Optimization of Power Analysis Using Neural Network

Zdenek Martinasek, Jan Hajny and Lukas Malina

Dpt. of Telecommunications, Brno University of Technology
Brno, Czech Republic
martinasek@feec.vutbr.cz
crypto.utko.feec.vutbr.cz

Outline

Crypto Research Group, Brno University of Technology, CZ

- Small group of cca 10 people,
- part of Department of Telecommunications, FEEC BUT in Brno, Czech Republic,
- equipped by SIX Research Centre,
- both basic and applied research,
- http://crypto.utko.feec.vutbr.cz/.

R&D in Cryptology and Computer Security

Basic research:

- provable cryptographic protocol design,
- light-weight cryptography,
- side channel cryptanalysis.

Implementation:

- smart-cards (Java, .NET, MultOS),
- mobile OS (iOS, Android),
- sensors, micro-controllers.

Main Characteristics of the Original Implementation

- PA based on two-layer perceptron network[1] (preparation of power patterns, training of the neural network, classification),
- the first experiment showed a success rate of **90%** for the first byte of AES secret key (`AddRoundKey` and `SubByte`),
- theoretical and empirical success rates were determined only to **80%** and **85%**, respectively,
- **these results were not sufficient enough**,
- other negative characteristics were revealed during the testing,
- optimization of the method above was realized **to increase the success rate of classification**.

[1] MARTINÁSEK, Z.; ZEMAN, V. Innovative Method of the Power Analysis. Radioengineering, 2013, vol. 22, no. 02, p. 586-594. ISSN: 1210- 2512.

Our Contribution

- Proposal of the optimization of the original power analysis method using the neural network,
- implementation of the proposed optimization,
- comparison the results of the optimized method with the original implementation,
- highlighting the positive and negative characteristic,
- verification of original method with standard 10-fold cross-validation,
- comparison of the results of both implementations using cross-validation..

Optimization Proposal - Preparation of Power Patterns

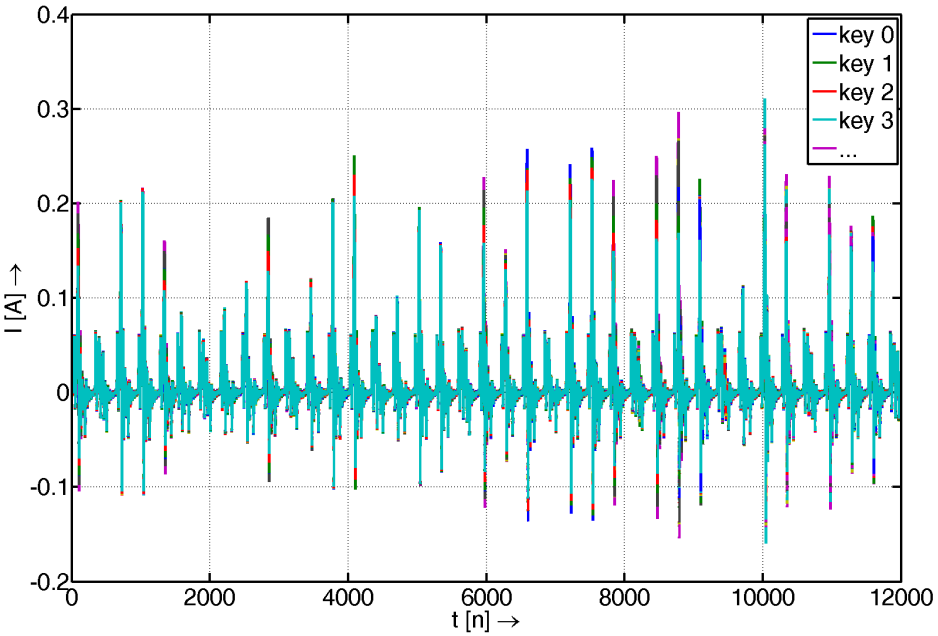
- The optimization using calculation of the average trace and the subsequent calculation of the difference power traces,
- denote $P[i, n]$ as power traces corresponding to every secret key value, where $n = \{0, \ldots, s\}$ is discrete time, and i represents all possible secret key byte values from 0 to 255,
- an average trace $\bar{A}$ can be calculate as:

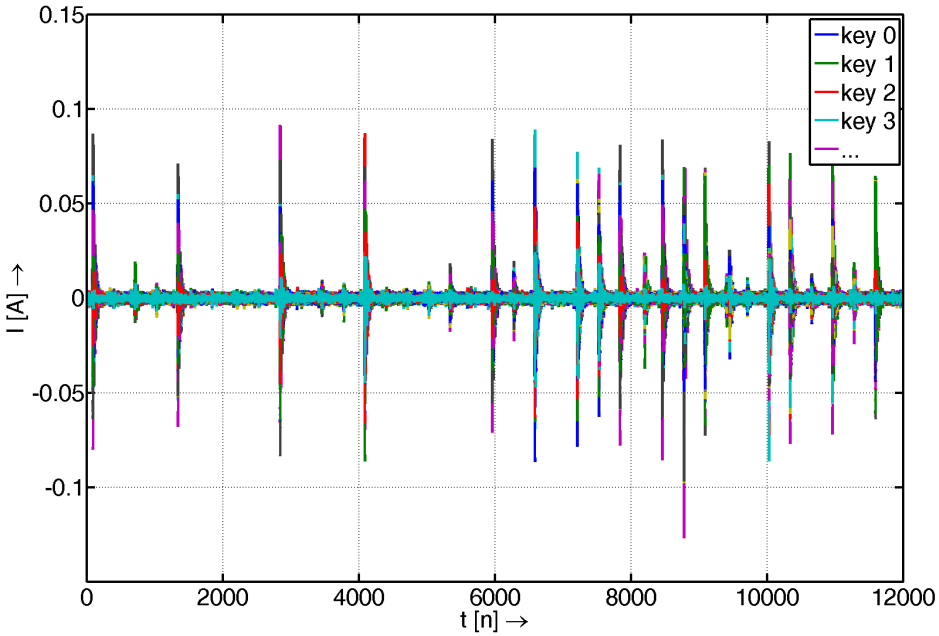
$$\bar{A}[n] = \frac{1}{256} \sum_{i=0}^{255} P[i, n]. \quad (1)$$

- training patterns for the optimized implementation are calculated as a subtraction:

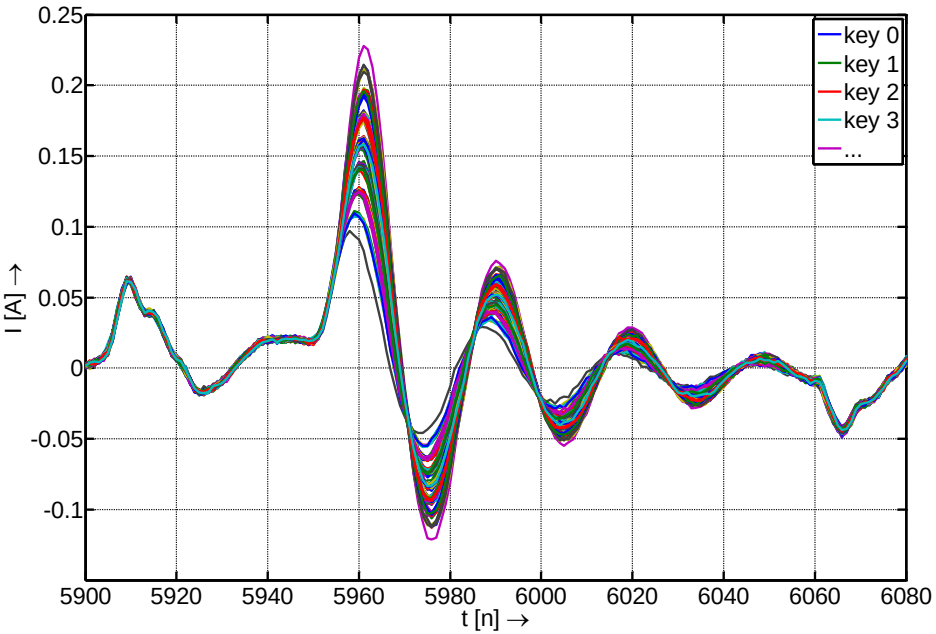
$$P_D[i, n] = \bar{A}[n] - P[i, n] = \frac{1}{256} \sum_{i=0}^{255} P[i, n] - P[i, n]. \quad (2)$$

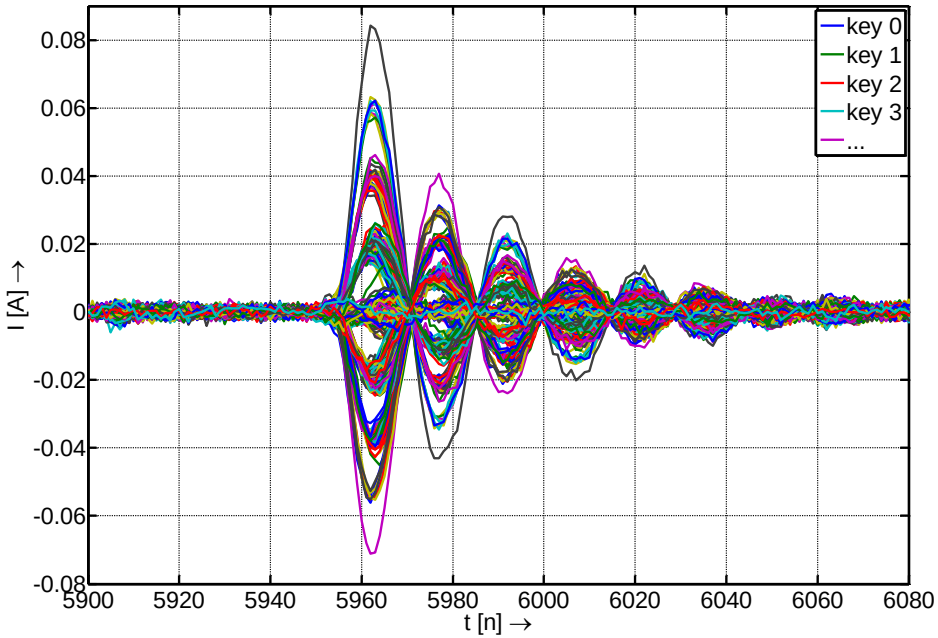
Comparison of Resulting Power Patterns





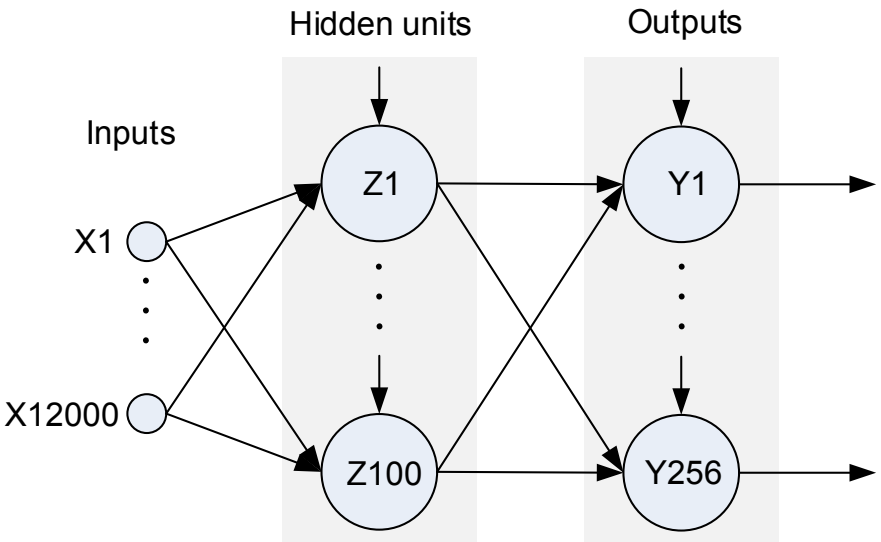
Detail of Power Patterns





Created Neural Network

- The neural network was created in MATLAB using the neural network toolbox,
- two-layer perceptron (MLP) was used,
- training set was realized by using 3×256 power traces, back propagation learning algorithm.



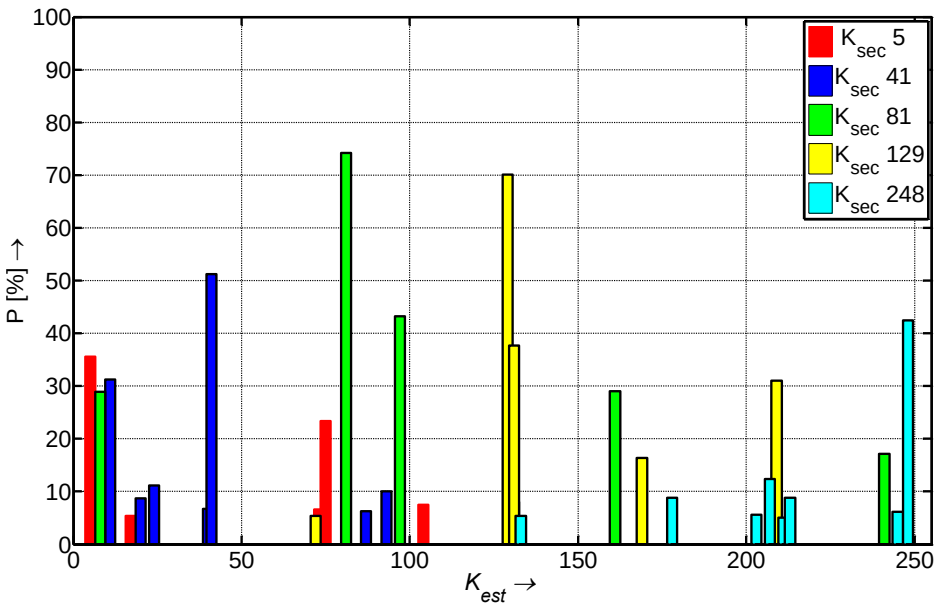
Comparison of Classification Results

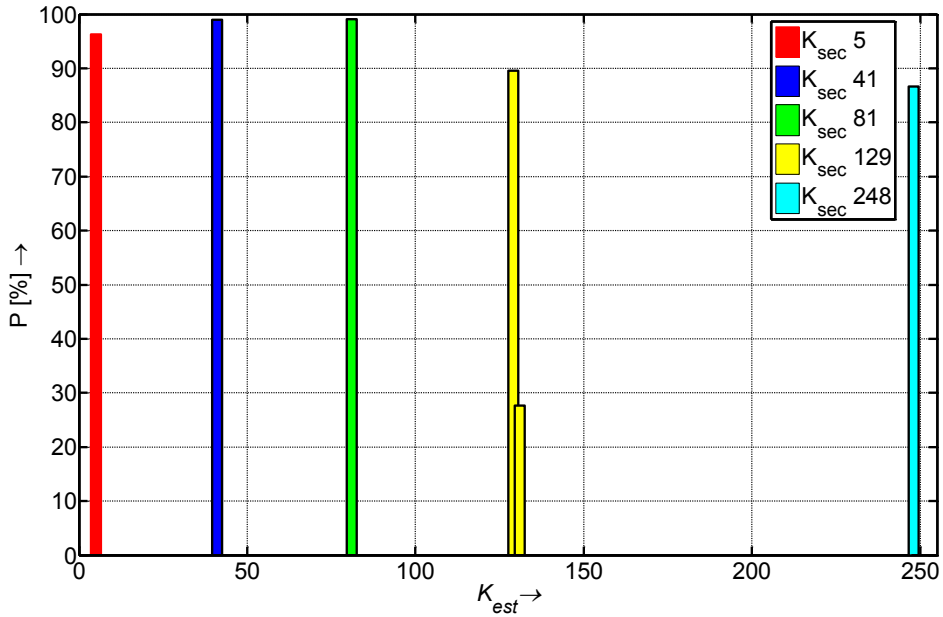
- A new set of 256 power traces corresponding to all secret key value was measured,
- whole set was subsequently classified.

$K_{sec} \downarrow$	Original implementation **R**				Optimized implementation $\mathbf{R_D}$				
⋮	...	...	...	...	...	...	...	...	...
2	0.00%	0.00%	6.46%	...	0.00%	0.00%	92.86%	0.00%	...
1	0.00%	66.42%	0.00%	...	0.00%	99.87%	0.00%	0.00%	...
0	36.77%	0.00%	0.00%	...	98.23%	0.00%	0.00%	0.00%	...
$K_{est} \rightarrow$	0	1	2	...	0	1	2	3	...

Probability Vector for Five Secret Keys

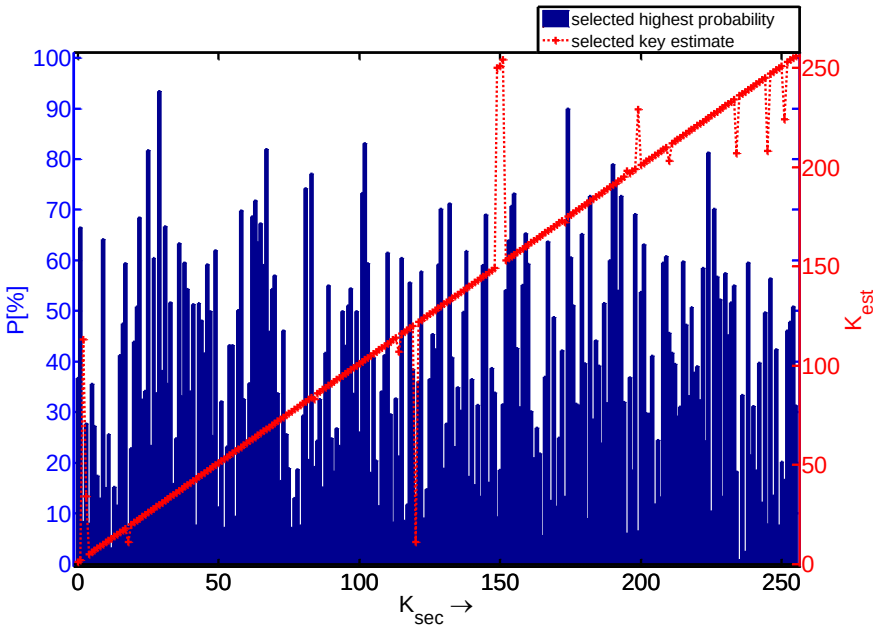
- Probability of correct key estimates is increased and the other possible key estimates are suppressed (negative?).

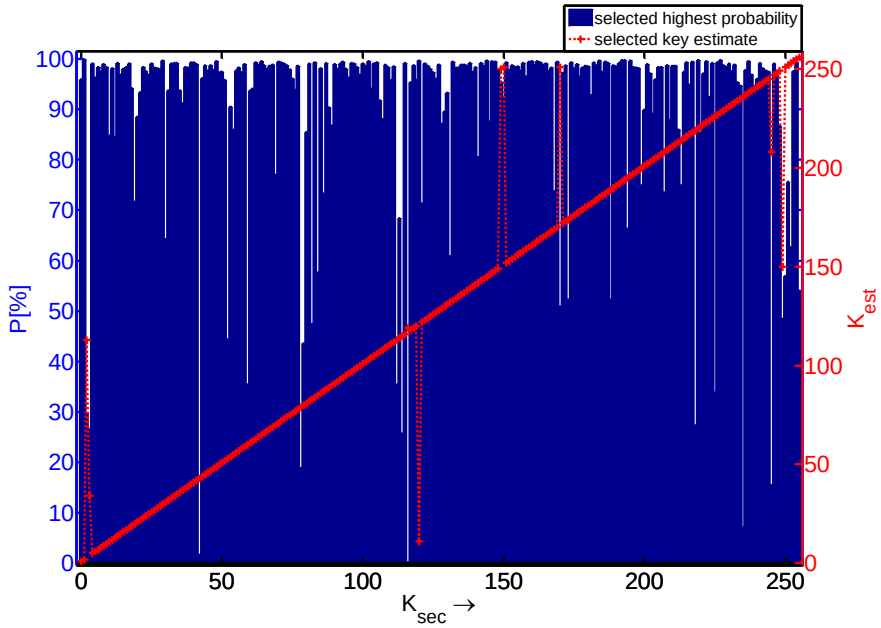




The Highest Selected Probabilities

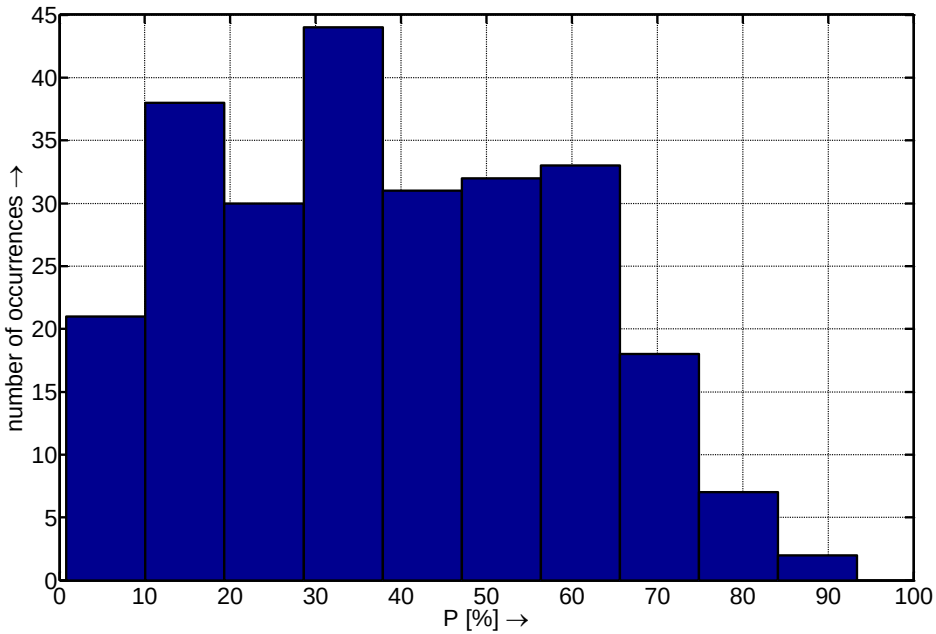
- Investigation of all selected key estimates,
- theoretical success rate 80% was calculated in the original implementation.

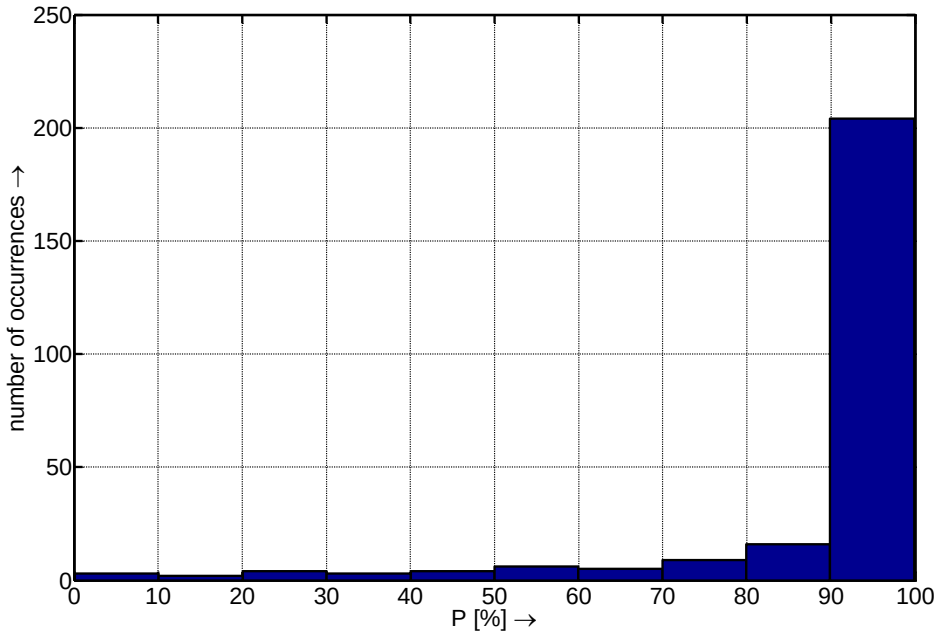




Histograms of Highest Probabilities

- The results confirm the increase of the maximum probabilities,
- number of keys potentially predisposed to wrong classification is reduced.





Cross-validation

- 2,560 power traces, 10 power traces for each key value,
- 10-fold cross-validation, 9 training traces and 1 testing in every step of validation,
- template attack: 256 templates, 9 interesting points.

Step of cross-validation	1	2	3	4	5	6	7	8	9	10	$\overline{err}$	Success rate [%]
Template $err[-]$	11	13	7	6	12	7	8	7	4	9	8.4	96.71
Original method $err[-]$	10	5	12	17	8	17	13	14	7	12	11.5	95.71
Optimized method $err[-]$	0	0	0	0	1	0	1	0	0	0	0.2	**99.92**

Conclusion

- Optimization of the power analysis based on multi-layer perceptron using preprocessing,
- the optimization allowed a significant improvement of the classification results,
- probability of correct key estimates was increased and the other possible key estimates were suppressed,
- total suppression of alternative probabilities might have negative effect,
- the original method and the optimized method were compared using the typical 10-fold cross-validation,
- the optimized method is able to reveal the secret key value with almost 100% success rate.

Thank you for attention!

martinasek@feec.vutbr.cz
crypto.utko.feec.vutbr.cz

This research work is funded by the Ministry of Industry and Trade of the Czech Republic, project FR-TI4/647. Measurements were run on computational facilities of the SIX Research Center, registration number CZ.1.05/2.1.00/03.0072.