Institut
Mines-Télécom

Template Attacks, Optimal Distinguishers & Perceived Information Metric

Cryptarchi – June 29-30, 2015 – Leuven

Sylvain Guilley*, Annelie Heuser*,
Olivier Rioul* and François-Xavier Standaert**

*Telecom ParisTech, **UCL

Overview

Outlines

Motivation

- Consolidate state-of-the-art about optimal distinguishers with a deeper look on the probability estimation

Motivation

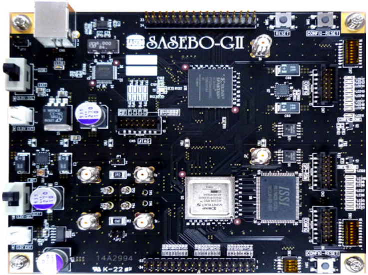
- Consolidate state-of-the-art about optimal distinguishers with a deeper look on the probability estimation
- Perceived Information (PI): information-theoretic metric quantifying the amount of leakage
- Show that PI is related to maximizing the success rate through the *Maximum a posteriori probability* (MAP)

Motivation

- Consolidate state-of-the-art about optimal distinguishers with a deeper look on the probability estimation
- Perceived Information (PI): information-theoretic metric quantifying the amount of leakage
- Show that PI is related to maximizing the success rate through the *Maximum a posteriori probability* (MAP)
- Use the *maximum likelihood* (ML) to derive MIA and the (experimental) template attack in case of profiling

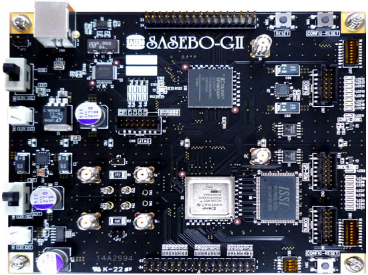
Motivation

Profiling device



$\hat{\mathbb{P}}$ for an estimation offline

Attacking device

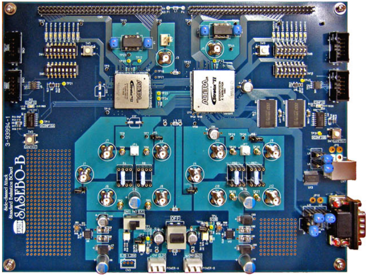


$\tilde{\mathbb{P}}$ estimated online on-the-fly

$\rightarrow$ $\mathbb{P}$ exact probability

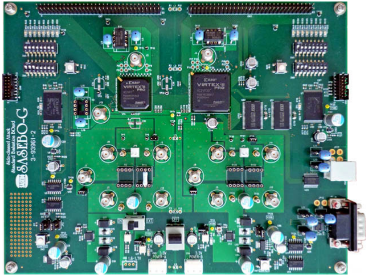
Motivation

Profiling device



$\hat{\mathbb{P}}$ for an estimation offline

Attacking device



$\tilde{\mathbb{P}}$ estimated online on-the-fly

$\rightarrow \mathbb{P}$ exact probability


l'union fait la force
Eendracht maakt macht
Einigkeit macht stark

Notations

- secret key k^* deterministic but unknown
- m independent measurements $\mathbf{x} = (x_1, ..., x_m)$ and independent and uniformly distributed inputs $\mathbf{t} = (t_1, ..., t_m)$
- leakage model $\mathbf{y}(k) = \varphi(f(k, \mathbf{t}))$, where φ is a device specific leakage function and f maps the inputs to an intermediate algorithmic state
- $\mathbf{x} = \mathbf{y}(k^*) + \mathbf{n}$ with independent noise $\mathbf{n}$

Perceived information

Idea [Renauld et al., 2011]

- Metric quantifying degraded leakage models
- Testing models against each other, e.g., from the true distribution against estimations
- Generalization of mutual information

Perceived information

Idea [Renauld et al., 2011]

- Metric quantifying degraded leakage models
- Testing models against each other, e.g., from the true distribution against estimations
- Generalization of mutual information

Ideal case

- the distribution $\mathbb{P}$ is known
- PI is MI

$$MI(K;X,T) = H(K) + \sum_k \mathbb{P}(k) \sum_t \mathbb{P}(t) \sum_x \mathbb{P}(x|t,k) \log_2 \mathbb{P}(k|t,x)$$

Perceived information

Profiled case

- the distribution $\mathbb{P}$ is known
- test a profiled model $\hat{\mathbb{P}}$ against $\mathbb{P}$

$$PI(K;X,T) = H(K) + \sum_k \mathbb{P}(k) \sum_t \mathbb{P}(t) \sum_x \mathbb{P}(x|t,k) \log_2 \hat{\mathbb{P}}(k|t,x)$$

Perceived information

Profiled case

- the distribution $\mathbb{P}$ is known
- test a profiled model $\hat{\mathbb{P}}$ against $\mathbb{P}$

$$PI(K;X,T) = H(K) + \sum_k \mathbb{P}(k) \sum_t \mathbb{P}(t) \sum_x \mathbb{P}(x|t,k) \log_2 \hat{\mathbb{P}}(k|t,x)$$

Real case

- the distribution $\mathbb{P}$ is unknown
- test a profiled model $\hat{\mathbb{P}}$ against an online estimated model $\tilde{\mathbb{P}}$

$$\hat{PI}(K;X,T) = H(K) + \sum_k \mathbb{P}(k) \sum_t \mathbb{P}(t) \sum_x \tilde{\mathbb{P}}(x|t,k) \log_2 \hat{\mathbb{P}}(k|t,x)$$

($\tilde{\mathbb{P}}$ estimated with non-parametric estimators, e.g. each x has $\tilde{\mathbb{P}} = \frac{1}{m}$)

Outlines

Maximum a posteriori probability

MAP

The optimal distinguishing rule is given by the *maximum a posteriori probability (MAP)* rule

$$\mathcal{D}(\mathbf{x}, \mathbf{t}) = \arg\max_{k} \ \mathbb{P}(k|\mathbf{x}, \mathbf{t})$$

Maximum a posteriori probability

MAP

The optimal distinguishing rule is given by the *maximum a posteriori probability (MAP)* rule

$$\mathcal{D}(\mathbf{x}, \mathbf{t}) = \arg\max_{k} \; \mathbb{P}(k|\mathbf{x}, \mathbf{t})$$

With the help of Bayes' rule...

$$\mathbb{P}(k|\mathbf{x}, \mathbf{t}) = \frac{\mathbb{P}(\mathbf{x}|k, \mathbf{t}) \cdot \mathbb{P}(k)}{\mathbb{P}(\mathbf{x}|\mathbf{t})} = \frac{\mathbb{P}(\mathbf{x}|k, \mathbf{t}) \cdot \mathbb{P}(k)}{\sum_k \mathbb{P}(k)\mathbb{P}(\mathbf{x}|\mathbf{t}, k)}$$

Relation between MAP and PI

- Profiling scenario
- Profiled model $\hat{\mathbb{P}}$, model $\tilde{\mathbb{P}}$ estimated online on-the-fly

Relation between MAP and PI

- Profiling scenario
- Profiled model $\hat{\mathbb{P}}$, model $\tilde{\mathbb{P}}$ estimated online on-the-fly
- $\hat{\mathbb{P}}(k|\mathbf{x}, \mathbf{t}) \propto \prod_{i=1}^{m} \hat{\mathbb{P}}(k|x_i, t_i)$

Relation between MAP and PI

- Profiling scenario
- Profiled model $\hat{\mathbb{P}}$, model $\tilde{\mathbb{P}}$ estimated online on-the-fly
- $\hat{\mathbb{P}}(k|\mathbf{x},\mathbf{t}) \propto \prod_{i=1}^{m} \hat{\mathbb{P}}(k|x_i,t_i)$

We start by maximizing MAP:

$$\begin{aligned}\arg\max_k \; \hat{\mathbb{P}}(k|\mathbf{x},\mathbf{t}) &= \arg\max_k \prod_{i=1}^{m} \hat{\mathbb{P}}(k|x_i,t_i) \\ &= \arg\max_k \prod_{x,t} \hat{\mathbb{P}}(k|x,t)^{m\tilde{\mathbb{P}}_k(x,t)}\end{aligned}$$

where $\tilde{\mathbb{P}}_k(x,t) = \tilde{\mathbb{P}}(x,t|k)$ is the "counting" estimation (online) of x and t that depends on k. Now taking the log_2 gives:

$$= \arg\max_k \sum_{x,t} \tilde{\mathbb{P}}_k(x,t)\log_2 \hat{\mathbb{P}}(k|x,t)$$

Relation between MAP and PI (cont'd)

$$= \arg\max_k \sum_{x,t} \tilde{\mathbb{P}}_k(x,t) \log_2 \hat{\mathbb{P}}(k|x,t)$$

$$= \arg\max_k \sum_{x,t} \tilde{\mathbb{P}}(x,t|k) \log_2 \hat{\mathbb{P}}(k|x,t)$$

$$= \arg\max_k \sum_{t} \tilde{\mathbb{P}}(t) \sum_{x} \tilde{\mathbb{P}}(x|t,k) \log_2 \hat{\mathbb{P}}(k|x,t)$$

Relation between MAP and PI (cont'd)

$$= \arg\max_k \sum_{x,t} \tilde{\mathbb{P}}_k(x,t) \log_2 \hat{\mathbb{P}}(k|x,t)$$

$$= \arg\max_k \sum_{x,t} \tilde{\mathbb{P}}(x,t|k) \log_2 \hat{\mathbb{P}}(k|x,t)$$

$$= \arg\max_k \sum_{t} \tilde{\mathbb{P}}(t) \sum_{x} \tilde{\mathbb{P}}(x|t,k) \log_2 \hat{\mathbb{P}}(k|x,t)$$

Taking the average over k and adding $H(K)$ gives $\hat{PI}(K;X,T) =$

$$H(K) + \sum_k \mathbb{P}(k) \sum_t \mathbb{P}(t) \sum_x \tilde{\mathbb{P}}(x|t,k) \log_2 \hat{\mathbb{P}}(k|x,t)$$

Relation between MAP and PI (cont'd)

PI $\Leftrightarrow$ MAP

$\hat{P}I$ (real case) is the expectation of the MAP over the keys

Relation between MAP and PI (cont'd)

PI $\Leftrightarrow$ MAP

$\hat{P}I$ (real case) is the expectation of the MAP over the keys

Profiled case

If we have an infinite number of traces to estimate $\tilde{\mathbb{P}} \to \mathbb{P}$ then we recover PI(K;X,T)

Relation between MAP and PI (cont'd)

PI $\Leftrightarrow$ MAP

$\hat{P}I$ (real case) is the expectation of the MAP over the keys

Profiled case

If we have an infinite number of traces to estimate $\tilde{\mathbb{P}} \to \mathbb{P}$ then we recover PI(K;X,T)

Ideal case

If we have an infinite number of traces to estimate $\tilde{\mathbb{P}} \to \mathbb{P}$ and $\hat{\mathbb{P}} \to \mathbb{P}$ then we recover MI(K;X,T)

Assumptions for ML

The leakage model follows the…

Markov condition

The leakage $\mathbf{x}$ depends on the secret key k only through the computed model y(k). Thus, we have the Markov chain

$$(k,t) \to y = \varphi(f(t,k)) \to x$$

Related to the EIS [Schindler et al., 2005] assumption.

- Markov condition: invariance of conditional probabilities
- EIS assumption: invariance of images under different subkeys

Maximum Likelihood Attack

Maximum Likelihood Attack

Assuming we have $y(k) = \varphi(f(t,k))$ that follows the Markov condition, then the optimal distinguishing rule is given by the maximum likelihood (ML) rule

$$\mathcal{D}(\mathbf{x},\mathbf{t}) = \arg\max_{k} \ \mathbb{P}(\mathbf{x}|\mathbf{y})$$

Proven and investigated in [Heuser et al., 2014]

Maximum Likelihood Attack

Similarly, as in the previous derivation we have:

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

Taking the $\log_2$ gives us:

$$\arg\max_k \sum_{x,y} \tilde{\mathbb{P}}(x,y) \log_2 \mathbb{P}(x|y)$$

Now we add the cross entropy term that does not depend on a key guess k:

$$-\sum_{x,y} \tilde{\mathbb{P}}(x,y) \log_2 \mathbb{P}(x).$$

Maximum Likelihood Attack

This results into:

$$\arg\max_k \sum_{x,y} \tilde{\mathbb{P}}(x,y) \log_2 \frac{\mathbb{P}(y|x)}{\mathbb{P}(y)}$$

Maximum Likelihood Attack

This results into:

$$\arg\max_{k} \sum_{x,y} \tilde{\mathbb{P}}(x,y) \log_2 \frac{\mathbb{P}(y|x)}{\mathbb{P}(y)}$$

In practice...

- $\mathbb{P}$ is most likely not known perfectly by the attacker
- So it is either estimated offline (leading to $\hat{\mathbb{P}}$)
- Or it is estimated online “on-the-fly" (leading to $\tilde{\mathbb{P}}$)

TELECOM ParisTech

Maximum Likelihood Attack

Profiled

If $\hat{\mathbb{P}}$ is estimated offline on a training device, we get

$$\arg\max_k \sum_{x,y} \tilde{\mathbb{P}}(x,y) \log_2 \frac{\hat{\mathbb{P}}(y|x)}{\hat{\mathbb{P}}(y)}$$

Which is the *template attack* [Chari et al., 2002]

i.e. a distinguisher resulting from the MAP with

- A priori knowledge on the key distribution
- & assumting the Markov condition

Maximum Likelihood Attack

Profiled

If $\hat{\mathbb{P}}$ is estimated offline on a training device, we get

$$\arg\max_k \sum_{x,y} \tilde{\mathbb{P}}(x,y) \log_2 \frac{\hat{\mathbb{P}}(y|x)}{\hat{\mathbb{P}}(y)}$$

Which is the *template attack* [Chari et al., 2002]

Non-Profiled

If $\tilde{\mathbb{P}}$ is estimated online on a the device under attack, we get

$$\arg\max_k \sum_{x,y} \tilde{\mathbb{P}}(x,y) \log_2 \frac{\tilde{\mathbb{P}}(y|x)}{\tilde{\mathbb{P}}(y)}$$

Which is the *Mutual Information Analysis* [Gierlichs et al., 2008]

Outlines

Conclusion

- Maximizing the PI = optimizing the MAP attacks (on average over the keys)
- ML is a alternative to MAP (no penalty if keys are uniform)
- Maximum likelihood attacks correspond to
  - template attacks when probabilities are estimated offline ($\hat{\mathbb{P}}$)
  - MIA when probabilities are estimated online “on-the-fly" ($\tilde{\mathbb{P}}$)
- All attacks work by "testing" a model (estimated offline or online "on-the-fly") against fresh samples
- For profiled attacks, a (well estimated) more accurate model always help / for non-profiled ones, simpler (easier to estimate online “on-the-fly") models can be better

Thank you!

Questions?

annelie.heuser@telecom.paristech.fr

References I

Chari, S., Rao, J. R., and Rohatgi, P. (2002).
Template Attacks.
In *CHES*, volume 2523 of *LNCS*, pages 13–28. Springer.
San Francisco Bay (Redwood City), USA.

Gierlichs, B., Batina, L., Tuyls, P., and Preneel, B. (2008).
Mutual information analysis.
In *CHES, 10th International Workshop*, volume 5154 of *Lecture Notes in Computer Science*, pages 426–442. Springer.
Washington, D.C., USA.

References II

Heuser, A., Rioul, O., and Guilley, S. (2014).
Good Is Not Good Enough - Deriving Optimal Distinguishers from Communication Theory.
In Batina, L. and Robshaw, M., editors, *Cryptographic Hardware and Embedded Systems - CHES 2014 - 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings*, volume 8731 of *Lecture Notes in Computer Science*, pages 55–74. Springer.

References III

Renauld, M., Standaert, F.-X., Veyrat-Charvillon, N., Kamel, D., and Flandre, D. (2011).
A Formal Study of Power Variability Issues and Side-Channel Attacks for Nanoscale Devices.
In *EUROCRYPT*, volume 6632 of *LNCS*, pages 109–128. Springer.
Tallinn, Estonia.

Schindler, W., Lemke, K., and Paar, C. (2005).
A Stochastic Model for Differential Side Channel Cryptanalysis.
In LNCS, editor, *CHES*, volume 3659 of *LNCS*, pages 30–46. Springer.
Edinburgh, Scotland, UK.