Lejla Batina[1], Benedikt Gierlichs[1], Kerstin Lemke-Rust[2]
[1] KU Leuven, Esat - Cosic, Belgium
[2] T-Systems GEI GmbH, Germany

Comparative Evaluation of

Rank Correlation based DPA

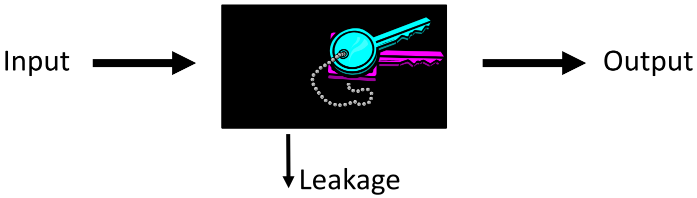
on an AES Prototype Chip

Outline

- Side-Channel Attacks and Differential Power Analysis (DPA)
- Overview of DPA variants
- Spearman Rank Correlation Coefficient
- Experiments and results
- Conclusion

Side-Channel Attacks

- Physical attacks ≠ Cryptanalysis
  (gray box, physics) (black box, maths)
- Do not tackle the algorithm's mathematical security



- Side-Channel leakage through:
  Timing, Power, EM, Light, Sound, Temperature, etc
- Observe physical quantities in the device's vicinity and use additional information during cryptanalysis

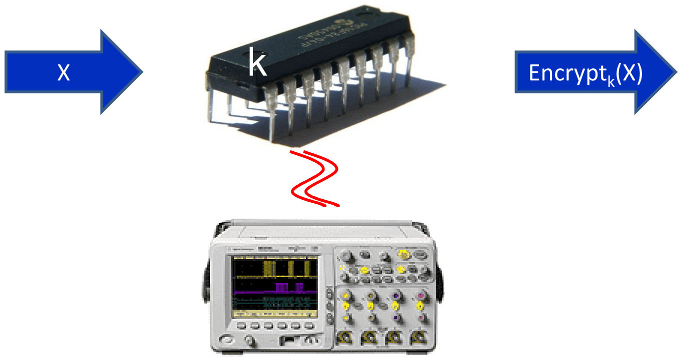
Principle is nothing new...

"Breaking into a vault is hard because one has to solve a single, very hard problem."

"Divide et impera"

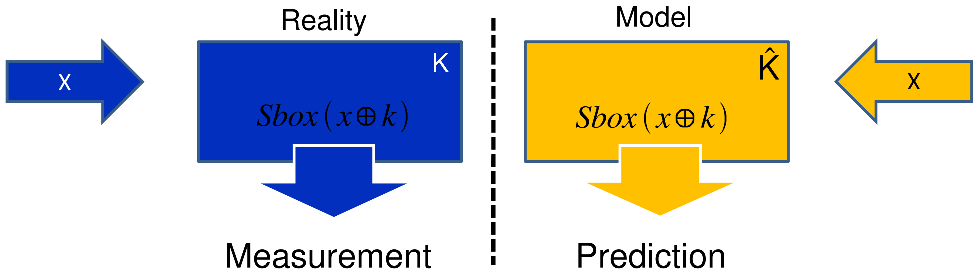
"Things are different if it is possible to solve many small problems instead..."

Power Analysis



- Measure power consumption during cryptographic computation
- Use measurements as Side-Channel to obtain information about device's internal state
  - Attack any chosen intermediate result of computation, e.g. attack the first round of a block cipher!

Differential Power Analyis



- Take measurements during encryption of several inputs x
- Choose a sensitive intermediate result, e.g. $Sbox(x \oplus k)$
- For each possible sub-key candidate $\hat{K}$, e.g. one byte
  - Predict Side-Channel leakage for all measurements
  - Apply statistics to test whether prediction and measurement match
- The key candidate that matches best is most likely correct

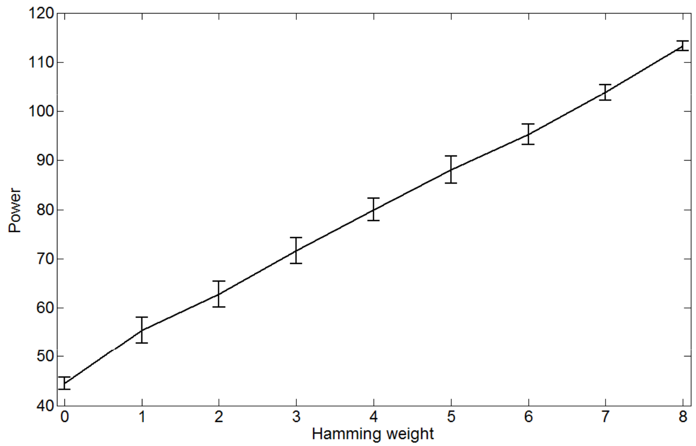
DPA Overview

- 1999: Single-bit DPA with Difference of Means test
  - Predict a single bit of the intermediate result: 0 vs 1
  - Divide power measurements in two sets and check the difference between the mean values of the two sets
  - Wrong key guess: measurements are assigned to sets at random and the means of the sets will not differ
  - Correct key guess: the measurements are partitioned correctly and the means of the sets differ
- 2000: Multi-bit DPA with Difference of Means test
  - Similar, but predict several or all bits of the intermediate result
  - Use 'most different' sets, e.g. '0000' and '1111', for test and discard all other measurements
  - Better Signal to Noise ratio (SNR) but more measurements required

DPA Overview cont'd

- 2004: Power model and Pearson Correlation
  - Predict several or all bits of the intermediate result(s)
  - Choose a power model, often linear in Hamming weight (HW), e.g. $P = a \cdot HW(data1 \oplus data2) + b$
  - Predict power consumption for all measurements and estimate (linear) Pearson correlation coefficient
  - Wrong key guess: the predictions are random and do not correlate with the measurements
  - Correct key guess: the predictions are 'correct' and correlate with the measurements
  - This is the 'default-attack' and maybe the state-of-the-art
    + Efficient and robust
    - Meaningful power model required

Power consumption of smart card µC



- Different HWs clearly distinguishable, almost perfectly linear
- Good SNR, relatively small standard deviation

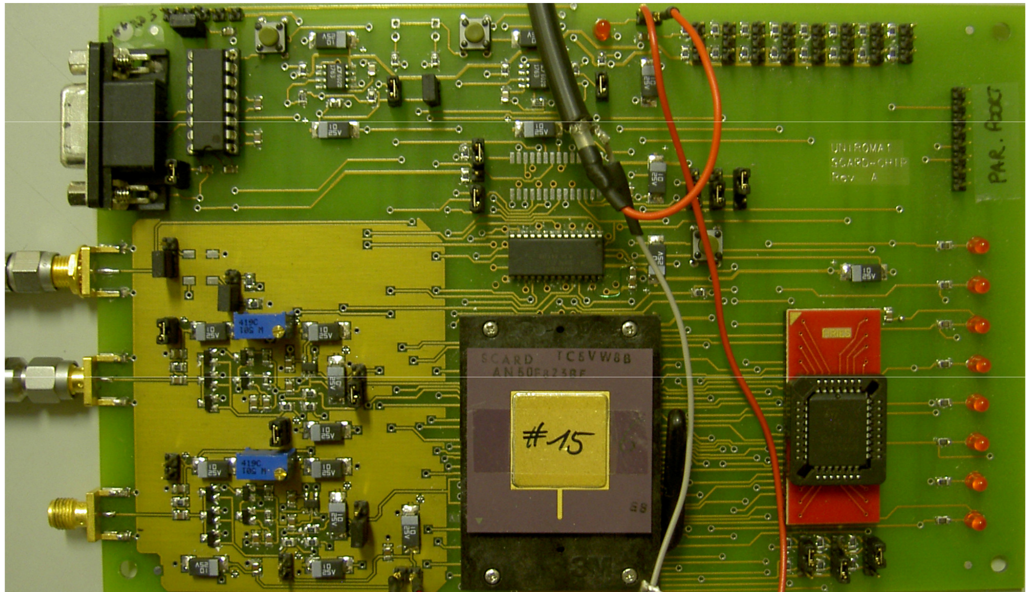
DPA Overview cont'd

- 2003, 2005: DPA attacks with device profiling
  - Consider a more powerful adversary
  - Adversary has unrestricted access to a training device in his control
  - Training device is identical to target device w.r.t. to Side-Channel
  - Profiling step: characterize the Side-Channel leakage of the training device; generate 'fingerprints' of different key dependencies
  - Classification step: obtain measurements from target device and use the 'fingerprints' to classify them (maximum likelihood)
  - Most powerful attack from an information theoretic point of view

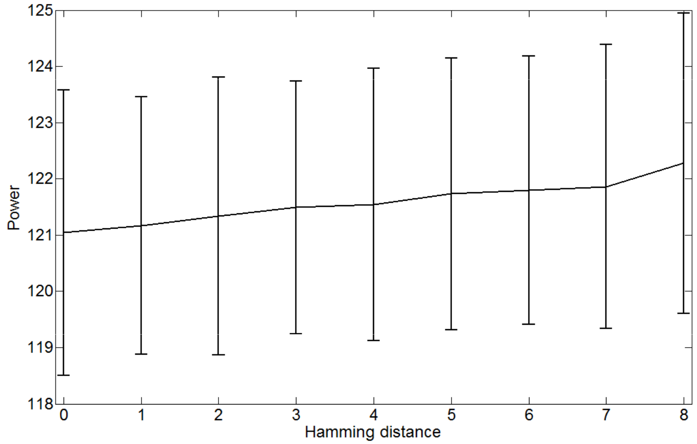
AES hardware module

- Side Channel Attack Resistant Design (SCARD) project
- By European Commission
- 5 European universities, 4 industrial partners
- The SCARD chip:
  - Focus here on 8051 µC with AES coprocessor in .13 µm sCMOS technology

AES hardware module



Power consumption of the AES hardware module in sCMOS



- Not strictly linear but monotonically increasing
- Bad SNR, relatively large standard deviation

Observations

- Power consumption behaviour not strictly linear
- Linear power model ➔ sub-optimal attack
- Confirms the widespread intuition that power models are
  - Not generic
  - Device dependent
  - A major issue in Differential Power Analysis

Non parametric statistics

- Example: Spearman Rank Correlation Coefficient
- Essentially: Pearson correlation coefficient on ranked data
  - Each data is assigned a rank, i.e. its position in the ordered set
  - Ranking reduces the information to > = <
  - Pearson correlation coefficient on ranked data captures also non-linear relations as long as they are monotonic
- For formulae and details, see the paper

Experimental design and metrics

- Power measurements of 50k encryptions
- Key recovery attack for one byte of the AES key
- Compare attack efficiency of
  - Single-bit and multi-bit DPA
  - Pearson correlation coefficient
  - Spearman correlation coefficient
  - Template Attack and Stochastic Model

- Successful attack := key byte correctly recovered
- Efficiency metric := number of measurements required
- Success rates computed from 500 experiments

Experimental Results

No.	DoM	Pearson Corr.	Sp. Rank. Corr.	Template Attack	Stochastic Model
500	-	13.6%	39.6%	15.6%	41.4%
1000	-	29.8%	77.8%	31.8%	73.4%
2000	-	64.2%	99.0%	63.2%	96.8%
3000	-	84.0%	100.0%	82.4%	100%

- Proposed attack outperforms other attacks considered in almost all settings
- In particular
  - Single and multi bit DPA: no positive results at all
  - Pearson correlation coefficient requires up to 3 times more measurements
  - Attacks with device profiling step are less efficient!

Conclusions

- Non-parametric statistics as a new class of Side-Channel distinguishers
- Experimental evaluation of the efficiency and comparison to the state-of-the-art
- Two main observations:
  - Spearman rank correlation outperforms Pearson correlation
    - Power model that is linear in the bitflips is suboptimal
    - Observation naturally bound to this specific target platform
  - Attacks with profiling step do not perform significantly better
    - Requires further investigation

Thanks for your attention!

Questions

?

benedikt.gierlichs@esat.kuleuven.be