Combined Distinguishers to Enhance the Accuracy and Success of Side Channel Analysis

M. Ali Vosoughi
Department of Electrical and Computer Engineering
University of Rochester
Rochester, New York, 14627
mvosough@ur.rochester.edu

Selçuk Köse
Department of Electrical and Computer Engineering
University of Rochester
Rochester, New York, 14627
selcuk.kose@rochester.edu

***Abstract*—For the first time, the combination of mutual information analysis and correlation power analysis is proposed to enhance the accuracy and success rate of side channel analysis. Using the k-nearest-neighborhood (KNN) algorithm, correlation power analysis is combined with mutual information analysis to classify various possible keys to two classes of correct and wrong keys. The advantage of the combination of the distinguishers is two fold. First, the accuracy of the estimation is enhanced due to availability of multiple possible values for the correct key. Second, the number of measurements required to disclose the correct key is reduced by combining the distinguishers. The effectiveness of combined distinguisher is verified by extensive simulations. The number of measurements required to perform a side channel attack with a success rate of 90% is improved, respectively, by 20% and 49%, as compared to individual correlation power analysis and mutual information analysis.**

***Index Terms*—Side channel analysis, side channel distinguisher, mutual information analysis, correlation power analysis, combined side channel distinguisher.**

I. Introduction

SIDE channel attacks are an important class of attacks on cryptographic circuits to obtain the secret key through the analysis of leakage information, such as power consumption, electromagnetic emanations, and timing information. With an effective side channel attack, the secret key can be recovered in couple of minutes with inexpensive equipment, while a supercomputer can take 149 trillion years to break the secret key with the brute force search [1]. Since the first side channel attack in 1996 [2], physical vulnerabilities of the devices have been utilized to endanger the privacy of the electronic devices, such as mobiles and computers [3]–[6]. Various tools, called as distinguishers, have been proposed for the analysis of side channel information to disclose the secret key of the cryptographic circuits. Mutual information analysis (MIA) [7] and correlation power analysis (CPA) [8] are among the distinguishers used for analysis of side channel information. In CPA, linear statistical dependencies between the side channel information and a hypothetical model, which is predicted by the attacker for the side channel output, are detected, while in MIA, any statistical and functional dependencies between the actual leakage and hypothetical model are detected [9].

The optimality of each distinguishers is based on the noise distribution of the leakage information and the accuracy of the hypothesized power model predicted by an attacker. Each of the distinguishers has the conditions under which the distinguisher is optimal [10]. For the noise with Gaussian distribution and the known hypothetical power model, the CPA is optimal distinguisher [10], and the correct key will be obtained with the less number of measurements as compared to the MIA. However, if the actual power consumption of the circuit differs from the hypothetical power model of the attacker, or if the noise distribution deviates from the Gaussian distribution, the optimality of the CPA will be compromised by deviations from the assumptions, and the MIA will be optimal as compared to CPA in the presence of the unknown power model, or non-Gaussian noise distribution [11]. These circumstances may occur in the presence of countermeasures [12]–[19] or nanoscale nonidealities in the CMOS device [11], [20]. Thus far, choosing one of the distinguishers has been the final solution to the side channel problems. Although a distinguisher can achieve the optimal performance for side channel analysis, the results obtained from various distinguishers are not necessarily the same. The differences in the results of the distinguishers offer potentially complementary information about the side channel leakage, which can be used to improve the efficiency and accuracy of the side channel analysis.

In this paper, the combination of normalized mutual information [11] and Pearson's correlation using k-nearest-neighborhood (KNN) algorithm [21] is proposed to enhance the accuracy of the side channel analysis, and to reduce the number of measurements to disclose the key (MTD). Normalized mutual information is analogous of Pearson's correlation in information theory and varies in $[0, 1]$ range [22]. Using KNN algorithm, the data are classified based on the similarities, which is determined by the distance from the nearest neighbors [21]. To the best of the knowledge of the authors, the combination of the distinguishers has not been published in any previous work, and this is the first work on the combination of the distinguishers.

The rest of the paper is as follows. In Section II, the combination of two distinguishers is proposed to increase the

This work was supported in part by the National Science Foundation CAREER Award under Grant CCF-1350451, in part by the National Science Foundation Award under Grant CCF-1421988, and in part by the Cisco Research Award.

978-1-7281-0397-6/19/$31.00 ©2019 IEEE

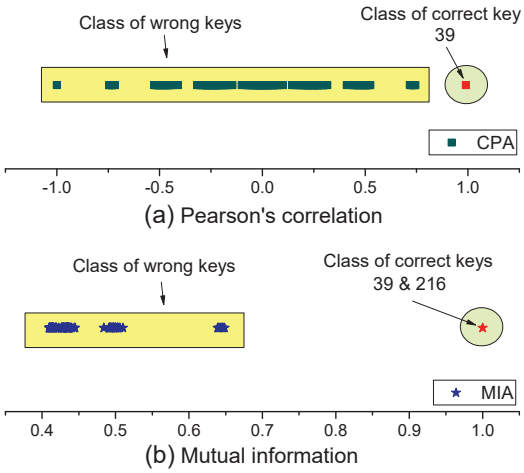

Fig. 1. The classification of keys to the classes of correct and wrong keys is shown for (a) CPA and (b) MIA. Because the process of finding the probability density function in simulations is symmetric, the value of mutual information for the possible keys and their complements are the same.

accuracy and reduce the MTD of the side channel analysis. In Section IV, the proposed method is evaluated in detail for four practical scenarios and the effectiveness of the proposed method in reducing the MTD and increasing the accuracy is verified. Finally, conclusions are offered in Section IV.

II. Proposed Combined Distinguisher

A side channel distinguisher can be viewed as a classifier, which divides possible keys into the classes of correct key(s) and wrong keys [23]. As shown in Fig. 1a and 1b, the problem of finding the correct key by CPA and MIA can be considered as a classification problem with two classes. Given the MIA and CPA as distinguishers, and the fact that these distinguishers may carry complementary information, various techniques can be applied to combine them [24]. In Algorithm. 1 the proposed combined distinguisher to recover the set of a secret key(s) (indicated by $\mathcal{K}^*$) from the set of candidate keys $\mathcal{K}$ and leakage trace $\mathcal{L}$ for MIA (indicated by μ) and CPA (indicated by ρ) is summarized.

Algorithm 1 Proposed algorithm for combined distinguisher

Input: $\mathcal{K} = \{k_1, \ldots, k_{|\mathcal{K}|}\}$, $\mathcal{L} = \{l_1, \ldots, l_{|\mathcal{L}|}\}$
Output: Secret key(s) $\mathcal{K}^*$
Initialization :
1: For each k_i compute ρ_i and μ_i
2: $\mathcal{X} \leftarrow \{(\rho_i, \mu_i), i = 1, \ldots, |\mathcal{K}|$
3: $d_i \leftarrow$ compute 1-NN distance for each $x_i \in \mathcal{X}$
4: $\mathcal{K}^* \leftarrow pickIndex(max(d))$
5: **return** $\mathcal{K}^*$

Reducing the number of MTD and decreasing sensitivity of the secret key recovery resultant to the accuracy of the leakage model are two principal motivations for combining distinguishers. By combining the distinguishers, the secret key can be obtained using a fewer number of measurements as compared to the use of individuals of the distinguishers thanks to additive information given by different distinguishers; however, the results may not be obtained by individual distinguisher with the same MTD.

Due to the imprecision in the leakage model and computations of CPA and MIA, the use of individual distinguisher may lead to a wrongly distinguished key [11], [13]. Utilizing the combined distinguisher reduces the sensitivity of the hypothesis test to the imprecision of the leakage model and computation errors.

III. Evaluation of the Proposed Method

Simulations are performed in MATLAB where machine learning and pattern recognition toolboxes are used [25]. Prazen window density estimation is used for probability density estimation to use in the mutual information and entropy calculations [26]. For simulations, the simultaneous computation of normalized mutual information and Pearson's correlation is used on side channel leakage information. The utilization of the scatter plot facilitates the inspection of patterns in the information obtained from the side channel by MIA and CPA. The pattern seen in Fig. 2 is the effect of increasing noise and increasing the number of measurements on the classification boundaries. As shown in Fig. 2a, Fig. 2b, and Fig. 2c, as the noise increases, the scattering of the mutual information and Pearson's correlation values is increased, and the boundary between the classes of possible correct keys and wrong keys is obscured. As the number of measurements is increased (Fig. 2d, Fig. 2e, and Fig. 2f), the boundary between the classes will be more clear, and the classification of the keys will be possible to the class of correct key(s) and the class of wrong keys. In order to confirm the improvement in the accuracy and success of the side channel analysis using the combined distinguisher, four scenarios corresponding to four practical cases are considered. These scenarios are based on literature for the optimal conditions for each of the distinguishers [11].

1) Known model and Gaussian noise: The hypothetical power model by an attacker is linearly correlated with the processed information in the cryptographic circuit in the absence of countermeasure and conditional variations in the power consumption, under these circumstances, distribution of noise will be Gaussian due to the central limit theorem [9], [27]. In the simulations shown in Fig. 2, known hypothetical leakage model and noise with Gaussian distribution are used. As shown in Fig. 2d, Fig. 2e, and Fig. 2f, CPA outperforms the MIA and converges faster (with lower number of measurements) to the correct key. When the number of measurements increase further, both MIA and CPA converges to the same result as the correct key.

2) Unknown model and Gaussian noise: Different encoding techniques, such as bus-invert coding [28], [29], are typically employed in modern digital circuits to reduce the power dissipation and increase the reliability. Due to the conditional variations in the power consumption of the circuit with bus-inverting method, hypothetical power model will be different from the linear power consumption model [28]–[30]. This

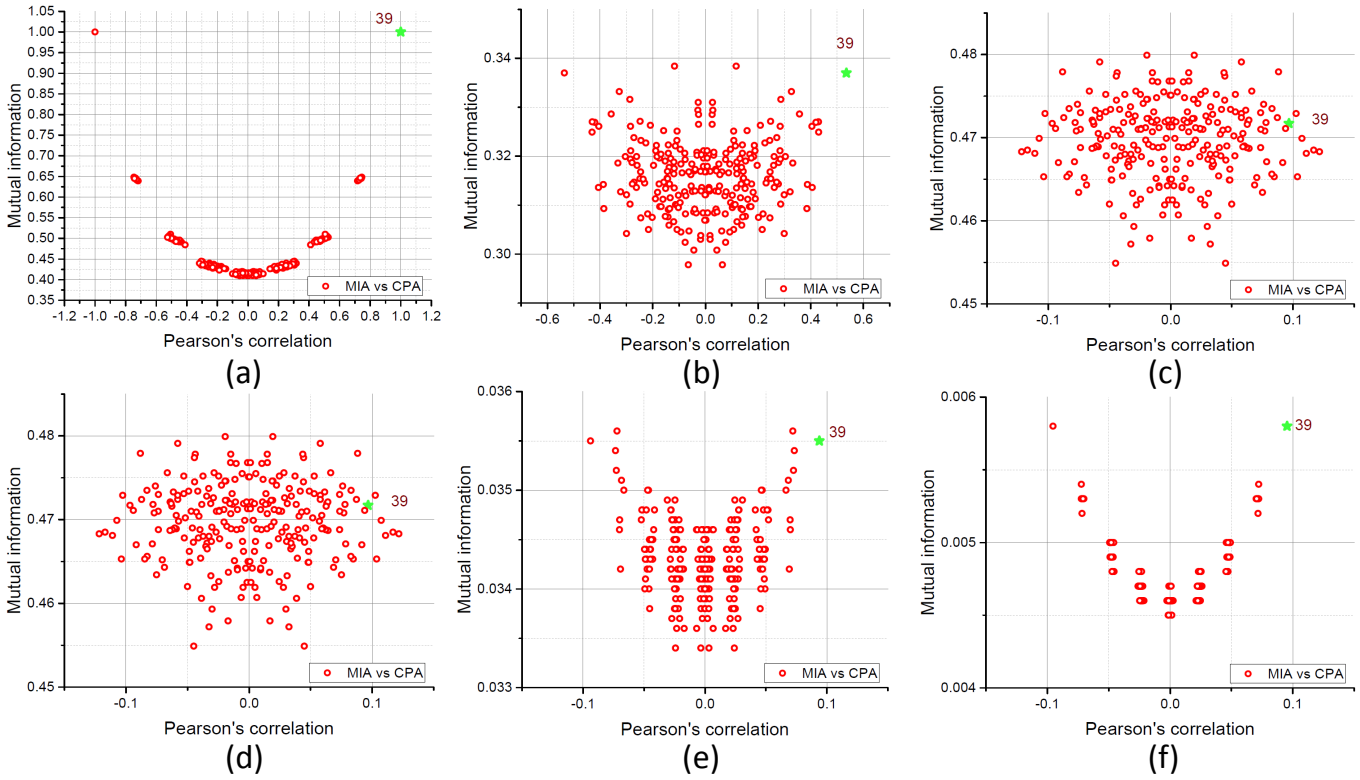

Fig. 2. Effect of increasing the noise and increasing the number of measurements on the scattering behavior of the distinguishers. The hypothetical power consumption model is identical with the actual leakage power of the cryptographic circuit, and the noise distribution is Gaussian. For the fixed number of plaintexts (=500), noise is increased from (a) $\sigma_N = 0\ mV$ to (b) $\sigma_N = 5\ mV$ and (c) $\sigma_N = 30\ mV$. For the fixed noise $\sigma_N = 30\ mV$, number of measurements is increased from (d) 500 plaintexts to (e) 50,000 and (f) 500,000 plaintexts.

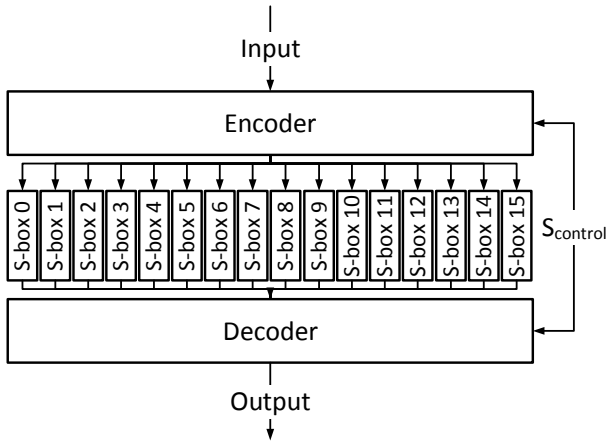

Fig. 3. S-box of AES with an encoder and decoder for reducing the power consumption and increasing the security of the memory interface [29], [30]. $S_{control}$ is related to the additional bus(es) for controlling the coding process.

technique is used in the design of a secure memory interface in [29] and to design a low-power S-box of an AES in [30], as shown in Fig. 3. In Fig. 4 the simulated case is for an S-box of AES with bus-invert coding, where the complement of data is transferred when the Hamming weight of the data on the bus exceeds four. While CPA converges to the wrong key, the MIA converges to the correct key(s), and the result of combined distinguisher is dominated by MIA. As shown in Fig. 4c, CPA converges to the wrong key (the correct key appears with the least correlation in the CPA), while using the proposed method, the correct key is one of the two keys with the highest probability.

3) Known model and non-Gaussian noise: In presence of measurement noise, quantization noise in sampling devices, and spikes, the final noise distribution of the leakage information can be non-Gaussian [10]. The simulation for non-Gaussian noise (uniform with $\sigma_N = 30\ mV$) with a known power model is shown in Fig. 5, where the MIA outperforms the CPA [11].

4) Unknown model and a combination of Gaussian and non-Gaussian noise: In the presence of quantization noise, spikes, measurement inaccuracies, and conditionally variable power consumption for the cryptographic circuit (*e.g.* bus-invert coding technique), the noise distribution will be combination of Gaussian and non-Gaussian, while the hypothetical power consumption model will be unknown to the attacker. Then MIA is expected to outperform CPA for the unknown power model and non-Gaussian noise distribution, while the CPA is expected to outperform the MIA for the Gaussian noise assumption of the side channel leakage. The competitive situation for obtaining the secret key using the CPA and MIA is when the MTD for CPA and MIA are close, as shown in Fig. 6. Using the combination of distinguishers, in presence of this competitive situation for distinguishers, the MTD for success rate of 90% is improved, respectively, by 20% and 49% as compared to the individual CPA and MIA, as shown

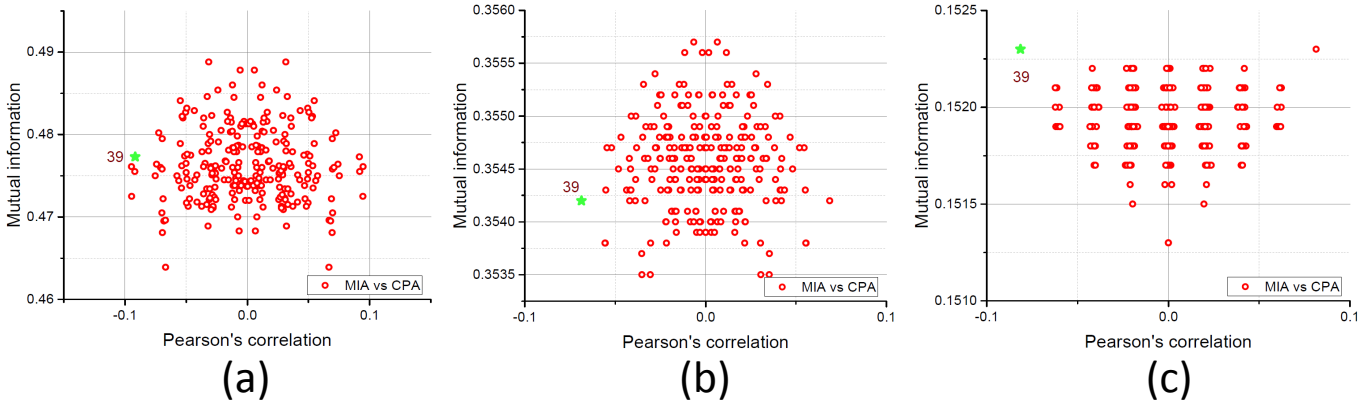

Fig. 4. Noise distribution is Gaussian with $\sigma_N = 5\ mV$, but the leakage information from the circuit does not have a linear relationship with the hypothetical power model, and the hypothetical power model is unknown to the attacker. The simulated case is for an 8 bits S-box with bus inverting.

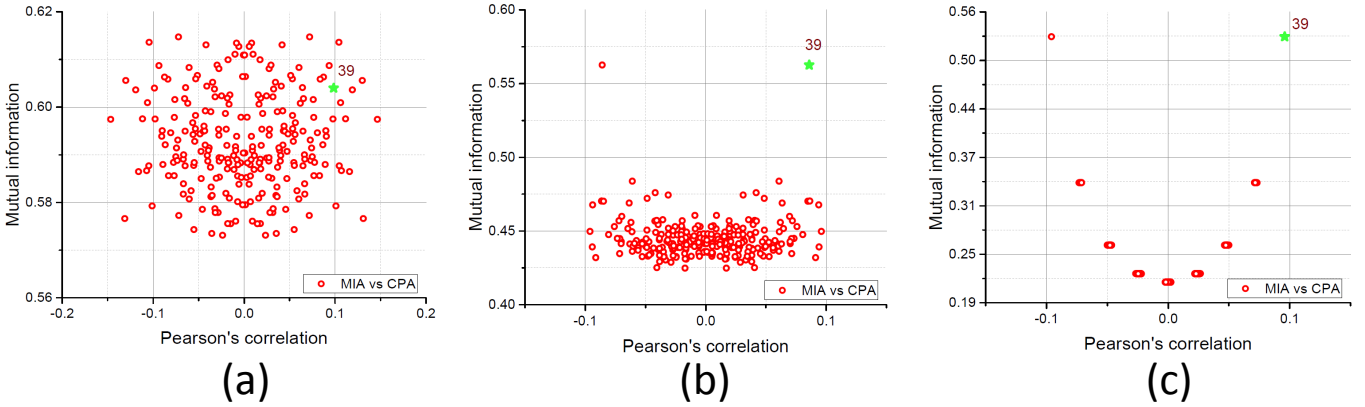

Fig. 5. Uniform noise with $\sigma_N = 30\ mV$ and known power model. MIA outperforms the CPA, and converges faster to the correct key. By increasing the number of measurements, the key is obtained through both the CPA and MIA.

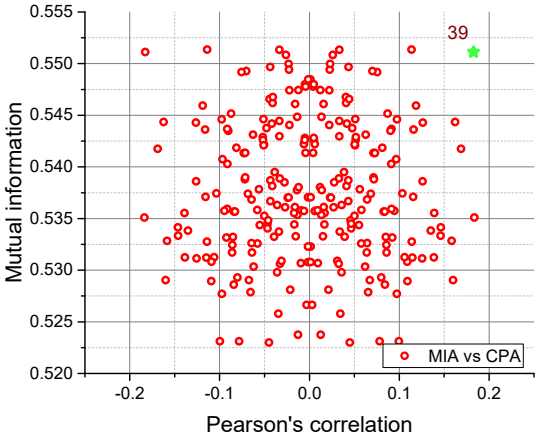

Fig. 6. The competitive situation for obtaining key using the CPA and MIA. Hypothetical power model is unknown to the attacker. Noise is combination of uniform distribution with $\sigma_{N1} = \pm 0.5\ mV$ and Gaussian distribution with $\sigma_{N2} = \pm 25\ mV$.

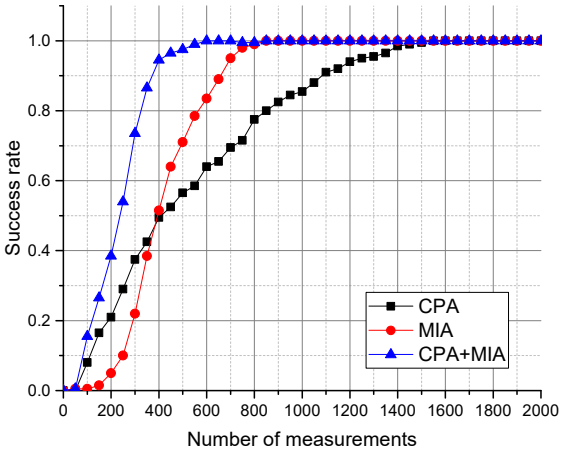

Fig. 7. Success rate of MIA, CPA, and combination of distinguishers with proposed method is shown. The MTD for success rate for combined distinguisher is improved by 20% and 49%, as compared to the use of individual CPA and MIA. Noise is with $\sigma_N = 26mV$, and noise distribution is combination of Gaussian and uniform distribution. The bus-inverting method is applied to the S-box, where the complement of data are processed when the Hamming weight of data exceeds four.

in Fig. 7.

IV. Conclusion

A technique is proposed to combine the MIA and CPA distinguishers using the k-nearest-neighborhood algorithm, Pearson's correlation, and normalized mutual information. The advantages of the proposed method are to reduce the number of MTD and the sensitivity of the hypothesis test to the imprecision of the leakage model and computational errors as compared to sole distinguisher. The effectiveness of the proposed technique is confirmed using extensive simulations for various practical scenarios. The MTD for the success rate of 90% with combined distinguisher is improved, respectively, by 20% and 49%, as compared to CPA and MIA distinguishers.

REFERENCES

[1] M. Arora, "How Secure is AES Against Brute Force Attacks?," *EE Times*, Vol. 5, No. 7, May 2012.

[2] P. Kocher, "Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems," *Annual International Cryptology Conference*. Springer Berlin Heidelberg, pp. 104–113, August 1996.

[3] W. Yu and S. Köse, "Security Implications of Simultaneous Dynamic and Leakage Power Analysis Attacks on Nanoscale Cryptographic Circuits," *Electronics Letters*, Vol. 52, No. 6, pp. 466–468, 2016.

[4] E. Karimi, Z. H. Jiang, Y. Fei, and D. Kaeli, "A Timing Side-Channel Attack on a Mobile GPU," *IEEE 36th International Conference on Computer Design*, pp. 67–74, October 2018.

[5] N. Takbiri, A. Houmansadr, D. L. Goeckel, and H. Pishro-Nik, "Matching Anonymized and Obfuscated Time Series to Users' Profiles," *IEEE Transactions on Information Theory*, Vol. 65, No. 2, pp. 724–741, February 2019.

[6] S. K. Khatamifard, L. Wang, S. Köse, and U. R. Karpuzcu, "A New Class of Covert Channels Exploiting Power Management Vulnerabilities," *IEEE Computer Architecture Letters*, Vol. 17, No. 2, pp. 201–204, July 2018.

[7] B. Gierlichs, L. Batina, P. Tuyls, and B. Preneel, "Mutual Information Analysis," *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, pp. 426–442, August 2008.

[8] E. Brier, C. Clavier, and F. Olivier, "Correlation Power Analysis with a Leakage Model," *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, pp. 16–29, August 2004.

[9] A. Moradi, N. Mousavi, C. Paar, and M. Salmasizadeh, "A Comparative Study of Mutual Information Analysis Under a Gaussian Assumption," *International Workshop on Information Security Applications*. Springer, pp. 193–205, August 2009.

[10] A. Heuser, O. Rioul, and S. Guilley, "Good Is Not Good Enough," *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, pp. 55–74, September 2014.

[11] È. De Chèrisey, S. Guilley, A. Heuser, and O. Rioul, "On the Optimality and Practicability of Mutual Information Analysis in Some Scenarios," *Cryptography and Communications*, Vol. 10, No. 1, pp. 101–121, January 2018.

[12] M. J. Azhar, F. Amsaad, and S. Köse, "Duty-Cycle-Based Controlled Physical Unclonable Function," *IEEE Transactions on Very Large Scale Integration Systems*, Vol. 26, No. 9, pp. 1647–1658, September 2018.

[13] N. Bruneau *et al.*, "Stochastic Collision Attack," *IEEE Transactions on Information Forensics and Security*, Vol. 12, No. 9, pp. 2090–2104, September 2017.

[14] W. Yu and S. Köse, "False Key-Controlled Aggressive Voltage Scaling: A Countermeasure Against LPA Attacks," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 36, No. 12, pp. 2149–2153, December 2017.

[15] W. Yu and S. Köse, "Security-Adaptive Voltage Conversion as a Lightweight Countermeasure Against LPA Attacks," *IEEE Transactions on Very Large Scale Integration Systems*, Vol. 25, No. 7, pp. 2183–2187, July 2017.

[16] W. Yu and S. Köse, "Exploiting Voltage Regulators to Enhance Various Power Attack Countermeasures," *IEEE Transactions on Emerging Topics in Computing*, Vol. 6, No. 2, pp. 244–257, April 2018.

[17] W. Yu and S. Köse, "A Lightweight Masked AES Implementation for Securing IoT Against CPA Attacks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 64, No. 11, pp. 2934–2944, November 2017.

[18] W. Yu and S. Köse, "Implications of Noise Insertion Mechanisms of Different Countermeasures Against Side-Channel Attacks," *IEEE International Symposium on Circuits and Systems*, pp. 1–4, May 2017.

[19] W. Yu, O. A. Uzun, and S. Köse, "Leveraging On-Chip Voltage Regulators as a Countermeasure Against Side-Channel Attacks," *Design Automation Conference*, pp. 1–6, June 2015.

[20] M. A. Vosoughi, H. Torun, and G. Dundar, "Noise Analysis in Switched Capacitor Amplifier Based Sensors," *New Generation of Circuits and Systems*, pp. 249–252, September 2017.

[21] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Chapter 2, pp. 61–64, Academic Press, 2008.

[22] A. Strehl and J. Ghosh, "Cluster Ensembles — A Knowledge Reuse Framework for Combining Multiple Partitions," *Journal of Machine Learning Research*, Vol. 3, pp. 583–617, December 2002.

[23] S. Picek *et al.*, "Side-Channel Analysis and Machine Learning: A Practical Perspective," *2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 4095–4102, May 2017.

[24] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On Combining Classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 3, pp. 226–239, March 1998.

[25] M. Chen, "Pattern Recognition and Machine Learning Toolbox," *MATLAB Central File Exchange*, 2016.

[26] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2016.

[27] X. Standaert, E. Peeters, G. Rouvroy, and J. Quisquater, "An Overview of Power Analysis Attacks Against Field Programmable Gate Arrays," *Proceedings of the IEEE*, Vol. 94, No. 2, pp. 383–394, February 2006.

[28] M. R. Stan and W. P. Burleson, "Bus-Invert Coding for Low-Power I/O," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 3, No. 1, pp. 49–58, March 1995.

[29] L. Benini *et al.*, "Energy-Efficient Data Scrambling on Memory-Processor Interfaces," *Proceedings of the International Symposium on Low Power Electronics and Design*, pp. 26–29, August 2003.

[30] J. Zhang, Q. Zuo, and T. Zhang, "Reducing the Power Consumption of the AES S-Box by SSC," *International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 2226–2229, September 2007.