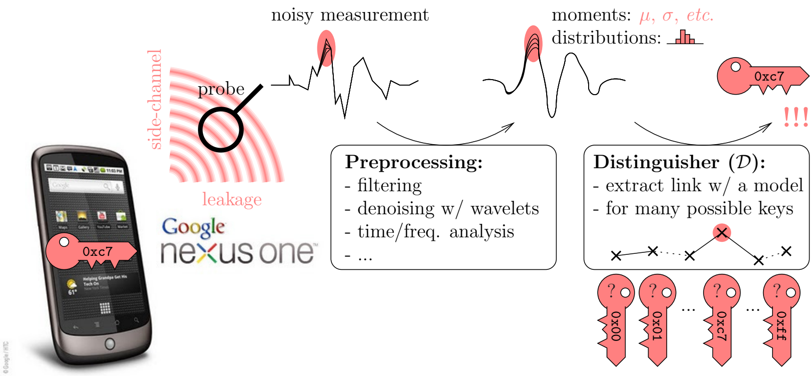
Time-Frequency Analysis for second order attacks

Pierre BELGARRIC, Shivam BHASIN, Nicolas BRUNEAU, Jean-Luc DANGER, Nicolas DEBANDE, Sylvain GUILLEY, Annelie HEUSER, Zakaria NAJM and Olivier RIOUL
< firstname.lastname@TELECOM-ParisTech.fr >

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

Side-Channel Analysis on Embedded Systems [3]



Context

Side-Channel Attacks pose a serious threat to embedded cryptography.

Countermeasures

- Extrinsic:
  - Noise addition
  - Delay insertion
- Intrinsic:
  - Hiding the power [1]
  - Data Masking [2]

Data Masking

Goal

Make the leakage and the intermediate values independent.

Principle

Use random values.

Threat

The masking can be defeated using "High Order attacks".

Presentation Outline

First order Masking: principle

- Aim: making the intermediate value independent of the leakages
- The sensitive variable Z is randomly split into two shares:

$$(P_0 = Z \perp M, P_1 = M)$$

P_0 is the masked variable and $\perp$ is an invertible operation

- Boolean masking is based on exclusive-or ($\mathtt{xor}$) operation:

$$(P_0 = Z \oplus M, P_1 = M)$$

Second order CPA

Idea

Combining (centered product [4] for example) the leakage of the first share and the leakage of the second share.

Software

In software the two shares are manipulated sequentially
$\Rightarrow$ leak in two different times t_0 and t_1.

- $\mathcal{L}(t_0)$ the leakage of the first share
- $\mathcal{L}(t_1)$ the leakage of the second share

$\Rightarrow$ How to perform 2O-CPA without knowing t_0 and t_1?

Second order CPA

Exhaustive search

- Try all the possible couples
- Test them by performing $\mathcal{O}(n^2)$ CPA

Find the good couple

- Sophisticated method to find this couple by Reparaz et al. [5]
- Only one attack

Preprocessing

- Fast way to combine the points by Waddle and Wagner [7]
- The size of the input and the output of the function is equal
- Univariate second order CPA

$\Rightarrow$ Find a fast way to combine the points.

Autocorrelation

To avoid the $\mathcal{O}(n^2)$ complexity Waddle and Wagner propose at CHES '04 [7] the FFT 2DPA.

- Combine the leakage of a window $\mathcal{L}$ using the autocorrelation:

$$(\mathcal{L} \star \mathcal{L})(t) = \sum_{t' \in \mathbb{Z}_n} \mathcal{L}(t') \cdot \mathcal{L}(t' + t)$$

- Compute this using the FFT and the theorem:

$$(\mathcal{L} \star \mathcal{L})(t) = \sqrt{n} \cdot \mathrm{IDFT}\left[\overline{\mathrm{DFT}\left[\mathcal{L}\right]} \cdot \mathrm{DFT}\left[\mathcal{L}\right]\right]$$

And then perform a DPA. $\Rightarrow \mathcal{O}(n \log_2 n)$ complexity

Presentation Outline

Case study

Measurements

- All the traces are derived from DPA contest v4 [6]
- ATMega163 8-bit smartcard

Algorithm

- Rotating Sbox masking
- RSM: Fourth degree masking scheme where the same mask is XORed to one plaintext byte (T) and to some S-box output (corresponding to another plaintext byte T')

Measurements

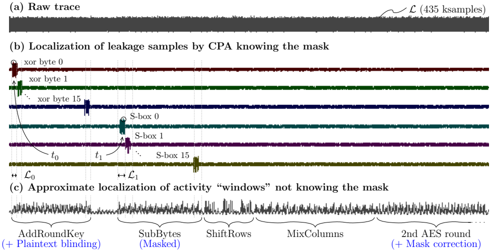


Figure 1: Analyses on traces of the DPA contest

- $\mathcal{L}_0$: the windows in which the share #0 ($T \oplus M$) is expected to leak. $\mathcal{M}_0 = w_H(T \oplus M) - 4$ the model of this leak.
- $\mathcal{L}_1$: the windows in which the share #1 ($\texttt{Sbox}[T' \oplus K]$) is expected to leak. $\mathcal{M}_1 = w_H(\texttt{Sbox}[T' \oplus K] \oplus M) - 4$ the model of this leak.

Auto/Crosscorrelation

AutoCorrelation

Concatenate the two window $\mathcal{L}_{01} = concat(\mathcal{L}_0, \mathcal{L}_1)$ and compute the autocorrelation on $\mathcal{L}_{01}$. Let call that auto-corr.

CrossCorrelation

- Combine the leakage of a window $\mathcal{L}_0$ and the window $\mathcal{L}_1$ using the cross-correlation.
- Compute this using the FFT and the theorem:

$$(\mathcal{L}_0 \star \mathcal{L}_1)(t) = \sqrt{n} \cdot \mathrm{IDFT}\left[\mathrm{DFT}\left[\overline{\mathcal{L}_0}\right] \cdot \mathrm{DFT}\left[\mathcal{L}_1\right]\right]$$

- Call this method x-corr

And then perform a CPA using $\mathcal{M}_{01} = \mathbb{E}[(\mathcal{M}_0 \cdot \mathcal{M}_1)|T, T', K]$ for prediction function. $\Rightarrow \mathcal{O}(n \log_2 n)$ complexity

Commentary

(a) Sample leaking the mask in $\mathcal{L}_0$ (b) Sample leaking the mask in $\mathcal{L}_1$

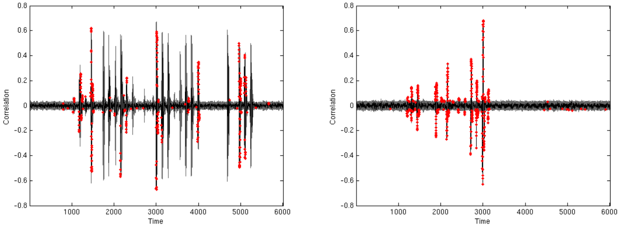


Figure 2: Correlation knowing the mask M

As wee see in this example the shares leaks in *many peaks* in time domain, and will have a *common signature* in frequency domain.

New methods in frequency domain

Preprocessing

New preprocessing methods with the properties:

- Stay in frequency domain
- Computed in $\mathcal{O}(n \log_2 n)$

Attack phase

Compute $\mathcal{O}(n)$ CPA with $\mathcal{M}_{01} = \mathbb{E}[(\mathcal{M}_0 \cdot \mathcal{M}_1) | T, T', K]$ for prediction function.

New methods

$\Rightarrow$ Five new methods that respect these properties.

Frequency domain

Stay in the frequency domain, not compute IFFT

Concatenate windows

- Compute $|\mathrm{DFT}\,[\mathcal{L}_{01}]\,|^2$
- Call this method concat-dft
  $\Rightarrow$ complexity $\mathcal{O}(n \log_2 n)$

Two windows

- Compute $|\overline{\mathrm{DFT}\,[\mathcal{L}_0]} \cdot \mathrm{DFT}\,[\mathcal{L}_1]\,|$
- Call this method window-dft
  $\Rightarrow$ complexity $\mathcal{O}(n \log_2 n)$

We go back in $\mathbb{R}$ using the absolute value $\Rightarrow$ lose the information phase.

DHT

The discrete Hartley transforms of a sequence $Y \in \mathbb{R}^n$ in another sequence: $\mathrm{DHT}[Y] \in \mathbb{R}^n$ such as:

$$\mathrm{DHT}[Y](f) = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} Y(t) \cdot (\cos(2\pi ft/n) + \sin(2\pi ft/n))$$

Property

- Compute the DHT using the DFT with: $\mathrm{DHT}[Y](f) = \Re\mathrm{e}\, \mathrm{DFT}[Y](f) - \Im\mathrm{m}\, \mathrm{DFT}[Y](f)$. $\Rightarrow$ with complexity $\mathcal{O}(n \log_2 n)$
- Real number
- $\mathrm{DHT}[\mathrm{DHT}[Y]] = Y$ without any loss of information

High order CPA with the DHT

Adapt the methods of the DFT with the DHT.

Concatenate windows

- Compute $|\mathrm{DHT}\left[\mathcal{L}_{01}\right]|^2$
- Call this method concat-dht
  $\Rightarrow$ complexity $\mathcal{O}(n \log_2 n)$

Two windows

- Compute $|\mathrm{DHT}\left[\mathcal{L}_{0}\right] \cdot \mathrm{DHT}\left[\mathcal{L}_{1}\right]|$
- Call this method window-dht
  $\Rightarrow$ complexity $\mathcal{O}(n \log_2 n)$

Heuristic method

Method mixing attack and point combining to perform "complex" 2O-CPA.
For example:

$$\max(|\rho((\Re e(\mathrm{DFT}[\mathcal{L}_{01}])^2), \mathcal{M}_{01})|, |\rho((\Im m(\mathrm{DFT}[\mathcal{L}_{01}])^2), \mathcal{M}_{01})|)$$

Call this method max-corr

Positive point

Can give good results.

Negative point

Maybe more data depend.

Presentation Outline

1. Masking scheme and High order attacks
2. New preprocessing methods
3. Empirical results

CPA and 2O-CPA

(a) Prediction function $= \mathcal{M}_1^m$ (b) Prediction function $= \mathcal{M}_{01}$

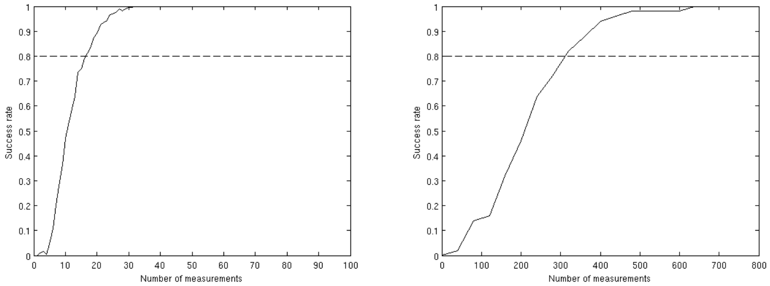


Figure 3: Success rate of (a) univariate CPA attack knowing the mask and (b) bi-variate 2O-CPA attack on knowing (t_0, t_1)

Low window size

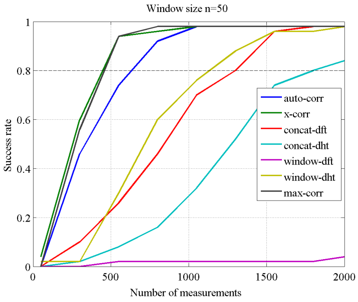


Figure 4: Success rate when using a small window size

Medium window size

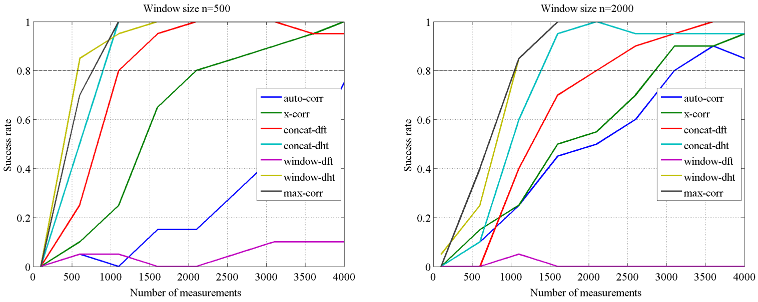
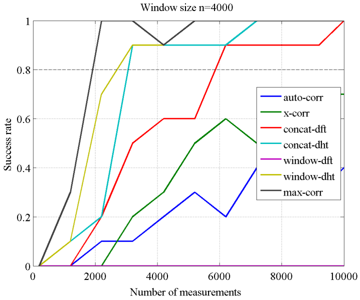


Figure 5: Success rate when using a medium window size

Large window size



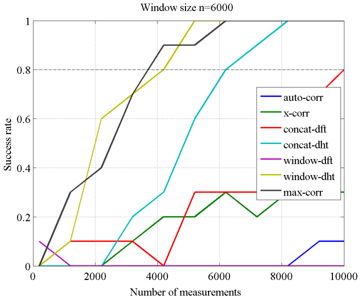


Figure 6: Success rate when using a large window size

Frequency

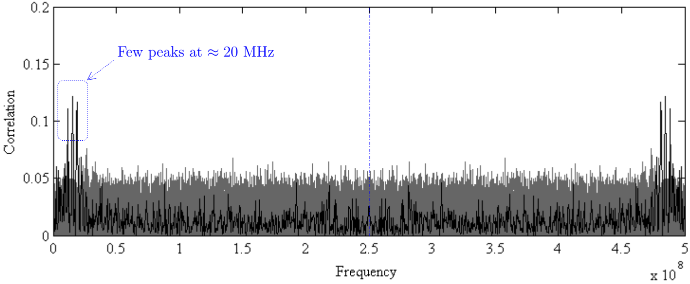


Figure 7: Correlation coefficient on a 2O-CPA on concat-dft in frequency domain when using $n = 6000$ and 10000 traces (we recall that the sampling rate is $F_S = 500$ Msample/s)

Results Summary

Table 1: Comparison of performance of proposed methods against attack efficiency.

Window Size	Best Attack	Number of traces for SR $\geq$0.8
50	x-corr	450
200	x-corr	750
500	window-dht	550
2000	window-dht max-corr	550
4000	max-corr	1950
6000	max-corr	3000

Conclusion

Results

- $\oplus\oplus$ Reduce the complexity from $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log_2 n)$.
- $\ominus$ Increase the number of traces needed.

Next steps

- Try these methods on High Order attacks.
- Try on different leaks.

Thanks for your attention.

References

[1] Jean-Luc Danger, Sylvain Guilley, Shivam Bhasin, and Maxime Nassar.
Overview of Dual Rail with Precharge Logic Styles to Thwart Implementation-Level Attacks on Hardware Cryptoprocessors, — *New Attacks and Improved Counter-Measures* —.
In *SCS*, IEEE, pages 1–8, November 6–8 2009.
Jerba, Tunisia. DOI: 10.1109/ICSCS.2009.5412599.

[2] Louis Goubin and Jacques Patarin.
DES and Differential Power Analysis. The "Duplication" Method.
In *CHES*, LNCS, pages 158–172. Springer, Aug 1999.
Worcester, MA, USA.

[3] Sylvain Guilley, Olivier Meynard, Maxime Nassar, Guillaume Duc, Philippe Hoogvorst, Houssem Maghrebi, Aziz Elaabid, Shivam Bhasin, Youssef Souissi, Nicolas Debande, Laurent Sauvage, and Jean-Luc Danger.
Vade Mecum on Side-Channels Attacks and Countermeasures for the Designer and the Evaluator.
In *DTIS (Design & Technologies of Integrated Systems)*, IEEE. IEEE, March 6-8 2011.
Athens, Greece. DOI: 10.1109/DTIS.2011.5941419 ; Online version:
http://hal.archives-ouvertes.fr/hal-00579020/en/.

[4] Emmanuel Prouff, Matthieu Rivain, and Régis Bevan.
Statistical Analysis of Second Order Differential Power Analysis.
IEEE Trans. Computers, 58(6):799–811, 2009.

[5] Oscar Reparaz, Benedikt Gierlichs, and Ingrid Verbauwhede.
Selecting Time Samples for Multivariate DPA Attacks.
In Emmanuel Prouff and Patrick Schaumont, editors, *CHES*, volume 7428 of *Lecture Notes in Computer Science*, pages 155–174. Springer, 2012.

[6] TELECOM ParisTech SEN research group.
DPA Contest (4^{th} edition), 2013–2014.
http://www.DPAcontest.org/v4/.

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

Time-Frequency Analysis for second order attacks

Pierre BELGARRIC, Shivam BHASIN, Nicolas BRUNEAU, Jean-Luc DANGER, Nicolas DEBANDE, Sylvain GUILLEY, Annelie HEUSER, Zakaria NAJM and Olivier RIOUL
<firstname.lastname@TELECOM-ParisTech.fr>

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