Comparison between Side-Channel Analysis Distinguishers

Houssem Maghrebi, Olivier Rioul, Sylvain Guilley, Jean-Luc Danger

Comparison between Side-Channel Analysis Distinguishers

Houssem Maghrebi, Olivier Rioul, Sylvain Guilley, and Jean-Luc Danger

Institut Mines-Télécom, Télécom-ParisTech, CNRS LTCI (UMR 5141)
{houssem.maghrebi,olivier.rioul,
sylvain.guilley,jean-luc.danger}@telecom-paristech.fr

Abstract. Side-channel analyses allow to extract keys from devices whatever their length. They rely on tools called "distinguishers". In this paper, we intend to compare two generic distinguishers *per se*: we provide a characterization environment where all the implementation details are equal, hence a fair comparison.

In the field of distinguishers that use a model, the notion of equivalence between distinguishers has already been studied in some seminal works [6,13]. However, no such work has been carried out for generic distinguishers, that work on observable values distributions rather than on their values themselves. In this paper, we set up simulations that aim at showing experimentally that two generic distinguishers are different. Then, we develop a theory to actually prove that one distinguisher is better than the other.

Keywords: Information Theoretic (IT) metrics, Probability/Cumulative Density Function (PDF/CDF), Kolmogorov-Smirnov Analysis (KSA), Inter-class Kolmogorov-Smirnov Analysis (IKSA), Masking.

1 Introduction

Smart cards play a crucial role in many security systems. These devices typically operate in hostile environments and, therefore, the data they contain might be relatively easily compromised. For example, their physical accessibility sometimes allows a number of very powerful attacks against their implementation. During the last decade, side-channel attacks in general, and power analysis attacks in particular, have shaken the belief in the security of smart cards. Kocher *et al.* showed in their pioneering article [10] that a smart card that is unprotected against power analysis attacks can be broken without difficulty. The core idea of side-channel attacks is to compare some key-dependent predictions of the physical leakages with actual measurements, in order to identify which prediction (or key) is the most likely to have given rise to the measurements. In practice, it requires both to be able to model the leakages with a sufficient precision in order to build the predictions, and to have a good comparison tool, thereafter referred to as a *distinguisher*, to efficiently extract the keys.

In 2008, Mutual Information Analysis (MIA) [7] has been proposed as a new side-channel distinguisher. MIA aims at genericity in the sense that it is expected

to lead to successful key recoveries with as little assumptions as possible about the leaking devices it targets. Previous works [14,19,23] demonstrated that the estimation of probability density functions for these key-dependent models is of decisive importance to the performance of MIA in practice.

The authors of [19] suggested an alternative distinguisher that do not require explicit density estimation: the Kolmogorov-Smirnov test. It is a non-parametric statistical test to distinguish between distributions by computing the absolute difference between their cumulative distribution. Reference [24] explores the effectiveness and efficiency of the Kolmogorov-Smirnov Analysis (KSA) in the context of SCA and compare it to the MIA in a number of relevant scenarios ranging from unprotected to masked implementations.

All the distinguishers listed above compare the key-dependent predictions of the physical leakages *versus* actual measurements. Our approach in this paper consists in comparing the conditional leakages between themselves (pairwise) in order to efficiently recover the secret key. We name this approach "*inter-class*". We provide a methodology to fairly compare two SCA distinguishers based on simulations.

The remainder of this article is organized as follows. The definition of the state-of-the-art and inter-class metric is given in section 2. This section contrasts the principle of inter-class metrics with other metrics. In section 3, a fair framework to evaluate and compare distinguishers is given. We applied this methodology to compare the KSA and the inter-class KSA (*aka* IKSA). These theoretical results are then validated by simulations in section 4. Section 5 concludes the paper and gives some perspectives for future works.

Our Contributions
This paper presents three novel contributions. First, we propose the notion of "inter-class" metrics, which allows to build a new distinguisher for SCA aimed to be efficient when exploiting several kinds of leakages. The originality of this new test is that it does a pairwise comparison between the key-dependent leakage classes. Second, we apply this notion to the Kolmogorov-Smirnov test which yield the Inter-class Kolmogorov-Smirnov Analysis (IKSA). In order to compare two SCA distinguishers, we propose a simulation-based "fair" framework which takes into account the different errors of estimation tools used in simulation process. Third, we present several experiments to compare IKSA to KSA using this framework, where simulated attacks are performed against unprotected and protected AES with Boolean masking. Attacks' simulation in section 4 confirm that the IKSA compares favorably to KSA and that IKSA is non-equivalent to KSA, even when masking is applied to ensure some protection.

2 Mutual and Inter-class Distinguishers

2.1 Notations

We use capital letters, like Z, to denote a random variable (RV), calligraphic letters, like $\mathcal{Z}$, to denote its support (set of possible values), and lowercase letters,

like z, for its realizations. The expectation of Z is denoted by $\mathbb{E}[Z]$. The Hamming weight of z is written as $HW(z)$. We use the following notations.

- X: a RV that represents the leakage (*e.g.* the measured current drawn by a cryptographic device);
- K: the cryptographic key;
- Z: the input or the output of the cryptographic device (*i.e.* its plaintext or ciphertext);
- $Y = \psi(Z, K)$: a sensitive variable used internally, that depends both on Z (known by the attacker) and K (unknown by the attacker). We assume that this sensitive variable Y can be computed exhaustively from K by the attacker and that it causes the leakages; put differently, when the key guess is correct, X and Y are dependent.

Side-channel analysis consists in estimating whether X and Y are dependent for every key guess, *i.e.*, for every value $K = k$. The analysis is said to be *sound* if the greatest dependence is obtained for the correct value of the key, noted $k^\star$. In this case, the key can be extracted successfully from the device. In practice, the values taken by X are noisy, because they consist in physical measurements and because the link between X and Y is imperfect (it might involve other variables, yielding algorithmic noise). Therefore, many couples (X, Y) are required for the 2^n estimations (for each value of K) to find the correct key, where n is the bit-width of K.

2.2 Inter-class Notion

Distinguishers can be defined based on the analysis of values or of distributions.

- Examples of distinguishers based on values: DPA [10], CPA [5], stochastic [16], DCA [1].
- Distinguishers based on distributions: MIA [2], KSA [24], *etc.*

The distinguishers based on values can be considered weaker than those based on distributions. A justification is that there exist some distributions (*e.g.* the log-normal distribution) that are not uniquely determined by their moments. Distinguishers based on distributions are referred to as *information-theoretic* and have been acknowledged as more generic.

Several "distances" $D(\cdot\,;\cdot)$ are known to measure the dependency between two distributions, such as Kullback-Leibler (KL) divergence, Hellinger distance, or Kolmogorov-Smirnov (KS) distance. In the sequel, we focus on KS test, because it has been investigated recently and constitutes an interesting competitor to the (already much discussed) mutual information based attacks.

The distance between distributions $D(\cdot\,;\cdot)$ is used to build distinguishers in two different ways:

1. (marginal-to-conditional approach) $D(X|Y;X)$, which yields the KSA distinguisher,
2. (inter-class approach) $D(X|Y;X|Y')$, where Y' is an independent copy of Y, which yields its inter-class counterpart, called IKSA.

3 Comparison Methodology

3.1 Frameworks

In this section, we analyze previous comparison frameworks, highlight possible limitations and motivate for a new setting. The first proposed evaluation framework is [17] basically suggests to use a leakage metric to quantify the maximal chance that an optimal attacker would have to extract secrets. This metric represents a vulnerability analysis, for an attacker might not be able to turn the leakage into a successful attack. For the comparison of attacks, *i.e.* of distinguishers, [17] suggests metrics like o-th order success rate (with $o \in [\![1, 2^n]\!]$) or guessing entropy. In another framework [23,24], the distance to the nearest rival is employed; it is the same definition as previously termed "Correlation Contrast" in [3]. Many other metrics can be invented, such as the signal (distinguisher expected value for the correct key $k^\star$) to noise (distinguisher variance over incorrect keys $k \in \mathbb{F}_2^n \backslash \{k^\star\}$) ratio [8] or the norm-2 of the characterized coefficients in a stochastic profiling [9].

Recent analyses [23] suggest pitfalls in the evaluation methodologies for distinguishers. Errors can arise from many sources:

- **Estimation Bias:** the estimator does not converge to the correct value. For instance, the MIA with few bins for the PDF estimation can have a square bias significantly larger than its variance.
- **Estimation Algorithm:** it can approximate the data. Whatever the kernels used in PDF constructions [14], the binning of the observed side-channel reduces its accuracy.
- **Success Rate Error:** it is a random variable, that has its own variance.
- **Sampling Errors:** the random variables are not drawn a sufficient number of times and thus do not obey to their law. As a rule of thumb, estimations are incorrect if a discrete RV has been measured a fewer number of times than the size of its set of possible values.

In the sequel, we intend to compare KSA [24] and IKSA on a fair basis.

3.2 The Kolmogorov-Smirnov as SCA Distinguisher

In a first stage of the SCA attack, an adversary has to estimate the leakage probability density functions (PDFs) for different key-dependent models. In a second stage, this adversary has to test the dependence of these models with actual measurements. The problem of modeling a PDF from random samples of a distribution is a well studied problem in statistics, referred to as PDF estimation. A number of solutions exist, ranging from simple histograms to kernel density estimation [7,14] or data clustering [20].

Interestingly, an explicit PDF estimation is not always necessary and there also exist statistical tools to compare two PDFs directly from their samples. The Kolmogorov-Smirnov (KS) test is typical of such non-parametric distinguishers.

In the context of SCA, the KSA test has been mentioned first in [19] as a non-parametric statistical test to distinguish between distributions. Then, [24] explores the effectiveness and efficiency of the Kolmogorov-Smirnov Analysis (KSA) and compare it with the Mutual Information Analysis (MIA). It is mainly used as a one-sample test where it allows the comparison of the frequency distribution of a sample to some known distribution, such as a Gaussian distribution, it can also be used as a two-sample test. As a two-sample test KSA distance compares the distributions of values in the two data vectors X_1 and X_2 of length n_1 and n_2, respectively. The null hypothesis for this test is that X_1 and X_2 have the same distribution. The alternative hypothesis is that they have different distributions. The KSA distance is a simple measure which is defined as the maximum value of the absolute difference between two cumulative distribution functions (CDFs): $D_{\mathrm{KSA}} = \sup_{x \in \mathcal{X}} |F_{X_1}(x) - F_{X_2}(x)|$, where F_{X_1} and F_{X_2} are the empirical CDFs (*aka* ECDFs). By definition a (univariate) ECDF is a step function. It is the proportion of observed values of a RV, that are less than or equal to some value. We can write it as: $F_X(x) = \frac{1}{N} \sum_{i=1}^{N} I_{x_i \leq x}$. In this formula, the tuple $\{x_i\}_{i \in [\![1,N]\!]}$ denotes the values realized by the RV X. The function I is an indicator, which is equal to one when the enclosed expression is true, and zero otherwise. Like MIA, the KSA distinguisher measures the maximum distance between the leakage (measurements) X and the hypothesis-dependent conditional observations $X \mid Y$:

$$D_{\mathrm{KS}} = \mathbb{E}_Y \sup_{x \in \mathcal{X}} \left| F_X(x) - F_{X|Y}(x) \right| \quad . \tag{1}$$

The KSA returns the largest difference when the key is correct, *i.e.* when $k = k^\star$.

In contrast to KSA, IKSA consists in comparing the conditional leakages between themselves, pairwise. The Inter-class KSA distinguisher can write as:

$$D_{\mathrm{IKSA}} = \frac{1}{2} \cdot \mathbb{E}_{Y,Y'} \sup_{x \in \mathcal{X}} \left| F_{X|Y}(x) - F_{X|Y'}(x) \right| \quad , \tag{2}$$

where Y' is an independent copy of Y. The $1/2$ factor makes up for double counts $((Y, Y') \leftrightarrow (Y', Y))$.

3.3 Increasing the Fairness of the Estimations

We try here to eliminate or at least bound the errors listed in Sec. 3.1.

- The KS distance is shown to be unbiased by the Glivenko-Cantelli theorem [22], (and furthermore there is a uniform convergence). This is never true for entropy estimators (for instance, all the estimation methods presented in [14] are biased).
- We use an estimation algorithm that keeps the data unchanged (see Eqn. (1) and (2)); Our estimation for KSA is the same as that of Whitnall, Oswald and Mather [24].
- We quantity the success rate error. An upper bound of the variance of the success rate error is shown below to behave as $1/\sqrt{N}$, where N is the number of experiments (also called "number of queries" in [17]).

– We consider attacks with a noise large enough for the success rate to be well below 100% for a number of queries smaller than the size of its definition set.[1]

3.4 Bounding the Success Rate

Let S_i denote i.i.d. Bernoulli variables that take binary values in $\{0, 1\}$ with probabilities p and $1 - p$, where p is the success probability. The success rate is defined as $SR = \frac{1}{N} \sum_{i=1}^{N} S_i$ and has expectation $\mathbb{E}[SR] = p$, *i.e.*, SR is an unbiased estimator of the success probability. According to the strong law of large numbers the success rate converges to p almost surely: $SR \xrightarrow{a.s.} p$. In addition, $\mathbb{E}[SR] = p$, *i.e.* SR is an unbiased estimator of the success rate. Now, the standard deviation of SR is easily computed:

$$\sigma(SR) = \sqrt{\frac{1}{N^2} \cdot N \cdot \sigma^2(S_j)} = \sqrt{\frac{p \cdot (1-p)}{N}} \quad . \tag{3}$$

Thus, the estimation error on the success rate is maximized when p is close to $1/2$, and is minimized when p is almost equal to 0 or 1.

In practice, one wishes to compare the success rates of two distinguishers by examining the values of intermediate p (*i.e.* $p \approx 1/2$). Note that there is a uniform majoration $\sigma(SR) \leqslant \frac{1}{2\sqrt{N}}$, but the error bars can be a function of p and N. The criterion for analyzing experiments will be that errors bars never overlap. Otherwise (see Fig. 1 for $N = 10$), more experiments must be done, so as to reach a situation such as Fig. 1 for $N = 200$. The exact number of experiments depends on the distinguishers to be relatively characterized. The closer they are in success rate, the more experiments are required.

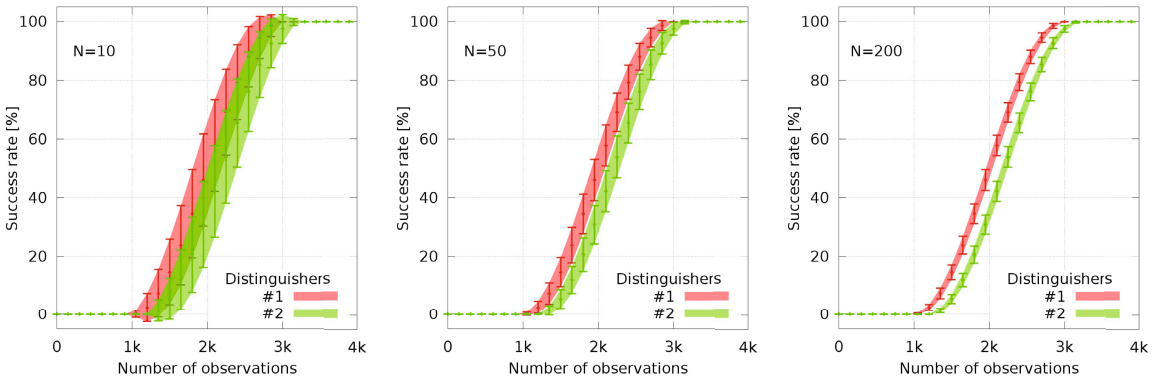

Fig. 1. Examples of success rates errors (Eqn. (3)) for various numbers of experiments

[1] For instance, it can be seen in Fig. 2 that for the unprotected (resp. Boolean masked) AES, the number of traces to recover the key successfully with probability $> 80\%$ is about $2,000$ (resp. $70,000$), which is significantly greater than the number of possible plaintexts (*i.e.* $2^n = 256$) for $\sigma \geqslant 8$.

4 Simulation Results

In this section, we perform several attack experiments to compare KSA and IKSA. Our methodology allows to observe how the different attacks behave against unprotected reference and a masking scheme, and to compare their resistance for different noise's standard deviations.

In what follows, we consider a model in which the leakage variable X is expressed as a deterministic leakage function ϕ of the intermediate variable Y with an independent additive noise N.

Target Leakage: We list hereafter the leakages we consider and the underlying leaking variables:

- 1^{st}-order leakage of an unprotected implementation: $X = \phi(Y) + N$;
- 2^{nd}-order leakage of 1^{st}-order Boolean masking scheme [18]: $X = \phi(Y \oplus M) + \phi(M) + N$, where the mask M is a uniformly distributed RV.

The leakage measurements have been simulated as samples of the random variables X with $\phi = HW$[2] and assuming an additive white Gaussian noise $N \sim \mathcal{N}(0, \sigma^2)$. For both attacks, the sensitive variable Y was chosen to be an AES S-box output of the form $S(Z \oplus k^\star)$, where $S : \mathbb{F}_2^8 \to \mathbb{F}_2^8$ is `SubBytes`, Z is uniformly distributed over $\mathbb{F}_2^8$, and represents a varying plaintext byte and $k^\star \in \mathbb{F}_2^8$ represents the key byte to recover.

Side-Channel Distinguishers: We apply KSA and IKSA such as described in previous sections. The guess key k is tested by estimating $D_{\mathrm{KSA}}(X; \hat{\phi}(Y(k)))$ and $D_{\mathrm{IKSA}}(X; \hat{\phi}(Y(k)))$, respectively, where $\hat{\phi}$ is the prediction function. We select the Hamming weight function as prediction function in our simulations.

Attack Simulation Results: For each investigated context, we compute the first-order success rate of the attacks, over a set of 200 independent experiments for several noise standard deviation values. For comparison purposes, we compute the same metric for other univariate distinguishers: MIA, DPA [4], CPA, VPA [11] and 2O-CPA [21]. Figure 2 summarizes the number of leakage measurements required to observe a success rate of 90% in retrieving $k^\star$ for those SCA attacks. This figure is the compilation of success rates curves obtained for different values of the noise standard deviation (see examples in Fig. 3).

The results presented in Fig. 2 show the significant gain of number of measurements needed induced by IKSA compared to KSA attack. Our new distinguisher compares favorably to KSA: the IKSA attack outperforms the KSA attack when targeting the unprotected implementation or even when the Boolean masking scheme is used for the protection. As expected, CPA performs well in both scenarios since the dependency between the leakage and the model is linear. But,

[2] Assuming Hamming weight leakage model is realistic for implementations on simple microcontrollers [12].

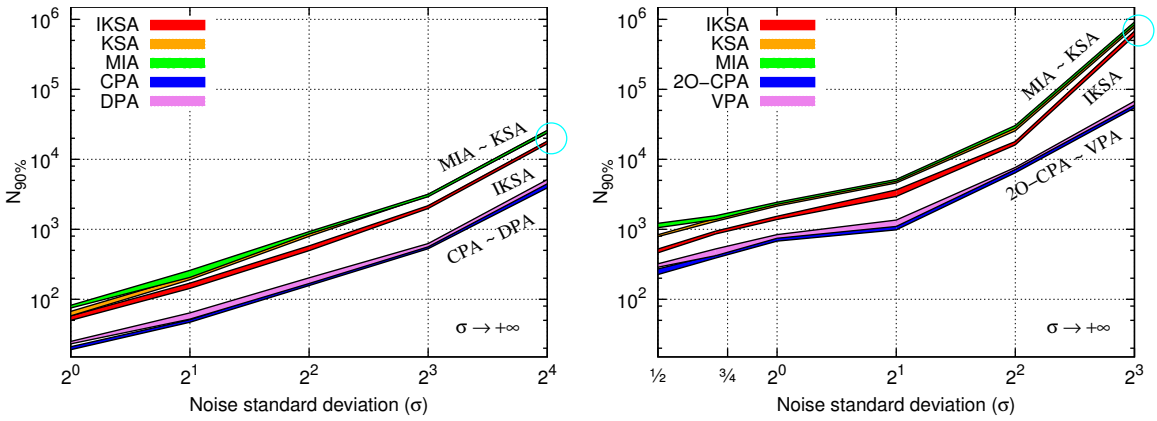

Fig. 2. Evaluation of $N_{90\%}$, the number of messages to achieve a success rate greater than 90%, according to the noise standard deviation when attacking unprotected (*left*) and Boolean masking (*right*) AES implementation

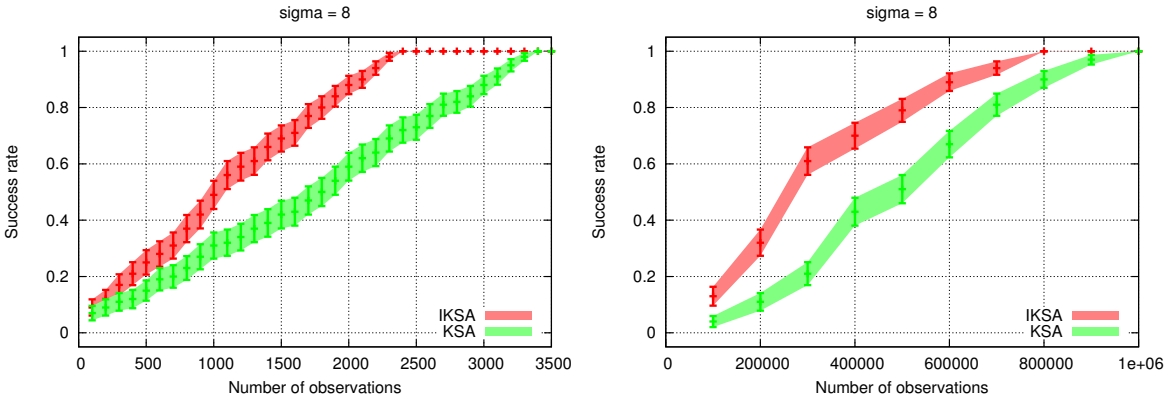

Fig. 3. Success rate of both IKSA and KSA distinguishers when attacking one substitution box of an unprotected AES (*left*) and of a Boolean masking scheme (*right*)

we like to stress that we focus in this paper only on information-theoretic distinguishers which are generic.

In [13], a notion of asymptotic equivalence (noted "∼") for side-channel distinguishers is introduced: two distinguishers are said *equivalent* if the number of traces to overcome a given success rate (say 90%) decreases when the noise variance increases. For example, the likelihood and the Pearson correlation are equivalent in this sense. A look at $N90\%$ curves in Fig. 2 shows that other univariate distinguishers exhibit a similar equivalence law:

- DPA ∼ CPA on an unprotected implementation (*left*);
- 2O-CPA ∼ VPA on a first-order masked implementation (*right*);
- KSA ∼ MIA on both implementations (already proved in [24]).

However, *IKSA and KSA are not equivalent*. The difference between IKSA and KSA ∼ MIA is materialized in Fig. 2 as a circle in cyan color. To the best of our knowledge, it is the first time that two distinguishers that do not become

equivalent in the sense of [13] are put forward. Incidentally, we note that this conclusion could not have been derived mathematically under the usual Gaussian approximation, because under this approximation equivalence holds as $\sigma \to +\infty$. This tends to show that the mutual and inter-class approaches are of a different kind, even in a mono-variate context.

5 Conclusions and Perspectives

In this paper, we have introduced the new "inter-class" concept to distinguish between various partitionings. We applied this concept to the Kolmogorov-Smirnov distance, resulting in IKSA. We also proposed a simulation-based fair framework to compare the two distinguishers KSA and IKSA. Our framework takes in account the different sources of errors estimations. We used this framework to compare KSA to IKSA using the success rate metric. Security metrics are clearly in favor of IKSA even when the implementation is unprotected or protected using a first-order Boolean masking countermeasure (with a linear leakage model).

An interesting question for the future work is to give a theoretical proof of the soundness of the distinguishers. Also, we endeavour to find a mathematical explanation why IKSA outperforms KSA for usual leakage functions.

References

1. Batina, L., Gierlichs, B., Lemke-Rust, K.: Differential Cluster Analysis. In: Clavier, C., Gaj, K. (eds.) CHES 2009. LNCS, vol. 5747, pp. 112–127. Springer, Heidelberg (2009)
2. Batina, L., Gierlichs, B., Prouff, E., Rivain, M., Standaert, F.-X., Veyrat-Charvillon, N.: Mutual Information Analysis: a Comprehensive Study. J. Cryptology 24(2), 269–291 (2011)
3. Benoît, O., Peyrin, T.: Side-Channel Analysis of Six SHA-3 Candidates. In: Mangard, S., Standaert, F.-X. (eds.) CHES 2010. LNCS, vol. 6225, pp. 140–157. Springer, Heidelberg (2010)
4. Bévan, R., Knudsen, E.W.: Ways to Enhance Differential Power Analysis. In: Lee, P.J., Lim, C.H. (eds.) ICISC 2002. LNCS, vol. 2587, pp. 327–342. Springer, Heidelberg (2003)
5. Brier, É., Clavier, C., Olivier, F.: Correlation Power Analysis with a Leakage Model. In: Joye, M., Quisquater, J.-J. (eds.) CHES 2004. LNCS, vol. 3156, pp. 16–29. Springer, Heidelberg (2004)
6. Doget, J., Prouff, E., Rivain, M., Standaert, F.-X.: Univariate side channel attacks and leakage modeling. J. Cryptographic Engineering 1(2), 123–144 (2011)
7. Gierlichs, B., Batina, L., Tuyls, P., Preneel, B.: Mutual Information Analysis. In: Oswald, E., Rohatgi, P. (eds.) CHES 2008. LNCS, vol. 5154, pp. 426–442. Springer, Heidelberg (2008)
8. Guilley, S., Hoogvorst, P., Pacalet, R.: Differential Power Analysis Model and some Results. In: Kluwer (ed.) Proceedings of WCC/CARDIS 2004, Toulouse, France, pp. 127–142 (August 2004), doi:10.1007/1-4020-8147-2_9
9. Heuser, A., Schindler, W., Stöttinger, M.: Revealing side-channel issues of complex circuits by enhanced leakage models. In: Rosenstiel, W., Thiele, L. (eds.) DATE, pp. 1179–1184. IEEE (2012)

10. Kocher, P.C., Jaffe, J., Jun, B.: Differential Power Analysis. In: Wiener, M. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 388–397. Springer, Heidelberg (1999)
11. Maghrebi, H., Danger, J.-L., Flament, F., Guilley, S.: Evaluation of Countermeasures Implementation Based on Boolean Masking to Thwart First and Second Order Side-Channel Attacks. In: SCS, Jerba, Tunisia, November 6-8, pp. 1–6. IEEE (2009), doi:10.1109/ICSCS.2009.5412597
12. Mangard, S., Oswald, E., Popp, T.: Power Analysis Attacks: Revealing the Secrets of Smart Cards. Springer (December 2006) ISBN 0-387-30857-1, http://www.springer.com/, http://www.dpabook.org/
13. Mangard, S., Oswald, E., Standaert, F.-X.: One for All - All for One: Unifying Standard DPA Attacks. Information Security, IET 5(2), 100–111 (2010) ISSN: 1751-8709; Digital Object Identifier: 10.1049/iet-ifs.2010.0096
14. Prouff, E., Rivain, M.: Theoretical and Practical Aspects of Mutual Information Based Side Channel Analysis. In: Abdalla, M., Pointcheval, D., Fouque, P.-A., Vergnaud, D. (eds.) ACNS 2009. LNCS, vol. 5536, pp. 499–518. Springer, Heidelberg (2009)
15. Rogaway, P. (ed.): CRYPTO 2011. LNCS, vol. 6841, pp. 2011–2031. Springer, Heidelberg (2011)
16. Schindler, W., Lemke, K., Paar, C.: A Stochastic Model for Differential Side Channel Cryptanalysis. In: Rao, J.R., Sunar, B. (eds.) CHES 2005. LNCS, vol. 3659, pp. 30–46. Springer, Heidelberg (2005)
17. Standaert, F.-X., Malkin, T.G., Yung, M.: A Unified Framework for the Analysis of Side-Channel Key Recovery Attacks. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 443–461. Springer, Heidelberg (2009)
18. Standaert, F.-X., Rouvroy, G., Quisquater, J.-J.: FPGA Implementations of the DES and Triple-DES Masked Against Power Analysis Attacks. In: FPL. IEEE, Madrid, Spain (August 2006)
19. Veyrat-Charvillon, N., Standaert, F.-X.: Mutual Information Analysis: How, When and Why? In: Clavier, C., Gaj, K. (eds.) CHES 2009. LNCS, vol. 5747, pp. 429–443. Springer, Heidelberg (2009)
20. Veyrat-Charvillon, N., Standaert, F.-X.: Generic Side-Channel Distinguishers: Improvements and Limitations. In: Rogaway [15], pp. 354–372
21. Waddle, J., Wagner, D.: Towards Efficient Second-Order Power Analysis. In: Joye, M., Quisquater, J.-J. (eds.) CHES 2004. LNCS, vol. 3156, pp. 1–15. Springer, Heidelberg (2004)
22. Wellner, J.A.: A Glivenko-Cantelli theorem and strong laws of large numbers for functions of order statistics. Ann. Statist. 5(3), 473–480 (1977)
23. Whitnall, C., Oswald, E.: A Comprehensive Evaluation of Mutual Information Analysis Using a Fair Evaluation Framework. In: Rogaway [15], pp. 316–334
24. Whitnall, C., Oswald, E., Mather, L.: An Exploration of the Kolmogorov-Smirnov Test as a Competitor to Mutual Information Analysis. In: Prouff, E. (ed.) CARDIS 2011. LNCS, vol. 7079, pp. 234–251. Springer, Heidelberg (2011)