Security Research
Between Attack and Design

Annelie Heuser[1,4], Michael Kasper[2,4], Werner Schindler[3,4], Marc Stöttinger[1,4]

1: TU Darmstadt: Integrated Circuits and Systems Lab, Darmstadt

2: Fraunhofer SIT

3: Bundesamt für Sicherheit in der Informationstechnik (BSI), Bonn

4: CASED (Center for Advanced Security Research Darmstadt)

Bochum, June 2011

Outline of the talk

- Motivation
- Stochastic approach (brief reminder)
- Several variants
- Design applications
    - Interpretation of a β-characteristic
    - Symmetry considerations
- Conclusion

Constructive side-channel analysis

- A successful attack proves that there is a weakness but this information may not suffice to fix the vulnerability.
- <u>Desirable:</u> constructive attacks, which
  - point to the source of the leakage
  - (ideally) quantify the leakage

  → support target-oriented (re-)design

  → allow interaction between attack and design

The stochastic approach

- target: block cipher
- exploits power measurements at several time instants $t_1 < t_2 < \ldots < t_m$
- The measurement values are interpreted as values that are assumed by random variables.
- The stochastic approach combines engineers' expertise with efficient stochastic methods from multivariate statistics.

Bundesamt
für Sicherheit in der
Informationstechnik

CASED

The stochastic model (basic variant)

target algorithm: block cipher (e.g., AES; no masking)

$x \in \{0,1\}^p$ (known) part of the plaintext or ciphertext

$k \in \{0,1\}^s$ subkey **[AES: (typically) s = 8]**

t time instant

$$I_t(x,k) = h_t(x,k) + R_t$$

Random variable
(depends on x and k)

quantifies the random-
ness of the side-channel
signal at time t

deterministic part
= leakage function
(depends on x and k)

Random variable
$E(R_t) = 0$

Noise (centered)

The stochastic model (masking case)

$x \in \{0,1\}^p$ (known) part of the plaintext or ciphertext

$z \in M$ masking value

$k \in \{0,1\}^s$ subkey [AES: (typically) s = 8]

$t \in \{t_1,t_2,...,t_m\}$ time instant

$$I_t(x,z;k) = h_t(x,z;k) + R_t$$

Random variable
(depends on x,z,k)

quantifies the random-
ness of the side-channel
signal at time t

deterministic part
= leakage function
(depends on x,z,k)

Random variable
$E(R_t) = 0$

Noise (centered)

Profiling, Step 1 (I)

- Fix a subkey $k \in \{0,1\}^s$.
- The unknown function

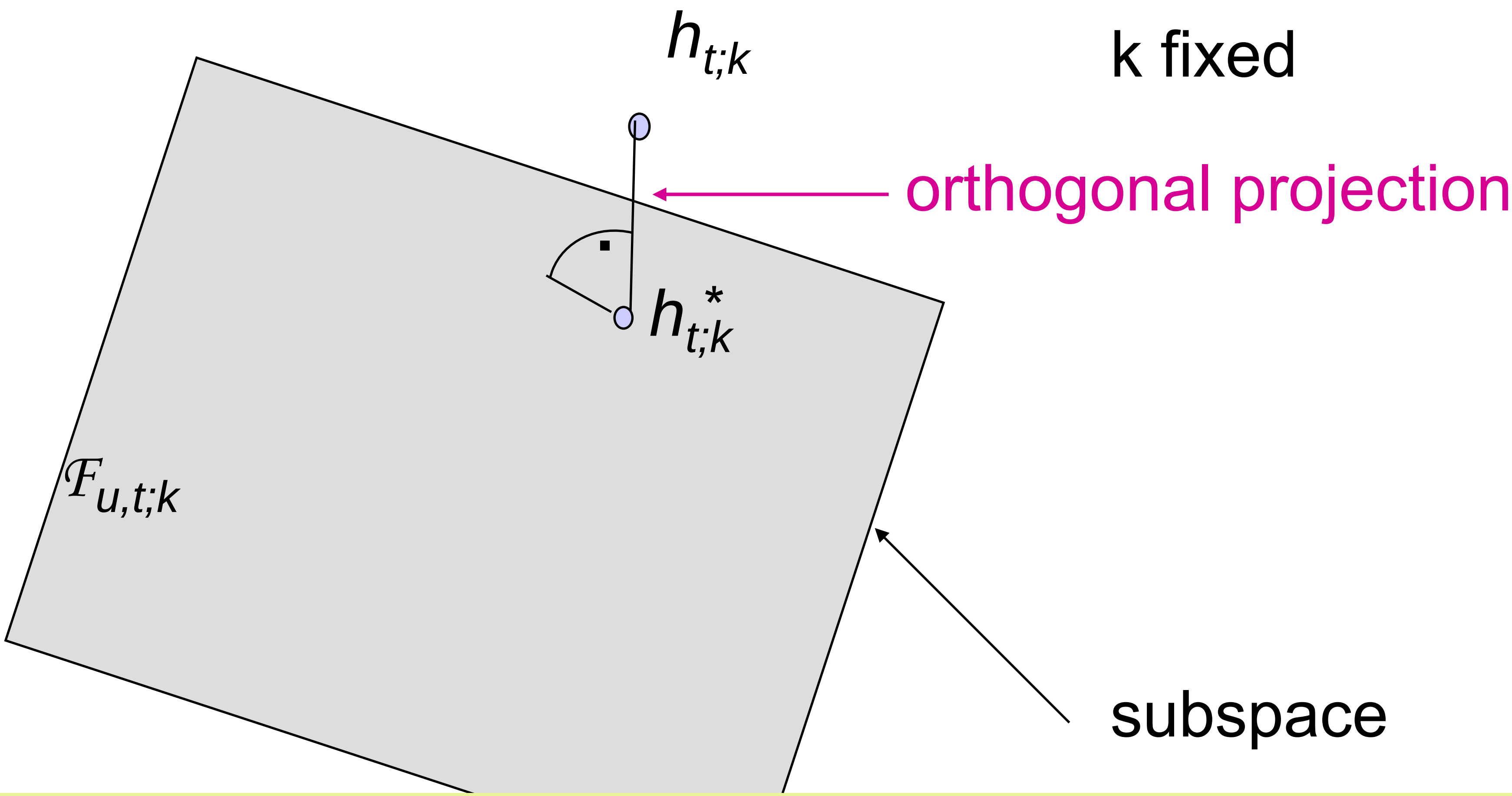
  $$h_{t;k}\colon \in \{0,1\}^p \times M \times \{k\} \to R,\ h_{t;k}(x,z;k) := h_t\,(x,z;k)$$

  is interpreted as an element of a high-dimensional real vector space $\mathcal{F}$. In particular, $\dim(\mathcal{F}) = 2^p\,|M|$.
- Goal: Approximate $h_{t;k}$ by its image $h^*_{t;k}$ under the orthogonal projection onto a suitably selected low-dimensional vector subspace $\mathcal{F}_{u,t;k}$

Bundesamt für Sicherheit in der Informationstechnik

CASED

Geometric illustration



k fixed

The image $h^*_{t,k}$ is the best approximator of $h_{t;k}$ in $\mathcal{F}_{u,t;k}$

Bundesamt
für Sicherheit in der
Informationstechnik

CASED

Profiling, Step 1 (II)

$$\mathcal{F}_{u,t;k} := \{h' : \{0,1\}^p \times \underset{\substack{\uparrow \\ \text{(only for masking)}}}{M} \times \{k\} \to R \mid \sum_{j=0}^{u-1} \beta'_{j,t;k}\, g_{j,t;k} \text{ with } \beta'_{j,t;k} \in R\}$$

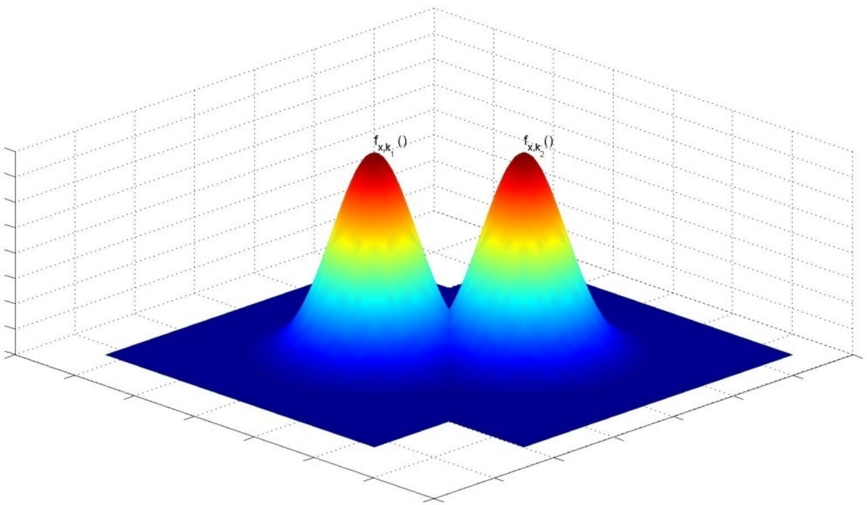
with basis functions $g_{j,t;k} : \{0,1\}^p \times M \times \{k\} \to R$

The basis $g_{0,t;k},\ldots,g_{u-1,t;k}$ shall be selected under consideration of the attacked device.

The estimation of $h^*_{t,k}$ can completely be moved to the low-dimensional subspace $\mathcal{F}_{u,t;k}$, which reduces the number of measurements to a small fraction.

Profiling, Step 2

- Estimate the covariance matrix C of the noise vector $(R_{t1}, \ldots, R_{tm})$ (multivariate normal distribution)
- $\rightarrow$ probability densities $\tilde{f}_{x,z;k}(\cdot)$ for $(I_{t\_1}(x,z;k), \ldots, I_{t\_m}(x,z;k))$.



Example (m = 2)

Attack phase (analogous to template attacks)

- Perform N_3 measurements on the target device
- (Maximum-likelihood principle) Decide for that subkey k*, which maximises

$$\prod_{j=1}^{N_3} \sum_{z' \in M} \Pr(Z_j = z') \tilde{f}_{x_j, z'; k}(i_j)$$

- For non-masked implementations this formula simplifies to

$$\prod_{j=1}^{N_3} \tilde{f}_{x_j; k}(i_j)$$

Bundesamt für Sicherheit in der Informationstechnik

CASED

Principal Component Analysis (PCA)

- Profiling, Step 2: The covariance matrix C may be 'almost' singular. This might cause large estimation errors for C^{-1}.
- Thus it is often advisable to consider only the information 'contained' in a subspace of R^m that is spanned by the eigenspaces of C, which belong to its largest eigenvalues: i.e., consider $(P^t\ C\ P)$ for a suitable transformation matrix P instead of C.
- Typically, only 1-3 eigenvalues are 'relevant'.

Bundesamt
für Sicherheit in der
Informationstechnik

CASED

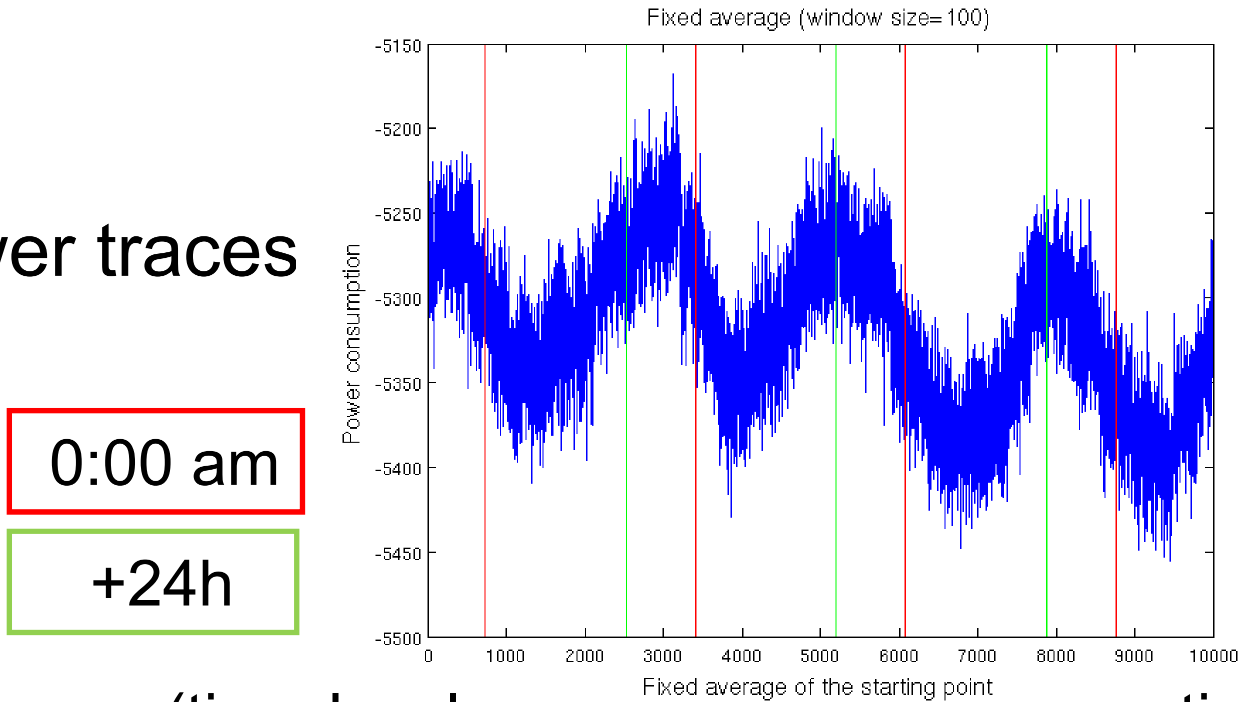
Variants of the stochastic approach (I)

Besides the “standard method” described on the previous slides there exist further variants:

- **Profiling with unknown masking values:** possible but less efficient than with knowledge of the masking values since all time instants have to be handled simultaneously
- **Attacking without profiling:** possible

Variants of the stochastic approach (II)

- Within long measurement series the environmental conditions might change, influencing the power consumption and thereby violating the (silent) assumption of having identical conditions all the time.
- Example:
  dpa-v2 power traces



(time-local average power consumption)

Bundesamt
für Sicherheit in der
Informationstechnik

CASED

Variants of the stochastic approach (III)

- Problem: There is a drifting offset that affects the effectiveness of profiling based attacks (template attacks, stochastic approach).
- Solution: new variant of the stochastic approach that tolerates those effects (submitted).

Newest results (dpa-v2 traces)

- Criterion: partial success rate > 80 %:
  - Contest winner: 5890 traces, 2nd rank: 7515 traces
  - Our new method: (checked only at the public base!): 4117 traces
- Criterion: global success rate > 80 %
  - Contest winner: 7061 traces, 2nd rank: 10666 traces
  - Our new method (checked only at the public base!): 6705 traces
- Note: The best results were gained with a 93-dimensional subspace.

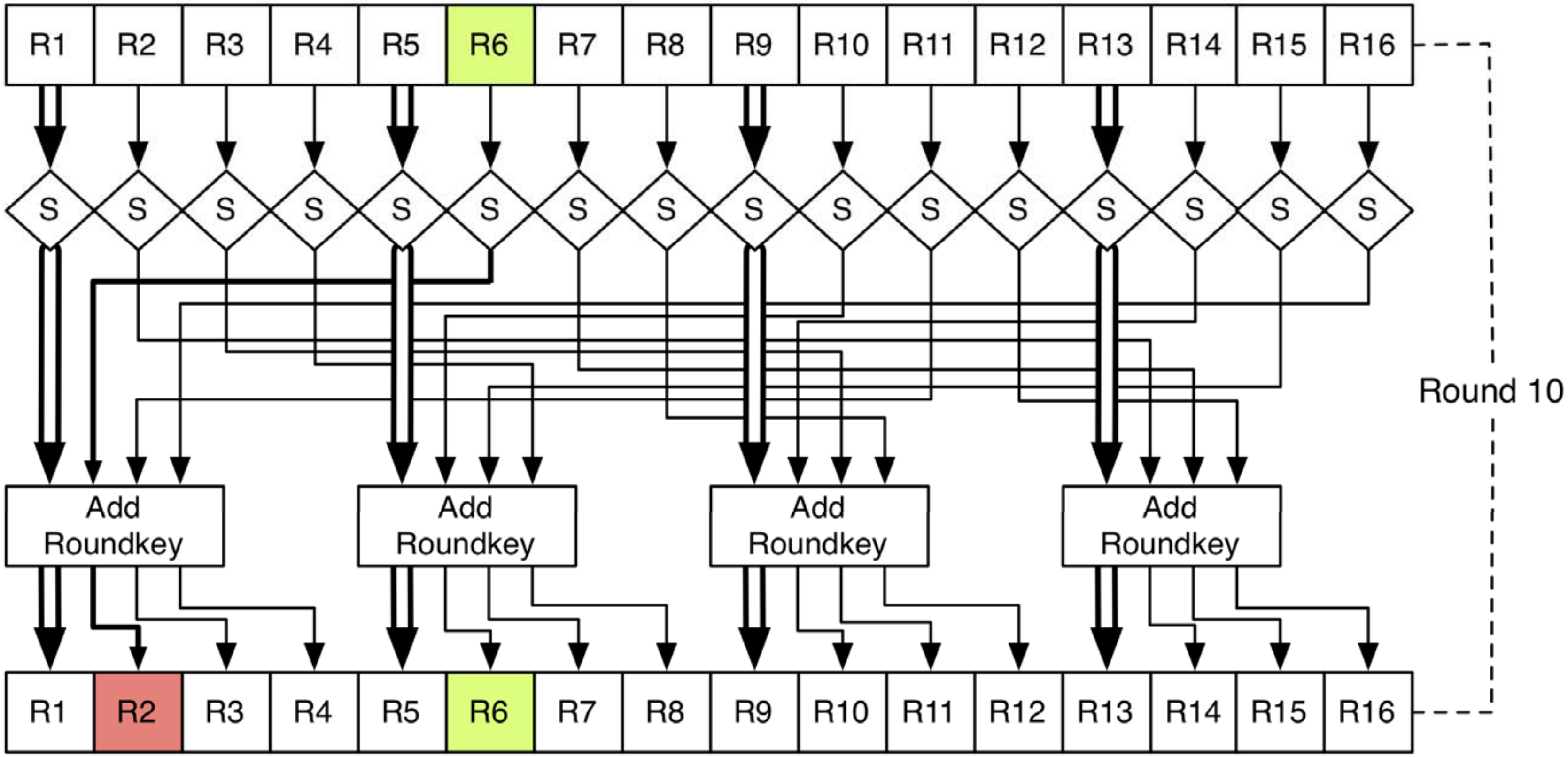
Bundesamt
für Sicherheit in der
Informationstechnik

CASED

Note

- For design purposes only profiling step 1 is relevant.

All experiments were performed on the **SASEBO FPGA evaluation board**



AES implementation on an FPGA

Target: key byte $k_{(2)} \in \{0,1\}^8$ in Round 10

$R_{(x)}$ value of byte register x after Round 10

9-dimensional vector subspace:

$g_{0,t;k(2)}((R_{(2)},R_{(6)}),k_{(2)}) = 1$

$g_{j,t;k(2)}((R_{(2)},R_{(6)}),k_{(2)}) = 2((R_{(6)} \oplus S^{-1}(R_{(2)} \oplus k_{(2)}))_j - 0.5)$

for $1 \le j \le 8$

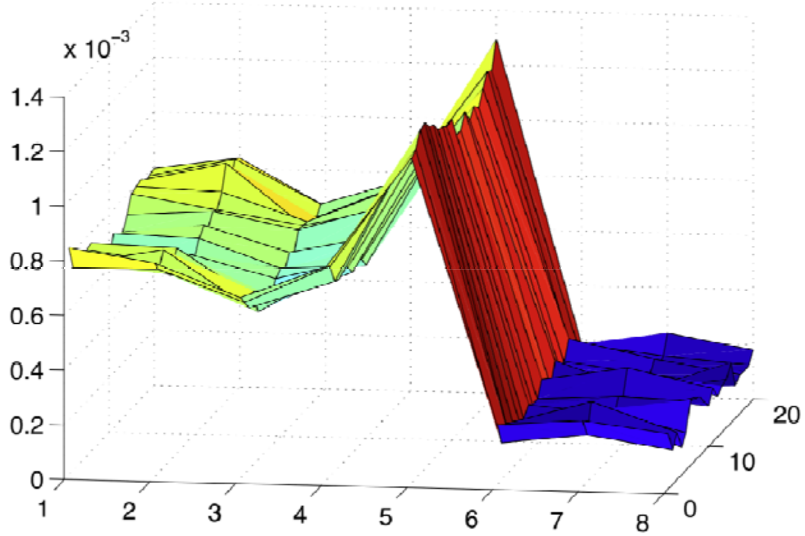
NOTE: Factor ‘2’ and summand ‘ – 0.5 ‘ imply that the basis is orthonormal.

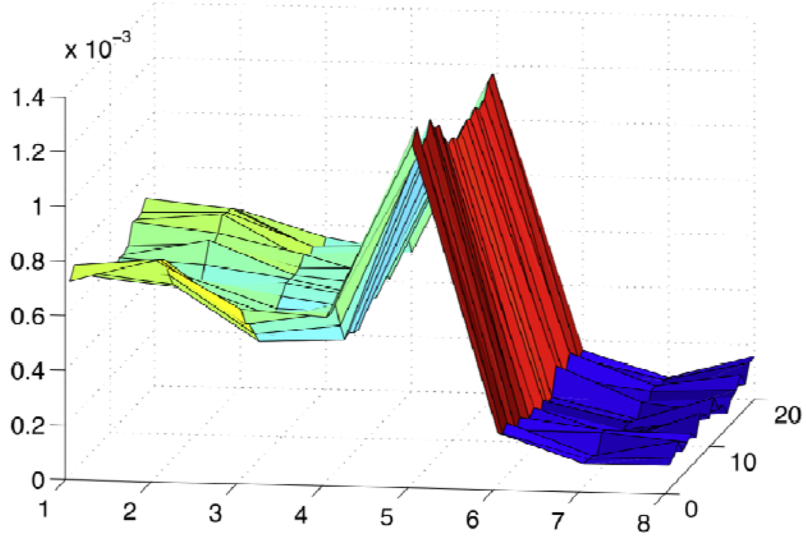
- Let

$$h^*_{t;k}(x,k) = \sum_{j=0}^{u-1} \beta^*_{j,t;k} g_{j,t;k}(x,k)$$

- A large coefficient $|\, \beta^*_{j,t;k} \,|$ for some j > 0 means that the ‘direction’ of the basis vector $g_{j,t;k}$ has considerable impact on the subkey-dependent part of the leakage $h_{t;k}$.

β-characteristics of an S-Box design (FPGA, TBL)





AES TBL, $k_{(1)} = 19$:

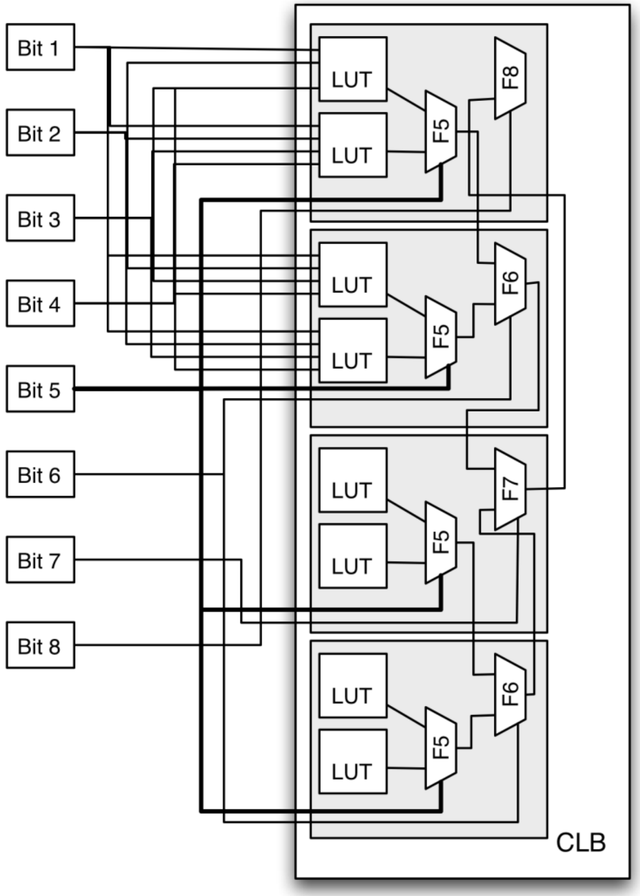
$|\beta_1|,\ldots,|\beta_8|$ for $t_1,\ldots,t_{20}$

AES TBL, $k_{(1)} = 209$:

$|\beta_1|,\ldots,|\beta_8|$ for $t_1,\ldots,t_{20}$

plots for two different keys and several time instants

Brief analysis of the implementation (FPT 2010)



- Part of the SBox implementation after the synthesis process and the Place & Route process
- The first layer of the multiplexer network is switched by the 5th bit
- Different propagation delays caused by LUT to the multiplexer produces data-dependent glitches.
- This implies bit-specific higher power consumption.

(The design was synthesized for the Virtex-II pro family.)

Remark

- Generally speaking, higher-dimensional subspaces $\mathcal{F}_{u,t;k}$ may provide more precise leakage models (work in progress).
- Another important question remains: Is the choice of the basis vectors appropriate?

Leakage model $\mathcal{B}$

The basis vectors

$$g_{0,t;k(2)}\,((R_{(2)},R_{(6)}),k_{(2)}) = 1$$
$$g_{j,t;k(2)}\,((R_{(2)},R_{(6)}),k_{(2)}) = 2((R_{(6)} \oplus S^{-1}(R_{(2)} \oplus k_{(2)}))_j - 0.5) \quad \text{for } 1 \le j \le 8$$

depend on $((R_{(2)},R_{(6)}),k_{(2)})$ only through

$\Phi((R_{(2)},R_{(6)}),k_{(2)}) := R_{(6)} \oplus S^{-1}(R_{(2)} \oplus k_{(2)})$

(symmetry assumption $\mathcal{B}$)

The same symmetry property holds for

$h^*_{t,k(2)}\,((R_{(2)},R_{(6)}),k_{(2)})$ and $\tilde{h}^*_{t,k(2)}\,((R_{(2)},R_{(6)}),k_{(2)})$

Bundesamt für Sicherheit in der Informationstechnik

CASED

Alternate leakage model $\mathcal{A}$

The basis vectors

alternate 9-dimensional vector subspace:

$$g'_{0,t;k(2)}((R_{(2)},R_{(6)}),k_{(2)}) = 1$$

$$g'_{j,t;k(2)}((R_{(2)},R_{(6)}),k_{(2)}) = 2\,((S^{-1}(R_{(2)} \oplus k_{(2)}))_j - 0.5) \quad \text{for } 1 \leq j \leq 8$$

depend on $((R_{(2)},R_{(6)}),k_{(2)})$ only through

$$\Phi'((R_{(2)},R_{(6)}),k_{(2)}) := S^{-1}(R_{(2)} \oplus k_{(2)})$$

(alternate symmetry assumption $\mathcal{A}$)

Symmetries

If symmetry assumption $\mathcal{B}$ is true then $\beta^*_{j,t;k'} = \beta^*_{j,t;k''}$ for all k',k''.

Analogously, if symmetry assumption $\mathcal{A}$ is true then $\beta'^*_{j,t;k'} = \beta'^*_{j,t;k''}$ for all k',k''.

Note: In case of a (perfect) symmetry it suffices to estimate $h^*_{t,k}$ for any single key k.

Bundesamt
für Sicherheit in der
Informationstechnik

CASED

Symmetry metric

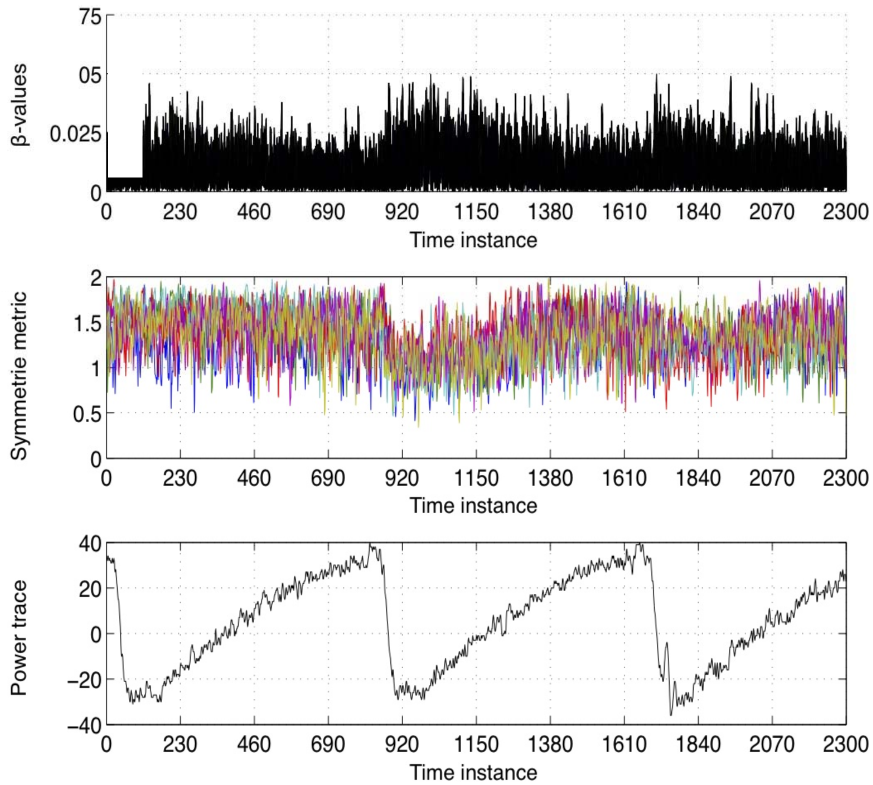
The ratio

$$\frac{2\sqrt{\sum_{j>0}\left(\tilde{\beta}_{j,t;k'}-\tilde{\beta}_{j,t;k''}\right)^2}}{\sqrt{\sum_{j>0}\tilde{\beta}_{j,t;k'}^{\,2}}+\sqrt{\sum_{j>0}\tilde{\beta}_{j,t;k''}^{\,2}}}$$

provides a symmetry metric.

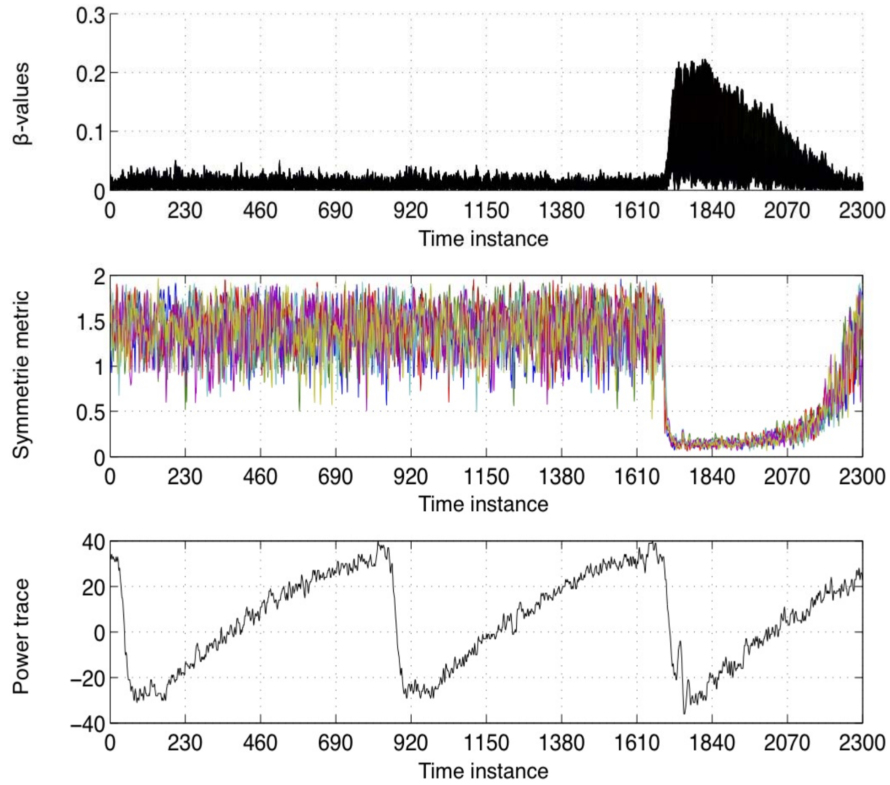
This metric has several interesting properties.
In particular, it is invariant under all orthonormal bases of $\mathcal{F}_{u,t;k}$ with first vector $g_{0,t;k}=1$.
The data-independent coefficient $\tilde{\beta}_{0,t;k}$ is neglected.

Empirical Results (DSD 2011, to appear)



Round 10

leakage model $\mathcal{A}$



Round 10

leakage model $\mathcal{B}$

Conclusion

- The stochastic approach is useful for many applications.
- It is a powerful attack method but it can also serve as a tool to support the design of secure implementations (constructive side-channel analysis).
- In particular, the β-characteristic allows to draw conclusions on specific features of an implementation, and leakage models can be verified or falsified.