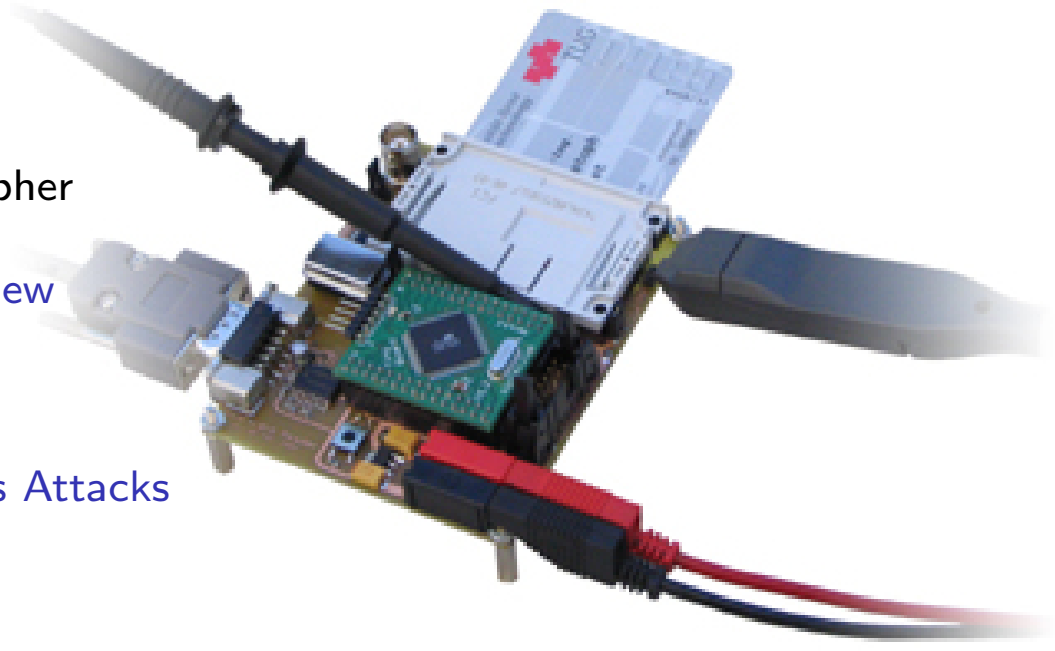
List decoding of RM(1,m) codes and Multi-linear Power Analysis attacks (MLPA)

Ilya Dumer, Rafael Fourquet, Grigory Kabatiansky,
Pierre Loidreau, Thomas Roche, Cédric Tavernier

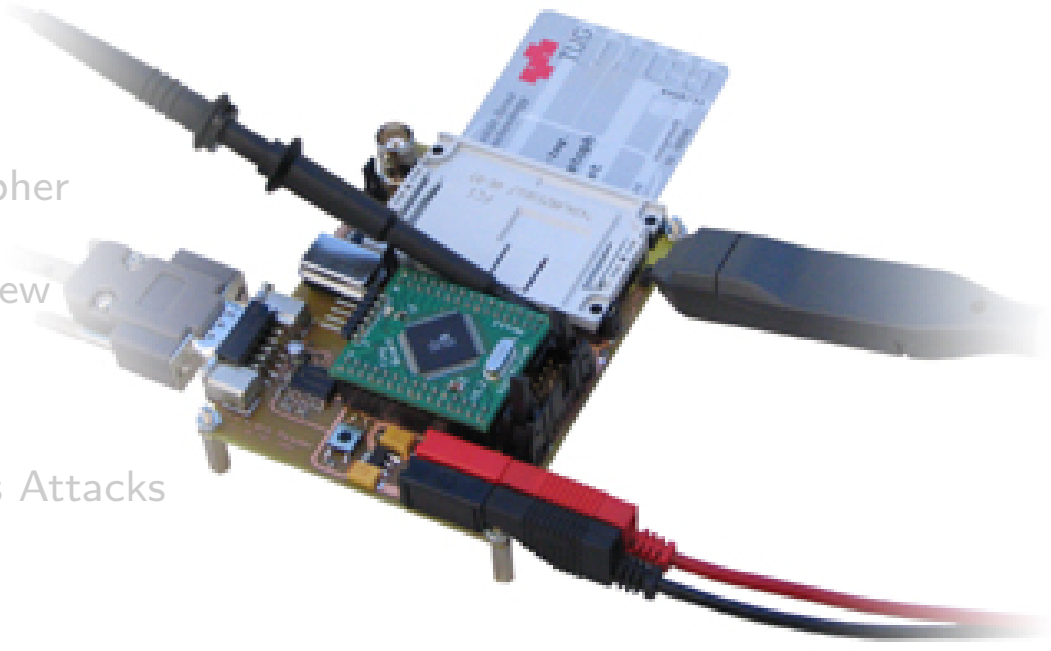
University of Riverside, CA, USA.
Research institute IPPI, Moscow, Russia.
University Paris VIII, Paris, France.
CELAR, Bruz
Laboratoire LIG, Grenoble, France.
Communications and Systems, Le Plessis Robinson, France.

29 mai 2009

Plan



Plan



Reed-Muller code properties

Definition of $RM(1, m)$

- $RM(1, m) = \{f \in GF(2)^{(1)}[x_1, x_2, \cdots, x_m]\}$;
- Usual representation : $(f(0), f(1), \cdots, f(2^m - 1))$;
- Boolean representation : $f = f_1 x_1 \oplus f_2 x_2 \oplus \cdots \oplus f_m x_m$
- code of lenght $n = 2^m$ and minimal distance $d = n/2$.

Classical Problem

Given a Boolean function g, we want to construct the list $\{f \in RM(1, m) \mid d_H(f, g) \leq n(1/2 - \epsilon)\}$, which is equivalent to $L_g(\epsilon) = \{f \in RM(1, m) \mid I^{(g)}(f) = \sum_{x \in GF(2)^m} (-1)^{f(x) \oplus g(x)} \geq 2\epsilon n\}$.

Johnson Bound

In fact $\|L_g(\epsilon)\| \leq \frac{1}{4\epsilon^2}$

List Decoding Algorithms

A simple idea

$2\epsilon n \leq |I^{(g)}(f)| \leq \sum_{s \in GF(2)^{m-i}} |\sum_{r \in GF(2)^i} (-1)^{g(r,s) \oplus f^{(i)}(r)}|$ where

$f^{(i)} = f_1 x_1 \oplus \cdots \oplus f_i x_i$.

Screnning process : we suggest f_i and we check if the inequality is satisfied.

$\Rightarrow L_g^{(i)}(\epsilon) = \{f \in RM(1,i) \mid \sum_s |\sum_{r \in GF(2)^i} (-1)^{g(r,s) \oplus f(r)}| \geq 2\epsilon n\}$.

In fact

$$M = \|L_g^{(i)}(\epsilon)\| \leq \frac{1}{4\epsilon^2}. \text{ With } E = L_g^{(i)}(\epsilon)$$

$$4n\epsilon^2 M \leq \sum_{a \in E} \sum_{b \in E} \sum_s |\sum_{r \in GF(2)^i} (-1)^{g(r,s) \oplus a^{(i)}(r) \oplus g(r,s) \oplus b^{(i)}(r)}| \leq n.$$

Complexity

Worst case complexity

The complexity of this algorithm is in $\mathcal{O}(n \log_2^2(\epsilon))$ [I Du 07].

The complexity of the prob. version is in $\mathcal{O}(m^2/\epsilon^6)$ [Kaba 04].

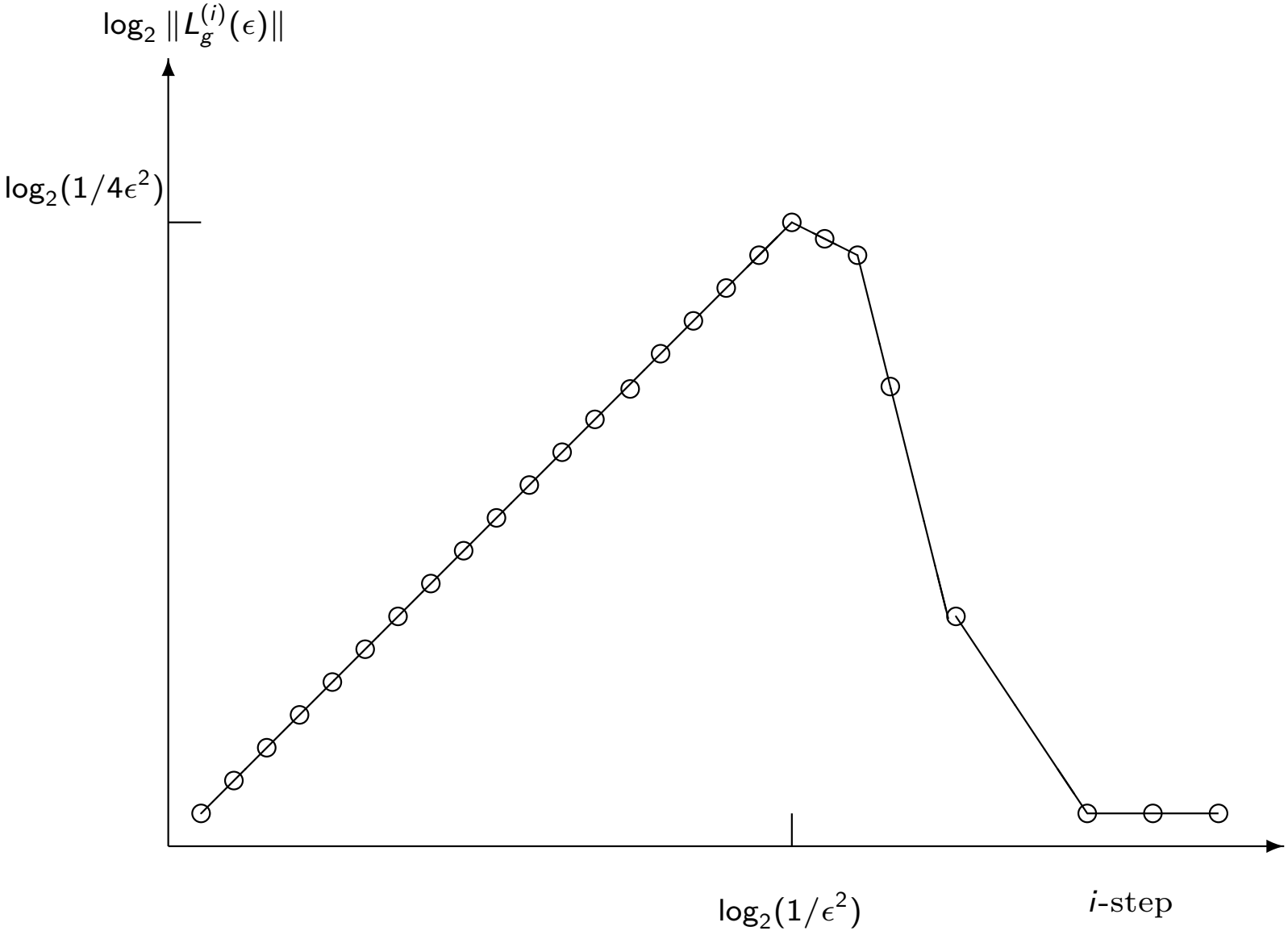
The size of the result can be of size $m/2\epsilon^2$, thus optimal complexity could be in $\mathcal{O}(m/\epsilon^2)$.

Optimal complexity

In fact Goldreich and Levin algorithm : $\mathcal{O}(m/\epsilon^4)$.

I. Dumer, G. Kabatiansky and C. Tavernier, not yet published : $\mathcal{O}(m/\epsilon^2)$.

Behaviour



Improvement and complexity

Idea

Computing the first steps by a FFT.

Complexity of the $\log_2(c/\epsilon^2)$ first steps is in $\mathcal{O}(\epsilon^{-4}\log_2(1/\epsilon^2))$

A complexity in $\mathcal{O}(m/\epsilon^2)$ should improve pratically the former algorithms.

Approximation of a bloc cipher

Analyzing a problem

A block cipher f can be seen as a vectorial function $f : GF(2)^l \times GF(2)^k \mapsto GF(2)^l$. For linear cryptanalysis, we have to find relation

$$\langle \alpha, X \rangle \oplus \langle \beta, f(X, K) \rangle \oplus \langle \mu, K \rangle = 0$$

that hold with the highest probability as possible $1/2 + \epsilon$.

An interpretation of the problem

By fixing β, we fall in a problem list decoding of the first order Reed-Muller code, we have to decode the noisy codeword $\langle \beta, f(X, K) \rangle$.

Results for 8 rounds of DES

Bias $\times 10^4$	Linear Combination		
-2.49	$P_H[15] \oplus P_L[0, 7, 18, 24, 31]$	$\oplus$	$K[4, 9, 13, 31, 33, 41, 44, 52, 54]$
4.86	$P_H[15] \oplus P_L[0, 7, 18, 24, 27, 31]$	$\oplus$	$K[4, 9, 13, 31, 33, 41, 44, 47, 52, 54]$
-4.68	$P_H[15] \oplus P_L[0, 7, 18, 24, 28]$	$\oplus$	$K[4, 9, 15, 31, 33, 41, 44, 52, 54]$
4.81	$P_H[15] \oplus P_L[0, 7, 18, 24, 27, 28]$	$\oplus$	$K[4, 9, 15, 31, 33, 41, 44, 47, 52, 54]$
-2.18	$P_H[15] \oplus P_L[0, 7, 18, 24, 27, 28, 29, 31]$	$\oplus$	$K[9, 13, 15, 31, 33, 41, 44, 47, 52, 54]$
-3.67	$P_H[15] \oplus P_L[0, 7, 18, 24, 27, 28, 31]$	$\oplus$	$K[4, 9, 13, 15, 31, 33, 41, 44, 47, 52, 54]$
-4.59	$P_H[15] \oplus P_L[0, 7, 18, 24, 30]$	$\oplus$	$K[4, 9, 30, 31, 33, 41, 44, 52, 54]$
2.63	$P_H[15] \oplus P_L[0, 7, 18, 24, 27, 30]$	$\oplus$	$K[4, 9, 30, 31, 33, 41, 44, 47, 52, 54]$
2.3	$P_H[15] \oplus P_L[0, 7, 18, 24, 29, 30, 31]$	$\oplus$	$K[9, 13, 30, 31, 33, 41, 44, 52, 54]$
2.69	$P_H[15] \oplus P_L[0, 7, 18, 24, 27, 29, 30, 31]$	$\oplus$	$K[9, 13, 30, 31, 33, 41, 44, 47, 52, 54]$
3.77	$P_H[15] \oplus P_L[0, 7, 18, 24, 30, 31]$	$\oplus$	$K[4, 9, 13, 30, 31, 33, 41, 44, 52, 54]$
3.23	$P_H[15] \oplus P_L[0, 7, 18, 24, 27, 30, 31]$	$\oplus$	$K[4, 9, 13, 30, 31, 33, 41, 44, 47, 52, 54]$
2.43	$P_H[15] \oplus P_L[0, 7, 18, 24, 27, 28, 29, 30]$	$\oplus$	$K[9, 15, 30, 31, 33, 41, 44, 47, 52, 54]$
-3.33	$P_H[15] \oplus P_L[0, 7, 18, 24, 28, 30]$	$\oplus$	$K[4, 9, 15, 30, 31, 33, 41, 44, 52, 54]$
-3.13	$P_H[15] \oplus P_L[0, 7, 18, 24, 28, 29, 30, 31]$	$\oplus$	$K[9, 13, 15, 30, 31, 33, 41, 44, 52, 54]$
4.52	$P_H[15] \oplus P_L[0, 7, 18, 24, 28, 30, 31]$	$\oplus$	$K[4, 9, 13, 15, 30, 31, 33, 41, 44, 52, 54]$
2.05	$P_H[15] \oplus P_L[7, 18, 24, 27, 31]$	$\oplus$	$K[4, 9, 13, 31, 33, 41, 44, 47, 52]$
2.48	$P_H[15] \oplus P_L[7, 18, 24, 27, 28, 30, 31]$	$\oplus$	$K[4, 9, 13, 15, 30, 31, 33, 41, 44, 47, 52]$
4.82	$P_H[15] \oplus P_L[7, 18, 24, 31]$	$\oplus$	$K[4, 9, 13, 31, 33, 41, 44, 52]$
2.05	$P_H[15] \oplus P_L[7, 18, 24, 27, 31]$	$\oplus$	$K[4, 9, 13, 31, 33, 41, 44, 47, 52]$
2.49	$P_H[15] \oplus P_L[7, 18, 24, 28, 29, 31]$	$\oplus$	$K[9, 13, 15, 31, 33, 41, 44, 52]$
-3.4	$P_H[15] \oplus P_L[7, 18, 24, 27, 28, 31]$	$\oplus$	$K[4, 9, 13, 15, 31, 33, 41, 44, 47, 52]$
3.55	$P_H[15] \oplus P_L[7, 18, 24, 29, 30]$	$\oplus$	$K[9, 30, 31, 33, 41, 44, 52]$
-2.31	$P_H[15] \oplus P_L[7, 18, 24, 27, 30]$	$\oplus$	$K[4, 9, 30, 31, 33, 41, 44, 47, 52]$
2.28	$P_H[15] \oplus P_L[7, 18, 24, 27, 28, 29, 30]$	$\oplus$	$K[9, 15, 30, 31, 33, 41, 44, 47, 52]$
5.83	$P_H[15] \oplus P_L[7, 18, 24, 27, 28, 29, 30, 31]$	$\oplus$	$K[9, 13, 15, 30, 31, 33, 41, 44, 47, 52]$

TAB.: Ciphertext bits combination : $C_L[12, 16] \oplus C_H[7, 18, 24]$

Soft information on the key

We remark that we have a soft information on certain point of the linear function
$H(X) = K_4X_1 + (K_9 + K_{31} + K_{33} + K_{41} + K_{44} + K_{52})X_2 + K_{13}X_3 + K_{15}X_4 + K_{47}X_5 + K_{54}X_6$

Given a sample of $(X, f(X, K))$, let $s_0(i) = \#\{X \| H(\lambda_i) = 0\}$ and $s_1(i) = \#\{X \| H(\lambda_i) = 1\}$.

Let $y(\lambda_i) = s_0(i) \log_2\left(\frac{1/2-\epsilon_i}{1/2+\epsilon_i}\right) + s_1(i) \log_2\left(\frac{1/2+\epsilon_i}{1/2-\epsilon_i}\right)$

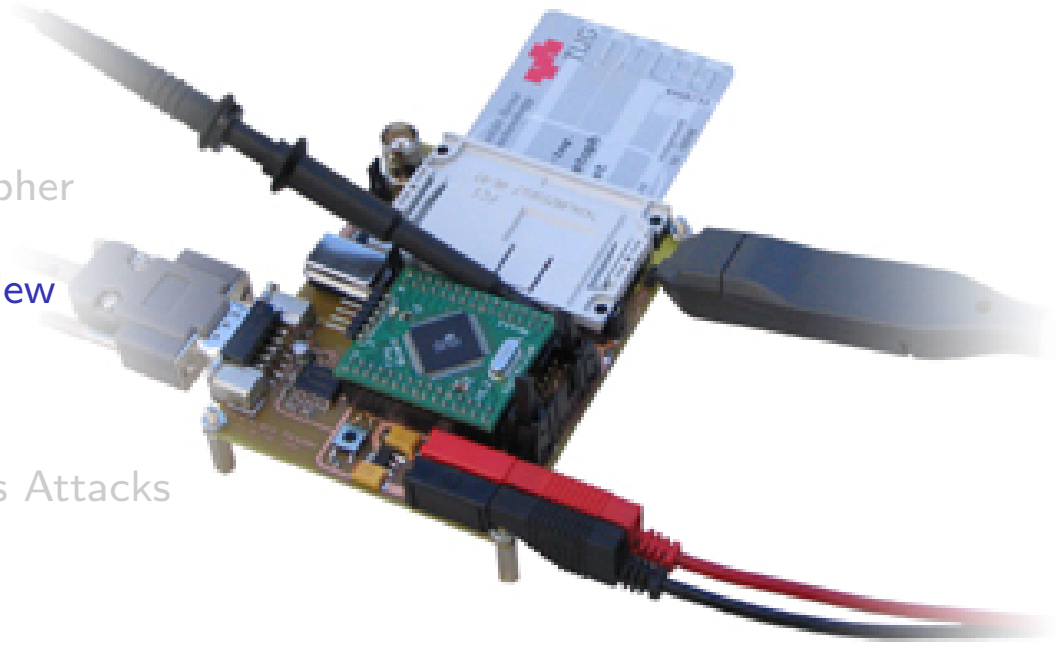
If λ does not correspond to a obtained relation, we set $y(\lambda) = 0$

Thus we have to decode the vector $(y(\lambda))_\lambda$

$\Rightarrow$ Determine H s.t. $\sum_\lambda y(\lambda)(-1)^{H(\lambda)}$ is max.

We reconstruct 6 bits of information with a complexity $\approx 2^{20}$

Plan



Generalities ([Koch 99])

Fundamental idea

The power consumption of a CMOS circuit is correlated with the operation it performs.

Power consumption leakage : starting points

- Observable only at synchronized points (e.g. data transfers through buses or stored in registers).
- Two models
  - Hamming Weight leakage : Number of 1-bits being processed at a given time.
  - Hamming Distance leakage : Number of modified bits during "k" clock cycles.

Generalities ([Koch 99])

Fundamental idea

The power consumption of a CMOS circuit is correlated with the operation it performs.

Power consumption leakage : starting points

- Observable only at synchronized points (e.g. data transfers through buses or stored in registers).
- Two models
  - Hamming Weight leakage : Number of 1-bits being processed at a given time.
  - Hamming Distance leakage : Number of modified bits during "k" clock cycles.

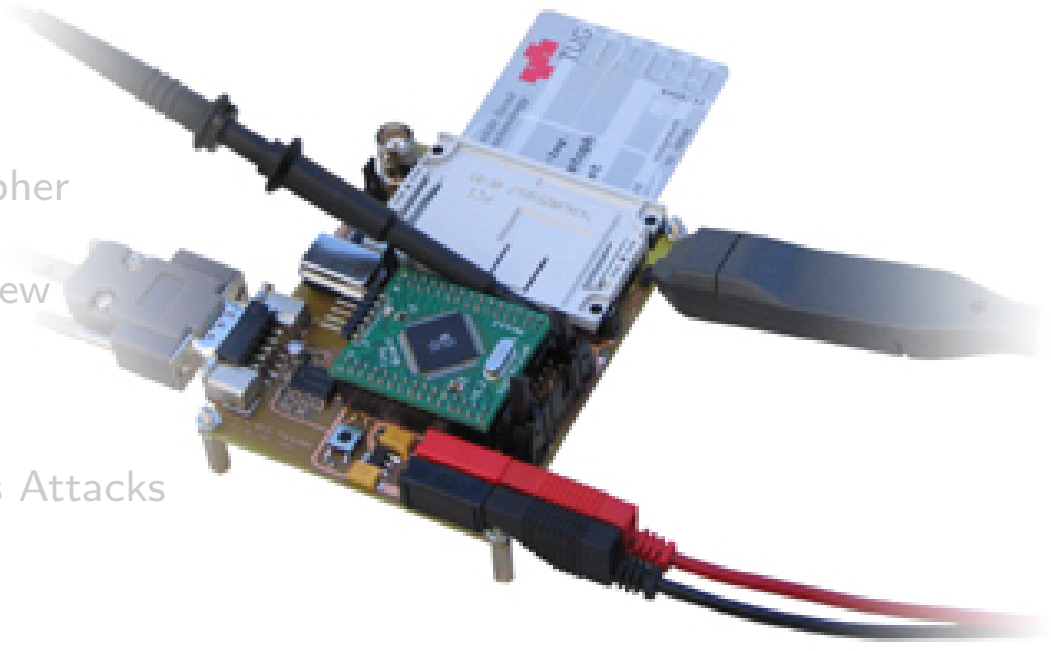
SPA/DPA/HO-DPA ([Koch 99])

SPA : System's power consumption directly observed at a given time.
Allows to break the target device when the execution path depends on the processed data.

DPA : SPA at a given time + statistical analysis.
Require to express an intermediate data value as a function of input bits and few key bits (less than 32 key bits).

HO-DPA : DPA involving several intermediate data values.

Plan



The implementation is safe if one can shut down all information leakages :

Suppress synchronization elements. (buses and registers)

and/or Randomize the data processed.
(masking techniques [Akka 01, Akka 03, Akka 04, Lv 05])

and/or Add random useless computations.

and/or balanced dynamic dual-rail gates designs.

and/or ...

Feasible ?

Maybe ... But at which cost ?

e.g. "Three 32-Bit Random Masks and Six Additional S-Boxes are the Minimal Cost for a Secure DES Implementation" [Lv 05]

The implementation is safe if one can shut down all information leakages :

Suppress synchronization elements. (buses and registers)

and/or Randomize the data processed.
(masking techniques [Akka 01, Akka 03, Akka 04, Lv 05])

and/or Add random useless computations.

and/or balanced dynamic dual-rail gates designs.

and/or ...

Feasible ?

Maybe ... But at which cost ?

e.g. "Three 32-Bit Random Masks and Six Additional S-Boxes are the Minimal Cost for a Secure DES Implementation" [Lv 05]

Glued Blocks

Solution

Concentrate on the firsts and lasts rounds.
i.e. no information leak during these critical rounds

$\Rightarrow$ No observable intermediate value is dependent to less than 32 key bits.

Enough ?

Introducing a new power analysis attack we show that it is **not**.

Glued Blocks

Solution

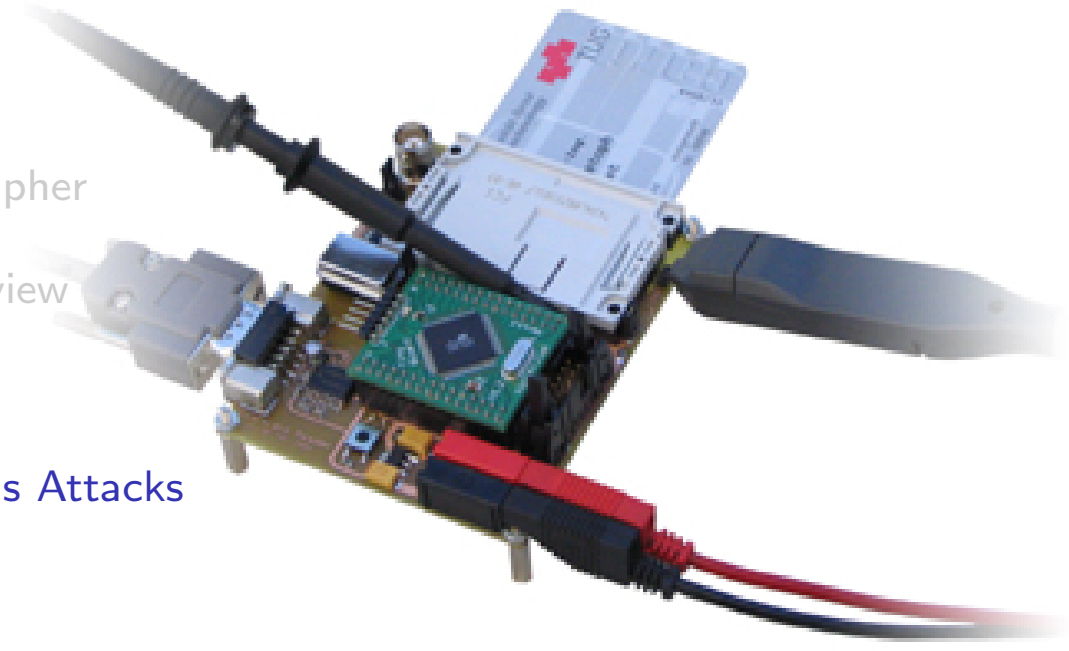
Concentrate on the firsts and lasts rounds.
i.e. no information leak during these critical rounds

$\Rightarrow$ No observable intermediate value is dependent to less than 32 key bits.

Enough ?

Introducing a new power analysis attack we show that it is **not**.

Plan



Linear approximations

Key idea

By the means of classical linear cryptanalysis of block cipher [Mats 93] one can approximate intermediate values as a function of inputs bits and few key bits even when it is strictly dependent to more than 32 key bits.

Formal Definition

Let us denote $|K|$, $|P|$, $|C|$ respectively the bit-lengths of key, plaintext and ciphertext. Let us consider a vector Π of length $|P|$, κ of length $|K|$ and Γ of length $|C|$ and a bit b.
Π, κ, Γ and b define a linear approximation of bias ϵ over the symmetric cipher if and only if :

$$\Pr_{P,K}(< P, \Pi > \oplus < K, \kappa > \oplus b =< C(P,K), \Gamma >) \geq 1/2 + \epsilon$$

Multi-linear cryptanalysis approach

Optimize the attack assuming we can find n "good shaped" linear approximations.
i.e. Few key bits (k) involved in the linear approximations' union.

Algorithm steps :

1. Proceed as in a classical DPA attack for every linear approximation.
2. Recover the sub-key by Reed-Muller decoding from the resulting word (on 2^k bits) as if coming from a noisy and erasure channel.

Correlation analysis

The measured quantity is evaluated with respect to the power consumption model (HW or HD) and some basic analysis of power consumption on the target board.

Results on real traces (dpa-contest) are made based on a (very) rough estimate of the Hamming distance.

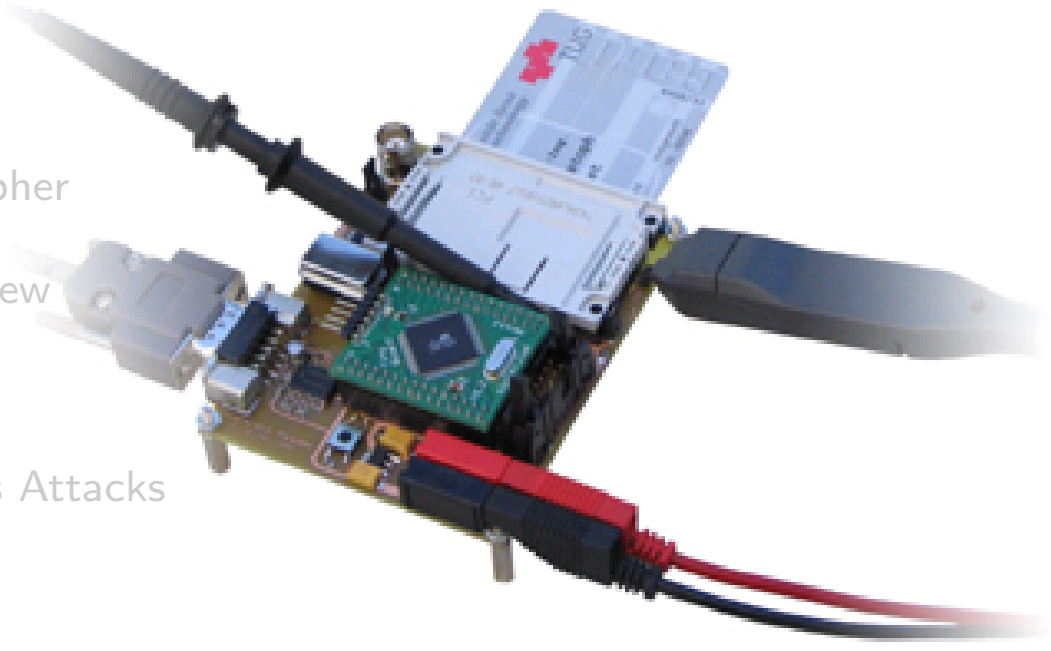
Twin board

Getting the linear approximations from a twin board

i.e. Chosen plaintexts and keys

- Approximations directly linked to the leaked information.
  much more accurate.
- No need to choose a power consumption model.
- No need to know the target block cipher.
  still need to know where/when to attack.

Plan



Information leakage = Hamming weight or Hamming distance

Cipher	Model	rounds	# linear equ.	# key bits	# Plaintexts	Pr(Success)
DES	HW	1	349	30	2^{10}	0.79
DES	HW	1	349	48	2^{12}	0.99
DES	HW	2	728	6	2^{9}	0.97
DES	HW	2	728	48	2^{12}	0.95
DES	HW	3	164	12	2^{17}	0.96
DES	HW	3	164	27	2^{20}	0.99
DES	HD	2	27	16	2^{14}	0.71
DES	HD	2	27	16	2^{16}	0.99
AES	HW	Last	1410	128	2^{10}	0.80
AES	HW	Last	1410	128	2^{11}	0.99

TAB.: Simulation Results

Information leakage from the dpa-contest traces

From traces "secmatv1_2006_04_0809"
http://www.dpacontest.org/

Cipher	rounds	# linear equ.	# key bits	# traces
DES	1	84	∼20	1000
DES	1	84	45	20000
DES	2	163	∼10	1000
DES	2	163	47	36000

TAB.: Attack on DPA-contest traces Results

Important remark :

> *The best PA attacks (i.e. DPA) are extremely empirical and not fully understood. It is our belief that with the growth of our understanding of how, when and which information is leaked, better counter-measure techniques will emerge. After which PA attacks shall get closer to the classical cryptanalysis, MLPA is an example of such rapprochement.*

Next Steps on MLPA :

- Other block ciphers.
- Better linear approximations.
- Unknown block cipher attack.

Questions ?

[*Akka* 01] M.-L. Akkar and C. Giraud. "An Implementation of DES and AES, Secure against Some Attacks". In : Çetin Kaya Koç, D. Naccache, and C. Paar, Eds., *CHES*, pp. 309–318, Springer, 2001.

[*Akka* 03] M.-L. Akkar and L. Goubin. "A Generic Protection against High-Order Differential Power Analysis". In : T. Johansson, Ed., *FSE*, pp. 192–205, Springer, 2003.

[*Akka* 04] M.-L. Akkar, R. Bevan, and L. Goubin. "Two Power Analysis Attacks against One-Mask Methods". In : B. K. Roy and W. Meier, Eds., *FSE*, pp. 332–347, Springer, 2004.

[*Jako* 98] T. Jakobson. "Cryptanalysis of Block Ciphers with Probabilistic Non-linear Relations of Low Degree". In : H. Krawczyk, Ed., *CRYPTO*, pp. 212–222, Springer, 1998.

[*Koch* 99] P. Kocher, J. J. E, and B. Jun. "Differential Power Analysis". In : , pp. 388–397, Springer-Verlag, 1999.

[*Koet* 03] R. Koetter and A. Vardy. "Algebraic soft-decision decoding of Reed-Solomon codes". *IEEE Transactions on Information Theory*, Vol. 49, No. 11, pp. 2809–2825, 2003.

[*Loid*] P. Loidreau, R. Fourquet, and C. Tavernier. "Finding good linear approximations of block ciphers and its application to cryptanalysis of reduced round DES". Can be found here : http ://ced.tavernier.free.fr/.

[*Lv* 05] J. Lv and Y. Han. "Enhanced DES Implementation Secure Against High-Order Differential Power Analysis in Smartcards". In : C. Boyd and J. M. G. Nieto, Eds., *ACISP*, pp. 195–206, Springer, 2005.

[*Mats* 93] M. Matsui. "Linear Cryptoanalysis Method for DES Cipher". In : *EUROCRYPT*, pp. 386–397, 1993.

[*Suda* 97] M. Sudan. "Decoding of Reed Solomon Codes beyond the Error-Correction Bound". *J. Complexity*, Vol. 13, No. 1, pp. 180–193, 1997.

The end.

references I

M.-L. Akkar and C. Giraud.
"An Implementation of DES and AES, Secure against Some Attacks".
In : Çetin Kaya Koç, D. Naccache, and C. Paar, Eds., *CHES*, pp. 309–318, Springer, 2001.

M.-L. Akkar and L. Goubin.
"A Generic Protection against High-Order Differential Power Analysis".
In : T. Johansson, Ed., *FSE*, pp. 192–205, Springer, 2003.

M.-L. Akkar, R. Bevan, and L. Goubin.
"Two Power Analysis Attacks against One-Mask Methods".
In : B. K. Roy and W. Meier, Eds., *FSE*, pp. 332–347, Springer, 2004.

references II

G. K. I. Dumer and C. Tavernier.
"List Decoding of the First Order Binary Reed Muller Codes".
Problems of Information Transmission, Vol. 43, No. 3, pp. 225–232, 2007.

G. Kabatiansky and C. Tavernier.
"List decoding with Reed Muller codes of order one".
In : *nine International Workshop On Algebraic and Combinatorial Coding Theory*, pp. 230–236, 2004.

P. Kocher, J. J. E, and B. Jun.
"Differential Power Analysis".
In : , pp. 388–397, Springer-Verlag, 1999.

references III


J. Lv and Y. Han.
"Enhanced DES Implementation Secure Against High-Order Differential Power Analysis in Smartcards".
In : C. Boyd and J. M. G. Nieto, Eds., *ACISP*, pp. 195–206, Springer, 2005.


M. Matsui.
"Linear Cryptoanalysis Method for DES Cipher".
In : *EUROCRYPT*, pp. 386–397, 1993.