When CPA and MIA go hand in hand

Shiqian WANG, **Thanh-Ha Le**, Maël Berthier

Morpho

24/02/2011

SAFRAN
Morpho

Context

- **SPACES: a French – Japanese project**
  - **Security evaluation of Physical Attack Cryptoprocessors in Embedded Systems**
  - **French: Telecom Paristech, LIP6, Morpho**
  - **Japanese: U. Tohoku, U. Kobe, AIST, UEC**

- **Some results of Shiqian Wang during the internship in Morpho**
  - **Implementation and evaluation of MIA approaches**
  - **August 2010 – February 2011**

Outline

- CPA and Extended CPA
- MIA
- Combination of CPA and MIA
- Conclusion

Correlation analysis

H_K : prediction corresponding to the key K of an intermediate value
$\varphi(H_K)$: selection function of H_K (ex: Hamming weight, Hamming distance)
W : side-channel observation

- **Correlation Power Analysis (CPA)**

$$\rho(W;\varphi(H_K)) = \frac{cov(W,\varphi(H_K))}{\sigma_W \sigma_{\varphi(H_K)}}$$

→ Linear dependency

Pearson's correlation coefficient

- **Other correlation approaches**
  - Spearman's Rank Correlation Coefficient
  - Kendall's Rank Correlation Coefficient (test a monotone relation)
  - Correlation of Distances

Extended CPA

Higher order correlation coefficients

- R.Blacher: « Indicateur de dependance entre deux variables aleatoires fournis par le developpement en serie de la densite de probabilite ». PhD thesis, Universite scientique et medicale de Grenoble (1983)
- R.Blacher: « Quelques applications des fonctions orthogonales en probabilte et statistique ». PhD thesis, Universite Joseph-Fourier Grenoble 1 (1990)

Extended CPA

- Given two orthogonal polynomial families (P_i) (i=0,1, 2…), (P'_j) (j=0,1, 2…), and two random variables X, Y. If X and Y are independent then for all (i,j) (i, j >0)

$$\rho_{i,j} = \frac{cov(P_i(X), P'_j(Y))}{\sigma_{P_i(X)}\sigma_{P'_j(Y)}} = 0$$

- Exploiting the dependency in different orders

Examples of Legendre and Hermite polynomials:

Legendre polynomials:

$$P_0 = 1;$$
$$P_1 = X;$$
$$(n+1) * P_{n+1} = (2n+1) * X * P_n - n * P_{n-1}$$

Hermites polynomials:

$$H_0 = 1;$$
$$H_1 = X;$$
$$H_{n+1} = X * H_n - n * H_{n-1}$$

Extended CPA

Alg. 1 ECPA side-channel distinguisher

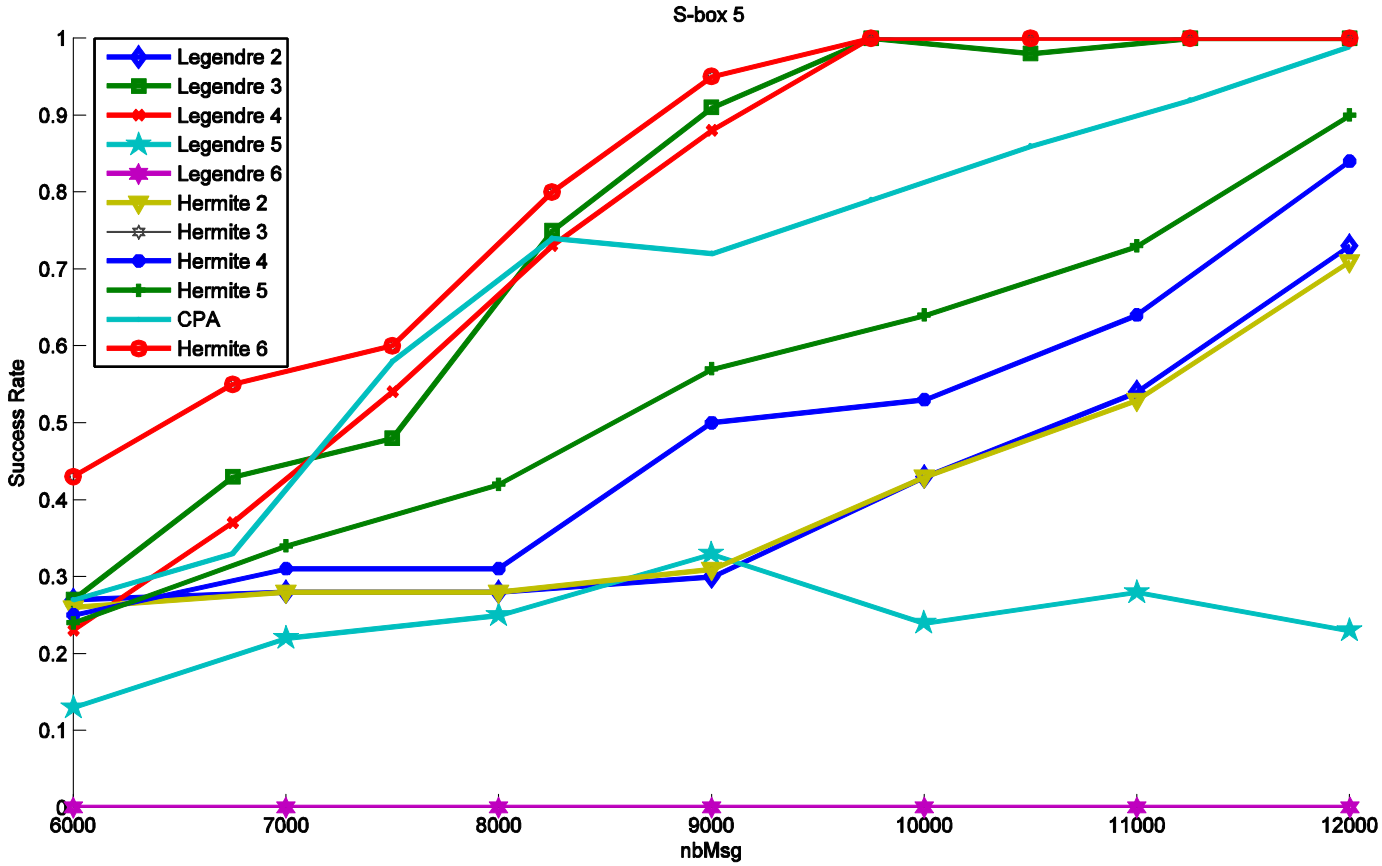
Require: plaintext vector X, consumption leakage vector C, *order* or the correlation coefficients.

Ensure: the guess of the good key: *keyguess*.

```
`// Create a vector Score of length 256 to stock the values of Score(HW_{k*}, C) for
k* ∈ [1, 2...256]
Score ← zeros(1,256);
for i = 1 : order do
   Generate the polynomial P_i (Legendre or Hermite);
end for
for key = 1 : 256 do
   Compute HW_key(X);
   for i = 1 : order do
      for j = 1 : order do
         Score(key) = Score(key) + cov(P_i(HW_key), P_j(C)) / (σ_{P_i(HW_key)} σ_{P_j(C)});
      end for
   end for
end for
return key such that Score(key) is the maximum of the vector Score.
```

Extended CPA: result

Signals from DPA Contest 2
Sbox 5



Mutual Information Analysis

H_K : prediction corresponding to the key K of an intermediate value
$\varphi(H_K)$: selection function of H_K (ex: Hamming weight, Hamming distance)
W : side-channel observation

Correlation Power Analysis (CPA)

$$\rho(W;\varphi(H_K)) = \frac{cov(W,\varphi(H_K))}{\sigma_W \sigma_{\varphi(H_K)}}$$

Linear dependency

! Correlation coefficient: easy to compute

Mutual information Analysis (MIA)

$$I(W;\varphi(H_K)) = \sum_{x\in W, y\in\varphi(H_K)} \mathbf{Pr}[W=x,\varphi(H_K)=y]\log\frac{\mathbf{Pr}[W=x,\varphi(H_K)=y]}{\mathbf{Pr}[W=x]\mathbf{Pr}[\varphi(H_K)=y]}$$

Linear dependency
Nonlinear dependencies

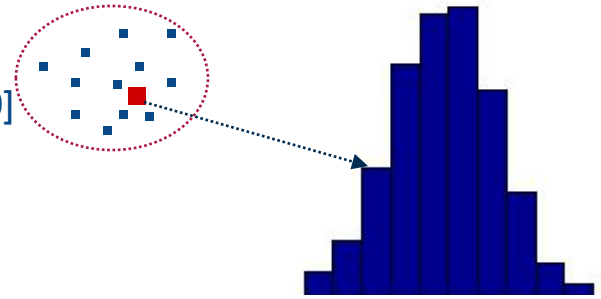
! $\mathbf{Pr}[W=x,\varphi(H_K)=y]$ & $\mathbf{Pr}[W=x]$ not evident to estimate

Some non-parametric approaches

- **Histogram-based Estimator**

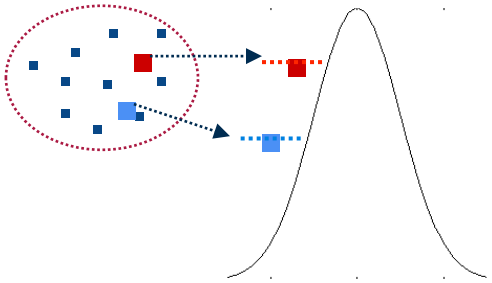
  [Gierlichs08, Veyrat09, Prouff09, Moradi09, Gierlichs10, Venelli10, Flament10]

  - **Simple**
  - **Number of bins, value distribution into bins ?**



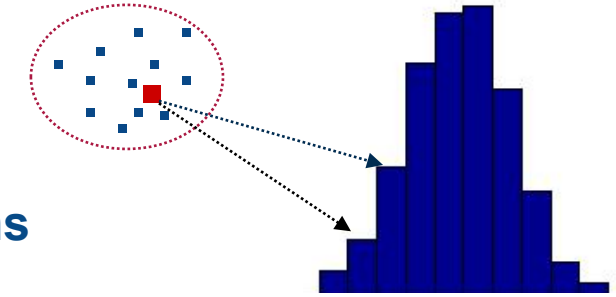
- **Kernel Density Estimator** [Veyrat09, Prouff09]

  - **Neighbouring zones defined by a Kernel function**
  - **Summing up these zones to build the estimator**



- **B-Spline Estimator** [Venelli10]

  - **A generalized histogram-based method**
  - **1 data point in several bins using B-Spline functions**



A comparison

Signals from DPA Contest 1

Method	Average success rate with 500 Msg	Nb of Msg for a 80% success rate
DPA	98%	300
CPA	100%	350
Spearman	98%	350
Kendall	100%	300
Correlation of distances	96%	500
MIA equidistant histogram	76%	F
MIA equiprobable histogram	75%	F
MIA kernel	80%	F
MIA KNN	$< 20\%$	F
MIA cumulant	97%	400
MIA linear B-Splines	85%	F
MIA quadratic B-Splines	94%	500
MIA bins and interpolations	70%	F
MIA adaptive partitioning	62%	F
CvM distance	40%	F
KS distance	$< 30\%$	F

Combinaison of CPA and MIA

$$Score = I(HW_k, C) \; * g(\rho(HW_k, C)) * f(\rho_e(HW_k, C))$$

MIA CPA Extended CPA

f et g: two increasing functions

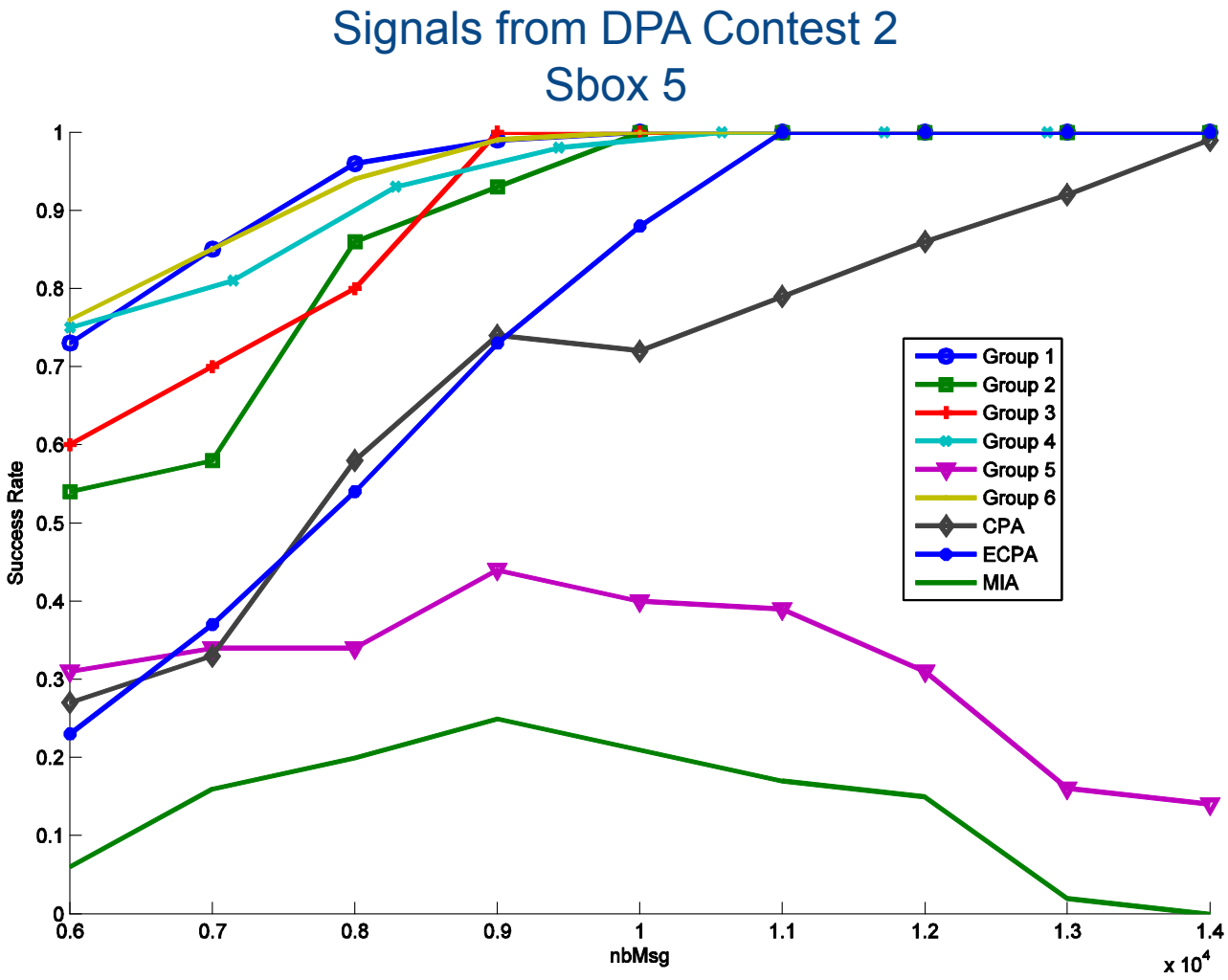
ρ_e : computed with the family « Legendre 4 ».

Histogram-based MIA

Six groups of (*f* , *g*) are tested:

- Group 1: *f(x)=x, g(x)=1* (combination of MIA and E-CPA)
- Group 2: *f(x)=1,g(x)=x* (combination of MIA and CPA)
- Group 3: *f(x)=x, g(x)=x* (combination of 3 methods)
- Group 4: *f(x)=* $\frac{1+x}{1-x}$, *g(x)=1* (combination of MIA and HOCPA)
- Group 5: *f(x)=1, g(x)=* $\frac{1+x}{1-x}$ (combination of MIA and CPA)
- Group 6: *f(x)=* $\frac{1+x}{1-x}$, *g(x)=* $\frac{1+x}{1-x}$ (combination of 3 methods)

Combinaison of CPA and MIA: result



Conclusion

- **Two new ideas**
  - **Extended CPA**
  - **Combination of CPA and MIA**
- **Need solid mathematical backgrounds for optimizing our propositions**
  - **Choice of orthogonal polynomials**
  - **Choice of combination**
- **Link between Extended CPA and MIA?**

Thank you for your attention

thanh-ha.le@morpho.com

SAFRAN
Morpho

Orthogonal polynomials

Definition 1: Probability mesure μ

The requirements for a function μ to be a probability measure on a probability space are that:

- μ must return results in the unit interval [0, 1], returning 0 for the empty set and 1 for the entire space.
- μ must satisfy the countable additivity property i.e. for all countable collections E_i of pairwise disjoint sets:
  $\mu(\cup E_i) = \sum_i \mu(E_i)$

Definition 2: Inner product associated to a probability mesure μ

We define an inner product $< .|. >_\mu$ associate to a probability mesure μ between two mesurable real functions f and g by

$$< f|g >_\mu = \int f(x)g(x)d\mu(x) = E_\mu[f * g] \qquad (3)$$

Orthogonal polynomials

Definition 3: Orthogonal polynomial family associated to a given probality mesure μ
A sequence of polynomials $(P_i)_{i=0,1,2...}$ forms an orthogonal polynomial family if

- $deg(P_i)$ =i for i=0,1,2...
- for $i \neq j, < P_i | P_j >_\mu = 0$

When μ stands for a uniform distribution on [-1,1], we obtain an orthogonal family of Legendre polynomials [8]. When μ stands for a normal Gaussian distribution , we obtain an orthogonal family of Hermite polynomials [1] [9].

8. R.Blacher: Indicateur de dépendance entre deux variables aléatoires fournis par le développement en série de la densité de probabilité. PhD thesis, Université scientifique et médicale de Grenoble (1983)
9. R.Blacher: Quelques applications des fonctions orthogonales en probabilté et statistique. PhD thesis, Université Joseph-Fourier Grenoble 1 (1990)