Side-channel Analysis of Lightweight Ciphers: Current Status and Future Directions

Annelie Heuser Stjepan Picek
Sylvain Guilley Nele Mentens

IRISA/CNRS, Rennes, France

KU Leuven, ESAT/COSIC and iMinds, Kasteelpark Arenberg 10, bus 2452, B-3001 Leuven-Heverlee, Belgium

TELECOM-ParisTech, Paris, France and Secure-IC S.A.S., Rennes, France

Lightweight Cryptography Workshop 2016
October 17 – 18, 2016

Outline

1 Introduction

2 Background

3 Empirical Evaluation

4 Conclusions

Outline

1 Introduction

2 Background

3 Empirical Evaluation

4 Conclusions

Side-channel analysis

- Side-channel attack (SCA) represents a powerful category of attacks against cryptographic devices.
- SCA analyzes physical leakage that is unintentionally emitted during cryptographic operations in a device (e.g., power consumption, electromagnetic emanation).
- This side-channel leakage is statistically dependent on intermediate processed values involving the secret key, which makes it possible to retrieve the secret from the measured data by maximizing some statistical distinguisher.
- There exist non-profiled and profiled attacks.

Non-profiled side-channel analysis

- Correlation Power Analysis (CPA) is one of the most common non-profiled side-channel distinguishers that is also integrated in common criteria evaluations.
- For CPA in order to reveal the secret key k^* the attacker makes hypothetical predictions $Y(k)$ depending on a key guess k on the deterministic part of the leakage.
- For example, for each key hypothesis $k \in \mathbb{F}_2^n$ one has:

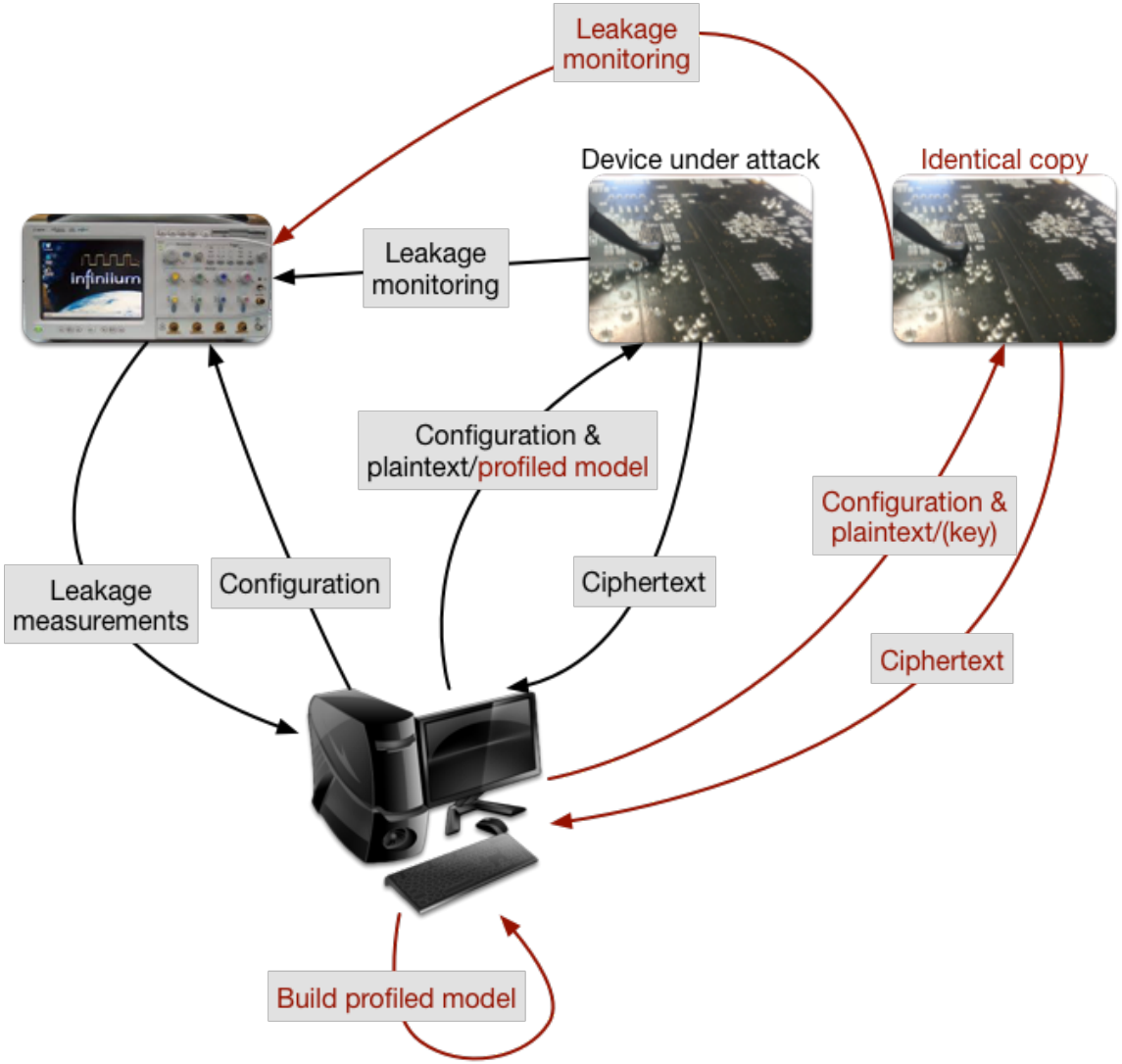
$$Y(k) = \mathrm{HW}(\mathtt{Sbox}[T \oplus k]).$$

- Given a set of Q leakage measurements $X_1, \ldots, X_Q$ corresponding to $T_1, \ldots, T_Q$ plaintexts, the attacker computes the correlation between the measurements and the hypothetical model.

Profiled side-channel analysis

- Profiled side-channel distinguishers assume that the attacker is able to possess an additional device to the one he want to attack, on which he has the freedom of nearly full control.
- Profiled attacks have a prominent place as the most powerful among side-channel attacks.
- Template attack is a well-known real-world attack that is also the most powerful attack from the information theoretic perspective.
- Machine learning (ML) also belongs to profiled attacks category (note that it is possible to run non-profiled ML for SCA).

Profiled side-channel analysis



Motivation

- Side-channel attacks on AES are well explored topic.
- However, lightweight ciphers are much less investigated from SCA perspective.
- There are many papers considering countermeasures, but very rare conducting extensive side-channel analysis.
- Moreover, there are many lightweight ciphers that are potentially interesting to explore.

Motivation

- When considering countermeasures, one must always ask how expensive are they and will it be possible to implement them in various constrained environments.
- Therefore, we investigate here what is the SCA resilience of a number of ciphers that do not have countermeasures.
- Still, they resilience differs due to inherent resistance of ciphers against SCA.
- Some of the measures describing that resilience are modified transparency order and **confusion coefficient**.

Motivation

- From one perspective, SCA for lightweight ciphers should be easier due to a lower number of possible classes when considering profiled attacks.
- On the other hand, since nonlinearity for S-boxes usually used in lightweight ciphers (4×4) can be maximally equal to 4, the difference between the input and the output of an S-box is much smaller than for instance in the case of AES.
- Therefore, one could conclude that SCA for lightweight ciphers must be more difficult than for standard ciphers.

Outline

1 Introduction

2 Background

3 Empirical Evaluation

4 Conclusions

- Ciphers with 4-bit S-box:
  - KLEIN
  - PRESENT
  - PRIDE
  - RECTANGLE
  - Mysterion
- Ciphers with 8-bit S-box:
  - AES
  - Zorro
  - Robin

- Our main targets are the weaknesses arising in software implementations on serial microprocessors.
- In these applications the Hamming weight (HW) and the Hamming distance (HD) leakage model are most commonly found in practice.
- More precisely, the loading and storing of data in memory (e.g., `S-box` calls) is usually causing a HW leakage, whereas the register updating (e.g., writing of intermediate round states) is causing HD leakage.
- Typically the HD is less significant than the HW, which is why we concentrate on a specific memory operation.

- The success rate of CPA depends on the number of measurements, SNR, and confusion coefficients $\kappa(k^*, k)$
- They describe the relationship between different hypothetical predictions $Y(k)$:

$$\kappa(k^*, k) = \mathbb{E}\left\{\left(\frac{Y(k^*) - Y(k)}{2}\right)^2\right\},$$

  where the expectation is taken over the plaintext/cipertext T.
- Considering $Y(k) = \mathrm{HW}(\mathtt{Sbox}[T \oplus k])$ the confusion coefficients depend on the choice of the S-box.

Introduction to ML

- Machine learning (ML) is a subfield of computer science that evolved from the study of pattern recognition and computational learning theory.
- Algorithms extract information from data, however, they also learn a model to discover something about the data in the future.
- Today, there exists a plenitude of ML algorithms to choose from.

Machine Learning

A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured with P, improves with experience E.

Outline

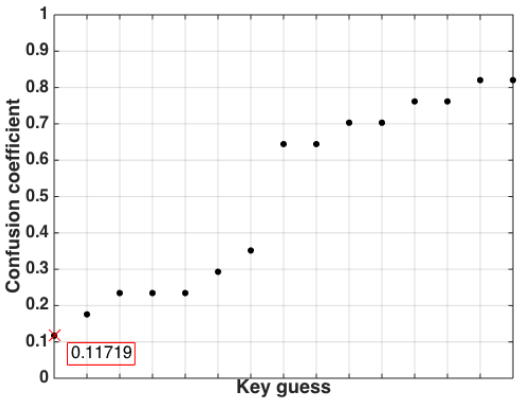
1 Introduction

2 Background

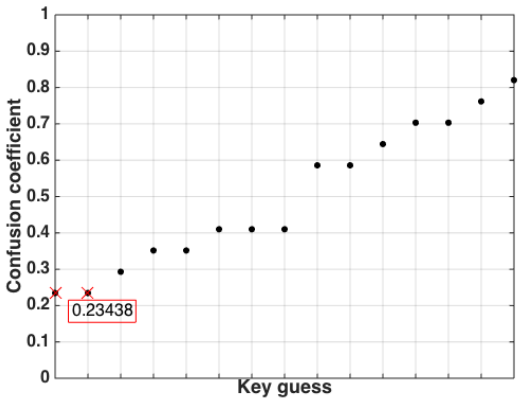
3 Empirical Evaluation

4 Conclusions

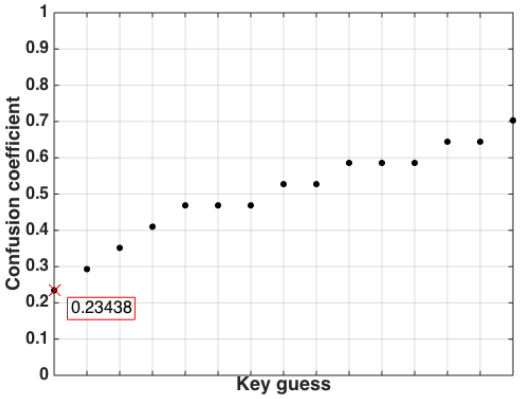
Confusion coefficient



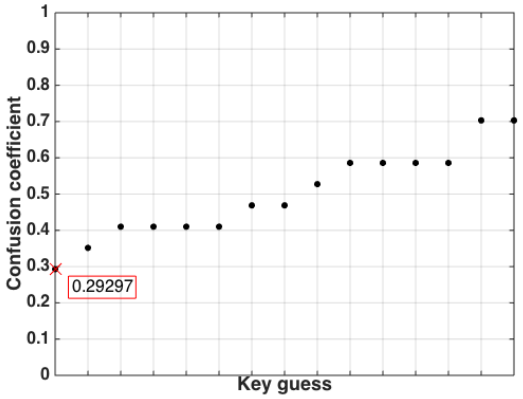
(a) KLEIN



(b) PRESENT

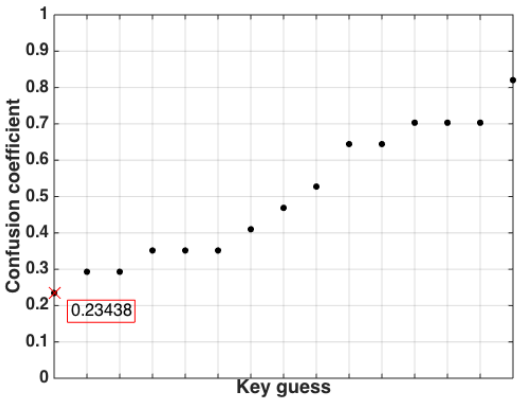


(c) PRIDE

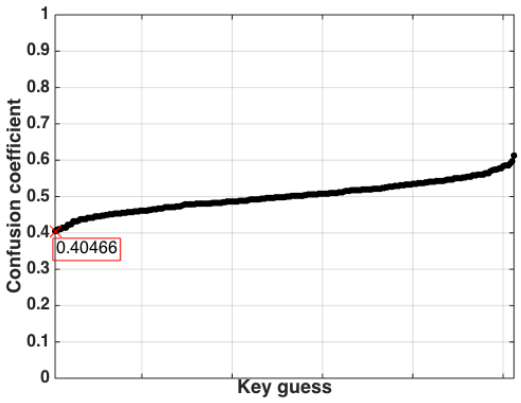


(d) Mysterion

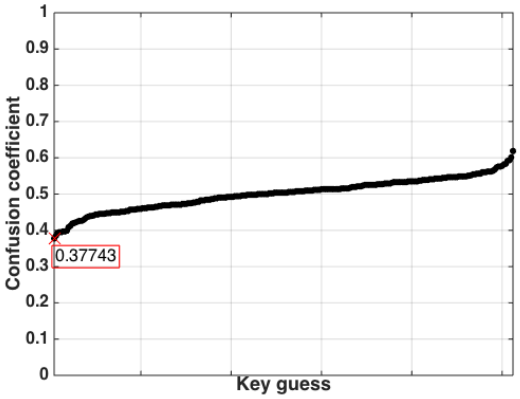
Confusion coefficient



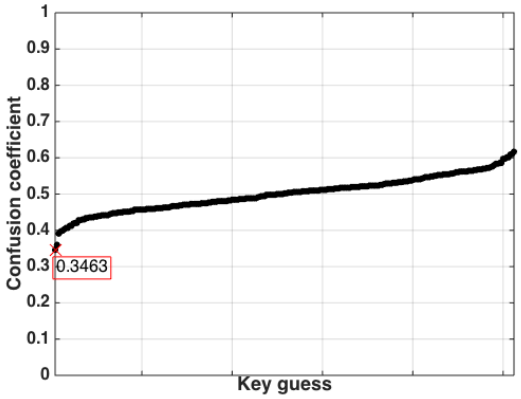
(a) RECTANGLE



(b) AES



(c) ZORRO



(d) Robin

Confusion coefficient

- The higher the minimum of the confusion coefficient the lower the side-channel security
- The smaller the variance of the confusion coefficient the lower the side-channel security

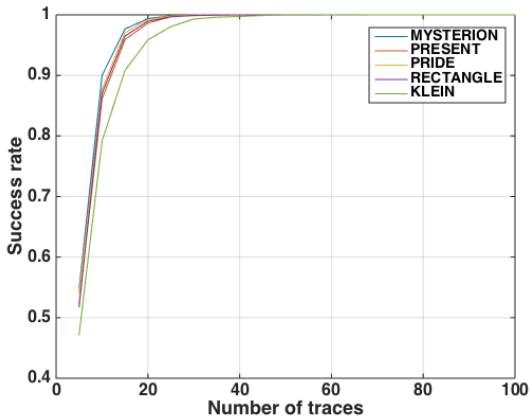
Table: Properties of $\kappa(k^*, k)$ (4-bit S-boxes)

	KLEIN	PRESENT	PRIDE	Mysterion	RECTANGLE
$\mathrm{var}(\kappa(k^*, k))$	0.071	0.038	0.018	0.015	0.035
$\min_k \kappa(k^*, k)$	0.117	0.234	0.234	0.292	0.234

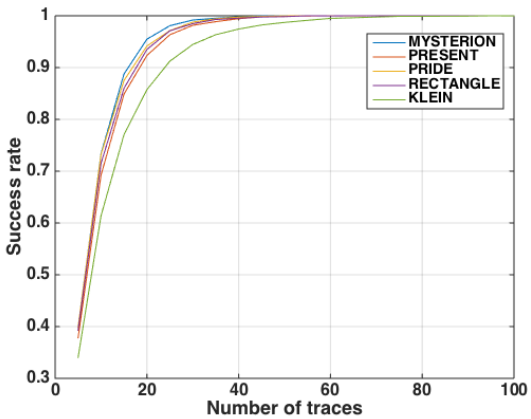
Table: Properties of $\kappa(k^*, k)$ (8-bit S-boxes)

	AES	ZORRO	Robin
$\mathrm{var}(\kappa(k^*, k))$	0.0017	0.0019	0.0023
$\min_k \kappa(k^*, k)$	0.4046	0.3774	0.3462

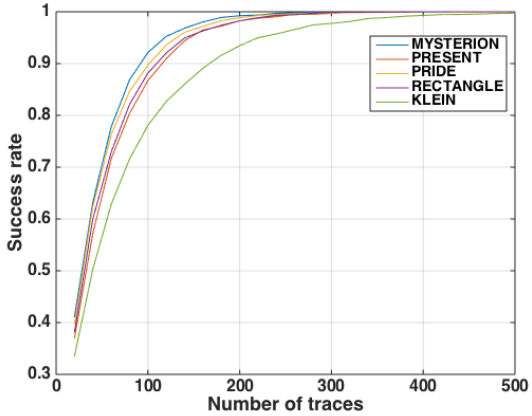
CPA



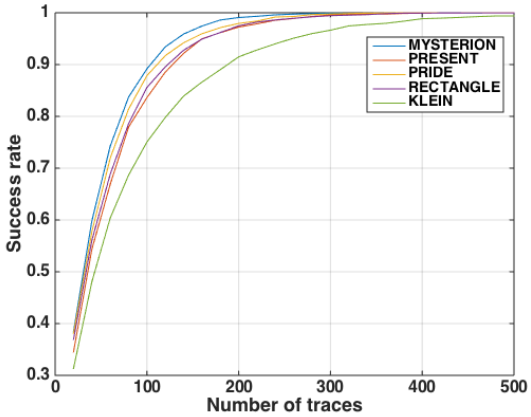
(a) $\sigma = \sqrt{1/2}$, SNR $= 2$



(b) $\sigma = 1$, SNR $= 1$

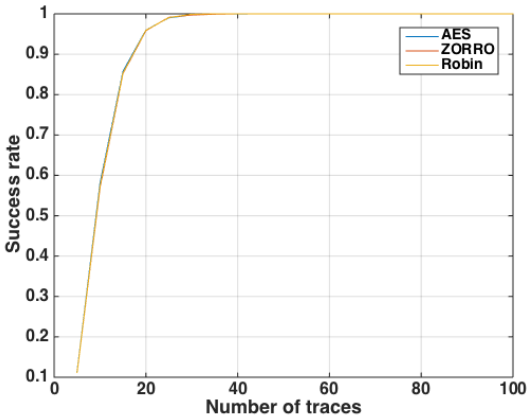


(c) $\sigma = \sqrt{8}$, SNR $= 1/8$

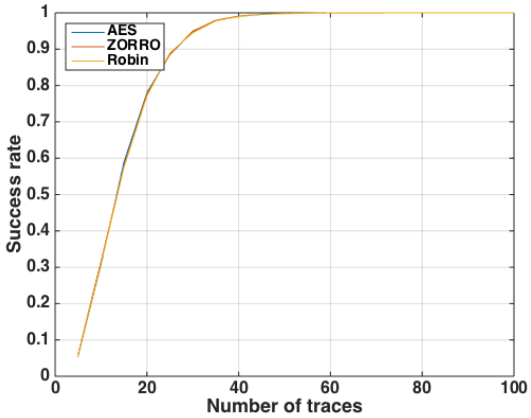


(d) $\sigma = 4$, SNR $= 1/16$

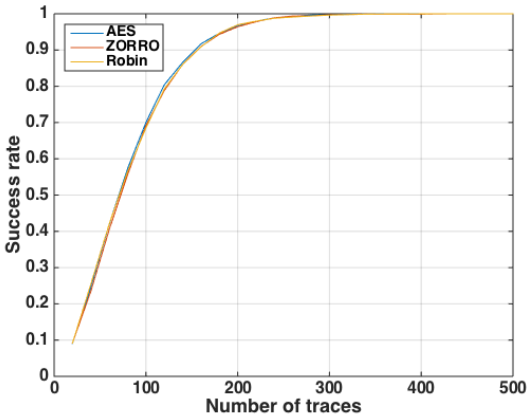
CPA



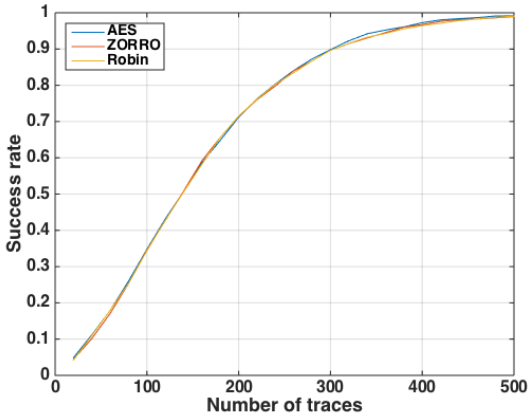
(a) $\sigma = 1$, SNR $= 2$



(b) $\sigma = \sqrt{2}$, SNR $= 1$



(c) $\sigma = 4$, SNR $= 1/8$



(d) $\sigma = \sqrt{32}$, SNR $= 1/16$

Introduction to ML

- Machine learning – Naive Bayes (also corresponds to template attack since we work with only a single feature), C4.5, and Multi-layered Perceptron.
- Training (profiling) and testing (attacking) phases with a data ratio of 2:1.
- The sizes of datasets are 10 000, 30 000, and 50 000.
- We compare PRESENT and AES.

Results

Table: Testing results for PRESENT

5 classes									
Algorithm	10k			30k			50k		
	σ_1	σ_3	σ_5	σ_1	σ_3	σ_5	σ_1	σ_3	σ_5
NB	51.27	38.55	37.12	51.17	38.57	37.1	51.04	38.92	37.81
C4.5	50.06	38.82	37.03	51.05	38.16	37.19	50.72	38.73	37.59
MLP	51.27	39.12	37.03	51.07	38.47	37.31	50.57	39	38
16 classes									
Algorithm	10k			30k			50k		
	σ_1	σ_3	σ_5	σ_1	σ_3	σ_5	σ_1	σ_3	σ_5
NB	41.55	19.94	12.06	42.62	18.68	13.86	41.72	18.53	14.04
C4.5	40.73	14.85	11.79	41.88	15.79	12.05	41.9	16.08	12.76
MLP	40.67	19.3	11.15	41.4	18.3	14.15	40.82	18.24	13.85

Results

Table: Testing results for AES

9 classes									
Algorithm	10k			30k			50k		
	σ_1	σ_3	σ_5	σ_1	σ_3	σ_5	σ_1	σ_3	σ_5
NB	27.67	27.63	28.18	27.07	27.04	27.52	27.94	27.93	28.04
C4.5	27.76	26.91	27.64	27.07	26.77	27.26	27.94	27.94	28.15
MLP	27.64	27.64	27.21	27.03	27.03	27.47	27.93	27.93	28.33
256 classes									
Algorithm	10k			30k			50k		
	σ_1	σ_3	σ_5	σ_1	σ_3	σ_5	σ_1	σ_3	σ_5
NB	38.33	12.67	7.42	37.43	13.04	8.23	38.84	13.29	8.47
C4.5	34.88	9.67	7.69	35.71	10.94	7.18	36.25	10.98	7.04
MLP	35.21	10.94	7.11	37.27	13	7.85	38.67	13.2	8.05

Outline

1 Introduction

2 Background

3 Empirical Evaluation

4 Conclusions

Conclusions

- In the case of non-profiled attacks, we see that the 8-bit S-boxes from AES, ZORRO, and Robin performed similar and we further cannot conclude that the 4-bit S-boxes are generally easier to attack than the 8-bit S-boxes.
- Ranking 4-bit: Mysterion, PRIDE, RECTANGLE, PRESENT, and most resistant KLEIN.
- When considering profiled attacks, our results show that attacking PRESENT is somewhat easier than attacking AES, the difference mainly stemming from the varying number of classes in one or other scenario.
- Still, that difference is not so apparent as one could imagine.
- This leaves us with a conclusion that attacking lightweight ciphers is not easy, but care should be taken if we consider attackers as powerful as for instance for the AES case.

Questions?

Thanks for your attention!

Q?