Spectrum Analysis of Cryptographic Modules to Counteract Side-Channel Attacks

Takeshi Sugawara [#1], Yu-ichi Hayashi [*1], Naofumi Homma [#2], Takaaki Mizuki [*2],
Takafumi Aoki [#3], Hideaki Sone [*3], Akashi Satoh [†1]

[#] *Graduate School of Information Sciences, Tohoku University*
[1, 2]{sugawara, homma}@aoki.ecei.tohoku.ac.jp, [3]aoki@ecei.tohoku.ac.jp
[*] *Cyberscience Center, Tohoku University*
[1, 2, 3] yu-ichi@mail.tains.tohoku.ac.jp, tm-paper@rd.isc.tohoku.ac.jp, sone@isc.tohoku.ac.jp
[†] *National Institute of Advanced Industrial Science and Technology*
[1] akashi.satoh@aist.go.jp

***Abstract*— This paper proposes a spectrum analysis method to identify the frequency band that can be used to perform power analysis attacks against cryptographic modules. The proposed method conducts *Differential Power Analysis* (DPA) in the frequency domain instead of the time domain. The result is then used to identify the frequency bands containing information leakage. The performance of the proposed method is examined through experiments using a field programmable gate array implementation of the standard block cipher *Advanced Encryption Standard*. We show that a noise filter designed using the results of the proposed method can be an effective countermeasure to power analysis.**

Key words: Side-Channel Attack, Cryptographic module, Differential Power Analysis, Noise reduction

I. INTRODUCTION

It is widely known that high frequency currents released from digital circuits cause simultaneous switching noise in the power/ground plane of the printed circuit board (PCB) [1]. As a result, many studies in the field of electromagnetic compatibility (EMC) have been devoted to understanding the mechanism behind the switching noise as well as developing techniques to reduce its effect [2, 3].

On the other hand, cryptography research takes a different view. Switching noise in cryptographic modules (software/hardware implementations of cryptographic algorithms) can be used to compromise the security of these modules. The attack is based on the fact the switching noise actually contains information about the module's behavior. The attack is usually referred as power analysis [4, 5]. It is a type of *side-channel attack*—a class of attack which exploits unintentional information leakage (*side-channel information*) from the cryptographic modules.

Power analysis is considered to be a dangerous threat. The cryptographic algorithms are usually considered to be very secure as they are theoretically very difficult to break. However, the original design of the algorithm does not take into account information leakage from actual implementation. Therefore, the cryptographic modules can be broken by exploiting the side-channel information.

Most countermeasures against power analysis [5] are usually implemented at the LSI and algorithm levels. They provide significant resistance to attack. However, they usually entail a considerable performance overhead. It is also difficult to apply them to modules consisting of off-the-shelf components.

For these reasons, countermeasures at the PCB level are more favorable in some cases. One such countermeasure involves noise reduction techniques that are well studied in the EMC field [2] and have a smaller overhead than conventional methods. The noise reduction is usually implemented in the form of filters that suppress a specific frequency band. Therefore, it is essential for designers to identify the frequency band containing the information leakage.

This paper presents a spectrum analysis method to determine the frequency band containing valuable information. We assume that the relevant frequency band is buried in the spectrum of measured switching noise. The proposed method performs *Differential Power Analysis* (DPA) in the frequency domain (instead of the time domain) and the result is used to identify the frequency band appropriate for power analysis. The performance of the proposed method is demonstrated through DPA experiments with the standard block cipher *Advanced Encryption Standard* (AES) [6] module. We show that the DPA is difficult to perform when the relevant frequency band is removed using digital filters.

II. DIFFERENTIAL POWER ANALYSIS

In most cases, secret information contained in a measured waveform is buried in other noise components. In that case, it is difficult to extract valuable information from a single waveform. DPA, proposed by Kocher et al. [4], extracts information from multiple waveforms based on a statistical

method. Although the term "Differential" means a specific statistical method, DPA represents a broad class of power analyses which utilize multiple waveforms [5].

In the following, we describe an improved DPA called *Correlation Power Analysis* (CPA) [7]. This method is used throughout this paper.

A. Correlation Power Analysis

The flow diagram of CPA is shown in Fig. 1. The *plaintexts* and *secret key* are kept secret from the attacker and the attacker's goal is to retrieve this secret information. When N plaintexts, namely $P_1 \ldots P_N$, are encrypted, the attacker obtains the corresponding ciphertexts $C_1 \ldots C_N$. At the same time, the attacker measures the power waveforms $W_1(t) \ldots W_N(t)$ of the target cryptographic module. In many cases, the power waveform is measured at the power/ground pins of the device in the form of current consumption, using instruments such as a digital oscilloscope. The attacker can calculate power estimates $E_1 \ldots E_N$ from $C_1 \ldots C_N$ based on a predicted partial key and a power model. Assuming that the power model is valid, the power estimates $E_1 \ldots E_N$ and the measured power $W_1(t) \ldots W_N(t)$ are correlated if the partial key prediction is correct. Therefore, the attacker can distinguish a correct prediction from wrong predictions. The candidate with the highest correlation is considered to be the correct key after all the possible partial keys are examined.

The efficiency of CPA can be compared with a *brute force attack* in terms of the size of the search space. We assume AES with a 128-bit key as the target. In a *brute force attack*, an exhaustive search over 2^{128} ($\approx 10^{38.5}$) key candidates is required. This means the attack is practically infeasible. For CPA, the 128-bit key can be determined from an 8-bit partial key (i.e. byte-wise search). Therefore, the space is reduced to $2^8 \times 128/8$ ($\approx 10^{3.6}$) candidates, which can be feasibly searched.

B. DPA in the Frequency Domain

In DPA (including CPA), it is implicitly assumed that the waveforms are captured at the exact moment of the cryptographic computation, so that they are precisely aligned in time. In reality, it is difficult to obtain a trigger signal precisely synchronized to the cryptographic computation. As a result, a displacement error is introduced into the measurements. If the error is significant, it is known that the attack becomes more difficult to perform as it requires more waveforms [5]. Therefore, there are some countermeasures which induce artificial time variations [5]. At the same time, there are some advanced attacks which compensate for the misalignment to defeat such countermeasures [5, 8].

In particular, DPA in the frequency domain was proposed in [9]. This method uses frequency spectra instead of time-domain waveforms. The frequency spectrum is obtained from the measured waveform using the *Discrete Fourier Transform* (DFT). Since the frequency spectrum is invariant under time shifts, the time displacement error can be ignored in the attack. In this work, we employ this method in order to identify the frequency band containing information leakage.

III. Frequency Band Identification Method

In the following, we assume that the target cryptographic module has a 128-bit datapath and can be examined in 8-bit sections. This assumption can be naturally extended to any other datapath length and search length.

The assumption suggests that the attack estimates only 8 bits in the 128-bit datapath. Therefore, the remaining 120 bits act as a noise generator (i.e., algorithmic noise [10]). In the following, we refer to the components of measured waveforms that correlate to the power estimates as *signal*, while the remaining components are called *noise*.

If we assume a designer who has complete information about the system (i.e., circuit architecture and secret information), he or she can obtain more precise power estimation by considering all 128 bits. Then, the designer can calculate the correlation between the waveforms and the precise power estimation, and produce a correlation-time graph. The correlation is more accurate because of the improved signal-to-noise ratio. The calculated correlation can be used to estimate the number of waveforms required for DPA [11]. This method will be referred to as *known-answer evaluation.*

Our proposed method conducts the *known-answer evaluation* in the frequency domain. A flow diagram of the proposed method is shown in Fig. 2. First, the designer calculates precise power estimates $E'_1 \ldots E'_N$ over the 128-bit datapath using the known secret key and obtained ciphertexts $C_1 \ldots C_N$. At the same time, the measured waveforms $W_1 \ldots W_N$ are converted to frequency spectra $S_1(f) \ldots S_N(f)$ using DFT. Finally, the designer calculates the correlation between $E'_1 \ldots E'_N$ and $S_1 \ldots S_N$, producing a correlation-frequency graph. A frequency band with high correlation is considered to be a significant frequency band (i.e., one containing information leakage).

IV. Experiments

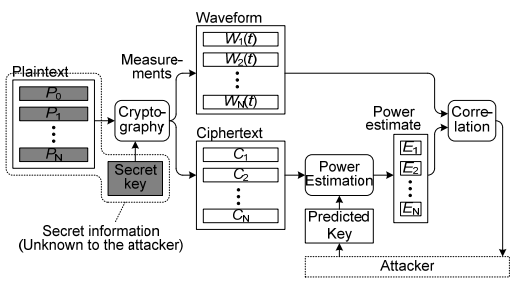

Fig. 1 Flow of the attack.

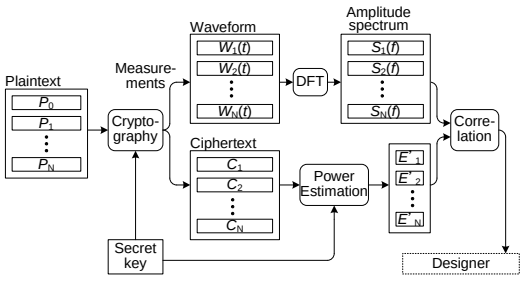

Fig. 2 Flow of the evaluation.

A. Waveform Acquisition

The proposed method is examined through experiment. The experimental setup and conditions are shown in Fig. 3 and Table 1, respectively. The measurement system consists of the *Side-channel Attack Standard Evaluation Board* (SASEBO) [12], a digital oscilloscope, and a PC. SASEBO involves two Field Programmable Gate Arrays (FPGAs), FPGA1 and FPGA2. FPGA1 is configured as the AES circuit [13], while FPGA2 is configured as the control and communication circuits. In these configurations, plaintexts are fed from the PC to FPGA1 via FPGA2. During the encryption, voltage variation on a 1Ω resistor inserted between the ground pin of FPGA1 and the ground plane of the board is measured using the oscilloscope. The oscilloscope is triggered using a signal from the I/O pin of FPGA1. The measurement is repeated for 30,000 different plaintexts, and the corresponding 30,000 waveforms are stored. One example of the measured power waveform and its corresponding frequency spectrum are shown in Figs. 4 and 5, respectively. In Fig. 4, the encryption process starts at around 140 ns and finishes after 11 clock cycles or 458 ns (=11×1/24 MHz). Fig. 4 also shows that the measured voltage rises during the encryption process.

B. Identification of Significant Band

Results of the *known-answer evaluation* and the proposed method (i.e. *known-answer evaluation* in frequency domain) are shown in Figs. 6 and 7, respectively. In these analyses, the correct power estimate is calculated by choosing AES final round as the target and the *Hamming distance model* [7] as the power model. For comparison, a wrong power estimate with wrong key prediction is emulated by randomly shuffling the index of the correct power estimate. Correlations for both power estimates are shown in Figs. 6 and 7. In Fig. 6, a maximum correlation value around 0.35 is achieved by the correct power estimate. On the other hand, the correlation with the wrong power estimate is almost zero. The correlation function starts to rise at around 550 ns and gradually converges to zero. This result agrees with the fact that the AES final round is being processed at this time. In Fig. 7, the maximum correlation value of 0.22 is lower than that of Fig. 6. This is because we applied DFT to the whole waveform $0 < t < 1000$ ns (i.e. 4000-point DFT) in which $0 < t < 550$ ns is not correlated to the power estimate (see Fig. 6). Therefore, components that are uncorrelated to the power estimate are spread over the frequency spectrum, and result in low correlation. Although we can enhance the correlation by applying DFT over a shorter time period, the result in Fig. 7 is sufficient to identify frequency bands with a significant correlation. There is strong correlation for 0-50 MHz, and relatively weak correlation for 50-500 MHz. The proposed method claims that a frequency band with a high correlation contains information related to the cryptographic operation.

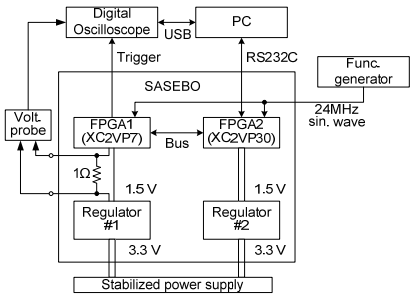

Fig. 3 Experimental setup.

Table 1. Experimental conditions.

	Setup
Oscillo-scope	Agilent MSO6104A @ 4.0 GSa/s
Voltage Probe	Agilent A1130A with SMA probe head (up to 1.5 GHz)
Operating Freq.	24MHz
Num. acquisition	30,000

C. DPA on the Filtered Waveform

In this section, we demonstrate that CPA becomes ineffective when the identified bands are removed. The measured waveforms have a strong correlation for 0-50 MHz and a weak correlation for 50-500 MHz. In order to remove these bands, Low-Pass Filters (LPFs) and High-Pass Filters

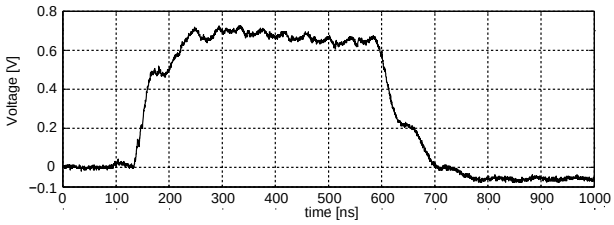

Fig. 4 Acquired waveform.

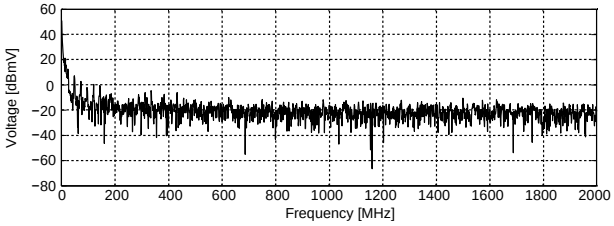

Fig. 5 Frequency spectrum of the acquired waveform.

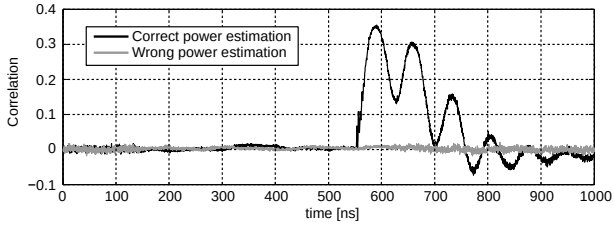

Fig. 6 Correlation in the time domain.

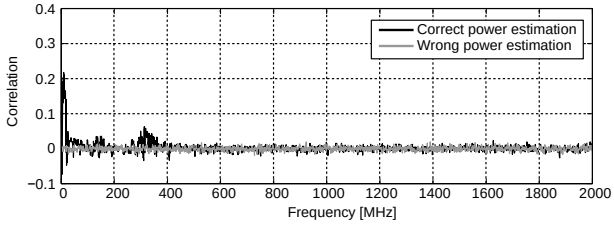

Fig. 7 Correlation in the frequency domain.

(HPFs) were designed and applied to the measured waveforms. We designed filters with cut-off frequencies f_c of 50 MHz (LPF1 and HPF1) and 500 MHz (LPF2 and HPF2). The frequency responses of the filters are shown in Fig. 8. The filters are applied using convolution in both forward and reverse directions to perform a zero-phase filtering [14].

CPA was applied in the time domain to (i) measured waveforms, and waveforms filtered using (ii) LPF1, (iii) HPF1, (iv) LPF2, and (v) HPF2. The results are shown in Fig. 9. The vertical axis represents the number of incorrectly predicted bytes. Since the length of the secret key is 16 bytes (=128 bits), the value ranges from 0 to 16 with 0 indicating the successful extraction of the whole key (i.e. completion of the attack). The results for waveforms of types (i), (ii), and (iv) have almost the same graph and the whole 128-bit key is successfully estimated using about 5,000 waveforms. These three analyses have the 0-50 MHz frequency band in common, which suggests that this low frequency band dominates these CPA results. This result agrees with the interpretation of the proposed method in which the 0-50MHz band shows a strong correlation (see Fig. 7).

If we apply HPF1 to the measured waveforms, the dominating 0-50 MHz band is omitted. By comparing the results of type (iii) waveforms with those of types (i), (ii), and (iv) in Fig. 9, we can see that more waveforms are required to retrieve as many key bytes, and even 30,000 waveforms were not sufficient to extract the whole key. This result clearly shows that CPA becomes difficult to perform after filtering the frequency band with a high correlation. However, the attack is still feasible if more traces are used. This is because the weak correlation for 50-500 MHz becomes dominant when the 0-50 MHz band is filtered out. Finally, not even a partial key could be extracted from waveforms of type (v) in which the 0-500 MHz band is filtered out. From these results, we can confirm that the proposed method successfully identified significant bands and DPA became difficult when those bands are filtered out.

V. Conclusion

A spectrum analysis method to identify significant frequency bands containing available side-channel information is proposed. We conduct CPA in the frequency domain under the known-key condition to identify significant frequency bands. The experimental results showed that DPA becomes ineffective when the significant bands are removed using digital filters. As a result, the proposed method opens up the possibility of designing an effective noise filter to counteract power analysis. We are now conducting further experiments under various conditions (e.g., other devices and clock frequencies) to investigate the characteristics of the significant frequency bands.

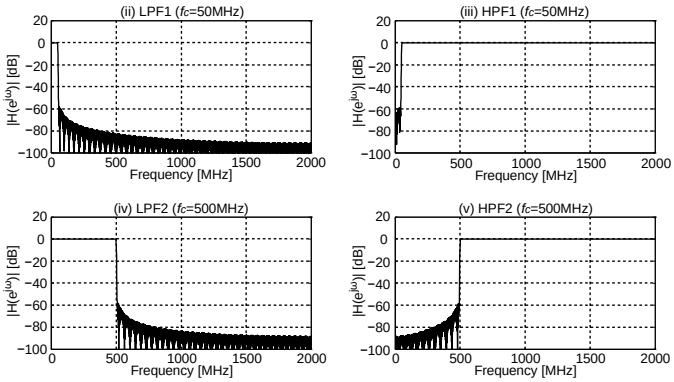

Fig. 8 Normalized frequency responses of the filters.

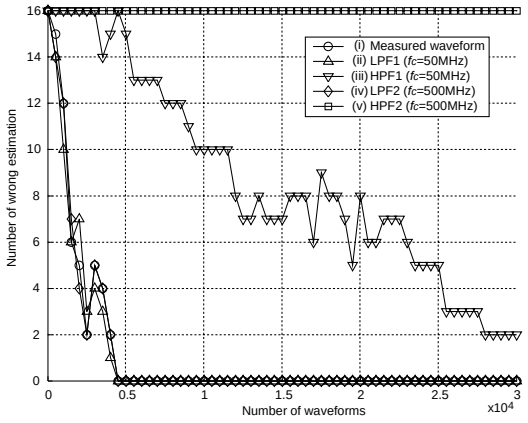

Fig. 9 CPA results.

References

[1] T. Sudo, H. Sasaki, N. Masuda and J. Drewniak, "Electromagnetic interference (EMI) of system-on-package (SOP)," IEEE Trans. Advanced Packaging, 27(2), pp. 304--314, 2004

[2] C. R. Paul, "Introduction to Electromagnetic Compatibility (Wiley Series in Microwave and Optical Engineering)," Wiley-Interscience, 2006

[3] J. Drewniak, F. Sha, T. Van Doren, T. Hubing, J. Shaw, "Diagnosing and Modeling Common-Mode Radiation from Printed Circuit Boards with Attached Cables." Electromagnetic Compatibility, 1995. Symposium Record. 1995 IEEE International pp. 465–470, Aug. 1995.

[4] P. Kocher, J. Jaffe, and B. Jun, "Differential Power Analysis," CRYPTO 1999, Lecture Notes in Computer Science, vol.1666, pp.388 – 397, Aug. 1999.

[5] S. Mangard, E. Oswald, T. Popp, "Power Analysis Attacks: Revealing the Secrets of Smart Cards," Springer-Verlag, 2007.

[6] NIST, "Advanced Encryption Standard (AES) FIPS Publication 197," Nov. 2001.

[7] E. Brier, C. Clavier, and F. Olivier, "Correlation Power Analysis with a Leakage Model," CHES 2004, LNCS 3156, pp. 16-29, 2004.

[8] N. Homma, S. Nagashima, Y. Imai, T. Aoki, and A. Satoh, "High-resolution Side-Channel Attack Using Phase-Based Waveform Matching," CHES2006, LNCS4249, pp. 187-200, 2006.

[9] H.C. Gebotys, S. Ho, and C.C. Tiu, "EM analysis of Rijndael and ECC on a Wireless Java-based PDA," CHES 2005, LNCS, vol.3659, pp.250 – 264, Aug. 2005.

[10] F. -X. Standaert, S. B. Őrs, and B. Preneel, "Power Analysis of an FPGA Implementation of Rijndael: Is Pipelining a DPA Countermeasure?" CHES 2004, LNCS3156, pp. 30-44, Aug. 2004.

[11] R. Watanabe, Y. Takahashi, T. Matsumoto, "Side Channel Attacks from Signal Lines of Cryptographic Modules －Part 2: Detailed Experimental Result－," In proc. CSS2008, Oct. 2008. (in Japanese)

[12] Research Center for Information Security, AIST, "Side-channel Attack Standard Evaluation Board (SASEBO)," http://www.rcis.aist.go.jp/special/ SASEBO/index-en.html

[13] "Cryptographic Hardware Project," Computer Structures Laboratory, Graduate School of Information Sciences, Tohoku University, http://www.aoki.ecei.tohoku.ac.jp/crypto/.

[14] Oppenheim, A.V., and R.W. Schafer, "Problem 5.39," in Discrete-Time Signal Processing, Prentice-Hall, 1989.