RESEARCH ARTICLE

Advanced Science Letters
Vol. 5, 1–5, 2012

Secret Recovery from Electromagnetic Emissions

Hongying Liu[1,*], Yibo Fan[2], and Satoshi Goto[1]

[1]*Graduate School of Information, Production and Systems, Waseda University, Kitakyushu, 8080135, Japan*
[2]*State-Key Lab of ASIC and System, Fudan University, Shanghai, 201203, China*

Electromagnetic emissions leak confidential data of cryptographic devices. The electromagnetic emission has been reported as an important side channel for cryptanalysis. Electromagnetic Analysis (EMA) exploits the external radiation of cryptographic devices during encryption to reveal secret keys. The performance of EMA depends on the acquired signals to a large extent. To protect the devices from attacks, noises are introduced in the side channel either by unintentional interference from surroundings or elaborate design from engineers. Thus the secret recovery becomes difficult and even unavailable. In this paper, we propose two signal processing techniques to counteract both of these noises. The bandpass filtering and independent component analysis are widely used in other areas. We demonstrate their applications to EMA against the encryption algorithms on application-specific integrated circuit. With these techniques, the secret keys are extracted successfully and rapidly.

Keywords: Electromagnetic Emissions, Side Channel Analysis (SCA), Electromagnetic Analysis (EMA), AES, Camellia.

1. INTRODUCTION

The issue of establishing secure systems has always been concerned. Due to the complexity of the computational systems, the attacks and protections vary. For example, the intrusion detection and defense are important for computer systems, while the non-invasive attacks have threatened the security of cryptographic devices.

A non-invasive attack involves close observation or manipulation of the device's operation. Unlike invasive attack involves depackaging cryptographic devices to directly access the internal components, this attack only exploits externally available information which is often unintentionally leaked, such as supply voltage and clock signal of the processor. Side channel analysis (SCA) belongs to this category. It exploits the information leaked from cryptographic devices during encryption or decryption to infer secrets.

Though a number of works have been done on electromagnetic emissions, such as the interference reduction by desynchronization methodology,[1] electromagnetic emissions have been reported as a powerful side channel of cryptographic devices, besides side channel information from timing, cache, power consumption etc. For example, according to news on Science (November 2008), two university students built a 40-EUR gadget and successfully obtained the data on RFID MIFARE Classic card.[2]

Electromagnetic analysis (EMA) is performed with low-cost sensors to extract secret information from these devices even at a distance. Thus the EM side channel leaks information which might not be available from power consumption. EMA has been actively investigated and studied as one of side channel cryptanalysis by researchers. The simple EMA (SEMA) and differential EMA (DEMA) were demonstrated.[3] Planar near-field cartography was used to enhance the correlation-based EMA by detecting the hot spot on the device.[4] The revealing of secrete keys largely depends on the acquired EM signals. In general, there are two types of noise that prevent a fast key exposure. One is the non-algorithm noise, which originates from external, intrinsic, sampling and quantization noise unintentionally, and the other is algorithm noise,[5] which results from the countermeasures added intentionally. For example, the signal may be displaced due to the random delays inserted into the encryption algorithm, or multiple encryption algorithms may run simultaneously, referred as simultaneous algorithm noise. Then the data dependent signals are hidden from detection. Several approaches have been investigated to reduce these noises. Le. et al.[6] adopt the fourth-order cumulant to decrease the non-algorithm noise. Homma et al.[7] apply the method of phase-based waveform matching to overcome the signal displacement. Because of the complexity and variations of the algorithm noise, there is few works deal with simultaneous algorithm noise.

Unlike the previous work, in this paper, we explore two signal processing techniques, which have been widely applied to other areas. They are studied and applied to EMA. Bandpass filtering is effective for non-algorithm noise reduction of EM signals. ICA can enhance the efficiency of EMA at the presence of simultaneous algorithm noise.

*Author to whom correspondence should be addressed.

The remainder of this paper is organized as follows. Section 2 describes some related preliminaries. Section 3 shows the measurement setup for the experiment. Section 4 presents the application of bandpass filtering in detail. The simultaneous algorithm noise and ICA are discussed in Section 5. Section 6 draws conclusions and suggests future work.

2. PRELIMINARIES

Advanced Encryption Standard (AES) was published by the National Institute of Standards and Technology (NIST) of United States in 2001. It has become one of the most popular symmetrical encryption algorithms. Camellia is a symmetrical key block cipher developed jointly by Mitsubishi Electric Corporation and Nippon Telegraph and Telephone Public Corporation (NTT in short) in 2000. The cipher has security levels and processing abilities comparable to AES.

The procedures for EMA which is similar to a typical SCA, includes: First, select the target for analysis. The target is certain point of the encryption of the algorithm. Second, measure the real leakage during the execution of encryption. Then make predictions about the leakage based on leakage model, such as Hamming weight, Hamming Distance etc. and analyze statistically with metrics such as correlation coefficient, difference of means etc. Finally reveal the keys.

Specifically, the algorithm for correlation-based EMA is calculating Hamming Distance[8] $\boldsymbol{H}_{i,R}$ between ciphertext $\boldsymbol{C}_i$ and the reference state $\boldsymbol{R}$ according to Eq. (1), where $\boldsymbol{HW}$ denotes Hamming weight, i is the number of sampling.

$$\boldsymbol{H}_{i,R} = \boldsymbol{HW}(\boldsymbol{R} \oplus \boldsymbol{C}_i) \tag{1}$$

$$\rho = \frac{N\sum P(C_i)H_{i,R} - \sum P(C_i)\sum H_{i,R}}{\sqrt{N\sum P(C_i)^2 - (\sum P(C_i))^2}\sqrt{N\sum H_{i,R}^2 - (\sum H_{i,R})^2}} \tag{2}$$

For a correct key, the operation is data dependent. Thereby, leakage $P(C_i)$ from EM signal has a linear relationship with Hamming Distance $\boldsymbol{H}_{i,R}$, and the correct key is the one that maximizes correlation coefficient ρ, given by Eq. (2).

Generally, the performance of EMA is assessed by success rate, which expresses the number of correct key guess among all the key bytes. In our work, we test the proposed two techniques by EMA against AES and Camellia implementations on Side-channel Attack Standard Evaluation Board-R (SASEBO-R).[9]

3. MEASUREMENT SETUP

Experiment environment is shown in Figure 1. A cryptographic LSI and a control FPGA are mounted on PCB of dimension 230 mm × 180 mm × 1.6 mm. The cryptographic cores use 0.13 μm TSMC standard library of CMOS process technology. From AES1 to AES4, the s-boxes are based on Look-up table, (Positive Polarity Reed Muler 1-stage) PPRM1, PPRM3 and the multiplicative inverse circuit with a composite field respectively. AES0 is similar to AES4 but with support of decryption. RS-232 and LAN interfaces are provided to communicate with the host PC. A sustentation, with scales in three dimensions, which is settled perpendicularly onto the baseplate, is used to control the height and the location of the sensor above the PCB. Additionally, a preamplifier with gain 50 dB is connected to EM sensor through coaxial cable to magnify weak EM signals before they are sent to oscilloscope. The oscilloscope is Agilent MSO 54832D. The EM sensor is a loop probe, with diameter 10 mm.

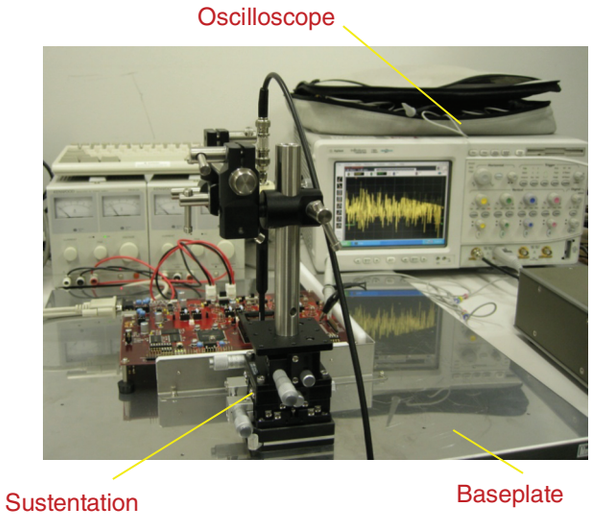

Fig. 1. Experimental environment.

Computer randomly generates 128-bit plaintext in groups, which are transmitted to FPGA through RS-232 serial ports, and then upon receipt of the plaintext, the FPGA controls LSI to implement AES encryption. The output of the encryption function "add round key" in the final round of AES is chosen as a target to analyze. The encryption proceeds with 10000 random plaintexts and a fixed but random 16-byte key (the final round): 28AFCE9F5AFFC8F1E054B352B0CE430E. Each measurement is repeated 30 times to calculate average and reduce accidental error. Then after signal processing, which is presented in the following section, the correlation-based EMA is performed.

4. EMA WITH BANDPASS FILTERING

This is a technique that passes frequencies within a certain range and rejects frequencies outside that range. It is always fulfilled by certain filter, which may be realized by hardware that appended to devices, such as a preamplifier, etc. The power spectral density (PSD) of one sampling signal from 0 Hz to 250 MHz is plotted in Figure 2. It shows that the power of EM signal distributes at a wide frequency though the working frequency is 24 MHz. Therefore, bandpass filtering based on software algorithm is necessary. The bandpass filtering is described by Eq. (3).

$$\boldsymbol{y}[t] = \sum_{i=0}^{N} \boldsymbol{b}_i \boldsymbol{x}[t-i] \tag{3}$$

where $\boldsymbol{x}[t]$ is input signal, $\boldsymbol{y}[t]$ is output signal, $\boldsymbol{b}_i$ is the coefficients of a filter, $\boldsymbol{N}$ is the order of filter.

Because it can maintain the frequency interval efficiently, we use the Hanning window function to design a filter in our work. The pass band of the filter is from 0 Hz to 40 MHz. Signals become smoother because that unrelated frequency component is attenuated. Thus it leads to an enhanced success rate. All the sub-keys are revealed in 2905 signals with filtering, while 3614 signals are needed without filtering.

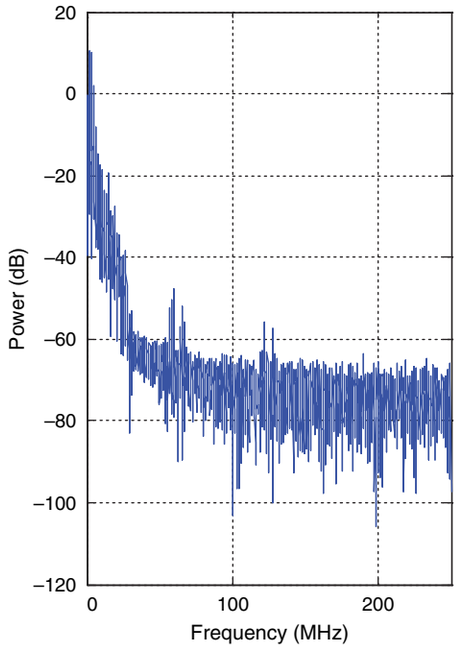

Fig. 2. PSD of sampling signal.

5. EMA WITH ICA

5.1. Simultaneous Algorithm Noise

In general, only one encryption module runs and the corresponding EM signals are measured and collected during an EMA. However, in order to hide the data-dependent information from attackers, multiple encryption modules may run simultaneously. This is known as simultaneous algorithm noise, which is an effective countermeasure that slower the key detection. For instance, multiple AES modules (AES0, AES1,…, AESn) and Camellia may run at the same time. On the ASIC, we activate module AES0-AES4 and Camellia simultaneously, record the mixed signals, and perform EMA. The 16 byte-keys are detected within 8291 signals, the evolution of the second byte key "AF" is shown in Figure 3. By contrast, only 3614 signals are needed when only AES0 runs.

5.2. Solutions

Aiming at probing into the possible solutions to reduce the number of signals at the presence of simultaneous algorithm noise, we studied the problem of blind source separation (BSS).

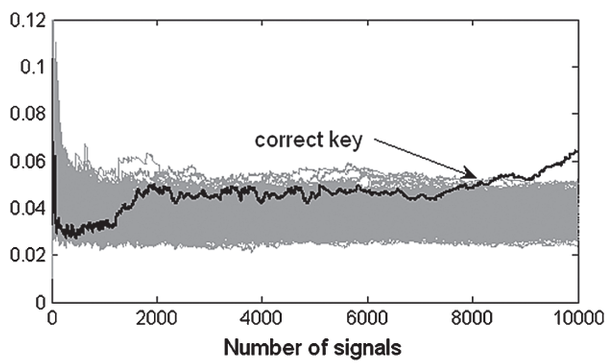

Fig. 3. Evolution of the second key byte: "AF".

Consider that there are a number of signals emitted by some physical objects or sources, such as the electric signals emitted by different areas of brain, the radio signals emanated by mobile phones or the speech signals etc. Then the sensors receive and record these signals in the form of a mixture of the original source signals. We are interested to find the original source signals from the mixture with little or no knowledge about the source signals.

ICA has been proved as an effective way to solve this problem. Because FastICA[10] has good performance for source separation and convergence, it has been one of the most popular algorithms. The model of ICA: Assume that we observe m linear mixed signals $\boldsymbol{X}$ of m independent components, shown by Eq. (4):

$$\boldsymbol{X} = \boldsymbol{AS} + \boldsymbol{N} \tag{4}$$

where $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2,\ldots, \boldsymbol{X}_m)^T$, is m mixed signals which are observed, $\boldsymbol{S} = (\boldsymbol{S}_1, \boldsymbol{S}_2,\ldots, \boldsymbol{S}_n)^T$, is the n source signals, $\boldsymbol{N} = (\boldsymbol{N}_1, \boldsymbol{N}_2,\ldots, \boldsymbol{N}_m)^T$, denotes the m noise vector, superscript T denotes transpose of matrix. All of these signals have sampling length L. Then, after estimating a matrix $\boldsymbol{W}$, the independent component can be obtained by: $\boldsymbol{S} = \boldsymbol{WX}$.

FastICA is based on a fixed-point iteration scheme to find a direction, i.e., a unit vector $\boldsymbol{W}$ such that the projection $\boldsymbol{W}^T\boldsymbol{X}$ maximizes nongaussianity. The algorithm is given by Eqs. (5)–(7).

$$\boldsymbol{Z} = \boldsymbol{QX} \tag{5}$$

$$\boldsymbol{W}^+ \leftarrow \boldsymbol{Zg}(\boldsymbol{Z}^T\boldsymbol{W}) - \boldsymbol{Wg}'(\boldsymbol{W}^T\boldsymbol{Z})\boldsymbol{O}_{Lx1} \tag{6}$$

$$\boldsymbol{W}^+ \leftarrow \boldsymbol{W}^+ / \left\| \boldsymbol{W}^+ \right\| \tag{7}$$

where $\boldsymbol{X}$ is centered and whitened to simplify the computation. $\boldsymbol{Q}$ is unitary matrix, $\boldsymbol{E}(\boldsymbol{ZZ}^T) = \boldsymbol{I}$ is satisfied, where $\boldsymbol{E}$ is the mathematical expectation, $\boldsymbol{g}$ is the non-linear contrast function (namely objective function), Vector $\boldsymbol{O}_{Lx1}$ has all values of one. And in Eq. (7), the normalization has been added to improve the stability. $\boldsymbol{g}'$ denotes the mathematical derivative.

The problem of simultaneous algorithm noise fits well with the above model. EMA is conducted when the details of encryption module is unknown to the attacker. One can only measure the leaked mixed signals. In the following experiments, FastICA is applied to the mixed signal to separate the most uncorrelated source of encryption.

5.3. Experiments

Experiment 1 A Mixture with 2 Encryption sources. We set the bits in the interface circuits through computer. AES0 and Camellia on the LSI execute simultaneously. Two mixed signals which are shown in Figures 4(a and b), with different plaintext and the same key are input to the FastICA algorithm. This leads to two separated signals: one is AES0 signal, and the other is Camellia signal. They are shown by Figures 4(d and f) respectively. Then AES0 and Camellia executes individually. These are supposed to be the source signals, which are plotted in Figures 4(c and e) respectively. With the same key and plaintexts as in Section 3, 10000 EM signals each with sampling length 2000, are recorded with oscilloscope during the execution of AES0 and Camellia. Then every two signals are input to FastICA algorithm, it yields

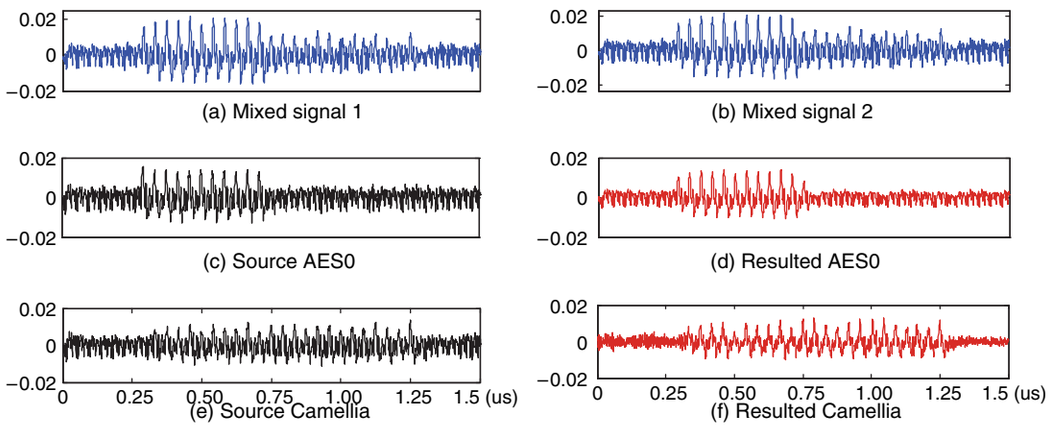

Fig. 4. The signals of two encryption source: AES0 and Camellia.

10000×2 separated signals in total, which has the same number of AES signals and Camellia signals. The 10000 resulted AES0 signals are used for EMA. All the key bytes are revealed within 5126 signals. This is faster than the mixed signals, which is with 8133 signals.

Experiment 2 A Mixture with 3 Encryption Sources. The situation becomes complex when 3 encryption sources are mixed. The encryption signals are recorded when AES0, AES1 and Camellia run simultaneously. Similar to the previous process, we use 3 mixed signals and attempt to obtain the 3 separated signals. However, the resulted signals are not clearly separated. Only one of the resulted signals has a greater correlation with the source Camellia. This indicates that Camellia has been separated successfully.

The explanations for these results are: because any one of the AES executions (AESi, $i = 0$–5) on LSI has a linear relation with Hamming Distance, the relation between different AES is not independent. The independence assumption of ICA is not satisfied. Thus the separation of different AES fails.

We substrate the resulted Camellia from the mixed signals, namely leave the mixture of AES0 and AES1. Then EMA is conducted with this mixture. The success rates are compared with the case of 3 mixed signals and shown in Figure 5. 8524 signals are needed to reveal all the key bytes for the original 3 mixed signals. Only 5207 signals are needed for the separated signals. The success rate is enhanced, though only Camellia is separated.

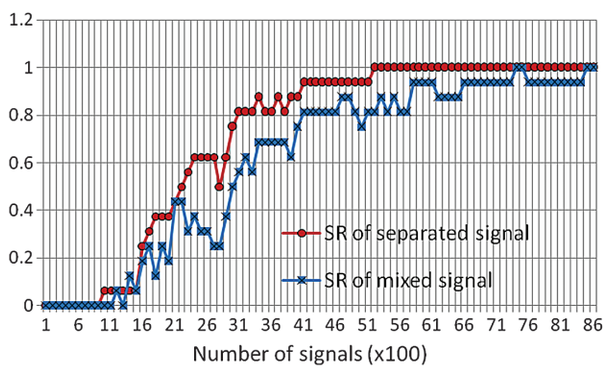

Fig. 5. Success rate (SR) of the mixed signal (three sources) and the separated signal.

Table I. The number of needed signals and correlations for each mixed encryption.

	Original signal		Separated signal		
Mixed type	No.	Corr.	No.	Corr.	Reduction rate (%)
AES1,2, C	Fails	0.0422	5,371	0.0996	46.3
AES1,3, C	7,012	0.0627	4,126	0.1098	41.1
AES1,4, C	6,411	0.0703	3,679	0.1327	42.6
AES2,3, C	Fails	0.0419	5,301	0.0921	47.0
AES2,4, C	9,835	0.0580	5,527	0.0908	43.8
AES3,4, C	7,164	0.0695	4,175	0.1162	41.7
AES1-4, C	8,291	0.0613	4,327	0.1204	47.8

Notes: No.: denotes the number of needed signals; Corr.: maximal correlation coefficients for key revealing; AES1,2, C: denotes the mixed type of "AES1, AES2 and Camellia"; Fails: keys can not be revealed within 10000 signals; Reduction rate: the number of needed signals for separated signal compared with original signal.

It also suggests that the mixed execution of AES0 and AES1 do not have much influence for the result of EMA.

Further experiments confirm this. The number of signals to reveal all the key bytes and the maximal correlation coefficient are listed in Table I. After the application of FastICA, the number of signals has been reduced by 41% at least with the separation of Camellia in all the above cases.

Experiment 3 A Mixture with More Than 3 Encryption Sources. From the hint of Experiment 2, we only need to separate Camellia from the mixed signals of multiple modules of AES executions and Camellia. Five signals, namely AES0-AES4 and Camellia are processed by FastICA, and then the separated Camellia is subtracted. We perform EMA with the resulted mixed signal. The number of signals used to reveal all the key bytes has been reduced 47.8%, which is listed in the last line of Table I.

All the above three groups of experiments indicate the successful application of the proposed ICA to EMA.

6. CONCLUSIONS

The main contribution of this work is that we propose two signal processing techniques and successfully apply them to EMA: bandpass filtering and ICA. When they are used properly, the secret recovery becomes easily. This is confirmed by the experiments of EMA against AES and Camellia implementation on ASIC. Several conclusions are elicited. Bandpass filtering is a general processing technique, which can attenuate the inference

from multiple frequency components. ICA is particularly effective to separate uncorrelated signals, which is fit for the mixed encryption implementations. With ICA, the countermeasure of simultaneous algorithm noise is greatly weakened. The mixed execution of different encryption can be bypassed with signal processing techniques. These results may also provide enlightment for the design of countermeasures. The implementations of mixed AES or AES with Camellia, without other countermeasures, are both vulnerable to side channel attacks, such as EMA. In the future, more advanced signal processing techniques will be investigated and studied. They will be applied to the evaluation of other countermeasures in order to improve the security of cryptographic devices.

Acknowledgment: This work was supported by "Global COE program" of MEXT and CREST of JST in Japan.

References and Notes

1. A. Nikos, L. Luciano, C. Fabio, and P. Davide, *Journal of Low Power Electronics* 6, 607 **(2010)**.
2. A. Cho, *Science* 322, 1322 **(2008)**.
3. K. Gandol, C. Mourtel, and F. Olivier, Eectromagnetic Analysis: Concrete Results, *Proceeding of CHES 2001, LNCS* **(2001)**, Vol. 2162, pp. 251–261.
4. D. Real, F. Valette, and M. Drissi, Enhancing Correlation Electromagnetic Attack Using Planar Near-field Cartography, *Proceeding. of Design, Automation and Test in Europe Conference and Exhibition* **(2009)**, pp. 628–633.
5. T. S. Messerges, E. A. Dabbish, and R. H. Sloan, *IEEE Transactions on Computer* 51, 541 **(2002)**.
6. T. H. Le, C. Servière, J. Cledière, and J.-L. Lacoume, *IEEE Trans. Inf. Forensic Security* 2, 710 **(2007)**.
7. N. Homma, S. Nagashima, Y. Imai, T. Aoki, and A. Satoh, *IEICE Transactions on Fundamentals* E91-A, **(2008)**.
8. E. Brier, C. Clavier, and F. Olivier, Correlation Power Analysis with A Leakage Model, *Proceeding of CHES 2004, LNCS* **(2004)**, Vol. 3156, pp. 16–29.
9. Research Center for Information Security (RCIS) of AIST, Side-Channel Attack Standard Evaluation Board (SASEBO). http://www. rcis. aist. go.jp/special/SASEBO/index-en. html.
10. A. Hyvärinen, *IEEE Transactions on Neural Networks* 10, 626 **(1999)**.

Received: 31 January 2011. Accepted: 30 March 2011.