# Performance Improvement of the DPA Attack based on Wavelet Denoising

Kyung-Won Song, You-Seok Lee, Hyoung-Nam Kim
*School of Electrical Engineering, Pusan National University, Busan, Korea*
hnkim@pusan.ac.kr

## Abstract

*A Differential Power Analysis (DPA) is known as a powerful method among physical attacks. However, its efficiency tends to be degraded by noise embedded in leaked power signals. To overcome this vulnerability of the DPA attack, we propose a noise-reduction method for the DPA with the use of wavelet denoising. Experimental results show that the proposed denoising method improves the attack efficiency of the DPA dramatically in terms of the number of signals required for the successful attack and reliability.*

**Keywords:** DPA, SCA, power analysis, wavelet denoising.

## 1. Introduction

Cryptographic systems might guarantee the safety mathematically but there have been some threatening attack methods to find an encryption key and steal information. The side channel attack (SCA), which exploits the physical weakness of the cryptographic device, is one of the most popular physical attack methods [1]. SCA methods are classified by the used physical characteristics of leaked signals such as timing information, power consumption, and electromagnetic emanation. The power analysis attacks, which use the power consumption signals as sources, have been actively studied. Among some types of power analysis, the differential power analysis (DPA) attack is a strong SCA method using the statistical characteristics [2] [3].

Although the DPA attack is powerful, the attack efficiency is prone to be degraded by embedded noise. The averaging process of the DPA may reduce the noise effect but lots of signals are still required for a successful attack. Since it is hard and inefficient to get much more signals only for noise reduction, it is desirable to find a pre-processing method to reduce noise effect with the limited number of given signals. To achieve this, we propose a pre-processing method to improve the attack efficiency with the use of the wavelet denoising.

This paper is organized as follows. In section 2, the basic concept of the DPA attack is explained. Section 3 describes the main idea of the proposed method. It focuses on reducing the unwanted noise with maintaining the signal characteristics. Experimental results of the DPA attack with and without the proposed attack method are given in section 4. Finally, in section 5, we conclude the paper.

## 2. Differential Power Analysis

A DPA attack, firstly introduced by P. Kocher et al. [2], is the way to find the cipher key of cryptographic systems by using the statistical characteristics of the power consumption signals. This technique has an advantage that an attack can be achieved without the knowledge of the detail structure about the cryptographic algorithms.

The attack starts by measuring $M$ number of power traces $W_j$ during $i$-th cryptographic operation from each random plaintext input. According to the DPA attack, the selection function $D(P_{lj}, K_j)$ produces a value of 0 or 1 corresponding to the inputs of a plaintext $P_{l,j}$ and a guessing key $K_j$. The power traces are split into two subsets according to the result of $D$. When the output of the selection function is 0, the power traces are classified into subset 0, and in case that the output is 1, those are contained to subset 1. Finally, the difference of the means of each subset is computed as follows:

$$\Delta_{i,j} = \frac{\sum_{l=1}^{M} D(P_{l,i}, K_j) W_l}{\sum_{l=1}^{M} D(P_{l,i}, K_j)} - \frac{\sum_{l=1}^{M} (1 - D(P_{l,i}, K_j)) W_l}{\sum_{l=1}^{M} (1 - D(P_{l,i}, K_j))}. \quad (1)$$

This work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MEST) (No. R01-2008-000-20987-0-(2008)).

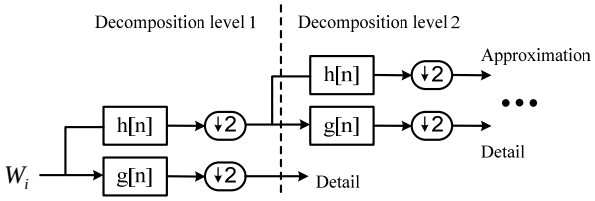


**Fig. 1: Block diagram of wavelet decomposition.**

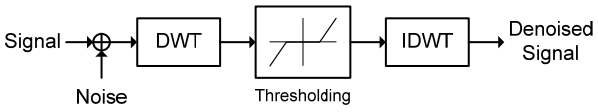


**Fig. 2: Block diagram of wavelet denoising.**

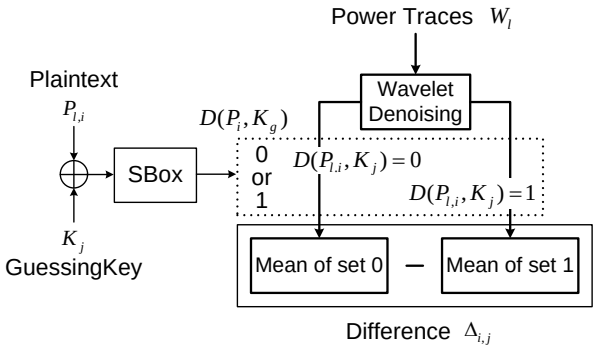


**Fig. 3: Block diagram of the proposed method.**

The basic concept of the DPA attack is that the power consumption manipulating 0-value bit is different from the case of 1-value bit. Therefore, if a correct cipher key is an input of the selection function, high or low power consumption signals are divided into different subset. Then the differential trace $\Delta_{i,j}$, the difference of mean of each subset, will have a meaningful peak named a DPA peak [2], [4]. On the other hand, since the power traces are divided into each subset at random, the difference of the means with a wrong guessing key will be small and no significant peaks will appear in the differential trace. Using this property, we can decide which key is used for encryption by observing that a DPA peak appears or not.

## 3. The Proposed Method

The attack efficiency of the DPA is highly correlated with how fast we can detect the correct key with given traces but the noise embedded in the measured signal disturbs the efficient attack. Therefore, many signals are required to suppress the noise effect in averaging process of the DPA. However, since it is hard to gather the much more signals, we need pre-processing methods to reduce noise effect. To achieve

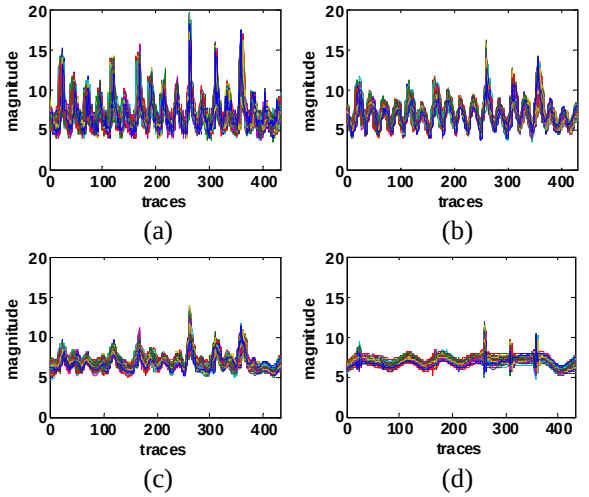


**Fig. 4: (a) Original power traces. (b) denoised signal by level 3. (c) denoised signal by level 4. (d) denoised signal by level 5.**

this, the trivial low-pass filter (LPF) can be adopted for a pre-processing procedure. However, the LPF blurs the peaks which are main components of the measured power traces and it may cause an undesirable loss of the crucial information of the signals.

To maintain the shapes of peaks and reduce the noise effect simultaneously, wavelet denoising can be a good candidate. The wavelet transform (WT) is a time-scale representation technique of a signal [5]. It is implemented as a digital filter by discrete wavelet transform (DWT) which is given as follows:

$$f(t) = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} d(k,l)2^{-k/2}\psi(\frac{t-2^k l}{2^k}), \qquad (2)$$

where $2^k$ and $2^k l$ are dilation and translation parameters that make the multiresolution analysis possible. As shown in Fig. 1, decomposition procedures are implemented with the low-pass filter h[n] and high-pass filter g[n], and filtered samples are down-sampled by 2. These procedures will be iterated as many as we want. Fig. 2 represents the process of getting the denoised signal. Adopting the WT to the power traces, the power traces are divided into approximation and detail components which are the output of the low-pass and high-pass filters, respectively. Unlike the LPF, the peak or edge information of the signal is preserved in detail components which are outputs of $g[n]$. As a result, these characteristics minimize the smoothing effect of the power traces by the LPF. Fig. 3 shows a block diagram of the proposed DPA attack with wavelet denoising. The power traces are treated by wavelet

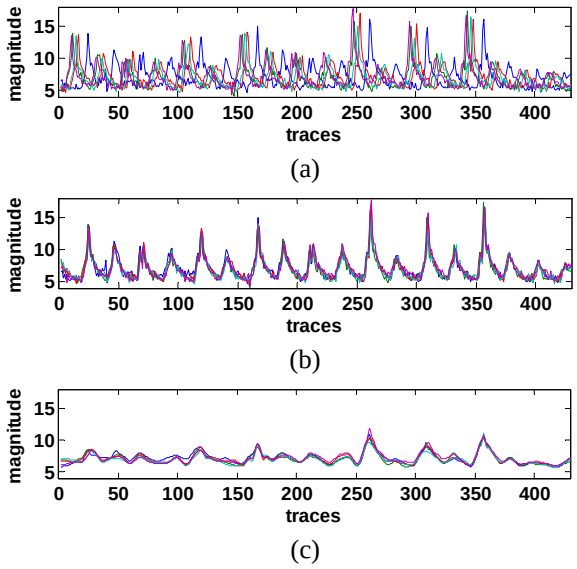


**Fig. 5: (a) Measured power traces. (b) result of alignment. (c) denoised power traces by the proposed method.**

denoising before being distributed into two subsets. After adopting the proposed pre-processing method for attacks, pure power traces are used as side channel signals as minimized a loss of the useful peak information, and they lead to the enhancement of the attack efficiency.

## 4. Experimental Results

To apply the wavelet denoising to the DPA, we should determine several parameters. First, the choice of mother wavelet is 'symlet' wavelet family. It has an order $N$ with filter length of $2N$, and we found the best performance with the order of 6. Next, we choose soft decision and universal thresholding. The universal thresholding takes the value of $\lambda_{UNIV} = \sqrt{2 \ln N}\sigma$, where $N$ and $\sigma$ denote a signal length and a noise variance, respectively [6]. Lastly, we selected the decomposition level. As shown in Fig. 4, we could observe the traces that unwanted peaks are smoothed minimizing the distortion of high peaks at the decomposition level 4.

**Table 1: The number of minimum traces for successful attack**

| SBox | Original | Proposed | SBox | Original | Proposed |
|---|---|---|---|---|---|
| 1 | 3,997 | 1,655 | 9 | 3,841 | 2,541 |
| 2 | 3,420 | 1,973 | 10 | Fail | 923 |
| 3 | 3,805 | 1,645 | 11 | Fail | 2,930 |
| 4 | 3,921 | 1,518 | 12 | Fail | 1,101 |
| 5 | 3,999 | 1,606 | 13 | 1,547 | 842 |
| 6 | Fail | 1,227 | 14 | Fail | 1,553 |
| 7 | 2,850 | 2,159 | 15 | Fail | 1,250 |
| 8 | 3,130 | 2,106 | 16 | 1,400 | 1,009 |

To evaluate the proposed method, we measured 4,000 power traces during the first round operation based on the Advanced Encryption Standard (AES) [7]. At first, it is necessary to align the signal against the embedded countermeasure like a random clock mechanism. Several numbers of traces are listed in Fig. 5 (a), and it shows the misalignment problem. To solve this problem, we divide the 1st power trace into 16 partitions corresponding to the each SBox operation. Then, we align the rest of 3,999 power trace by using the correlation with 1st one. Fig. 5 (b) shows the result of the alignment, and in fig. 5 (c) we can see the power traces after the wavelet denoising. To verify the performance of the proposed method, the DPA attack was carried out for each 16 SBox, respectively. Also, the DPA attack with the wavelet denoising was carried out repeatedly.

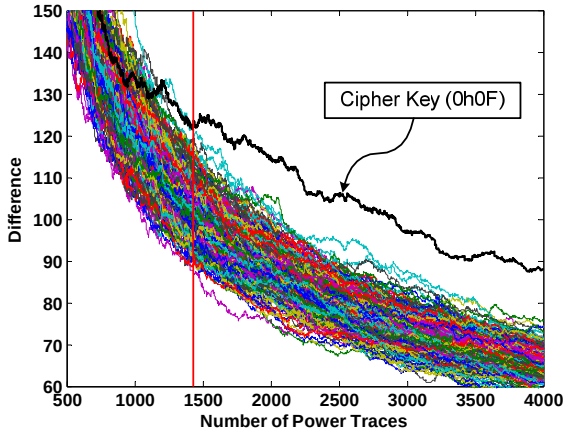


**Fig. 6: Difference curves of original DPA attack.**

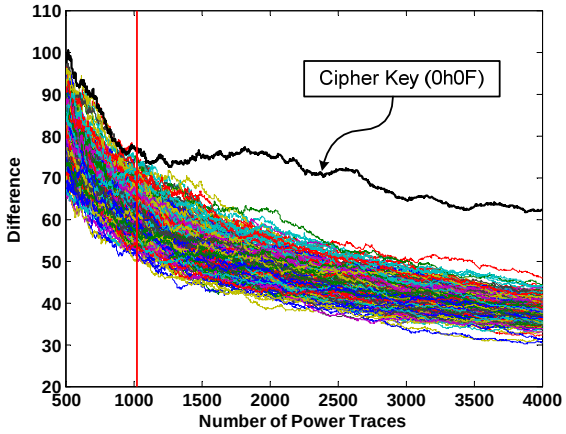


**Fig. 7: Difference curves of proposed method.**

The experimental results are listed in Table 1. The original DPA attack failed to find the cipher keys at 6 SBoxes within 4,000 traces. Besides, the results of 1st and 5th SBoxes, where the numbers of the used traces were nearly 4,000, might not guarantee the stable success in attacking. On the other hand, in the proposed method based on the wavelet denoising, all cipher keys were detected with much smaller traces used for the attack. Figs. 6 and 7 are the difference curves of 16th SBox with the original and the proposed attacks. After a certain point which marked as a red bar, the difference of the mean with the correct guessing key (0h0F) becomes outstanding, compared with left 255 guessing keys. Comparing with two figures we can be sure the performance improvement of the proposed method.

We also analyze the result in terms of reliability. We compute the 1st and 2nd maximum difference ratio (MDR) listed below:

$$\mathrm{MDR}_i = \frac{\max_{tr}(\Delta_{i,j}) - \max 2_{tr}(\Delta_{i,j})}{\max_{tr}(\Delta_{i,j})}, \qquad (3)$$

where subscript *tr* is the number of used traces, and $\max 2$ means the 2nd maximum value. If the MDR is large, we can easily find the cipher key because it means that a remarkably larger difference value exists certainly. Finally, the figure of merits (FOM) is derived the ratio of the proposed MDR to the original MDR. If the FOM is larger than 1, we simply evaluate that the proposed method manifests the high performance.

$$\mathrm{FOM}_i = \frac{\mathrm{MDR}_{i,proposed}}{\mathrm{MDR}_{i,origianl}}. \qquad (4)$$

Some results of the reliability with SBox 13 and 16 are listed in Table 2.

**Table 2: Figure of merits for the reliability**

| Traces | Figure of Merits | |
|---|---|---|
| | SBox 13 | SBox 16 |
| 2,000 | 2.7792 | 2.1612 |
| 2,500 | 2.7402 | 2.3012 |
| 3,000 | 1.9676 | 2.2375 |
| 3,500 | 1.5415 | 1.8819 |
| 4,000 | 1.5452 | 1.8712 |

Since the all investigated FOMs are larger than 1, it makes easy to determine the cipher key when we use the proposed method.

## 5. Conclusion

We proposed a noise-reduction method to improve the DPA attacks. By adopting the wavelet denoising we could reduce noise embedded in the signals with the minimization of meaningful information loss, resulting in the performance improvement in terms of the required number of traces or signals. The proposed method focuses on the pre-processing which applies only for the power traces. Therefore it can be helpful for the other side channel attack method, such as frequency-domain based DPA [8] or correlation power analysis [9].

Though we focused on stealing of hidden keys, it is expected that the proposed method could contribute to the study on developing countermeasures for a reliable encryption for security by finding out the inherent weaknesses of the cryptographic devices.

## References


[1] Paul C. Kocher, "Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems," *Advances in Cryptology-Crypto*. 1996, LNCS 1109, pp 104-113

[2] P. Kocher, J. Jaffe, and B. Jun, "Differential Power Analysis," *in Proc. Advances in Cryptology (CRYPTO '99)*, pp. 388-397, 1999

[3] R. Bevan, E. Knudsen "Ways to Enhance Differential Power Analysis," *In proceedings of ICISC 2002*, LNCS 2587, pp.327-342, Springer-Verlag, 2003.

[4] T. S. Messerges, E. A. Dabbish and R. H. Sloan, "Examining Smart-Card Security under the Threat of Power Analysis Attacks," *IEEE Trans. on Computers*, vol. 51, N5, pp. 541-552, May 2002

[5] I. Daubechies, "Where do wavelets come from? – A personal point of view," *in Proc. of the IEEE*, Vol.84, No.4, April, 1996

[6] D. L. Donoho, I.M. Johnstone, G. Kerkyacharian, and D. Picardi, "Wavelet Shrinkage: Asymptopia?" *Journal of the Royal Statistical Society*, B, vol. 57, pp. 301-369, 1995

[7] Advanced Encryption Standard (AES), Federal Information Processing Standards Publication 197, 2001.

[8] C. Gebotys, S. Ho. And C.C. Tiu, "EM Analysis of Rijndael and ECC on a Wireless Java-Based PDA," *in Proceedings of CHES 2005*, LNCS 3659, pp. 350-264, Springer-Verlag, 2005.

[9] E. Brier, C. Clavier, and F. Olivier, "Correlation power analysis with a leakage model," *in Proceedings of CHES 2004*, LNCS 3156, pp. 16-29, 2004.