Information Leakage Discovery Techniques to Enhance Secure Chip Design

Alessandro Barenghi[1], Gerardo Pelosi[1,2], and Yannick Teglia[3]

[1] DEI – Dipartimento di Elettronica e Informazione
Politecnico di Milano, Via Ponzio 34/5, I-20133 Milano, Italy
{barenghi,pelosi}@elet.polimi.it
[2] DIIMM – Dipartimento di Ingegneria dell'Informazione e Metodi Matematici
Università degli Studi di Bergamo, Viale Marconi 5, I-24044 Dalmine (BG), Italy
gerardo.pelosi@unibg.it
[3] STMicroelectronics, ZI de Rousset BP2, 13106 Rousset Cedex, France
yannick.teglia@st.com

Abstract. Side channel attacks analyzing both power consumption and electromagnetic (EM) radiations are a well known threat to the security of devices dealing with sensitive data. Whilst it is well known that the EM emissions of a chip represent an information leakage stronger than the overall dynamic power consumption, the actual relation between the emissions and the computations is still a subject under exploration. It is important for the chip designer to be able to distinguish which portions of the measured EM emissions are actually correlated with the sensitive information. Our technique obtains a detailed profile of the information leakage, identifying which harmonic components carry the largest part of the it on the measured signals. It may be successfully integrated in a design workflow as a post-testing feedback from the prototype chip, in the form of additional constraints aimed at reducing the local wires congestion up to a point where the emissions are no longer sufficient to conduct an attack. The analysis allows the design of ad-hoc countermeasures (shields and/or EM jammers), which do not require architectural changes to the chip. We provide a validation of the proposed technique on a commercial grade ARM Cortex-M3 based System on Chip (SoC), executing a software implementation of AES-128. The proposed approach is more efficient than a search of the whole frequency spectrum, allowing to conduct a deeper analysis with the same timing constraints.

Keywords: Side-Channel Attacks, Embedded Systems Security, Differential Power Attacks, Differential Electromagnetic Attacks.

1 Introduction

A significant part of the security margin provided by a cryptographic device is represented by its resistance to side channel attacks. Side channel attacks aim at disclosing the secret key of cryptographic primitives, through measuring environmental parameters during their computation. Typical environmental parameters from which it is possible to extract information relative to the values being

C.A. Ardagna and J. Zhou (Eds.): WISTP 2011, LNCS 6633, pp. 128–143, 2011.

processed are: power consumption [6,10], electromagnetic radiation [1,9] and execution timing [5]. Depending on the environmental parameter being measured, the attack techniques are called respectively Differential Electromagnetic Analysis (DEMA), Differential Power Analysis (DPA) and Timing analysis (TA).

A regular Differential Electromagnetic Analysis (DEMA) or a DPA attack aims at modeling the variation of the EM emissions or the dynamic power consumption of a chip caused by the different inputs fed to the cryptographic primitive executed on it. Assuming all the values undergoing computation are known, it is possible to accurately predict the values of the aforementioned environmental parameters, but, since the value of the secret key is not known to the attacker, an alternative strategy is devised. The attacker creates a family of a-priori models, each one depending on a *key hypothesis*, i.e. an hypothetical value of a small part of the key. The most common way of modelling either the power consumption or the EM emissions of a digital circuit is taking into account either the Hamming Weight (HW) of the values being computed or the Hamming Distance between two subsequent values, during a part of the cryptographic primitive involving the secret key. After building the models, the attacker correlates each of the predicted results with each sample in the time series of an experimental measurement (*trace*) of the environmental parameter related to the execution of the targeted part of the cryptographic primitive. Since only one of the models will fit, it is possible for the attacker to deduce the right key hypothesis. In order to obtain a significant estimate of the relation between the models and the physical parameters being measured, a large number of traces are taken while employing different inputs (plaintexts or ciphertexts). Pearson's linear correlation coefficient is the most common figure of merit employed to assess the goodness of fit of the a-priori models against the actual measurements [3]. Pearson's correlation coefficient turns out to be rather effective in practice since the strength of both the EM-emissions and the dynamic power consumption depend linearly on the switching activity of the underlying circuitry [8].

The analysis of EM emissions has proven an effective side channel able to yield efficient attacks [1, 4, 12], although it implies a quadratic increase in the number of chip spots to be considered when compared to a DPA technique. A significant factor for this efficiency is the use of small probes with a consequent precise spatial localization of the sources of the measured signal.

The signals collected through these kinds of measurements have a high correlation with the data computed by the cryptographic primitive operation considered in the a-priori models. For example, in [13] the authors show that recording the emission traces over a particular spot of an FPGA programmed with an implementation of the AES block cipher (identified as the places where the S-Box function was synthesized) resulted in an effective reduction of the number of traces required. Therefore, this kind of enhancements may be expected also when the device under test is either an ASIC implementation or a general purpose CPU running the same cryptographic primitive. Moreover, an important advantage of EM analysis is the possibility to bypass common power analysis

countermeasures, such as voltage pumps, and to ignore the presence of static dummy cycles inserted to rebalance timing issues.

Within this context, we propose an enhancement of the testing methodology for a circuit in order to include evaluation of the resistance to EM side channel analysis as a design step. We propose an information leakage finding algorithm aiming at recognizing which harmonic components of the measured signals actually convey the significant part of the exploitable side channel information. The proposed algorithm is faster than an exhaustive brute force sweeping the whole frequency range, while preserving the same accuracy for real world scenarios. The proposed analysis enables to design countermeasures targeted to the specific leakage pattern of the device and may be conducted also on simulated traces without any change in the procedure. At the best of the authors' knowledge, this is not currently possible due to the lack of publicly available EM emission estimation tools from any pre-prototype description of the chip. The availability of such tools would allow a pre-prototyping evaluation of the EM leakage and would allow to properly tune the post place-and-route procedure in order to mitigate the EM leakage. This may be achieved through proper routing of the wires, which represent the most EM radiating part of the circuit.

The paper is organized as follows: Sect. 2 describes how the proposed methodology integrates within the current chip design workflow. Sect. 3 explains the proposed information leakage finding algorithm and provides insights on its inner workings. Sect. 4 reports a practical validation of the proposed technique on a commercial grade Cortex-M3 SoC running an industrial grade implementation of AES-128. Finally, Sect. 5 presents our conclusions.

2 EMA Analysis as a Design Phase

A typical digital circuit design flow is composed of a fixed chain of stages following the high level specification of the device, in the form of a netlist description of the chip. The first steps in tackling the transformation of a netlist into an accurate blueprint of the chip are: performing a preliminary consumption analysis on the design, and adding the Built In Self Test (BIST) additional logic required to perform functional testing of the circuit. After these steps, the chip description is accurate enough to perform a full placement through the wire routing and clock tree design process. The obtained description is accurate up to a full three dimensional representation of the design at single wire level, stored in the Graphic Database System II (GDSII) common interchange format in order to control the integrated circuit photolithographic etching. After completing the whole design of the chip, a first prototype of the actual device is realized and packaged in order to be sent to the testing stage. The prototype chips are still subject to a series of compliance tests among which the Electromagnetic compatibility (EMC) ones, aimed at ascertaining that the EM radiations of the device are not strong enough to disturb the regular functioning of neighbouring devices. Electromagnetic compatibility tests are oriented to obtain a quantitative measure of the radiated energy, regardless of the information which may

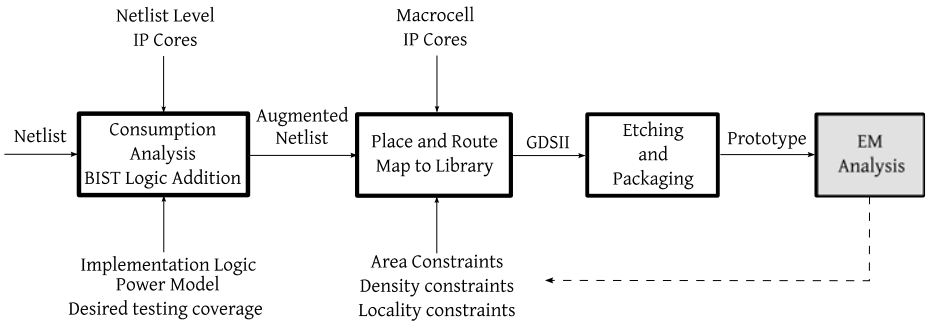


Fig. 1. Description of a typical digital chip design flow. The proposed EM analysis stage is highlighted in gray

be carried by the wavefronts. In particular, since the die emits significantly less than the bonding wires which connect it to the package pins, its emissions are usually regarded as harmless by the common EM compatibility standards. The security testing methodology proposed in this paper can easily be integrated in the EMC compatibility testing phase, since it requires the same equipment to be performed, and does not add a prohibitive amount of time to be spent at the workbench. The target of the EM analysis is to provide a more accurate information on both the spatial location and the informative content of the EM radiation of the silicon die, through checking if there is a viable side channel for attacks. After performing the proposed analysis, it is possible for the designer to employ the gathered information for introducing countermeasures during the place-and-route (p&r) step. In particular, it is possible to either exploit the free space on the top layer of the chip, after all the p&r operations have been performed, to introduce a grounded metallic shield over the most radiative zones or to reroute partially the wires in order to avoid excessive local congestions. A further possible countermeasure is the introduction of a jamming resonator tuned on the frequencies which carry sensitive informations out of the chip. Such a resonator may be easily realised as a simple tuned wire antenna and does not need to interact with other circuits related to the chip. Thus it is possible to design it without having any concerns on the actual chip architecture thus, helping a late-stage introduction, with only a negligible area overhead.

2.1 Electromagnetic Emission Analysis

The first step in the testing methodology is to obtain a map of the intensity of the electromagnetic emissions of the silicon die. In a region of space close to the chip surface, it is possible to model the EM emitting components as a set of wires lying on the die plane. Since the radiated field of a single wire is emitted perpendicularly to direction of the current flow, a probe constituted by a wire coil placed

parallel to the die surface will not be sensitive to crosstalk from nearby wires[1], thus resulting in a reliable measure of the EM field intensity per underlying area unit. As a consequence, the parts of the chip which will be radiating more strongly are the ones characterized by a high wiring density. At the moment, the routing tools are not considering excessive wiring density as a problem because the power estimate is usually done before the wires are placed[2]. Consequently, it is possible to have strongly radiating zones which are not de-congestioned automatically by the tool, resulting in EM radiating hotspots. Round coil probes are already in use to perform EMC testing[3] on packaged chips, and may be used to perform the mapping of the EM emissions as well, thus enabling equipment reuse during the testing methodology. It is possible to obtain a precise mapping of the intensity of the EM emissions of a chip per area unit during the computation of a software cryptographic primitive through recording the field emitted in a spot and repeating the measurements while sweeping with the probe all over the die surface.

A kind of chip areas which may be of particular interest to be mapped are the so-called *glue logic* areas: sections of the chip where the placement tool is allowed complete freedom over the component and wire layout thus possibly causing large wire skeins. Since the recorded emission signals provide the evolution of the field intensity over time, it also is possible for the designer to locate exactly which part is emitting through checking either where the chip was active during a certain time instant or which instruction is being executed on a mapped CPU. This is particularly interesting in order to focus the analysis only on the instructions of the running cryptographic primitive dealing with the secret values thus, avoiding unnecessary concerns about strong EM emissions in unrelated time instants. We point out that the EM testing is performed in a white-box environment, where the designer knows all the implementation details of the chip, including the software running on the general purpose CPU in case the algorithm is not directly implemented as an ASIC. This kind of analysis is the one warranting the strongest security on the final product, since it already assumes that the attacker is able to know all the details of the device he will target, thus considering the position of utmost advantage for him. Indeed, motivated attackers may apply hardware reverse engineering techniques in order to fully reconstruct the structure of a chip[4]. After the strongest emitting spots of the chip have been located, a set of traces T is collected on top of them in order to proceed to the frequency analysis of the EM radiation.

[1] In a circular wire coil, placed parallel to the surface of the chip, the induced voltage drop at the ends of the wire is proportional derivative of the sensed magnetic flux ($\Delta V = -\frac{\mathrm{d}\phi}{\mathrm{d}t}$).

[2] Cadence Design Systems, Inc., *Physical Prototyping–Key to Nanometer SoC Chip Design*, Whitepaper. Dec. 2010, `http://www.cadence.com`

[3] International Electrotechnical Commission, *IEC/TS 61967-3 ed1.0*, ISO-Standard, Dec. 2010, `http://webstore.iec.ch/webstore/webstore.nsf/artnum/035659`

[4] Chipworks, *Report Library & Technical Competitive Analysis*, Technical Report, Dec. 2010, `http://www.chipworks.com/Report_search.aspx`

Algorithm 1. Leaked Information Finding Algorithm

Globals: T: set of traces, b: branching factor,
γ: confidence level of the correlation attack

Input: δ: frequency interval on which traces are evaluated; $|\delta|$ denotes the length of the frequency interval

Output: L: list of pairs $(\delta,\ n_\delta)$, where δ is a frequency interval used to set up a filter for the measured traces, and n_δ is the minimum number of filtered traces so that the correlation attack succeed with the given confidence level. Initially, $L \leftarrow \emptyset$

```
1  begin
2      if (|δ| ≥ η N_threshold) then
3          T_δ ← Filter(δ, T)
4          n_δ ← CorrelationAttack(γ, T_δ)
5          L ← L ∪ { (δ, n_δ) }
6          if (n_δ = ⊥) then
7              return
8          else
9              {δ_0, ..., δ_{b−1}} ← SplitUp(δ)
10             for i from 0 to b−1 do
11                 Call Algorithm 1(δ_i)
```

3 Information Finding Algorithm

In order to automatically determine the harmonic components of the recorded signals that actually carry the exploitable information, we devised an information finding method reported in Alg. 1. This computation is intended to improve the information leakage characterization by lowering the ratio between the energies $\mathcal{E}_f$ and $\mathcal{E}_t$ of the filtered and unfiltered version of any trace, respectively. The output of the algorithm provides all the information required to build an optimum multi-bandpass filter in order to both maximize the aforementioned energy ratio, and discard the signal components not related to the key-dependent computation. The effect on leakage estimation efficiency and precision is measured through considering the decrease in the minimum number of traces needed to carry out a successful attack with a reasonable confidence margin. The key idea is to split the spectrum in equally sized shares and filter the EM traces with a Finite Impulse Response (FIR) filter whose bandpass keeps only one share at a time. The shape of the filtering window is driven by the necessity of having a maximum flat bandpass while retaining moderate aliasing in the time domain and a reasonable roll-off in the frequency domain. Reasonable choices are either a Chebyshev (type II) window or a tapered cosine window [11]. Subsequently, a series of attacks on the filtered traces is performed to understand which is the minimum number of measurements needed to distinguish, with a reasonable statistic confidence, the correct key hypothesis from the other ones. In order to perform a computationally efficient search in the frequency space, Alg. 1 exploits a b-ary split search strategy. The algorithm employs the previously

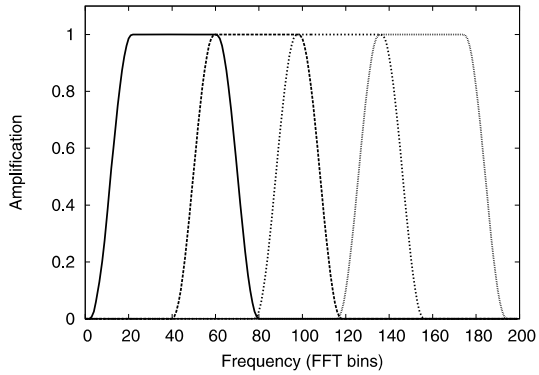


Fig. 2. Overlapping windows generated by the search algorithm. This example depicts the amplitude of four tapered cosine windows, as it could happen on the first run of the algorithm with $b = 4$.

collected trace set T as a test bench for the efficiency of the attacks. Let N be the number of samples in the time series of each trace or, equivalently, the number of *bins* of the Fast Fourier Transform (FFT) of the trace, and η the frequency interval corresponding to a single bin (η=2 f_{Ny}/N). The frequency interval below which considering the energy of the signal is still appropriate may be defined as $\eta\ N_{threshold}$, $N_{threshold} > 1$, therefore the FFT bins of each trace may be thought as a sequence of B=$N/N_{threshold}$ slices. The algorithm receives the frequency interval δ, on which traces are evaluated, as a parameter. If the length of the frequency interval in input, $|\delta|$, is greater or equal than $\eta\ N_{threshold}$, a digital filter with such a support is applied to each trace in the set T, thus obtaining a new set of filtered traces T_δ (lines 1–3). Given the set T_δ, a correlation attack on the filtered traces is carried out in order to obtain either the minimum number of traces n_δ (which enable to recover the secret key) or a null value $\perp$ (in case of a failure of the attack) (line 4). Subsequently, the pair (δ, n_δ) is inserted in a global list L in order to record the shortest frequency intervals where the correlation attack either succeeded or failed (line 5). In case the figure of merit n_δ is different from $\perp$ (line 8), the interval δ is split up in b shares, with a 50% mutual overlap, (as depicted in Fig. 2) and the same procedure is recursively called on each share (lines 8–11), otherwise the algorithm returns from the call.

At the end of the execution, the list L will contain both the shortest frequency intervals for which the correlation attack is able to recover the secret key and the largest frequency intervals for which the measured traces T do not provide enough information to retrieve the key. The use of a larger trace set may lead the algorithm to spot more leaking intervals than the ones obtained with less traces at the cost of a longer computation. After obtaining the output of the algorithm, the designer is able to exploit the information to design an ad-hoc filter which will remove all the harmonic components not containing any relevant information. In order to design the filter, the designer may choose to keep all the parts of the spectrum where the attack has succeeded with the number of traces at his

disposal. A more restrictive choice is to keep only the harmonic components where the attack succeeds with a number of traces smaller than the one needed for an attack with unfiltered traces. The rationale behind this choice is the fact that the stricter filtering will yield a higher ratio between the energy of the filtered signal $\mathcal{E}_f$ and the total energy of the signal $\mathcal{E}_t$, while retaining most of the informative content. On the other hand, discarding harmonic components which are still carrying some information, although more polluted than the original unfiltered signal, may be detrimental to the analysis, in case the leakage is not concentrated in a precise number of slices. In the following sections we will refer to the former spectrum slices as the *good* ones, while the latter will be indicated as the *acceptable* ones.

3.1 Complexity Analysis

Let the running time of the algorithm be expressed as the number of FILTER, and CORRELATIONATTACK operations being executed. Assuming the shortest possible frequency interval ($\eta\ N_{threshold}$) for the application of every digital filter and correlation attack, a linear scan of the spectrum $[0, B\ \eta\ N_{threshold}]$ would have a temporal complexity equal to $\Theta(B)$. It is possible to obtain a significant reduction of the computational effort needed to detect the leaking components of the signal through exploiting their sparsity and clustering over the whole spectrum, since real world scenarios commonly exhibits such a behaviour.

The best case of Alg. 1 happens when the useful information in each measured trace is concentrated in at most 1 out of b frequency sub-intervals for each call of the recursive procedure, thus giving a computational complexity of $\mathcal{T}(B)=\Omega(b \log_b B)$.

The worst case of Alg. 1 gives an upper bound to the temporal complexity $\mathcal{T}(B)$ and corresponds to a balanced configuration of the b-tree where at leaves level each group of sub-intervals has at least a slice where the attack succeeds. We note that the worst case condition implies the information is uniformly spread on the entire spectrum. Although this scenario is highly unlikely, it is possible to mitigate the additional computational complexity that a tree-based search algorithm would imply in such a case. A sensible trade-off is to modify the algorithm, so that the execution will halt in case the slices reach a size of $b\ N_{threshold}$, while all the attacks are still successful. This yields a temporal complexity proportional to $\mathcal{T}(B)=O(\sum_{i=0}^{\lceil \log_b B \rceil - 1} b^i)=O(\frac{B-1}{b-1})$ at the cost of a reduction of a factor b in the precision of the analysis, while in turn avoiding the only case where the algorithm is slower than a linear scan.

The average-case running time requires a more accurate analysis. Intuitively, the information is very clustered over the entire frequency domain, thus the average running-time is expected to be much closer to the best case than to the worst case for most part of the practical cases. Let $p(j)$ the probability of mounting a successful CORRELATIONATTACK when considering the harmonic components of the traces in T on a single slice of the frequency spectrum:

$[\ j\ \eta\ N_{threshold},\ (j+1)\ \eta\ N_{threshold}\ )$, where $j \in \{0,\ \ldots,\ B-1\}$, consequently the probability that the attack does not succeed is: $1 - p(j)$.

Let X_j be the indicator random variable associated with such an event. Hence, $X_j=1$ if the CorrelationAttack is successful through filtering the traces on the interval $[\ j\ \eta\ N_{threshold},\ (j+1)\ \eta\ N_{threshold}\ )$, $j \in \{0,\ \ldots,\ B-1\}$. Assume each call of Alg. 1 to be bound to a b-tree node, corresponding to a determined share of the frequency spectrum B. Denote with $X_{j_t}^{[h-t]}$ the indicator random variable associated with the event of executing a successful CorrelationAttack when considering the harmonic components of the traces in T on the frequency interval: $[\ j_t\ b^{h-t}\ \eta\ N_{threshold},\ (j_t+1)\ b^{h-t}\ \eta\ N_{threshold}\ )$, where $j_t \in \{0,\ \ldots,\ B/b^{h-t}-1\}$, $t \in \{0,\ \ldots,\ h\}$, $h=\lceil \log_b B \rceil$. Therefore, $X_{j_h}^{[h]},\ X_{j_{h-1}}^{[h-1]},\ \ldots,\ X_{j_0}^{[0]}$ may denote the random variables (from the leaf level to the root level) associated with each node of the aforementioned b-tree, respecting the following relations:

$$
\begin{aligned}
\Pr(X_{j_h}^{[h]} = 1) &= p(j_h); && j_h \in \{0, \ldots, B-1\} \\
\Pr(X_{j_{h-1}}^{[h-1]} = 1) &= \sum_{j_h = j_{h-1} \cdot b}^{j_{h-1} \cdot b + b - 1} \Pr(X_{j_h}^{[h]} = 1); && j_{h-1} \in \{0, \ldots, \tfrac{B}{b} - 1\} \\
\Pr(X_{j_{h-2}}^{[h-2]} = 1) &= \sum_{j_{h-1} = j_{h-2} \cdot b}^{j_{h-2} \cdot b + b - 1} \Pr(X_{j_{h-1}}^{[h-1]} = 1); && j_{h-2} \in \{0, \ldots, \tfrac{B}{b^2} - 1\} \\
&\ldots && \ldots \\
\Pr(X_{j_1}^{[1]} = 1) &= \sum_{j_2 = j_1 \cdot b}^{j_1 \cdot b + b - 1} \Pr(X_{j_2}^{[2]} = 1); && j_1 \in \{0, \ldots, \tfrac{B}{b^{h-1}} - 1\} \\
\Pr(X_{j_0}^{[0]} = 1) &= \sum_{j_1 = j_0 \cdot b}^{j_0 \cdot b + b - 1} \Pr(X_{j_1}^{[1]} = 1); && j_0 \in \{0\}
\end{aligned}
$$

The probability density function $p(0), p(1), \ldots$ in the above formula must be either estimated or modelled taking into account (i) the specific operation of the targeted cryptographic primitive, (ii) the hardware design of the target device and (iii) the physical characteristics of the environmental parameter measured by the attack. In common practical cases such as the analysis of EM emissions, the harmonic components carrying information are usually restricted to a relatively small bandwidth, since it is reasonable to assume that the resonating conductors will have reasonably close impedences. The same narrow band consideration may be made for power consumption measurements, since a synchronous circuit dissipates most of the dynamic power at each clock edge, thus resulting in a large part of the informative signal being concentrated on the same frequencies. Therefore, the probability density function $p(0), p(1), \ldots$ it is expected to be highly clustered (i.e., there are only a few $p(k) \neq 0$, $k \in \{0, \ldots, B-1\}$). Such a probability density function results in a low number of branches requiring a full depth exploration. Thus the proposed algorithm is faster than a linear scan in the best and average case and as fast as the linear scan in the worst case.

4 Experimental Validation

4.1 Workbench

The device under exam was a commercial grade Cortex-M3 based SoC[5], endowed with on die SRAM and Flash memory, both coupled to the CPU, and USB, RS-232 and GPIO interfaces. The Cortex macrocell is synthesized together with 10+ other IP cores in a single block of glue logic, thus it is not possible to identify through optical inspection any of the components, nor to infer the placement of any of IP cores through looking at die surface with an optical microscope. In order to get as close as possible to the die surface during the measurements, the top of the chip package was removed through a combination of nitric and sulphuric acid. The device under profiling was mounted on a regular development board and affixed to a gas suspended X–Y moving table controlled by the same computer gathering the data from the oscilloscope. We chose to map the chip area by moving in 100μm steps and covering the whole zone to be mapped scanning it line by line. The equipment employed to collect all the measures was a LeCroy WavePro 7300a digital oscilloscope[6] sampling at $f_{s,DEMA}$=10Gsample/s (thus resulting in a Nyquist limit for the sampled components at $f_{Ny}=f_s/2$=5GHz), and an EM-profiling oriented Langer ICR H probe[7] made of an horizontally oriented coil with an inner diameter of 150μm, placed roughly at 0.8mm from the die surface. The signal picked up by the probe was amplified by a low noise differential amplifier and fed directly into the oscilloscope sampling channel. We targeted the `load` operation executed for the first look up in the S-Box. The start of the acquisition was triggered by the device under test through the use of a GPIO pin asserted by the enciphering program before the start of the execution of the first AES-128 round. All the recorded traces was obtained through averaging 16 measurements taken with the same settings and the same plaintext in order to reduce the environmental noise.

In order to provide comparative results on Alg. 1, we also conducted a power consumption measurements campaign. The measurements were collected with a LeCroy Waverunner WavePro 7100A with a maximum sampling rate of 20Gsamples/s with a LeCroy AP034 differential probe[8] connected to a 2Ω shunt inserted on the only power supply line available for the Cortex-M3 SoC. The sampling rate was set to $f_{s,DPA}$=5Gs/sec (f_{Ny}=2.5GHz) in order to provide a sound safety margin on the sampling of fast dynamics which may be useful during the analysis. All the signals measured are recorded in an 8-bit per sample raw format and all signal treatment on a Core i7 920 running 64-bit Linux Gentoo 2010.1.

[5] ARM, *Cortex-M3 Processor*, Technical Specifications, Dec. 2010, `http://www.arm.com/products/processors/cortex-m/cortex-m3.php`

[6] LeCroy, *WavePro 7000 Series*, Technical Specifications, Dec. 2010, `http://www.lecroy.com/france/press/articles/images/WavePro7000_DS.pdf`

[7] Langer EMV-Technik, *IC Test Systems–Near Field Microprobes (ICR probes)*, Technical Specifications, Dec 2010, `http://www.langer-emv.de/en/products/ic-measurement/`

[8] LeCroy, *AP034 Differential Probe*, Technical Specifications, Dec. 2010, `http://www.lecroy.com/Options/`

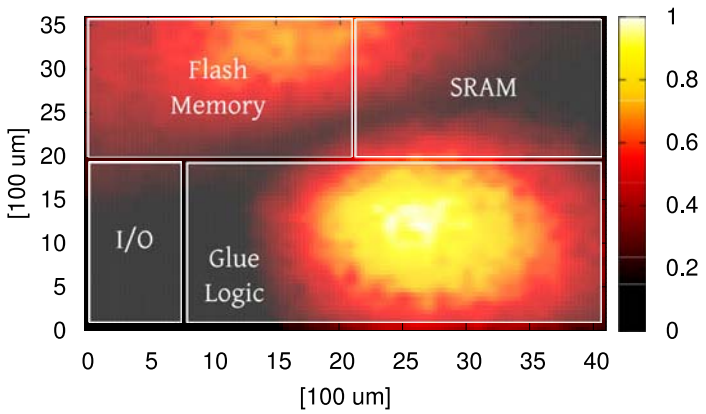


Fig. 3. Intensity of the Cortex-M3 SoC EM emissions during the time lapse when the first round key addition of AES-128 is performed. Clearer zones represent a stronger EM activity. The magnitude of the current measured by the probe has been normalized for visual enhancement.

4.2 Experimental Results

Employing the described setup, the silicon die of the device under attack was fully mapped in order to determine which components were emitting and which parts of the device logic were most active during the computation. The collected traces were processed in order to obtain a temporal sequence of maps of the emissions of the chip through adding the values of the emitted signal for 50 consecutive samples at once. The result of this preprocessing was a movie depicting the evolution of the emission during the whole running time of the AES-128, with a time accuracy of 5ns per frame. Through the knowledge of the code running on the chip, and thanks to the synchronization provided by the trigger raised by the Cortex-M3 board, it was possible to locate when the CPU was doing active computation at the beginning of the first AES round thus, obtaining the frame depicted in Fig. 3. Figure 3 depicts the amount of emitted EM radiation, measured as the intensity of the current running through the probe, where clearer colours indicate a higher EM activity. The overlain boxes point out which areas of the chip are optically recognizable. Through examining the map, it is possible to distinguish which zone of the glue logic is occupied by the Cortex-M3 core, thanks to the higher radiation caused by the ongoing switching activity. The second zone having non negligible radiating activity is the flash memory: this activity can be ascribed to the ongoing instruction fetch operation, performed in pipeline with the CPU execution phase. A possible cause of the lower radiation activity shown by the flash memory is the metal tiling added at manufacturing time to flatten the photolithographic layers which partially shields the emissions. Nonetheless, the memory electromagnetic activity is still strong enough to be measured by the probe. Placing the probe directly above the center of the hotspot (at the bottom right of the map) we collected 1100 traces of the EM emissions of the chip during

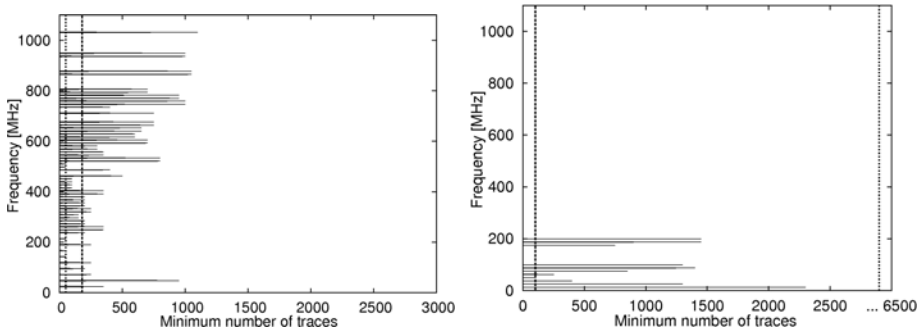


Fig. 4. Minimum number of traces necessary to perform an attack in Alg. 1 considering a single slice. Figure (a) depicts the results for the EM traceset, while Fig.(b) for the power consumption traceset. The leftmost dashed line indicates the least number of traces employed per slice, the rightmost the number of traces necessary with the unfiltered traces. The omitted part of the spectrum has no slices where the attack succeeds.

the execution of the AES-128 algorithm employed in the previous phase, while changing the input plaintexts. Each of the 1100 measurements is obtained as the average of 16 measurements taken with the same plaintext in input to the cipher. In order to perform efficiency analyses also on the power consumption traces, 10000 traces of the power consumption of the same chip were collected. Each recorded trace was the result of the averaging of 64 executions of the first round of AES-128 with the same plaintext.

During the execution of Alg. 1, all the attacks were performed considering the Hamming weight of the output of one byte of the S-Box as the emission intensity model of the observable value. The branching factor of the b-tree was set to b=20 and the branching depth employed was 2. The precision achieved for the EM spectrum leakage detection was $N_{threshold,DEMA} \cdot \eta_{DEMA}$=80·1.25MHz=100MHz, while for the power spectrum was $N_{threshold,DPA} \cdot \eta_{DPA}$=250·0.2MHz=50MHz. Figure 4(a) shows the minimum number of traces required to successfully perform an attack on a specific slice of the spectrum (indicated on the y-axis). For the sake of clarity, all the slices where the attack does not succeed have been represented having 0 as the minimum number of traces (instead of the maximum number available for each traceset). The two vertical dashed lines, for each picture, represent the minimum number of traces required to perform an attack keeping only a single slice of the spectrum (leftmost line) and the number of traces required to perform an attack with the unfiltered traceset (rightmost line). Both figures show that the effective part of the spectrum carrying information is rather small and, in particular, concentrated towards the low end of the spectrum. In particular, the power traces (Fig. 4(b)) show only two zones containing significant information for the analysis, while the EM spectrum (Fig. 4(a)) is richer in terms of leaking components. We note that many components of the EM spectrum contain more

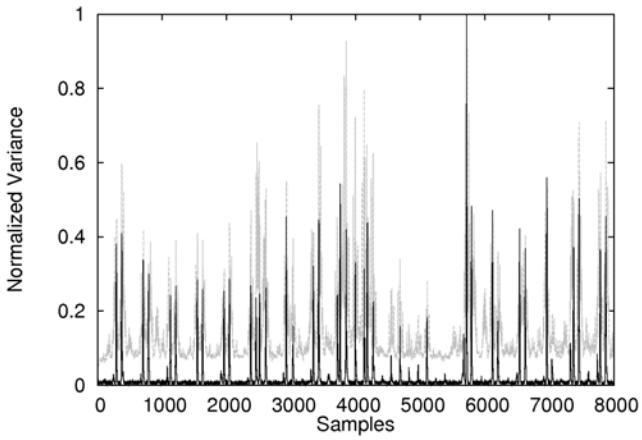
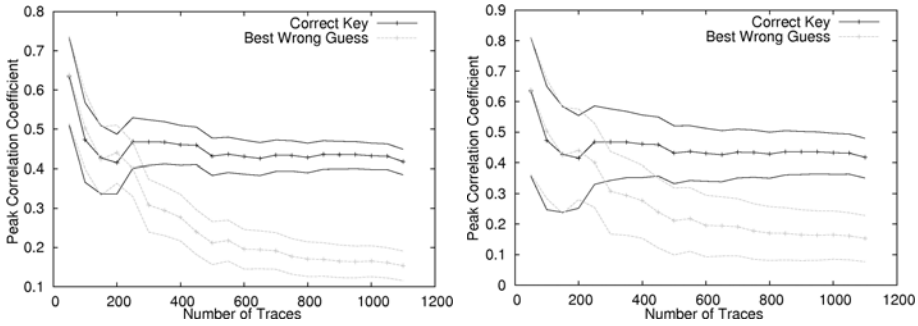


Fig. 5. Time-wise variance of the whole traceset for EM emission before (grey) and after (black) the filtering with the filter encompassing both *good* and *acceptable* slices



(a) (b)

Fig. 6. Correlation analysis of filtered EM traces assuming a confidence level γ=80% (a), and γ=99% (b)

noise than information with respect to the unfiltered traces. Indeed, a significant number of attacks succeed with a greater number of traces than the one needed for the unfiltered ones (this is indicated by the horizontal bars exceeding the second dashed threshold).

Figure 5 depicts the variance of the whole EM trace set, computed sample-wise. The black plot depicts the variance of the traces where a filter keeping both good and acceptable slices has been applied, while the grey plot depicts the variance of the unfiltered traces. It can be easily seen that evicting all the frequency components found to be unrelated by the Alg. 1 yields a time series where it is easy to spot which time instants have a large variation among different inputs, and are thus expected to leak a significant amount of information. This allows the designer to spot the other temporal locations where the information is leaked by the same physical location of the chip. After obtaining the two filters,

Table 1. Comparison among the efficiency of the filter construction technique on EM and power consumption traces

	Statistical Confidence	No Filtering	Good Slices		Accept.+Good Slices	
		Min. Num. of Traces	Min. Num. of Traces	$\mathcal{E}_f/\mathcal{E}_t$	Min. Num. of Traces	$\mathcal{E}_f/\mathcal{E}_t$
DEMA	80%	310	240	-63.8 dB	210	-61.6 dB
	99%	660	400		410	
DPA	80%	5800	300	-62.4 dB	450	-55.8 dB
	99%	5800	800		800	

one containing only the good slices, the other containing both the good and the acceptable ones, we compared the attacks run with both the filtered and the raw traces. In order to obtain a robust evaluation of the precision of the characterization performed, we chose *the minimum number of traces* necessary for the attack to succeed as a figure of merit. Since this figure is dependent on a point estimate of a statistical value (Pearson's coefficient) it is important to take into account its level of confidence in order to properly evaluate the results. This, in turn, implies that, instead of comparing the correlation coefficient of the correct key with the best guess among the wrong ones (i.e., the most likely error for an attacker), we checked when the two confidence intervals for the two values are disjoint. This occurs when the value of the estimate of the correlation coefficient for the correct key is above the one for the best mistake an attacker will make, with a statistical significance (80% and 99%, respectively, in Tab. 1) given by the width of the confidence interval. An example of the trend of the correlation coefficient when employing a growing number of traces is presented in Fig. 6. The figure depicts how the estimate of the correlation coefficients of filtered traces stabilizes with respect to the number of traces employed to perform the attack: the value of the figure of merit can be directly read (240 traces at 80% confidence level (see Fig. 6(a)) and 410 at 99% confidence level (see Fig. 6(b)). The correlation values for the DPA attacks follow the same pattern, except for the required number of traces which is higher. This may be attributed to the large amount of uncorrelated power consumption which happens on the SoC during the computation.

Table 1 reports the quantitative improvements obtained through the application of the profiling technique to both DEMA and DPA. The first row of the table shows how employing the automatically designed filters improves the efficiency of the attack on a set of measurements. The eviction of the part of the frequency spectrum unrelated to the observed value reduces the number of measurements needed to detect the leakage by 22.5% employing only dense zones and by 32% employing also the sparse ones thus, enhancing the quality of the analysis of the radiated emissions. The quantity of noise removed by the filtering is particularly relevant: the ratio between the energy of the filtered signal $\mathcal{E}_f$ and the one of the raw acquisition $\mathcal{E}_t$ is in the -60 dB range, implying that only a

small part of the radiated emission is actually correlated with the critical computation. Nonetheless, EM attacks are still able to succeed, if more resources are devoted to take a large number of measurements from the chip. The second row of the table reports the gains when the automatic filtering design methodology is applied to power traces. The results suggest that also the correlation analysis on the power consumption signals benefits from employing proper filtering on the measured signals. The number of measurements is reduced by an order of magnitude, coherently with the fact that the power traces, taking into account the consumption of the whole chip, are expected to contain more content unrelated to the attack. In both cases, the enhancement in the efficiency allows the designer to take into account the real entity of the threat, which would have been masked by the environmental and systematic noise. One particular effect is that, while DPA attacks benefit from employing only harmonic components where the obtained information is dense, DEMA attacks perform better when including also sparse ones. This may be ascribed to the fact that the leakage in power consumption is concentrated in a few harmonic components [2].

The running time of the algorithm was sensibly lower than the linear scanning of the spectrum for both cases. In particular, the analysis of the EM traces required only 20 attacks at the first level and 40 at the second level of the b-tree (thus 60 calls to CorrelationAttack instead of 400), while the analysis of the power traces required 20 attacks at the first level and 20 at the second (40 calls versus the 400 needed for a linear scan). Taking into account the time for a single attack the overall running time of Alg. 1 was of 2.5 hours (against a 16.6 for a full linear scan) for the power consumption profiling and of 19 minutes for the EM profiling (against a 3 hours and 10 minutes long linear scan).

5 Conclusion

In this work we proposed a new technique able to obtain a characterization of the information of the EM leakage and demonstrated its viability employing an ARM Cortex-M3 chip running an implementation of AES-128. The proposed algorithm is able to obtain a precise characterization of the harmonic components of the side channel measurements (up to a 1/400th of the measured bandwidth in our experiments), within an acceptable time frame on a single desktop. We note that the information obtained from the spatial and frequency profiling of the EM traces allows the designer to introduce ad-hoc countermeasures to the information leakage. This results either in savings in terms of shielded area or in the introduction of non-architecturally invasive active countermeasures into the chip to selectively choke up the EM emitted information. As future developments for reinforcing the security of cryptographic devices, through employing signal processing techniques, we plan to investigate topics about blind source separation (BSS) methods [7]. After a proper mapping of the targeted device, these techniques provide an interesting tool to separate the signal components bound to the cryptographic primitive computation from the signals emitetted from other active parts of the device.

References

1. Agrawal, D., Archambeault, B., Rao, J.R., Rohatgi, P.: The EM side-channel(s). In: Kaliski Jr., B.S., Koç, Ç.K., Paar, C. (eds.) CHES 2002. LNCS, vol. 2523, pp. 29–45. Springer, Heidelberg (2003)
2. Barenghi, A., Pelosi, G., Teglia, Y.: Improving First Order Differential Power Attacks Through Digital Signal Processing. In: Elçi, A., Makarevich, O.B., Orgun, M.A., Chefranov, A., Pieprzyk, J., Bryukhomitsky, Y.A., Örs, S.B. (eds.) SIN, pp. 19–29. ACM, New York (2010)
3. Brier, E., 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)
4. Gebotys, C.H., White, B.A.: EM analysis of a Wireless Java-based PDA. ACM Trans. Embedded Comput. Syst. 7(4) (2008)
5. Kocher, P.C.: Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems. In: Koblitz, N. (ed.) CRYPTO 1996. LNCS, vol. 1109, pp. 104–113. Springer, Heidelberg (1996)
6. 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)
7. Le, T.H., Clediere, J., Serviere, C., Lacoume, J.L.: How can Signal Processing benefit Side Channel Attacks? In: Workshop on Signal Processing Applications for Public Security and Forensics 2007. SAFE 2007, pp. 1–7. IEEE, Los Alamitos (2007)
8. Mangard, S., Oswald, E., Popp, T.: Power Analysis Attacks: Revealing the Secrets of Smart Cards (Advances in Information Security). Springer-Verlag New York, Inc., Secaucus (2007)
9. Mangard, S.: Attacks on Cryptographic ICs Based on Radiated Emissions. In: Proceedings of Austrochip, pp. 13–16 (2003)
10. Messerges, T.S., Dabbish, E.A., Sloan, R.H.: Investigations of Power Analysis Attacks on Smartcards. In: WOST 1999: Proceedings of the USENIX Workshop on Smartcard Technology, p. 17. USENIX Association, Berkeley (1999)
11. Oppenheim, A.V., Schafer, R.W., Buck, J.R.: Discrete-Time Signal Processing, 2nd edn. Prentice-Hall, Englewood Cliffs (1999)
12. Peeters, E., Standaert, F.X., Quisquater, J.J.: Power and Electromagnetic Analysis: Improved Model, Consequences and Comparisons. Integr. VLSI J. 40(1), 52–60 (2007)
13. Réal, D., Valette, F., Drissi, M.: Enhancing Correlation Electromagnetic Attack Using Planar Near-Field Cartography. In: DATE, pp. 628–633. IEEE, Los Alamitos (2009)