Masking against Side-Channel Attacks: a Formal Security Proof

Matthieu Rivain

Joint work with Emmanuel Prouff

EUROCRYPT 2013 – May 27th

CryptoExperts

WE INNOVATE TO SECURE YOUR BUSINESS

Outline

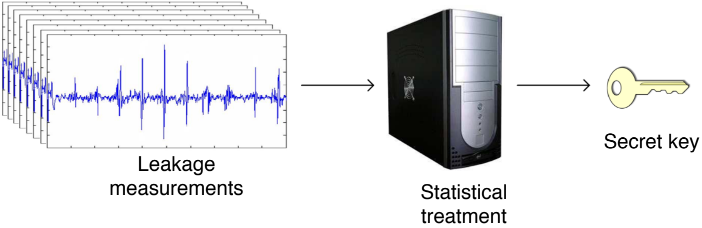
CryptoExperts

Outline

CryptoExperts

Side-Channel Attacks

- Attacks exploiting physical information leakage
  - timing [Kocher. CRYPTO'96]
  - power consumption [Kocher et al. CRYPTO'99]
  - electromagnetic emanations [Gandolfi et al. CHES'01]



CryptoExperts

Masking

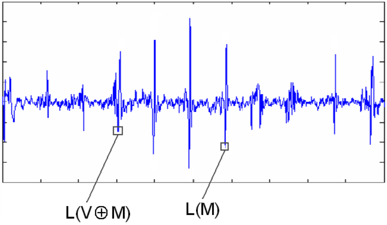
- [Chari et al. CRYPTO'99] [Goubin-Patarin. CHES'99]
- Apply secret sharing to internal variables
- A sensitive variable x is shared into $d + 1$ variables

$$x_0 \oplus x_1 \oplus \cdots \oplus x_d = x$$

- Computing on each share separately

Masking Schemes

- A lot of *first-order* masking schemes have been published
  - [Kocher et al. US Patent 1999] [Goubin-Patarin. CHES'99] [Messerges. FSE'00] [Akkar-Giraud. CHES'01] [Blomer et al. SAC'04] [Oswald et al. FSE'05] [Prouff et al. CHES'06] [Prouff-Rivain. WISA'07]
- Used in current smart cards products
- Limitation: vulnerable to *second-order SCA*



Masking Schemes

- Increasing masking order
  - $\Rightarrow$ increasing attack order
  - $\Rightarrow$ increasing attack difficulty
- Soundness [Chari et al. CRYPTO'99]
  - Noisy leakage model: $L_i \sim x_i + \mathcal{N}(\mu, \sigma^2)$
  - Distinguishing $\big((L_i)_i | x = 0\big)$ from $\big((L_i)_i | x = 1\big)$ takes q samples:

$$q \geq cst \cdot \sigma^d$$

- Higher-order masking schemes
  - [Rivain-Prouff. CHES'10] [Kim et al. CHES'11] [Carlet et al. FSE'12] [Coron et al. FSE'13]
- Limitation: no security proof against an adversary using the whole leakage of the computation

Physically Observable Cryptography

- [Micali-Reyzin. TCC'04]
- Framework for leaking computation
- Assumption: *Only Computation Leaks (OCL)*
- Computation divided into subcomputations $y \leftarrow \mathsf{C}(x)$
- Each subcomputation leaks a function of its input $f(x)$

Leakage Functions

- Leakage-Resilience model [Dziembowski-Pietrzak. STOC'08]
  - bounded-range leakage functions

$$f : \{0,1\}^n \to \{0,1\}^\lambda \qquad \text{with} \quad \lambda \ll n$$

- Leakage model for circuits [Faust et al. EUROCRYPT'10]
  - computationally bounded leakage functions: $f \in \mathcal{AC}^0$
    (computable by a circuit of constant depth)
  - noisy leakage functions: $f(x) = x \oplus \varepsilon$
    with ε being some sparse error vector

CryptoExperts

Limitations

- In practice the leakage is far bigger than n bits ($\lambda \gg n$)

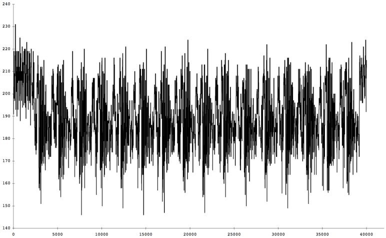
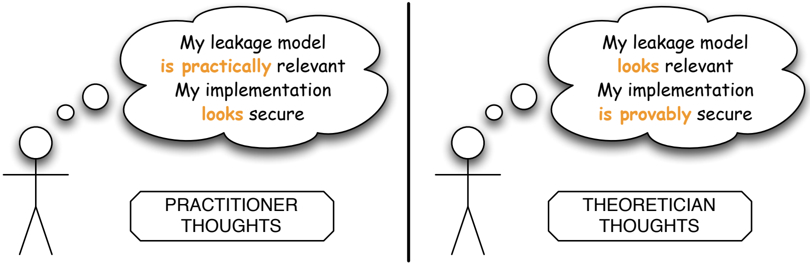


Figure: Power consumption of a DES computation.

- The leakage result from the switching activity of logic gates
  - it can hardly be modeled by an $\mathcal{AC}^0$ function
  - noise can hardly be modeled as the xor of an error vector

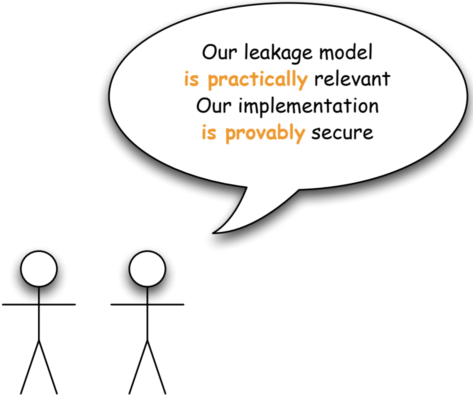
State of the Art

- Lack of practically relevant leakage models
- Masking widely used without formal proof



Our Goal

- A step toward:



CryptoExperts

Outline

Our Contribution

- Leakage model
  - OCL assumption [Micali-Reyzin. TCC'04]
  - subcomputations = elementary calculations
    (a few CPU intructions, small inputs)
- New class of noisy leakage functions
  - $f(x)$ implies a *bounded bias* in the distribution of x

Our Contribution

- Formal security proof for a block cipher computation
  - negligible entropy loss on the key (w.r.t. masking order)

- Need for a *leak-free component* (for mask refreshing)

$$\underbrace{\boldsymbol{x} = (x_0, x_1, \ldots, x_d)}_{\bigoplus_i x_i = x} \longmapsto \underbrace{\boldsymbol{x}' = (x'_0, x'_1, \ldots, x'_d)}_{\bigoplus_i x'_i = x}$$

with $(\boldsymbol{x} \mid x)$ and $(\boldsymbol{x}' \mid x)$ mutually independent.

Outline

CryptoExperts

Notion of Bias

- Bias of X given $Y = y$:

$$\beta(X|Y = y) = \|\mathrm{P}[X] - \mathrm{P}[X|Y = y]\|$$

with $\|\cdot\| =$ Euclidean norm.

- Bias of X given Y:

$$\beta(X|Y) = \sum_{y \in \mathcal{Y}} \mathrm{P}[Y = y]\, \beta(X|Y = y)\ .$$

- Related to MI by:

$$\mathrm{MI}(X; Y) \leq \frac{N}{\ln 2} \beta(X|Y) \qquad \text{(with } N = |\mathcal{X}|\text{)}$$

Model of Leaking Computation

- Every elementary calculation leaks a *noisy function* of its input
  - noise modeled by a fresh random tape argument
- f adaptively chosen by the adversary in $\mathcal{N}(1/\psi)$

$$\beta\big(X|f(X)\big) < \frac{1}{\psi}$$

- ψ is some *noise parameter*
- Capture any form of noisy leakage
- Assumtpion: ψ can be set by the designer (linear in the security parameter)

CryptoExperts

Outline

CryptoExperts

Overview of the Proof

- Consider a SPN computation

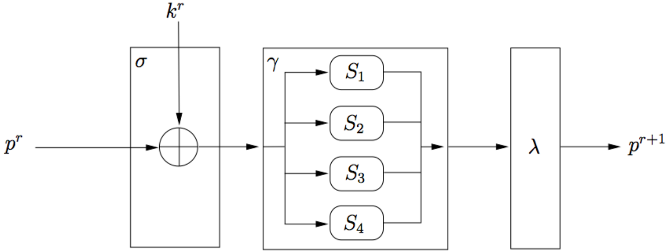


Figure: Example of SPN round.

Overview of the Proof

- Classical implementation protected with masking

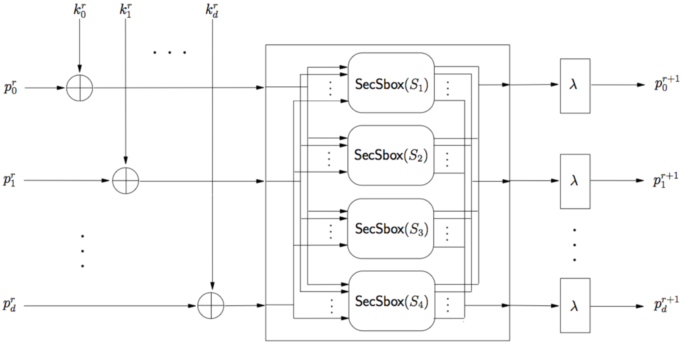


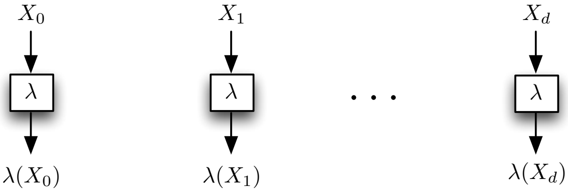
Figure: Example of SPN round protected with masking.

S-Box Computation

- [Carlet et al. FSE'12]
- Polynomial evaluation over GF(2^n)
- Two types of elementary calculations:
  - linear functions (additions, squares, multiplication by coefficients)
  - multiplications over GF(2^n)

Linear Functions

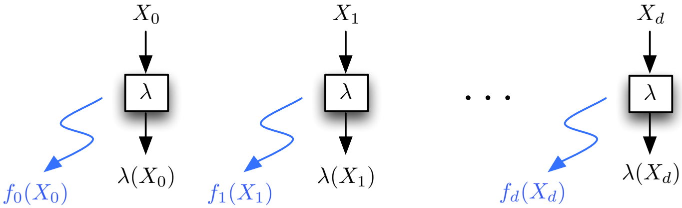
- Given a sharing $X = X_0 \oplus X_1 \oplus \cdots \oplus X_d$



- Apply mask-refreshing on output sharing

Linear Functions

- Given a sharing $X = X_0 \oplus X_1 \oplus \cdots \oplus X_d$



- Apply mask-refreshing on output sharing

Linear Functions

- For $f_0, f_1, \ldots, f_d \in \mathcal{N}(1/\psi)$, we show

$$\beta\big(X \big| f_0(X_0), f_1(X_1), \ldots, f_d(X_d)\big) \leq \frac{N^{\frac{d}{2}}}{\psi^{d+1}} \ .$$

- Taking $\psi \sim N^{\frac{1}{2}}\,\omega$ we get

$$\mathrm{MI}\big(X;\ (f_0(X_0), f_1(X_1), \ldots, f_d(X_d))\big) \leq \frac{1}{\omega^{d+1}}$$

- Result in accordance with [Chari et al. CRYPTO'99]

Multiplications

- Given two sharings $A = \bigoplus_i A_i$ and $B = \bigoplus_i B_i$

$$A \times B = \left(\bigoplus_i A_i\right)\left(\bigoplus_i B_i\right) = \bigoplus_{i,j} A_i B_j$$

- First step: cross-products

$$\begin{array}{cccc} A_0 \times B_0 & A_0 \times B_1 & \cdots & A_0 \times B_d \\ A_1 \times B_0 & A_1 \times B_1 & \cdots & A_1 \times B_d \\ \vdots & \vdots & \ddots & \vdots \\ A_d \times B_0 & A_d \times B_1 & \cdots & A_d \times B_d \end{array}$$

Multiplications

- Given two sharings $A = \bigoplus_i A_i$ and $B = \bigoplus_i B_i$

$$A \times B = \Big(\bigoplus_i A_i\Big)\Big(\bigoplus_i B_i\Big) = \bigoplus_{i,j} A_i B_j$$

- First step: cross-products

$$\begin{matrix} A_0 \times B_0 & A_0 \times B_1 & \cdots & A_0 \times B_d \\ A_1 \times B_0 & A_1 \times B_1 & \cdots & A_1 \times B_d \\ \vdots & \vdots & \ddots & \vdots \\ A_d \times B_0 & A_d \times B_1 & \cdots & A_d \times B_d \end{matrix}$$

$$\begin{matrix} f_{0,0}(A_0, B_0) & f_{0,1}(A_0, B_1) & \cdots & f_{0,d}(A_0, B_d) \\ f_{1,0}(A_1, B_0) & f_{1,1}(A_1, B_1) & \cdots & f_{1,d}(A_1, B_d) \\ \vdots & \vdots & \ddots & \vdots \\ f_{d,0}(A_d, B_0) & f_{d,1}(A_d, B_1) & \cdots & f_{d,d}(A_d, B_d) \end{matrix}$$

Multiplications

- We have $A = g(X)$ and $B = h(X)$ where $X =$ s-box input
- Bias given cross-product leakages:

  For $f_{i,j} \in \mathcal{N}(1/\psi)$ we show

  $$\beta\Big(X \big| (f_{i,j}(A_i, B_j))_{i,j}\Big) \leq 2N^{\frac{3d+7}{2}} \left(\frac{\lambda_1 d + \lambda_0}{\psi}\right)^{d+1}$$

  with $\lambda_1 \in [1;2]$ and $\lambda_2 \in [1;3]$.
- Taking $\psi \sim N^{\frac{3}{2}}(\lambda_1 d + \lambda_0)\omega$ we get

  $$\mathrm{MI}\Big(X;\ (f_{i,j}(A_i, B_j))_{i,j}\Big) \leq \frac{1}{\omega^{d+1}}$$

- The noise parameter must be roughly multiplied by d

Multiplications

- Second step: refreshing
- Apply on each column and one row of

$$\begin{matrix} A_0 \times B_0 & A_0 \times B_1 & \cdots & A_0 \times B_d \\ A_1 \times B_0 & A_1 \times B_1 & \cdots & A_1 \times B_d \\ \vdots & \vdots & \ddots & \vdots \\ A_d \times B_0 & A_d \times B_1 & \cdots & A_d \times B_d \end{matrix}$$

- We get a fresh $(d+1)^2$-sharing of $A \times B$

$$\begin{matrix} V_{0,0} & V_{0,1} & \cdots & V_{0,d} \\ V_{1,0} & V_{1,1} & \cdots & V_{1,d} \\ \vdots & \vdots & \ddots & \vdots \\ V_{d,0} & V_{d,1} & \cdots & V_{d,d} \end{matrix}$$

Multiplications

- Third step: summing rows

$$Z_i \leftarrow V_{i,0} \oplus V_{i,1} \oplus \cdots \oplus V_{i,d}$$

- Takes d elementary calculations (XORs) per row:

$$\begin{aligned} T_{i,1} &\leftarrow V_{i,0} \oplus V_{i,1} \\ T_{i,2} &\leftarrow T_{i,1} \oplus V_{i,2} \\ &\vdots \\ T_{i,d} &\leftarrow T_{i,d-1} \oplus V_{i,d} \end{aligned}$$

  (with $Z_i = T_{i,d}$)

- Then $(Z_0, Z_1, \ldots, Z_d)$ is a sharing of $A \times B$
  - Apply mask-refreshing

CryptoExperts

Multiplications

- Third step: summing rows

$$Z_i \leftarrow V_{i,0} \oplus V_{i,1} \oplus \cdots \oplus V_{i,d}$$

- Takes d elementary calculations (XORs) per row:

$$\begin{aligned}
T_{i,1} &\leftarrow V_{i,0} \oplus V_{i,1} &&\rightsquigarrow f_{i,1}(V_{i,0}, V_{i,1})\\
T_{i,2} &\leftarrow T_{i,1} \oplus V_{i,2} &&\rightsquigarrow f_{i,2}(T_{i,1}, V_{i,2})\\
&\vdots && \vdots\\
T_{i,d} &\leftarrow T_{i,d-1} \oplus V_{i,d} &&\rightsquigarrow f_{i,d}(T_{i,d-1}, V_{i,d})
\end{aligned}$$

(with $Z_i = T_{i,d}$)

- Then $(Z_0, Z_1, \ldots, Z_d)$ is a sharing of $A \times B$
  - Apply mask-refreshing

Multiplications

- For $f_{i,j} \in \mathcal{N}(1/\psi)$ we show

$$\beta\big(X | F_0(Z_0), F_1(Z_1), \dots, F_d(Z_d)\big) \leq N^{\frac{3d+5}{2}} \left(\frac{2}{\psi}\right)^{d+1}$$

where $F_i(Z_i) = \big(f_{i,1}(V_{i,0}, V_{i,1}), f_{i,2}(T_{i,1}, V_{i,2}), \dots, f_{i,d}(T_{i,d-1}, V_{i,d})\big)$

- Taking $\psi \sim 2N^{\frac{3}{2}}\omega$ we get

$$\mathrm{MI}\big(X;\ (F_0(Z_0), F_1(Z_1), \dots, F_d(Z_d))\big) \leq \frac{1}{\omega^{d+1}}$$

CryptoExperts

Putting everything together

- Several subsequences of elementary calculations
- Each provides some leakage L_t about $X_t = g_t(M, K)$
- L_t are mutually independent given (M, K)

$$\mathrm{MI}\big((M, K); (L_1, L_2, \ldots, L_T)\big) \leq \sum_{t=1}^{T} \mathrm{MI}(X_t; L_t) \leq \frac{T}{\omega^{d+1}}$$

Outline

1. Introduction and Previous Works
2. Our Contribution
3. Model of Leaking Computation
4. Overview of the Proof
5. Conclusion and Perspectives

CryptoExperts

Conclusion and Perspectives

Conclusion:

- New practically relevant leakage model
- Formal security for masking against SCA

Perspectives and open issues:

- Practical estimation of the noise parameter ψ
- Relax proof assumptions:
  - fixed noise parameter
  - no leak-free component

Conclusion and Perspectives

- What about efficiency?