TELECOM ParisTech

Institut Mines-Télécom

# On the Optimality of Mutual Information Analysis for Discrete Leakages

Cryptarchi – June 29-30, 2015 – Leuven

Éloi de Chérisey*, Annelie Heuser**, Sylvain Guilley** and Olivier Rioul**

* ENS Cachan, **Telecom ParisTech

# Outline

# Outline

| Éloi de CHÉRISEY *is with* | Sylvain GUILLEY *is also with* | Annelie HEUSER *is PhD fellow at* | Olivier RIOUL *is also Prof at* |
| --- | --- | --- | --- |
| ENS CACHAN | SECURE-IC THE TRUSTED COMPUTING COMPANY | Google | l'X ÉCOLE POLYTECHNIQUE UNIVERSITÉ PARIS-SACLAY |

## Research thread: “*Distinguishing distinguishers*”

- CHES 2014: known model + large Gaussian noise

  $\Longrightarrow$ **CPA is optimal**
- ASICRYPT 2014: idem with masking

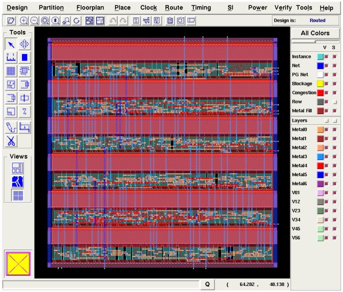
  $\Longrightarrow$ **HO-CPA is optimal**
- CRYPTARCHI 2015: unknown model

  $\Longrightarrow$ **MIA is optimal**

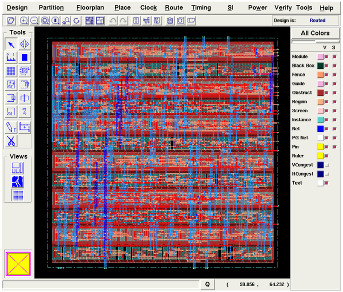
# Motivations

1. How to attack when
   - no leakage model is available?
   - no profiling is possible?

   (e.g., balanced dual-rail circuits)



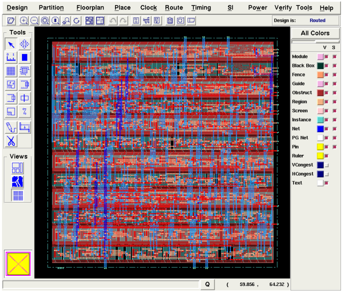
# Motivations




1. How to attack when
   - no leakage model is available?
   - no profiling is possible?

   (e.g., balanced dual-rail circuits)

# Motivations

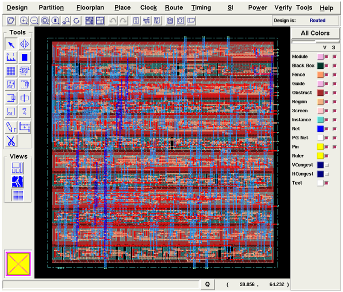



1. How to attack when
   - no leakage model is available?
   - no profiling is possible?

   (e.g., balanced dual-rail circuits)

   → Mutual Information Analysis (MIA) [Gierlichs et al., CHES 2008]

# Motivations




1. How to attack when
   - no leakage model is available?
   - no profiling is possible?

   (e.g., balanced dual-rail circuits)

   → Mutual Information Analysis (MIA) [Gierlichs et al., CHES 2008]
2. Is it possible for MIA to be:
   - Optimal?
   - Effective?
   - Efficient?

## Results

We show that:

- MIA is“optimal” in that it is equivalent to a “universal” maximum likelihood (U-ML)
- MIA can even outperform CPA
- the computational complexity of MIA can be significantly reduced

# Outline

## Leakage model

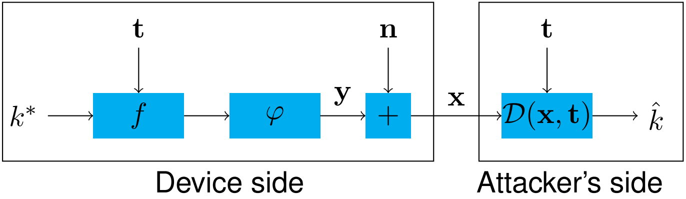
Assumptions:

- secret key $k^*$: deterministic but unknown
- uniformly distributed text bytes $\mathbf{t} = (t_1, \ldots, t_m)$: random but known
- $m$ i.i.d. noisy measurements $\mathbf{x} = (x_1, \ldots, x_m)$

$$\mathbf{x} = \mathbf{y}(k^*) + \mathbf{n}$$

- leakage model:

$$\mathbf{y}(k) = \varphi(f(k, \mathbf{t}))$$



# Markov Chain Property

## Markov Condition

The leakage $\mathbf{x}$ depends on the secret key $k^*$ only through the computed model $\mathbf{y}(k^*) = \varphi(f(k^*, \mathbf{t}))$

Thus the *conditional distribution* $\mathbb{P}(\mathbf{x}|\mathbf{y})$ *depends on the key only through the value of* $\mathbf{y}$.

## Example

- $x_i = w_H(k^* \oplus t_i) + n_i$, where $w_H$ is the Hamming weight
- $x_i = \varphi(S(k^* \oplus t_i)) + n_i$, where $S$ is a substitution box

# Empirical Distribution (Histogram)

## Assumptions

Leakage is discrete (or discretized *via* binning).

Counting occurrences of each value of $x$ and $y$ gives the estimate:

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

## Definition (Empirical Mutual Information)

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

# Best Distinguisher

The optimal distinguisher $\mathcal{D}(\mathbf{x}, \mathbf{y})$ and associated distinguishing rule $\hat{k} = \arg\max_{k \in \mathcal{K}} \mathcal{D}(\mathbf{x}, \mathbf{y})$ maximizes success:

**Definition (Average Success Rate)**

$$SR = \frac{1}{2^n} \sum_{k=0}^{2^n-1} \mathbb{P}_k(\hat{k}(\mathbf{X}, \mathbf{T}) = k)$$

This is a theoretical definition since $\mathbb{P}_k(\mathbf{x}, \mathbf{t})$ is not known perfectly.

# Maximum likelihood

**Theorem (Optimal Distinguisher = Maximum Likelihood)**

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

*Recall* $\mathbf{y} = \varphi(f(k, \mathbf{t}))$ *depends on the key hypothesis* $k$*.*

Proved by [Heuser et al., in CHES 2014]

Optimality holds for perfectly known leakage model $\varphi$

# Universal Maximum Likelihood

Universal = from the data without prior information.

**Definition (Universal Maximum Likelihood)**

$$\hat{k} = \arg\max_{k \in \mathcal{K}} \tilde{\mathbb{P}}(\mathbf{x}|\mathbf{y})$$

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

Loss in performance due to empirical probability estimation.

# Universal Maximum Likelihood

**Theorem (MIA = U-ML)**

*The universal ML key estimate is equivalent to MIA! [Gierlichs et al. CHES 2008]*

$$\hat{k} = \arg\max_{k \in \mathcal{K}} \tilde{\mathbb{P}}(\mathbf{x}|\mathbf{y}) = \arg\max_{k \in \mathcal{K}} \tilde{I}(X;Y)$$

This surprising theoretical result shows that MIA is “universally” optimal.

## Proof of the Theorem

Rearrange empirical probability:

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

where

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

Now take the logarithm:

$$\log \tilde{\mathbb{P}}(\mathbf{x}|\mathbf{y}) = \sum_{x\in\mathcal{X},y\in\mathcal{Y}} m\tilde{\mathbb{P}}(x,y)\log(\tilde{\mathbb{P}}(x|y))$$

We recognize entropy!

$$\tilde{H}(X|Y) = -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} \tilde{\mathbb{P}}(x,y)\log(\tilde{\mathbb{P}}(x|y)) \quad /\ldots$$

## Proof of the Theorem (cont'd)

In other words:

$$\tilde{\mathbb{P}}(\mathbf{x}|\mathbf{y}) = 2^{-m\tilde{H}(X|Y)}$$

Thus maximizing $\tilde{\mathbb{P}}(\mathbf{x}|\mathbf{y})$ amounts to minimizing $\tilde{H}(X|Y)$, i.e., maximizing

$$\tilde{I}(X,Y) = \tilde{H}(X) - \tilde{H}(X|Y)$$

since $\tilde{H}(X)$ does not depend on $k$.

Conclusion:

$$\boxed{\arg\max_{k\in\mathcal{K}} \tilde{\mathbb{P}}(\mathbf{x}|\mathbf{y}) = \arg\max_{k\in\mathcal{K}} \tilde{I}(X;Y)}$$

$\square$

# Outline

# Case-Study

Consider the following leakage : $X = Y + N$ with $Y = \varphi(f(T \oplus K^*))$ where $Y = \pm 1$ and $N \sim \mathcal{U}(\pm\sigma)$ (uniformly distributed with values $\pm\sigma$).

## Assumptions

- $T, K \in \mathbb{F}_2^n$
- $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is an Sbox function applied to a xor and $\varphi$ is a one-bit selection function.
- $\sigma \in \mathbb{N}$ is an integer

## Example

Dynamic Voltage Scaling (DVS) used as a countermeasure.

# Case-Study: Distinguisher

## Theorem (Optimal Distinguisher)

*One easily finds*

$$\mathcal{D}_{opt}(\mathbf{x}, \mathbf{t}) = \arg\max_{k \in \mathcal{K}} \mathbb{P}(\mathbf{x}|\mathbf{y}) = \arg\max_{k \in \mathcal{K}} \frac{1}{2^m} \prod_{i=1}^{m} \begin{cases} 1 & \text{if } x - \varphi(f(t,k)) = 0 \\ 1 & \text{if } x - \varphi(f(t,k)) = 2\sigma \\ 0 & \text{otherwise} \end{cases}$$

To be compared to

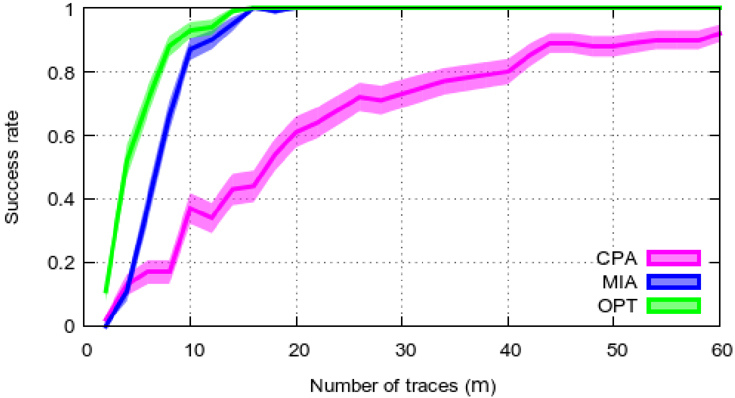
i CPA:

$$\mathcal{D}_{CPA}(\mathbf{x}, \mathbf{t}) = \arg\max_{k \in \mathcal{K}} \frac{\widetilde{\text{cov}}(X, Y)}{\tilde{\sigma}_X \tilde{\sigma}_Y}$$

ii MIA :

$$\mathcal{D}_{MIA}(\mathbf{x}, \mathbf{t}) = \arg\max_{k \in \mathcal{K}} \tilde{I}(X, Y)$$

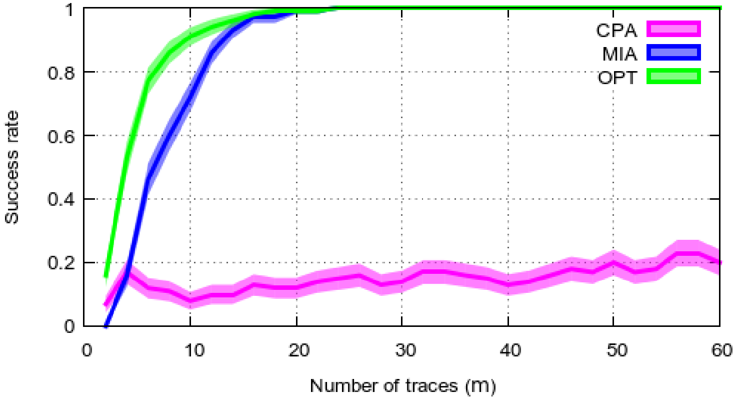
# Results for MIA and CPA ($\sigma = 1$)






TELECOM ParisTech

# Results for MIA and CPA ($\sigma = 4$)





TELECOM ParisTech

# Outline

# Size of the Bins

## Why Binning?

Due to noise, $\mathbf{x}$ assumes real values: $p(\mathbf{x}|\mathbf{y})$ as a pdf that must be estimated using bins as $\tilde{\mathbb{P}}(\mathbf{x}|\mathbf{y})$

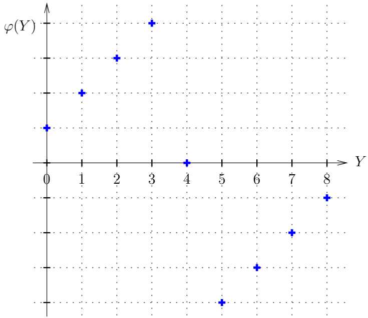
- what is the optimal size of bins?
- does it depend of the number of traces?

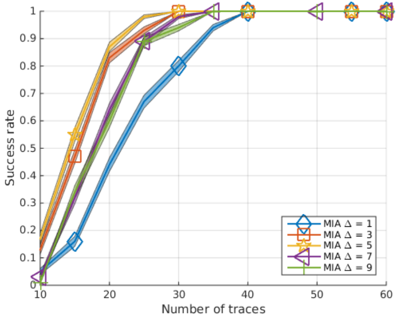
## Leakage Example

### Example

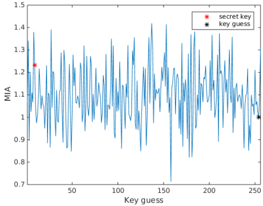
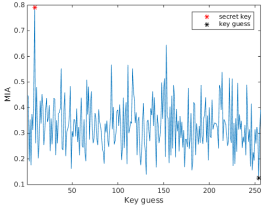
$\mathbf{y} = \varphi(f(\mathbf{y}(k^*, \mathbf{t})))$ with

- $f(k, \mathbf{t}) = \mathsf{Sbox}(\mathbf{t} \oplus k^*)$
- $\varphi = \psi(w_H(\bullet))$ where $\psi$ is a non-linear bijective function s.t. $\mathsf{Cov}(Y, \varphi(Y)) = 0$



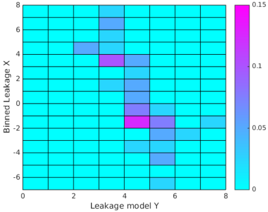
# Leakage Example (cont'd)

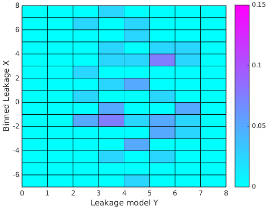
| | MIA using 30 traces |
|---|---|
| $\Delta = 1$ |  |
| $\Delta = 5$ |  |

# Leakage Example (cont'd)

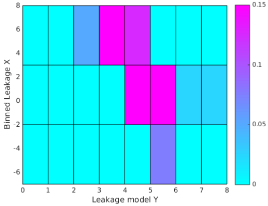
| | Correct key guess | False key guess |
|---|---|---|

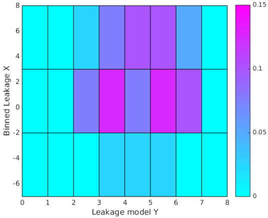
$\Delta = 1$




$\Delta = 5$



## Fast MIA Algorithm Principle

See [Victor Lomné et al. in ASIACRYPT 2013] for CPA

### Principle

The pmf is given by

$$\tilde{\mathbb{P}}(y, x) = \sum_{t \mid \varphi(f(t,k))=y} \tilde{\mathbb{P}}(t, x)$$

Implement this once and for all in place of empirical probability $\tilde{\mathbb{P}}(t, x)$.

# Fast MIA Algorithm Flow

**input** : $\mathbf{x}$ a set of $m$ traces which take discrete values, $\mathbf{t}$ a corresponding set of $m$ plaintexts/ciphertexts

**output**: $(\tilde{I}(\mathbf{x}, \mathbf{y}(k)))_{k \in \mathcal{K}}$

`// From x and t, build an hash table` $\mathsf{PMF}[t][x]$ `(i.e., an histogram);`

**for** $i \in \{1, \ldots, m\}$ **do**
    $\mathsf{PMF}[t_i][x_i] \mathrel{+}= 1$;
**end**

**for** $x \in \mathcal{X}$ **do**
    $\mathsf{P}(x) = 0$ ; `// P(x) holds` $m\tilde{\mathbb{P}}(x)$
    **for** $t \in \mathbb{F}_2^n$ **do**
        $\mathsf{P}(x) \mathrel{+}= \mathsf{PMF}[t][x]$ ;
    **end**
**end**

# Fast MIA Algorithm Flow

**for** $k \in \mathcal{K}$ **do** `// Key enumeration`
- $\forall x \in \mathcal{X}, y \in \mathcal{Y}, \mathsf{P}(x,y) = 0$ ; `//` $\mathsf{P}(x,y)$ `holds` $m\tilde{\mathbb{P}}(x,y)$
- **for** $t \in \mathbb{F}_2^n$ **do**
  - **for** $x \in \mathcal{X}$ **do**
    - $\mathsf{P}(x, \varphi(f(k,t))) \mathrel{+}= \mathsf{PMF}[t][x]$ ; `//` $y = \varphi(f(k,t))$
  - **end**
- **end**
- $\tilde{I}(\mathbf{x}, \mathbf{y}(k)) = 0;$
- **for** $y \in \mathcal{Y}$ **do**
  - $\mathsf{P}(y) = 0$ ; `//` $\mathsf{P}(y)$ `holds` $m\tilde{\mathbb{P}}(y)$
  - **for** $x \in \mathcal{X}$ **do**
    - $\mathsf{P}(y) \mathrel{+}= \mathsf{P}(x,y);$
  - **end**
  - **for** $x \in \mathcal{X}$ **do**
    - **if** $\mathsf{P}(x,y) \neq 0$ **then** `//` $\mathsf{P}(x,y) \neq 0 \implies (\mathsf{P}(x) \neq 0 \wedge \mathsf{P}(y) \neq 0)$
      - $\tilde{I}(\mathbf{x}, \mathbf{y}(k)) \mathrel{+}= \frac{\mathsf{P}(x,y)}{m} \log_2 \frac{m\mathsf{P}(x,y)}{\mathsf{P}(x)\mathsf{P}(y)}$ ;
    - **end**
  - **end**
- **end**

**end**

**return** $(\tilde{I}(\mathbf{x}, \mathbf{y}(k)))_{k \in \mathcal{K}}$

# Fast MIA Algorithm Flow

**for** $k \in \mathcal{K}$ **do** `// Key enumeration`

- $\forall x \in \mathcal{X}, y \in \mathcal{Y}, \mathsf{P}(x,y) = 0$ ; `//` $\mathsf{P}(x,y)$ `holds` $m\tilde{\mathbb{P}}(x,y)$
- **for** $t \in \mathbb{F}_2^n$ **do**
  - **for** $x \in \mathcal{X}$ **do**
    - $\mathsf{P}(x, \varphi(f(k,t))) \mathrel{+}= \mathsf{PMF}[t][x]$ ; `//` $y = \varphi(f(k,t))$
  - **end**
- **end**
- $\tilde{I}(\mathbf{x}, \mathbf{y}(k)) = 0$;
- **for** $y \in \mathcal{Y}$ **do**
  - $\mathsf{P}(y) = 0$ ; `//` $\mathsf{P}(y)$ `holds` $m\tilde{\mathbb{P}}(y)$
  - **for** $x \in \mathcal{X}$ **do**
    - $\mathsf{P}(y) \mathrel{+}= \mathsf{P}(x,y)$;
  - **end**
  - **for** $x \in \mathcal{X}$ **do**
    - **if** $\mathsf{P}(x,y) \neq 0$ **then** `//` $\mathsf{P}(x,y) \neq 0 \implies (\mathsf{P}(x) \neq 0 \wedge \mathsf{P}(y) \neq 0)$
      - $\tilde{I}(\mathbf{x}, \mathbf{y}(k)) \mathrel{+}= \mathsf{P}(x,y) \log_2 \frac{\mathsf{P}(x,y)}{\mathsf{P}(y)}$ ; `// no longer a mutual information, but OK...`
    - **end**
  - **end**
- **end**

**end**

**return** $(\tilde{I}(\mathbf{x}, \mathbf{y}(k)))_{k \in \mathcal{K}}$

# Conclusions

- MIA is equivalent to Universal ML
- MIA can be close to optimal
- Binning size is crucial
- Fast MIA

# Thank you!

Questions?

# References


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.

Annelie Heuser, Olivier Rioul, and Sylvain Guilley.
Good Is Not Good Enough - Deriving Optimal Distinguishers from Communication Theory.
In Lejla Batina and Matthew Robshaw, 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, 2014.

Victor Lomné, Emmanuel Prouff, and Thomas Roche.
Behind the Scene of Side Channel Attacks.
In Kazue Sako and Palash Sarkar, editors, *ASIACRYPT (1)*, volume 8269 of *Lecture Notes in Computer Science*, pages 506–525. Springer, 2013.