

Analysis of the Algebraic Side-Channel Attacks

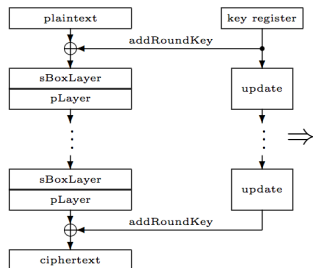
Jean-Charles Faugère Christopher Goyet Guénaél Renault
équipe SALSA, CNRS/INRIA/LIP6/UPMC, THALES

THALES



UPMC
PARIS UNIVERSITÉS

Algebraic cryptanalysis



$$\left\{ \begin{array}{l} x_1x_2 + x_1k_2 + x_1 + x_2k_1 + x_3 + x_4s_4 + \\ s_1s_4 + s_3s_4 + s_3 + s_4k_4 + s_4 + k_1k_2 + k_1 + k_3, \\ x_4 + s_1s_3 + s_2 + s_4 + k_4 + 1, \\ \vdots \\ s_{21} + s_{52}y_{124} + s_3y_{124} + y_{121}y_{124} + y_{121} + \\ y_{123}y_{124} + y_{124}k_{122} + y_{124}k_{123} + y_{124} + k_{121} \end{array} \right.$$



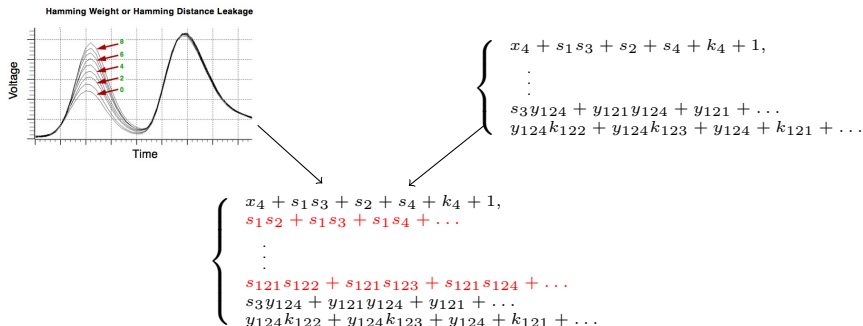
Solving



find the secret key

Algebraic Side-Channel Attacks (ASCA)

New kind of attacks recently by Renault, Standaert and Veyrat-Charvillon (CHES 2009, Inscrypt2009) mixing **Power Analysis** and **algebraic cryptanalysis**



main idea of ASCA

- 1 Online Phase: physical leakages measures
- 2 Offline Phase: algebraic attack
 - modeling cipher and additional information by a system of equations
 - solving this system



Algebraic Side-Channel Attacks

Renauld, Standaert, Inscrypt 2009



Algebraic Side-Channel Attacks on the AES: Why Time also Matters in DPA

Renauld, Standaert, Veyrat-Charvillon, CHES 2009



Algebraic Methods in Side-Channel Collision Attacks and Practical Collision Detection

Bogdanov, Kizhvatov, Pyshkin, Indocrypt 2008



Blind Differential Cryptanalysis for Enhanced Power Attacks

Handschuh, Preneel, Selected Areas in Cryptography 2006



Multi-Linear cryptanalysis in Power Analysis Attacks

Roche, Tavernier, 2009



...

Algebraic Side-Channel Attacks

Interesting aspects

- require much less observations than a DPA
- solving step seems very **fast** (with a SAT-solver)
- can deal with masking countermeasure

Algebraic Side-Channel Attacks

Interesting aspects

- require much less observations than a DPA
- solving step seems very **fast** (with a SAT-solver)
- can deal with masking countermeasure

However, the effectiveness depends on

- the device used and the quality of the trace
- the leakage model
- the amount of available information
- the shape of the system of equations (cipher modeling)
- the **heuristics** used in the **SAT-solver**
- ...

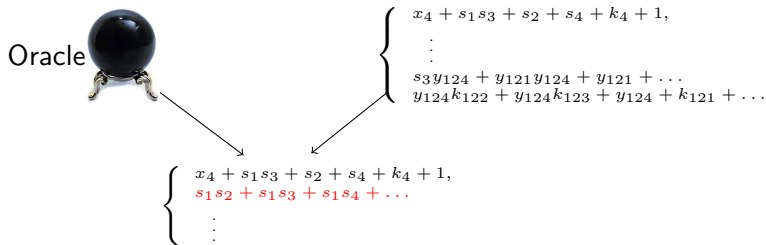
↪ very difficult to explain and predict results of experiments

Main goal: analysis of algebraic phase

in order to explain the effectiveness of the solving step

Main goal: analysis of algebraic phase

in order to explain the effectiveness of the solving step



Our analysis of algebraic phase

- impact of the oracle model?
- how many oracle queries are needed?
- some queries more valuable than others?
- which cipher intermediate operations to target?

So, we need a more stable and predictable solving method than Sat-solver
without heuristics \implies Gröbner basis

Main goal: analysis of algebraic phase

Oracle model:

- Oracle gives 8-bits Hamming weights of output layers
- assumed error-free

PRESENT	PRESENT+Oracle
Sat-Solver = ∞ ❌	Sat-Solver \simeq 1s ✓ (CHES 2009)
Gröbner basis = ∞ ❌	Gröbner basis (F4) \simeq 20min ✓ (our work)

∞ : more than one day of computation

Sat-Solver = Heuristics \Rightarrow ~~analysis~~

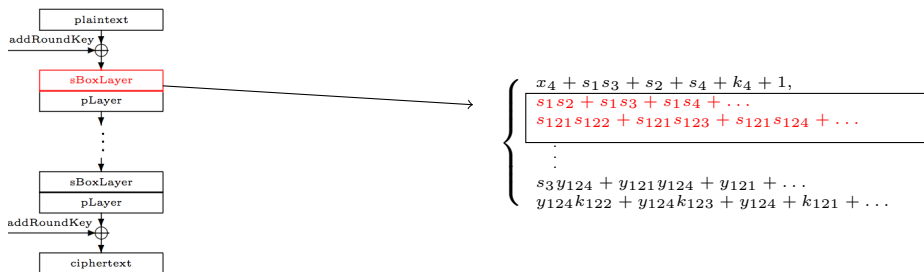
Gröbner basis = Algebraic resolution \Rightarrow theoretical analysis

Global to local study

Global to local study

- S-boxes are the only nonlinear part of many block ciphers
- They give the resistance against algebraic attacks

Main criterion to evaluate the algebraic resistance of a block cipher is the **Algebraic Immunity** of the S-boxes



⇒ We start to study the S-boxes

Algebraic Immunity (Carlet, Courtois, ...)

Main criterion for algebraic attack = **Algebraic Immunity**

Notations

- Let $S : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ be a n -bits S-box.
- X_1, \dots, X_n and Y_1, \dots, Y_n respectively its input and output bits.
- $F_i(X_1, \dots, X_n, Y_1, \dots, Y_n)$, $1 \leq i \leq n$ are the functions defining S

Definition of Algebraic Immunity (Ars, Courtois, Carlet, ...)

Let $I_S = \langle \{F_i(X_1, \dots, X_n, Y_1, \dots, Y_n), X_i^2 - X_i, Y_i^2 - Y_i, i \in \{1 \dots n\}\} \rangle$.
Then the **Algebraic Immunity** of S is defined by

$$AI(S) = \min\{\deg(P), P \in I_S \setminus \{0\}\}$$

The **number** of such lowest degree relations is also an important invariant

Algebraic Immunity (Carlet, Courtois, ...)

How to compute the **Algebraic Immunity** for a given S-box S ?

It is enough to compute a Gröbner basis with the **DRL order** of

$$I_S = \langle \{F_i(X_1, \dots, X_n, Y_1, \dots, Y_n), X_i^2 - X_i, Y_i^2 - Y_i, i \in \{1 \dots n\}\} \rangle$$

Indeed, we have

Prop

The reduced Gröbner basis G_S of I_S with respect to a graded order contains a linear basis of the lowest relations of S (i.e. the polynomials $P \in I_S$ such that $\deg(P) = AI(S)$).

Example with the AES S-box

The Algebraic Immunity of the inverse function over \mathbb{F}_{2^8} (e.g. AES S-box) equals **2**. Indeed, the inverse function is represented by a set of 39 quadratics equations over \mathbb{F}_2 (Courtois 2002)

A new notion of Algebraic Immunity

ASCA context \Rightarrow consider **leakage information**

Notations

For every value ℓ of the leakage model, we denote

- $E_\ell(X_1, \dots, X_n, Y_1, \dots, Y_n)$ the equations representing the leakage information ℓ
- $I_\ell = \langle E_\ell(X_1, \dots, X_n, Y_1, \dots, Y_n) \cup \{F_i(X_1, \dots, X_n, Y_1, \dots, Y_n), X_i^2 - X_i, Y_i^2 - Y_i, i \in \{1 \dots n\}\} \rangle$

Definition of Algebraic Immunity with Leakage

The lowest degree relations in I_ℓ are called **Algebraic Immunity With Leakage** ℓ of the S-box S . It is denoted by $AI_L(S, \ell)$ and the number of such relations is denoted by $\#AI_L(S, \ell)$.

Algebraic Immunity with Leakage: HW example

Assumption : leakage L of S gives

- HW of input value
- HW of output value
- $\ell = (w_{in}, w_{out})$

\Rightarrow the ideal I_ℓ contains at least 2 independent **linear polynomials**:

$$X_1 + \cdots + X_n + (w_{in} \bmod 2) \in I_\ell$$

$$Y_1 + \cdots + Y_n + (w_{out} \bmod 2) \in I_\ell$$

Results

\forall S-box S , and $\forall \ell \in \{0, \dots, n\}^2$

$$AI_L(S, \ell) = 1$$

$$\#AI_L(S, \ell) \geq 2$$

Are these two linear polynomials **linearized** our S-Box?

HW example ($\ell = (w_{in}, w_{out})$)

\Rightarrow the ideal I_ℓ contains at least these 2 independent **linear polynomials**:

$$X_1 + \cdots + X_n + (w_{in} \bmod 2) \in I_\ell$$

$$Y_1 + \cdots + Y_n + (w_{out} \bmod 2) \in I_\ell$$

Does not help enough for solving our system:

- no linear relation between input and output
- substitution layer is always **nonlinear**

But now, we know that leakages may give rise to linear equations!!
Is there any other more interesting?

HW example ($\ell = (w_{in}, w_{out})$)

Trivial example: $w_{in} = 0$

\forall S-box S , if $w_{in} = 0$ then $X_1 = X_2 = \dots = X_n = 0$
and the Y_i are given by

$$Y_1, \dots, Y_n = S(0, \dots, 0) = y_1, \dots, y_n$$

$\#AI_L(S, \ell) = 2n$ is **maximal** with this case and
the corresponding S-box is **completely described** by linear relations

HW example ($\ell = (w_{in}, w_{out})$)

Trivial example: $w_{in} = 0$

\forall S-box S , if $w_{in} = 0$ then $X_1 = X_2 = \dots = X_n = 0$
and the Y_i are given by

$$Y_1, \dots, Y_n = S(0, \dots, 0) = y_1, \dots, y_n$$

$\#AI_L(S, \ell) = 2n$ is **maximal** with this case and
the corresponding S-box is **completely described** by linear relations

PRESENT S-box example: $\#AI_L(S, (w_{in}, w_{out}))$

$w_{in} \backslash w_{out}$	0	1	2	3	4	5	6	7	8
0					16				
1					9				
2			15	15	8	13	15		
3			9	5	9	5	9		
4	16	15	14	2	11	3	12	13	16
5		13	13	2	7	10	11	13	
6		15	12	15	7	15	14		
7			13		13				
8			16						

A lot of
interesting linear
equations can
appear, depending
on the leakage
value

Another invariant

Definition

\forall S-box S , \forall leakage value ℓ
we define

$$\begin{aligned} N_S(\ell) &= \#\{x \in \mathbb{F}_2^n \text{ s.t. leakage of } S(x) = \ell\} \\ &= \#V(I_\ell) \end{aligned}$$

Another invariant

Definition

\forall S-box S , \forall leakage value ℓ
we define

$$\begin{aligned} N_S(\ell) &= \#\{x \in \mathbb{F}_2^n \text{ s.t. leakage of } S(x) = \ell\} \\ &= \#V(I_\ell) \end{aligned}$$

Prop

Let n the bus size of S . If $AI_L(S, \ell) = 1$ and $N_S(\ell)$ is non-zero then

$$\#AI_L(S, \ell) \geq 2n + 1 - N_S(\ell)$$

$N_S(\ell)$ small \rightsquigarrow a lot of linear relations between input and output

Take a look at PRESENT S-box

Assumptions : 8-bits bus and **Hamming weight** leakage model

$w_{in} \backslash w_{out}$	0	1	2	3	4	5	6	7	8
0					16				
1					9				
2			15	15	8	13	15		
3			9	5	9	5	9		
4	16	15	14	2	11	3	12	13	16
5		13	13	2	7	10	11	13	
6		15	12	15	7	15	14		
7			13		13				
8			16						

Figure: $\#AI_L(S, w_{in}, w_{out})$

$w_{in} \backslash w_{out}$	0	1	2	3	4	5	6	7	8
0					1				
1					8				
2			2	2	18	4	2		
3			8	12	8	20	8		
4	1	2	3	24	7	22	6	4	1
5		4	4	16	12	8	8	4	
6		2	6	2	12	2	4		
7			4		4				
8			1						

Figure: $N_S(w_{in}, w_{out})$

Observations

- confirm that small $N_S \Rightarrow$ large $\#AI_S$
- Most of leakages give a lot of linear relations:
 $\mathbb{E}(\#AI_L) = 7, 9$
- We are now able to sort leakages by relevance

Take a look at PRESENT S-box

Assumptions : 8-bits bus and **Hamming weight** leakage model

$w_{in} \backslash w_{out}$	0	1	2	3	4	5	6	7	8
0					16				
1					9				
2			15	15	8	13	15		
3			9	5	9	5	9		
4	16	15	14	2	11	3	12	13	16
5		13	13	2	7	10	11	13	
6		15	12	15	7	15	14		
7			13		13				
8			16						

Figure: $\#AI_L(S, w_{in}, w_{out})$

$w_{in} \backslash w_{out}$	0	1	2	3	4	5	6	7	8
0					1				
1					8				
2			2	2	18	4	2		
3			8	12	8	20	8		
4	1	2	3	24	7	22	6	4	1
5		4	4	16	12	8	8	4	
6		2	6	2	12	2	4		
7			4		4				
8			1						

Observations

- confirm that small $N_S \Rightarrow$ large $\#AI_S$
- Most of leakages give a lot of linear relations:
 $\mathbb{E}(\#AI_L) = 7, 9$
- We are now able to sort leakages by relevance

Take a look at PRESENT S-box

Assumptions : 8-bits bus and **Hamming DISTANCE** leakage model

Definition:

$$d = HD(x, S(x)) = HW(x \oplus S(x))$$

HD model :

- $AI_L(d) = 1$
- $\#AI_L(d) \geq 1$
- $\mathbb{E}(\#AI_L) = 2, 3$

d	0	1	2	3	4	5	6	7	8
$N_S(d)$	0	0	16	56	81	64	30	8	1
$\#AI_L(S, d)$	0	0	10	3	1	1	1	9	16

Figure: HD model and PRESENT S-Box

Much less than in HW model

↪ predict that solving will be much more difficult in this case

Global Study

Solving strategy

- triangular structure \rightarrow blocks of equations (Layers, SBoxes, ...)
- blocks corresponding to Sboxes \rightarrow Gröbner basis of I_ℓ
- polynomial system modeling PRESENT partly linearized

Results:

Successive Gröbner basis computation (F4)

\rightarrow better control on the degree

\rightarrow better solving strategy

Criterion of success

Attack with following assumptions is explained:

- a very simple SPN block cipher : PRESENT
- Oracle gives **8-bits Hamming weights** of output layers
- assumed error-free

Because of:

- $AI_L = 1$
- $\mathbb{E}(\#AI_L) = 7,9$
- $\mathbb{P}(\#AI_L \geq 8) \approx \frac{1}{2}$

⇒ Expected linear relations for one substitution layer ≈ 64

Criterion of success

Attack with following assumptions is explained:

- a very simple SPN block cipher : PRESENT
- Oracle gives **8-bits Hamming weights** of output layers
- assumed error-free

Because of:

- $AI_L = 1$
- $\mathbb{E}(\#AI_L) = 7,9$
- $\mathbb{P}(\#AI_L \geq 8) \approx \frac{1}{2}$

⇒ Expected linear relations for one substitution layer ≈ 64

Why this attack still work with weaker ASCA assumptions?

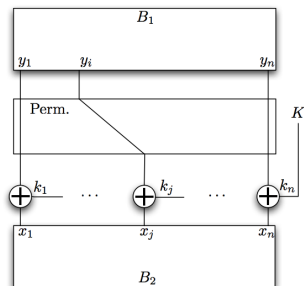
- with leakages in **only 3 or 4 rounds?**
- in **unknown plaintext/ciphertext** scenario?

Few consecutive leakages or unknown P/C

Going back to the local study:

$N_S(\ell)$ small \Rightarrow a lot of linear relations

$N_S(\ell)$ very small (≤ 6) \Rightarrow fixed input/output bits!!



\rightsquigarrow subkey bits easily deduced

Experiments - Conclusion

Experiments

Experiments performed against PRESENT and AES

Analysis supported by experiments:

- | | GB |
|---------------------------------------|----|
| • reject of leakages with large N_S | ✓ |
| • reject of leakages with small N_S | ✗ |
| • no consecutive leaked rounds | ✗ |
| • importance of the model: HD example | ✗ |

Experiments

Experiments performed against PRESENT and AES

Analysis supported by experiments:

	GB	SAT-solver
• reject of leakages with large N_S	✓	✓
• reject of leakages with small N_S	✗	✗
• no consecutive leaked rounds	✗	✗
• importance of the model: HD example	✗	✗

Analysis is valid with both Gröbner basis **and** SAT-solver

Conclusion

- New notion of Algebraic Immunity
- Good understanding of influence of leakage information
 - ▶ Results of experiments are explained
 - ▶ Leakages informations can be sorted by importance

Perspectives

- Identify resistant S-boxes against ASCA and others cryptanalysis (current work with Claude Carlet)
- Study more realistic oracle models
- Dealing with errors