Analysis Of Variance and CPA in SCA

S. Tiran[1], G. Reymond[2], J.B. Rigaud[3], D. Aboulkassimi[2], G. Ducharmes[4], P. Maurine[2]

[1] LIRMM
[2] CEA
[3] EMSE
[4] EPS - Institut de Mathématiques et de Modélisation de Montpellier

Is CPA a wise choice in the context of SCA?

CPA :

- Easy to compute
- Fast
- Only detect linear relations

MIA :

- Can detect any kind of relations
- Requires a good choice of hyper-parameters*
- Time consuming

* Mathieu Carbone, Sébastien Tiran, Sébastien Ordas, Michel Agoyan, Yannick Teglia, G. Ducharme, and Philippe Maurine. On adaptive bandwidth selection for ecient MIA.

Analysis Of Variance in SCA

- **Introduced** by F. Standaert and B. Gierlichs in [1] and further analysed in [2]
- **No general conclusion** could be drawn
- Seems to **give similar results** to the other attacks on the devices that were tested

[1] François-Xavier Standaert, Benedikt Gierlichs, and Ingrid Verbauwhede. Partition vs. comparison side channel distinguishers: An empirical evaluation of statistical tests for univariate side-channel attacks against two unprotected cmos devices.
[2] Lejla Batina, Benedikt Gierlichs, and Kerstin Lemke-Rust. Differential cluster analysis.

Analysis Of Variance in SCA

- **More generic** distinguisher than Pearson correlation
- Can detect any kind of **relation on the means**
- No hyper-parameter to select

Is the **Analysis Of Variance (AOV)** a good alternative to CPA?

AOV principle

- The one-way AOV allows to study the behavior of a random variable $\mathcal{L}$ according to the values of a factor $\mathcal{H}$, taking H distinct values denoted h.
- A sample of n_h values of $\mathcal{L}$, noted $\{L_{h,1},...,L_{h,n_h}\}$ is observed at each value h.

AOV principle

- The total sum of squares is defined as :

$$SS_{tot} = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \left(L_{h,i} - \bar{L} \right)^2$$

- The decomposition can be written as :

$$SS_{tot} = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \left(L_{h,i} - \bar{L}_{h.} \right)^2 + \sum_{h=1}^{H} n_h \left(\bar{L}_{h.} - \bar{L} \right)^2$$

$$SS_{tot} = SS_{err} + SS_{treat}$$

$$R^2_{aov} = 1 - \frac{SS_{err}}{SS_{tot}}$$

Comparing AOV to Pearson correlation

- The coefficient of determination, denoted R^2, indicates how well data fit a statistical model.
- In case of a simple linear regression, R^2_{reg} is simply the square of the **Pearson correlation coefficient** used in **CPA**.
- What is the difference between R^2_{aov} and R^2_{reg}?

Comparing AOV to Pearson correlation

- **When H=2**, the AOV is the same as a squared student t test, and $R^2_{aov} = R^2_{reg}$ (the square of Pearson correlation coefficient).

  In this case, CPA is also equivalent to Kocher's DPA*.
- **When H>2**, R^2_{aov} **is equivalent to** R^2_{reg} if the points $\{(h, \bar{L}_{h.}), h=1...H\}$ fall on a **straight line**.

* Stefan Mangard, Elisabeth Oswald, and François-Xavier Standaert. One for all-all for one: unifying standard differential power analysis attacks.

Comparing AOV to Pearson correlation

Comparison between AOV and CPA :

- Both should give similar results when the leakage is linear
- AOV should provide better results than CPA in non-linear cases

AOV remains less generic than MIA, because it cannot detect links in moments higher than the mean.

Experimental Results

Comparison of two test-cases :

- DPA contest v2
- AES-128 designed with a 65nm Low Power High Threshold Voltage CMOS technology

Leakage profiling

- What is the **shape of the leakage** on these two cases?
  - **Least squares** method to find the polynomial that fits the best the data.
- What is the **degree of the polynomial** that best represents the data?

Leakage profiling

The **AIC (Akaike Information Criterion)** estimates the **degree of the polynomial fitting the best the data**.

The AIC is a trade-off between godness of fit and complexity.

Its expression is :

$$AIC = 2(N+1) + n \times \ln\left(\frac{SS_{reg}}{n}\right)$$

with N the degree of the polynomial, n the number of points, SS_{reg} the regression sum of squares.

Leakage profiling

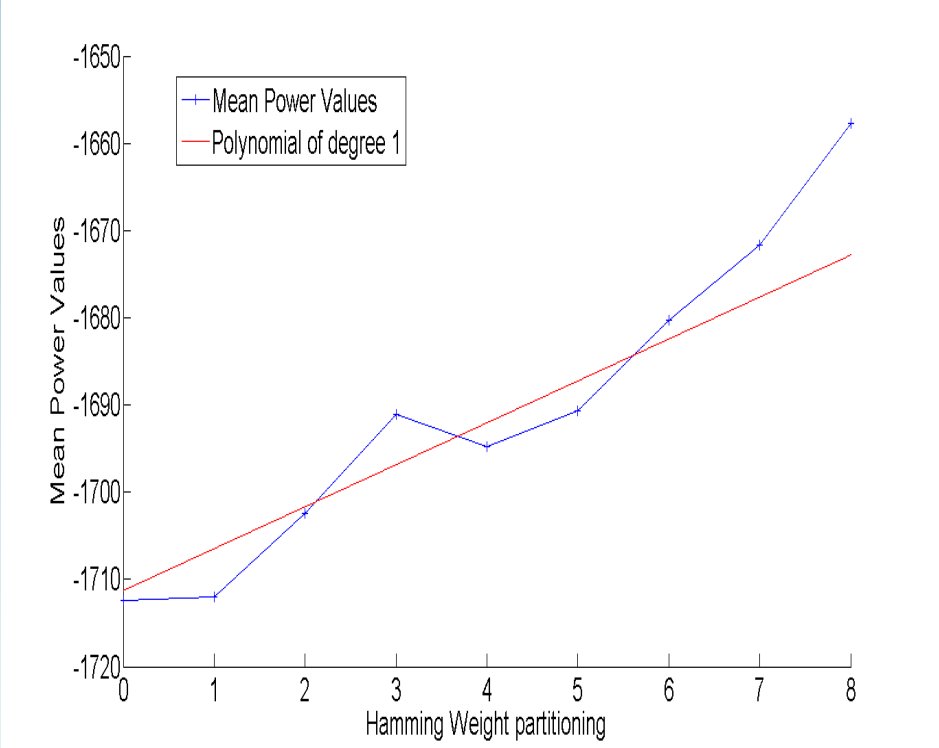
Context : known key, HD model, HW partitionning

For both test-cases, and for each S-box, the degree of the polynomial that minimizes the **AIC** is searched :

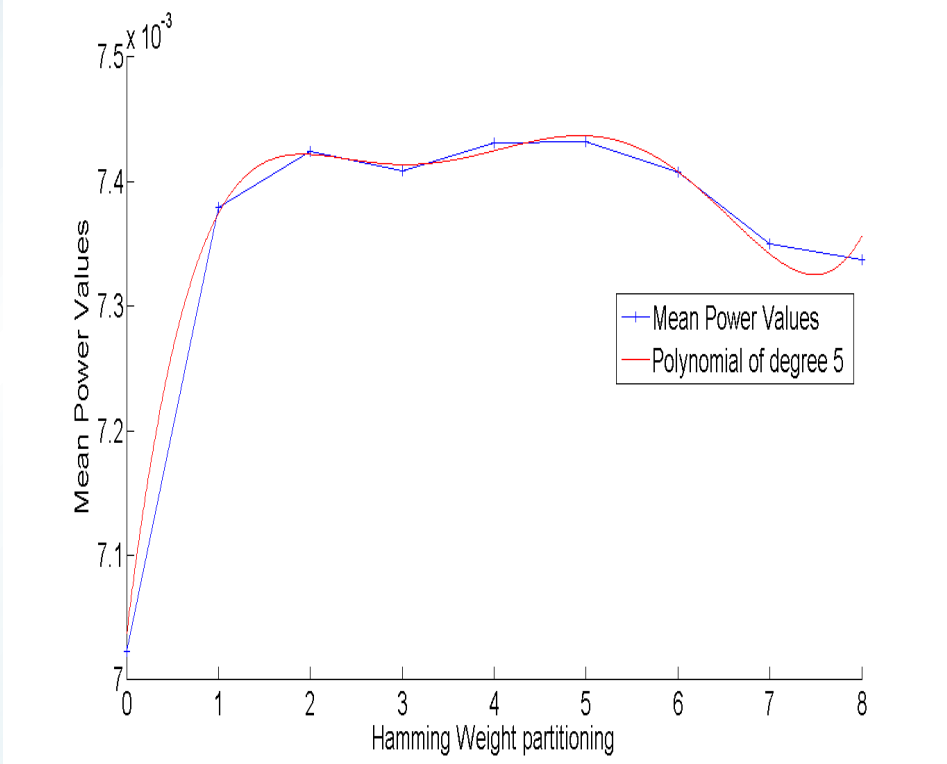
S-box	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
DPA-v2 AES	2	1	1	2	1	1	2	1	3	1	1	2	3	1	1	1
65nm AES	4	5	6	5	0	5	6	6	4	8	5	5	8	5	5	5

- **DPA-v2 AES : leakage close to linearity**
- **65nm AES : leakage far from linearity**

Leakage profiling

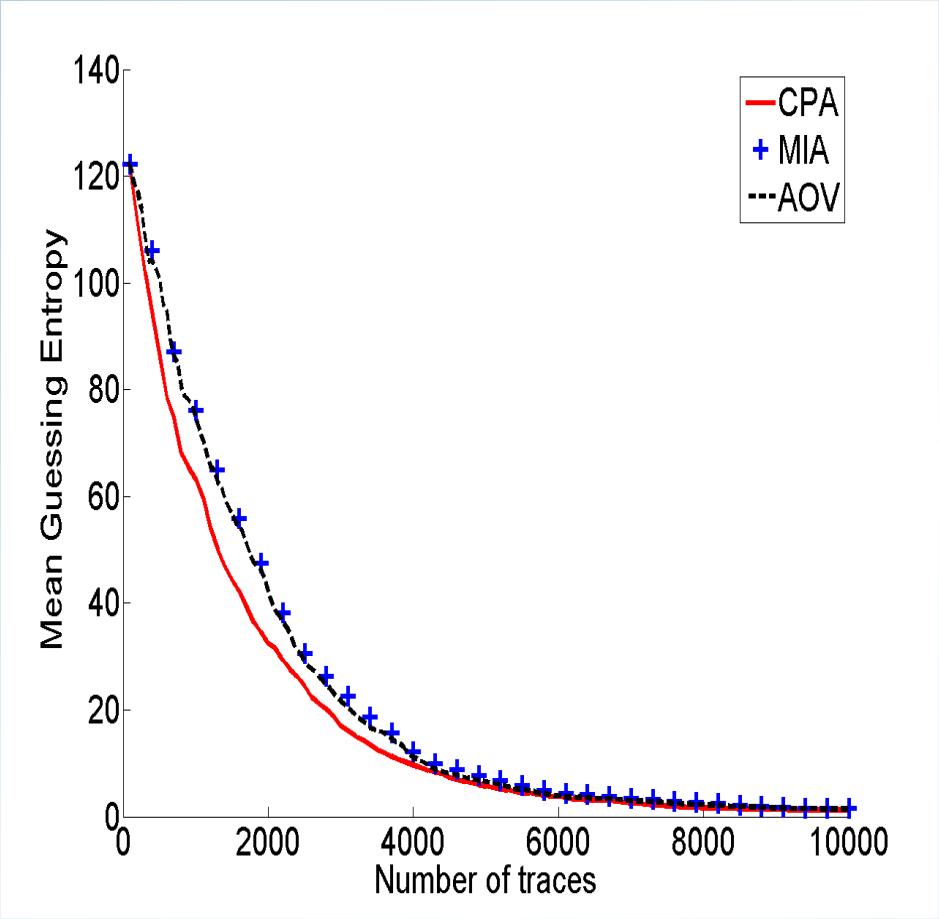


DPAv2-AES (20000 traces) - Polynomial fitting the best the leakage according to the AIC

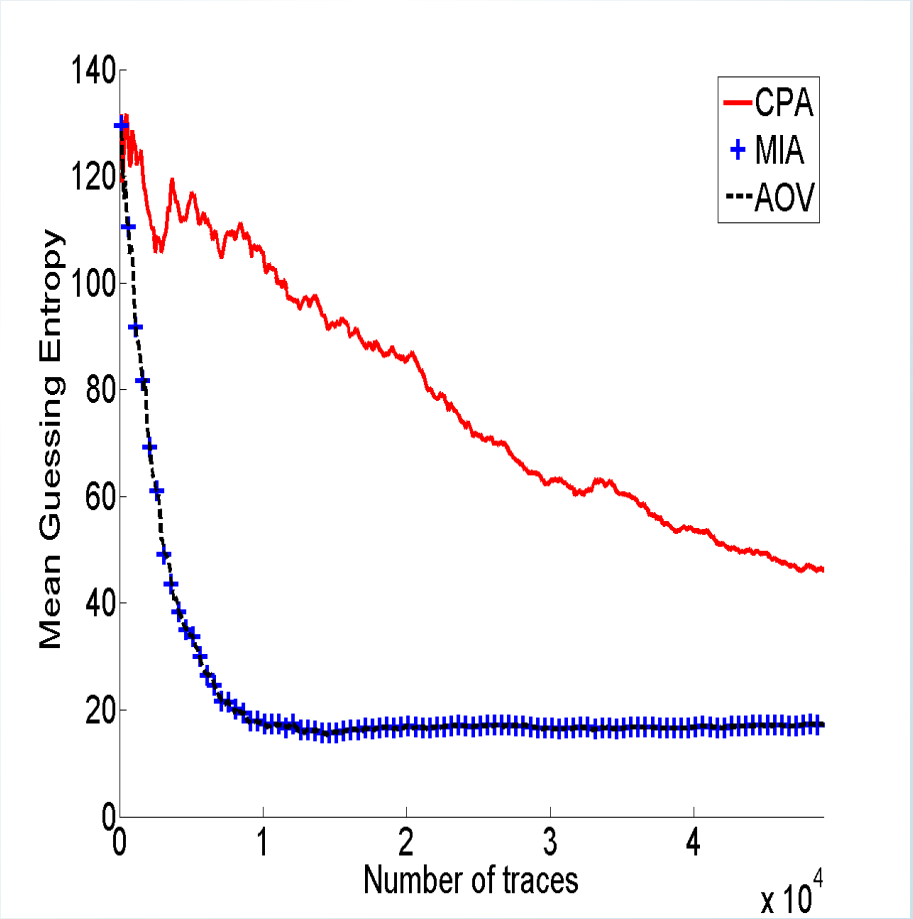


65nm-AES - Polynomial fitting the best the leakage according to the AIC

CPA vs AOV : number of traces

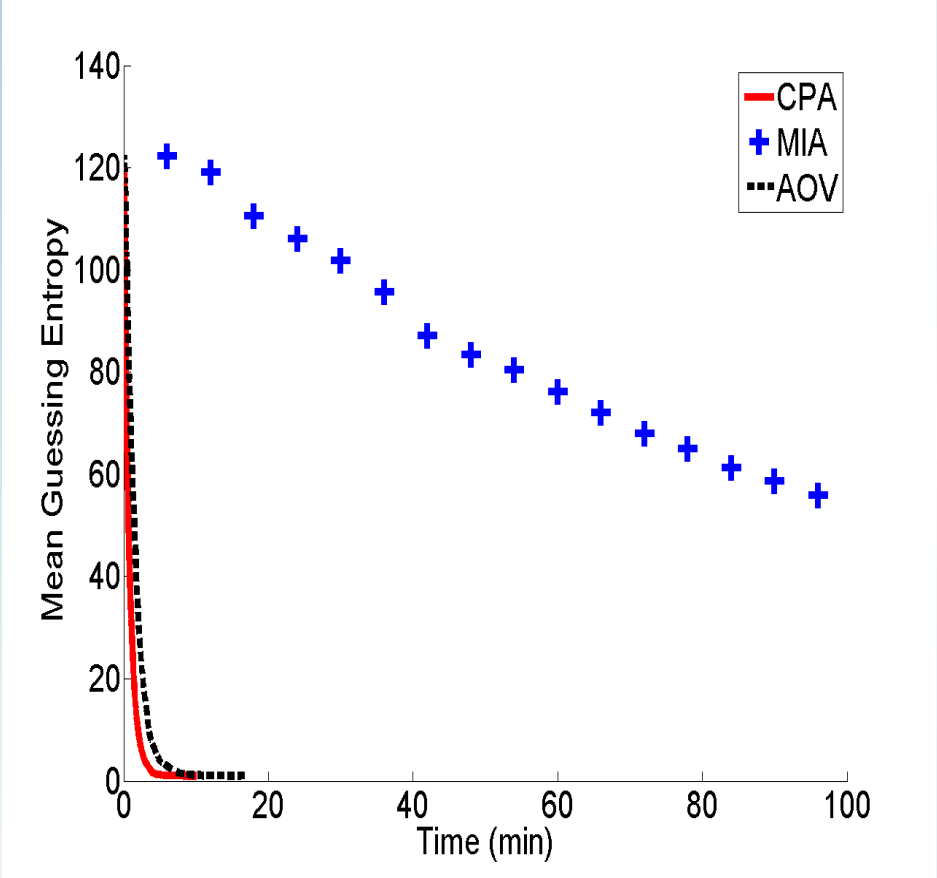


Mean Guessing Entropy obtained with CPA, AOV and MIA after the processing of 20000 traces of the DPAv2-AES

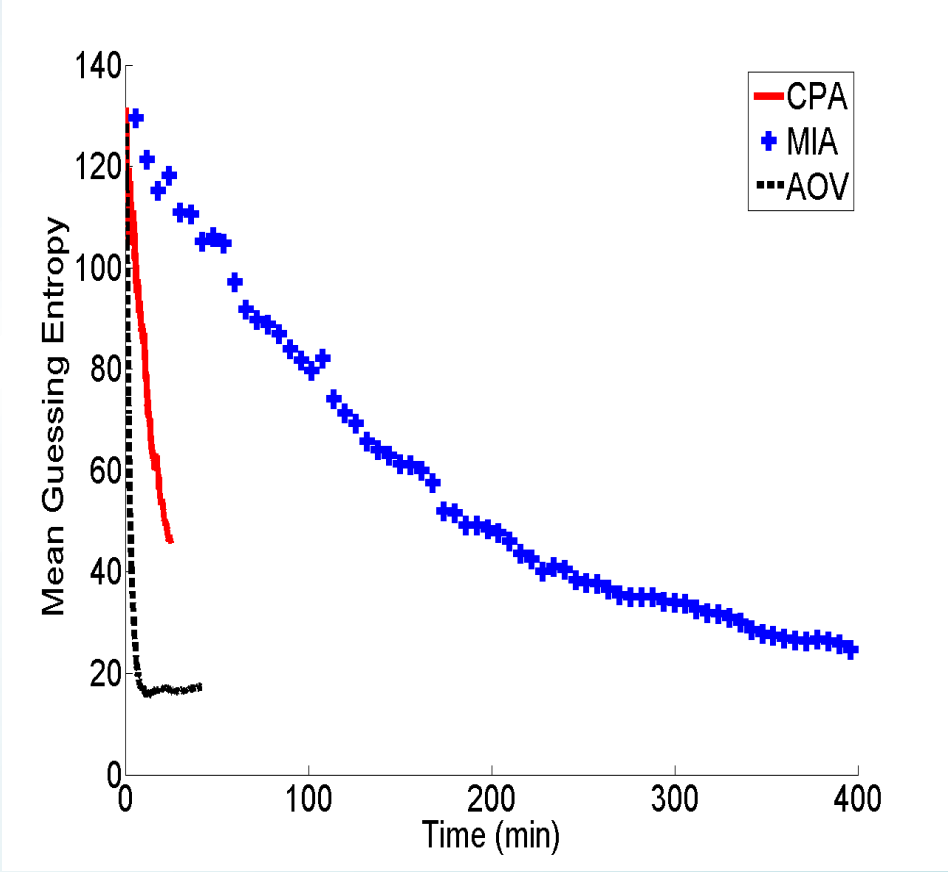


Mean Guessing Entropy obtained with CPA, AOV and MIA after the processing of 60000 traces collected above the 65nm-AES

CPA vs AOV : computation time



Mean Guessing Entropy obtained with CPA, AOV and MIA after the processing of 20000 traces of the DPAv2-AES, depending on the computation time



Mean Guessing Entropy obtained with CPA, AOV and MIA after the processing of 60000 traces collected above the 65nm-AES, depending on the computation time

Conclusion

AOV should be prefered to CPA :

- It provides **similar results** in case of a **linear leakage**
- It can **outperform** CPA in case of a **non-linear leakage** that is still on the means
- Its **computational cost is similar**

AOV remains **less generic than MIA**, but doesn't require the choice of hyperparameters, and is less time consuming.