Guessing a Secret Cryptographic Key from Side-Channel Leakages

Wei Cheng[1], Olivier Rioul[1], and Sylvain Guilley[1,2]

[1] LTCI, Télécom Paris, Institut Polytechnique de Paris, France

[2] Secure-IC S.A.S., 35510 Cesson-Sévigné, France

Email: wei.cheng@telecom-paristech.fr

Abstract

We experiment relative merits of information-theoretic metrics such as guessing entropy, conditional Shannon or Rényi entropies vs. success probability, in the problem of guessing a cryptographic key form a leakage in some practical cryptosystems, with Hamming weight leakage model in additive (Gaussian) measurement noise.

This is ongoing work with Sylvain Guilley (Telecom Paris, Secure-IC) and Olivier Rioul (Telecom Paris)

Keywords. *Guessing entropy, Conditional Shannon entropy, Rényi entropy, Success probability*

Information-theoretic metrics

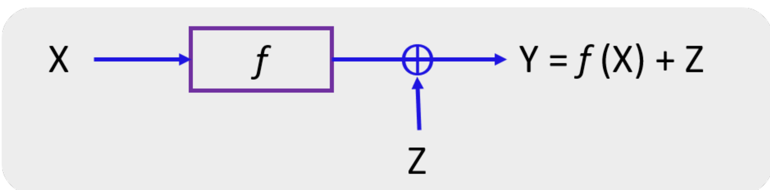

Figure 1: Leakage model: secret X, noise Z and leakage Y

Let X be a *discrete* random variable with probability distribution $p(x)$. Without loss of generality we may suppose that $X \in \{1, 2, \ldots, n, \ldots\}$ with respective probabilities $p_1, p_2, \ldots, p_n, \ldots$. Let $Y = f(X) + Z$ be additional information (*leakage*) about X. If noise Z is present, Y is a continuous r.v. with density $p(y)$, while in the noiseless case ($Z = 0$), Y is discrete with distribution $p(y)$. The attacker knows Y and guesses X. We have the following metrics:

- **(Conditional) Guessing entropy**: letting $p_k = p(x = k)$, $k = 1, 2, \ldots, n, \ldots$, we have the (conditional) guessing entropies $G(X)$ and $G(X|Y)$ as:

$$G(X) = \sum_k k p_{(k)}, \qquad G(X|Y) = \sum\!\!\!\!\!\!\int p(y) G(X|Y = y) \tag{1}$$

where the probabilities are arranged in decreasing order $p_{(1)} \geq p_{(2)} \geq \cdots \geq p_{(n)} \geq \cdots$.

- **(Conditional) Shannon Entropies**:

$$\begin{aligned} H(X) &= \sum_{x \in \mathcal{X}} p(x) \log_2 \frac{1}{p(x)} \\ H(X|Y) &= \sum\!\!\!\!\!\!\int_{y \in \mathcal{Y}} p(y) \sum_{x \in \mathcal{X}} p(x|y) \log \frac{1}{p(x|y)} \end{aligned} \tag{2}$$

- **(Conditional) Arimoto-Rényi Entropies**:

$$\begin{aligned} H_\alpha(X) &= \frac{\alpha}{1-\alpha} \log \Big(\sum_x p(x)^\alpha\Big)^{1/\alpha} \\ H_\alpha(X|Y) &= \frac{\alpha}{1-\alpha} \log \sum\!\!\!\!\!\!\int_{y \in \mathcal{Y}} p(y) \Big(\sum_x p(x|y)^\alpha\Big)^{1/\alpha} \end{aligned} \tag{3}$$

- **(Conditional) Success probability**:

$$P_s(X) = \max_x p(x), \qquad P_s(X|Y) = \sum\!\!\!\!\!\!\int_{y \in \mathcal{Y}} p(y) \max_x p(x|y) \geq P_s(X) \tag{4}$$

Guessing X with Noiseless Hamming Weight Leakages

Hamming weight leakage model $f = w_H$ is one of the most general leakage model used in side-channel analysis. Particularly, hardware implementations leak bits in parallel, hence the leakage is the sum of the registers state bits, that is the Hamming weight of the register contents.

Let $Y = w_H(X)$ where w_H is the Hamming weight function, in the noiseless case ($Z = 0$). We choose $|\mathcal{X}| = M = 2^n$ for the sake of calculation.

$$p(x) = \frac{1}{2^n}, \qquad p(y) = \frac{\binom{n}{y}}{2^n}, \qquad p(x|y) = \frac{\mathbf{1}_{y = w_H(x)}}{\binom{n}{y}} \tag{5}$$

We focus on quantifying the reduction of uncertainty of X knowing Y. Thus,

- **(Conditional) Guessing entropy**:

$$\begin{aligned} G(X) &= \sum_k p_k = \sum_{k=1}^{2^n} k \cdot \frac{1}{2^n} = \frac{2^n + 1}{2} \\ G(X|Y) &= \sum_y \mathbb{P}(y) \sum_x x \cdot \mathbb{P}(x|y) = \frac{1}{2} + \frac{1}{2^{n+1}} \binom{2n}{n} \approx \frac{1}{2}\left(1 + \frac{2^n}{\sqrt{\pi n}}\right) \end{aligned} \tag{6}$$

- **(Conditional) Shannon Entropies**:

$$\begin{aligned} H(X) &= \sum_x p(x) \log \frac{1}{p(x)} = \log 2^n = n \\ H(X|Y) &= -\sum_{x,y} p(x,y) \log p(x|y) = 2^{-n} \sum_y \binom{n}{y} \cdot \log \binom{n}{y} \end{aligned} \tag{7}$$

- **Conditional Rényi Entropies**:

$$H_\alpha(X|Y) = \frac{\alpha}{1-\alpha} \log \sum_y p(y) \left(\sum_x p(x|y)^\alpha\right)^{\frac{1}{\alpha}} = \frac{\alpha}{\alpha - 1}\left(n - \log \sum_y \binom{n}{y}^{\frac{1}{\alpha}}\right) \tag{8}$$

- **Conditional Success probability**:

$$P_s(X|Y) = \mathbb{E}_Y \max_x p(x|Y) = \frac{M'}{M} = \frac{n+1}{2^n} \tag{9}$$

Numerical Results on Noiseless Leakages

By upper bound from Fano's inequality and lower bound $H(X|Y) \geq \varphi^*(P_s(X|Y))$ where $\varphi^*(s) = \lfloor \frac{1}{s} \rfloor (s \lceil \frac{1}{s} \rceil - 1) \log \lfloor \frac{1}{s} \rfloor + \left(1 - \lfloor \frac{1}{s} \rfloor (s \lceil \frac{1}{s} \rceil - 1)\right) \log \lceil \frac{1}{s} \rceil$ and $H_\alpha(X|Y) \geq \frac{\alpha}{1-\alpha} \log \phi^*_\alpha(P_s(X|Y))$, where $\phi^*_\alpha(s) = \left(\lceil \frac{1}{s} \rceil s - 1\right) \lfloor \frac{1}{s} \rfloor^{1/\alpha} + \left(1 - \lfloor \frac{1}{s} \rfloor \left(\lceil \frac{1}{s} \rceil s - 1\right)\right) \lceil \frac{1}{s} \rceil^{\frac{1-\alpha}{\alpha}}$ (by Sason et al. [1]), we numerically show the conditional Shannon and Rényi entropies of X as Fig. 2. Specifically, the upper bound of Rényi entropy is highly dependent on the α. With α much larger than 1.0, the marked region is much smaller than the region with $\alpha < 1.0$.

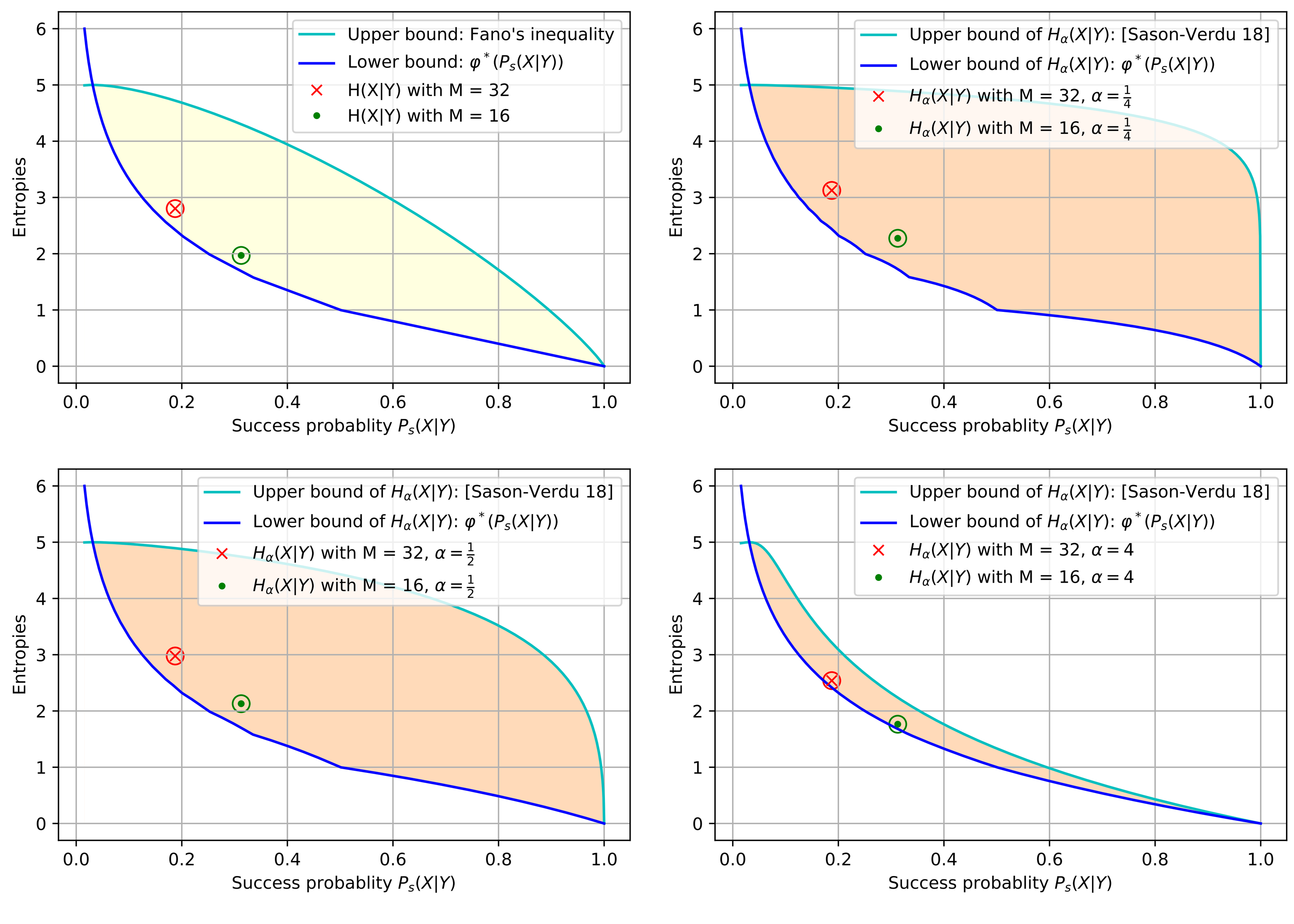

Figure 2: Conditional Shannon and Rényi Entropies of X with Hamming weight leakages

Guessing X with Noisy Hamming Weight Leakages

In fact, noise is the intrinsic part in the side-channel leakages, like power consumption and electromagnetic radiations. Thus we consider the noisy leakages in a classic way by assuming the noise is the additive white Gaussian noise (AWGN), which is a basic noise model to mimic the effect of many random processes.

We assume that $Z \sim \mathcal{N}(0, \sigma^2)$ of standard normal density $\varphi(z)$ which is a nonincreasing function of $|z|$. Thus we have:

$$\begin{aligned} p(x) &= \frac{1}{M}, \qquad p(y) = \sum_x p(x)p(y|x) = \frac{1}{M} \sum_x \varphi(y - f(x)) \\ p(y|x) &= \varphi(y - f(x)), \qquad p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{\varphi(y)}{\sum_{x'} \varphi(y - f(x'))} \end{aligned} \tag{10}$$

In addition, maximum conditional probability of success is computed as follows.

$$\begin{aligned} P_s(X|Y) &= \mathbb{E} \max_x p(x|Y) = \int \left(\frac{1}{M} \sum_{x'} \varphi(y - f(x'))\right) \times \frac{\varphi(\min_x |y - f(x)|)}{\sum_{x'} \varphi(y - f(x'))} \mathrm{d}y \\ &= \frac{1}{M} \int \varphi(y - f(x^*(y)))\, \mathrm{d}y \quad (\text{where } x^*(y) = \arg\min_x |y - f(x)|) \\ &= \frac{M'}{M} - 2\frac{M' - 1}{M} Q\left(\frac{\Delta/2}{\sigma}\right) \qquad (\text{where } Q(x) = \frac{1}{2} erfc\left(\frac{x}{\sqrt{2}}\right)) \\ H(X|Y) &= H(X) - h(Y) + h(Y|X) = \log M + \frac{1}{2} \log(2\pi e \sigma^2) - \int p(y) \log \frac{1}{p(y)} \mathrm{d}y \end{aligned} \tag{11}$$

Numerical Comparison with Lower and Upper Bounds of $G(X|Y)$

We present here six upper and lower bounds of guessing entropy of X by knowing its Hamming weight leakages. Interestingly, Bostas's upper bound is the best one which is identical to guessing entropy, which in the Hamming weight leakage scenarios.

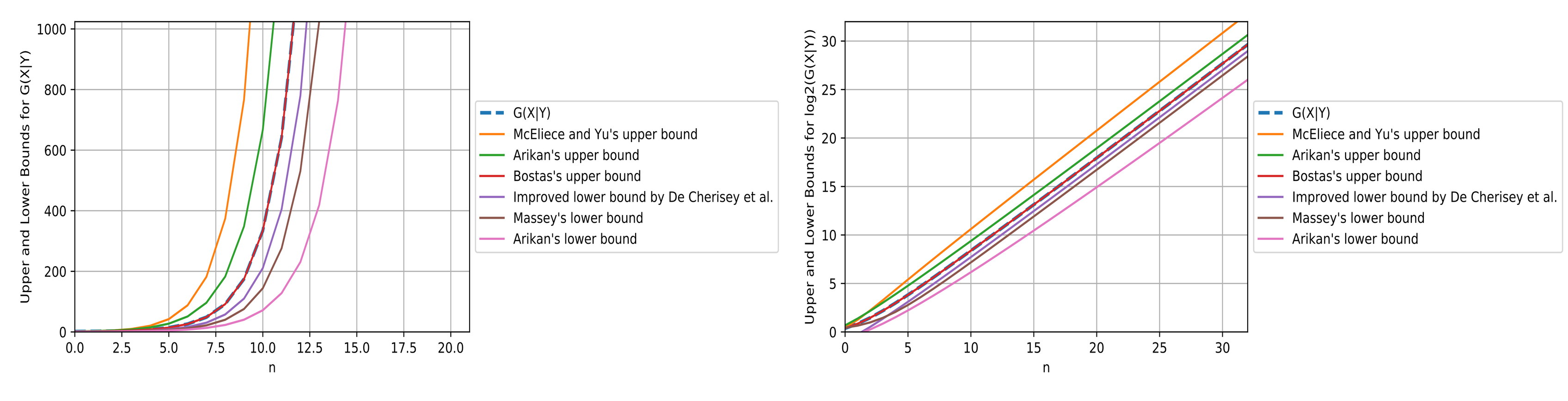

Figure 3: Comparison of six upper and lower bounds of $G(X)$

Preliminary Conclusions

We present two scenarios of guessing a secret X with Hamming weight leakages. Specifically, with small $M = 2^n$, this type of leakage has much more impact on the conditional entropies, which are the common cases in embedded systems. This explains why the Divide-and-Conquer attacks work in side-channel analysis. However, with large M, such as $M = 2^{128}$ for the AES-128 cryptographic key, the Hamming weight of whole key is of very little help for the attacker.

References

[1] I. Sason and S. Verdú, "Improved bounds on lossless source coding and guessing moments via Rényi measures," *IEEE Trans. Information Theory*, vol. 64, no. 6, pp. 4323–4346, 2018. [Online]. Available: https://doi.org/10.1109/TIT.2018.2803162

[2] E. de Chérisey, S. Guilley, O. Rioul, and P. Piantanida, "Best information is most successful: Mutual information and success rate in side-channel analysis," *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, vol. 2019, no. 2, pp. 49–79, 2019. [Online]. Available: https://doi.org/10.13154/tches.v2019.i2.49-79

[3] T. van Erven and P. Harremoës, "Rényi divergence and Kullback-Leibler divergence," *IEEE Trans. Information Theory*, vol. 60, no. 7, pp. 3797–3820, 2014. [Online]. Available: https://doi.org/10.1109/TIT.2014.2320500

[4] S. Verdú, "α-mutual information," in *2015 Information Theory and Applications Workshop, ITA 2015, San Diego, CA, USA, February 1-6, 2015*, 2015, pp. 1–6. [Online]. Available: https://doi.org/10.1109/ITA.2015.7308959