

# Multi-Linear cryptanalysis in Power Analysis MLPA

Thomas Roche<sup>1,2</sup> and Cédric Tavernier<sup>2</sup>

<sup>1</sup> Laboratoire d'Informatique de Grenoble, 51 av. Jean Kuntzmann, 38330  
Montbonnot-Saint-Martin, France, [Thomas.Roche@imag.fr](mailto:Thomas.Roche@imag.fr)

<sup>2</sup> CS, Communication&Systems, 22 avenue Galilée, 92350 Le Plessis Robinson,  
France, [Cedric.Tavernier@c-s.fr](mailto:Cedric.Tavernier@c-s.fr)

**Abstract.** Since the Differential Power Analysis attacks (DPA, HO-DPA) were introduced by P. Kocher et al. it became necessary to develop sound and efficient countermeasures. Nowadays most embedded cryptographic primitives integrate one or several of these countermeasures (e.g. masking techniques, asynchronous designs, balanced dynamic dual-rail gates designs, noise adding, power consumption smoothing, etc. ...). This document presents new power analysis attacks using multi-linear approximations (MLPA attacks) that still work even when the symmetric cipher implementation is resistant to DPA and HO-DPA attacks.

**Keywords:** Power Analysis, multi-linear cryptanalysis, Reed-Muller codes.

## 1 Introduction

Since the discovery of Differential Power Analysis (DPA) and High Order Differential Power Analysis (HO-DPA) attacks in 1998 [1], the urge to develop resistant hardware implementations of symmetric ciphers has not ceased. Many countermeasures were developed: masking techniques, asynchronous designs, balanced dynamic dual-rail gates designs, noise adding, power consumption smoothing, etc. ... None of them assure, at the same time, very good security against DPA or HO-DPA and low overhead in terms of time and space (which of course is highly critical when considering smartcards). For instance, Jiqiang Lv and Yongfei [2] stated considering unique masking methods against HO-DPA: "Three 32-Bit Random Masks and Six Additional S-Boxes are the Minimal Cost for a Secure DES Implementation" (in such solution the six additional S-Boxes are generated at each new encryption).

Even though they give good theoretical resistances, Those countermeasures are so costly one would like to apply them only when the DPA-like attacks are possible. As a matter of fact, DPA attacks are only applicable when intermediate values of the encryption function are dependent on less than 32 key-bits. Hence, depending on the cipher diffusion function, after some rounds, there is no more need for countermeasures against DPA.

The contribution of this paper is twofold: First we present new Power Analysis Attacks, based on linear approximations, which can actually apply even

when DPA attacks are not feasible. i.e. when the leaked information from consumption measurement is only derived from 32 key-bits dependent intermediate data values. Then we give the current results given by MLPA attacks on some simulations and on some real consumption traces from the DPA-contest. Hence pragmatically proving the practicability of our attack (Section 2). Secondly, we show how those attacks could be ameliorated by disassociation with theoretical consumption models (Hence being closer to the reality) and could be used to attack unknown symmetric ciphers (Section 3).

## 2 (M)LPA Attacks

(M)LPA attacks correspond strictly to Linear [3] and Multi-Linear cryptanalysis [4] in the side-channel world.

### 2.1 Approach justification: Linear approximations

A linear approximation is a boolean linear function that takes plaintext and key bits as input and outputs a combination of ciphertext bits.

Let us denote  $|K|$ ,  $|P|$ ,  $|C|$  respectively the bit-lengths of key, plaintext and ciphertext. Let us consider a vector  $\Pi$  of length  $|P|$ ,  $\kappa$  of length  $|K|$  and  $\Gamma$  of length  $|C|$  and a bit  $b$ .  $\Pi$ ,  $\kappa$ ,  $\Gamma$  and  $b$  define a linear approximation of bias  $\epsilon$  over the symmetric cipher if and only if :

$$\Pr_{P,K}(\langle P, \Pi \rangle \oplus \langle K, \kappa \rangle \oplus b = \langle C(P, K), \Gamma \rangle) \geq 1/2 + \epsilon$$

where  $\langle x, y \rangle$  is the scalar product of two vectors  $x$  and  $y$  of same length over  $GF(2)$  (vector of bits). Given such a linear equation, Matsui showed that a high probability of success to recover the involved key bits in the equation using linear cryptanalysis would require a data-complexity (i.e. number of plaintext-ciphertext pairs) of  $N = 1/\epsilon^2$ .

It was shown in [5] that the use of several linear approximations involving the same key bits would improve the attack performances. Thus, given  $n$  linearly independent approximations of respective bias  $\epsilon_j, j \in \{1; \dots; n\}$ , the data-complexity of the attack would be reduced to  $N \approx 1/(\sum_{j=1}^n \epsilon_j^2)$ .

In a very recent paper, Loidreau et al. [6], studied the problem of finding all the linear approximations with a given bias of a given Boolean function. They showed the equivalence between this problem with a fixed output mask ( $\Gamma$  fixed) and a list decoding problem in the first order Reed-Muller code: it is done by polynomial reconstruction with queries in a highly noisy channel. They were then able to find good linear approximations up to 8 rounds of DES.

The use of approximations instead of a classic DPA selection function is a key point in our attack. An approximation can evaluate data values that are strictly dependent to more than 32 key-bits without involving all of them (relaxing on the certainty of the function allows to decrease the number of variables), hence we will be able to attack where DPA-like attacks are powerless: a cipher implementation where have been suppressed all registers or bus accesses — or any

countermeasure that would disable information leakage — up to the necessary number of rounds after which all the manipulated bits are dependent to more than 32 key-bits.

## 2.2 The MLPA attack

Power analysis attacks are based on consumption models (for CMOS circuits), the most common are the Hamming Weight (HW) and the Hamming Distance Model (HD). In the HW model the power consumption is related to the hamming weight of the data manipulated. Usually more accurate, in the HD model the consumption is related to the to the hamming weight of difference of two successive manipulated data (e.g. difference : bitwise xor).

The idea in LPA or MLPA attacks is to use linear approximations directly approximate the Hamming weight of a register since this is the quantity which is the most correlated to what is being measured. Thanks to the work of Loidreau et al. in [6], it is possible to find linear approximations of  $\langle H(C(P, K)), \Gamma_H \rangle$  for any chosen vector  $\Gamma_H$ . Where  $H()$  denotes the Hamming weight function on a vector of bits and  $\Gamma_H$  is a vector of length  $\log_2(|C|)$  with respect to the notations of section 2.1.

If we assume the HW or the HD model to be close enough to reality, then the use of linear approximations on the hamming weight value of a register (or a bus) should lead to attacks even when the manipulated values are dependent to more than 32 key-bits (which is impossible for DPA-like attacks).

Furthermore, the use of several linear approximations in a Multi-linear Power Analysis attack is possible, hence ameliorating the attack performances. As presented in [7, 6] in the context of classical multi-linear cryptanalysis, one can consider the recovering of some key-bits as the decoding problem of a code whose length is equal to the number of available linear relations and over a memoryless channel whose capacity depends on the respective biases of the linear approximations. Let us consider a set of  $n$  linear relations of biases  $\epsilon_l, l = 1, \dots, n$  with a form as follow :

$$\langle P, \Pi_l \rangle \oplus \langle H(C(P, K)), \Gamma_{H_l} \rangle \oplus b_l = \langle K, \kappa_l \rangle$$

where the set of vectors  $\kappa_l, l = 1 \dots n$  are such that a limited number  $k$  of key bits are involved in the equations (in practice less than 32 bits) and form a matrix of rank  $k$ . The idea is to reconstruct a code word  $y$  of length  $2^k$  from a noisy and erased codeword  $\tilde{y}$  wich is close enough to  $y$ , to be able to decode it in the first Reed-Muller code. The attack algorithm is detailed an online and more complete version of this article [8], its complexity is comparable to classical multi-linear cryptanalysis. Furthermore, the algorithm used to find the approximations is also of the same order ( $O(1/\epsilon^2)$ ).

## 2.3 Results

The results obtained using the above described attacks on the DES and AES cipher are of two kinds. First a theoretical validation of the attack (simulations)

and then a practical validation done on real power consumption traces. Table 1 and Table 2 summarize some of the results, in these tables, ”# linear equ.” refers to the total number of linear approximations found for the attack, not all of them have been useful, ”# Plaintext” or ”# Traces” refers to the data complexity of the attack and ”Pr(Success)” refers to the probability of success of the attack in simulation to retrieve ”# key bits” bits of the secret key.

**Attack simulations** By the means of Loidreau et al.’s work on finding linear approximations, up to three rounds can be approximated with good enough biases for the hamming weight of an intermediate data value. The simulation has been done considering that the cipher implementation leakage gives the hamming weight of the intermediate data values. Hence, in the HW model, the linear approximations evaluate the hamming weight of the round register (after respectively 1, 2 and 3 rounds), in the HD model, the linear approximations evaluate the hamming weight of the differences of the round register between two execution (two different plaintexts). Let note here that in a chosen plaintext attack, the HW model results correspond to an HD model. The attack on AES has been

Cipher	Model	rounds	# linear equ.	# key bits	# Plaintexts	Pr(Success)
DES	HW	1	349	48	$2^{12}$	0.99
DES	HW	2	728	48	$2^{12}$	0.95
DES	HW	3	164	27	$2^{20}$	0.99
DES	HD	2	27	16	$2^{16}$	0.99
AES	HW	Last	1410	128	$2^{11}$	0.99

**Table 1.** Simulation Results

done on the last round since it does not contains the `MixColumn` function.

**Attack on DPA-contest traces** Thanks to the DPA contest, power consumptions traces are freely available. The attack has been launched on the contest traces (`secmatv1_2006_04_0809`) that yield about 80000 power consumption traces. The linear approximations evaluate the hamming weight of the difference of data stored in the implementation register ( $LR$ ) (see [9] for more details on the DES implementation). For reasons of space, the attack description and setup are not described in details here, the interested reader can refer to the Annexe of an online and more complete version of this article [8].

Cipher	rounds	# linear equ.	# key bits	# traces
DES	1	84	20	1000
DES	1	84	45	20000
DES	2	163	10	1000
DES	2	163	47	36000

**Table 2.** Attack on DPA-contest traces Results

### 3 A Black-box approach

As we have seen in the results section, the attack described in section 2 works fine by approximating the power consumption as the Hamming weight of the difference of successive manipulated data (with the DPA-contest traces).

Let us now assume that we have access to a twin device where we can put arbitrary chosen keys, it would be possible to run the algorithm that search linear approximations directly on the twin device as a pre-processing phase of our attack. As the algorithm is run on a Boolean function as a black box, using the consumption measurement as output value of our Boolean function might render the attack even more efficient than in the model presented above since it does not have to rely on theoretical models anymore (HD or HW).

Further more, it is then possible to mount unknown cipher attacks since no knowledge of the symmetric cipher is needed except for its SPN structure (the hardware device is seen as a black box from which the consumption leakage are the outputs).

### 4 Conclusion and future work

The results shown in section 2.3 prove the feasibility of the MLPA attacks and so even when DPA-like attacks are not feasible.

The next step of this study is to set the practical attack using a twin board.

### References

1. Kocher, P., E, J.J., Jun, B.: Differential power analysis, Springer-Verlag (1999) 388–397
2. Lv, J., Han, Y.: Enhanced des implementation secure against high-order differential power analysis in smartcards. In Boyd, C., Nieto, J.M.G., eds.: ACISP. Volume 3574 of Lecture Notes in Computer Science., Springer (2005) 195–206
3. Matsui, M.: Linear cryptanalysis method for des cipher. In: EUROCRYPT. (1993) 386–397
4. Jr., B.S.K., Robshaw, M.J.B.: Linear cryptanalysis using multiple approximations. In Desmedt, Y., ed.: CRYPTO. Volume 839 of Lecture Notes in Computer Science., Springer (1994) 26–39
5. Biryukov, A., Cannière, C.D., Quisquater, M.: On multiple linear approximations. In Franklin, M.K., ed.: CRYPTO. Volume 3152 of Lecture Notes in Computer Science., Springer (2004) 1–22
6. Loidreau, P., Fourquet, R., Tavernier, C.: Finding good linear approximations of block ciphers and its application to cryptanalysis of reduced round des. In: Workshop on Coding Theory and Cryptography, Ullensvang, Norvège. (2009)
7. Gerard, B., Tillich, J.P.: On linear cryptanalysis with many linear approximations. Technical report (2007)
8. Roche, T., Tavernier, C.: Multi-linear cryptanalysis in power analysis attacks, mlpa. (Can be found here : <http://arxiv.org/abs/0906.0237>)
9. Guilley, S., Hoogvorst, P., Pacalet, R.: A fast pipelined multi-mode des architecture operating in ip representation. *Integration* **40**(4) (2007) 479–489