Current Mask generation: an analogical circuit to thwart DPA attacks

Daniel Mesquita[1]♦, Jean-Denis Techer[1], Lionel Torres[1], Michel Robert[1], Gaston Cambon[1], Gilles Sassatelli, Fernando Moraes[2]

[1]*LIRMM – Université Montpellier II – France*
{mesquita, techer, torres, robert, cambon}@lirmm.fr

[2]*PUCRS – Porto Alegre - Brazil*
moraes@inf.pucrs.br

Abstract

This work addresses the leakage information problem concerning cryptographic circuits. Physical implementations of cryptographic algorithms may let escape some side channel information, like electromagnetic emanations, temperature, computing time, and power consumption. With these information, an attacker can retrieve the data that is being computed, like cryptographic keys. This paper proposes a novel method to thwart DPA attacks, based on power consumption control. As main goal, this approach not requires any modification on the cryptographic algorithm, the messages or keys.

1. Introduction

The main objective of cryptographic systems is to allow the communication between two agents, among an insecure channel, with privacy. To accomplish this task, modern cryptographic algorithms uses complex mathematical functions and large keys. In this context, "large" means a number sequence with a range between 128 and 2048 bits.

Cryptographic algorithms commonly are classified in two categories: symmetric and asymmetric. The symmetric ones use the same key to encrypt and to decrypt messages. That supposes a secure channel to accomplish the key exchange, but secret key based algorithms are very performing. On the other hand, asymmetric crypto algorithms uses a pair of keys, mathematically dependent, where one key remains secret, and the other must be published. This kind of algorithms can be used to perform digital signatures and authentication schemes. However, public key algorithms are less performing that secret key ones.

Actually, the two classes of algorithms are commonly combined. With a public key algorithm a secure channel can be established. First of all, the users have they origin ascertained with the authentication protocol. Then, they can exchange the symmetric algorithm's secret key, by encrypting it with the asymmetric algorithm. So, the users can communicate securely.

This idea can be applied to a cellular-to-cellular communication, to a web based video conference, and many other context. Among these, a very growing trend is the embedded crypto system, like smartcards to ID or credit cards. For instance, in France, each credit card has a memory and a crypto processor. This secure device runs the RSA [1] (asymmetric) and the 3-DES[1] [2] (symmetric) algorithms. Nowadays there is 45 million of this kind of credit cards, and in the next years, secure smartcards can become an European standard [3].

RSA and AES are crypto algorithms that are proven as being mathematically robust under some conditions. However, the weaknesses of such algorithms are frequently based on implementation problems. Factors like bad random number generation and others can compromise the whole system security. Concerning hardware implementations, even a careful designer cannot avoid a specific class of cryptanalysis.

The hardware devices implementing cryptographic algorithms (processors, ASIC, FPGA and others), may leak some information, like electromagnetic emanations, computing time and power consumption. By analyzing one or more of these information, an attacker can relate the leaked data with the device's internal state, and so, with the secret key. This kind of attack is called Side Channel Attack (SCA).

The SCA most famous is the Differential Power Analysis (DPA) [4]. The DPA attack is very efficient and relatively low cost. Power analysis principle is based on the current consumption to compute logical 0 (zeros) and logical 1 (ones), that is different for each case. Differential Power Analysis enables an intruder to extract secret keys and information from smartcards, which can be used to create fraudulent transactions, generate counterfeit digital cash or perform content piracy. DPA eavesdrops on the fluctuating electrical power consumption of the microprocessors at the heart of these devices, and uses advanced statistical methods

♦ This work has been partially supported by the brazilian agence CAPES (Project Nº 0276-02/2)

[1] In the next months, all CB will be substituted by CBs with the AES algorithm

to extract cryptographic keys and other secrets. Although DPA attacks currently require a high level of technical skill in several fields to implement, they can be repeated using a few thousand dollars worth of standard equipment, and can often break a device in a few minutes.

After a while, some efficient algorithmic countermeasures have been presented, but most of them rely on the modification at the algorithm level, to avoid the correlation between the power consumption, the message and key data. Our original approach simplifies this task by masking power consumption, without any algorithmic modification.

This paper is organized as follows: Section 2 describes the DPA attack. Section 3 shows previous and related works on DPA countermeasures. Section 4 presents the new method to avoid DPA attacks, and conclusions are discussed and future works shown in Section 5.

2. DPA Attack

DPA attacks use statistical techniques to determine secret keys from complex, noisy power consumption measurements [4]. For a typical attack, an adversary repeatedly samples the target device's power consumption through each of several thousand cryptographic computations with the same key. These power traces can be collected using high-speed analogical-to-digital converters, using digital storage oscilloscopes. Figure 1 illustrates this method.

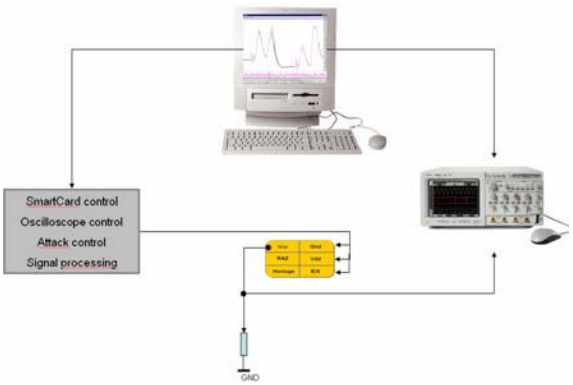

Figure 1 - A DPA attack platform

Because it's widespread use, the DES algorithm is used to explain a DPA attack. DES executes in 16 steps, called rounds. In each round, a transformation F is performed on 32 bits. This F function uses eight non-linear transformations from 6 bits to 4 bits. Each of such transformations is called S-Box.

First, it is needed to make some measures (1000 samples, for instance) from the first (or the last) round of DES computation. After that the 1000 curves are stocked, and an average curve (AC) is calculated.

Secondly, the first output bit (b) of the attacked S-box is observed. This b bit depends only of the 6 bits from the secret key. Then, the attacker can make an hypothesis on the involved bits. He computes the expected values for b; this enables to separate the 1000 inputs into two categories: those giving $b=0$ and those giving $b=1$.

Thirdly, the attacker computes the average curve AV' corresponding to inputs of the first category. If AV' and AV have a difference much greater than the standard deviation of the measured noise), it means that the chosen values for the 6 key bits are correct. But, if AV' and AV do not show any visible difference, the second step must be repeated with another hypothesis for the 6 key bits.

Afterwards, the second and third steps must be repeated with a target bit b in the second S-box, then in the third, and so on, until the eight S-Box. As a result, the attacker can obtain the 48 bits of the secret key. Finally, the remaining 8 bits can be retrieved by exhaustive search.

More details of DPA attacks against DES can be found in the reference [6].

3. Related work

The countermeasures that have been developed against DPA attacks until now can be classified in two families. The first group is composed by the algorithmic countermeasures. The basic idea from references [5], [6], [7] and [8] is to randomize the intermediate results that are produced during the computation of a cryptographic algorithm. Classical DPA attacks are impracticable if these countermeasures are well implemented. But these randomizations are quite expensive to implement for non-linear operations as they are used in algorithms like DES and AES. Furthermore, the algorithmic approach does not provide sufficient protection against high-order DPA attacks [19]. As consequence, this kind of method needs complementary hardware countermeasures.

The hardware method to counteract DPA attacks differs expressively from the algorithmic one. For the hardware approach the intermediate results of the cryptographic algorithm computation are not affected. As an alternative, the contribution of the hardware approach is to hide the attackable part of the power consumption with different noises. The noise addition has a direct relation with the needs of measurement. It does not avoid DPA attacks, but makes it quite more difficult. The effectiveness of the countermeasures against DPA is due to the fact that cryptographic devices are typically protected by a combination of algorithmic and hardware techniques, or only the hardware one [9].

In order to decrease the correlation between data inputs and the power consumption of a given circuit, we must be able to increase the samples needed in DPA. Two major hardware countermeasures in this sense have been proposed. The first one concerns the reduction of the signal-to-noise ratio (SNR). For definition of SNR we call I_c the current consumption of the attacked circuit at a given moment t. In is the current noise caused by the hardware countermeasure. So, the current consumption

can be written as $I_{total,\ t} = I_c + I_n$. The k variable is the signal attenuation caused by the I_n current. The SNR definition can be viewed in Equation (1).

$$SNR = 20 \times \log\left(\frac{I_c}{kR}\right) \quad (1)$$

The lower SNR is, the lower is the correlation between the correct hypothetical current consumption and the real power consumption of the device. To reduce SNR there are some works that use special logic to minimize the data dependency of the current consumption.

In [10] and [11] the balanced dual-rail logic is proposed. The basic idea is that a logic gate must consume an equivalent power, independently from the incoming input values. The SNR is reduced by this data-independent switching of the standard cells. Unfortunately, the experiments show that this goal is only partially reached. Dual-rail approach is not sufficient to guarantee a complete data independent power signature. One potential problem is that the gate loads may differ due to differences in routing. The design of each dual-rail gate must ensure equal input pin loads and balanced power usage. To achieve this, the process of grouping cells in the placement must be done carefully, which implies a high development effort. Besides that, the final circuit with dual-rail logic takes about tree times the area and two times the consumption of the original circuit.

The second hardware approach to prevent DPA attacks is to reduce the correlation between input data and power consumption by randomly disarrange the moment of time at which the attacked intermediate result is computed. If the time tc is different in every power trace, the correlation between the hypothetical power consumption and the real one is highly reduced. The countermeasure proposed by [12] lies on the insertion of random delays. The method described in [9] counteracts the DPA by using power-managed blocks to mask the power consumption. Both approaches, with the [13] and [14] works, difficult the DPA attack. But, as shown in [15], even if a direct calculation of the maximum probability of a given power consumption occurring at a given time is not practical, it is always possible to approximate it empirically based on a software model of the countermeasure.

This work gives a trend to mask the power consumption not by randomizing the consumption or creating noise but by generating, at the transistor level, a constant consumption. It is a little similar with the work proposed by Adi Shamir in [17], concerning the approach's level of abstraction. But the circuit described in [17] considers only if the attacker probes the Vcc, because the Gnd line remains vulnerable. Also, the two capacitors proposed are too big to be integrated (100nF). As explained in next session, our circuit masks the consumption even if the attack occurs in the Vcc or in the Gnd line.

4. Current Mask Generation technique

Based on the decreasing Signal to Noise Ratio idea, we conceived an analogical circuit able to mask the real power consumption from a cryptographic circuit. The main goal of this approach is to increase the security of the crypto devices without any modification of the cryptographic algorithm implemented. In addition, no special standard cells are required, unlike the dual-rail approach.

Our technique tries to mask the power consumption by normalizing the current consumed by the cryptographic circuit (*CC*). This task is accomplished by an analogical circuit called Current Mask Generator (*CMG*), which's role is to maintain the total current constant (from an external view). To design the *CMG*, firstly some measures were made, in order to establish the CC's peak of current consumption. Once this value detected, the objective is to remain at this peak, even if the CC's consumption is lower than it.

The CMG is composed basically of a high-swing current mirror, a follower circuit, and a small capacitance. As can be depicted from

Figure 2, the CMG acts aside of the CC, this outlines that any change in the cryptographic circuit is required.

Still in Figure 2, the current mirror acts imposing a fixed current (I2). I2 is given by the w coefficients from P1 and P0, as can be viewed in Equation (2) and is equals to the CC's peak of current consumption.

$$I_2 = \left(\frac{wP_1}{wP_0}\right) \quad (2)$$

The cryptographic circuit consumes a I_c current. When $I_c = I_2$, it means that the *CC* consumes all current furnished, and the *CMG* must stand by. Otherwise, when the *CC* not requires all I_2 current, then the circuit follower performs a feedback-loop to consume a current I_L so that $I_L = I_2 - I_c$.

In fact, the circuit follower plays as a voltage generator. The operational amplifier receives a tension from the mirror and compares it with a reference voltage (i.e. V_{ext}).). If the cryptographic circuit consumes an amount of current less than I_2, the voltage at the Op-Amp input will be lower than the reference voltage. Then the output of the Op-Amp will send 0 to the P_4 transistor. So, it will consume an I_L current, that is the difference between I_2 and I_c. When the *CC* consumes at the peak (i.e. $I_c = I_2$), the Op-Amp sends an 1 to the P_4, switching off the transistor, because it is no longer necessary to drain current.

Finally, the 10µF capacitor's function is to give some time to the feedback-loop react. Also, the capacitor smoothes the tension, what have a beneficial effect to the consumption masking

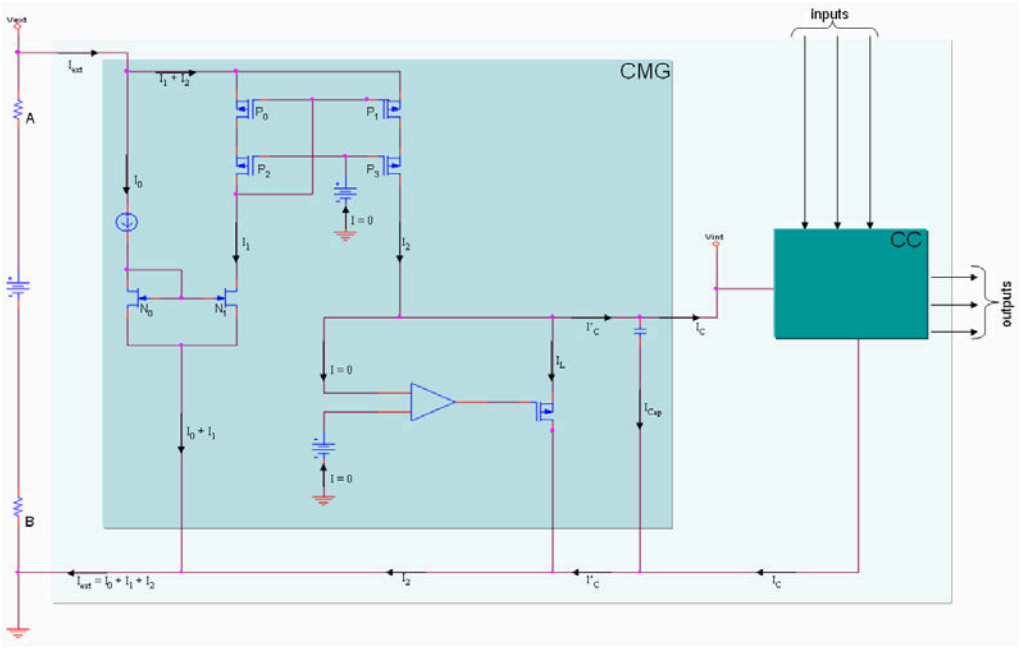

Figure 2 - The CMG circuit in detail[2]

To validate the *CMG* method, it is used a DES S-box, to play the role of *CC*. Then, the S-box was simulated to determine the current consumption worst case. The Figure 3 shows a peak consumption about 6mA.

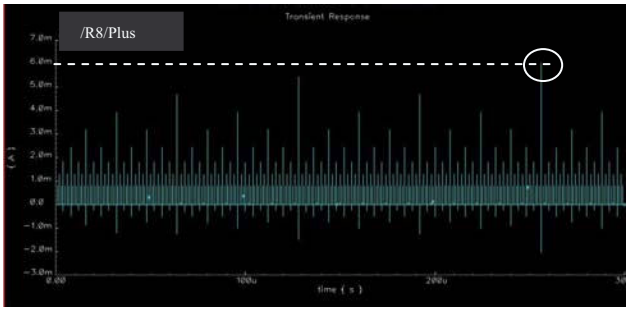

Figure 3 - The S-box current consumption

Many simulations were made for different data scenarios. As can be viewed in Figure 4, the *CMG* works efficiently, masking the *CC* current consumption, and making DPA attacks a very difficult task. The signal */R8/Plus* is the current consumed by the CC and the signal */R4/Plus* is the masked signal.

The Figure 5 shows, from a top-down view, the current consumption that can be plotted from the external Vdd or Gnd, the current consumption of the cryptographic circuit, de data input called a1 and the data input called a0. Analyzing the consumption reported to data input, Figure 5 shows that even with an 1 or a zero, or two ones, or two zeros as entries, the consumption viewed at the attackers side remains the same.

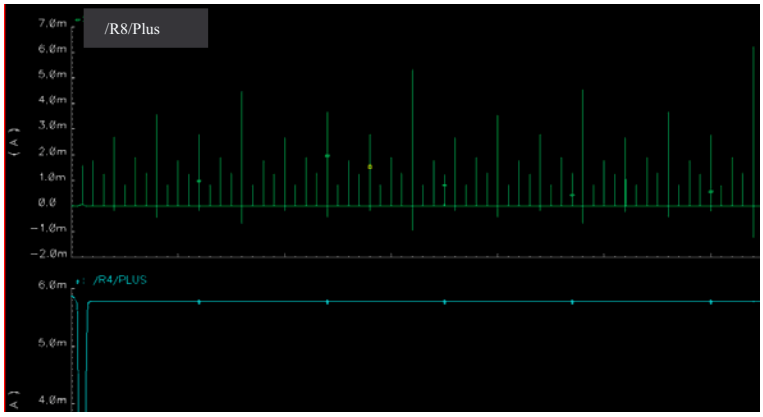

Figure 4 - The CC's current consumption and the current provided by the CMG circuit.

[2] Note that The CMG and the CC are not in the same scale

To define the difficulty to make a DPA attack, some parameters must be considered. The first one is the Signal to Noise Ratio. Contrary to a normal multimedia application, were the designer search to increase the SNR, by decreasing the noise as much as possible, the CMG approach intends the opposing: decrease the signal.

Figure 5 - The current provided by the CMG circuit, the current consumed by the CC, and some data input

Figures 4 and 5 shows glitches on the masked signal */R4/Plus*. If a instant is zoomed, the same pattern found in */R8/Plus* is repeated in */R4/Plus*. It signifies that the system is not perfect. But if the values of each signal are considered (see Figure 6), it is clear that the CMG attenuates the current by a factor $k \approx 20$.

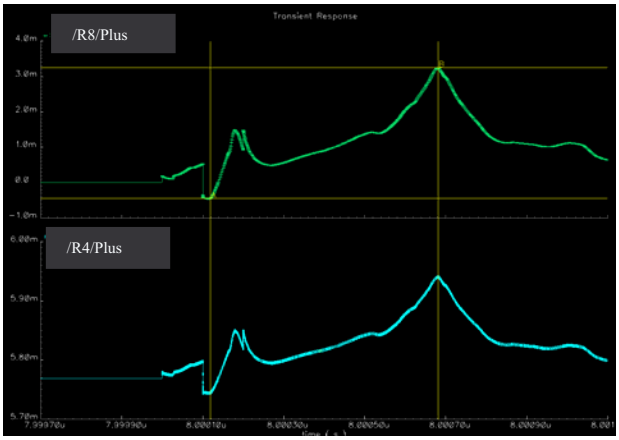

Figure 6 - The signal attenuation reached by the CMG

This k factor is obtained by measuring the */R8/Plus* signal's difference between its peak and its minor value (in the Figure 6, the points B and A), which done an CC_{delta}. Then, the process is repeated for the */R8/Plus* signal, obtaining a CMG_{delta} value. So,

$$k = \left(\frac{CC_{delta}}{CMG_{delta}} \right). \quad (3)$$

To view the CMG attenuation, the Signal to Noise Ratio show in the Equation (1) must be expanded:

$$SNR_{CMG} = 20 \times \log\left(\frac{I_c}{kN} \right) = 20 \times \log\left(\frac{I_c}{N}\right) - 20 \times \log(k) \quad . (4)$$

$$SNR_{CMG} = SNR - 20 \times \log(k)$$

With the equation (4), and regarding Figure 7 for the given example, the current viewed by an attacker is smoothed by 25db. It means that the observed signal could be drowned into the noise (Figure 7 (b)).

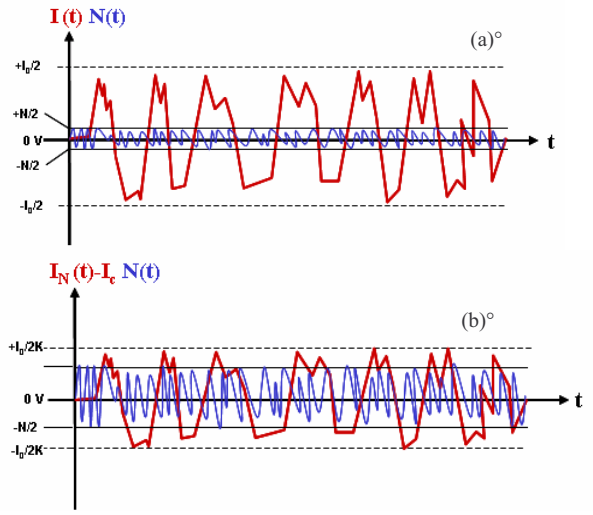

Figure 7 - Normal power consumption and noise (a) and the power consumption with the CMG, immersed into noise (b)

5. Conclusion

The presented work improves the robustness of cryptographic circuits against DPA attacks.

In this paper we have proposed a low level solution, which has as major contribution the fact that no changes are needed into the cryptographic circuit.

A problem with this approach is the integration of the capacitance. There are two solutions in study. The first one concerns the capacitance integration in a SIP technology [18]. This is feasible, but we do not made any costs study of this trend. On the other hand, maybe with the improvement of the feedback-loop and the current generator, i.e. with a more efficient mechanism, a large capacitance will not be required.

As another drawback, our proposal is not a low power solution. But it remains interesting for applications like credit cards, set-top boxes, phone cards and others where the low-power for cryptographic applications is less essential. In banking operations, like cash transactions, the cryptographic operation is not used all the time and the whole user operation is not so time-consuming that justifies a low power approach. The most important in this case is the security.

In spite of the drawbacks, the CMG approach remains interesting. The poor Signal to Noise Ratio generated by the CMG circuit makes a DPA attack very difficult.

6. References

[1] Rivest, R., Shamir, A., et al. "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems". *ACM Communications*, vol 21. pp 120-126. 1978.

[2] – . " Data Encryption Standard (DES)". Federal Information Processing Standards Publications (FIPS PUBS) Nº 46-3. http://csrc.nist.gov/publications/fips/fips46-3/fips46-3.pdf. EUA.October 25, 1999..

[3] Groupement des Cartes Bancaires CB. " Les cartes Bancaires en Nombres 2004". http://www.cartes-bancaires.com/FR/info/communiques/2005/DPchiffresCB2004.pdf. Paris, march 2005.

[4] Kocher, P., Jaffe J., et al. " Differential Power Analysis : Leaking Secrets ". Advances in Cryptology: Proceedings of CRYPTO'99, Vol 1666, Springer-Verlag,. pp. 388-397. 1999.

[5] Messerges, T. S., Dabbish E. A., et al. " Power Analysis of Modular Exponentiation in Smartcards ". Cryptographic Hardware and Embedded Systems - CHES 199. Lecture Notes in Computer Science, Vol. 1717, Springer, ISBN: 3-540-66646-X. Pp 144-157, 1999.

[6] Goubin, L., Patarin, J. "DES and Differential Power Analysis – The "duplication" method". Cryptographic Hardware and Embedded Systems - CHES 1999. Lecture Notes in Computer Science, Vol. 1717, Springer, ISBN: 3-540-66646-X. Pp 158-172, 1999.

[7] Trichina, E., De Seta, D. Et al. " Simplified Adaptive Multiplicative Masking for AES ". Cryptographic Hardware and Embedded Systems - CHES 2002. Lecture Notes in Computer Science, Vol. 2523, Springer, ISBN: 3-540-00409-2. Pp 187-197, 2003.

[8] Golic, J. D., Tymen, C. "Multiplicative masking and Power Analysis of AES". Cryptographic Hardware and Embedded Systems - CHES 2002. Lecture Notes in Computer Science, Vol. 2523, Springer, ISBN: 3-540-00409-2. Pp 198-212, 2003.

[9] Benini, L., Macii, A., et al. " Energy-aware design techniques for differential power analysis protection ". Design Automation Conference – DAC 2003. Anaheim, USA. June, 2003.

[10] Saputra, H. Vijaykrishnan, N., et al. " Masking behavior of DES encryption ". Design, Automation and Test Europe – DATE 2003. ACM-Sigda, ISBN 0-7695-1471-5. Munich, Germany, 2003.

[11] Simon M., Ross A., et al. "Balanced Self-Checking Asynchronous Logic for Smart Card Applications", Microprocessors and Microsystems Journal, 27(9). Elsevier, ISSN 0141-9331. pp 421-430, October 2003.

[12] Clavier, C., Coron, J-S., et al. " Differential Power Analysis in the presence of hardware countermeasures ". Cryptographic Hardware and Embedded Systems - CHES 2000. Lecture Notes in Computer Science, Vol. 1965, Springer, ISBN: 3-540-41455-X. Pp 252-263, 2000.

[13] Irwin, J., Page D., et al. " Instruction stream mutation for non-deterministic processors. Internation conference on Application Specific Systems, Architectures and Processors – ASAP 2002. IEEE press. Pp 286-295. 2002

[14] May, D., Muller H. L., et al. "Non-deterministic processors". Information security and privacy – ACISP 2001. Lecture Notes in computer Science, volume 2119. Springer ISBN 3-540-42300-1. pp 115-129. Sydney, Australia. July 2001.

[15] Mangard, S. "Hardware countermeasures against DPA – a statistical analysis of their effectiveness". Topics in Cryptology – CT-RSA 2004. Lecture Notes in Computer Science, Vol. 2964, Springer, ISBN: ISBN 3-540-20996-4. pp. 222 – 235. San Francisco, USA. February 2004.

[16] Fouque, P.-A., Muller F., et al. "Defeating Countermeasures Based on Randomized BSD Representations". Cryptographic Hardware and Embedded Systems - CHES 2004. Lecture Notes in Computer Science, Vol. 3156, Springer, ISBN: 3-540-22666-4 pp. 312 - 327. Cambridge, EUA. 2004.

[17] Shamir, A. "Protecting smart cards from passive power analysis with detached power supplies". Cryptographic Hardware and Embedded Systems - CHES 2000. Lecture Notes in Computer Science, Vol. 1965, Springer, ISBN: 3-540-41455-X. Pp 71-77, 2000.

[18] Tummala, R. and Madisetti, V. "System on Chip or System on Package?" IEEE Design and Test of Computers Review. Vol. 16, N. 2. IEEE Press. ISSN: 0740-7475. pp 48-56, April-June 1999.

[19] P. Kocher, J. Jaffe, B. Jun. "Introduction to Differential Power Analysis and Related Attacks". Technical Report, Cryptography Research Inc., 1998. Available from http://www.cryptography.com/dpa/technical/index.html.