A fair experimental evaluation of distance correlation side-channel distinguisher

Petr Socha, Vojtěch Miškovský, Martin Novotný
Czech Technical University in Prague
Faculty of Information Technology
Czech Republic
{petr.socha,vojtech.miskovsky,martin.novotny}@fit.cvut.cz

***Abstract*—Side-channel attacks pose a severe threat to cryptographic implementations, allowing the attacker to recover secret information based on physical observations of the cryptographic device. Correlation Power Analysis is considered to be one of the most powerful attacks in the non-profiled scenario. In this paper, we consider the distance/Brownian correlation instead of the traditionally used Pearson coefficient. We give a fair comparison of our novel approach attacking AES on three different FPGA platforms and we discuss the distance correlation potential in the context of side-channel analysis.**

***Index Terms*—Side-Channel Analysis, Embedded Security, Internet of Things, Correlation Power Analysis, Non-linear Correlation**

I. Introduction

In today's IoT and Industry 4.0 era, embedded systems are becoming a natural part of our living environment. To ensure security and privacy, various authentication, authorization, and encryption schemes and the underlying cryptographic primitives must be implemented. While these are considered secure under the cryptanalyst's black-box model, they may be vulnerable to physical attacks such as side-channel power analysis when implemented improperly [1], [2].

These attacks exploit the fact that the instantaneous power consumption of the cryptographic device is data-dependent. In the non-profiled side-channel attack scenario, the first phase of the attack is typically sampling the power consumption during multiple encryptions, while capturing the input or output data. In the second phase, the attacker forms consumption hypotheses based on the captured data, for every key value (attacking a part of the key at a time, e.g., in the case of AES, a byte), and statistically selects – distinguishes – the most probable hypothesis based on the real sampled consumption.

Different side-channel distinguishers were proposed over time. The original Differential Power Analysis [1] proposes partitioning the power traces according to a single bit in a predicted key-dependent working variable, for every key hypothesis. Assuming the correct hypothesis, the two partitions should be distinguishable, which is done by searching for the greatest point-wise difference of means. Correlation Power Analysis [3], [2] is a similar attack based on searching for a significant point-wise Pearson correlation of the real power consumption and the predicted consumption, using, e.g., Hamming weight/distance power leakage model. Mutual Information Analysis [4], [5] is a more generic approach working with only a few leakage assumptions, based on approximating the mutual information between the hypothetical and the real power consumption. Kolmogorov-Smirnov Analysis [5], [6] is another example of a generic side-channel distinguisher. In order to reduce the leakage assumptions in correlation-based attacks, rank correlation, namely Spearman correlation, is evaluated in [7]. Recently, a non-profiled side-channel distinguisher based on deep learning was proposed [8]. The usage of the non-parametric distance-based statistics, such as distance/Brownian correlation [9], [10], in the side-channel analysis context, was suggested as an alternative in [11], and it was recently used for attacking a digital multiplier [12].

In this paper, we first briefly describe the Correlation Power Analysis attack. Next, we present the non-parametric multivariate distance correlation coefficient, and we propose a multivariate clock-wise side-channel distinguisher. We evaluate the distinguisher by attacking AES [13] encryption implementations on three different FPGA platforms, using success rate and guessing entropy metrics [14].

II. Distance Correlation Power Analysis

In this section, we first describe the Correlation Power Analysis attack using the Pearson correlation coefficient. Then we describe the distance correlation. Finally, we propose a novel multivariate clock cycle-wise side-channel distinguisher.

A. Correlation Power Analysis

In this subsection, we describe the Correlation Power Analysis [2] attack. First, let us define the hat symbol $\hat{N}$ for any variable $N \in \mathbb{N}$ as $\hat{N} = \{1, 2, ..., N\}$ for simplicity.

In the first phase of the attack, N power traces are measured, e.g., using an oscilloscope, during N encryptions of uniform random plaintexts. Each power trace consists of power consumption sampled in S points in time. The N power consumption samples at a given time s can be considered N samples from a random variable $P_s(n) \in \mathbb{R}$, where $s \in \hat{S}, n \in \hat{N}$.

In the second phase, the attacker predicts power consumption for each of the N encryptions, and for each of the K key hypotheses (for AES key byte, $K = 256$). The N hypothetical power predictions for a given hypothetical key k can be considered N samples from a random variable $H_k(n) \in \mathbb{R}$,

This research has been supported by the grant VJ02010010 of the Ministry of the Interior of the Czech Republic, "Tools for AI-enhanced Security Verification of Cryptographic Devices" in the program Impakt1 (2022-2025).

where $k \in \hat{K}, n \in \hat{N}$. The random sample $(P_s(n), H_k(n))$ represent paired data, for $s \in \hat{S}, k \in \hat{K}, n \in \hat{N}$.

Finally, the Pearson correlation coefficient is computed for every $s \in \hat{S}, k \in \hat{K}$:

$$\rho_{P_s,H_k} = \frac{\mathrm{Cov}(P_s, H_k)}{\sqrt{\mathrm{Var}(P_s)\mathrm{Var}(H_k)}}, \tag{1}$$

and the key is selected, e.g., as

$$\arg\max_k \quad |\rho_{P_s,H_k}|. \tag{2}$$

Note that the Pearson correlation coefficient may provide suboptimal results when the correlated variables are not normal or their relationship is not linear [7].

B. Distance Correlation

Distance correlation, equal to Brownian correlation, is a multivariate generalization and extension of the Pearson product-moment correlation coefficient [9], [10]. The distance correlation between $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$ is the number $\mathcal{R}(X, Y)$ defined as

$$\mathcal{R}^2(X,Y) = \begin{cases} \frac{\mathcal{V}^2(X,Y)}{\mathcal{V}^2(X)\mathcal{V}^2(Y)}, & \text{if } \mathcal{V}^2(X)\mathcal{V}^2(Y) > 0; \\ 0, & \text{if } \mathcal{V}^2(X)\mathcal{V}^2(Y) = 0; \end{cases} \tag{3}$$

where $\mathcal{V}(X, Y)$ is distance covariance, and $\mathcal{V}(X)$ is distance variance. To define these, let us first define a distance matrix.

For random sample $(X(n), Y(n))$, $n \in \hat{N}$, of N i.i.d. random vectors from joint distribution, we compute the Euclidean distance matrices $(a_{kl}) = |X(k) - X(l)|_p$ and $(b_{kl}) = |Y(k) - Y(l)|_q$, where $k, l \in \hat{N}$. Furthermore, we compute the double centered Euclidean distance matrices as

$$(A_{kl}) = a_{kl} - \overline{a_{k\cdot}} - \overline{a_{\cdot l}} + \overline{a_{\cdot\cdot}}, \tag{4}$$

where $k, l \in \hat{N}$, and

$$\overline{a_{k\cdot}} = \frac{1}{N}\sum_{l\in\hat{N}} a_{kl}, \quad \overline{a_{\cdot l}} = \frac{1}{N}\sum_{k\in\hat{N}} a_{kl}, \quad \overline{a_{\cdot\cdot}} = \frac{1}{N^2}\sum_{k,l\in\hat{N}} a_{kl}. \tag{5}$$

Define (B_{kl}) similarly for $k, l \in \hat{N}$. The sample distance covariance $\mathcal{V}(X, Y)$ is then defined as

$$\mathcal{V}^2(X,Y) = \frac{1}{N^2}\sum_{k,l\in\hat{N}} A_{kl}B_{kl}, \tag{6}$$

and sample distance variance as $\mathcal{V}^2(X) = \mathcal{V}^2(X, X)$.

Note that the distance correlation is applicable to random variables of arbitrary and unequal dimensions. Also, unlike the Pearson correlation coefficient, the distance correlation is equal to zero if and only if the variables are independent.

C. Clock Cycle-wise Correlation Power Analysis

While the distance correlation can be used in the same univariate fashion as the Pearson correlation coefficient, it allows for further extensions. Suppose we have power traces containing S samples, measured during C clock cycles. Instead of considering the power traces to be S random variables $P_s(t) \in \mathbb{R}$, we may consider them to be C random variables $Q_c(t) \in \mathbb{R}^{S/C}$. Using distance correlation, these can be correlated to hypotheses $H_k(t)$ for every $c \in \hat{C}$, $k \in \hat{K}$, and the key can be selected once again as

$$\arg\max_k \quad \mathcal{R}(Q_c, H_k). \tag{7}$$

Note that while this is a multivariate side-channel distinguisher, it uses sample points within a single clock cycle. Therefore, it exploits univariate side-channel leakage resulting in a univariate attack, unlike more complex bivariate/multivariate attacks on protected implementations.

III. Experimental Evaluation

A. Methodology

We evaluate and compare three correlation-based side-channel distinguishers:

- Pearson correlation, point-wise,
- distance correlation, point-wise,
- distance correlation, clock cycle-wise,

attacking AES-128 encryption on three FPGA platforms:

- Evariste-III [15] system, with Altera Cyclone III (65 nm),
- Sakura-G [16] board, with Xilinx Spartan 6 (45 nm),
- DPABoard [17] board, with Xilinx Artix 7 (28 nm).

The voltage drop over the FPGA core is sampled using Picoscope 6404D oscilloscope. We aim our attack at the last round working register Hamming distance leakage. The point-wise correlation is computed for all sample points, capturing entire encryption, with no prior points-of-the-interest analysis. The clock cycle-wise correlation is computed for all the 11 clock cycles.

The distinguishers are compared using success rate and guessing entropy metrics [14]. Assume all the K possible key candidates sorted according to the respective distinguisher, with the most probable key candidate on the first position; and define $pos \in \hat{K}$ as a position of the correct key candidate. Success rate is then defined as $\mathrm{Succ} = \Pr(pos = 1)$, i.e., the probability of the correct key being successfully revealed. Guessing entropy is defined as $\mathrm{GE} = \mathbb{E}(pos)$, i.e., the expected position of the correct key guess.

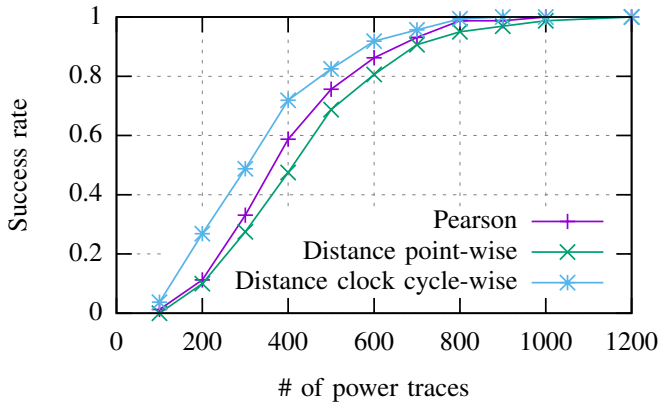
The presented results, for each platform, are based on 50 independent data sets and averaged over all 16 bytes of the AES key.

B. Results

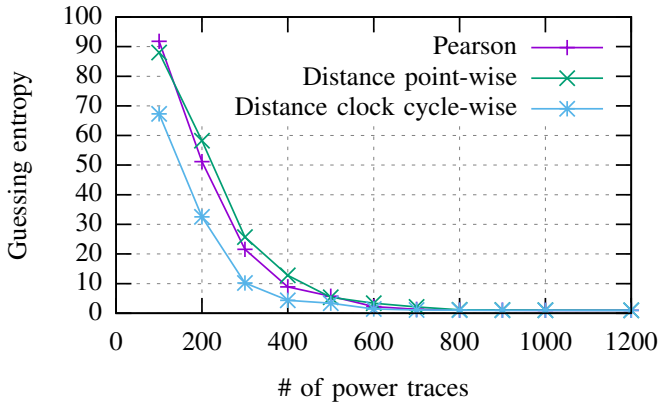
Figure 1 shows results for Evariste III + Altera Cyclone III FPGA, Figure 2 shows results for Sakura-G, i.e., Xilinx Spartan 6 FPGA, and Figure 3 shows results for DPABoard, i.e., Xilinx Artix 7 FPGA. While point-wise distance correlation performs the same or slightly worse than Pearson correlation, the cycle-wise analysis shows better results on all three platforms, most notably on Altera Cyclone III. On the Sakura-G board, the results are comparable for all three distinguishers.

C. Discussion and Future Work

Distance correlation has shown to be at least as useful, in the side-channel analysis context, as the product-moment Pearson correlation coefficient. We expect it to perform even better in specific scenarios where a significantly non-linear relationship appears, such as attacking some ASICs [7].

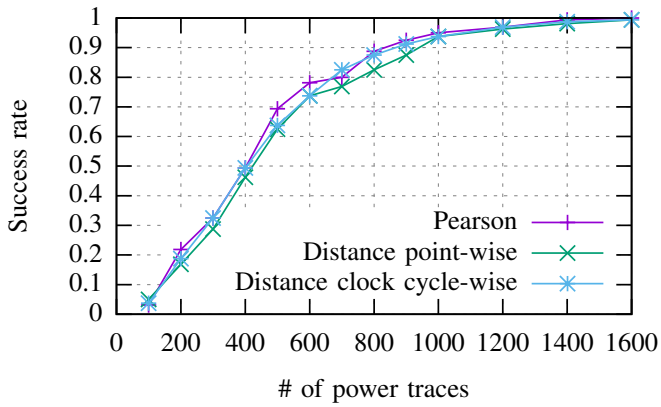


(a) Average Success rate.

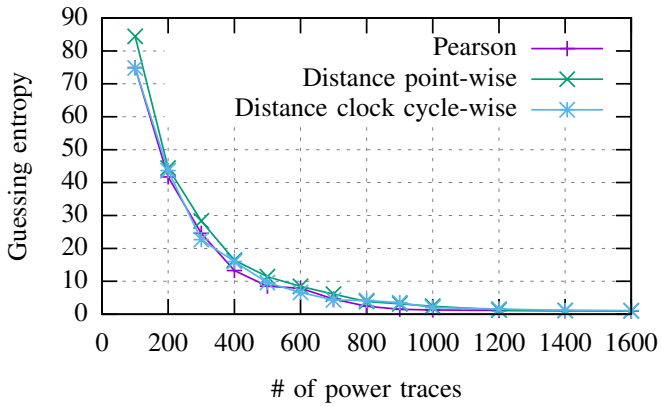


(b) Average Guessing entropy.

Figure 1: Evariste III + Altera Cyclone III.

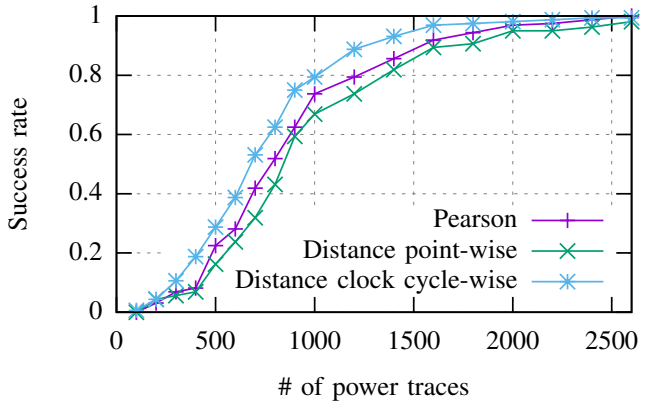


(a) Average Success rate.

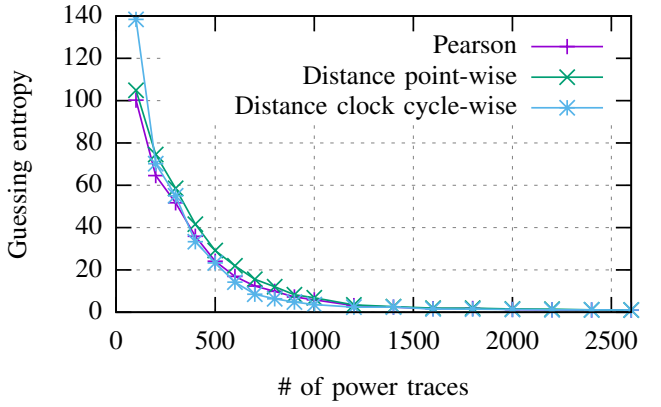


(b) Average Guessing entropy.

Figure 2: Sakura-G (Xilinx Spartan 6).



(a) Average Success rate.



(b) Average Guessing entropy.

Figure 3: DPABoard (Xilinx Artix 7).

The main pitfall of the distance correlation is the computational and memory complexity when using the described distance matrix approach, quadratic with the number of samples (i.e., power traces), in comparison with linear complexity of the product-moment correlation [18], rendering it unpractical for more complex scenarios, such as attacking protected implementations, where significantly more power traces are required due to the noise amplification effect. A possible solution to this problem could be using a different approach to approximating the distance covariance or reducing the overall attack complexity. On the other hand, the computational complexity grows only linearly with the sample dimensions.

Given the multivariate nature of the distance correlation, it could be considered for multivariate leakage exploitation, e.g., effectively attacking protected (masked) cryptographic implementations. Moreover, given the fact that the distance correlation is equal to zero if and only if the variables are independent, its ability to measure both linear and non-linear dependence, and its generally weak assumptions on the examined variables, it could serve as a basis for leakage assessment methodology.

IV. Conclusion

In this paper, we have given a fair experimental comparison of the Pearson and distance correlation coefficients as side-channel distinguishers attacking AES on three different FPGA platforms. Furthermore, we have proposed and evaluated a clock cycle-based distinguisher. We have shown that the distance correlation is a powerful alternative to the widely known Pearson correlation. Moreover, we have discussed the distance correlation properties and we have proposed further use cases worth future evaluation.

Acknowledgment

Petr Socha is a member of the student research team of the internal Czech Technical University (CTU) project No. SGS20/211/OHK3/3T/18. Computational resources were supplied by the project "e-Infrastruktura CZ" (e-INFRA CZ LM2018140) supported by the Ministry of Education, Youth and Sports of the Czech Republic.

References

[1] P. Kocher, J. Jaffe, and B. Jun, *Differential Power Analysis*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 388–397.

[2] E. Brier, C. Clavier, and F. Olivier, "Correlation power analysis with a leakage model," in *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, 2004, pp. 16–29.

[3] B. den Boer, K. Lemke, and G. Wicke, "A dpa attack against the modular reduction within a crt implementation of rsa," in *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, 2002, pp. 228–243.

[4] B. Gierlichs, L. Batina, P. Tuyls, and B. Preneel, "Mutual information analysis," in *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, 2008, pp. 426–442.

[5] N. Veyrat-Charvillon and F.-X. Standaert, "Mutual information analysis: How, when and why?." in *CHES*, vol. 5747. Springer, 2009, pp. 429–443.

[6] C. Whitnall, E. Oswald, and L. Mather, "An exploration of the kolmogorov-smirnov test as a competitor to mutual information analysis," in *International Conference on Smart Card Research and Advanced Applications*. Springer, 2011, pp. 234–251.

[7] L. Batina, B. Gierlichs, and K. Lemke-Rust, "Comparative evaluation of rank correlation based dpa on an aes prototype chip," in *International Conference on Information Security*. Springer, 2008, pp. 341–354.

[8] B. Timon, "Non-profiled deep learning-based side-channel attacks with sensitivity analysis," *IACR Transactions on Cryptographic Hardware and Embedded Systems*, pp. 107–131, 2019.

[9] G. J. Székely, M. L. Rizzo, N. K. Bakirov *et al.*, "Measuring and testing dependence by correlation of distances," *The annals of statistics*, vol. 35, no. 6, pp. 2769–2794, 2007.

[10] G. J. Székely and M. L. Rizzo, "Brownian distance covariance," *The annals of applied statistics*, pp. 1236–1265, 2009.

[11] C. Whitnall and E. Oswald, "A cautionary note regarding the usage of leakage detection tests in security evaluation," Cryptology ePrint Archive, Report 2019/703, 2019, https://eprint.iacr.org/2019/703.

[12] J. Kundrata, D. Fujimoto, Y. Hayashi, and A. Barić, "Comparison of pearson correlation coefficient and distance correlation in correlation power analysis on digital multiplier," in *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*. IEEE, 2020, pp. 146–151.

[13] J. Daemen and V. Rijmen, *The design of Rijndael: AES-the advanced encryption standard*. Springer Science & Business Media, 2013.

[14] F.-X. Standaert, T. Malkin, and M. Yung, "A unified framework for the analysis of side-channel key recovery attacks." in *Eurocrypt*, vol. 5479. Springer, 2009, pp. 443–461.

[15] N. Bochard, C. Marchand, O. Pet'ura, L. Bossuet, and V. Fischer, "Evariste iii: A new multi-fpga system for fair benchmarking of hardware dependent cryptographic primitives," in *Workshop on Cryptographic Hardware and Embedded Systems, CHES 2015*, 2015.

[16] H. Guntur, J. Ishii, and A. Satoh, "Side-channel attack user reference architecture board sakura-g," in *Consumer Electronics (GCCE), 2014 IEEE 3rd Global Conference on*. IEEE, 2014, pp. 271–274.

[17] M. Bartík and J. Buček, "A low-cost multi-purpose experimental fpga board for cryptography applications," in *Advances in Information, Electronic and Electrical Engineering (AIEEE), 2016 IEEE 4th Workshop on*. IEEE, 2016, pp. 1–4.

[18] P. Socha, V. Miškovský, H. Kubátová, and M. Novotný, "Optimization of pearson correlation coefficient calculation for dpa and comparison of different approaches," in *Design and Diagnostics of Electronic Circuits & Systems (DDECS), 2017 IEEE 20th International Symposium on*. IEEE, 2017, pp. 184–189.