A Fuzzy C-means algorithm to detect cryptographic signatures on DPA/DEMA attacks

Plínio Finkenauer Junior, Vinícius Valduga de Almeida Camargo, Vitor Lima, Marilton S. Aguiar, Rafael I. Soares
Technology Development Center - CDTec
Federal University of Pelotas - UFPel
Pelotas, Brazil
{pfinkenauer, vvacamargo, vgdlima, marilton, rafael.soares}@inf.ufpel.edu.br

***Abstract*—Side Channel Attacks (SCA) poses a severe threat to the security of cryptographic devices. SCA exploits the physical leaked information from a system during its cryptographic procedure to unveil the secret key. Differential Power Analysis (DPA) and Differential Electromagnetic Analysis (DEMA) are effective examples which benefit from this approach. Nevertheless, these attacks are susceptible to countermeasures based on random processing and, so, they require temporal alignment to be successful. This work proposes a preprocessing stage to filter and detect the electromagnetic radiation signature from the traces of a cryptographic device. The proposal includes signal processing techniques and the Fuzzy C-means algorithm. A study case with a hardware implementation of the DES algorithm submitted to DEMA attack is presented. Results point to the viability of the proposed method, demonstrating an average decrease of 4.9 times in the number of traces needed for a successful attack when compared to a threshold approach.**

***Index Terms*—Side Channel Attacks; Cryptanalysis; Machine Learning; Clustering; Signal Processing**

I. Introduction

System-on-Chip (SoC) developed for security applications relies upon encryption mechanisms to maintain the integrity and confidentiality of the system. The hardware implementing the cryptographic algorithm has been proved to leak side-channel information [1]. In view of this evidence, a class of cryptanalysis is proposed aiming the investigation of such knowledge. This class, known as Side Channel Attacks (SCA), consists of a series of strategies that exploit physical vulnerabilities leaked from a cryptography device. Based on this, SCA attacks take advantage of the dependency of the processed data by the system with this leaked information to reveal the cryptographic key [2]. Examples of such properties are processing time, power and the electromagnetic radiation from the circuit. Since these measures depend on the intern use of the secret key, the adversary can produce an efficient attack to retrieve this key and reveal confidential data.

Differential Power Analysis (DPA) is one of the most successful models of attack to modern cryptographic systems since it is passive and non-invasive [3]. That means, the attack explores the physical behavior of the device, and does not leave shreds of evidence of it. DPA operates by monitoring the power consumption from the target device. The power consumption of a system is acquired by observing the current consumption of the circuit during its operation [4]. Once these measurements, also referred as power traces, are collected, the data processed by the device can be correlated with its power consumption through statistical analysis in order to unveil the cryptographic key.

Differential Electromagnetic Analysis (DEMA) applies a similar procedure proposed by DPA, considering the electromagnetic radiation emitted by the system. The information obtained by the DPA attack is identified likewise in the DEMA since the electric charges flow yields a magnetic field equivalent to the power consumption observed [5]. The analysis presented in this paper was performed with electromagnetic traces, but are also applicable to power traces, for the case of a DPA attack.

An efficient DEMA attack requires the acquisition of an elevated number of power traces to be able to perform the correlation analysis [1], [6]. Furthermore, it is essential that the power traces are aligned in the time domain in order to establish a relation between the data processed and the assessed physical quantity [4]. In this way, the deliberated insertion of misalignments in the temporal scope represents a strong countermeasure because it makes it more difficult to recognize the cryptographic processing signature from the target device [7]. The identification of this target sequence aims to improve the attack, delimiting strategic regions and attenuating noise. A typical approach to extract the target sequence consists in defining two parameters: a threshold and a starting point for observing each trace. The first parameter denotes an estimate for the amplitude that differentiates the signature from noise [8].

In this context, this paper proposes a strategy to detect the target sequence of the power traces, through the application of signal processing techniques and unsupervised learning. The experiments are performed in electromagnetic traces obtained from an FPGA prototype of a DES (Data Encryption Standard) implementation proposed in [9]. After applying the methodology to extract the target signature, the resulting traces set are submitted to the DEMA attack flow.

The paper is structured as follows: Section II describes the proposed methodology for detecting the target sequence. Section III presents and discusses experimental results, while Section IV concludes this paper and gives directions for further work.

II. Methodology

Power or electromagnetic traces obtained from cryptographic systems with countermeasures demand a preprocessing stage previously to the correlation analysis to be able to unveil the secret key. An approach to introduce randomness and noise during the cryptographic procedure aims to conciliate the Globally Asynchronous and Locally Synchronous (GALS) design style with random frequency clock, as demonstrated in [9]. This architecture implements a pipeline where each stage generates locally its clock signal with different frequencies. During the execution, each stage chooses randomly the frequency which causes the misalignment in the time domain during the traces acquisition [10].

Fig. 1 shows an example of an electromagnetic trace obtained during the encryption in the GALS architecture with two stages, processing eight rounds of the DES algorithm in each one. The area highlighted by dashed line represents the first stage, which denotes the target signature for the attack and, consequently, the sequence to be detected and extracted. The arrows depict the high amplitude peaks indicating the noise induced by the trigger generated by the architecture to synchronize itself with the oscilloscope, highlighting the beginning and end of the significant data acquisition.

The proposed approach to identify the target signature is based on the application of a clustering technique. Clustering analysis is a machine learning procedure, whose purpose is to partition a data set into homogeneous agroupments, referred to as clusters. These clusters are formed so that the elements of the same agroupment present great similarity with each other and high distinction from elements of a different agroupment. A clustering technique is an unsupervised learning method since no label is given to it, being the algorithm itself responsible for identifying patterns in the data. Since clustering is not an invariable task, it is often necessary to adjust the data and the model parameters until the results achieve the desired response [11]. In this way, a preprocessing stage is conceived to clean the data to be handled by the clustering algorithm.

A. Preprocessing Stage

The target signature for the attack is comprised between the consumption peaks produced by the oscilloscope trigger. In this way, the first step involves eliminating the outside region by applying an exponential function in the trace to accentuate amplitude differences. From this resultant trace, the time period in which the maximum amplitude occurs is selected, and if this mark is less than the trace length divided by four, it is defined as the initial point of the trace containing the signature. After this primary reduction of the trace, the time period of maximum amplitude is identified again, denoting the final point of the trace. If the first picked value was bigger than one-quarter of the trace length, the steps described above should be done inversely. These values were empirically stipulated, through an exploratory analysis from a small subset of traces, and the obtained results show that this procedure is successful for all traces.

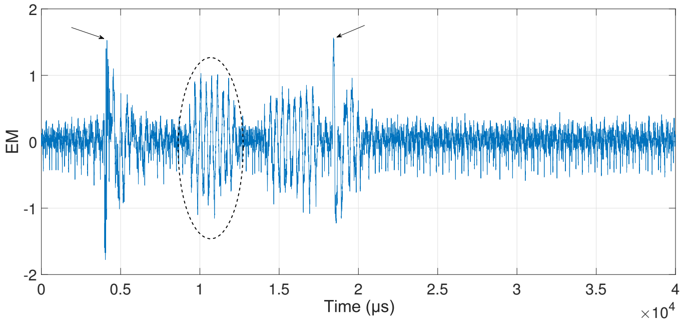


Fig. 1. Electromagnetic trace exposing a cryptographic procedure with different operation frequencies.

Subsequent to the reduction of the trace size, two signal processing techniques are used in order to emphasize the cryptographic signature and improve the signal to noise ratio (SNR). Since the DPA attack defines the first stage of the DES algorithm as the target of the attack, any other region is treated as noise during the detection phase.

The first filter applied is the Simple Moving Average (SMA), which smooths the shape of the trace and salients possible patterns in a time series [12]. In this context, SMA represents a low-pass filter. SMA is characterized as the average formed by a stipulated number of previous values of a dataset and is defined as

$$SMA_t = \frac{x_t + x_{t-1} + ... + x_{t-n+1}}{n} \tag{1}$$

where n denotes the number of observations used in the equation and t the time period where it occurred [13].

The final step in the preprocessing stage involves submitting the resultant trace to the hyperbolic tangent (*tanh*) function. The application of the *tanh* function accents the rounds of the DES algorithm since its outputs are assumed to be in the range of -1 to 1. The *tanh* is a common activation function for the learning process in neural networks [14] and is given by

$$tanh(x) = \frac{senh(x)}{cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2}$$

The application of both techniques described above allows a significant reduction of the noise while maintaining the response of the original signal. Algorithm 1 summarizes the steps adopted in the preprocessing stage for the dataset.

B. Clustering Stage

The chosen unsupervised technique for the clustering stage was the Fuzzy C-means (FCM) [15]. FCM works by iteratively advancing the centroids to its right location until convergence is achieved. The centroid of a cluster characterizes the mean of all its points. Since FCM is derived from fuzzy logic, each element partitioned can belong to more than one cluster [16]. The FCM quantifies these partitioned elements by their degree of belonging to each cluster. This value varies from 0 to 1 according to the similarity with the cluster, being 1 the most similar possible. FCM method is characterized by a quadratic

Algorithm 1: Pseudo code for the preprocessing stage.

```
Input: Traces set $T_{i...n}$
Output: Sub-traces for the clustering stage
for each $i \in T$ do
    traceX = $exp(T_i)$;
    max = $index(max(traceX))$;
    if $max < (length(trace)/4)$ then
        init = $max$;
        subTrace = $T_i[init :]$;
        final = $index(max(subTrace))$;
        subTrace = $T_i[init : final]$;
    else
        final = $max$;
        subTrace = $T_i[: final]$;
        init = $indice(max(subTrace))$;
        subTrace = $T_i[init : final]$;
    end
    end
    subTrace = $SMA(subTrace, n = 150)$;
    traceF= $tanh(subTrace)$;
end
```

object function, denominated as J_m, which must be minimized and is defined by

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} \mu_{ij}^{m} \left\| x_i - c_j \right\|^2 \tag{3}$$

where μ_{ij} represents the degree of belonging of the point x_i in the cluster j, c_j express the clusters centroids, and the fuzzy coefficient m denotes the level of fuzziness of the resulting classification.

The number of clusters (k) is an user-defined parameter, and its appropriated value is arbitrary. To find an optimal k, the adopted criterion was the sum of squared error for a different number of clusters. This technique, also referred as elbow method [17], is based on the Within-Cluster Sum of Squares (WCSS) and provides a visual estimate for detecting the right number of clusters. Figure 2 presents the elbow method implemented on the WCSS measure for 2 to 25 clusters in the trace produced by the preprocessing stage. The plot exhibits and elbow shape near the solution with $k = 5$, where the WCSS value is not decreasing drastically. Thus, this point was selected as the ideal number of clusters and the trace was able to be submitted to the FCM algorithm.

Fig. 3 shows an example of an electromagnetic trace applied to the proposed flow. In Fig. 3 (a) is presented the sub-trace resulting from the preprocessing stage which represents the input to the clustering algorithm. The highlighted yellow area exhibits the first stage of the DES algorithm with its eight rounds, representing the same underlined region of Fig. 1. Fig. 3 (b) presents the application of FCM with five clusters in the trace of Fig. 3 (a). Each color represents a different agroupment. The blue region indicates the two processing stages from DES, where the first eight rounds can be clearly counted. This cluster is picked by selecting the indexes in the smaller one. The final step involves detaching the DES stages by calculating the absolute distance between two subsequent indexes. Then, the first stage is dissociated by selecting the point immediately before that in which the distance is greater than an empirical value, determined by the traces behavior.

III. Results and Discussion

A dataset containing 100,000 electromagnetic traces extracted from a GALS architecture with two stages executing a DES algorithm was submitted to the proposed methodology. Each stage operates with a local clock with frequency randomly defined between 38MHz and 60 MHz [9]. DEMA attack is performed in these traces to verify the vulnerability of the sub-traces resulting from the application.

The targets of the attack are the outputs of the eight Substitution Boxes (SBoxes) in the first round of execution of the DES algorithm contained in the first stage of the GALS architecture. A Sbox is an essential function in the DES, being responsible for concealing the relationship between the cryptographic key and the original message. The DEMA attack analyzes differential traces from each Sbox, where their peak indicates the correct key. The main metric to evaluate the quality of an attack is the minimum number of traces required to unveil the cryptographic key [8]. The lower this number, more efficient is the attack.

Table I presents the attack results, where the values correspond to the necessary number of traces to achieve a successful attack. It is observed that the needed number of traces decreases in every SBox when compared to the attack performed on the traces extracted by the threshold strategy. The reduction varies from 3.4 up to 11.1 times in the number of traces for DEMA attack to converge. It is important to salient that SBoxes 5 and 8, which did not converge to the right key, exhibit a disparity relating to the others, due to problems during the acquisition of the electromagnetic traces [8]. Thus, the average number of traces to recover the cryptographic key only considered the six remaining SBoxes.

As a consequence, it is stated that the FCM generated clusters appropriated to the problem examined. Through an

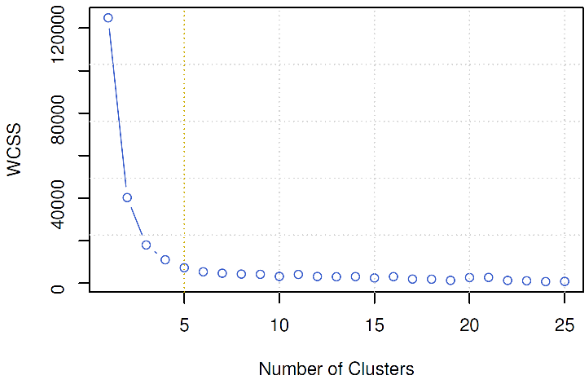
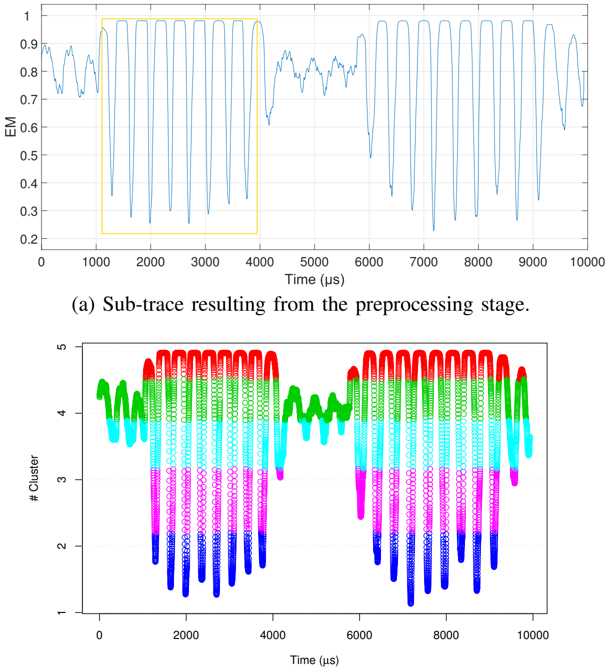


Fig. 2. WCSS value according to the number of clusters.

TABLE I
NUMBER OF TRACES REQUIRED BY S-BOX FOR DEMA ATTACK TO CONVERGE.

DEMA	SBox 1	SBox 2	SBox 3	SBox 4	SBox 5	SBox 6	SBox 7	SBox 8	Mean	SD
Threshold	15026	41236	77111	18337	NC	64164	33944	NC	41736.28	21108.96
FCM	2266	10285	22816	3103	NC	9528	3067	NC	8510.36	7149.56

Note: NC = not converged; SD = standard deviation.



(a) Sub-trace resulting from the preprocessing stage.

(b) Sub-trace grouped with five clusters through FCM.

Fig. 3. Electromagnetic trace from a GALS architecture having two stages executing with local frequency submitted to the proposed methodology.

exploratory data analysis, in order to define the parameters for the preprocessing stage, the presented methodology can be expanded to a different set of traces, whether itself contains power or electromagnetic measurements. The results of the DEMA attack show the influence of the solution by pointing to a cryptographic key retrieve in 4.9 times fewer traces than the threshold approach. Moreover, the extraction of the target information potentially decreases the computation cost for performing the attack.

IV. CONCLUSION

The performance of DEMA and DPA attacks is associated with the time domain alignment of the electromagnetic or power traces acquired. Thus, a countermeasure adopted consists in the insertion of randomness and noise, as well as the asynchronous operation of cryptographic modules. This paper introduces a procedure for detection and extraction of the target sequence from measurements obtained from a cryptosystem having a local clock with frequency randomly defined. The method comprises signal processing techniques and an unsupervised learning algorithm. It allows the identification of interest region leading to a more efficacious attack. The impact of the proposed method is evaluated by performing the DEMA attack. Results demonstrate that the proposed methodology is substantially more effective than a threshold strategy in unveiling the secret key.

As future work, it is intended to verify the effect of other clustering methods, such as *k*-means and *k*-medoids. Besides, the refinement and abstraction of parameters empirically identified can be further explored.

REFERENCES

[1] P. Hodgers, N. Hanley and M. O'Neill, "Pre-processing power traces with a phase-sensitive detector," in IEEE International Symposium on Hardware-Oriented Security and Trust (HOST), Austin, TX, 2013, pp. 131-136.

[2] P. C. Kocher, "Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and other Systems," in Advances in Cryptology CRYPTO 96, Springer, LNCS 1109, pp. 104-113.

[3] P. Kocher, J. Jaffe, and B. Jun, "Differential Power Analysis," in Advances in Cryptology,, LNCS 1666, Springer, 1999, pp. 388-397.

[4] S. Mangard, E. Oswald, and T. Popp, Power Analysis Attacks: Revealing the Secrets of Smart Cards (Advances in Information Security). Secaucus, NJ: Springer-Verlag New York, 2007.

[5] C. H. Gebotys, C. C. Tiu and X. Chen, "A countermeasure for EM attack of a wireless PDA," in International Conference on Information Technology: Coding and Computing (ITCC'05) - Volume II, 2005, pp. 544-549 Vol. 1.

[6] Q. Tian and S. A. Huss, "A General Approach to Power Trace Alignment for the Assessment of Side-Channel Resistance of Hardened Cryptosystems," in 8th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Piraeus, 2012, pp. 465-470.

[7] S. Mangard, "Hardware Countermeasures against DPA: A Statistical Analysis of Their Effectiveness", Topics in Cryptology CT-RSA 2004, pp. 222-235, 2004.

[8] R. Lellis, R. I. Soares and A. Souza, "An energy-based attack flow for temporal misalignment coutermeasures on cryptosystems," in IEEE International Symposium on Circuits and Systems (ISCAS), Baltimore, MD, 2017, pp. 1-4.

[9] R. Soares, N. Calazans, F. Moraes, P. Maurine and L. Torres, "A Robust Architectural Approach for Cryptographic Algorithms Using GALS Pipelines," IEEE Design & Test of Computers, vol. 28, no. 5, pp. 62-71, Sept.-Oct. 2011.

[10] L. Loder, A. de Souza, M. Fay and R. Soares, "Towards a framework to perform DPA attack on GALS pipeline architectures," in 27 Symposium on Integrated Circuits and Systems Design, Aracaju, 2014, pp. 1-7.

[11] D. Márquez, A. Fred, A. Otero, C. García and P. Flix, "Introducing Negative Evidence in Ensemble Clustering Application in Automatic ECG Analysis," in International Workshop on Similarity-Based Pattern Recognition, pp. 54-69, 2015.

[12] S. W. Smith, "The scientist and engineer's guide to digital signal processing", 2nd ed, San Diego: California Technical Publishing, 1999.

[13] J.E. Hanke and D. W. Wichern, "Business Forecasting", 8th ed., New Jersey: Pearson Prentice Hall, 2005.

[14] S. Haykin, "Neural networks and learning machines", 3rd ed., New York: Prentice Hall, 2009, 907 p.

[15] J. C. Dunn, "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters," J. Cybern, vol. 3, no. 3, pp. 3257, 1974.

[16] J. Bezdek, "Objective Function Clustering' in Pattern Recognition with Fuzzy-Objective Function Algorithms, Springer: Boston, pp. 43-93, 1981.

[17] D. Ketchen Jr. and C. Shook, "The application of cluster analysis in strategic management research: an analysis and critique", Strategic Management Journal, vol. 17, no. 6, pp. 441-458, 1996.