High-Resolution Side-Channel Attack Using Phase-Based Waveform Matching

Naofumi Homma[†], Sei Nagashima[†], Yuichi Imai[†], Takafumi Aoki[†], and Akashi Satoh[‡]

[†]Graduate School of Information Sciences, Tohoku University
6-6-05, Aramaki Aza Aoba, Aoba-ku, Sendai-shi 980-8579, Japan
{homma, nagasima, imai}@aoki.ecei.tohoku.ac.jp, aoki@ecei.tohoku.ac.jp
[‡]IBM Research, Tokyo Research Laboratory, IBM Japan, Ltd.
1623-14 Shimo-tsuruma, Yamato-shi, Kanagawa, 242-8502, Japan
akashi@jp.ibm.com

Abstract. This paper describes high-resolution waveform matching based on a Phase-Only Correlation (POC) technique and its application for a side-channel attack. Such attacks, such as Simple Power Analysis (SPA) and Differential Power Analysis (DPA), use a statistical analysis of signal waveforms (e.g., power traces) to reduce noise and to retrieve secret information. However, the waveform data often includes displacement errors in the measurements. The use of phase components in the discrete Fourier transforms of the waveforms makes it possible to estimate the displacements between the signal waveforms with higher resolution than the sampling resolution. The accuracy of a side-channel attack can be enhanced using this high-resolution matching method. In this paper, we demonstrate the advantages of the POC-based method in comparison with conventional approaches through experimental DPA and Differential ElectroMagnetic Analysis (DEMA) against a DES software implementation on a Z80 processor.

Keywords: side-channel attacks, DPA, DEMA, cryptographic module, waveform matching, phase-only correlation.

1 Introduction

Cryptanalysis based on side-channel information is of major concern for the evaluation of tamper-resistant devices. When a cryptographic module performs encryption or decryption, secret parameters correlated to the intermediate data being processed can be leaked via power dissipation [1], electromagnetic radiation [2], or operating times as side-channel information. These are now essential issues for designers of smartcards and other embedded cryptosystems.

In general, a side-channel attack requires a statistical analysis of waveforms (e.g., power traces) to reduce noise and to retrieve secret information. The important assumption here is that each waveform is captured by a digital measuring device at the exact moment as the corresponding cryptographic computation. However, it is almost impossible to time exactly when the data was captured for cryptographic modules in actual applications, because there is no trigger signal precisely synchronized with the cryptographic computation. For example, wireless devices and smartcards often have no internal clock generator, or devices using PLLs will not have any external clock synchronized with the internal clock. Even if a trigger signal is available, it often contains jitter-related deviations from the true timing of the encryption process. As a result, the measured waveforms always include displacement errors. The displacement errors are usually smaller than the sampling interval, but may cause significant loss of the secret information when the waveforms are averaged together, unless there is exact alignment during the statistical analysis.

Some approaches dealing with the displacements in waveforms were proposed [3], [4]. In a theoretical model, Differential Power Analysis (DPA) with the fast Fourier transform of the power waveforms is introduced to correct the displacement errors [3]. Reference [4], on the other hand, demonstrated a practical approach to analyze Rijndael and ECC on a Java-based wireless PDA. The reported methods were performed in the frequency domain, and thus it would be very difficult to use them in collaboration with other side-channel attacks in the time domain.

Addressing the displacement problem, we propose a high-resolution waveform matching method using a Phase-Only Correlation (POC) function. POC techniques have been successfully applied to high-accuracy image matching tasks [5]-[8]. The POC function employs phase components in the discrete Fourier transforms of waveforms, and makes it possible to determine displacement errors between signal waveforms with high noise tolerance by using the location of the correlation peak. By fitting the analytical model of the correlation peak to the actual numerical data, we can evaluate the displacement errors with a higher resolution than the sampling resolution. The waveform matching can be available directly for a wide variety of side-channel attacks in the time domain against real-world applications.

In this paper, we describe a high-resolution side-channel attack using POC-based waveform matching, and demonstrate its advantages in comparison with conventional methods through experimental analysis of DPA and Differential ElectroMagnetic Analysis (DEMA) against a DES software implementation on a Z80 processor. The essence of the proposed method is to use the POC-based waveform matching as a preprocessing step followed by standard analysis. In this experiment, the side-channel information is monitored with a digital oscilloscope for various sampling rates. The differential analysis with the POC-based matching shows better results in comparison with the conventional attacks for all of the sampling rates.

2 High-resolution waveform matching using Phase-Only Correlation

2.1 Phase-based waveform matching

Consider two signal waveforms, $f(n)$ and $g(n)$, where we assume that the index range is $n = -M, \cdots, M$ for mathematical simplicity, and hence the length of waveforms $N = 2M + 1$. Let $F(k)$ and $G(k)$ denote the Discrete Fourier Transforms (DFTs) of the two waveforms. $F(k)$ and $G(k)$ are given by

$$F(k) = \sum_{n=-M}^{M} f(n) W_N^{kn} = A_F(k) e^{j\theta_F(k)}, \tag{1}$$

$$G(k) = \sum_{n=-M}^{M} g(n) W_N^{kn} = A_G(k) e^{j\theta_G(k)}, \tag{2}$$

where $W_N = e^{-j\frac{2\pi}{N}}$, $A_F(k)$ and $A_G(k)$ are amplitude components, and $e^{j\theta_F(k)}$ and $e^{j\theta_G(k)}$ are phase components.

The cross-phase spectrum (or normalized cross spectrum) $R_{FG}(k)$ is defined as

$$R_{FG}(k) = \frac{F(k)\overline{G(k)}}{\left|F(k)\overline{G(k)}\right|} = e^{j\theta_{FG}(k)}, \tag{3}$$

where $\overline{G(k)}$ denotes the complex conjugate of $G(k)$ and $\theta_{FG}(k) = \theta_F(k) - \theta_G(k)$. The POC function $r_{fg}(n)$ is the Inverse Discrete Fourier Transform (IDFT) of $R_{FG}(k)$ and is given by

$$r_{fg}(n) = \frac{1}{N} \sum_{k=-M}^{M} R_{FG}(k) W_N^{-kn}. \tag{4}$$

If there is a similarity between two waveforms, the POC function gives a distinct sharp peak. (When $f(n) = g(n)$, the POC function becomes the Kronecker delta function.) If not, the peak drops significantly. The height of the peak can be used as a good similarity measure for waveform matching, and the location of the peak shows the translational displacement between the two waveforms.

Now consider $f_c(t)$ as a waveform defined in continuous space with a real number index t. Let δ represents a displacement of $f_c(t)$. So, the displaced waveform can be represented as $f_c(t - \delta)$. Assume that $f(n)$ and $g(n)$ are spatially sampled waveforms of $f_c(t)$ and $f_c(t - \delta)$, and are defined as

$$f(n) = f_c(t)|_{t=nT}, \tag{5}$$

$$g(n) = f_c(t - \delta)|_{t=nT}, \tag{6}$$

where T is the sampling interval and the index range is given by $n = -M, \cdots, M$. For simplicity, we assume $T = 1$. The cross-phase spectrum $R_{FG}(k)$ and the POC function $r_{fg}(n)$ between $f(n)$ and $g(n)$ will be given by

$$R_{FG}(k) = \frac{F(k)\overline{G(k)}}{\left|F(k)\overline{G(k)}\right|} \simeq e^{j\frac{2\pi}{N}k\delta}, \tag{7}$$

$$r_{fg}(n) = \frac{1}{N} \sum_{k=-M}^{M} R_{FG}(k) W_N^{-kn}$$

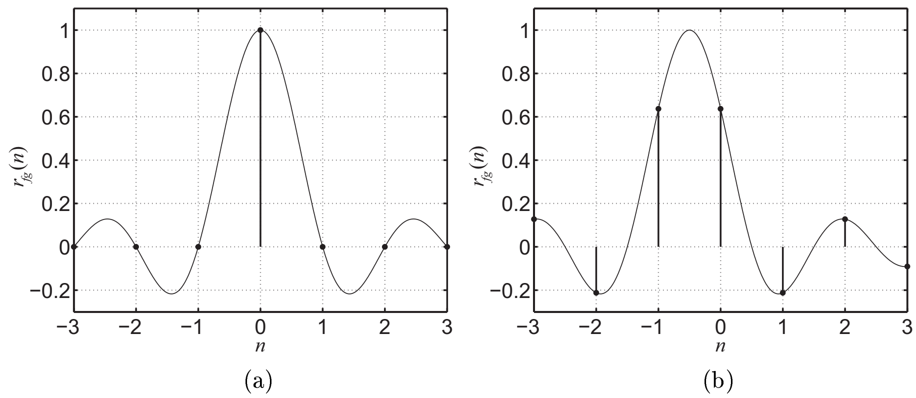


Fig. 1. POC functions: (a) for the case $\delta = 0$, (b) for the case $\delta = 0.5$.

$$\simeq \frac{\alpha}{N} \frac{\sin \left\{ \pi \left(n + \delta \right) \right\}}{\sin \left\{ \frac{\pi}{N} \left(n + \delta \right) \right\}}, \tag{8}$$

where $\alpha = 1$. The above Eq. (8) represents the shape of the peak for the POC function between the same waveforms that are slightly displaced with each other. This equation gives a distinct sharp peak. The peak position δ of the POC function corresponds to the displacement between the two waveforms. We can prove that the peak value α decreases (without changing the shape of the function itself), when small noise components are added to the original waveforms. Hence, we assume $\alpha \leq 1$ in practice. For the waveform matching task, we evaluate the similarity between the two waveforms by the peak value α, and estimate the displacement by the peak position δ.

By calculating the POC function for two waveforms $f(n)$ and $g(n)$, we can obtain a numerical value of $r_{fg}(n)$ for each discrete index n, where $n = -M, \cdots, M$. Fig. 1 shows the POC functions around the correlation peaks when (a) $\delta = 0$ and (b) $\delta = 0.5$, where the black dots indicate the discrete data values from $r_{fg}(n)$. We use Eq. (8) (the closed-form peak model of the POC function) directly for estimating the peak position by function fitting. Fig. 1 also shows these examples, where the solid lines represent the estimated shapes of the POC functions. Thus, it is possible to find the location of the peak that may exist between sampling intervals by fitting the peak model to the calculated data around the correlation peak, where α and δ are fitting parameters. Note here that we can use other types of functions, such as a Gaussian function or a quadratic function, for the function fitting.

2.2 Preliminary evaluation

Consider two waveforms $f(n)$ and $g(n)$, and an estimated displacement δ. The waveform matching finally calculates $g'(n)$ by shifting $g(n)$ by an amount corresponding to δ. For example, this waveform shifting is done by the phase rotation

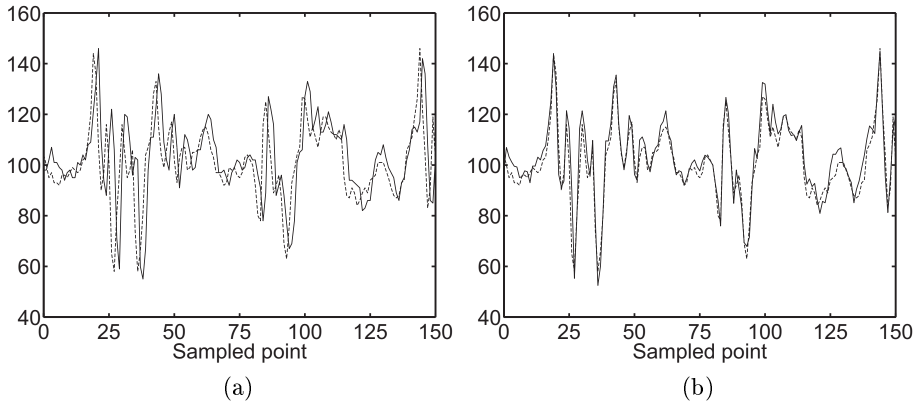


Fig. 2. Example of POC-based waveform matching: (a) input waveforms $f(n)$ and $g(n)$, (b) $f(n)$ and displacement-normalized waveform $g'(n)$.

of the waveform in the frequency domain. Let $G'(k)$ denotes the DFT of $g'(n)$. $G'(k)$ will be given by

$$G'(k) \simeq G(k)e^{j\frac{2\pi}{N}k\delta}. \tag{9}$$

Therefore, $g'(n)$ is given by

$$g'(n) = \frac{1}{N}\sum_{k=-M}^{M} G'(k)W_N^{-kn}. \tag{10}$$

We can also implement the waveform shifting with various interpolation techniques, such as bicubic interpolation.

Fig. 2 shows an example of the POC-based waveform matching, where the two waveforms are power traces from a microprocessor captured by using a trigger signal at the times of the same computation. Due to the trigger jitter, there is a displacement error between these waveforms as shown in Fig. 2(a). Using the POC-based waveform matching, we can obtain the displacement $\delta = 1.5555$. Fig. 2(b) shows two waveforms after the waveform shifting. Thus, the proposed method can be used to match waveform positions with higher resolution than the sampling resolution.

Fig. 3 shows examples of the POC function and the ordinary correlation function, where we use the two waveforms shown in Fig. 2. We observe that the POC function provides a sharp peak in comparison to the ordinary function. The sharp peak typically exhibits good discrimination properties.

To evaluate the sharpness of a correlation peak, we consider the Peak-to-Sidelobe Ratio (PSR) between a central region around at the peak and the residual region (sidelobe region). PSR is determined as PSR $= (peak - mean)/std$, where *peak* is the correlation peak value, and *mean* and *std* are the mean and standard deviation in the sidelobe region [9]. In this example, the PSR values of

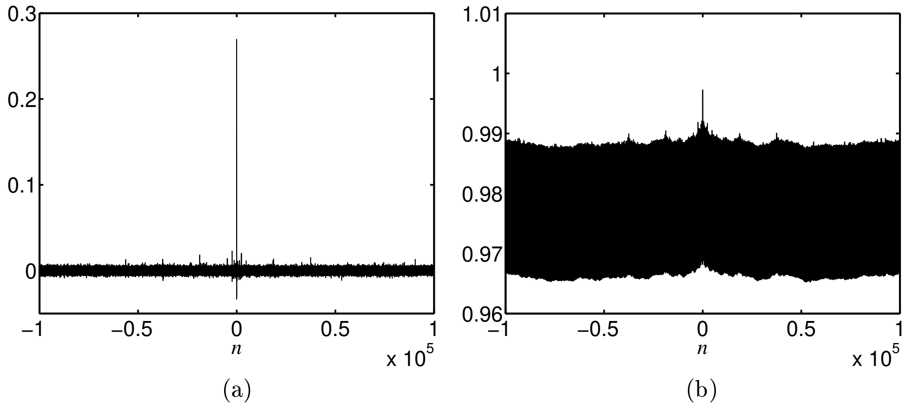


Fig. 3. Examples of the POC function and the ordinary correlation function between two similar waveforms: (a) POC function, (b) ordinary correlation function.

the POC function and the ordinary function are 104.76 and 6.54, respectively. This suggests that the POC function exhibits much higher discrimination than the ordinary correlation function.

3 Side-channel attacks with POC-based waveform matching

3.1 Basic concept

The proposed waveform matching is used as a preprocessing step followed by standard analysis. Fig. 4 shows an overview of the proposed DPA with the POC-based waveform matching. We first collect power traces by sampling power consumption for a series of encryptions of different plaintexts. In the following experiment, a trigger signal is used as in the conventional DPAs for simplicity, and the measured waveforms are initially aligned at the trigger. However, the proposed method can get the same alignment of the waveforms without using a trigger signal. After gathering a number of power traces, we use the POC-based matching for the precise alignment of the waveforms. In the matching step, we select any one of the waveforms as a reference, and then evaluate the displacement errors between the other waveforms and the reference. The POC-based matching considered here includes the advanced techniques described in Appendix A. Finally, we resample each waveform according to the evaluated displacement.

After the waveform matching using POC, the standard analysis is performed. First, we divide the waveforms into two groups according to one bit output from a selection function calculated by guessing the secret key. If the guess is correct, a noticeable difference is found between the two averaged waveforms, but no significant difference appears for a wrong guess that gives no correlation between the selection function and the secret key.

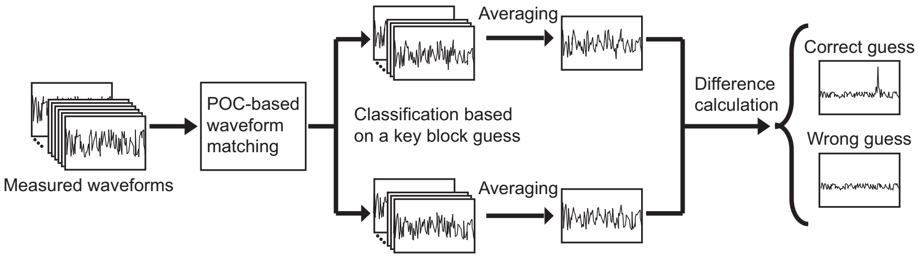


Fig. 4. Proposed differential analysis with POC-based waveform matching.

3.2 Experimental conditions

We applied the POC-based matching technique to DPAs and to DEMAs against a DES software implementation on a Zilog Z80 processor (8 MHz). For the selection functions, we focus on the S-box computation in the 16-th (final) round. DES has eight 6-bit-input and 4-bit-output S-boxes, and thus $4 \times 8 = 32$ selection functions using the S-box output can be formed. For each selection function, we have $2^6 = 64$ key candidates derived from the 6-bit S-box input.

Fig. 5 shows the INSTAC-8 CPU board [10] designed for the side-channel attack experiment, and the measurement points on the board. The power consumption of the processor was monitored as the voltage drop caused by a resistor inserted between the Z80 ground pin and the ground plane of the board. The electromagnetic radiation was also monitored over the Z80 processor as illustrated in Fig. 6. We used a trigger signal synchronized with the beginning of round 15, and obtained four sets of waveforms at sampling rates of 100 MSa/s (millions of samples per second), 200 MSa/s, 400 MSa/s, and 1 GSa/s. Fig. 7 is the measured waveform at 400 MSa/s. The capture range of waveforms is from 4.22 ms to 4.24 ms after the trigger signal, which contains all of the operations of S-box 1 to S-box 8. Two sets of 1,000 waveforms (power and electromagnetic) were measured during encryption of 1,000 random plaintexts with a fixed key. The subkey values from S-box 1 to S-box 8 at the round 16 were fixed as 21, 16, 31, 35, 9, 51, 51, and 48 in decimal, respectively.

3.3 Experimental results

Fig. 8 shows the evaluated displacements for the 1,000 waveforms sampled at 200 MHz (5 ns/point), where the horizontal and vertical axes indicate the waveform index and the displacement value, respectively. The waveforms contain relatively large displacement errors even though they were captured by using a trigger signal synchronized to the system clock, which was generated by the board. Fig. 9 shows the correlation peaks between each waveform and a reference. The peak between two waveforms was about 0.2 due to the different plaintexts and noise. (The peak between identical waveforms is 1.) However, we can identify the peak position clearly since the POC function gives a distinct sharp peak as shown

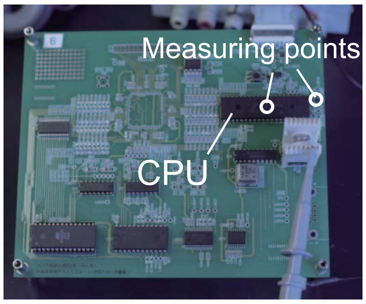


Fig. 5. Evaluation board (INSTAC-8).

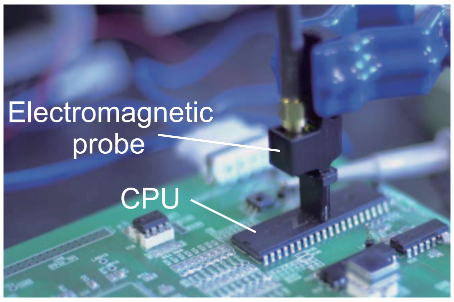


Fig. 6. Electromagnetic probing.

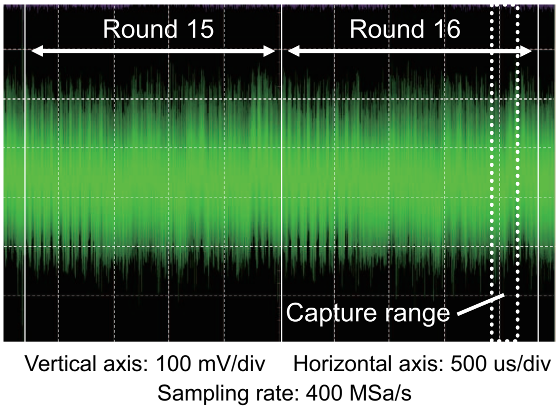
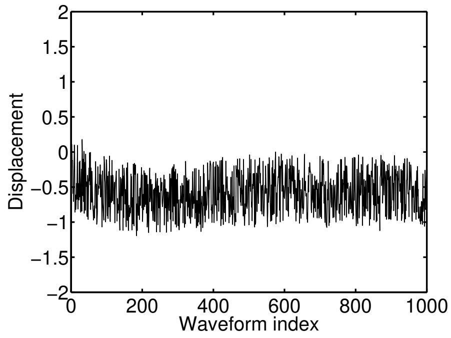


Fig. 7. Example of measured waveform.

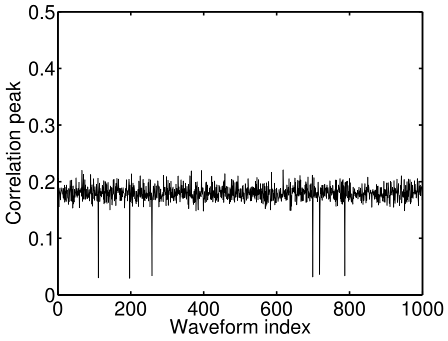
in Fig. 10(a). Fig. 9 also shows that there are some small peaks among the waveforms. Fig. 10(b) shows one of the corresponding POC functions. We found that the waveforms at a low peak value were quite different in shape from the reference waveform. The proposed POC-based analysis can easily detect this kind of inaccurately measured waveform, and thus adverse effects on the statistical analysis can be prevented by removing them in an averaging process.

Fig. 11 illustrates the results of the conventional DPA and the proposed DPA, where both DPAs have used the same set of waveforms sampled at 200 MHz. These results were obtained by evaluating 64 possible keys with one out of the four selection functions of S-box 1. When the DPA succeeds, the highest peak appears in the averaged waveform indicating the correct key, but the conventional DPA in Fig. 11(a) gives many high false peaks for incorrect keys. In contrast, the proposed DPA clearly indicates the true peak with the correct key as shown in Fig. 11(b). In this experiment, we confirmed that the proposed DPA consistently increased the peak signal and reduced the noise at all four of the sampling rates, 100 MSa/s, 200 MSa/s, 400 MSa/s, and 1 GSa/s.



(a)

Fig. 8. Estimated displacements.



(b)

Fig. 9. Correlation peak values.

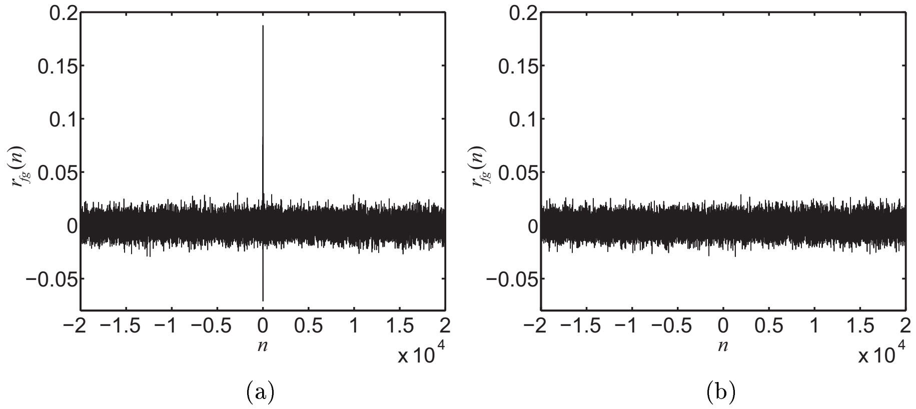


Fig. 10. POC functions: (a) for the case $\alpha \approx 0.2$, (b) for the case $\alpha \approx 0.02$.

Fig. 12 compares the error rates of the conventional DPA and those of the proposed DPA for different numbers of waveforms, where the vertical axis indicates the number of error bits. In other words, Fig. 12 shows the number of selection functions that could not distinguish a correct key from the incorrect keys by investigating the highest peak. If no secret key bit was obtained, the number of errors is 32 bits. The error rate comparisons between the conventional and proposed DEMA are also shown in Fig. 13. The sampling rate of 1 GSa/s is high enough to attack the slow 8-MHz processor by using conventional DPA and DEMA, but the proposed method has clear computational advantages at the sampling rates of 200 MSa/s and 400 MSa/s as shown in Figs. 12 and 13. For example, the proposed DEMA at 200 MSa/s requires less than half the number of waveforms to achieve the 50 % error rate in comparison with the conventional DEMA.

Tables 1 and 2 show the DPA results using 1,000 waveforms at 200 MSa/s and 400 MSa/s, respectively. The 4-bit output from each S-box S_i was used for four

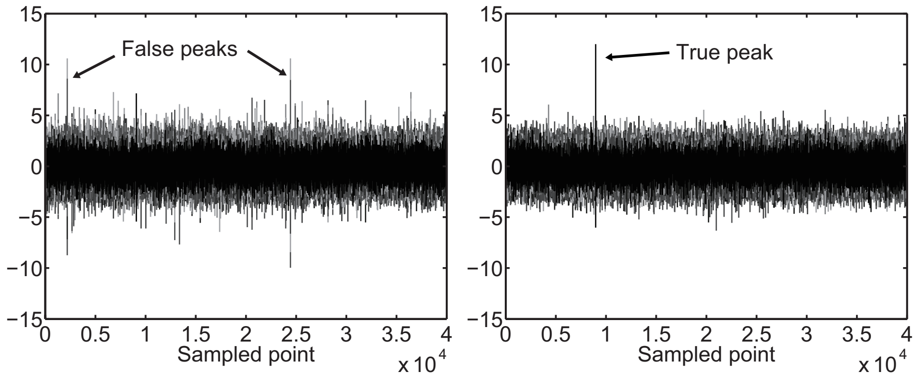


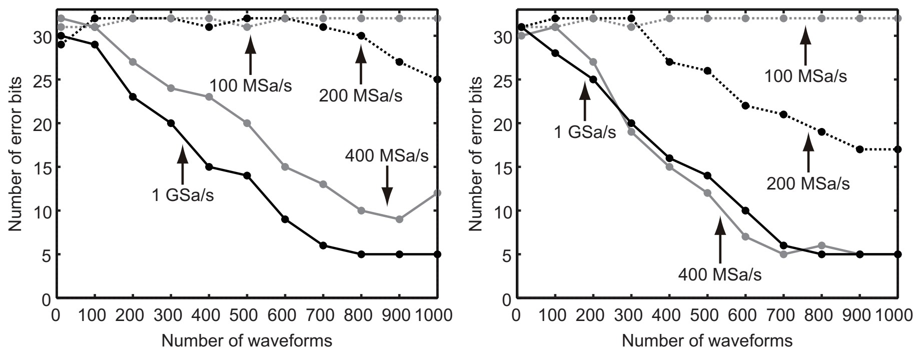
Fig. 11. Example of DPAs: (a) conventional DPA, (b) proposed DPA.

selection functions, and thus four estimations were made for each 6-bit subkey that was XORed to 6 bits of S-box input data. Therefore, four 6-bit possible keys were obtained for each S-box in the tables, and the shaded boxes indicate the correctly guessed keys. If two or more of the values in an S-box column are the same, then there is a very high probability that we can obtain the correct subkey by majority vote. As shown in Table 1, the proposed DPA found five out of eight subkeys at 200 MSa/s while the conventional approach found none. In the 400 MSa/s measurements, the proposed DPA determined all of the subkeys, as shown in Table 2. It is important to note that both DPAs used exactly the same waveform data, and the POC pre-process is simply applied to the captured waveforms before statistical analysis. Therefore, our method can be applied to many varieties of side-channel attacks, such as SPA, SEMA, DPA, and DEMA, to improve the precision of the key estimations.

4 Conclusions

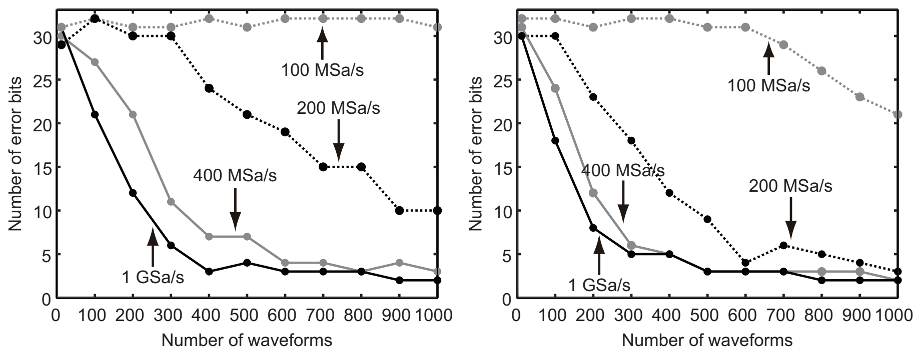
In this paper we proposed a high-resolution waveform matching method based on the POC technique and described its application to side-channel attacks. The POC-based matching method makes it possible to evaluate the displacement between signal waveforms with higher resolution than the sampling resolution. In addition to the waveform alignment, we can remove inaccurately measured waveforms by detecting significant drops of correlation peaks, eliminating their adverse effects on the statistical analysis.

The advantage of proposed method over the conventional methods was demonstrated by experimental DPAs and DEMAs against a DES software implementation. The results showed that the proposed method improved the accuracy of the differential analysis. A high success rate of finding correct subkeys was obtained even at a low sampling rate where the conventional attacks failed. At higher sampling rates, the proposed analysis requires fewer plaintexts to obtain



(a) (b)

Fig. 12. Error rates for various sampling rates: (a) conventional DPA, (b) proposed DPA.



(a) (b)

Fig. 13. Error rates for various sampling rates: (a) conventional DEMA, (b) proposed DEMA.

the same error rate as the conventional analysis. As a result, we confirmed that the POC-based waveform matching can be used efficiently for both DPA and DEMA on software implementations without any drawbacks. Applications to other side-channel attacks and on other platforms such as FPGAs and ASICs remain for future study.

The important feature of the proposed method is its capability to enhance any side-channel analysis independently of the cryptographic algorithms, implementations (software or hardware), and kind (power or electromagnetic) of side-channel information. In experiments with a Z80 board, we used a trigger signal synchronized with the cryptographic operations for simplicity, but the POC-based matching does not require this for aligning a number of power traces. Therefore, our approach is very efficient for attacking cryptographic modules in

Table 1. Estimation results of DPAs at 200 MSa/s

Conventional DPA							
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
21	16	31	35	5	51	51	48
51	5	44	33	27	54	32	11
15	22	43	58	33	13	26	60
8	45	54	14	11	11	60	20

Proposed DPA							
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
21	16	31	35	9	51	51	48
38	16	31	21	12	49	51	48
11	25	53	58	2	26	8	7
10	16	31	14	50	51	19	20

Table 2. Estimation results of DPAs at 400 MSa/s

Conventional DPA							
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
21	16	31	35	9	51	51	48
21	16	31	53	47	51	51	39
11	25	31	35	9	26	8	51
55	16	31	14	9	32	51	36

Proposed DPA							
S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
21	16	31	35	9	51	51	48
21	16	31	35	47	51	51	48
21	25	31	35	9	26	8	48
21	16	31	14	9	51	51	48

actual applications, even where no trigger signal or no internal clock can be observed.

In addition, the proposed method can defeat some countermeasures creating distorted waveforms with random delays, dummy cycles, or unstable clocking. In this paper, the proposed waveform matching was used for relatively long waveforms, in which the number of sample points is from 20,000 to 200,000. Even when the waveforms have small numbers of sample points, the POC-based technique exhibits good discrimination properties. For example, high-accuracy block matching of small images (e.g., 33×33 pixels) have been implemented using the POC-based techniques [11]. The block matching technique can also be effective for waveforms. Thus, POC-based block matching would easily cancel out distortion components in waveforms. We are now conducting research to develop advanced waveform matching techniques and to investigate their utility in attacks against the waveform-distortion countermeasures.

Acknowledgments. We would like to thank Prof. M. Yamaguchi of Tohoku University for his important advice about electromagnetic measurements. We would also like to thank K. Degawa for his generous support.

References

1. P. Kocher, J. Jaffe, and B. Jun, "Differential power analysis," *CRYPTO 1999, Lecture Notes in Computer Science*, vol. 1666, pp. 388 – 397, Aug. 1999.
2. K. Gandolfi, C. Mourtel, and F. Olivier, "Electromagnetic analysis: Concrete results," *CHES 2001, Lecture Notes in Computer Science*, vol. 2162, pp. 251 – 261, May 2001.
3. J. Waddle and D. Wagner, "Towards efficient second-order power analysis," *CHES 2004, Lecture Notes in Computer Science*, vol. 3156, pp. 1 – 15, Aug. 2004.
4. H. C. Gebotys, S. Ho, and C. C. Tiu, "EM analysis of Rijndael and ECC on a wireless Java-based PDA," *CHES 2005, Lecture Notes in Computer Science*, vol. 3659, pp. 250 – 264, Aug. 2005.
5. Q. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition," *IEEE Transactions Pattern Analysis & Machine Intelligence*, vol. 16, pp. 1156 – 1168, Dec. 1994.
6. K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi, "High-accuracy subpixel image registration based on phase-only correlation," *IEICE Transactions on Fundamentals of Electronics, Communicati ons and Computer Sciences*, vol. E86-A, pp. 1925 – 1934, Aug. 2003.
7. K. Ito, H. Nakajima, K. Kobayashi, T. Aoki, and T. Higuchi, "A fingerprint matching algorithm using phase-only correlation," *IEICE Transactions on Fundamentals of Electronics, Communicati ons and Computer Sciences*, vol. E87-A, pp. 682 – 691, Mar. 2004.
8. K. Takita, A. M. Muquit, T. Aoki, and T. Higuchi, "A sub-pixel correspondence search technique for computer vision applications," *IEICE Transactions on Fundamentals of Electronics, Communicati ons and Computer Sciences*, vol. E87-A, pp. 1913 – 1923, Aug. 2004.
9. B. V. Kumar, *Correlation Pattern Recognition.* Cambridge University Press, 2005.
10. T. Matsumoto, S. Kawamura, K. Fujisaki, N. Torii, S. Ishida, Y. Tsunoo, M. Saeki, and A. Yamagishi, "Tamper-resistance standardization research committee report," *The 2006 Symposium on Cryptography and Information Security*, pp. 1 – 6, Jan. 2006.
11. A. M. Muquit, T. Shibahara, and T. Aoki, "A high-accuracy passive 3D measurement system using phase-based image matching," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E89-A, pp. 686 – 697, Mar. 2006.

A Advanced techniques for high-resolution waveform matching

Listed below are important considerations for high-resolution waveform matching.

A.1 Windowing to reduce boundary effects

Due to the DFT's periodicity, a waveform can be considered to "wrap around" at an edge, and therefore discontinuities, which are not supposed to exist in real world, occur at every edge in DFT computations. We reduce the effect of a discontinuity at a waveform border by applying a window function to the input waveforms $f(n)$ and $g(n)$. For example, we can employ a Hanning window defined as

$$w(n) = \frac{1 + \cos(\frac{\pi n}{M})}{2}. \tag{11}$$

The use of window functions is especially useful when the length of waveforms is short.

A.2 Spectral weighting technique to reduce aliasing and noise effects

For natural waveforms, typically the high frequency components may have less reliability (low S/N ratio) compared with the low frequency components. We could improve the estimation accuracy by applying a low-pass-type weighting function $H(k)$ to $R_{FG}(k)$ in frequency domain and eliminating the high frequency components having low reliability. The simplest weighting function $H(k)$ is defined as

$$H(k) = \begin{cases} 1 & |k| \leq U \\ 0 & \text{otherwise} \end{cases}, \tag{12}$$

where U is an integer satisfying $0 \leq U \leq M$. The cross-phase spectrum $R_{FG}(k)$ is multiplied by the weighting function $H(k)$ when calculating the IDFT. Then the modified $r_{fg}(n)$ will be given by

$$\begin{aligned} r_{fg}(n) &= \frac{1}{N} \sum_{k=-M}^{M} R_{FG}(k) H(k) W_N^{-kn} \\ &\simeq \frac{\alpha}{N} \frac{\sin\left\{\frac{V}{N}\pi\,(n+\delta)\right\}}{\sin\left\{\frac{\pi}{N}\,(n+\delta)\right\}}, \end{aligned} \tag{13}$$

where $V = 2U + 1$. When using the spectral weighting technique, Eq. (13) should be used for function fitting instead of Eq. (8). The main lobe of the POC function is extended by the spectral weighting technique.

Note that we can use any other weighting functions according to the reliability of the frequency components. If we use a weighting function, we need to change the peak model for function fitting correspondingly. The peak model can be calculated by the IDFT of the product of the weighting function and the cross-phase spectrum in Eq. (7).

A.3 Band-limited POC function

Another important technique for eliminating the high frequency components of waveforms is to use a band-limited POC function [7].

Assume that the range of the inherent frequency band is given by $k = -K, \cdots, K$, where $0 \leq K \leq M$. (The parameter K may be automatically detected by waveform processing.) The band-limited POC function is defined as

$$\begin{aligned} r_{fg}^{K}(n) &= \frac{1}{L} \sum_{k=-K}^{K} R_{FG}(k) W_L^{-kn} \\ &\simeq \frac{\alpha}{L} \frac{\sin\left\{\pi\,(n+\delta')\right\}}{\sin\left\{\frac{\pi}{L}\,(n+\delta')\right\}}, \end{aligned} \tag{14}$$

where $L = 2K + 1$, $n = -K, \cdots, K$ and $\delta' = \frac{L}{N}\delta$. Therefore, the displacement is given by $\delta = \frac{N}{L}\delta'$. The maximum value of the correlation peak of the band-limited POC function is always normalized to 1 and is not dependent on the frequency band size L. In practice, we can combine the band-limited POC function with the above spectral weighting technique.