Evaluation of Countermeasures Hardware Implementation to Thwart Side Channel HO-DPA

Jean-Luc DANGER, Sylvain GUILLEY, Florent FLAMENT,
Houssem MAGHREBI
< jean-luc.danger@TELECOM-ParisTech.fr >

Institut TELECOM / TELECOM-ParisTech
CNRS – LTCI (UMR 5141)

CryptArchi-09, Prague — Friday June 26th, 2009.

Presentation Outline

Masking: principle [1, 2, 3, 6]

- Aims at making power consumption constant
- The internal variables are shared: $(m\ ,\ x_m = x\ \theta\ m)$
  x_m is the masked variable and θ is an inversible operation
  - Boolean masking is based on exclusive-or (xor) operations:

$$x_m = x \oplus m\,,$$

  - Arithmetic masking is made with modulus operation on a finite field:

$$x_m = x + m\ (mod\ n) \quad \text{or}$$
$$x_m = x * m\ (mod\ n)$$

- Theoretically provable against first-order attack [4]
- But many possible Side Channel Attacks on Hardware implementations [9, 5, 8]

Linear Function

- $f(x \oplus m) = f(x) \oplus f(m)$.
- $f(x)$ is rebuild from $f(x \oplus m)$ et $f(m)$.

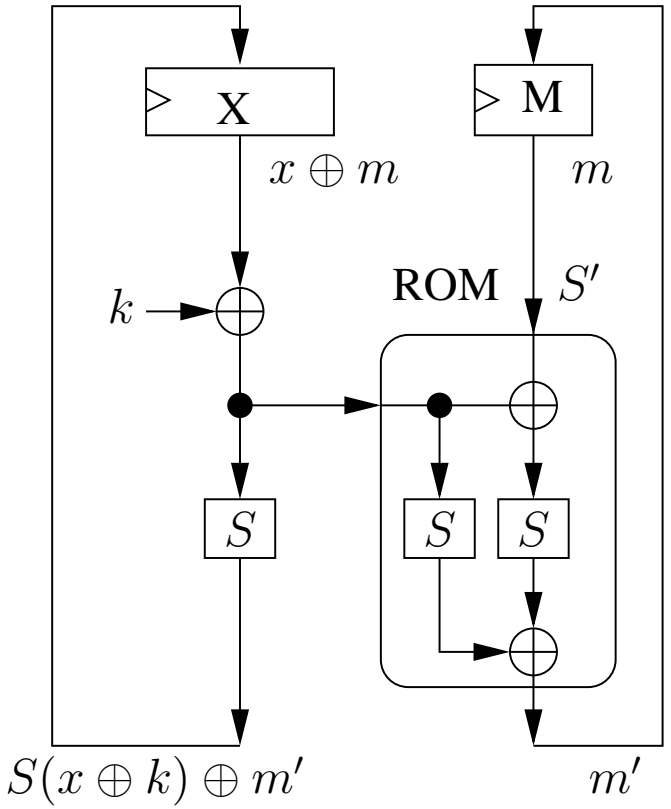
Non-Linear Function (NL)

- $S(x_m \oplus k) = S(x \oplus m \oplus k) = S(x \oplus k) \oplus m'$
- $m' = S'(x_m \oplus k, m)) = S'(x \oplus m \oplus k, m)$
  - $\Rightarrow$ demasking is necessary before the NL function
  - $\Rightarrow$ masking is necessary after the NL function
- The robustness of a masking implementation depends on the way the NL function is implemented.

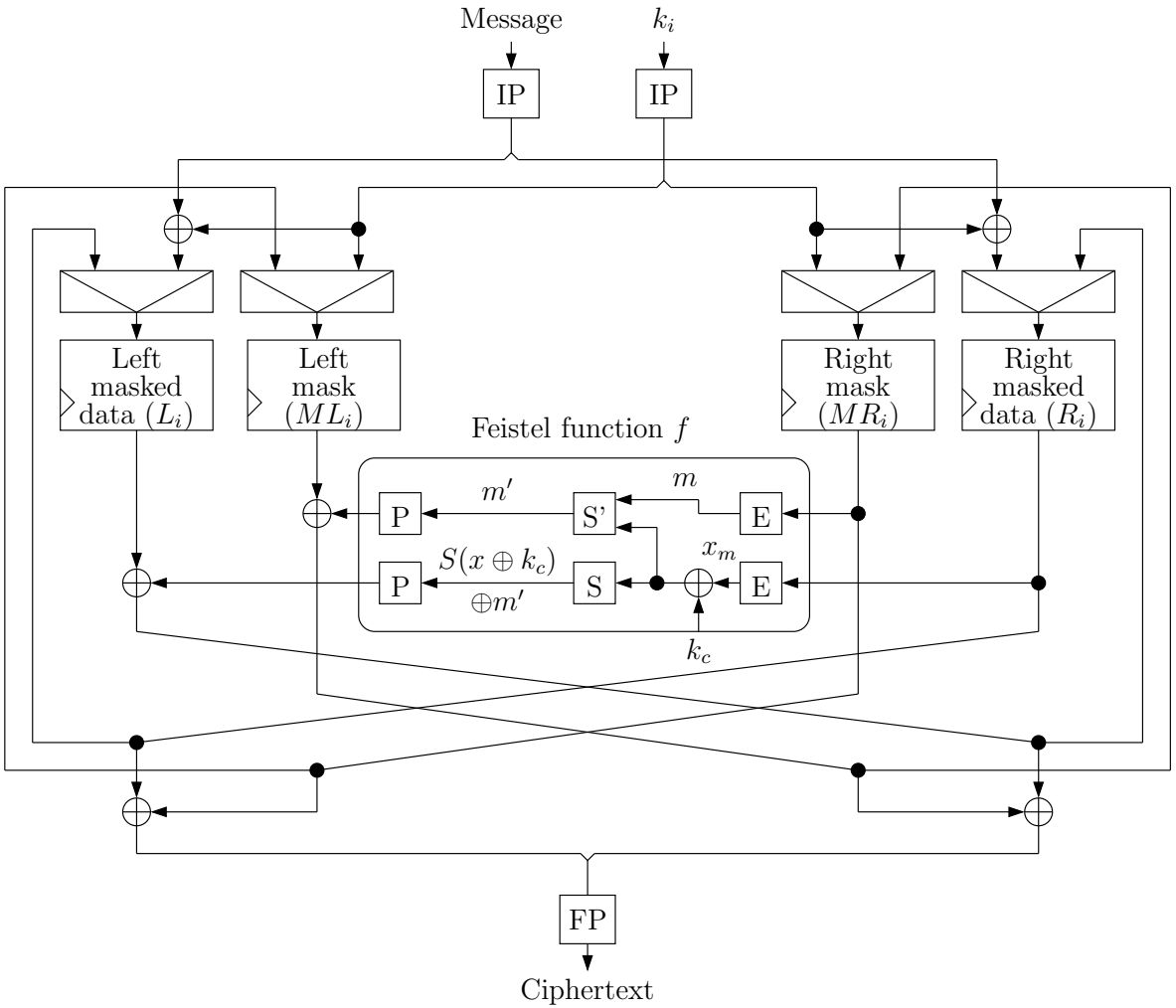
Presentation Outline

ROM masking

- Needs at least one sbox of 2^{2n} words
- Complexity is reduced if an algebric form is possible
  - in AES inverse in GF(2^8) $S(x) = x^{-1}$ in GF(2^8) $\Rightarrow\ S(x_m) = (x * m)^{-1} = S(x) * S(m)$
  - But zero-value attacks are possible then
  - Algebric expressions have to be refined [7]



Example: DES


Message
k_i
IP
IP
Left masked data (L_i)
Left mask (ML_i)
Right mask (MR_i)
Right masked data (R_i)
Feistel function f
m'
m
P
S'
E
$S(x \oplus k_c)$
x_m
P
S
E
$\oplus m'$
k_c
FP
Ciphertext

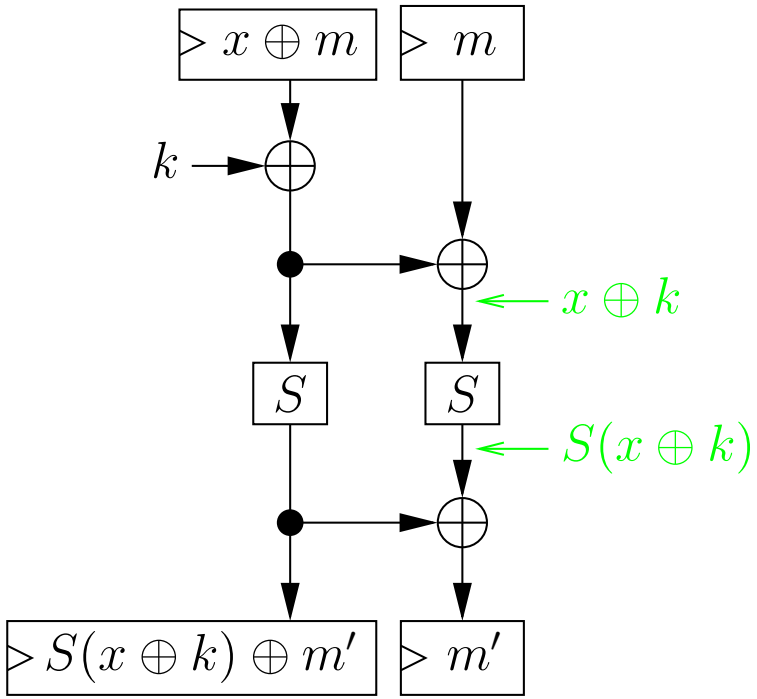
Masking by Universal Sbox Masking (USM)

Generic structure and Reduced complexity

- Only 2 sboxes of 2^n blocks

Drawbacks

- Demasked data at the sbox I/Os
- $\Rightarrow$ "Shallow" attack



Presentation Outline

DPA attack results

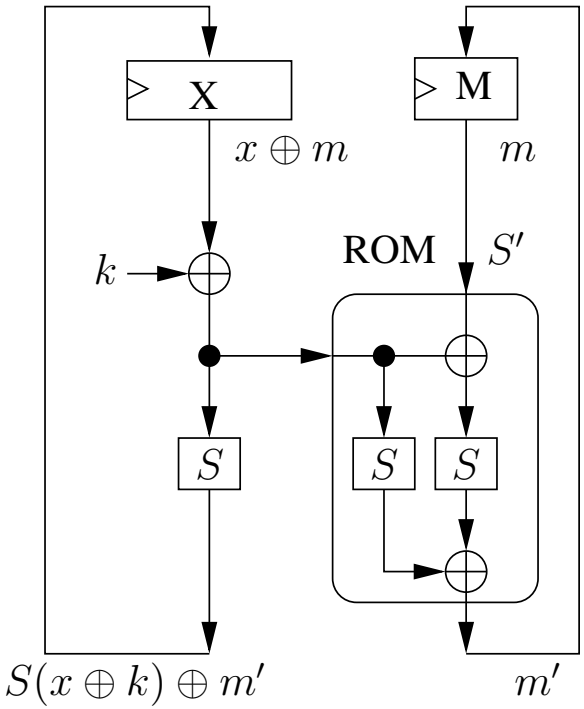
Module \ SBox #	S1	S2	S3	S4	S5	S6	S7	S8
Unprotected DES								
Measurements To Disclose	2974	2635	997	3317	965	2034	1803	1133
Maximal correlation in mV	0,82	1,12	1,23	0,95	1,98	1,5	1,34	1,69
SNR @ Disclosure	5,76	6,78	5,58	6,57	7,27	6,69	5,24	8,34
USM implementation								
Measurements To Disclose	20657	43513	11347	11779	16012	23517	94944	23998
Maximal correlation in mV	0,19	0,18	0,28	0,21	0,19	0,19	0,08	0,18
SNR @ Disclosure	5,40	5,41	5,38	5,21	7,85	4,77	2,93	5,82
ROM implementation								
Measurements To Disclose	–	–	–	–	–	–	–	–
Maximal correlation in mV	–	–	–	–	–	–	–	–
SNR @ Disclosure	–	–	–	–	–	–	–	–

First-order attack conclusions

- USM implementation is sensitive to first order attack
- ROM implementation is robust but
- Requires big size ROMs if no algebric expression is possible
- Is ROM implementation sensitive to HO-DPA ?

Presentation Outline

Zero-offset attack



From Waddle *et al.* [10], Peeters *et al.* [8].

Activity

$$A = HW[(x \oplus m) \oplus (S(x \oplus k) \oplus m')] + HW[m \oplus m']$$

- The register data Hamming distance is:

$$\Delta(x) = x \oplus S(x \oplus k)$$

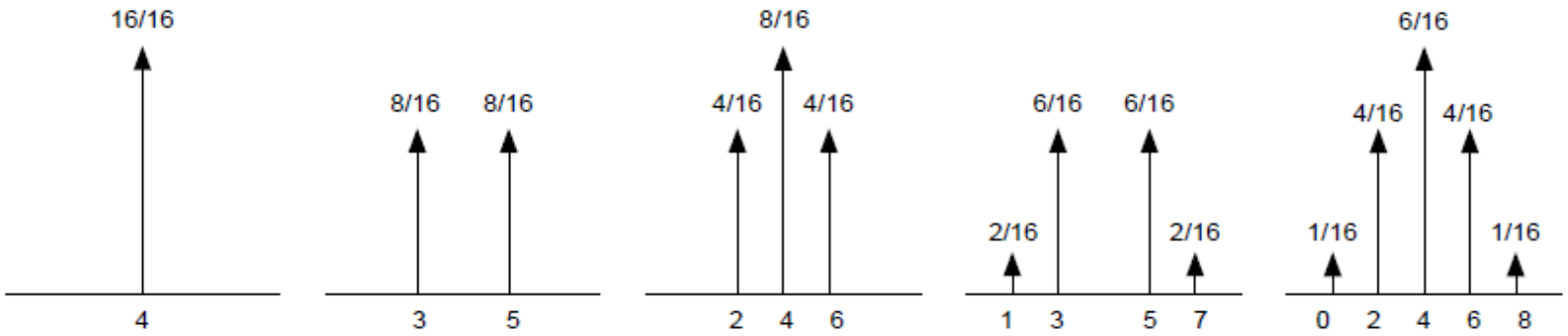
- The register mask Hamming distance is:

$$\Delta(m) = m \oplus m'$$

- Then:

$$A = HW[\Delta(x) \oplus \Delta(m)] + HW[\Delta(m)]$$

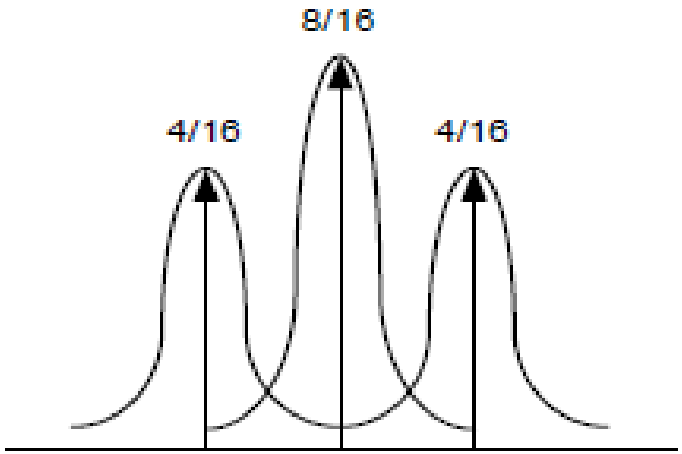
Power consumption distribution



- The probability density function $P[A|\Delta(x)]$
- Knowing $\Delta(x)$ we know the probability density function of masking an activity $P[A|\Delta(x)]$)

Noise effect

- The noise comes from others SBoxes and the environment
- It is assumed to be Gaussian



Attack principle

1. Apply n plaintext message (x_i, $i \in [1, n]$) and collect n observations of power consumption (traces A_i)
2. Make assumptions about the key k_j with $j \in [0, 63]$ and obtain for each key assumption the $\Delta(x)$ values:

$$\begin{cases} \Delta(k_0) = \Delta(x_0, k_0), \Delta(x_1, k_0), \cdots, \Delta(x_n, k_0) \\ \Delta(k_1) = \Delta(x_0, k_1), \Delta(x_1, k_1), \cdots, \Delta(x_n, k_1) \\ \qquad\qquad\qquad\qquad\qquad\qquad \cdots \\ \Delta(k_{63}) = \Delta(x_0, k_{63}), \Delta(x_1, k_{63}), \cdots, \Delta(x_n, k_{63}) \end{cases}$$

3. For each $\Delta(k_i)$ compute

$$P[A|\Delta(k_i)] = \prod_{j=0}^{n} P[A = A_j|\Delta(k_i, x_j)]$$

4. Apply the maximum likelihood approach: the correct key corresponds to the maximum probability $P[A|\Delta(k_i)]$

Optimized Attack for simulation

The simulated attack is performed as follow:

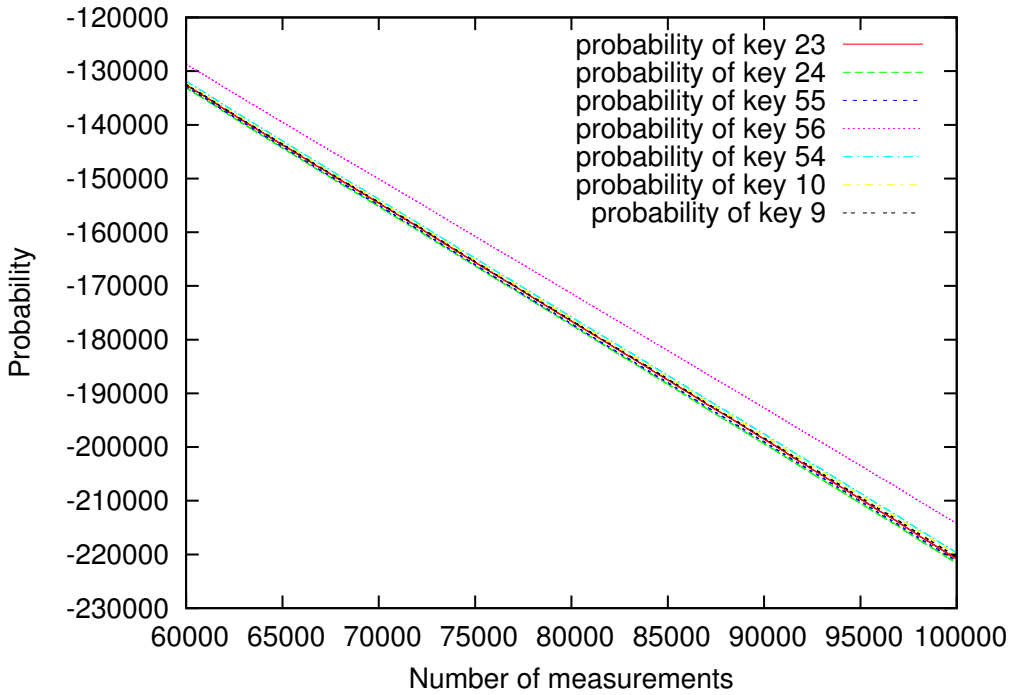
- From n plain text x_i calculate the observation vector which corresponds to the activity (A_i) with $i \in [1, n]$
- For each hypothesis of key k_i compute $P[A = A_i|\Delta(k_i, x_j)]$
- Since the product of the probability falls quickly to 0, use the logarithmic domain to have

$$P[A|\Delta(k_i)] = \sum_{j=0}^{n} (\ln P[A = A_j|\Delta(k_i, x_j)])$$

- The correct guess of the key is the argument of the maximum probability $P[A|\Delta(k_i)]$

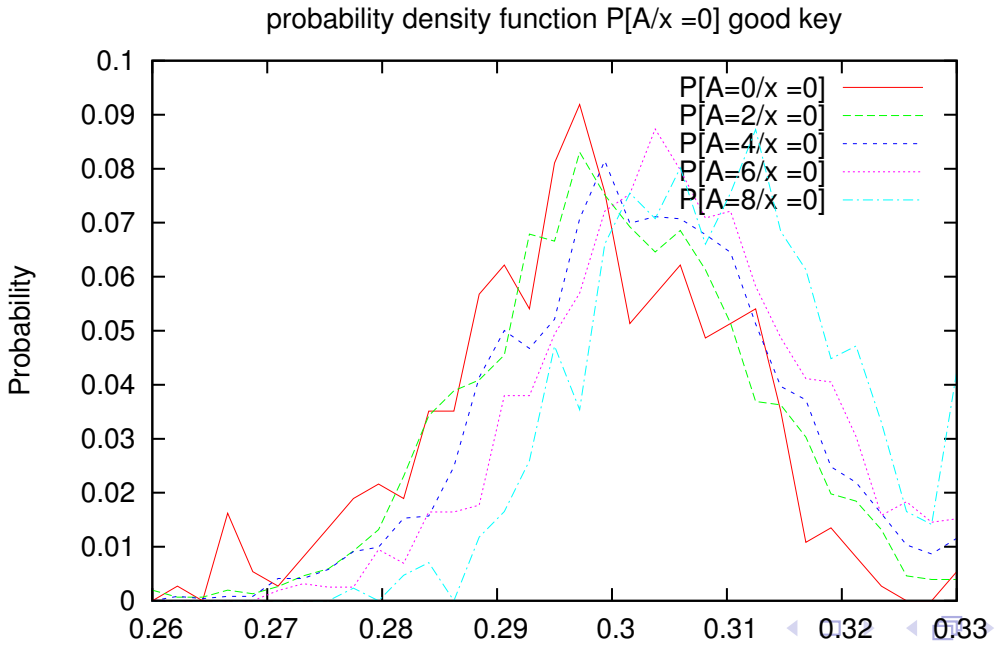
Attack on simulated traces

- Result attack on the first round of the DES ROM implementation



Real Attack Feasibility Analysis

- Target = StratixII SASEBO board
- Using a known mask and key we can calculate the real probability density function
- We sort the gaussians within the mask state in order to analysis the attack feasibility (100000 traces)



Feasibility Analysis

- Mean and variance of each gaussian

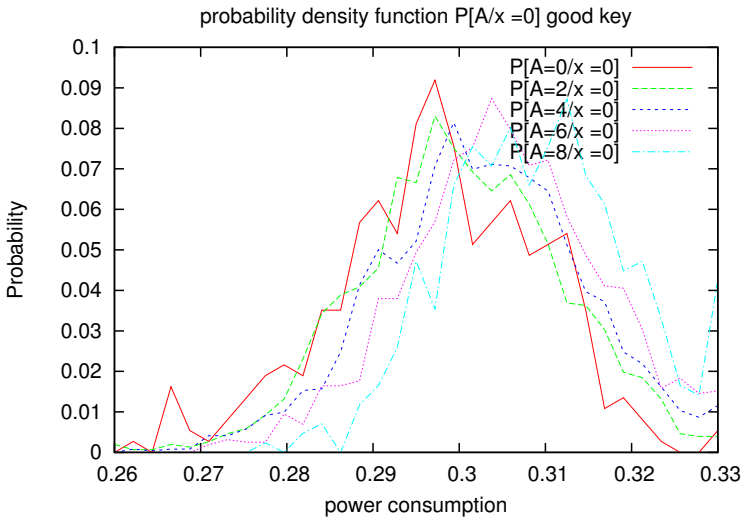
gaussian \ statistical characteristic	mean	variance
$\Delta(x) = 0$ **and** $\Delta(m) = 0$	0.298697656377	0.000142274310396
$\Delta(x) = 0$ **and** $\Delta(m) = 2$	0.301273303439	0.00013476465848
$\Delta(x) = 0$ **and** $\Delta(m) = 4$	0.303599821346	0.00013910645988
$\Delta(x) = 0$ **and** $\Delta(m) = 6$	0.306287380606	0.00012627529337
$\Delta(x) = 0$ **and** $\Delta(m) = 8$	0.310170751696	0.00011457955989

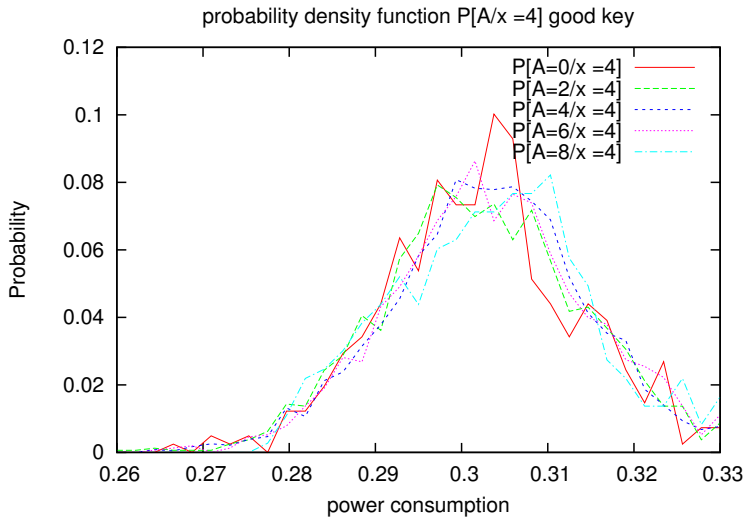
- The noise variance is very important.
- $\Rightarrow$ The ML attack needs an accurate profiling for every $\Delta(x)$.
- Can we take advantage of the mean move (hence the variance difference) between the $\Delta(x)$ pdfs ?

Real Attack by pdf analysis

Using the good key and knowing the mask:

- If $\Delta(x) = 0$: the gaussians have different mean;
- If $\Delta(x) = 4$: the gaussians have the same mean.

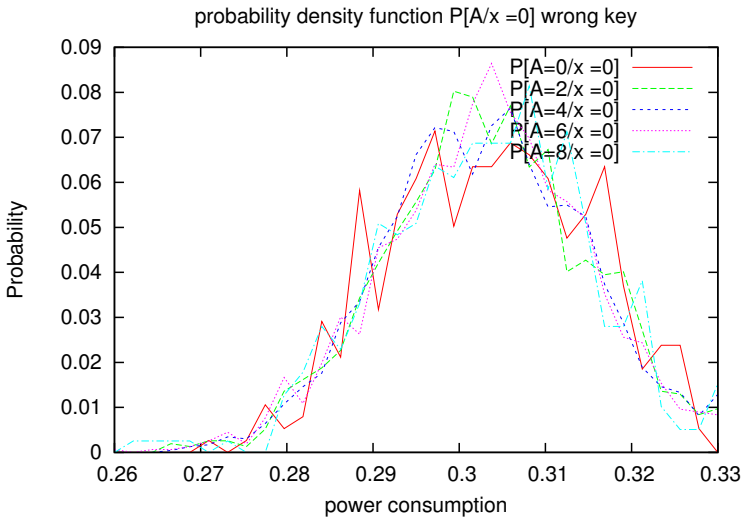


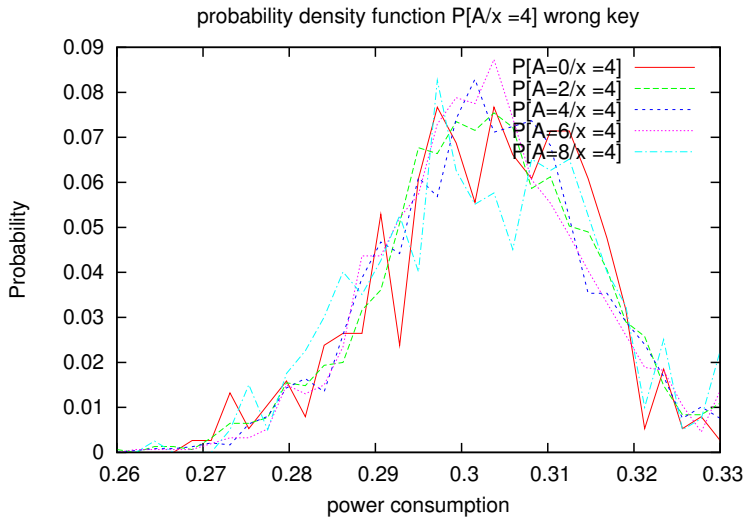


Real Attack by pdf analysis

Using the bad key and knowing the mask:

- If $\Delta(x) = 0$: the gaussians have the same mean;
- If $\Delta(x) = 4$: the gaussians have the same mean.





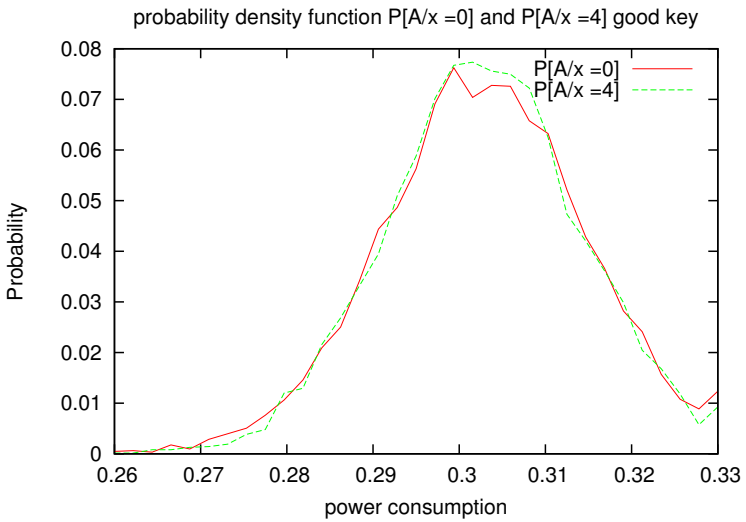
	$\Delta(x) = 0$	$\Delta(x) = 4$	**variance difference**
Good Key	0.000140736	0.000128611	0.00012125
Bad Key	0.000134715	0.000137386	-0.00002671

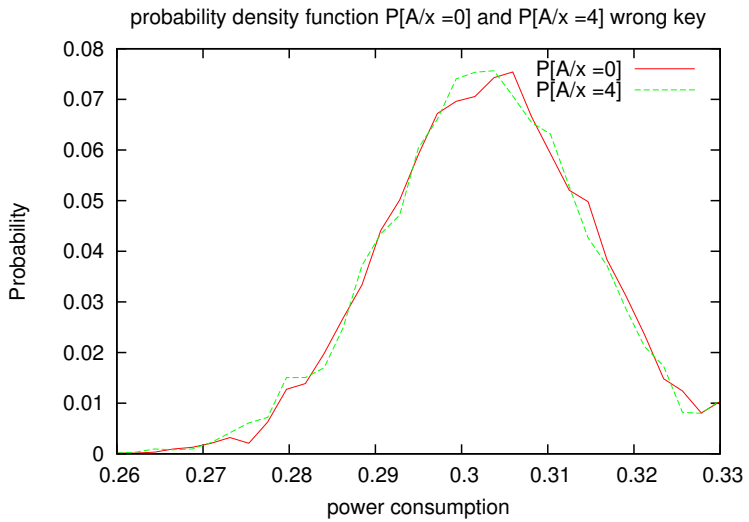
- A new kind of HO-DPA based on the variance difference

Attack example

we use two keys:

- Right key: 6b65796b65796b65
- wrong key: 014080014001a110





The difference of variance of the good key is important compared to the wrong one

Attack by pdf analysis

Algorithm:

- From n plain text x_i Calculate the observation vector which corresponds to the activity (A_i) with i in [1,n]
- For each hypothesis of key k_i
  - Sort the activity (A_i) within $\Delta(x_j, k_i)$ equal to 0 or 4
  - Compute the difference of the variance between $\Delta(x_j, k_i)= 0$ and $\Delta(x_j, k_i)= 4$
- The correct guess of the key corresponds to the maximum of the variance difference.

Presentation Outline

- ROM Masking implementation is robust against first-order DPA (theory verified)
- Real measurements show that HO-DPA is feasible with a reasonable amount of traces (100K)
- The attack is based on pdf analysis
- HO-DPA Attack possible improvements:
  - Preprocessing : Noise reduction (Kalman filters, EM algorithm,...)
  - Use of Principal components

References

[1] Mehdi-Laurent Akkar and Christophe Giraud.
An Implementation of DES and AES Secure against Some Attacks.
In LNCS, editor, *Proceedings of CHES'01*, volume 2162 of *LNCS*, pages 309–318. Springer, May 2001.
Paris, France.

[2] S. Chari, C. Jutla, J. Rao, and P. Rohatgi.
Towards Sound Approaches to Counteract Power-Analysis Attacks.
In *CRYPTO*, volume 1666 of *LNCS*, August 1999.
ISBN: 3-540-66347-9.

[3] Louis Goubin and Jacques Patarin.
DES and differential power analysis.
In *CHES*, LNCS, pages 158–172. Springer, Aug 1999.

[4] J.Blomer, J.Guajardo, and V. Krummel.
Provably Secure Masking of AES.
In LNCS, editor, *Proceedings of SAC'04*, volume 3357, pages 69–83. Springer, August 2004.
Waterloo, Canada.

[5] Stefan Mangard and Kai Schramm.
Pinpointing the Side-Channel Leakage of Masked AES Hardware Implementations.
In *CHES*, volume 4249 of *LNCS*, pages 76–90. Springer, 2006.
PDF.

[6] Thomas S. Messerges.
Securing the AES Finalists Against Power Analysis Attacks.
In *Fast Software Encryption'00*, pages 150–164. Springer-Verlag, April 2000.
New York.

[7] Elisabeth Oswald, Stefan Mangard, Norbert Pramstaller, and Vincent Rijmen.
A Side-Channel Analysis Resistant Description of the AES S-box.
In LNCS, editor, *Proceedings of FSE'05*, volume 3557 of *LNCS*, pages 413–423. Springer, February 2005.

Paris, France.

[8] Éric Peeters, François-Xavier Standaert, Nicolas Donckers, and Jean-Jacques Quisquater.
Improved Higher-Order Side-Channel Attacks With FPGA Experiments.
In *CHES*, volume 3659 of *LNCS*, pages 309–323. Springer-Verlag, 2005.

[9] Stefan, Elisabeth Oswald, and Thomas Popp.
Power Analysis Attacks: Revealing the Secrets of Smart Cards.
Springer, December 2006.
ISBN 0-387-30857-1, http://www.dpabook.org/.

[10] Jason Waddle and David Wagner.
Towards Efficient Second-Order Power Analysis.
In *CHES*, volume 3156 of *LNCS*, pages 1–15. Springer, 2004.
PDF.