Efficient Selection of Time Samples for High-Order DPA with Projection Pursuits

François Durvaux, François-Xavier Standaert, Nicolas Veyrat-Charvillon, Jean-Baptiste Mairy, Yves Deville

Université catholique de Louvain – Belgium

Outline

1. Motivations
2. PPs on unprotected implementations
3. PPs on masked implementations
4. Achievements & future work

Outline

Context

"Naive" side-channel attack:

- leakage traces with N_s (thousands) time samples
- each time sample is tested independently
- only a small portion of the leaked information is considered

Two big questions:

1. how to combine for a better exploitation?
2. how to extend to masked implementation?

Previous Work

1. Lower dimensional subspace projection:
   - PCA (Archambeau et al. – CHES 2006)
   - LDA (Standaert et al. – CHES 2008)

   (+) eigenvectors (projection) $\Rightarrow$ Points-Of-Interest (POI)
   (−) objective functions target first-order leakages
   (product traces often not applicable)

2. POI selection of masked implementations:
   - "educated guess" (Oswald et al. – CT-RSA 2006)
   - "no need to test all the key candidates"
     (Reparaz et al. – CHES 2012)

   (+) detection speed is increased (search space is reduced)
   (−) still relies on exhaustive search on pairs of points

This Work

Projection pursuits:

- project on lower dimensional subspace
- dimensions selected to maximize an objective function f_{obj}
- improvement of the projection is tracked when applying small random perturbations

(+) can deal with any projection function

(+) can deal with any objective function

But it's a heuristic:

(−) convergence not guaranteed

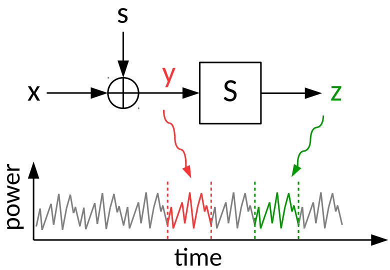
(−) complexity may vary with the context

Outline

Unprotected Setup

Case-study:

- AES S-box
- 8-bit AVR - 20MHz
- $N_s = 1500$



- target: key addition and S-box intermediate values
- profiled setting: 100 traces for each y values $\rightarrow l_y^i(t)$ with $i \in [1; 100]$ and $t \in [0; N_s - 1]$

Instantiation

Projection:

- $\boldsymbol{\alpha}$ as the projection profile
- first-order $\rightarrow$ weighted sum: $\lambda_y^i = \sum_{t=0}^{N_s-1} \alpha(t) \cdot l_y^i(t)$

Objective functions:

1. SNR ($\sim$ LDA):
   - $\hat{\mu}_y = \hat{\mathsf{E}}_i(\lambda_y^i)$, $\quad \hat{\sigma}_y^2 = \hat{\mathsf{var}}_i(\lambda_y^i)$, $\quad$ for $i = 1 \rightarrow 100$
   - $f_{obj} = \hat{\mathsf{var}}_y(\hat{\mu}_y) / \hat{\mathsf{E}}_y(\hat{\sigma}_y^2)$
2. profiled CPA:
   - $\hat{\mu}_y = \hat{\mathsf{E}}_i(\lambda_y^i)$, $\quad$ for $i = 1 \rightarrow 50$
   - $\boldsymbol{\lambda}_y \leftarrow \lambda_y^i$, $\quad$ for $i = 51 \rightarrow 100$
   - $f_{obj} = \hat{\rho}(\hat{\mu}_Y, \boldsymbol{\lambda}_Y)$

Algorithm Description

$\boldsymbol{\alpha} = \text{initialize()}$

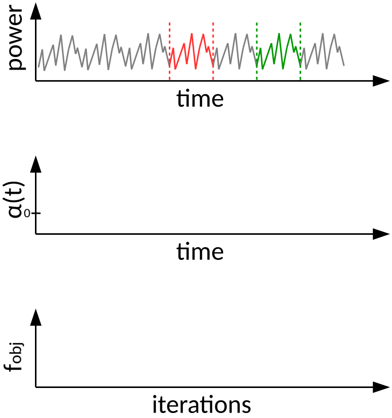
for $j = 1 \rightarrow N_r$

$r = \text{rand\_index}(\text{N}_\text{s})$

$\alpha_m = \underset{\alpha(\text{r})}{\text{arg\_max}}(@\text{f}_{\text{obj}}, \boldsymbol{\alpha}, \text{r})$

$\alpha(r) = \alpha_m$

endfor



Algorithm Description

$\boldsymbol{\alpha} = \text{initialize}()$

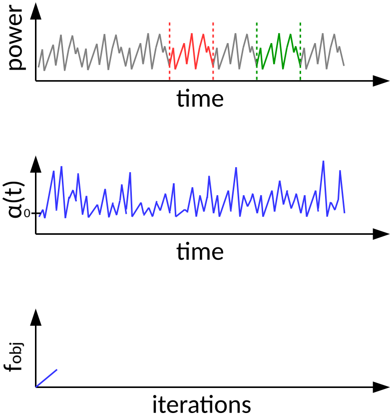
for $j = 1 \rightarrow N_r$

$r = \text{rand\_index}(\mathrm{N_s})$

$\alpha_m = \underset{\alpha(\mathrm{r})}{\text{arg\_max}}(@\mathrm{f_{obj}}, \boldsymbol{\alpha}, \mathrm{r})$

$\alpha(r) = \alpha_m$

endfor



Algorithm Description

$\boldsymbol{\alpha} = \text{initialize}()$

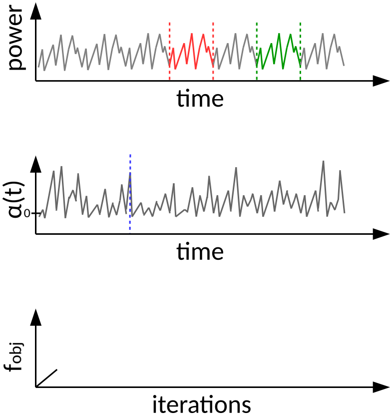
for $j = 1 \rightarrow N_r$

$r = \text{rand\_index}(\mathsf{N_s})$

$\alpha_m = \underset{\alpha(\mathsf{r})}{\text{arg\_max}}(@\mathsf{f_{obj}}, \boldsymbol{\alpha}, \mathsf{r})$

$\alpha(r) = \alpha_m$

endfor



Algorithm Description

$\boldsymbol{\alpha} = \text{initialize}()$

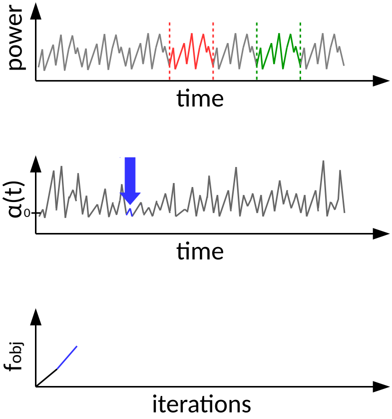
for $j = 1 \rightarrow N_r$

$r = \text{rand\_index}(\mathrm{N_s})$

$\alpha_m = \underset{\alpha(r)}{\text{arg\_max}}(@\mathrm{f_{obj}}, \boldsymbol{\alpha}, \mathrm{r})$

$\alpha(r) = \alpha_m$

endfor



Algorithm Description

$\boldsymbol{\alpha} = \text{initialize}()$

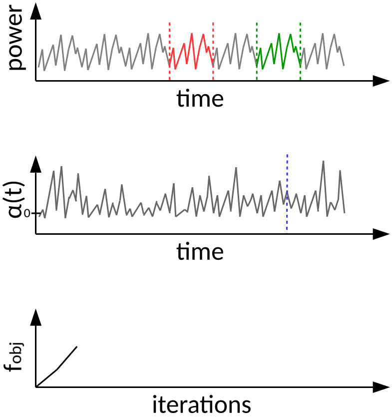
for $j = 1 \rightarrow N_r$

$r = \text{rand\_index}(\mathsf{N_s})$

$\alpha_m = \underset{\alpha(\mathsf{r})}{\text{arg\_max}}(@\mathsf{f_{obj}}, \boldsymbol{\alpha}, \mathsf{r})$

$\alpha(r) = \alpha_m$

endfor



Algorithm Description

$\boldsymbol{\alpha} = \text{initialize}()$

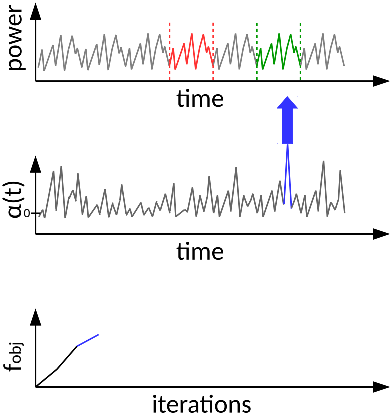
for $j = 1 \rightarrow N_r$

$r = \text{rand\_index}(\text{N}_\text{s})$

$\alpha_m = \underset{\alpha(\text{r})}{\text{arg\_max}}(@\text{f}_{\text{obj}}, \boldsymbol{\alpha}, \text{r})$

$\alpha(r) = \alpha_m$

endfor



Algorithm Description

$\boldsymbol{\alpha} = \text{initialize}()$

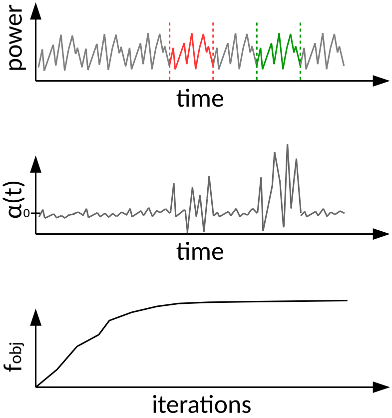
for $j = 1 \rightarrow N_r$

$r = \text{rand\_index}(\mathrm{N_s})$

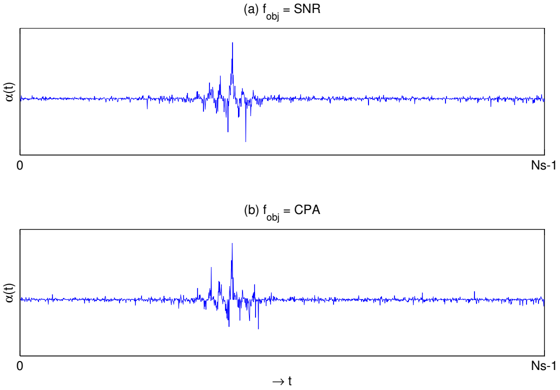
$\alpha_m = \underset{\alpha(\mathrm{r})}{\text{arg\_max}}(@\mathrm{f_{obj}}, \boldsymbol{\alpha}, \mathrm{r})$

$\alpha(r) = \alpha_m$

endfor

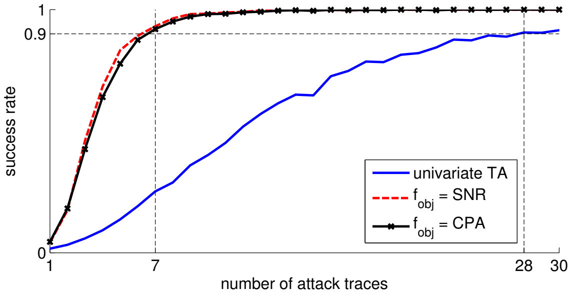


Projection Profiles



- Zoom-in on the S-box $\alpha(t)$'s

Success Rates

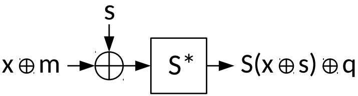


Outline

Protected Setup

Case-study: same but...

- pre-computed tables mask implementation
- first order secure
- $N_s = 30000$
- 50 traces for each y
- masks (m, q) unknown



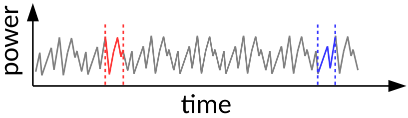
Leakage Traces

Masked S-box:

(1) precomputation

for $i = 0 \rightarrow 255$

$S^*(i) = S(i \oplus m) \oplus q$



(2) execution

$z_q = S^*(y_m)$

with $y_m = y \oplus m$ and $z_q = z \oplus q$

- leakages from q and z_q (or m and y_m) must be manipulated simultaneously!
- need to find the POIs
- directly applying the previous algorithm won't work

Projection

Two possibilities:

1. centered-product (usual candidate):
   - (+) information in the mean*
   - (−) very sensitive to noise (noise variance multiplied)

2. weighted sum (our choice):
   - (±) information in the variance*
   - (+) less sensitive to noise (noise variance added)

Yet, still too much noise if all the samples are considered simultaneously $\Rightarrow$ window-based search

* Standaert et al., "The World is Not Enough: [. . .]", Asiacrypt 2010

Objective Function

Moments against Moments Profiled Correlation (MMPC):

- similar to profiled CPA
- d^{th}-order m_y^d moments are compared

In our case $d = 2$:

- $\hat{m}_y^2 = \hat{\mathrm{var}}_i(\lambda_y^i), \quad$ for $i = 1 \rightarrow 25$
- $\tilde{m}_y^2 = \hat{\mathrm{var}}_i(\lambda_y^i), \quad$ for $i = 26 \rightarrow 50$
- $f_{obj} = \hat{\rho}(\hat{m}_Y^2, \tilde{m}_Y^2)$

Hard detection threshold possible $\rightarrow T_{det} = 0.2$ ($> 3\sigma$ for ρ with 256 elements, see Fisher's transformation)

Algorithm Description

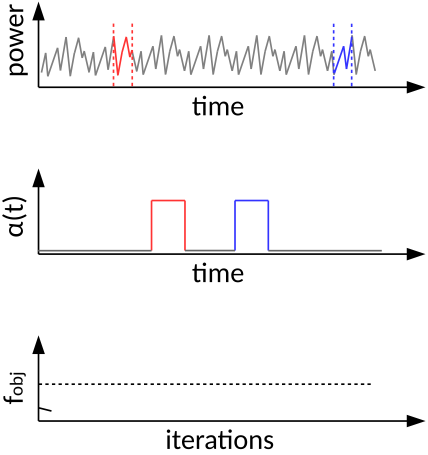
Local search:

- 2 windows of length W_{len}:

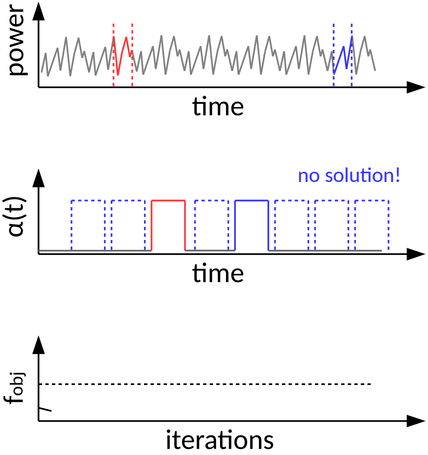
$$\alpha(t) = \begin{cases} 1, & \text{if within a window} \\ 0, & \text{otherwise} \end{cases}$$

- phase 1: find_solution
  - windows repeatedly placed at random locations
  - evaluated neighbours: windows translations
  - until T_{det} is crossed
- phase 2: improve_solution
  - small local perturbations
  - evaluated neighbours: translations and W_{len} modifications

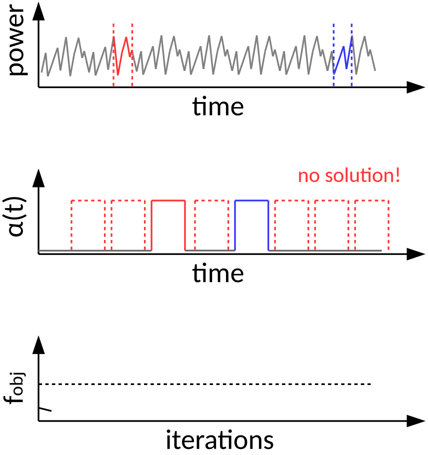
Local Search



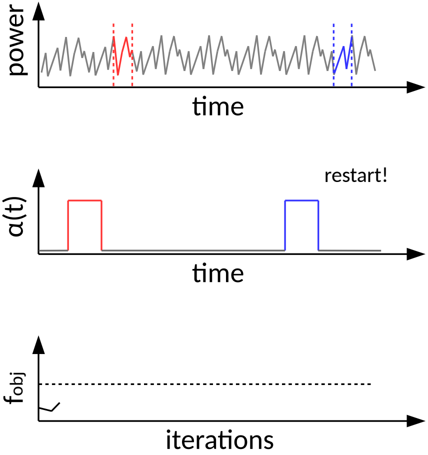
Local Search



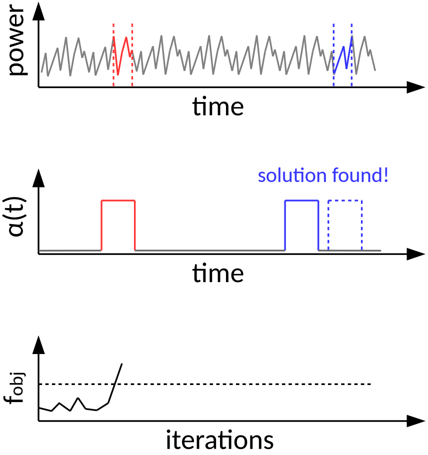
Local Search



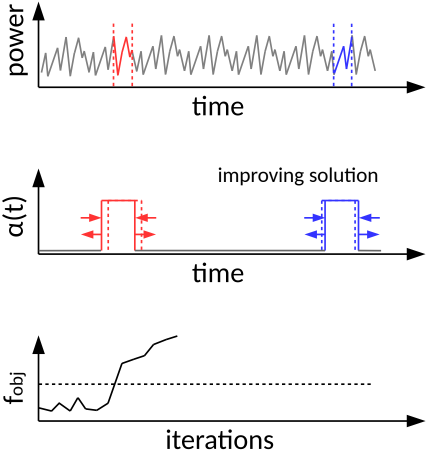
Local Search



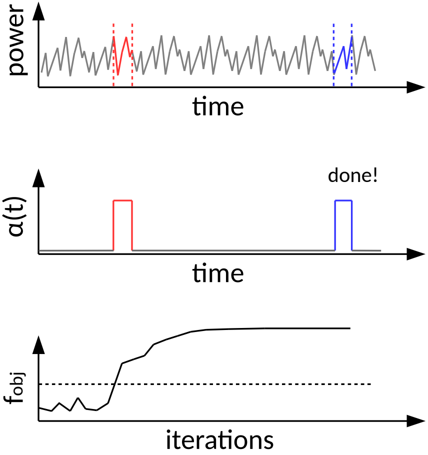
Local Search



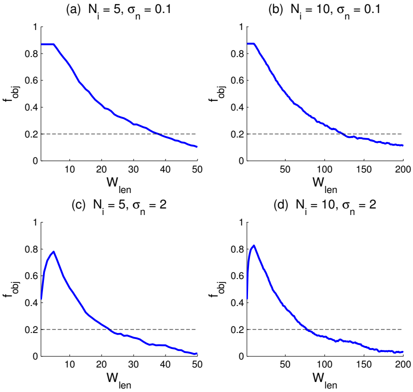
Local Search



Local Search



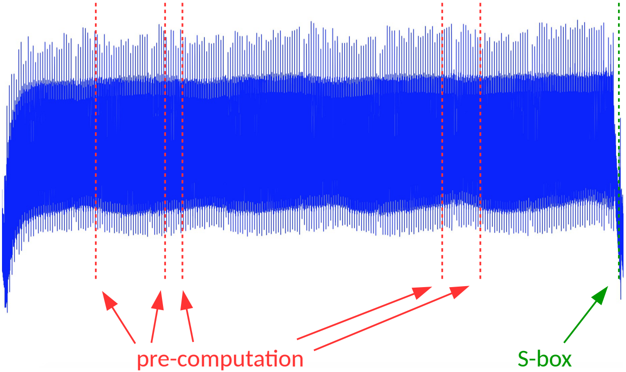
Parameters



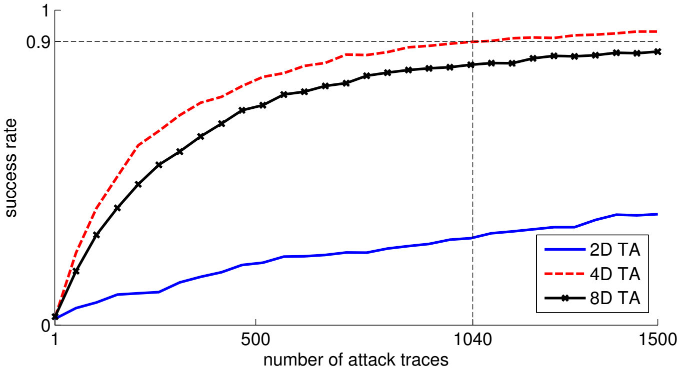
- if $W_{len} \nearrow$ more traces are needed to estimate m_y^2
- trade-off between convergence speed and number of traces

Detected POIs

- $N_s = 30000, \ W_{len} = 25$



Success Rates



Outline

Achievements

This work:

- unprotected case: highlights POIs and combines them
- protected case: POIs detection
- instantiation and convincing results on real measurements

(+) generic tool: any projection function, any f_{obj}

(+) faster than exhaustive search (a constant factor function of W_{len})

(+) easy extension to higher orders of masking

(−) heuristic:

  - parameters to set (sometimes tricky)
  - convergence no guaranteed
  - varying complexity

Future Work

Masked implementations:

- other projection functions
- other objective functions
- other projection profiles than (1,0)
- non-profiled version based on Reparaz et al. observation (already done actually)

Future Work

Masked implementations:

- other projection functions
- other objective functions
- other projection profiles than (1,0)
- non-profiled version based on Reparaz et al. observation (already done actually)

Thank you!

Efficient Selection of Time Samples for High-Order DPA with Projection Pursuits

François Durvaux, François-Xavier Standaert, Nicolas Veyrat-Charvillon, Jean-Baptiste Mairy, Yves Deville

Université catholique de Louvain – Belgium