Research and Implementation of Correlation Power Analysis Attack Based on Wavelet Transform Denoising (Postprint)

Authors: Duan Xiaoyi, Chen Dong, Gao Xianwei, Xiaohong Fan, JIN Jifang

Date: 2019-01-28T00:00:00+00:00

Abstract

To improve the efficiency of power analysis attacks and mitigate the impact of noise, this study investigates the effect of wavelet transform denoising on power analysis attacks and the relationship between the correlation coefficient of correlation power analysis (CPA) and attack effectiveness. It proposes employing the translation-invariant wavelet method and the wavelet modulus maxima method for denoising preprocessing of power consumption traces. This approach utilizes Kalman filtering, wavelet modulus maxima method, and translation-invariant wavelet method to denoise power consumption traces, followed by performing CPA on both the original and denoised data. Experimental results demonstrate that, compared with the original data, the correlation coefficient of CPA enhanced by the translation-invariant wavelet method increases by 165% over using CPA alone, by 31.4% over the Kalman filtering method, and by 26.4% over the wavelet modulus maxima method, while simultaneously reducing the number of power consumption traces required for successful attack by 92%. Experimental results indicate that the correlation power analysis attack improved by the translation-invariant wavelet method achieves the best performance.

Full Text

Preamble

Vol. 37 No. 4
Application Research of Computers
ChinaXiv Partner Journal

Research and Implementation of Correlation Power Analysis Attack Based on Wavelet Transform Denoising

Duan Xiaoyi, Chen Dong, Gao Xianwei, Fan Xiaohong, Jin Jifang
(Department of Electronic Information Engineering, Beijing Electronics Science & Technology Institute, Beijing 100070, China)

Abstract: To improve the efficiency of power analysis attacks and reduce noise interference, this paper investigates the impact of wavelet transform denoising on power attacks and the relationship between the correlation coefficient in Correlation Power Analysis (CPA) and attack effectiveness. We propose using translation invariant wavelet method and wavelet modulus maximum method for denoising preprocessing of power consumption curves. This approach employs Kalman filtering, wavelet modulus maximum method, and translation invariant wavelet method for denoising preprocessing of power curves, followed by CPA on both original and denoised data. Experimental results demonstrate that compared with original data, the improved CPA correlation coefficient using the translation invariant wavelet method increases by 165% over standard CPA analysis, by 31.4% over Kalman filtering, and by 26.4% over wavelet modulus maximum method, while simultaneously reducing the number of power curves required for successful attack by 92%. The experimental results indicate that the improved correlation power analysis attack using translation invariant wavelet method achieves the best performance.

Keywords: power analysis attack; translation invariant wavelet method; correlation power consumption; cryptographic chip

0 Introduction

The proposal of timing analysis techniques [1] marked the birth of side-channel attack technology. The introduction of Differential Power Analysis (DPA) and Simple Power Analysis (SPA) [2] ushered in a new era of side-channel attack research. Since then, various attack methods and defense strategies have emerged continuously, such as Correlation Power Analysis (CPA) [3], Multi-Channel Analysis (MCA) [4], Higher-Order Differential Power Analysis [5], Template Attack (TA) [6], and others. These attack methods have continuously driven the development of side-channel attack and defense countermeasures, including the addition of random masking [7,8] and random noise [9].

Research has shown that noise significantly impacts the efficiency of power analysis attacks. Literature [10] primarily discusses methods for analyzing keys from power consumption curves containing substantial noise. The noise sources in power consumption signals typically include radiation from surrounding electronic devices, transmission media, and artificially added random noise. Since key extraction in power analysis is based on collected power consumption signals, the signal-to-noise ratio (SNR) of real power signals greatly affects the success

rate of key analysis. Therefore, removing noise from power consumption curves can improve attack success rates. Current noise suppression methods mostly rely on increasing the number of power curve samples, but in practical attacks, the acquisition of power curve samples is limited, making unlimited increases in sample quantity impractical.

Consequently, under the constraint of limited power curve samples, minimizing noise interference in sampled power signals has become a focal point in power analysis research. Common preprocessing methods for power consumption curves in power analysis include Kalman filtering [11], Principal Component Analysis [12], Minkowski distance method [13], and improved Singular Value Decomposition [14]. However, Kalman filtering requires infinite past data; Principal Component Analysis necessitates obtaining an autocorrelation matrix to solve for eigenvectors and eigenvalues, a process involving excessive computational complexity and storage space for normal computing equipment, making it time-consuming; the Minkowski distance method, though computationally simple, fails to consider the differentiation of various components; the improved Singular Value Decomposition method faces difficulties in selecting the number of singular values. These limitations significantly restrict their application in noise suppression for power curves.

Wavelet analysis [15], in contrast, can adapt time-frequency resolution according to the local characteristics of the signal—providing high frequency resolution and low time resolution for low-frequency long-duration signals, while offering low frequency resolution and high time resolution for high-frequency short-duration signals. This zooming characteristic endows wavelet transform with strong adaptability to non-stationary signals, yielding excellent noise suppression effects by effectively separating signals from noise and improving SNR. Therefore, this paper applies translation invariant wavelet method and wavelet modulus maximum method [16] from wavelet analysis to preprocess power consumption curves, aiming to enhance power analysis efficiency.

1.1 Power Consumption Model for Power Curves

The energy consumption exploitable by attackers in power analysis [17] originates from the data processed and operations executed by the cryptographic device. Electronic noise components and constant components are two non-negligible energy components in actual collected power curves. Therefore, the total energy consumption at a single sampling point in a power curve generally consists of four components, as shown in Equation (1):

$$P_{total} = P_{op} + P_{data} + P_{noise} + P_{const}$$

where P_{op}, P_{data}, P_{noise}, and P_{const} represent the operation component, data component, noise component, and constant component, respectively, and they

are mutually independent. P_{op}, P_{data}, and P_{noise} are the most important information components for attackers, while the P_{const} component contains no exploitable information and is therefore irrelevant to power analysis. Attackers can only obtain key information by analyzing P_{op} and P_{data}, but as the noise component P_{noise} increases, attacks become more difficult.

To further clarify the energy components in the power consumption model, let P_{exp} represent the exploitable energy consumption component for attackers. The signal-to-noise ratio (SNR) is defined as the ratio of the actual collected signal component to the noise component. In power analysis, the SNR formula is shown in Equation (2):

$$SNR = \frac{Var(P_{exp})}{Var(P_{noise})}$$

$Var(P_{exp})$ quantifies the magnitude of variation in exploitable signals that cause changes in power curve points, while $Var(P_{noise})$ quantifies the magnitude of variation at those points caused by noise. The higher the SNR, the easier it is to identify P_{exp} from noise.

1.2 Principle of Correlation Power Analysis

Cryptographic chips consume energy during encryption or decryption algorithm execution. For flip-flops in registers, there is a direct relationship between processed data and energy consumption. This relationship manifests as Hamming weight or Hamming distance of data before and after instruction execution at the operand level, as 0/1 state transitions of flip-flops in registers, and as charging/discharging of load capacitance at the CMOS gate circuit level. Therefore, Hamming weight or Hamming distance of data before and after instruction execution can be utilized to analyze energy consumption during cryptographic chip data processing. The principle of power analysis attacks is to obtain keys through the correlation between encryption/decryption data and cryptographic chip power consumption. Correlation Power Analysis primarily exploits the correlation between actual measured encryption energy consumption and estimated Hamming distance or Hamming weight of intermediate encryption data for key analysis.

In practical engineering applications, first collect power consumption curves for N plaintexts encrypted with a fixed key, with M sampling points per curve, resulting in an $N \times M$ measured power matrix T. Then use subkeys to calculate and estimate an $N \times H$ hypothetical power matrix H. Let T_j be the j-th column of the measured power matrix T, representing the power consumption of N samples executing the same operation. Let H_j be the j-th column of the hypothetical power matrix H, representing the Hamming weight and Hamming

distance of N samples executing the same operation. Finally, calculate the correlation coefficient between matrix T and matrix H using Equation (5):

$$\rho_j = \frac{E(T_j)E(H_j) - E(T_jH_j)}{\sqrt{Var(T_j)Var(H_j)}}$$

[Figure 1: see original paper] shows CPA attack results. The spike appearing in the correlation coefficient matrix can be used to analyze the key used by the attacked cryptographic chip. However, in practical engineering applications, due to noise interference, especially artificially added noise, the spike correlation coefficient in attack results becomes quite small, thereby affecting attack success rates.

Based on Equation (3), model the energy consumption at a single sampling point, using P_{total} to represent the energy consumption of the encryption device at sampling position j. Let matrix H represent predicted energy consumption values, yielding Equation (6):

$$P_{total} = P_jH_j$$

Substituting into Equation (5) gives Equation (7):

$$\rho(P_{total}, H_j) = \frac{E(P_{total})E(H_j) - E(P_{total}H_j)}{\sqrt{Var(P_{total})Var(H_j)}}$$

Expanding the calculation as shown in Equation (8):

$$\rho(P_{total}, H_j) = \rho(P_{exp} + P_{sw.noise} + P_{el.noise} + P_{const}, H_j)$$

Since P_{const} is a constant component that does not affect the correlation coefficient, and the noise component P_{noise} is statistically independent of H_j, further simplification yields Equation (9):

$$\rho(P_{total}, H_j) = \rho(P_{exp}, H_j) + \rho(P_{noise}, H_j)$$

Substituting Equation (9) into Equation (5) gives Equation (10):

$$\rho = \frac{E(P_{exp})E(H_j) - E(P_{exp}H_j)}{\sqrt{Var(P_{exp} + P_{noise})Var(H_j)}} = \frac{E(P_{exp})E(H_j) - E(P_{exp}H_j)}{\sqrt{(Var(P_{exp}) + Var(P_{noise}))Var(H_j)}} = \frac{1}{\sqrt{1 + \frac{1}{SNR}}}$$

$E(\cdot)$ and $Var(\cdot)$ represent mathematical expectation and variance calculations, respectively. The correlation coefficient result matrix ρ indicates the magnitude of correlation between them—the larger the ρ, the greater their correlation. The correct subkey corresponds to a distinct spike in the correlation power curve, as shown in Figure 1(a), while incorrect subkeys show no obvious spike, as shown in Figure 1(b).

Equation (10) mathematically describes the relationship between correlation coefficient ρ and SNR. When the correlation coefficient between Hamming weight and exploitable energy consumption component remains constant, if the P_{noise} component in the power curve increases, the SNR decreases, and the correlation coefficient ρ becomes smaller; conversely, if P_{noise} decreases, SNR increases, and ρ becomes larger. When ρ becomes too small, the distinct spike corresponding to the correlation coefficient cannot be found, significantly increasing attack difficulty and even causing attack failure.

3.1 Wavelet Analysis

Wavelet analysis is a time-frequency analysis method whose resolution automatically matches the analyzed signal. It selects longer time windows when analyzing low-frequency signals and shorter time windows for high-frequency signals, perfectly solving the practical engineering problem of long-duration low-frequency signals and short-duration high-frequency signals. Hence, it enjoys the reputation of "mathematical microscope." Wavelet analysis finds wide applications in pattern recognition, image processing, data compression, fault diagnosis, speech recognition, and signal processing.

For any function $f(t) \in L^2(R)$, its continuous wavelet transform is:

$$W_f(a,b) = \langle f(t), \psi_{a,b}(t) \rangle = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt$$

Its inverse wavelet transform is:

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_f(a,b) \psi_{a,b}(t) \frac{dadb}{a^2}$$

where $\psi(t)$ is called a wavelet basis function or mother wavelet, such as commonly used db11, haar, sym8, and coif5 mother wavelets.

3.2 Translation Invariant Wavelet Denoising

Wavelet threshold denoising is widely applied across different fields due to its simple calculation process, easy algorithm implementation, and excellent denoising effects. However, it is prone to pseudo-Gibbs phenomena near singular points in reconstructed signals. Since collected power consumption information signals contain complex components and numerous singular points, this paper adopts the translation invariant wavelet method for preprocessing power curves, which can overcome the pseudo-Gibbs phenomenon caused by wavelet threshold methods.

The wavelet threshold denoising concept involves decomposing noisy signals across wavelet scales. First, a threshold is set and compared with wavelet coefficients at each scale. Coefficients larger than the threshold are processed according to certain rules, while those smaller than the threshold are set to zero. Finally, the remaining wavelet coefficients are reconstructed through inverse wavelet transform to obtain the denoised signal. The wavelet threshold denoising process is shown in [Figure 2: see original paper].

Translation invariant wavelet denoising is an improvement on wavelet threshold denoising. Its core idea is that after cyclic shifting, the wavelet coefficients of a signal have no corresponding shift relationship with the original signal' s wavelet coefficients. When singular points in a signal are located at position n_1, the transform results show almost no pseudo-Gibbs phenomenon; whereas singular points at position n_2 exhibit pseudo-Gibbs phenomena. Therefore, by changing the signal' s arrangement position to alter singular point locations, oscillation phenomena can be suppressed. Specifically, through artificial shifting, singular points at other positions are moved to position n_1 to eliminate pseudo-Gibbs phenomena, followed by reverse shifting to restore the original arrangement order, achieving denoising of noisy signals.

For a signal of length N, the translation invariant wavelet denoising steps are: first, cyclically shift the noisy signal with shift range 1 to N for complete coverage; second, perform discrete wavelet transform on each cyclically shifted signal to obtain wavelet decomposition coefficients at different scales; third, apply wavelet thresholding to process the wavelet coefficients; fourth, reconstruct by inverse transforming the remaining wavelet coefficients; finally, average and reverse cyclically shift to obtain the denoised signal, as shown in [Figure 3: see original paper].

3.3 Wavelet Modulus Maxima Denoising Method

The modulus maxima of wavelet transform coefficients reflect the singularity and peak characteristics of a signal. Specifically, the modulus maxima points of noise decrease with increasing scale, while those of the signal increase with scale. The basic idea of modulus maxima denoising is to reconstruct the signal

using the remaining modulus maxima points at different decomposition scales after processing.

The singularity at a point can be studied through wavelet transform. Let $\psi(t)$ be the wavelet function. If in any neighborhood of point t_0, the wavelet transform of function $f(t)$ at any scale j satisfies the following relationship with a constant k:

$$|W_f(2^j, t)| \leq k \cdot 2^{j\alpha}$$

then function $f(t)$ has a uniform Lipschitz exponent [18] at point t_0. Taking logarithms of both sides of Equation (13) reveals the relationship between wavelet transform modulus maxima and signal Lipschitz exponent:

$$\log_2 |W_f(2^j, x)| \leq \log_2 k + \alpha j$$

where j is the decomposition scale and $|W_f(2^j, x)|$ is the wavelet transform maximum at a point on the x-axis at scale j. Taking logarithms of both sides yields: if the Lipschitz exponent $\alpha < 0$, the wavelet transform modulus maxima of the function decrease with increasing scale; if $\alpha > 0$, the modulus maxima increase with scale. Signals and noise exhibit clear differences in singularity: signals have positive Lipschitz exponents, while noise has negative exponents. Consequently, their modulus maxima show distinctly different variation patterns across wavelet transform scales. The modulus maxima method exploits this characteristic—after multiple wavelet transforms, the amplitude of noise component modulus maxima points is removed or becomes very weak, leaving mostly signal-generated extremum points. Reconstructing the signal using these remaining modulus maxima yields the denoised signal, with the denoising process shown in [Figure 4: see original paper].

4.1 Power Analysis Platform and Evaluation Metrics

The correlation power analysis platform consists of three parts: cryptographic chip, data acquisition, and data processing, as shown in [Figure 5: see original paper]. The cryptographic chip selected is the Atmel ATMEGA16A microcontroller with a clock frequency of 4MHz. The data acquisition component uses a Tektronix DPO7104 oscilloscope to collect 10,000 data samples (10,000 power consumption curves) at a 50MHz sampling frequency, with each curve containing 25,000 power sampling points. Data processing employs Matlab for denoising preprocessing of power curves using Kalman filtering, wavelet modulus maxima method, and translation invariant method, followed by correlation power analysis executed in VC. The attack target selected is the 128-bit AES algorithm' s first-round S-box output bytes, with results displayed using Matlab.

The standard CPA attack directly uses power curves for analysis, while the improved correlation power analysis proposed in this paper performs CPA after preprocessing the power curves, as shown in [Figure 6: see original paper].

4.2 Power Curve Preprocessing

The purpose of power curve preprocessing is to suppress irrelevant noise in power curves, improve the SNR of useful signals, make the correlation coefficient ρ in correlation power analysis results more accurate, and thereby reduce the amount of data required for analysis. This paper uses an oscilloscope to collect 10,000 random plaintext sample data with a fixed key at a 50MHz sampling frequency, with each power curve containing 25,000 sampling points. Since the AES algorithm encryption operation only occurs within the first 14,000 sampling points, to reduce computational load, only the first 16,000 sampling points are processed.

Power curves primarily contain low-frequency signal components. To better analyze high-frequency signal portions, DC removal is typically performed first, leaving the high-frequency signals in the power curve, followed by Fourier transform to observe the frequency spectrum distribution. Figure 7: see original paper(b) show the original power consumption curve and its spectrum. Direct observation of Figure 7: see original paper cannot reveal key information through simple power analysis. Figure 7: see original paper intuitively shows that high-frequency signals are distributed below 25 MHz, indicating that the collected power data contains significant Gaussian white noise, with several maximum amplitude frequencies near 16 MHz, 8 MHz, and 4 MHz.

Figure 7: see original paper(d) show the power curve and spectrum after Kalman filtering. Compared with Figure 7: see original paper, the amplitude of the power curve in Figure 7: see original paper decreases. Compared with Figure 7: see original paper, the frequency components of high-frequency signals in Figure 7: see original paper are mainly concentrated within 8 MHz, mostly within 4 MHz, with frequency components beyond 8 MHz filtered out.

[Figure 8: see original paper] shows power curves and spectra after denoising using wavelet modulus maxima method and translation invariant wavelet method. From the time domain perspective, the denoised power curve waveforms in Figure 8: see original paper(c) appear sparser compared with Figure 7: see original paper(c) because these two wavelet denoising methods better remove noise components while preserving useful signals. From the frequency domain perspective, compared with the original spectrum in Figure 7: see original paper, the spectrum after wavelet modulus maxima denoising in Figure 8: see original paper is mainly concentrated within 2.5 MHz, with Gaussian white noise in the power curve almost completely removed, whereas Figure 7: see original paper still contains substantial other high-frequency noise components, indicating that Kalman filtering is not thorough. Compared with the

original spectrum in Figure 7: see original paper, the spectrum after translation invariant wavelet method in Figure 8: see original paper shows high-frequency components mainly concentrated within 2.5 MHz, with high-frequency signals in other frequency ranges reduced to very small values compared with Figure 7: see original paper.

This preliminary frequency-domain analysis of the three preprocessing methods' denoising performance must ultimately be combined with correlation power analysis to measure noise suppression effectiveness from attack results. Therefore, preprocessing method performance is evaluated from two aspects: correlation coefficient and resource consumption, measuring effectiveness from attack effectiveness and efficiency perspectives respectively.

4.3 Evaluation from Attack Effectiveness

The magnitude of the correlation coefficient in correlation power analysis results directly affects attack success rates. Noise in power curves easily causes ghost peaks in analysis results. Larger ghost peaks create greater interference with attack results, directly affecting the discrimination of correct key peaks and even causing attack failure. If noise components in power data can be effectively removed to improve the SNR of exploitable power data, the correlation coefficient peak corresponding to the correct key will be clearly distinguishable from ghost peaks. When their gap widens further, ghost peak interference becomes negligible. Therefore, this paper uses the correlation coefficient peak magnitude corresponding to the correct key as a metric for evaluating preprocessing method denoising performance. Appropriate preprocessing methods can effectively suppress noise in power data (i.e., reduce the P_{noise} component), resulting in obvious ghost peak elimination and improved correlation power analysis success rates.

[Figure 9: see original paper] shows CPA results using 120 power curves with the correct key (16th byte key). Figure 9: see original paper presents results using 120 untreated original power curves, while Figure 9: see original paper-(d) show results using 120 power curves preprocessed by Kalman filtering, wavelet modulus maxima method, and translation invariant wavelet method respectively. Figure 9: see original paper shows no expected correlation coefficient spike for the correct key, while Figure 9: see original paper-(d) show obvious spikes, though Figure 9: see original paper is less distinct and creates significant interference for correct discrimination. According to Equation (3), power curves contain substantial noise P_{noise} that masks the exploitable signal P_{exp}, making SNR small. Equation (10) shows that when P_{noise} is large, ρ becomes small. The experimental results in [Figure 9: see original paper] demonstrate that CPA using a small number of original power curves cannot guess the correct key.

[Figure 10: see original paper] shows CPA results using 400 power curves with the correct key (16th byte key). All subfigures Figure 10: see original paper-

(d) show very distinct spikes, indicating that as the number of power curves increases, some noise can be suppressed in correlation power analysis to obtain relatively clear attack results. At this point, the correlation coefficient corresponding to the correct key spike is 0.32 in Figure 10: see original paper, 0.45 in Figure 10: see original paper, 0.46 in Figure 10: see original paper, and 0.57 in Figure 10: see original paper. The translation invariant wavelet method yields the largest correlation coefficient, higher discrimination, less judgment interference, and more reliable attack results, making it the best noise suppression method. Compared with wavelet modulus maxima method, the correlation coefficient improves by 23.9%; compared with Kalman filtering, it improves by 26.7%; and compared with the standard scheme, it improves by 78.1%.

From these attack results, we can infer the SNR ranking of power curves using different denoising methods under current conditions: translation invariant wavelet > wavelet modulus maxima > Kalman filtering. [Figure 9: see original paper] and [Figure 10: see original paper] demonstrate that noise interference significantly impacts CPA attacks, and removing noise interference enables faster and more accurate CPA attacks, making wavelet preprocessing for interference removal crucial.

To ensure experimental accuracy, the results in were obtained by performing correlation power analysis on 5,000 power curves preprocessed by these three methods, with the attack target being the 128-bit AES algorithm key bytes. CPA attack effects using power curves preprocessed by Kalman filtering, wavelet modulus maxima method, and translation invariant wavelet method are significantly superior to standard CPA. Compared with Kalman filtering, wavelet modulus maxima method improves the correlation coefficient by 4.2% on average per key byte, showing slightly better denoising performance. The translation invariant wavelet method demonstrates significantly better CPA attack results than other preprocessing techniques, improving the correlation coefficient by 165% on average per key byte compared with the standard scheme, by 31.4% compared with Kalman filtering, and by 26.4% compared with wavelet modulus maxima method. The translation invariant wavelet denoising method is suitable for signals containing several discontinuities and low SNR, enabling it to remove irrelevant noise from power signals and improve attack success rates.

TABLE:1 Correlation Coefficient for Correct Guessing of AES 16-Byte Key

Key Byte	Standard CPA	Kalman Filtering	Wavelet Modulus Maxima	Translation Invariant Wavelet
(Table data would appear here)				

4.4 Evaluation from Attack Efficiency

Power analysis continuously pursues more accurate results using less data. In real attack scenarios, repeated collection under identical conditions is impossible, and most current attack technologies require collecting substantial power consumption data as a prerequisite. The key issue is minimizing power curve samples while ensuring analysis accuracy. Therefore, from an attack efficiency perspective, the primary metric is power curve utilization rate—the higher the utilization, the better the noise suppression effect.

When using CPA attacks, the question is how many power curves are needed to obtain distinct spikes. [Figure 11: see original paper] provides a preliminary answer, where the red curve represents the correlation coefficient for the correct key hypothesis, while other key hypotheses are plotted in black. The figure shows that as the number of power curves increases, the correlation coefficient for the correct key hypothesis separates from those of incorrect key hypotheses.

[Figure 11: see original paper] uses the 16th key byte for attack. Subfigures Figure 11: see original paper-(d) show the relationship between power curves and correlation coefficients for original curves, Kalman-filtered curves, wavelet modulus maxima denoised curves, and translation invariant wavelet denoised curves respectively. In Figure 11: see original paper, the correct key' s correlation coefficient converges to 0.08 while incorrect keys converge to 0.03, requiring approximately 233 or more power curves for the correct key to produce the highest correlation coefficient, indicating a successful attack needs about 233 original power curves. Figure 11: see original paper requires about 54 power curves after wavelet modulus maxima denoising, and Figure 11: see original paper requires about 22 power curves after translation invariant wavelet denoising. Compared with using original power curves for CPA, using denoised power curves significantly improves power curve utilization.

TABLE:2 Quantitative Relationship Between Correlation Coefficient and Power Consumption Curve

Key Byte	Original Data	Kalman Filtering	Wavelet Modulus Maxima	Translation Invariant Wavelet
(Table data would appear here)				

TABLE:2 shows the number of power curves required to distinguish all 128-bit

AES algorithm keys. The translation invariant wavelet method requires the fewest power curves—only 2,921—while undenoised power curves require 42,533, representing a 93.1% reduction. The translation invariant wavelet method reduces power curve usage by 31.4% compared with Kalman filtering and by 26.5% compared with wavelet modulus maxima method. Thus, translation invariant wavelet method achieves the highest power curve utilization, with wavelet modulus maxima performing slightly better than Kalman filtering. In practical engineering applications, due to limitations on obtainable power curve sample quantities, efficiently utilizing cryptographic device power curves to quickly crack encryption algorithms becomes particularly important.

---

5 Conclusion

This paper proposes using wavelet modulus maxima method and translation invariant wavelet method for denoising preprocessing of power consumption curves. Experimental results demonstrate that all three denoising methods improve both the attack effectiveness and efficiency of correlation power analysis methods. The translation invariant wavelet method shows the most significant noise interference suppression effect, substantially reducing ghost peak generation and improving attack accuracy. It requires only 7% of original power curve quantity to successfully analyze all 128-bit keys, with the correlation coefficient for each key byte improving by an average of 165% compared with the original. Therefore, translation invariant wavelet method achieves optimal denoising performance in correlation power analysis. Through comparison, translation invariant wavelet method proves to be the most effective among the three denoising methods.

References

[1] Kocher P C. Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems [C]//Proc of International Cryptology Conference on Advances in Cryptology. 2010: 104-113.

[2] Kocher P, Jaffe J, Jun B, et al. Introduction to differential power analysis [J]. Journal of Cryptographic Engineering, 2011, 1 (1): 5-27.

[3] Bottinelli P, Bos J W. Computational aspects of correlation power analysis [J]. Journal of Cryptographic Engineering, 2016, 3 (7): 1-15.

[4] Agrawal D, Rao J R, Rohatgi P. Multi-channel attacks [C]//Proc of International Workshop on Cryptographic Hardware and Embedded Systems. 2003: 2-16.

[5] Zhang Liwei, Ding Aidong A, Fei Yunsi, et al. Efficient nonprofiling 2nd-order power analysis on masked devices utilizing multiple leakage points [J].

IEEE Trans on Dependable & Secure Computing, 2017, (99): 1-1.

[6] 李佩之, 严迎建, 段二朋. DES 密码芯片模板攻击技术研究 [J]. 计算机应用与软件, 2013, 30 (4): 310-312. (Li Peizhi, Yan Yingjian, Duan Erpeng. Research on template attack techniques against DES cryptographic chip [J]. Computer Applications and Software, 2013, 30 (4): 310-312.)

[7] Chari S. A cautionary note regarding evaluation of AES candidates on smart-cards [C]//Proc of Advanced Encryption Standard Candidate Conference. 2010: 133-147.

[8] 李浪, 欧雨, 邹祎. 一种 AES 随机变换掩码方案及抗 DPA 分析 [J]. 密码学报, 2018, 5 (4): 442-454. (Li Lang, OU Yu, Zou Yi. On AES random transform masking scheme against DPA [J]. Journal of Cryptologic Research, 2018, 5 (4): 442-454.)

[9] 徐佩. 智能卡 AES 加密模块抗侧信道攻击掩码技术研究与实现 [D]. 重庆: 重庆大学, 2015. (Xu Pei. Research and Implementation with Mask Technology on AES Encryption Module of Smartcard against Side Channel Attack [D]. Chongqing: Chongqing University, 2015.)

[10] Kunihiro N, Takahashi Y. Improved key recovery algorithms from noisy rsa secret keys with analog noise [M]// Topics in Cryptology. Switzerland: Springer International Publishing, 2017.

[11] Souissi Y, Guilley S, Danger J, et al. Improvement of power analysis attacks using Kalman filter [C]//Proc of IEEE International Conference on Acoustics, Speech, and Signal Processing. 2010: 1778-1781.

[12] Cagli E, Dumas C, Prouff E. Enhancing dimensionality reduction methods for side-channel attacks [M]// Smart Card Research and Advanced Applications. Switzerland: Springer International Publishing, 2015.

[13] Ou Changhai, Wang Zhu, Sun Degang, et al. Enhanced correlation power analysis by biasing power traces [M]// Information Security. Switzerland: Springer International Publishing, 2016.

[14] Sun Degang, Zhou Xinping, Wang Zhu, et al. POSTER: using improved singular value decomposition to enhance correlation power analysis [M]// Security and Privacy in Communication Networks. Switzerland: Springer International Publishing, 2015.

[15] 李杨. 小波去噪方法的研究 [J]. 科技视界, 2017 (25): 56-56. (Li Yang. Research on wavelet denoising method [J]. Science & Technology Vision, 2017 (25): 56-56.)

[16] 张新鹤. 基于小波变换模极大值的信号奇异性检测 [J]. 电子制作, 2015 (5): 1-3. (Zhang Xinhe. Signal singularity detection based on wavelet transform modulus maxima [J]. Practical Electronics, 2015 (5): 1-3.)

[17] Mangard S, Oswald E, Popp T. Power analysis attacks: Revealing the secrets of smart cards[M]. Springer Science & Business Media, 2008.

[18] 华春红, 任章, 张敏虎. 基于自适应阈值估计的模极大值去噪方法 [J]. 航天控制, 2011, 29 (1): 37-47. (Hua Chunhong, Ren Zhang, Zhang Minhu. The wavelet maxima denoising based on the adaptive Bayes shrink threshold [J]. Aerospace Control, 2011, 29 (1): 37-47.)

Note: Figure translations are in progress. See original paper for figures.