$

Multidimensional linear (ML) [Hermelin et al 08] :

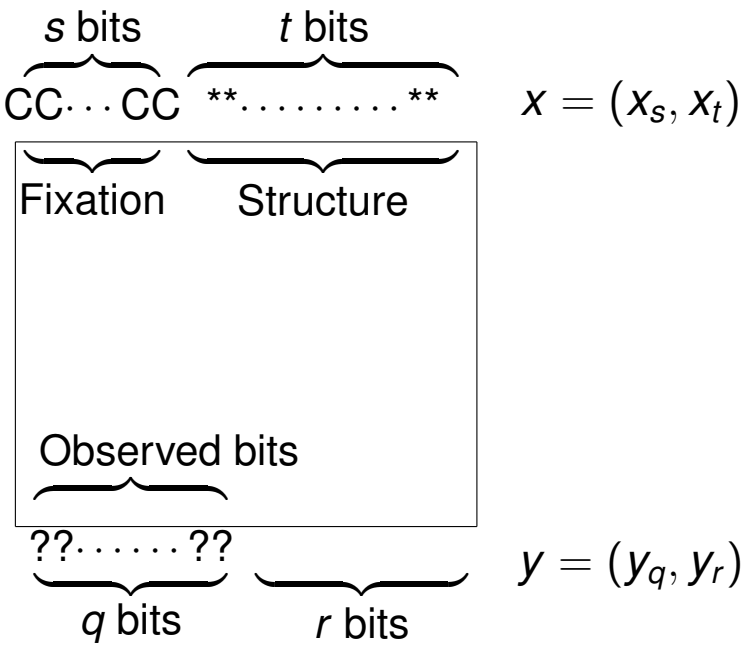
Set of masks $(u, v) \in U \times V \setminus \{0, 0\}$

Capacity : $C = \sum_{u \in U} \sum_{v \in V} \mathbf{cor}_x^2(u, v)$

Outline

Statistical Saturation Attack [Collard Standaert 09]

The distinguisher :



$D = 0$

for M values of $x_s \in \mathbb{F}_2^s$ **do**

  Initialize a table T of size 2^q

  for $S(= 2^t)$ values of $x_t \in \mathbb{F}_2^t$ **do**

    $(y_q, y_r) = E((x_s, x_t))$

    $T[y_q] += 1$

  for all $y_q \in \mathbb{F}_2^q$ **do**

    $D += T[y_q]^2$

Statistical Saturation Attack [Collard Standaert 09]

The distinguisher :

For M structures

For (all) elements in a structure

Count the number of occurrences of y_q

Compute the statistic

$D = 0$
for M values of $x_s \in \mathbb{F}_2^s$ **do**
  Initialize a table T of size 2^q
  for $S(= 2^t)$ values of $x_t \in \mathbb{F}_2^t$ **do**
    $(y_q, y_r) = E((x_s, x_t))$
    $T[y_q] += 1$
  for all $y_q \in \mathbb{F}_2^q$ **do**
    $D += T[y_q]^2$

Statistical Saturation Attack [Collard Standaert 09]

The distinguisher :

For M structures

For (all) elements in a structure

Count the number of occurrences of y_q

Compute the statistic

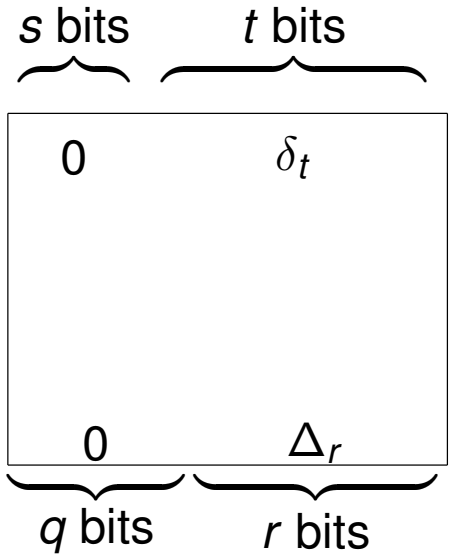
$D = 0$
for M values of $x_s \in \mathbb{F}_2^s$ **do**
  Initialize a table T of size 2^q
  for $S(= 2^t)$ values of $x_t \in \mathbb{F}_2^t$ **do**
    $(y_q, y_r) = E((x_s, x_t))$
    $T[y_q] += 1$
  for all $y_q \in \mathbb{F}_2^q$ **do**
    $D += T[y_q]^2$

The key-recovery attack :

Adding rounds,

- At the end $\Rightarrow$ time and memory complexity cost
- At the beginning $\Rightarrow$ also data complexity cost

Truncated Differential Distinguishers (1)



$x = (x_s, x_t)$

$y = (y_q, y_r)$

$D = 0$

for M values of $x_s \in \mathbb{F}_2^s$ **do**

  Create a table T of size S

  for $S(= 2^t)$ values of $x_t \in \mathbb{F}_2^t$ **do**

    $(y_q, y_r) = E((x_s, x_t))$

    $T[x_t] = y_q$

  for all pairs (x_t, x_t') **do**

    if $(T[x_t] \oplus T[x_t']) == 0$ **then**

      $D+ = 1$

Truncated Differential Distinguishers (1)

For M structures

For all elements in a structure

Store the partial ciphertexts

Count the number of pairs which have no difference in $\mathbb{F}_2^q$

$D = 0$
for M values of $x_s \in \mathbb{F}_2^s$ **do**
  Create a table T of size S
  for $S(= 2^t)$ values of $x_t \in \mathbb{F}_2^t$ **do**
    $(y_q, y_r) = E((x_s, x_t))$
    $T[x_t] = y_q$
  for all pairs (x_t, x_t') **do**
    if $(T[x_t] \oplus T[x_t']) == 0$ **then**
      $D+ = 1$

Truncated Differential Distinguishers (1)

For M structures

For all elements in a structure

Store the partial ciphertexts

Count the number of pairs which have no difference in $\mathbb{F}_2^q$

$D = 0$
for M values of $x_s \in \mathbb{F}_2^s$ **do**
  Create a table T of size S
  for $S(= 2^t)$ values of $x_t \in \mathbb{F}_2^t$ **do**
    $(y_q, y_r) = E((x_s, x_t))$
    $T[x_t] = y_q$
  for all pairs (x_t, x_t') **do**
    if $(T[x_t] \oplus T[x_t']) == 0$ **then**
      $D+ = 1$

Time complexity : dominated by the step consisting at verifying all pairs

$$Time \approx M \cdot S^2/2$$

TD with Less Time Complexity

- Checking if the ciphertext pairs have no difference in $\mathbb{F}_2^q$

Example :

- 4 ciphertexts : $(y_1, b_1)\ (y_2, b_2)\ (y_1, b_3)\ (y_4, b_4)$
  1 pair with equal y_q
- Previous algorithm : 6 comparisons

TD with Less Time Complexity

- Checking if the ciphertext pairs have no difference in $\mathbb{F}_2^q$

Example :

- 4 ciphertexts : $(y_1, b_1)\ (y_2, b_2)\ (y_1, b_3)\ (y_4, b_4)$
  1 pair with equal y_q
- Previous algorithm : 6 comparisons
- Counting the number of occurrences of each y_q :

	y_1	y_2	y_3
$T[y_q]$	2	1	1

and computing $D = \sum_q T[y_q](T[y_q] - 1)/2 = 1$

Complexity : storing the number of occurrences

Truncated Differential Distinguishers (2)

For M structures

For all elements in a structure

Count the number of occurrences of the partial ciphertexts

Compute the statistic

$D = 0$
for M values of $x_s \in \mathbb{F}_2^s$ **do**
  Create a table T of size 2^q
  for $S(= 2^t)$ values of $x_t \in \mathbb{F}_2^t$ **do**
    $(y_q, y_r) = E((x_s, x_t))$
    $T[y_q] += 1$
  for all $y_q \in \mathbb{F}_2^q$ **do**
    $D += T[y_q](T[y_q] - 1)/2$

Truncated Differential Distinguishers (2)

For M structures

For all elements in a structure

Count the number of occurrences of the partial ciphertexts

Compute the statistic

$D = 0$
for M values of $x_s \in \mathbb{F}_2^s$ **do**
  Create a table T of size 2^q
  for $S(= 2^t)$ values of $x_t \in \mathbb{F}_2^t$ **do**
    $(y_q, y_r) = E((x_s, x_t))$
    $T[y_q] += 1$
  for all $y_q \in \mathbb{F}_2^q$ **do**
    $D += T[y_q](T[y_q] - 1)/2$

Remark :

$$\sum_{y_q} T[y_q] \cdot (T[y_q] - 1)/2 = \frac{1}{2} \sum_{y_q} T[y_q]^2 - \overbrace{\sum_{y_q} T[y_q]}^{S}$$

For each structure, a capacity is computed : $\sum_{y_q} T[y_q]^2$

Truncated Differential Distinguishers (2)

For M structures

For all elements in a structure

Count the number of occurrences of the partial ciphertexts

Compute the statistic

$D = 0$
for M values of $x_s \in \mathbb{F}_2^s$ **do**
  Create a table T of size 2^q
  for $S(= 2^t)$ values of $x_t \in \mathbb{F}_2^t$ **do**
    $(y_q, y_r) = E((x_s, x_t))$
    $T[y_q] += 1$
  for all $y_q \in \mathbb{F}_2^q$ **do**
    $D += T[y_q](T[y_q] - 1)/2$

Remark :

$$\sum_{y_q} T[y_q] \cdot (T[y_q] - 1)/2 = \frac{1}{2} \sum_{y_q} T[y_q]^2 - \overbrace{\sum_{y_q} T[y_q]}^{S}$$

For each structure, a capacity is computed : $\sum_{y_q} T[y_q]^2$

- This distinguisher is the same as the SS distinguisher

Link between SS and ML distinguishers

Link [Leander 11] :

For a fixed $x_s \in \mathbb{F}_2^s$ we denote by $C(x_s)$ the capacity of the distribution of y_q :

$$C = 2^{-s} \sum_{x_s \in \mathbb{F}_2^s} C(x_s)$$

Link between SS and ML distinguishers

Link [Leander 11] :

For a fixed $x_s \in \mathbb{F}_2^s$ we denote by $C(x_s)$ the capacity of the distribution of y_q :

$$C = 2^{-s} \sum_{x_s \in \mathbb{F}_2^s} C(x_s)$$

SS is chosen plaintext (CP) attack
ML is known plaintext (KP) attack

Link between SS and ML distinguishers

Link [Leander 11] :
For a fixed $x_s \in \mathbb{F}_2^s$ we denote by $C(x_s)$ the capacity of the distribution of y_q :

$$C = 2^{-s} \sum_{x_s \in \mathbb{F}_2^s} C(x_s)$$

SS is chosen plaintext (CP) attack
ML is known plaintext (KP) attack

To summarize :

- SS attacks link mathematically with ML attacks
- SS attacks link algorithmically with TD attacks

Outline

Link between Differential and Linear Cryptanalysis

[Chabaud Vaudenay 94]

Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$

$$\mathbf{P}[\delta \to \Delta] = 2^{-m} \sum_{u \in \mathbb{F}_2^n} \sum_{v \in \mathbb{F}_2^m} (-1)^{u \cdot \delta \oplus v \cdot \Delta} \mathbf{cor}_x^2(u, v)$$

Link between Differential and Linear Cryptanalysis

[Chabaud Vaudenay 94]

Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$

$$\mathbf{P}[\delta \to \Delta] = 2^{-m} \sum_{u \in \mathbb{F}_2^n} \sum_{v \in \mathbb{F}_2^m} (-1)^{u \cdot \delta \oplus v \cdot \Delta} \mathbf{cor}_x^2(u, v)$$

This link in the literature :

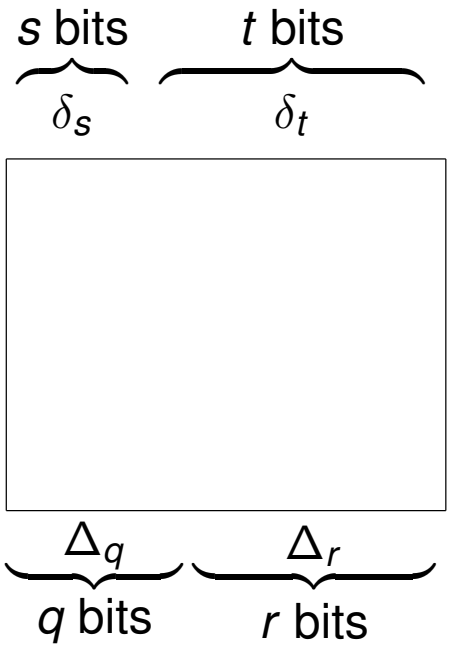
- almost bent (AB) functions are almost perfect non-linear (APN)
- ...
- [Blondeau Nyberg 13] : Computation of truncated differential probability using square correlations

Generalization of the Previous Link

For all $\delta_s \in \mathbb{F}_2^s$ and $\Delta_q \in \mathbb{F}_2^q$,

$$2^{-t} \sum_{\delta_t \in \mathbb{F}_2^t, \Delta_r \in \mathbb{F}_2^r} \mathbf{P}[(\delta_s, \delta_t) \xrightarrow{F} (\Delta_q, \Delta_r)] =$$

$$2^{-q} \sum_{u_s \in \mathbb{F}_2^s, v_q \in \mathbb{F}_2^q} (-1)^{u_s \cdot \delta_s \oplus v_q \cdot \Delta_q} \mathbf{cor}_x^2 \left((u_s, 0) \cdot x \oplus (v_q, 0) \cdot F(x)\right)$$

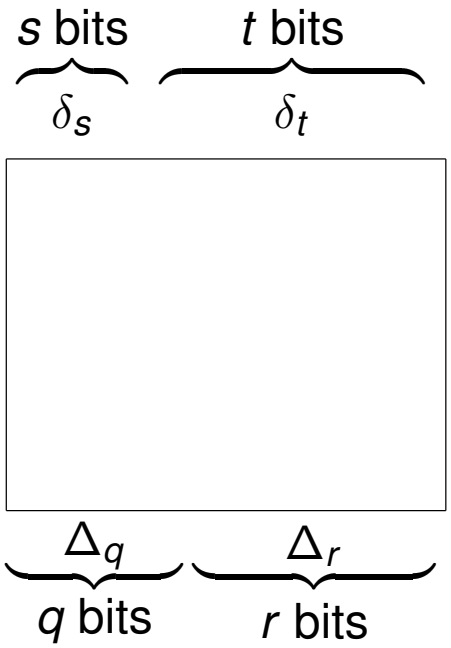


Generalization of the Previous Link

For all $\delta_s \in \mathbb{F}_2^s$ and $\Delta_q \in \mathbb{F}_2^q$,

$$2^{-t} \sum_{\delta_t \in \mathbb{F}_2^t, \Delta_r \in \mathbb{F}_2^r} \mathbf{P}[(\delta_s, \delta_t) \xrightarrow{F} (\Delta_q, \Delta_r)] =$$

$$2^{-q} \sum_{u_s \in \mathbb{F}_2^s, v_q \in \mathbb{F}_2^q} (-1)^{u_s \cdot \delta_s \oplus v_q \cdot \Delta_q} \mathbf{cor}_x^2 \left((u_s, 0) \cdot x \oplus (v_q, 0) \cdot F(x)\right)$$



If $\delta_s = 0$ and $\Delta_q = 0$

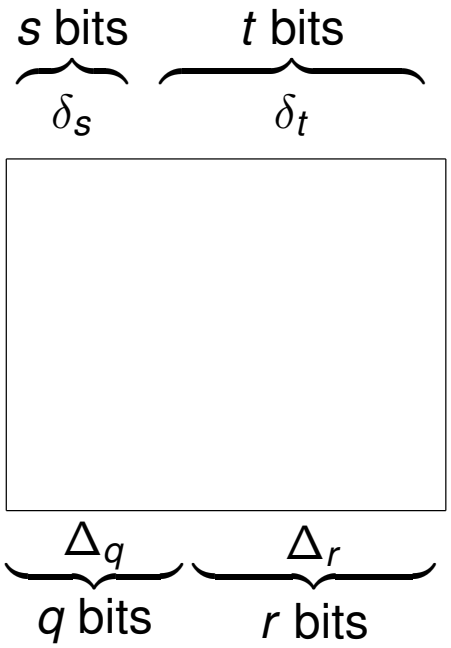
$$2^{-t} \sum_{\delta_t \in \mathbb{F}_2^t, \Delta_r \in \mathbb{F}_2^r} \mathbf{P}[(0, \delta_t) \xrightarrow{F} (0, \Delta_r)] = 2^{-q} \sum_{u_s \in \mathbb{F}_2^s, v_q \in \mathbb{F}_2^q} \mathbf{cor}_x^2 \left((u_s, 0) \cdot x \oplus (v_q, 0) \cdot F(x)\right)$$

Generalization of the Previous Link

For all $\delta_s \in \mathbb{F}_2^s$ and $\Delta_q \in \mathbb{F}_2^q$,

$$2^{-t} \sum_{\delta_t \in \mathbb{F}_2^t, \Delta_r \in \mathbb{F}_2^r} \mathbf{P}[(\delta_s, \delta_t) \xrightarrow{F} (\Delta_q, \Delta_r)] =$$

$$2^{-q} \sum_{u_s \in \mathbb{F}_2^s, v_q \in \mathbb{F}_2^q} (-1)^{u_s \cdot \delta_s \oplus v_q \cdot \Delta_q} \mathbf{cor}_x^2 \left((u_s, 0) \cdot x \oplus (v_q, 0) \cdot F(x)\right)$$



If $\delta_s = 0$ and $\Delta_q = 0$

$$\underbrace{2^{-t} \sum_{\delta_t \in \mathbb{F}_2^t, \Delta_r \in \mathbb{F}_2^r} \mathbf{P}[(0, \delta_t) \xrightarrow{F} (0, \Delta_r)]}_{p} = \underbrace{2^{-q} \sum_{u_s \in \mathbb{F}_2^s, v_q \in \mathbb{F}_2^q} \mathbf{cor}_x^2 \left((u_s, 0) \cdot x \oplus (v_q, 0) \cdot F(x)\right)}_{2^{-q}(C+1)}$$

$$p = 2^{-q}(C+1)$$

TD and ML

- ML : $[(u_s, 0), (v_q, 0)]_{u_s \in \mathbb{F}_2^s \setminus \{0\},\ v_q \in \mathbb{F}_2^q}$ with capacity C
- TD : $[(0, \delta_t), (0, \Delta_r)]_{\delta_t \in \mathbb{F}_2^t,\ \Delta_r \in \mathbb{F}_2^r}$ with probability p

$$p = 2^{-q}(C + 1)$$

TD and ML

- ML : $[(u_s, 0), (v_q, 0)]_{u_s \in \mathbb{F}_2^s \setminus \{0\},\ v_q \in \mathbb{F}_2^q}$ with capacity C
- TD : $[(0, \delta_t), (0, \Delta_r)]_{\delta_t \in \mathbb{F}_2^t,\ \Delta_r \in \mathbb{F}_2^r}$ with probability p

$$p = 2^{-q}(C + 1)$$

p^* : probability if we assume $\delta_t \neq 0$ $\qquad p = \frac{2^t - 1}{2^t} p^* + 2^{-t}$

TD and ML

- ML : $[(u_s, 0), (v_q, 0)]_{u_s \in \mathbb{F}_2^s \setminus \{0\},\ v_q \in \mathbb{F}_2^q}$ with capacity C
- TD : $[(0, \delta_t), (0, \Delta_r)]_{\delta_t \in \mathbb{F}_2^t,\ \Delta_r \in \mathbb{F}_2^r}$ with probability p

$$p = 2^{-q}(C + 1)$$

p^* : probability if we assume $\delta_t \neq 0$ $\qquad p = \frac{2^t - 1}{2^t} p^* + 2^{-t}$

Zero Correlation and Impossible Differential [Blondeau Nyberg 13] :

- Zero Correlation : $C = 0$
- Impossible Differential : $p^* = 0$ and $p = 2^{-t}$

If $t = q$: Zero Correlation is mathematically equivalent to Impossible Differential

TD and ML

- ML : $[(u_s, 0), (v_q, 0)]_{u_s \in \mathbb{F}_2^s \setminus \{0\},\ v_q \in \mathbb{F}_2^q}$ with capacity C
- TD : $[(0, \delta_t), (0, \Delta_r)]_{\delta_t \in \mathbb{F}_2^t,\ \Delta_r \in \mathbb{F}_2^r}$ with probability p

$$p = 2^{-q}(C + 1)$$

p^* : probability if we assume $\delta_t \neq 0$ $\qquad p = \frac{2^t - 1}{2^t} p^* + 2^{-t}$

Zero Correlation and Impossible Differential [Blondeau Nyberg 13] :

- Zero Correlation : $C = 0$
- Impossible Differential : $p^* = 0$ and $p = 2^{-t}$

If $t = q$: Zero Correlation is mathematically equivalent to Impossible Differential

CP versus KP?

Data Complexity of a Distinguishing Attack

$P_S = 50\%$ and $\varphi_a = \Phi^{-1}(1 - 2^{-a})$, with a the advantage

- Multidimensional Linear :

$$N^{ML} = \frac{2^{(s+q+1)/2}}{C} \cdot \varphi_a$$

- Truncated Differential :

$$N^{TD} = \frac{2^{-q+1}}{S \cdot (p - 2^{-q})^2} \cdot \varphi_a^2,$$

where S is the size of a structure

For $p = 2^{-q}(C + 1)$,

$$N^{TD} = \frac{2^{q+1}}{S \cdot C^2} \cdot {\varphi_a}^2$$

Data Complexity of a Distinguishing Attack

$P_S = 50\%$ and $\varphi_a = \Phi^{-1}(1 - 2^{-a})$, with a the advantage

- Multidimensional Linear :

$$N^{ML} = \frac{2^{(s+q+1)/2}}{C} \cdot \varphi_a$$

- Truncated Differential :

$$N^{TD} = \frac{2^{-q+1}}{S \cdot (p - 2^{-q})^2} \cdot \varphi_a^2,$$

where S is the size of a structure

For $p = 2^{-q}(C + 1)$,

$$N^{TD} = \frac{2^{q+1}}{S \cdot C^2} \cdot \varphi_a{}^2$$

Data Complexity of a Distinguishing Attack

$P_S = 50\%$ and $\varphi_a = \Phi^{-1}(1 - 2^{-a})$, with a the advantage

- Multidimensional Linear :

$$N^{ML} = \frac{2^{(s+q+1)/2}}{C} \cdot \varphi_a$$

- Truncated Differential :

$$N^{TD} = \frac{2^{-q+1}}{S \cdot (p - 2^{-q})^2} \cdot \varphi_a^2,$$

where S is the size of a structure

For $p = 2^{-q}(C + 1)$, $N^{TD} = \frac{(N^{ML})^2}{2^s \cdot S}$,

and if $S = 2^t$, $N^{TD} = 2^{-n} \cdot (N^{ML})^2$

On the SS attack on PRESENT [Collard Standaert 09]

- It is observed and then heuristically assumed that the capacity C decreases linearly with the number of rounds
- This estimatation of the capacity has been verified in [Leander 11] (divided approximatively by 2^3 at each round of the cipher)

On the SS attack on PRESENT [Collard Standaert 09]

- It is observed and then heuristically assumed that the capacity C decreases linearly with the number of rounds
- This estimatation of the capacity has been verified in [Leander 11] (divided approximatively by 2^3 at each round of the cipher)
- As in ML, it is assumed that the data complexity is proportional to $\frac{2^{(q+1)/2}}{C}$
- In [Kerckhof et al 11], the attack is verified experimentally
  - Theory and practice match for a small number of rounds
  - A sensible difference is noticed for the attack on 18 rounds of PRESENT

On the SS Attack on PRESENT

From the strong link between SS and TD cryptanalysis, we can show that :

- The data complexity estimate is correct when only one structure is used (small number of rounds)

$$N = \frac{2^{(q+1)/2}}{C}\varphi_a$$

- For more rounds (more structures) the data complexity is equal to

$$N = \frac{2^{q+1}}{S \cdot C^2}\varphi_a^2$$

- Using the SS model of [Collard Standaert 09] one can only perform an attack on 23 rounds (instead of on the 24 rounds originally claimed)

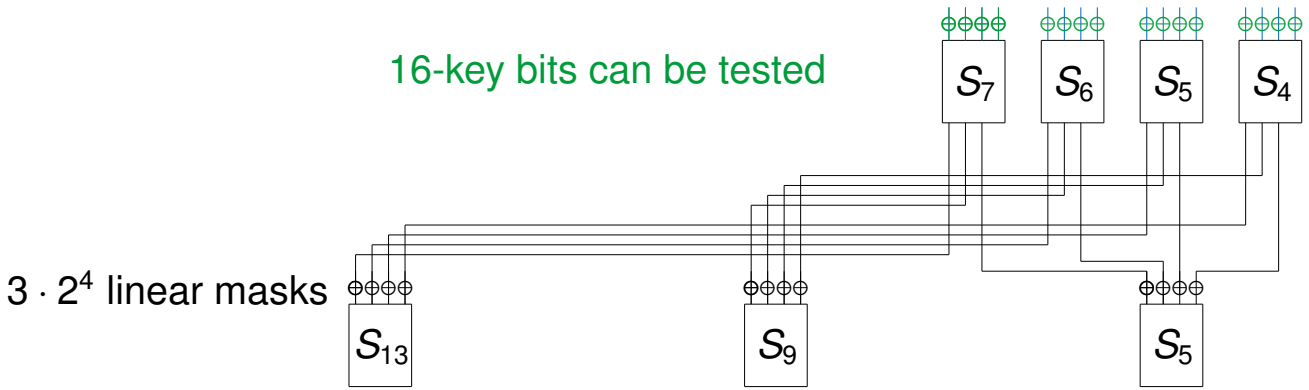
Outline

KP ML and CP TD attack : An Example on PRESENT

[Cho 09] :

- Distinguisher on 24 rounds
- KP ML attack on 26 rounds (inversion of the first and last round)

First round :



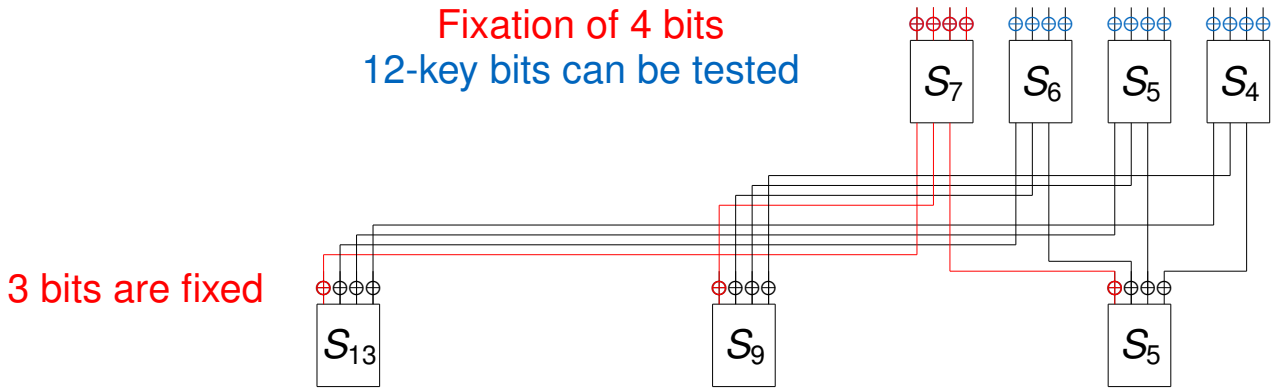
- KP ML $\Rightarrow$ Guess 16 bits

KP ML and CP TD attack : An Example on PRESENT

[Cho 09] :

- Distinguisher on 24 rounds
- KP ML attack on 26 rounds (inversion of the first and last round)

First round :



- KP ML $\Rightarrow$ Guess 16 bits

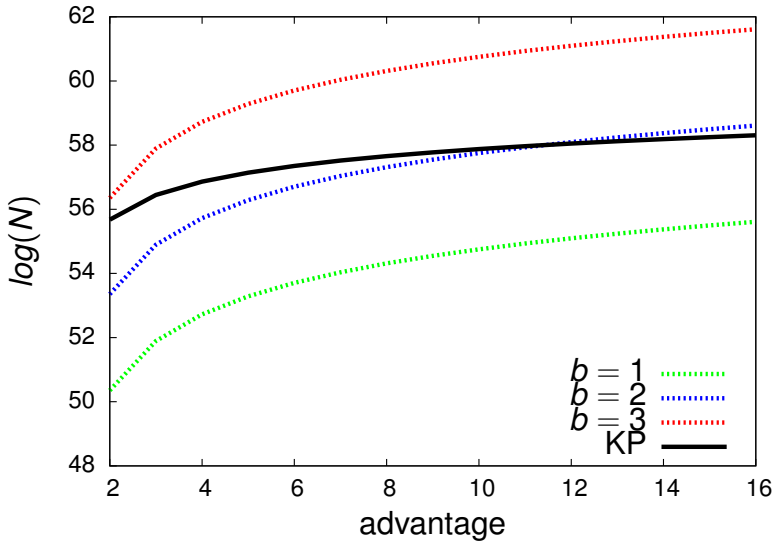
Using the link between TD and ML

- CP TD $\Rightarrow$ Guess $4, 8, 12, 16$ bits

Example of CP TD attack on 24 rounds of PRESENT

- Fixing $4b$ bits in the first round

Data Complexity :



- Depending of the size of the fixation, the data complexity of a CP ML attack can be smaller than for a KP ML attack

Example of CP TD attack on 24 rounds of PRESENT

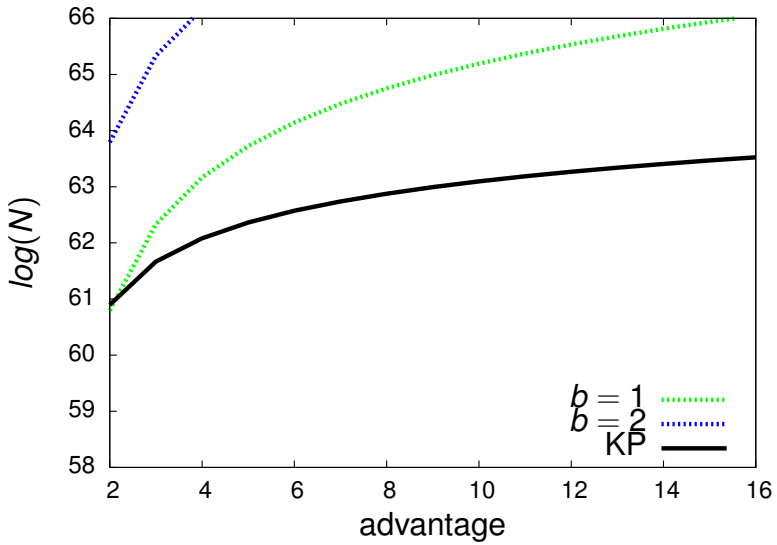
Fixing 4 bits :

Model	a	Data	Memory	Time_1	Time_2
CP	10	$2^{54.75}$	2^{29}	$2^{54.75}$	2^{70}
KP	5	$2^{57.14}$	2^{32}	$2^{57.14}$	2^{75}

Time_1: Complexity of the distillation phase
Time_2: Complexity of the search phase

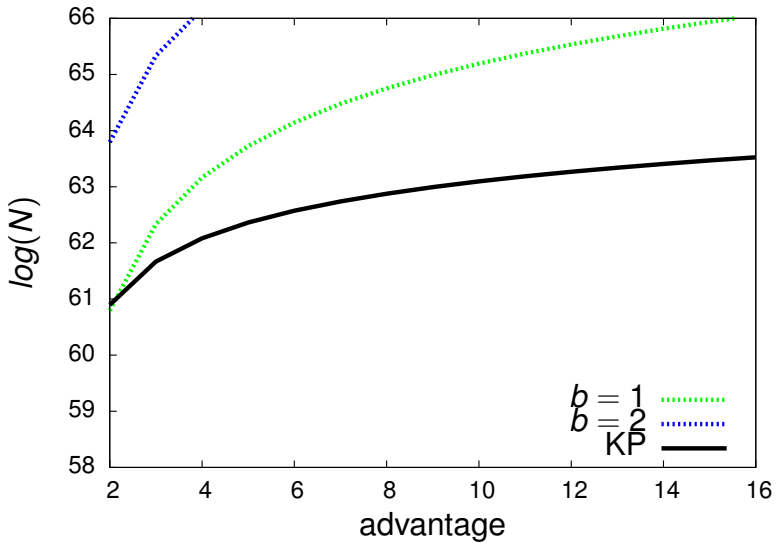
- Time and memory complexities of a CP TD attack can also be smaller than for a KP ML attack

Example of CP TD attack on 26 rounds of PRESENT



Model	a	Data	Memory	Time_1	Time_2
CP	4	$2^{63.16}$	2^{29}	$2^{63.16}$	2^{76}
KP	4	$2^{62.08}$	2^{32}	$2^{62.08}$	2^{76}

Example of CP TD attack on 26 rounds of PRESENT



Model	a	Data	Memory	Time_1	Time_2
CP	4	$2^{63.16}$	2^{29}	$2^{63.16}$	2^{76}
KP	4	$2^{62.08}$	2^{32}	$2^{62.08}$	2^{76}

- A CP TD attack on 26 rounds of PRESENT
- The previous differential-type attack was on 19 rounds

Remarks and Conclusions

- Every KP attack can be converted to a CP attack with same complexity
- The previous example illustrate that a KP ML attack can, up to some restrictions, be converted to a CP TD attack with smaller complexity
- We can have similar reasoning for other statistical attacks :

  For instance in [Bogdanov et al 12], a zero-correlation distinguisher is converted to a CP integral attack