Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 04-69

Information theoretic measures applied to power and electromagnetic traces measured from an FPGA performing an Elliptic Curve Point Multiplication [1]

E. Dewitte, B. De Moor and B. Preneel[2]

09 April 2004
Published in the Proceedings of the 25th Symposium on Information Theory in the Benelux

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/dewitte/reports/WICsymp2004.ps.gz*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SCD-SISTA, Kasteelpark 10, 3001 Leuven, Belgium, Tel. 32/16/32 10 35, Fax 32/16/32 19 70, WWW: *http://www.esat.kuleuven.ac.be/scd*. E-mail: *evelyne.dewitte@esat.kuleuven.ac.be*. Evelyne Dewitte is a research assistant with the I.W.T. (Flemish Institute for Scientific and Technological Research in Industry). Dr. Bart De Moor and Dr. Bart Preneel are full professors at the Katholieke Universiteit Leuven, Belgium. Research supported by Research Council KUL: GOA-Mefisto 666, GOA-Ambiorics, several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects, G.0240.99 (multilinear algebra), G.0407.02 (support vector machines), G.0197.02 (power islands), G.0141.03 (Identification and cryptography), G.0491.03 (control for intensive care glycemia), G.0120.03 (QIT), G.0452.04 (QC), G.0499.04 (robust SVM), research communities (ICCoS, ANMMM, MLDM); AWI: Bil. Int. Collaboration Hungary/ Poland; IWT: PhD Grants, GBOU (McKnow) Belgian Federal Government: Belgian Federal Science Policy Office: IUAP V-22 (Dynamical Systems and Control: Computation, Identification and Modelling, 2002-2006), PODO-II (CP/01/40: TMS and Sustainibility); EU: FP5-Quprodis; ERNSI; Eureka 2063-IMPACT; Eureka 2419-FliTE; Contract Research/agreements: ISMC/IPCOS, Data4s, TML, Elia, LMS, IPCOS, Mastercard.

Abstract

Side-channel attacks exploit information that leaks from a cryptographic device due to a specific implementation. Two important examples of side-channels are the power dissipation and the electromagnetic radiation. Though each of these channels is being well studied, the question remains whether combining multiple channels yields advantages such as faster exhibition of sensitive data. Of course, the more independent the channels, the more interesting combining them. Information theory not only presents us tools to measure (in)dependence. It also tells us how badly a certain channel is perturbed by noise. In calculating the entropy of our measurements per clock cycle, a surprising pattern in the plots was found.

INFORMATION THEORETIC MEASURES APPLIED TO POWER AND ELECTROMAGNETIC TRACES MEASURED FROM AN FPGA PERFORMING AN ELLIPTIC CURVE POINT MULTIPLICATION

E. Dewitte, B. De Moor, B. Preneel
Katholieke Universiteit Leuven
Department of Electrical Engineering, ESAT-SCD-SISTA/COSIC
Kasteelpark Arenberg 10, B-3001 Heverlee (Leuven), Belgium
{Evelyne.Dewitte, Bart.Demoor, Bart.Preneel}@esat.kuleuven.ac.be

Side-channel attacks exploit information that leaks from a cryptographic device due to a specific implementation. Two important examples of side-channels are the power dissipation and the electromagnetic radiation. Though each of these channels is being well studied, the question remains whether combining multiple channels yields advantages such as faster exhibition of sensitive data. Of course, the more independent the channels, the more interesting combining them. Information theory not only presents us tools to measure (in)dependence. It also tells us how badly a certain channel is perturbed by noise. In calculating the entropy of our measurements per clock cycle, a surprising pattern in the plots was found.

INTRODUCTION

The security of a cryptographic system depends not only on the mathematical analysis of the algorithm itself but also on the security of the implementation [?]. The past seven years a lot of research is done towards so-called *side-channel attacks*. These attacks are tailored at a specific implementation. They exploit the fact that while a device is running a cryptographic algorithm, sensitive information may leak through physical quantities of the device (the side-channels). Well-studied examples of side-channels are timing [?], power consumption [?, ?] or electromagnetic (EM) radiation of the working device [?, ?].
The idea to use information theory on side-channel measurements originated from the question whether it is possible to determine a lower bound on the number of measurements we should take to have a succesful attack. Each measurement is affected by noise; by using statistical tools one hopes to reduce the noise as far as possible. Statistics however always imply assumptions such as Gaussian noise,

mean-zero noise, ... We approached the question in the following way: noise introduces uncertainty about the outcome of the measurements at each sample time; let us try to catch this uncertainty in a mathematical way. The higher the uncertainty, the more noise and the more measurements are required! Information theory offers us a great tool to quantify this uncertainty: the entropy. At the same time it allows us to detect moments of low respectively high entropy in the running of the algorithm *and* to compare channels. We are still performing experiments to explore what information theory can reveal to us; a few surprising results however showed from the first tests.

EXPERIMENTAL SETUP

Elliptic curve cryptography (ECC) was proposed in the 1980's [?]. When compared with a classical cryptosystem as RSA, ECC offers advantages such as lower power consumption and shorter keys. We expect ECC to be used more and more in the future.

In a first setup we executed an EC point addition on a Xilinx Virtex FPGA board that was hand-made at Cosic. This EC point addition is realized by algorithm **??**. The EC point addition requires fourteen states and six temporary

Algorithm 1 EC point addition

Require: $P_1 = (x, y, 1, a),\ P_2 = (X_2, Y_2, Z_2, aZ_2^4)$
Ensure: $P_1 + P_2 = P_3 = (X_3, Y_3, Z_3, aZ_3^4)$

1.	$T_1 \leftarrow Z_2^2$	
2.	$T_2 \leftarrow xT_1$	
3.	$T_1 \leftarrow T_1 Z_2$	$T_3 \leftarrow X_2 - T_2$
4.	$T_1 \leftarrow yT_1$	
5.	$T_4 \leftarrow T_3^2$	$T_5 \leftarrow Y_2 - T_1$
6.	$T_2 \leftarrow T_2 T_4$	
7.	$T_4 \leftarrow T_4 T_3$	$T_6 \leftarrow 2T_2$
8.	$Z_3 \leftarrow Z_2 T_3$	$T_6 \leftarrow T_4 + T_6$
9.	$T_3 \leftarrow T_5^2$	
10.	$T_1 \leftarrow T_1 T_4$	$X_3 \leftarrow T_3 - T_6$
11.	$aZ_3^4 \leftarrow Z_3^2$	$T_2 \leftarrow T_2 - X_3$
12.	$T_3 \leftarrow T_5 T_2$	
13.	$aZ_3^4 \leftarrow \left(aZ_3^4\right)^2$	$Y_3 \leftarrow T_3 - T_1$
14.	$aZ_3^4 \leftarrow a\left(aZ_3^4\right)$	

registers. We restrained ourselves however to states 9, 10 en 11 (the X_3 register update). Simultaneously, we recorded the power consumption of the device and the electromagnetic radiation. While the power trace was registered fully, we

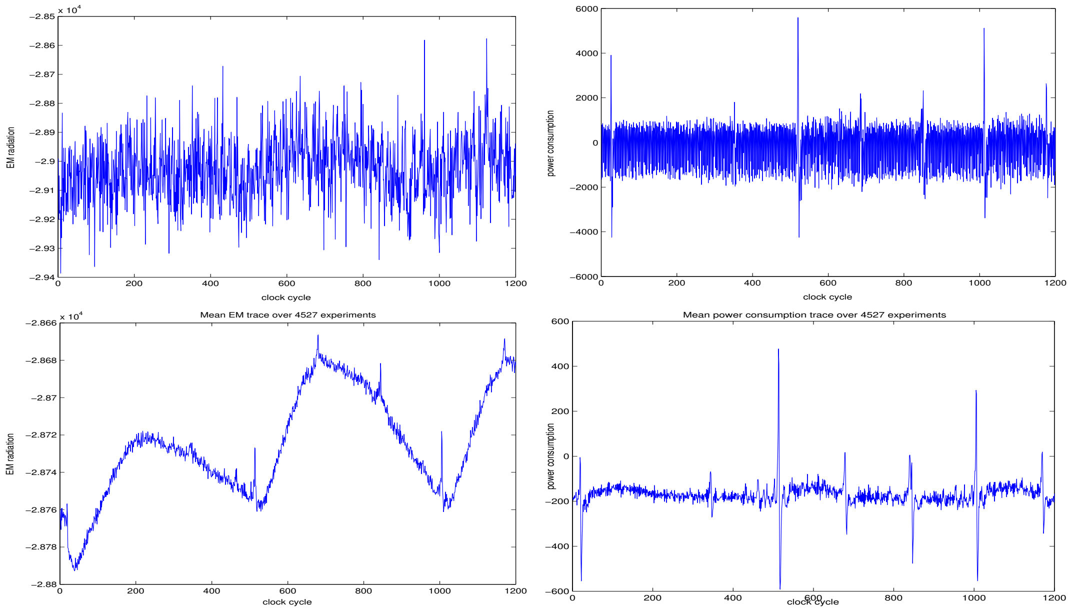

Figure 1: Above: Typical EM (left) and power traces (right). Below: the mean measured traces

only kept the above part of the EM signal because of the limited internal memory of our scope (Tektronix TDS714L). Typical traces captured in this way can be found in figures **??**.

SIDE-CHANNEL EFFECTS AT FIXED TIME POINTS

Let $Y_{em,i}$ for $1 \leq i \leq 1200$ be the electromagnetic radiation at clock cycle i in the running of the cryptographic algorithm. Due to noise, these physical quantities are random variables with respectively probability density functions (pdfs) f_i:

$$Y_{em\ ,\ i}\ \sim\ f_i\ \in\ \mathcal{F}\ , \tag{1}$$

where $\mathcal{F}$ is the family of pdfs for electromagnetic radiation. For each of the 1200 members of the family, we gathered 4527 observations as a sample from the unknown pdf.

Analogously we define $Y_{pow,j}$ for $1 \leq j \leq 1200$ to be the power consumption at instant j in the running of the cryptographic algorithm:

$$Y_{pow\ ,\ j}\ \sim\ g_j\ \in\ \mathcal{G}\ , \tag{2}$$

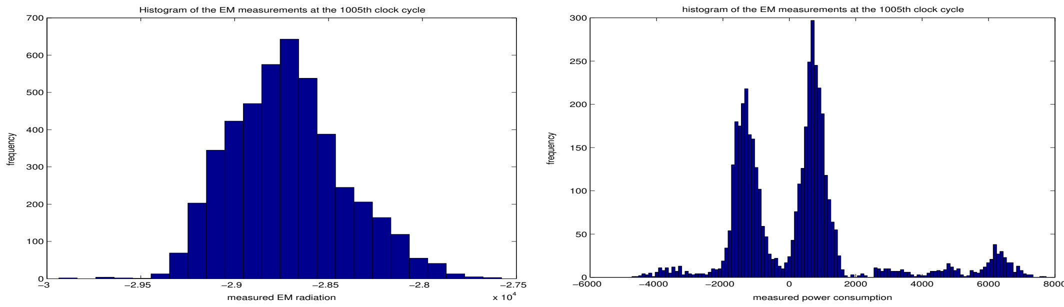

Figure 2: Histograms for the side-channel measurements in clock cycle 1005: the EM radiation (left) and the power dissipation (right)

where $\mathcal{G}$ is the family of pdfs for power dissipation.
Remark that all pdfs are unknown and have to be estimated. The estimators of f_i and g_j are denoted by $\hat{f}_i$ and $\hat{g}_j$. The resulting entropy values are influenced by the choice of estimator.

DENSITY ESTIMATION

We approximated the pdfs (**??**) and (**??**) of our continuous stochastic variables by means of a histogram technique. Hence we discretize the side-channels pretending we can only measure a finite number of discrete values per clock cycle. For each cycle, we estimated the pdf at equally-spaced points starting from the minimum measured value and choosing a fixed spacing δ. Consequently the bin-width is the same in all histograms (thus for both channels and all clock cycles); the number of bins increases with the range of measured values at a certain cycle. In figure **??** for instance, we compare the histograms found for the EM radiation respectively the power dissipation in clock cycle 1005 in the running of the cryptographic algorithm.

ESTIMATING THE ENTROPY OF A CONTINUOUS RANDOM VARIABLE

Shannon entropy We discretized the side-channels; initially we can use the estimated probabilities derived from the histogram to calculate the well-known *Shannon entropy*:

$$H(\hat{Y}_{em}\ ,\ _i) = -\sum_{k=1}^{M_i} \hat{f}_i(k)\log_2(\hat{f}_i(k)) \qquad 1 \leq i \leq 1200 \tag{3}$$

$$H(\hat{Y}_{pow}\ ,\ _j) = -\sum_{k=1}^{N_j} \hat{g}_j(k)\log_2(\hat{g}_j(k)) \qquad 1 \leq j \leq 1200\ , \tag{4}$$

where M_i and N_j denote the respective number of bins.

Differential entropy In a second scheme we respect the fact that the measurements $Y_{em,i}$ and $Y_{pow,j}$ represent continuous variables. The entropy we would like to calculate, is the *continuous* or *differential entropy*:

$$h(\hat{Y}_{em\ ,\ i}) \quad = \quad -\int_{-\infty}^{\infty} \hat{f}_i(y)\log_2(\hat{f}_i(y))dy \qquad 1 \leq i \leq 1200 \tag{5}$$

$$h(\hat{Y}_{pow\ ,\ j}) \quad = \quad -\int_{-\infty}^{\infty} \hat{g}_j(y)\log_2(\hat{g}_j(y))dy \qquad 1 \leq j \leq 1200\ . \tag{6}$$

Because histograms only evaluate a discrete number of points, we use the following straightforward approximation:

$$\hat{h}(\hat{Y}_{em\ ,\ i}) \quad = \quad -\frac{\delta_{em\ ,\ i}}{2}\{S(1) + S(M_i) + 2\sum_{k=2}^{M_i-1} S(k)\} \tag{7}$$

$$\hat{h}(\hat{Y}_{pow\ ,\ j}) \quad = \quad -\frac{\delta_{pow\ ,\ j}}{2}\{T(1) + T(N_j) + 2\sum_{k=2}^{N_j-1} T(k)\}\ , \tag{8}$$

where $S(k) = \hat{f}_i(k)\ \log_2(\hat{f}_i(k))$, $T(k) = \hat{g}_j(k)\ \log_2(\hat{g}_j(k))$ and δ denotes the space between the equally-spaced points in which the density estimate was evaluated.

THE ESTIMATED ENTROPY OF THE SIDE-CHANNEL EFFECTS PER CLOCK CYCLE

Finally we visualize the entropy calculations for our approximations: $H(\hat{Y}_{em\ ,\ i})$ and $H(\hat{Y}_{pow\ ,\ j})$ in figure **??**, $\hat{h}(\hat{Y}_{em\ ,\ i})$ and $\hat{h}(\hat{Y}_{pow\ ,\ j})$ in figure **??**. As expected, the estimations made by (**??**) and (**??**) are bigger than those based on the histogram technique. It is well-known that when the quantization is made finer and finer (i.e. smaller δ), the entropy keeps on increasing. In the limit (for $\delta \to \infty$), the entropy becomes infinitely large. This is due to the fact that continuous variables have an infinite number of possible outcomes. In practice, we can only use these estimates to compare the entropy of two continuous random variables. We then just have to make sure that the same discretization δ is used. What surprized us were the clear patterns that popped up in the entropy traces; especially for the power dissipation channel. An explanation may be that in states 10 and 11 of algorithm **??** the same type of instructions are executed: a multiplication in parallel with an addition. These instructions would then induce a same amount of

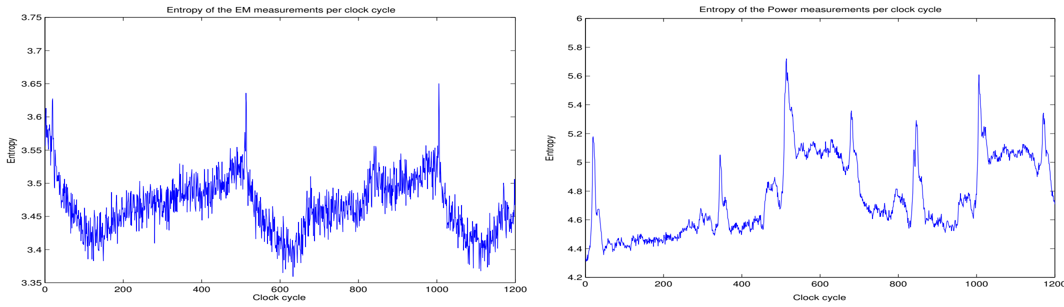

Figure 3: The histogram-based estimated entropy per clock cycle for both channels: EM (left) and power (right)

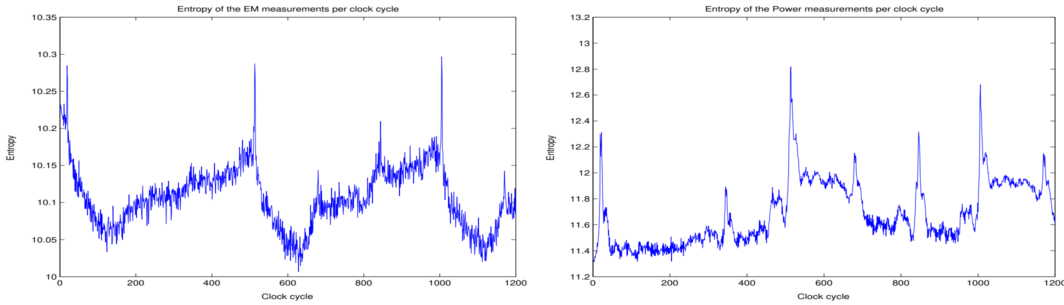

Figure 4: The approximated differential entropy per clock cycle for both channels: EM (left) and power (right)

entropy and thus uncertainty. Consequently, making entropyplots of side-channel measurements may present us with patterns in the plots enabling us to distinguish between the underlying instructions. Remark that no clear patterns were present in the original measurements.

CONCLUSIONS AND FUTURE WORK

We proposed to use information theoretic measures such as entropy and mutual information to quantify the uncertainty of side-channel measurements on the one hand and to assess the dependence of the channels on the other hand. While we are still busy interpreting the results, we surprisingly found clear patterns in the entropy traces. Similar patterns appear when the device is executing analogue instructions; thus making plots of the entropy per clock cycle may enable us to visually recognize patterns and hence instructions in the algorithm. In the original measurements no patterns could be found indicating the underlying instructions. Of course, prudence is called for as we have to test these suggestions on more datasets. Currently, more experiments are done in order to confirm or reject our hypothesis.

AKNOWLEDGEMENTS

Evelyne Dewitte is a research assistant with the I.W.T. (Flemish Institute for Scientific and Technological Research in Industry). Dr. Bart De Moor and Dr. Bart Preneel are full professors at the Katholieke Universiteit Leuven, Belgium. Research supported by Research Council KUL: GOA-Mefisto 666, GOA-Ambiorics, several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects, G.0240.99 (multilinear algebra), G.0407.02 (support vector machines), G.0197.02 (power islands), G.0141.03 (Identification and cryptography), G.0491.03 (control for intensive care glycemia), G.0120.03 (QIT), G.0452.04 (QC), G.0499.04 (robust SVM), research communities (ICCoS, ANMMM, MLDM); AWI: Bil. Int. Collaboration Hungary/ Poland; IWT: PhD Grants, GBOU (McKnow) Belgian Federal Government: Belgian Federal Science Policy Office: IUAP V-22 (Dynamical Systems and Control: Computation, Identification and Modelling, 2002-2006), PODO-II (CP/01/40: TMS and Sustainibility); EU: FP5-Quprodis; ERNSI; Eureka 2063-IMPACT; Eureka 2419-FliTE; Contract Research/agreements: ISMC/IPCOS, Data4s, TML, Elia, LMS, IPCOS, Mastercard.