TELECOM ParisTech

Institut Mines-Télécom

Confusing Information:
How Confusion Improves Side-Channel Analysis for Monobit Leakages

Cryptarchi 2018
June 17-19, 2018
Guidel-Plage, France

Eloi de Chérisey, Sylvain Guilley & Olivier Rioul

Télécom ParisTech, Université Paris-Saclay, France.

Contents

Contents

Motivation

- what is the exact link between side-channel distinguishers and the confusion coefficient for monobit leakages?

Motivation

- what is the exact link between side-channel distinguishers and the confusion coefficient for monobit leakages?
- re-derive it for DoM, CPA, KSA and derive it for MIA;

Motivation

- what is the exact link between side-channel distinguishers and the confusion coefficient for monobit leakages?
- re-derive it for DoM, CPA, KSA and derive it for MIA;
- is any sound distinguisher a function of the confusion coefficient (and noise) ?

Leakage Model

Definition (Leakage Sample)

$$X = Y(k^*) + N$$

where

$$Y(k) = f(k, T)$$

is the sensitive variable.

Notations:

- T a random plain or cyphertext;
- k^* the secret key;
- N some additive noise;
- f a deterministic function.

Assumptions

W.l.o.g. assume

- $Y(k) = \pm 1$ equiprobable:
  - zero mean $\mathbb{E}[Y(k)] = 0$ and unit variance $\mathbb{E}[Y(k)^2] = 1$
  - $\mathbb{P}(Y(k) = -1) = \mathbb{P}(Y(k) = 1) = 1/2$
- Gaussian noise $N \sim \mathcal{N}(0, \sigma^2)$.

Definition (Distinguisher)

Practical distinguisher : $\hat{\mathcal{D}}(k)$
Theoretical distinguisher : $\mathcal{D}(k)$.

$$\hat{k} = \arg\max \hat{\mathcal{D}}(k)$$

The estimated key maximizes $\mathcal{D}(k)$.
If sound, $\arg\max \hat{\mathcal{D}}(k) = k^*$.

Fei et al.'s Confusion Coefficient

After [Fei et al., 2014].

Definition (Confusion Coefficient)

$$\kappa(k, k^*) = \kappa(k) = \mathbb{P}(Y(k) \neq Y(k^*))$$

valid only for monobit leakages (DoM).

Confusion and Security

From [Heuser et al., 2014].

Theorem (Differential Uniformity)

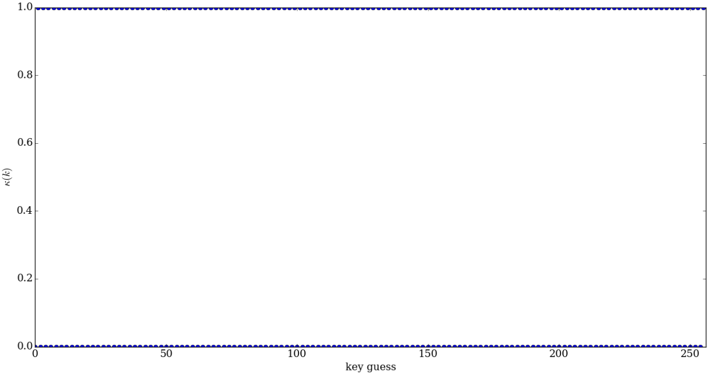
The differential uniformity of an S-box is linked with the confusion coefficient by:

$$2^{-n}\Delta_S - \frac{1}{2} = \max_{k \neq k^*} \left|\frac{1}{2} - \kappa(k)\right|$$

$\implies$ a “good” S-box should have confusion coefficient near $\frac{1}{2}$.

Illustration Without Permutation

Example with $Y(k) = T \oplus k \bmod 2$



$k^* = 54.$

Illustration for Random Permutation

Example with $Y(k) = \textcolor{red}{\mathrm{RP}}(T \oplus k) \bmod 2$

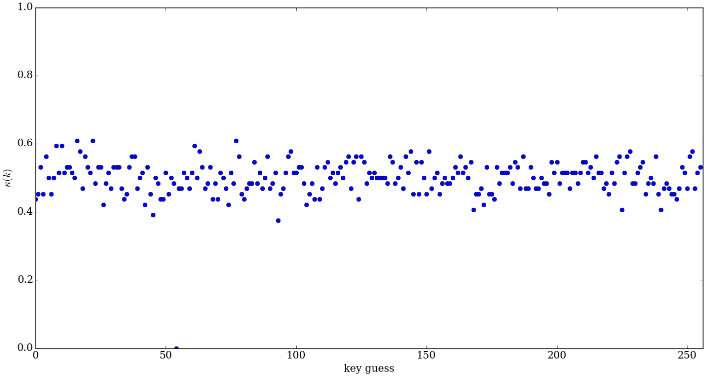
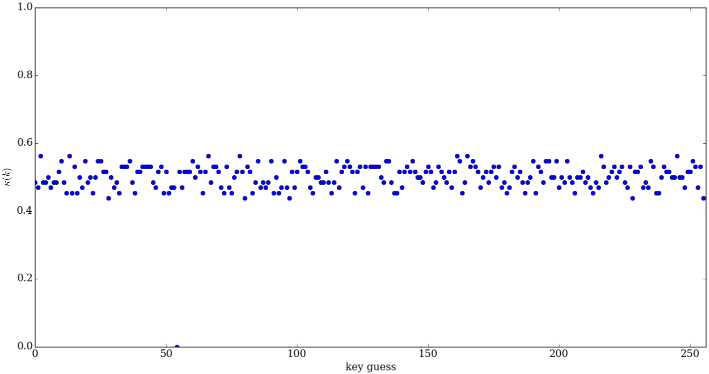


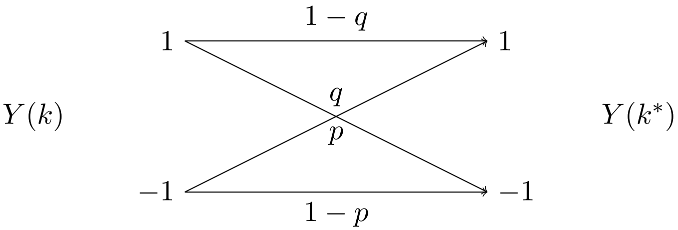
Illustration for AES S-box

Example with $Y(k) = \mathrm{S_{box}}(T \oplus k) \bmod 2$



Contents

A Confusion Channel from $Y(k)$ to $Y(k^*)$



Since $\mathbb{P}(Y(k^*) = -1) = (1-p)\mathbb{P}(Y(k) = -1) + q\mathbb{P}(Y(k) = 1) =$ $P(Y(k^*) = 1) = (1-q)\mathbb{P}(Y(k) = 1) + p\mathbb{P}(Y(k) = 1)$, we have:

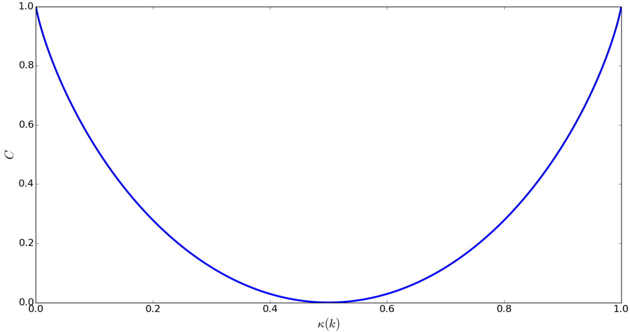
$$p = q = \kappa(k).$$

This is a binary symmetric channel (BSC).

Confusion Channel's Capacity

Since $Y(k)$ is equiprobable, the mutual information of the BSC equals its capacity:

$$C(k) = I(Y(k^*); Y(k)) = 1 - H_2(\kappa(k))$$



A General Result for any Distinguisher

Theorem (Monobit Leakage Distinguisher)

The theoretical distinguisher of any monobit leakage is a function of $\kappa(k)$ and σ.

Proof.

The theoretical distinguisher depends on the joint distribution of X and $Y(k)$:

$$\begin{aligned}\mathbb{P}(X, Y(k)) &= \mathbb{P}(Y(k^*) + N; Y(k)) = \mathbb{P}(Y(k)) \cdot \mathbb{P}(Y(k^*) + N \mid Y(k)) \\ &= \mathbb{P}(\mathcal{B}_{1/2}) \cdot \mathbb{P}(\mathcal{B}_{\kappa(k)} + N)\end{aligned}$$

where $N \sim \mathcal{N}(0, \sigma^2)$. □

Contents

Difference of Means (DoM)

Definition (DoM)

Practical distinguisher:

$$\hat{\mathcal{D}}(k) = \frac{\sum_{q/Y(k)=1} X_q}{\sum_{q/Y(k)=1} 1} - \frac{\sum_{q/Y(k)=-1} X_q}{\sum_{q/Y(k)=-1} 1}.$$

Theoretical distinguisher:

$$\mathcal{D}(k) = \mathbb{E}[XY(k)]$$

DoM Computation

We have:

$$\begin{aligned}\mathcal{D}(k) &= \mathbb{E}[XY(k)] \\ &= \mathbb{E}[(Y(k^*) + N)Y(k)] \\ &= \mathbb{E}[Y(k)Y(k^*)] \\ &= \mathbb{E}[2_{Y(k)=Y(k^*)} - 1] \\ &= 2(1 - \kappa(k)) - 1 \\ &= 1 - 2\kappa(k).\end{aligned}$$

Therefore:

$$\boxed{\mathcal{D}(k) = 2\Big(\frac{1}{2} - \kappa(k)\Big)}$$

Correlation Power Analysis (CPA)

Definition (CPA)

Practical distinguisher: Pearson coefficient

$$\hat{\mathcal{D}}(k) = \frac{|\hat{\mathbb{E}}[XY(k)] - \hat{\mathbb{E}}[X]\hat{\mathbb{E}}[Y(k)]|}{\hat{\sigma}_X \hat{\sigma}_{Y(k)}},$$

Theoretical distinguisher:

$$\mathcal{D}(k) = \frac{|\mathbb{E}[XY(k)] - \mathbb{E}[X]\mathbb{E}[Y(k)]|}{\sigma_X \sigma_{Y(k)}},$$

which is the correlation coefficient between X and $Y(k)$.

CPA Computation

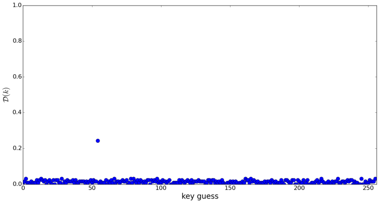
Since $\mathbb{E}[Y(k)] = 0$ and $\sigma_{Y(k)} = 1$, we have:

$$\mathcal{D}(k) = \frac{\mathbb{E}[XY(k)] - \mathbb{E}[X]\mathbb{E}[Y(k)]}{\sigma_X \sigma_{Y(k)}} = \frac{|\mathbb{E}[XY(k)]|}{\sigma_X}.$$

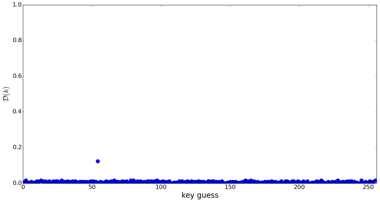
From the DoM computation and since $\sigma_X^2 = 1 + \sigma^2$, we have:

$$\boxed{\mathcal{D}(k) = \frac{2|1/2 - \kappa(k)|}{\sqrt{1+\sigma^2}}.}$$

Illustration for AES SubBytes w.r.t. Noise

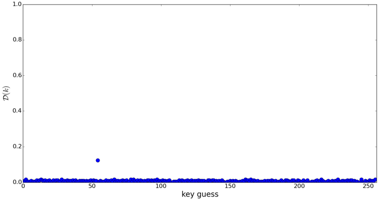


$\sigma = 4$

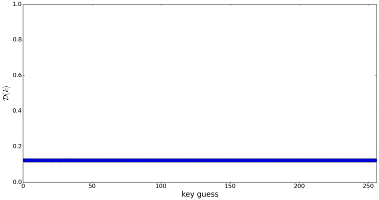


$\sigma = 8$

Illustration for $\sigma = 8$ w.r.t. SubBytes



AES Subbytes



no Subbyte

Kolmogorov-Smirnov Analysis (KSA)

Definition (KSA)

Practical Distinguisher:

$$\hat{\mathcal{D}}(k) = \mathbb{E}_{Y(k)} \| \hat{F}(x|Y(k)) - \hat{F}(x) \|_\infty$$

Theoretical Distinguisher:

$$\mathcal{D}(k) = \mathbb{E}_{Y(k)} \| F(x|Y(k)) - F(x) \|_\infty$$

where:

- $F(x)$ and $F(x \mid Y(k))$ the cummulative distibution functions of X and $X \mid Y(k)$.
- $\|f(x)\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|$.

KSA Computation

Theorem (KSA and Confusion [Heuser et al., 2014])

With our assumptions, we have:

$$\mathcal{D}(k) = \operatorname{erf}\left(\sqrt{\frac{\mathsf{SNR}}{2}}\right) \left|\frac{1}{2} - \kappa(k)\right|$$

where $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{-\infty}^{x} e^{-t^2} \mathrm{d}t$.

Mutual Information Analysis (MIA)

Definition (MIA)

Practical Distinguisher: $\hat{\mathcal{D}}(k) = \hat{I}(X; Y(k))$

Theoretical Distinguisher: $\mathcal{D}(k) = I(X; Y(k)) = h(X) - h(X|Y(k))$

Theorem (MIA Computation (Main result))

For a monobit leakage:

$$\mathcal{D}(k) = 2\log_2(e)\Big(\frac{1}{2} - \kappa(k)\Big)^2 f(\sigma).$$

where f is such that $f(\sigma) \to 1$ when $\sigma \to 0$ and $f(\sigma) \sim 1/\sigma^2$ as $\sigma \to \infty$.

Main Result: Sketch of the Proof

$$
\begin{aligned}
I(X;Y(k)) &= h(X) - h(X \mid Y(k)) \\
&= h(\mathcal{B}'_{1/2} + N) - H(\mathcal{B}'_{\kappa(k)} + N)
\end{aligned}
$$

Case 1: Very high SNR ($\sigma \to 0$)

$$
h(\mathcal{B}'_{1/2} + N) \approx H(\mathcal{B}'_{1/2}) + h(N)
$$

$$
H(\mathcal{B}'_{\kappa(k)} + N) \approx H(\mathcal{B}'_{\kappa(k)}) + h(N)
$$

$$
\mathcal{D}(k) \approx 1 - H(\mathcal{B}'_{\kappa(k)}) = 1 - H_2(\kappa(k))
$$

Second order Taylor expansion about 1/2:

$$
\boxed{\mathcal{D}(k) \approx 2\log(e)(1/2 - \kappa(k))^2}
$$

Main Result: Sketch of the Proof (Cont'd)

Case 2: Very low SNR ($\sigma \to +\infty$)
All signals behaves like Gaussian.

$$\begin{aligned}
\mathcal{D}(k) &= h(\mathcal{B}'_{1/2} + N) - h(\mathcal{B}'_{\kappa(k)} + N) \\
&\approx \frac{1}{2}\log_2(2\pi e(\sigma^2+1)) - \frac{1}{2}\log_2(2\pi e(\sigma^2 + 4\kappa(k)(1-\kappa(k))) \\
&= \frac{1}{2}\log_2 \frac{\sigma^2+1}{\sigma^2 + 4\kappa(k)(1-\kappa(k))} \\
&= -\frac{1}{2}\log_2 \frac{\sigma^2 + 1 + 4\kappa(k)(1-\kappa(k)) - 1}{\sigma^2+1} \\
&\approx \frac{\log_2(e)}{2}\frac{4\kappa(k)(1-\kappa(k)) - 1}{\sigma^2+1} = \boxed{2\log_2(e)\frac{(1/2-\kappa(k))^2}{\sigma^2}}
\end{aligned}$$

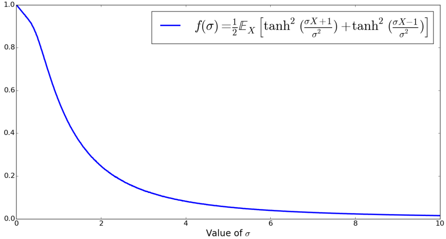
Main Result: Sketch of the Proof (Cont'd)

General Case: any SNR, $\kappa \approx 1/2$

Theorem

$$\mathcal{D}(k) = 2(\log_2 e)\Big(\frac{1}{2} - \kappa(k)\Big)^2 \frac{1}{2}\mathbb{E}_X\Big[\tanh^2(\frac{\sigma X + 1}{\sigma^2}) + \tanh^2(\frac{\sigma X - 1}{\sigma^2})\Big]$$

where $X \sim \mathcal{N}(0,1)$ is standard normal.



Contents

Conclusion

A unified view of side-channel distinguishers on monobit leakages:

- DoM $\frac{1}{2}(1/2 - \kappa(k))$;
- CPA $\frac{|1/2-\kappa(k)|}{1+\sigma^2}$;
- KSA $|1/2 - \kappa(k)|\mathrm{erf}\left(\sqrt{\frac{\mathsf{SNR}}{2}}\right)$;
- MIA $2(\log_2 e)(1/2 - \kappa(k))^2 f(\sigma)$.

TELECOM ParisTech

Institut Mines-Télécom

Confusing Information:
How Confusion Improves Side-Channel Analysis for Monobit Leakages

Cryptarchi 2018
June 17-19, 2018
Guidel-Plage, France

Eloi de Chérisey, Sylvain Guilley & Olivier Rioul

Télécom ParisTech, Université Paris-Saclay, France.

References I

Fei, Y., Ding, A. A., Lao, J., and Zhang, L. (2014).
A statistics-based fundamental model for side-channel attack analysis.
Cryptology ePrint Archive, Report 2014/152.
http://eprint.iacr.org/2014/152.

Heuser, A., Rioul, O., and Guilley, S. (2014).
A Theoretical Study of Kolmogorov-Smirnov Distinguishers — Side-Channel Analysis vs. Differential Cryptanalysis.
In Prouff, E., editor, *Constructive Side-Channel Analysis and Secure Design - 5th International Workshop, COSADE 2014, Paris, France, April 13-15, 2014. Revised Selected Papers*, volume

References II

8622 of *Lecture Notes in Computer Science*, pages 9–28. Springer.