Template Attacks in Principal Subspaces

Cédric Archambeau, **Eric Peeters**, François-Xavier Standaert and Jean-Jacques Quisquater

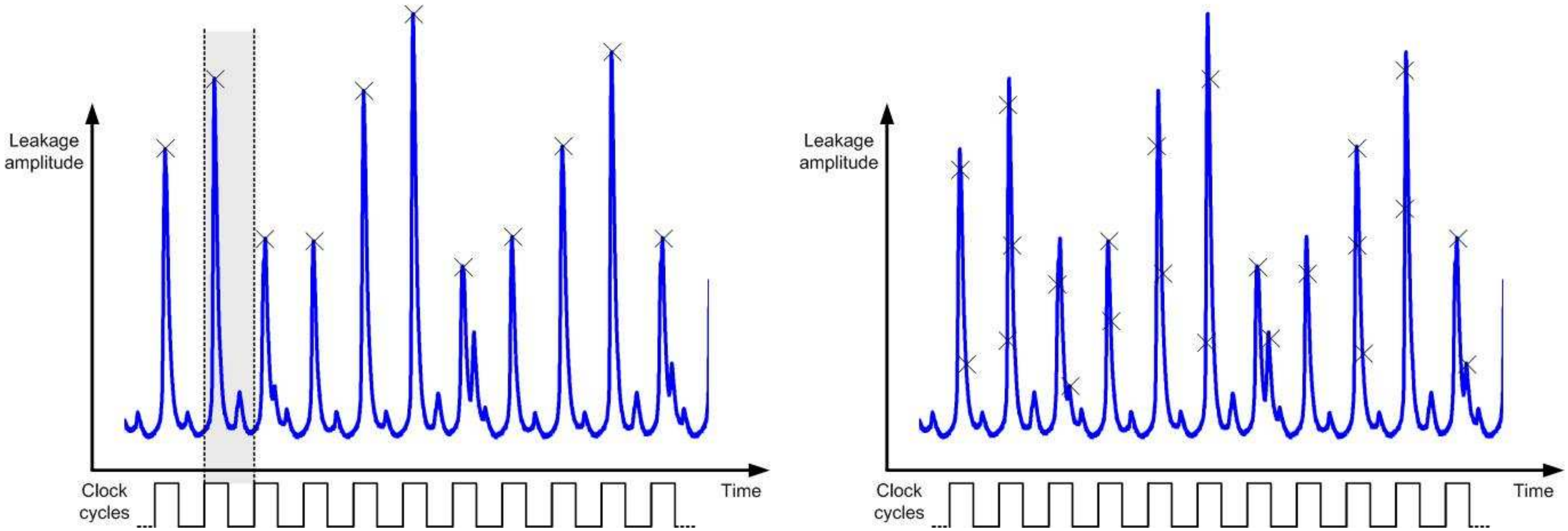
UCL Crypto Group, Université catholique de Louvain

- CHES 2006, Yokohama, Japan -

Outline

- Principle of template attacks
- Open issues in template attacks
- PCA in high dimensional space
- PSTA
- Results on RC4
- Results on AES
- Conclusions

Representation of side-channel information: univariate vs. multivariate approach



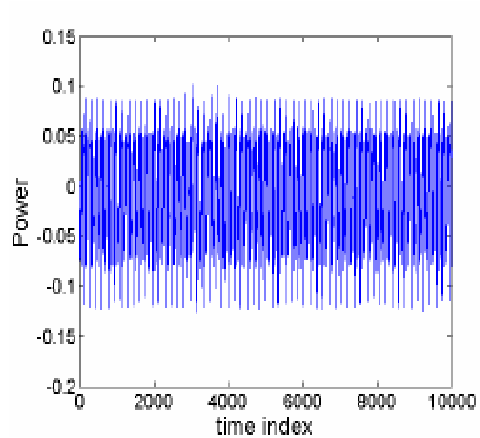
⇒ Most powerful adversary:

1. Take all **relevant** samples
2. and build a **multivariate** statistical model

Example: template attacks

[Chari *et al.*, 2002]

- *K* power traces



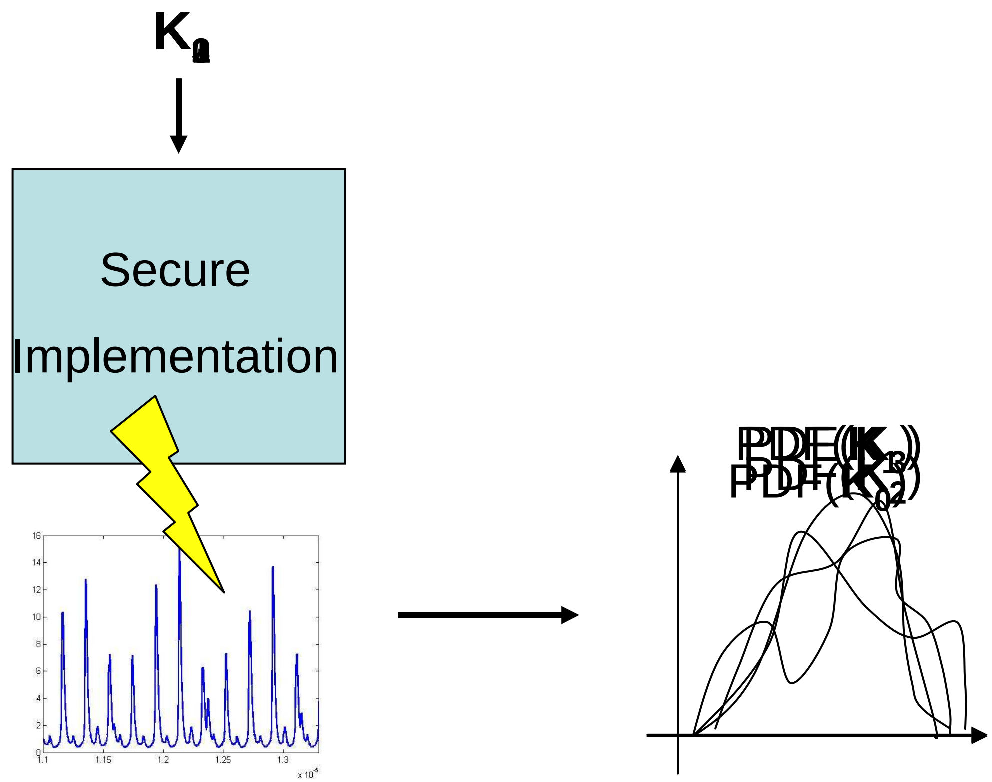
- Multivariate Gaussian noise model

$$\mathbf{P}(\boldsymbol{t}|s_k) = \mathcal{N}(\mathrm{t}|\boldsymbol{m}, \boldsymbol{S}) = \frac{1}{(2\pi)^{N/2}|\boldsymbol{S}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathrm{t}-\boldsymbol{m})^{\top}\boldsymbol{S}^{-1}(\mathrm{t}-\boldsymbol{m})\right\}$$

- Attack on new device

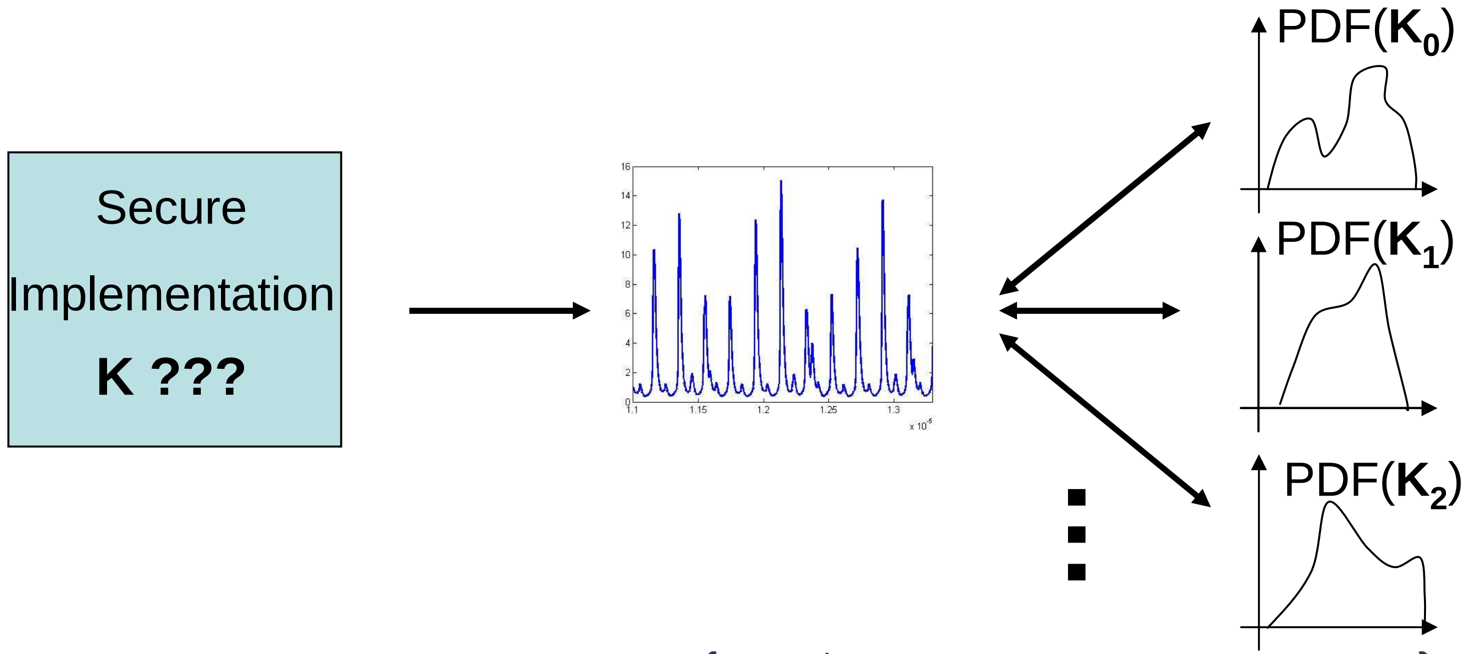
$$\widehat{s}_k = \underset{s_k}{argmax}\, \mathbf{P}(\boldsymbol{t}_{new}|s_k)\mathbf{P}(s_k)$$

Profiling phase



Any leakage:
power,
electromagnetic…

Attack phase



$$\widehat{s}_k = \underset{s_k}{argmax}\,\mathbf{P}(s_k|Observations)$$

Challenge: break a cipher using 1 single trace!

Open issues

1. **How to select the relevant samples?**
   - Look for the largest differences between the mean traces. [Chari et al. 2002, Rechberger et al. 2004]
   - Look for the largest cumulative differences.
   - Look for the samples with maximal variance.
2. **How to select window size?**
   - Clock cycle?
3. **How many samples are needed to attack?**

A trace ~ 10^5 samples

- Prohibitive memory usage!
- Automated way to reduce trace's size?

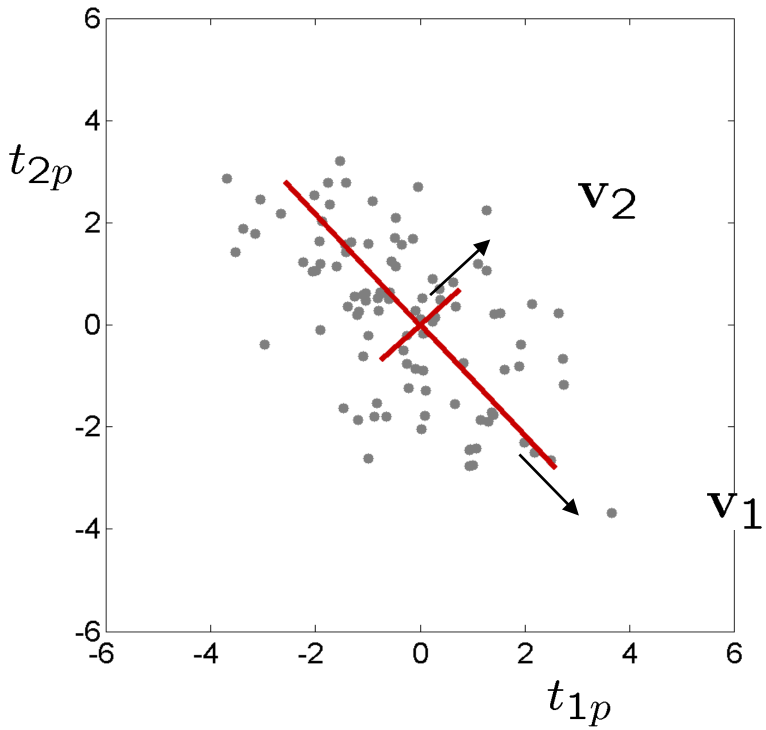
Hypothesis:

Information relies on amplitude of leakage signal (e.g. HW, HD models)

➔ Focus on instants where signal variability is maximal!

➔ 1 candidate: Principal Component Analysis

Principal Component Analysis



1. Rotate axes.
2. Discard irrelevant dimensions.

Find subspace that preserves maximal data variance!

Ordinary PCA

- Rotation matrix
  - Compute sample mean and covariance matrix

$$m = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{t}_k$$

$$S = \frac{1}{K}\sum_{k=1}^{K} (\boldsymbol{t}_k - m)(\boldsymbol{t}_k - m)^\top$$

  - Diagonalize sample covariance matrix (by eigendecomposition)

$$S\mathbf{V} = \mathbf{V}\boldsymbol{\Lambda}$$

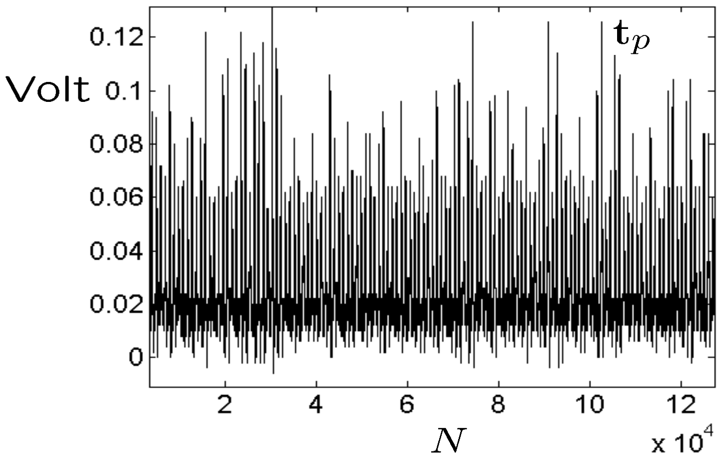
where $\mathbf{V}^\top\mathbf{V} = \mathbf{I}_K$

- Keep M eigenvectors corresponding to M largest eigenvalues

Principal directions

Variance in each direction

PCA in high dimensional data



- Practical limitations of PCA:
  - The complexity of an eigendecomposition is $O(N^3)$
  - $K << N$
- How to find the *K-1* first principal directions?
  - Eigendecomposition $\left(\frac{1}{K}\boldsymbol{T}_c^\top \boldsymbol{T}_c\right)\boldsymbol{U} = \boldsymbol{U}\boldsymbol{\Lambda} \quad \boldsymbol{T}_c \in \mathcal{R}^{K\times N}$
  - Covariance matrix $\left(\frac{1}{K}\boldsymbol{T}_c\boldsymbol{T}_c^\top\right)$
  - Left-multiplying by $\boldsymbol{T}_c$ gives $\boldsymbol{S}\left(\boldsymbol{T}_c\boldsymbol{U}\right) = \left(\boldsymbol{T}_c\boldsymbol{U}\right)\boldsymbol{\Lambda}$
  - Eigenvectors normalized $\boldsymbol{V} = \frac{1}{K}\left(\boldsymbol{T}_c\boldsymbol{U}\right)\boldsymbol{\Lambda}^{1/2}$

Principal Subspace Template Attacks

- Keep principal directions: M eigenvectors

$$\boldsymbol{V}_{1:M} = \frac{1}{\sqrt{K}} \left(\boldsymbol{T}_c \boldsymbol{U}_{1:M} \right) \boldsymbol{\Lambda}_{1:M}^{1/2}$$

- Parameters of multivariate Gaussian noise model:

$$\mathrm{P}\left(\boldsymbol{V}_{1:M}^{\top} \boldsymbol{t} | s_k \right) = \mathcal{N}\left(\boldsymbol{V}_{1:M}^{\top} \boldsymbol{t} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k \right)$$

where

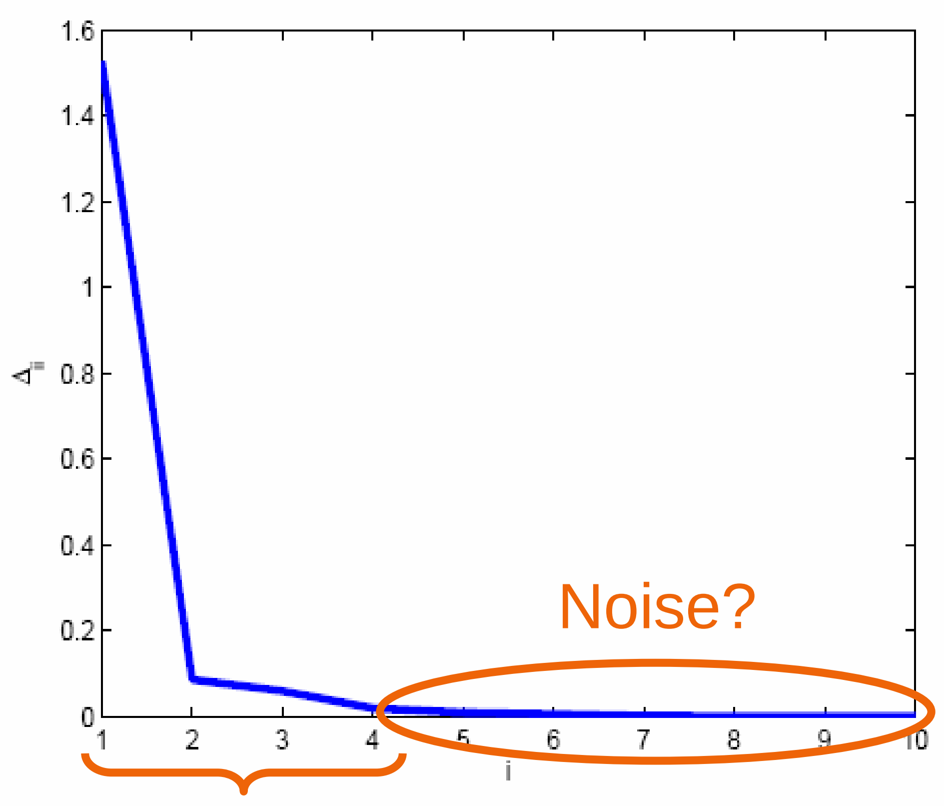
$$\begin{aligned} \boldsymbol{\mu}_k &= \boldsymbol{V}_{1:M}^{\top} \boldsymbol{m}_k \\ \boldsymbol{\Sigma}_k &= \boldsymbol{V}_{1:M}^{\top} \boldsymbol{S}_k \boldsymbol{V}_{1:M} \end{aligned}$$

- Attack in subspace:

$$\widehat{s}_k = \underset{s_k}{argmax}\, \mathbf{P}\left(\boldsymbol{V}_{1:M}^{\top} \boldsymbol{t}_{new} | s_k \right) \mathbf{P}(s_k)$$

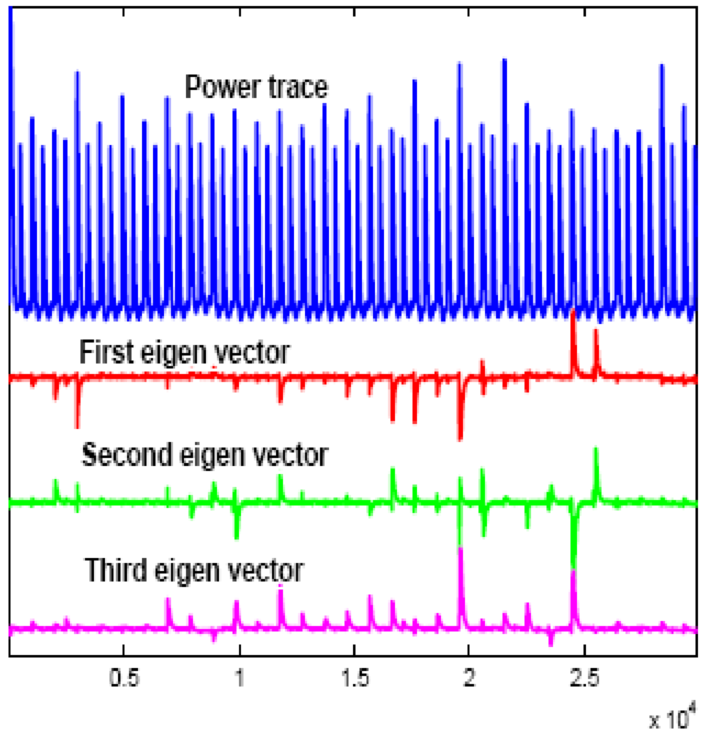
Results on RC4: ATmega88

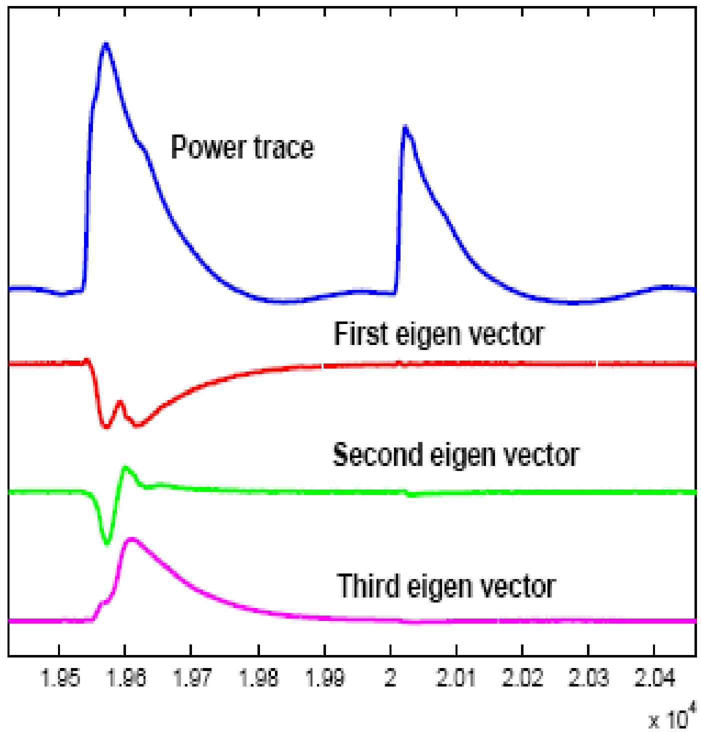
M eigenvalues?



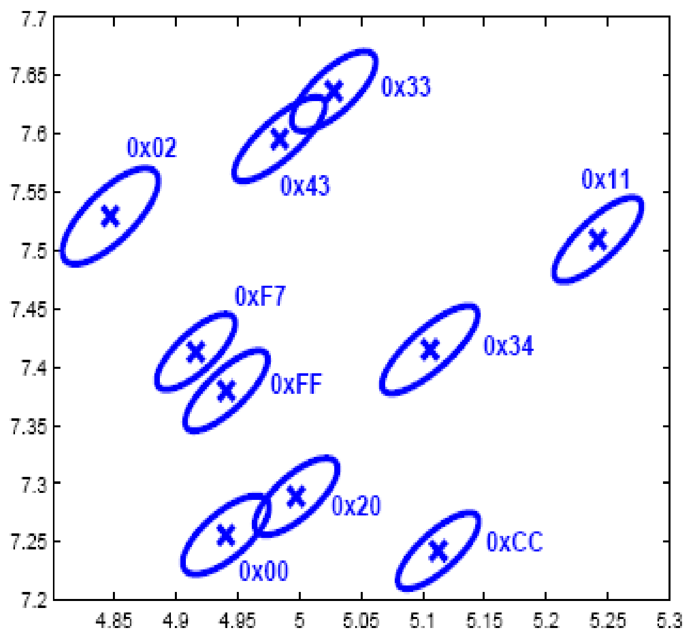
Information summurized in principal directions

Linear transformation in each direction = weighted sum

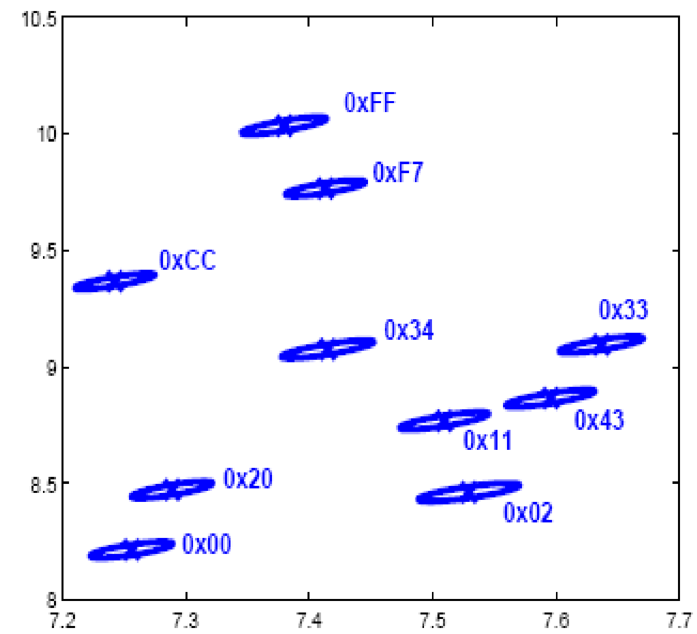




Classification of 10 keys



1st and 2nd directions

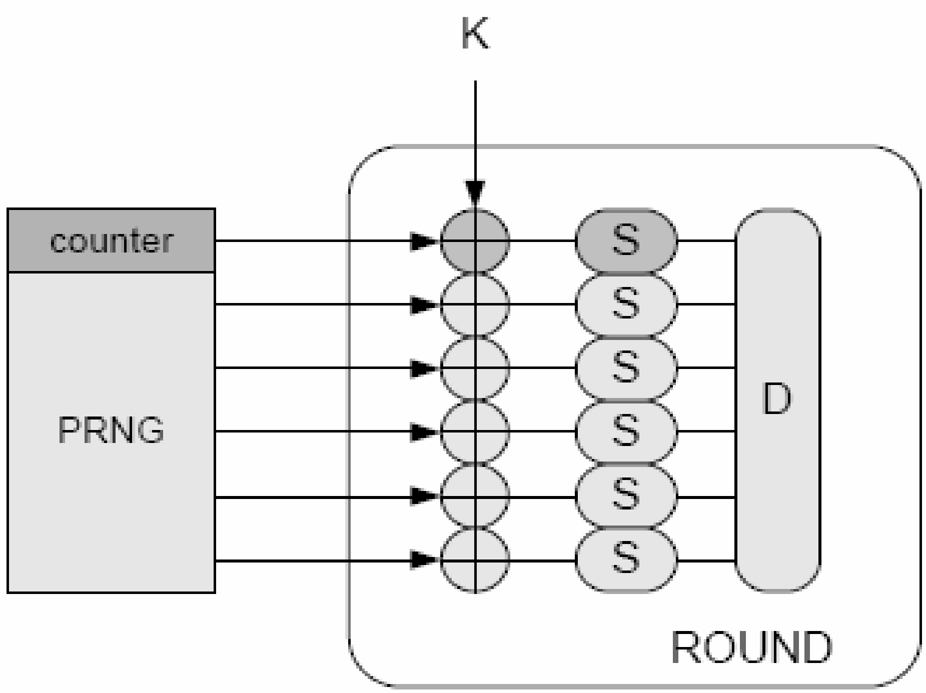


2nd and 3rd directions

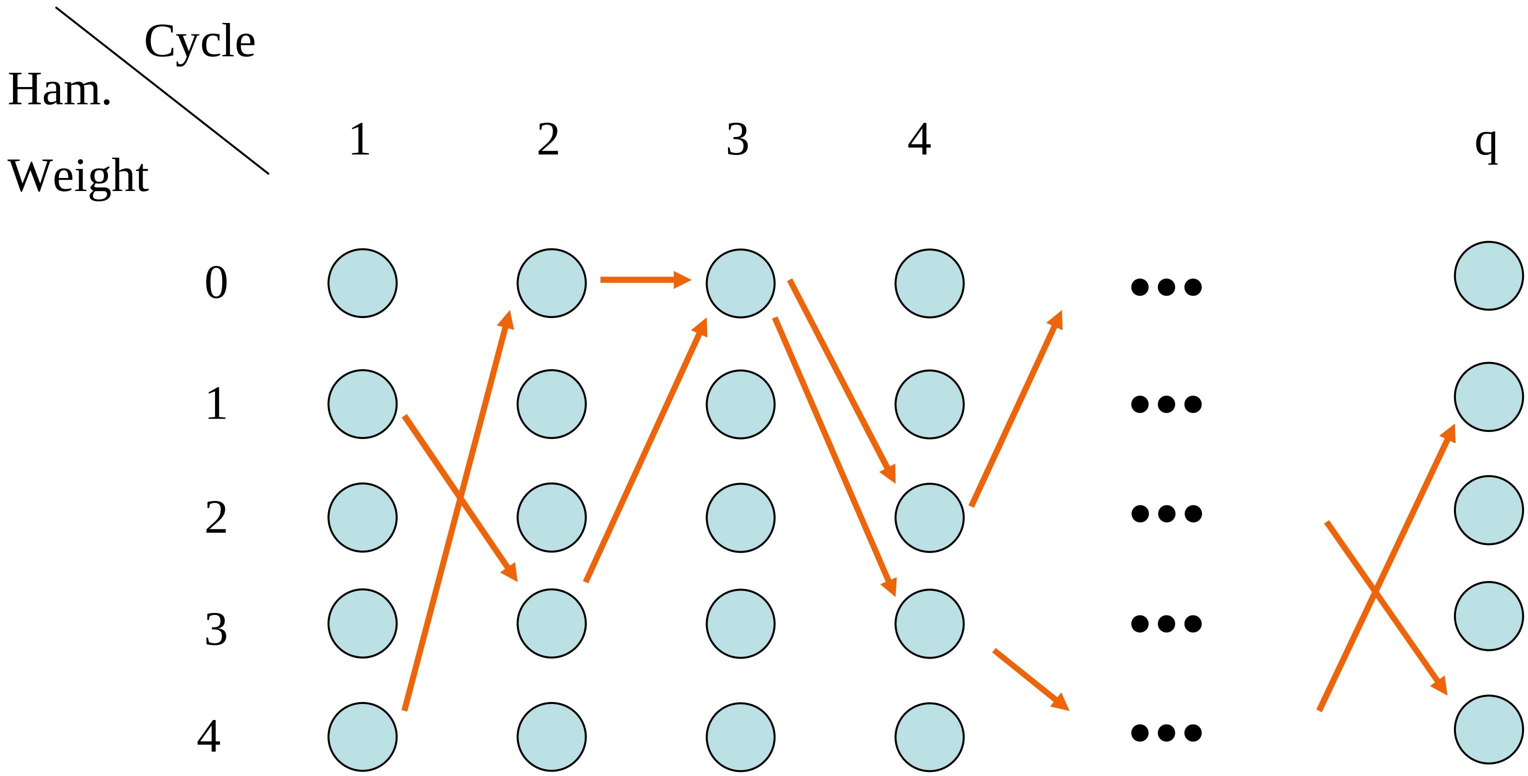
Classification rate: 99 % with 3 components

Results on AES Rijndael

- FPGA implementation on Spartan II



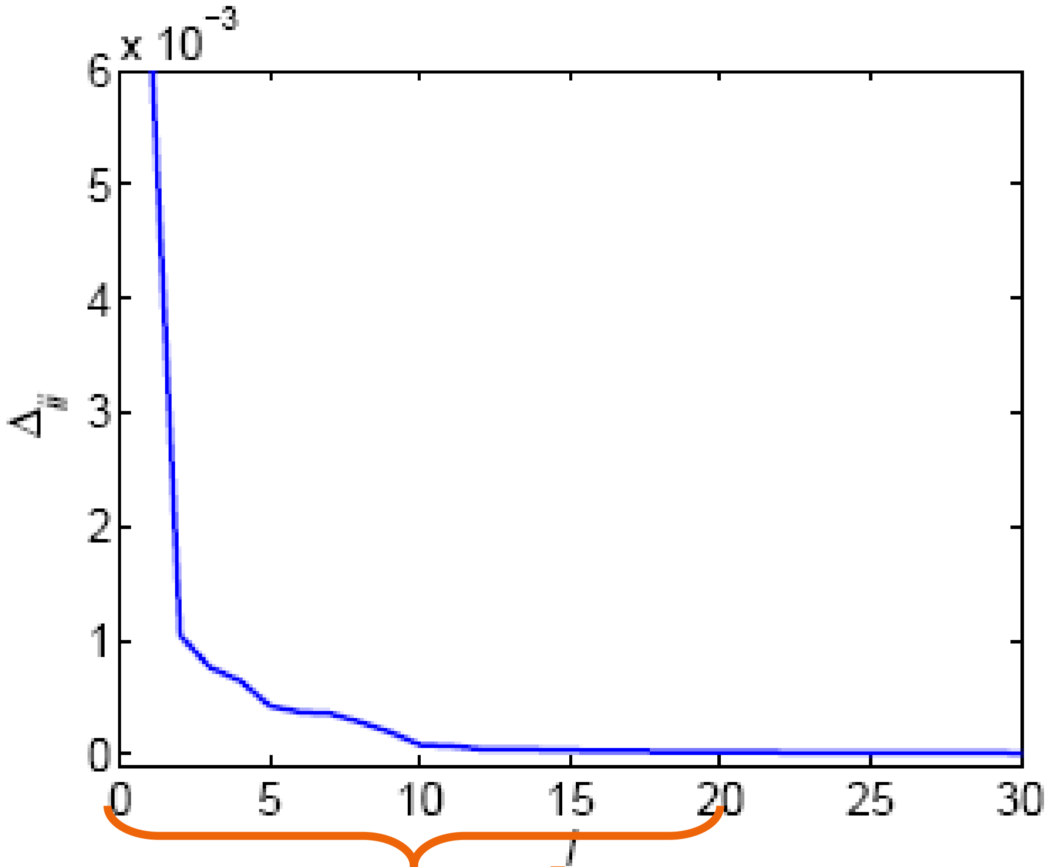
Each key candidate = ≠paths



s_k='000...0'

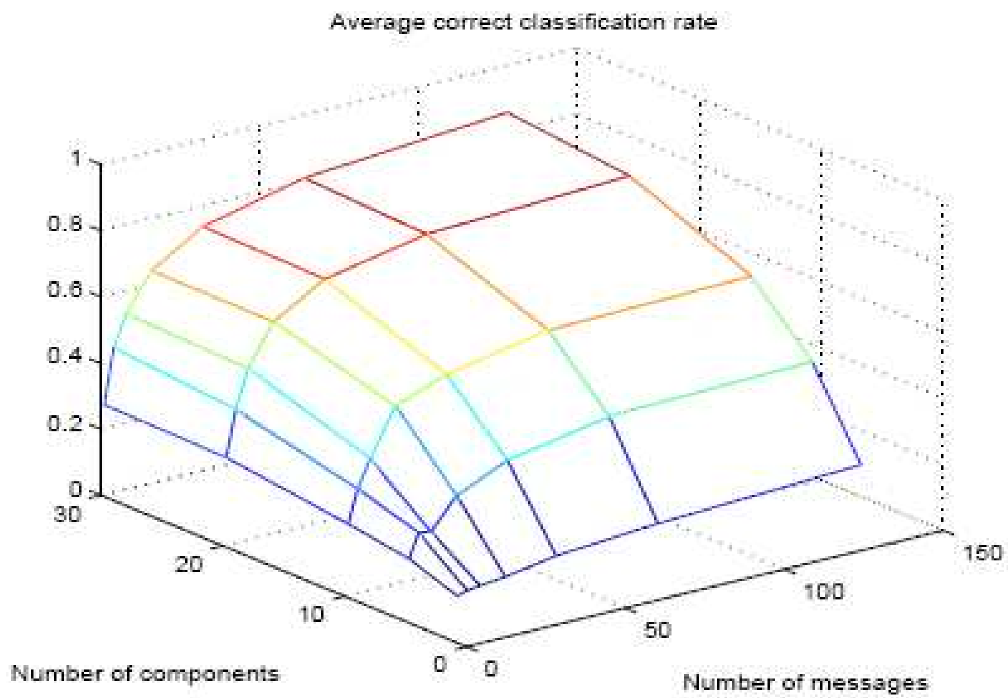
More directions needed

Eigenvalues?



Components kept

Classification rate



20 components and 128 encrypted messages:
86.7% on average
(vs. Previous attack's results: 500→ 2000 encr. mess.)

Conclusions

- PCA-based TA → principled approach for TA
- Relevant info → in a very few features (compression) automatically selected
- Maximal variance criterion → Starting hypothesis
- Succesfully applied to RC4 and AES
- Future work:
  - Optimal number of components and encrypted messages in case of the AES?
  - Behavior when noise process is important (or non-Gaussian)?

Questions?

How many principal directions?