

On optimality of MIA for unknown leakage models and related new practical results

François Bailly, Sylvain Guilley, Annelie Heuser, Olivier Rioul

▶ **To cite this version:**

François Bailly, Sylvain Guilley, Annelie Heuser, Olivier Rioul. On optimality of MIA for unknown leakage models and related new practical results. Lejla Batina; Matthew Robshaw. 16th Workshop on Cryptographic Hardware and Embedded Systems (CHES 2014), Sep 2014, Busan, South Korea. Springer, Lecture Notes in Computer Science, 8731, 2014, Cryptographic Hardware and Embedded Systems – CHES 2014 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings. hal-02300004

HAL Id: hal-02300004

https://telecom-paris.hal.science/hal-02300004

Submitted on 20 Aug 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the Optimality of Mutual Information Analysis

Institut Mines-Télécom
CNRS LTCI

Authors

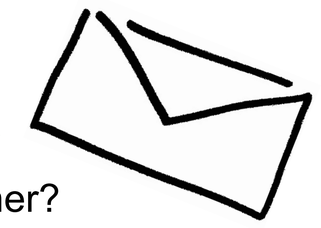
Francois Bailly
Sylvain Guilley
Annelie Heuser
Olivier Rioul

Annelie Heuser is a recipient of the Google Europe Fellowship in Privacy, and this research is supported in part by this Google Fellowship.

Motivation & State-of-the Art

- As the threat of side-channel attacks is well known, countermeasures are used for protection
- For example: Weakening the link between the measured leakage $\mathbf{x}$ and the sensitive variable $\mathbf{y}$
- Generic distinguisher cope with this scenario
  - Mutual Information Analysis (MIA) [1]
  - Kolmogorov-Smirnov distance (KSA) [2]
  - [3],…

Take home message!



- What is the *optimal* generic distinguisher?
- By applying the maximum likelihood principle, we derive the optimal generic distinguisher
- When the leakage has been quantified and probabilities are estimated from histograms, the optimal distinguisher's expression turns out to coincide with the mutual information analysis

[1] Benedikt Gierlichs, Lejla Batina, Pim Tuyls, and Bart Preneel. Mutual information analysis. In CHES, 10th International Workshop, volume 5154 of Lecture Notes in Computer Science, pages 426–442. Springer, August 10-13 2008. Washington, D.C., USA.
[2] Nicolas Veyrat-Charvillon and Fran ̧cois-Xavier Standaert. Mutual Information Analysis: How, When and Why? In CHES, volume 5747 of LNCS, pages 429–443. Springer, September 6-9 2009. Lausanne, Switzerland.
[3] N. Veyrat-Charvillon and F.-X. Standaert. Generic side-channel distinguishers: Improvements and limitations. In P. Rogaway, editor, CRYPTO, volume 6841 of Lecture Notes in Computer Science, pages 354–372. Springer, 2011.
[4] Julien Doget, Emmanuel Prouff, Matthieu Rivain, François-Xavier Standaert: Univariate side channel attacks and leakage modeling. J. Cryptographic Engineering 1(2): 123-144 (2011)

Universal Maximum Likelihood Equivalent to MIA

Notations & Assumptions

- Values are quantized (discrete leakage)

$$\hat{\mathbb{P}}(x|y) = \frac{\sum_{i=1}^{m} \mathbb{1}_{x_i=x,y_i=y}}{\sum_{i=1}^{m} \mathbb{1}_{y_i=y}} = \frac{\hat{\mathbb{P}}(x,y)}{\hat{\mathbb{P}}(y)}$$

$$\hat{\mathbb{P}}(x,y) = \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{x_i=x,y_i=y}$$

$$\hat{\mathbb{P}}(y) = \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{y_i=y}.$$

$$\hat{\mathbb{P}}(x) = \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{x_i=x}.$$

- Empirical Mutual Information

$$\hat{I}(\mathbf{x},\mathbf{y}) = \sum_{x\in\mathcal{X},y\in\mathcal{Y}} \hat{\mathbb{P}}(x,y)\log_2 \frac{\hat{\mathbb{P}}(x,y)}{\hat{\mathbb{P}}(x)\hat{\mathbb{P}}(y)}$$

- From the Maximum Likelihood it is known that maximizing the success rate amounts to select the key guess $\hat{k}$ that maximizes

$$\mathbb{P}(\mathbf{x}|\mathbf{y})$$

- In practice if no profiling is possible the conditional distribution is unknown
- Therefore, we need a *universal* version (computed from the available data without prior information)

$$\hat{k} = \arg\max_k \hat{\mathbb{P}}(\mathbf{x}|\mathbf{y}) = \arg\max_k \prod_{i=1}^{m} \hat{\mathbb{P}}(x_i|y_i)$$

- Universal Maximum Likelihood is equivalent to mutual information analysis

$$\hat{k} = \arg\max_k \hat{I}(\mathbf{x},\mathbf{y})$$

MIA is the optimal tool for key recovery when the model is unknown.

Proof sketch

Denoting $n_{x,y} = \sum_{i=1}^{m} \mathbb{1}_{x_i=x,y_i=y} = m\,\hat{\mathbb{P}}(x,y)$

$$\hat{\mathbb{P}}(\mathbf{x}|\mathbf{y}) = \prod_{i=1}^{m} \hat{\mathbb{P}}(x_i|y_i) = \prod_{x\in\mathcal{X},y\in\mathcal{Y}} \hat{\mathbb{P}}(x|y)^{n_{x,y}}$$

$$\hat{\mathbb{P}}(\mathbf{x}|\mathbf{y}) = \prod_{x\in\mathcal{X},y\in\mathcal{Y}} \hat{\mathbb{P}}(x|y)^{m\,\hat{\mathbb{P}}(x,y)} = 2^{-m\hat{H}(\mathbf{x}|\mathbf{y})}$$

where $\hat{H}(\mathbf{x}|\mathbf{y}) = \sum_{x\in\mathcal{X},y\in\mathcal{Y}} \hat{\mathbb{P}}(x,y)\log_2 \frac{1}{\hat{\mathbb{P}}(x|y)}.$

Empirical Future Work

- Experiments showing empirically the optimality of Mutual Information
- Especially, in comparison to Linear Regression Analysis [4]

Contact annelie.heuser@telecom-paristech.fr Web http://perso.telecom-paristech.fr/~heuser/