Journal of Cryptographic Engineering (2020) 10:269–274
https://doi.org/10.1007/s13389-020-00235-6

REGULAR PAPER

A remark on a success rate model for side-channel attack analysis

Andreas Wiemers[1]

Received: 30 January 2019 / Accepted: 19 June 2020 / Published online: 18 July 2020
© The Author(s) 2020

Abstract
The success rate is the most common evaluation metric for measuring the performance of a particular side-channel attack scenario. We improve on an analytic formula for the success rate.

Keywords Side-channel attacks · evaluation metric · success rate

1 Introduction

In [1], a general statistical model for side-channel attack analysis is proposed. Based on this model, one can calculate a success rate of an attack by numerical simulation. This success rate is the most common evaluation metric for measuring the performance of a particular attack scenario. In [5], it is stated: "Closed-form expressions of success rate are desirable because they provide an explicit functional dependence on relevant parameters such as number of measurements and signal-to-noise ratio which help to understand the effectiveness of a given attack and how one can mitigate its threat by countermeasures. However, such closed-form expressions involve high-dimensional complex statistical functions that are hard to estimate". In the following, we will derive an analytic formula for the success rate. Simulation experiments confirm that this analytic formula is a good approximation for the success rate for a wide class of leakage functions.

2 Leakage model

We consider the case of a side-channel attack against a typical block cipher. We assume that this block cipher consists of several rounds for encryption and decryption. In each round, the block cipher uses computations of substitution boxes of small size n (e.g., 6 bits for DES or n bits for AES), where the key is mixed with intermediate values.
We further restrict ourselves to the simplest setting:

- The attacker tries to find an n-bit subkey k_c of the S-Box computation in the first round of the block cipher. The input of this S-Box computation is of the form $p_w \oplus k_c$ with plaintext inputs p_w.
- We have m measurements. m is a multiple of $N = 2^n$, and all plaintext inputs p_w of this S-Box are equally distributed over these m measurements.
- The side-channel measurement is a trace of a certain number of points. We assume that the key-dependent leakage occurs in just one point of time which is known to the attacker.
- The measurement in this point of time is the sum of a deterministic signal and Gaussian noise. It can be written in the form

  $$\tilde{b}_w = \tilde{h}(p_w \oplus k_c) + \tilde{\tau}_w.$$

  $\tilde{h}$ is a deterministic function that only depends on the input $p_w \oplus k_c$ of the S-Box computation. $\tilde{h}$ is completely known to the attacker. $\tilde{\tau}_w$ describes the noise of the measurement. We assume that $\tilde{\tau}_w$ are realizations of m independent random variables $\tilde{T}_w$; each one is normally distributed with known expectation and variance. For ease of notation, we associate the sets $\{0, 1\}^n$ and $\{0, 1, \ldots, N-1\}$ by the 2-adic representation of an integer. We further assume

  $$E(\tilde{T}_w) = 0,\ V(\tilde{T}_w) = \sigma^2,$$
  $$\sum_{z=0}^{N-1} \tilde{h}(z) = 0,\ \sum_{z=0}^{N-1} \tilde{h}(z)^2 = N\tilde{\delta}^2.$$

- We can calculate the mean value of all $\tilde{b}_w$ with the same p_w. In the representation of $\tilde{b}_w$, this just reduces the vari-

✉ Andreas Wiemers
andreas.wiemers@bsi.bund.de

1 BSI, Bonn, Germany

ance of $\tilde{T}_w$. Additionally, by applying a constant factor to each $\tilde{b}_w$ we can normalize the representation of $\tilde{b}_w$. To this end, we get a representation in the form

$$b_w = h(w \oplus k_c) + \tau_w, w = 0, \ldots, N-1$$

with

$$E(T_w)=0, V(T_w)=1, \sum_{z=0}^{N-1} h(z) = 0, \sum_{z=0}^{N-1} h(z)^2 = N\delta^2.$$

If we start with the representation of $\tilde{b}_w$, the normalized representation b_w has parameter δ with

$$\delta^2 = \frac{m}{N}\frac{\tilde{\delta}^2}{\sigma^2}.$$

As in [1], we now apply the maximum likelihood attack: We compute the conditional probability density function of the observations b_w under each hypothesis k. We choose as the correct key that k which maximizes the probability density function. An easy calculation shows that we have to compare the values

$$\sum_{w=0}^{N-1} (b_w - h(w \oplus k))^2.$$

This can further be reduced to the values

$$\sum_{w=0}^{N-1} h(w \oplus k) b_w$$

since $\sum_{w=0}^{N-1} h(w \oplus k)^2$ does not depend on k. The success rate as defined in [1] is the probability that

$$\Pr(X_{k_c} > X_k \text{ for all } k \neq k_c)$$

where X_k is the random variable

$$X_k = \sum_{w=0}^{N-1} h(w \oplus k)(h(w \oplus k_c) + T_w).$$

This success rate can certainly be computed by numerical simulation of the T_w.

3 An approximation of the success rate

Let A be the N×N-matrix with entries $h(w \oplus k)$. The rows of A are

$$a_k = (h(k), \ldots, h(w \oplus k), \ldots, h((N-1) \oplus k)).$$

Let T be the random vector (as column) of length N with entries T_w. Let $d = A \cdot a_{k_c}^t$ with entries d_k. We define the set R of all vectors of length N with entries y_k that fulfill

$$y_k < y_{k_c} + N\delta^2 - d_k \text{ for all } k \neq k_c.$$

An easy calculation shows that the success rate can be written as

$$\Pr(X_{k_c} > X_k \text{ for all } k \neq k_c) = \Pr(A \cdot T \in R).$$

A is a symmetric matrix, and therefore there exists an orthonormal basis of eigenvectors $v_0, \ldots, v_{N-1}$ with corresponding eigenvalues $\lambda_0, \ldots, \lambda_{N-1}$ of A. T can be written in the basis of eigenvectors in the form

$$T = X_0 v_0 + \cdots + X_{N-1} v_{N-1}$$

where the X_i are independent random variables with standard normal distribution. The distribution of $A \cdot T$ is the image of the standard normal distribution under A. Each vector in the distribution of T is stretched in the direction of the eigenvectors of A with the corresponding eigenvalue as factor.

$$A \cdot T = \lambda_0 X_0 v_0 + \cdots + \lambda_{N-1} X_{N-1} v_{N-1}.$$

We easily compute

$$E(||A \cdot T||^2) = N^2\delta^2 = N\delta^2 E(||T||^2) = \lambda_0^2 + \cdots + \lambda_{N-1}^2.$$

For values like $n = 6$ or $n = 8$, $N = 2^n$ is a relatively large number, so that the typical vector in the distribution of $A \cdot T$ has square of norm $N^2\delta^2$. As a heuristic approximation for the success rate, we just replace the distribution of $A \cdot T$ by the normal distribution stretched by the constant factor $2^{n/2}\delta$:

1st approx. formula: $\Pr(2^{n/2}\delta \cdot T \in R)$.

In addition, we omit the influence of d and get

2nd approx. formula: $\Pr(T \in \tilde{R})$

where $\tilde{R}$ is the set of all vectors t_k that fulfill

$$t_k < t_{k_c} + 2^{n/2}\delta \text{ for all } k \neq k_c.$$

The last probability can be in fact computed as a two-dimensional integral

$$\Pr(T \in \tilde{R}) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}a^2) \left[\int_{-\infty}^{a+2^{n/2}\delta} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}t^2) \mathrm{d}t \right]^{N-1} da.$$

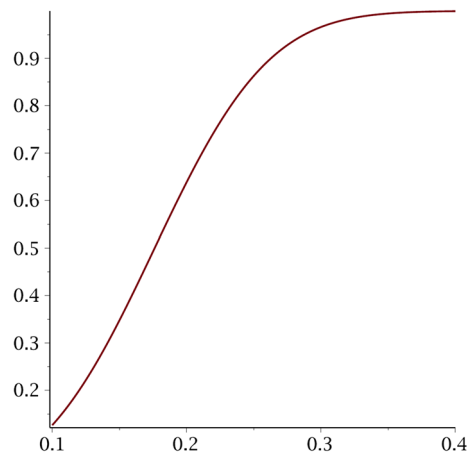

Fig. 1 Second approximating formula. Success rate as function in δ for $n = 8$

This expression only depends on δ, so that it can easily be listed for different δ by numerical methods. Figure 1 plots this approximated success rate as computed by MAPLE software for $n = 8$.

Remarks:

- If we start with the representation of $\tilde{b}_w$, the success rate as computed by the second approximating formula only depends on

$$\delta^2 = \frac{m}{N}\frac{\tilde{\delta}^2}{\sigma^2}.$$

- The approximating formulas are only valid if the eigenvalues do not vary too much. As an extreme example, we can consider the case that only one eigenvalue is large, whereas the others can be neglected. Let $\lambda_0 > 0$ be this large eigenvalue. Then, $A \cdot T$ is roughly distributed as $\lambda_0 X_0 v_0$. $\Pr(A \cdot T \in R)$ can be written as a one-dimensional integral over the random variable X_0.
- In our approach, we replaced the covariance matrix A^2 by a diagonal matrix. In effect, we treated X_k as independent random variables.
- $\Pr(T \in \tilde{R}) \geq \frac{1}{N}$ with equality for $\delta = 0$. The probability of $\frac{1}{N}$ for $\delta = 0$ follows from the symmetry of the set $\tilde{R}$.

4 More on the matrix *A*

The properties of the matrix A are used in the context of dyadic codes; see [2]. In [3], the matrix A is called dyadic matrix. Due to the structure of A, we can compute the eigenvectors of A explicitly: There are N GF(2)-linear functions L

$$L : \mathrm{GF}(2)^n \longrightarrow \mathrm{GF}(2).$$

For every L, $v_L = [(-1)^{L(w)}]_w$ is a vector of length N. For every k, we have

$$\sum_w h(k \oplus w)(-1)^{L(w)} = \sum_y h(y)(-1)^{L(y \oplus k)} = (-1)^{L(k)} \sum_y h(y)(-1)^{L(y)}.$$

Therefore, v_L is an eigenvector with eigenvalue $\sum_y h(y)(-1)^{L(y)}$. The rank of A is the number of nonzero eigenvalues.

5 Example: *h* depends on a single bit

Let S be the S-Box of the AES and G a fixed GF(2)-linear function. We assume that the leakage function h only depends on $G \circ S$, i.e., after normalization

$$h(w \oplus k) = \delta(-1)^{G(S(w \oplus k))}.$$

The eigenvalues of A are now

$$\sum_y h(y)(-1)^{L(y)} = \delta \sum_y (-1)^{G(S(y))}(-1)^{L(y)}.$$

With other words: The set of eigenvalues is exactly the Walsh spectrum of the Boolean function $G \circ S$ multiplied by δ. Each eigenvalue is a measure how good $G \circ S$ can be approximated by a linear function L. S is the composition of the inversion over $F = \mathrm{GF}(256)$ and an affine function. The Walsh spectrum of any function of the form $G \circ S$ is well known: It can be expressed by the so-called Kloosterman sums; see [4].

$$K(a) = \sum_{y \in F^x} (-1)^{tr(y^{-1} + ay)}$$

where $tr(y)$ denotes the trace of y over F. Any GF(2)-linear function $L : F \longrightarrow \mathrm{GF}(2)$ can be written as $L(y) = tr(ly)$ for exactly one $l \in F$. Therefore, we find $c \in F$ such that

$$G(S(y)) \oplus L(y) = tr(cy^{-1} \oplus ly) \text{ for all } y \in F^x$$

or

$$G(S(y)) \oplus L(y) = tr(cy^{-1} \oplus ly) \oplus 1 \text{ for all } y \in F^x.$$

Note that for $c \neq 0$

$$\sum_{y \in F^x} (-1)^{tr(cy^{-1} + ly)} = K(c \cdot l).$$

The distribution of the Kloosterman sums can be described by values of certain class numbers (see [4, Prop. 9.1]), which can be interpreted in terms of the Walsh spectrum.

6 Example: *h* depends on the Hamming weight of the input

In this example, h does not depend on the output of the substitution box, but on the Hamming weight of the input. After normalization, we can write

$$h(w \oplus k) = \delta g(w \oplus k) \text{ with } g(z) = \frac{1}{\sqrt{n}}((-1)^{z_1} + \cdots + (-1)^{z_n}), z = (z_1, \ldots, z_n).$$

In this case, A has exactly n eigenvectors with eigenvalues $\neq 0$ and these are given by the n linear projections

$$v_j = \frac{1}{2^{n/2}}[(-1)^{z_j}]_{z=(z_1,\ldots,z_n)}.$$

The eigenvalues of these n eigenvectors are equal to $\delta \frac{N}{\sqrt{n}}$. Since we have only a few eigenvalues $\neq 0$, we cannot expect that the second approximating formula is a good approximation in this case.
However, we can derive an exact formula for the success rate: Since h is a linear function, we have

$$\sum_{w=0}^{N-1} h(w \oplus k) b_w = \frac{\delta}{\sqrt{n}} \sum_{j=1}^{n} (-1)^{k_j} \left(\sum_{w=0}^{N-1} (-1)^{w_j} b_w. \right)$$

The sums in brackets do not depend on k, so that

$$\max_k \sum_{w=0}^{N-1} h(w \oplus k) b_w = \frac{\delta}{\sqrt{n}} \sum_{j=1}^{n} \mid \sum_{w=0}^{N-1} (-1)^{w_j} b_w \mid .$$

The maximum likelihood attack is therefore successful exactly in the event that

$$(-1)^{k_{c,j}} \left(\sum_{w=0}^{N-1} (-1)^{w_j} b_w \right) \geq 0 \text{ for all } j = 1, \ldots, n.$$

With other words: The success rate is the probability that the random variable Y_j fulfills

$$Y_j = (-1)^{k_{c,j}} \left(\sum_{w=0}^{N-1} (-1)^{w_j} (h(w \oplus k_c) + T_w) \right) \geq 0 \text{ for all } j = 1, \ldots, n.$$

Y_j is normally distributed with an expectation value $\frac{\delta N}{\sqrt{n}}$ and variance N. Since the covariance between Y_j and $Y_{\tilde{j}}$ is 0 for $j \neq \tilde{j}$, the success rate is given by the formula

$$\Pr(Y_j \geq 0,\ j{=}1, \ldots, n){=}\left[\int_{-\frac{\delta 2^{n/2}}{\sqrt{n}}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}t^2) dt \right]^n.$$

Table 1 Comparison of success rates, $n = 8$

	$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$
2nd approx. formula	0.13	0.64	0.97
$h = \delta(-1)^f$	0.12	0.62	0.96
$h = \delta g \circ P$	0.12	0.62	0.96
Hamming weight formula	0.07	0.33	0.69
$h = \delta g$	0.07	0.33	0.69

Table 2 Comparison of success rates, $n = 6$

	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
2nd approx. formula	0.25	0.53	0.79
$h = \delta(-1)^f$	0.24	0.50	0.75
$h = \delta g \circ P$	0.24	0.50	0.75
Hamming weight formula	0.17	0.34	0.55
$h = \delta g$	0.17	0.34	0.55

Table 3 Success rates in the case of masking ($m = 10 \cdot N^2$, input dependency, $n = 8$)

	$\sigma = 37.6$ $\delta = 0.1$	$\sigma = 26.6$ $\delta = 0.2$	$\sigma = 21.6$ $\delta = 0.3$
Simulation	0.07	0.33	0.70
Hamming weight formula	0.07	0.33	0.69

7 Simulation results

We computed the success rate for different n, h and δ by numerical simulation of the T_w. Table 1 compares the success rates for $n = 8$, and Table 2 the same for $n = 6$. In both tables, f is chosen as a random function $\mathrm{GF}(2)^n \longrightarrow \mathrm{GF}(2)$, but uniformly distributed. P is chosen as a random permutation on $\mathrm{GF}(2)^n$. g is the function from paragraph 6. We repeated the simulation 1000 times with different f and P, so that a mean is given in both tables.

We note that the second approximating formula and the Hamming weight formula from paragraph 6 give different values for identical δ, but both formulas match the numerical values very well. In all experiments, the numerical values in each of the 1000 repetitions were very close to the mean given in the tables. For $n = 8$ (Table 1), the empirical standard deviation was less than 0.004. For $n = 6$ (Table 2), the empirical standard deviation was less than 0.02.

8 Success rate in the case of masking

Similar to [5], we can apply the second approximating formula to the case of masking. For a concrete example, we adapt our leakage model in the following way:

Table 4 Success rates in the case of masking, ($m = N^2$, input dependency),$n = 8$

	$\sigma = 21.1$ $\delta = 0.1$	$\sigma = 14.8$ $\delta = 0.2$	$\sigma = 11.9$ $\delta = 0.3$
Simulation	0.07	0.33	0.71
Hamming weight formula	0.07	0.33	0.69

Table 5 Success rates in the case of masking, ($m = 10 \cdot N^2$, output dependency), $n = 8$

	$\sigma = 37.6$ $\delta = 0.1$	$\sigma = 26.6$ $\delta = 0.2$	$\sigma = 21.6$ $\delta = 0.3$
Simulation	0.13	0.62	0.96
2nd approx. formula	0.13	0.64	0.97

Table 6 Success rates in the case of masking, ($m = N^2$, output dependency),$n = 8$

	$\sigma = 21.1$ $\delta = 0.1$	$\sigma = 14.8$ $\delta = 0.2$	$\sigma = 11.9$ $\delta = 0.3$
Simulation	0.12	0.62	0.96
2nd approx. formula	0.13	0.64	0.97

- We have m measurements. m is a multiple of N, and all plaintext inputs p_w of this S-Box are equally distributed over these m measurements.
- There are exactly two points of time when meaningful leakages occur. Both points of time are known to the attacker. One leakage is mask-dependent; the other one is key-dependent, but on the input of an S-Box computation.
- The measurements can be written in the form

$$\tilde{b}'_w = \mu(p_w \oplus k_c \oplus m_w) + \tilde{\tau}'_w$$
$$\tilde{b}''_w = \mu(m_w) + \tilde{\tau}''_w.$$

μ is a centralized form of the Hamming weight, i.e.,

$$\mu(z) = (-1)^{z_1} + \cdots + (-1)^{z_n}.$$

$\tilde{\tau}'_w$ and $\tilde{\tau}''_w$ describe the noise of the measurement. We assume that $\tilde{\tau}'_w$ and $\tilde{\tau}''_w$ are realizations of $2m$ independent random variables $\tilde{T}'_w$, $\tilde{T}''_w$; each one is normally distributed with expectation 0 and variance σ^2. m_w describes the mask. m_w are the realizations of m independent uniformly distributed random variables M_w on GF(N).

We set

$$c_\nu = \frac{N}{m} \sum_{w, p_w = \nu} \tilde{b}'_w \tilde{b}''_w.$$

The sum is taken over $\frac{m}{N}$ realizations of independent random variable. For any fixed mask m_w, we compute

$$E((\mu(p_w \oplus k_c \oplus m_w) + \tilde{T}'_w)(\mu(m_w) + \tilde{T}''_w)) = \mu(p_w \oplus k_c \oplus m_w)\mu(m_w)$$

and

$$\begin{aligned} &V((\mu(p_w \oplus k_c \oplus m_w) + \tilde{T}'_w)(\mu(m_w) + \tilde{T}''_w)) \\ &= E((\mu(p_w \oplus k_c \oplus m_w) + \tilde{T}'_w)^2(\mu(m_w) + \tilde{T}''_w)^2) \\ &\quad -\mu(p_w \oplus k_c \oplus m_w)^2\mu(m_w)^2 \\ &= \sigma^2(\mu(p_w \oplus k_c \oplus m_w)^2 + \mu(m_w)^2) + \sigma^4. \end{aligned}$$

If $\frac{m}{N}$ is not too small, we approximate c_ν as realizations of N independent normally distributed random variables, each with expectation

$$\begin{aligned} &\frac{N}{m} \sum_{w, p_w = \nu} \mu(p_w \oplus k_c \oplus m_w)\mu(m_w) \\ &= \frac{N}{m} \sum_{w, p_w = \nu} \mu(\nu \oplus k_c \oplus m_w)\mu(m_w) \end{aligned}$$

and variance

$$\left(\frac{N}{m}\right)^2 \sum_{w, p_w = \nu} \left(\sigma^2(\mu(\nu \oplus k_c \oplus m_w)^2 + \mu(m_w)^2) + \sigma^4\right).$$

Again if $\frac{m}{N}$ is not too small, we approximate these sums by the expectation over the random variables M_w. An easy calculation shows

$$\frac{N}{m} \sum_{w, p_w = \nu} \mu(p_w \oplus k_c \oplus m_w)\mu(m_w) \approx \mu(\nu \oplus k_c)$$

and

$$\begin{aligned} &\left(\frac{N}{m}\right)^2 \sum_{w, p_w = \nu} \left(\sigma^2(\mu(\nu \oplus k_c \oplus m_w)^2 + \mu(m_w)^2) + \sigma^4\right) \\ &\approx \frac{N}{m}(2n\sigma^2 + \sigma^4). \end{aligned}$$

Since $\sum_z \mu(z)^2 = n \cdot N$, we can apply the leakage model of paragraph 2 with

$$\delta^2 = \frac{nm}{N(2n\sigma^2 + \sigma^4)}.$$

Given the measurements $\tilde{b}'_w, \tilde{b}''_w$, we directly compare the values

$$\sum_\nu \mu(\nu \oplus k)c_\nu$$

for different k and decide for the k with the largest value. For large m, we can expect that the success rate of this ad hoc attack only depends on $\delta^2 = \frac{nm}{N(2n\sigma^2+\sigma^4)}$.

Table 3 gives the success rates of this attack computed by numerical simulation and $n = 8$. We compare these success rates with the values for the example from paragraph 6 ($h = \delta g$). Since the numerical simulations are rather slow, we repeated the simulation only for a few instances. However, in all instances the values matched very well.

Table 4 gives similar data, but for $m = N^2$.

Remark:

The leakage in $\tilde{b}'_w$ depends on the input of an S-Box computation. We can certainly consider the case that the leakage depends on the output of an S-Box computation, i.e.,

$$\tilde{b}'_w = \mu(S(p_w \oplus k_c) \oplus m_w) + \tilde{\tau}'_w.$$

The computation is completely analog, but we expect that the second approximating formula applies. Tables 4 and 5 compare the numerical values for the success rate with the second approximating formula. Again, we computed only a few instances, but in all instances the values matched very well (Table 6)

Acknowledgements Open Access funding provided by Projekt DEAL. Open Access funding provided by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Fei, Y., Ding, A.A., Lao, J., Zhang L.: A statistics-based success rate model for DPA and CPA. J. Cryptogr. Eng. (2015)
2. Rajan, S.: Moon Ho Lee: Quasi-Cyclic dyadic codes in the Walsh-Hadamard transform domain. IEEE Trans. Inf. Theory **48**(8), 2406–2412 (2002)
3. Misoczki, R., Barreto, P.S.L.M.: Compact McEliece Keys from Goppa Codes. In: Jacobson, M.J., Rijmen, V., Safavi-Naini, R. (eds.) Selected Areas in Cryptography. SAC, : Lecture Notes in Computer Science, vol. 5867. Springer, Berlin (2009)
4. Lachaud, W.: Weights of the orthogonals of the extended quadratic binary Goppa codes, IEEE Trans. Inf. Theory (1990). https://doi.org/10.1109/18.54892
5. Guilley, S., Heuser, A., Rioul, O.: A Key to Success–Success Exponents for Side-Channel Distinguishers, Cryptology ePrint Archive: Report 2016/987

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.