International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.11, (2016), pp.1-12
http://dx.doi.org/10.14257/ijsip.2016.9.11.01

MEMD Interval Threshold Filter and Twin SVM for Electromagnetic Template Attack

Duan Li[1,2], Hong-xin Zhang[1,3*], Qiang Li[1], Xinjie Zhao[4] and Pengfei He[5]

[1]*School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[2]*School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo, 454000, China*
[3]*Beijing Key Laboratory of Work Safety Intelligent Monitoring (Beijing University of Posts and Telecommunications), Beijing 100876, China*
[4]*Institute of North Electronic Equipment, Beijing 100191, China*
[5]*School of Opto-electronic Information Science and Technology, Yantai University, Yantai 264005, China*
[1]*liduan@hpu.edu.cn*, [3]*hongxinzhang@263.net*

Abstract

To overcome the dependence on prior knowledge of traditional filtering algorithm, this paper proposes a novel non-parameter and adaptive multivariate empirical mode decomposition interval threshold (MEMD-IT) denoising approach for signal preprocessing of electromagnetic template attack(ETA). MEMD-IT can reduce the discontinuity induced by traditional MEMD-DT and aims to remove the Gaussian noise coupled into side-channel electromagnetic radiations. The proposed method and some other filters such as butterworth low-pass filter(BLPF), wavelet threshold denoising and MEMD direct threshold (MEMD-DT) denoising are applied to analyze the electromagnetic radiation traces intercepted while cipher device was implementing RC4 encryption algorithm. Furthermore, twin SVM multi-class classifier based ETA was performed to evaluate the attack results. In the same attack scenario, the highest predictive success rate for 9 Hamming Weights of the key reached 92%,90%,84%,83% and 73% for MEMD-IT, WFTF, MEMD-DT,WULETF and BLPF preprocessing methods, respectively. The experiment results indicate that our proposed scheme has a significant performance compared with the traditional ETA.

Keywords: *Electromagnetic template attack, Multivariate empirical mode decomposition (MEMD), Interval threshold denoising, Twin support vector machine(TWSVM),RC4 encryption*

1. Introduction

Machine learning has a powerful potential for performing electromagnetic template attack (ETA) which aims to recover the key-relevant intermediate value or even the key by exploiting the electromagnetic radiations from physical devices [1-4]. The signal-to-noise ratio (SNR) of side-channel physical leakage may significantly influence the predictive accuracy. Since Electromagnetic radiations are often corrupted by coupling interference and the noise of test instruments, also, there are electronic noise and switch noise, which is irrelevant to the key-relevant intermediate value in the noisy electromagnetic radiations [5]. Furthermore, the noise is related to encryption algorithm implementation, cryptographic device and

*Corresponding Author

ISSN: 2005-4254 IJSIP

measurement environment. Therefore effective adaptive electromagnetic radiations preprocessing is the key for side-channel analysis.

In references [6-7], the averaging operation are used to reduce noise. However, this method requires a large number of electromagnetic radiation traces. Messerges *et al.,* introduced another method which consists of filtering noise and using the multibit DPA attack to improve the SNR DPA signals [8]. Thanh-Ha Le *et. al.,* [5] proposed the fourth-order cumulant method to improve the performance of side-channel attacks. But the high order statistics method is insensitive to noise and needs a lot of traces.

Empirical mode decomposition(EMD), first introduced by Hang *et. al.,* in [9] ,has been widely used to analyze the non-stationary and non-linear signal processes by adaptively decomposing any signal into oscillatory components called intrinsic mode functions (IMFs). Recently, the statistical characteristics of white Gaussian noise and fractional Gaussian noise through EMD have been revealed in [10-12]. According to these characteristics, many EMD based denoising methods are provided to remove noise from observed data. Compared with the traditional methods, such as wavelet and Fourier transform, EMD denoising is an adaptive and more precise method. Flandrin *et. al.,* proposed that EMD denoising is superior to wavelet for fractional Gaussian noise [9]. Boudraa *et. al.,* have later proved that EMD filtering based on partial reconstruction of relevant modes performs in an adaptive way in [13]. But it is disastrous for noise removal when the selection of relevant modes is incorrect.

To alleviate both mode alignment and mode mixing, we propose multivariate empirical mode decomposition interval threshold (MEMD-IT) method to the Electromagnetic emanation preprocessing. Different from the EMD direct threshold (EMD-DT), interval threshold considers the zero-crossing interval as a whole to perform thresholding, which can improve the discontinuity of the reconstructed signal. Furthermore, the twin SVM (TWSVM) classifier which has the solid theoretical foundation [14] and four times faster computational complexity as compared to the traditional SVMs is applied to the electromagnetic template attack. The electromagnetic analysis scheme combined EMD interval threshold preprocessing with TWSVM classifier is implemented in the practical attack platform while cipher chip performing RC4 encryption algorithm.

The rest of this paper is as follows. In Section 2, we propose a criterion of selecting relevant modes for the MEMD-IT denoising method, and introduce the principle of selecting threshold, furthermore, the simulation results is analyzed. Electromagnetic radiations preprocessing based on MEMD-IT is performed and discussed in Section 3. In Section 4, a twin SVM multi-class classification is used to predict the Hamming Weight of the Key, several traditional popular denoising methods are investigated and compared in the practical attack scenario. Finally, conclusions are given in Section 5.

2. MEMD Based Denoising

All the IMFs decomposed by MEMD method include different frequency bands ranging from high to low. The high-frequency band IMFs contain the noise and the low-frequency band IMFs contain mostly the signal energy [15-16]. How to determine which IMFs are pure signal, pure noise, or contain both is a key issue to effectively perform the MEMD filter.

2.1. Criterion of Selecting Relevant Modes

Consider a noiseless signal $y(t)$ contaminated by an additive noise $n(t)$, the denoising aims to find an estimate $x(t)$ of the observed signal $y(t)$. For MEMD denoising, it is important to discriminate between relevant and irrelevant modes. MEMD denoising based on partial reconstruction is given by

$$x(t) = \sum_{i=k_{th}}^{L} h^{(i)}(t) + r_L(t) \tag{1}$$

The k_{th} can be determined by an estimation of correlation coefficient between the original data and decomposition modes. To calculate conveniently, Equation (1) can be rewritten as

$$x_m(t) = x(t) - \sum_{i=1}^{m} h^{(i)}(t) \tag{2}$$

where, $m = k_{th} - 1$.

The correlation coefficient between $x(t)$ and $x_m(t)$ is as follows:

$$\rho(m) = \frac{\sum_{t=1}^{N} x(t) x_m(t)}{\sqrt{\sum_{t=1}^{N} x^2(t)} \sqrt{\sum_{t=1}^{N} x_m^2(t)}} \tag{3}$$

where, N is the length of data. k_{th} is determined when the $\rho(m)$ starts smaller than some constant C, usually, $C \in [0.75, 0.85]$. In this study, C is set to 0.8. So k_{th} is given by:

$$k_{th} = \arg \underset{1 \le m \le L}{last} \{\rho(m) \ge 0.8\} + 1 \tag{4}$$

The "last" stands for the last value in series $\rho(m)$ bigger than 0.8.

As an example, a noisy signal decomposed synchronously with the other two signals is shown in Figure 1. To get the statistical value of k_{th}, 50 trials have been carried out on the above noisy signals with different SNR. As shown in Figure 2, the arrow points to k_{th} on different SNR that calculated from Formula (4).When SNR is -5 or -10, k_{th} equals to 2. Because k_{th} increases with the increase of SNR, the choice of relevant modes is important for signal denoising.

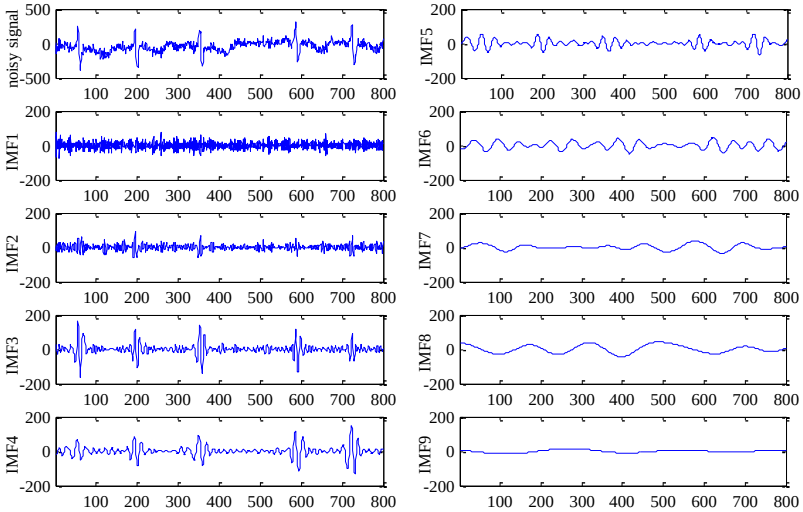

Figure1. The Noisy Signal and MEMD Decomposition of It

2.2. MEMD Interval Threshold Denoising

MEMD based denoising can be classified into two main categories, partial reconstruction and whole reconstruction. The partial reconstruction given in Equation (2), with the relevant modes kept and irrelevant modes discarded, usually miss some useful information in discarded modes especially when SNR is high. To

improve the performance of MEMD denoising, inspired by wavelet threshold method, the whole reconstruction with filtered modes is provided in this study.

2.2.1. Principle of Selecting Threshold: Threshold is a risk parameter to control the alternative change of deviation and variance. If estimate value is two small, the

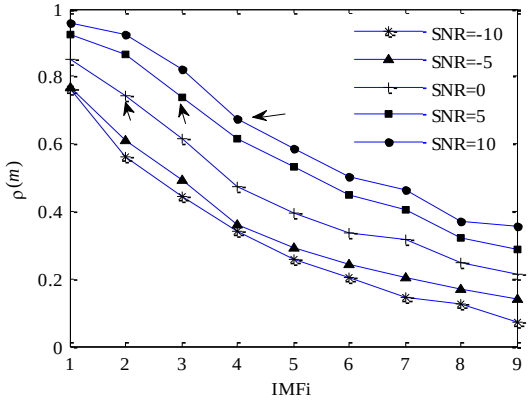

Figure 2. The Variation of k_{th} with SNR Change

reconstruction signal is very similar to noisy signal. On the contrary, it will give rise to under-fitting in signal reconstruction. As the electromagnetic radiation noise can be regarded as white Gaussian noise, the characteristic of decomposition by MEMD is that the power spectra of the other IMFs, apart from the first noise-only IMF, exhibit self-similar characteristics. As the statistic characteristic of Gaussian noise depends on Hurst exponent H [10], the energy of IMFs, V_k, decreases linearly in a semilog diagram, the form is:

$$V_k = \frac{V_1}{\beta}\rho^{-k}, k = 2,3,\cdots,L \tag{5}$$

Where, V_1 is noise energy, the standard deviation of the noise is estimated using a robust estimator on the first IMF:

$$\sigma_1 = \frac{median(\left|h^{(1)}(t)\right| : t = 1,\cdots,N)}{0.6745} \tag{6}$$

In reference [9], based on plenty of experimental results, Flandrin *et al.*, proposed that the appropriate value of β and ρ are 0.719 and 2.01,respectively.

The adaptive threshold of IMFs are given by:

$$T_k = \sqrt{\frac{V_k}{N} 2\ln N}, \quad k = 2,3,\cdots,L \tag{7}$$

where, N is signal length.

2.2.2. Interval Threshold Denoising: There are two ways for MEMD-denoising, hard threshold or soft threshold. The estimate noise $x(t)$ is :

$$x(t) = \sum_{i=M_1}^{M_2} h^{(i)}(t) + \sum_{i=M_2+1}^{L} h^{(i)}(t) + r_L(t) \tag{8}$$

$$h^{(i)}(t)=\begin{cases} h^{(i)}(t)\,, & \left|h^{(i)}(t)\right|>T_i \\ 0\ , & \left|h^{(i)}(t)\right|\le T_i \end{cases} \tag{9}$$

where
for hard thresholding and

$$h^{(i)}(t)=\begin{cases} \mathrm{sgn}(h^{(i)}(t))(\left|h^{(i)}(t)\right|-T_i)\,, & \left|h^{(i)}(t)\right|>T_i \\ 0\ , & \left|h^{(i)}(t)\right|\le T_i \end{cases} \tag{10}$$

for soft thresholding. Where T_i is threshold of the i-th mode. According to the criterion of selecting relevant modes, variables M_1 and M_2 in Equation (6) can be selected. The Equations (8) to (10) are MEMD direct threshold filter which will result in disastrous consequence for the continuity of the denoised signal. This is because the IMFs resemble an AM/FM modulated sinusoid with zero mean. As a result, it is guaranteed that, even in a noiseless case, in any interval $\mathbf{z}_j^{(i)}=[z_j^{(i)}\ z_{j+1}^{(i)}]$, the absolute amplitude of the i-th IMF, $i=1,\cdots,N$, will drop below any nonzero threshold in the proximity of the zero-crossings $z_j^{(i)}$ and $z_{j+1}^{(i)}$. So we can guess if the interval $\mathbf{z}_j^{(i)}$ is noise-dominant or signal-dominant based on the single extrema that correspond to this interval. If the signal is absent, the absolute value of this extrema will lie below the threshold. Alternatively, in the presence of strong signal, the extrema value can be expected to exceed the threshold. As a result, the MEMD interval hard thresholding and soft thresholding translate to:

$$h^{(i)}(\mathbf{z}_j^{(i)})=\begin{cases} h^{(i)}(\mathbf{z}_j^{(i)})\,, & \left|h^{(i)}(r_j^{(i)})\right|>T_i \\ 0\ , & \left|h^{(i)}(r_j^{(i)})\right|\le T_i \end{cases} \tag{11}$$

$$h^{(i)}(\mathbf{z}_j^{(i)})=\begin{cases} h^{(i)}(\mathbf{z}_j^{(i)})\dfrac{\left|h^{(i)}(r_j^{(i)})\right|-T_i}{\left|h^{(i)}(r_j^{(i)})\right|}, & \left|h^{(i)}(r_j^{(i)})\right|>T_i \\ 0\ , & \left|h^{(i)}(r_j^{(i)})\right|\le T_i \end{cases} \tag{12}$$

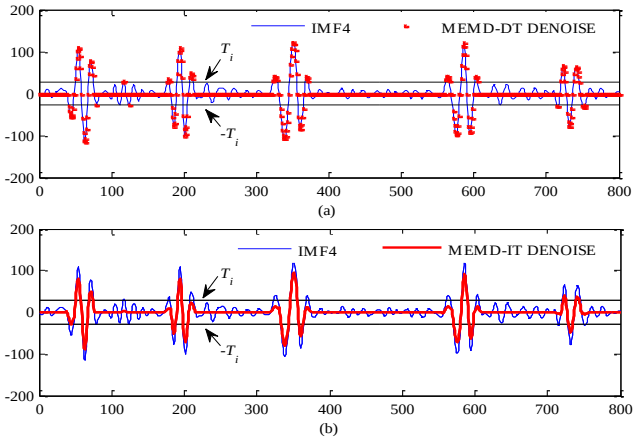

Figure 3. Denoising Results of Direct and Interval Thresholding

where, $h^{(i)}(\mathbf{z}_j^{(i)})$ indicates the samples from instant $z_j^{(i)}$ to $z_{j+1}^{(i)}$ of the i-th mode and $h^{(i)}(r_j^{(i)})$ is the extrema corresponding to this interval. As an example, the denoising

results of the IMF4 in Figure 1 based on MEMD-DT and MEMD-IT is shown in Figure 3 (a) and (b), respectively. As shown in Figure 3 (a), there are serious discontinuous because the signal value between interval $[-T_i\ T_i]$ are set zero. The interval filter in Figure 3(b), maintains continuity and smoothness of IMF4.

3. Electromagnetic Radiations Preprocessing Based on MEMD-IT

3.1. Electromagnetic Radiations Interception

In this work, taking a software implementation of RC4-8 on microcontroller, the attack is carried out to classify the hamming weights of the key. In the first stage of the RC4 algorithm, the original key is pushed to the Key repeatedly until the Key is filled up. So the electromagnetic radiation leakage during key expansion procedure is relevant to the original key.

As shown in Figure 4, Our experimental platform is an unprotected software implementation of the RC4 algorithm running on a STC89C52. The clock frequency of the microcontroller is set to 11.05926MHz. The electromagnetic radiations are captured by an EM5030 in near-field, and the data are sampled with Tektronix TDS5054 digital oscilloscope whose sampling frequency is set to 500MHz. The oscilloscope is triggered by microcontroller, and the sampled data are transmitted to PC synchronously. In this experiment, we collect radiation traces while microcontroller running RC4 algorithm 7200 times with random keys. In the following study, the training and testing samples are the feature vectors extracted from the 7200 traces.

3.2. Preprocessing Based on MEMD-IT

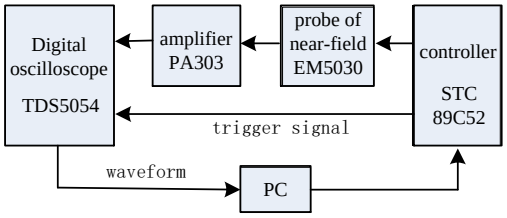

Figure 4. The Schematic of Electromagnetic Radiations Acquisition

3.2.1. Noise Analysis of Electromagnetic Emanations: During radiations interception, electromagnetic radiation traces are easily affected by intrinsic and external noise. Intrinsic noise is due to physical fluctuations in circuits, such noise can be distinguished into at least four different types: thermal noise, shot noise, 1/f noise, and generation-recombination noise [17]. External noise is generated by external sources, such as measuring equipment or environment conditions.

All fluctuating currents and voltages generated in electrical devices have a probability density function of Gaussian form since the fluctuating quantity is the sum of a large number of independent random variables. Thus, the intrinsic noise is Gaussian. Similarly, we consider the external noise as Gaussian noise.

3.2.2. MEMD-IT Filtering Experiments: Figure 5, is part of electromagnetic radiations while cryptographic chip executing key expansion procedure of RC4 algorithm. As shown in Figure 5, different electromagnetic emanation traces corresponding to different Hamming Weight of the key are compared. To ensure that the microcontroller perform the same operations at the same time, high precision alignment of the traces is also important for ETA. After selecting the

pattern, least squares algorithm is used in pattern matching to realize signal alignment. Figure 5(a), and (b), are raw traces and alignment traces, respectively. The smoothing traces after denoising based on MEMD-IT are shown in Figure 5(c).The intermediate value (the Hamming Weight of the key) can be distinguished by the small magnitude difference of the electromagnetic radiations shown in Figure 5(c).Therefore, if there are enough sample traces, machine learning method can be used to predict the intermediate value of the encryption key.

4. Electromagnetic Side-Channel Attack Analysis Based on TWSVM

Machine learning has a powerful potential for performing the template attack of cryptographic device. In this work, TWSVM is used to predict the 9 Hamming Weights of the key which can reduce the cipher key search space effectively.

4.1. Feature Selection of the Traces

As machine learning methods are difficult to deal with the high dimensionality characteristic of the side-channel leakage signals. Feature extraction is important for classification. In this study, Pearson correlation approach is used to select feature points. It filters out wrong features and irrelevant components in a trace vector in order to avoid confusing the classifiers.

The principle of the method is expressed by Formula (13) as follows:

$$\rho(i) = \frac{\mathrm{cov}(x(i), y)}{\sqrt{\mathrm{var}(x(i) \cdot \mathrm{var}(y)}} \qquad -1 \le \rho \le 1 \tag{13}$$

where cov(·,·) represents the covariance, var(·) represents the variance, $x(i)$ is a vector composed of the i-th component of all the training trace. And y is the vector which contains the class labels that are the hamming weights corresponding to the keys. So $\rho(i)$ is the correlation coefficient between i-th component of traces and assumption power consumption.

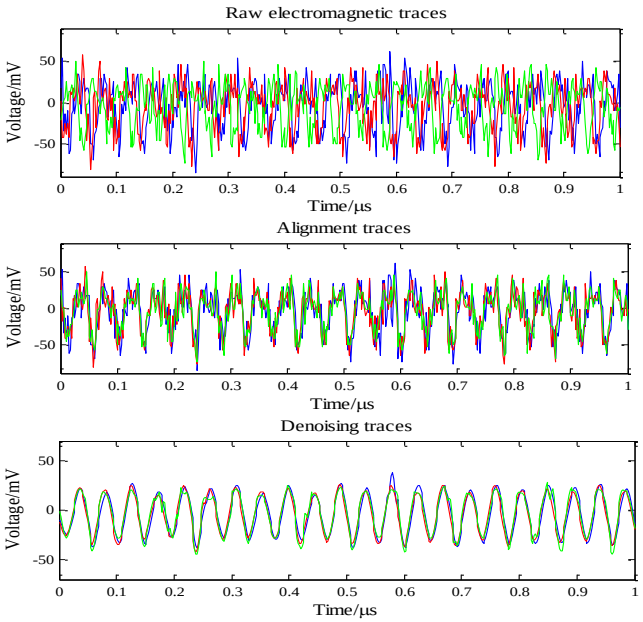

Figure 5. Preprocessing of Electromagnetic Radiations

From the 7200 vectors extracted from electromagnetic traces, we selected 1800 vectors (200 vectors from each hamming weight of original keys, the hamming weight from 0 to 8) for training and randomly choose 900 vectors (hamming weight from 0 to 8) from the remainder for testing.

The number of feature points we can extract is determined by the correlation coefficient $\rho(i)$ we set. The lower correlation coefficient $\rho(i)$ we set, the larger number of feature points we can get. Considering the running time and predictive rate of the proposed predictive model, we set $\rho(i) \geq 0.65$. In this study, several filters such as Butterworth low-pass filter (BLPF) with 80MHz cutoff frequency, wavelet unbiased likelihood estimator threshold filter(WULETF),wavelet fixed threshold filter(WFTF),MEMD direct threshold (MEMD-DT) and interval threshold (MEMD-IT) denoising are applied to the preprocessing of the intercepted electromagnetic radiations. Three level decomposition and db3 wavelet function are used for wavelet filter. Using Pearson correlation approach, the feature selection results of the training traces are shown in table1. Set $\rho(i) \geq 0.65$, the number of feature points selected is 73 if we use MEMD-IT filter for preprocessing, which is larger than other filters. Next is wavelet fixed threshold filter, we can get 65 feature points based on Pearson correlation approach. The results of Butterworth low-pass filter will be greatly influenced by the choice of cutoff frequency. Therefore, it has poor adaptivity.

4.2. Side-Channel Analysis Based on TWSVM

4.2.1. The Mathematical Model of TWSVM: The traditional SVM involves the solution of a single quadratic programming problem (QPP).This can be time-consuming for datasets with large number of features. Also, the SVM involves obtaining the predicted label using a single maximum-margin hyperplane. This technique is based on the intuition that a better prediction can be obtained by using a formulation which allows for non-parallel, as well as more than one hyperplanes.

Table 1. Feature Selection Results of Different Preprocessing Methods

$\rho(i)$ / Feature Points / Filters	0.6	0.61	0.62	0.63	0.64	0.65	0.66	0.67	0.68	0.69	0.7	0.71	0.72	0.73
MEMD-IT	2031	1240	744	348	105	73	24	18	8	7	2	0	0	0
MEMD-DT	51	20	13	12	3	2	0	0	0	0	0	0	0	0
WULETF	31	27	24	20	15	10	5	5	3	2	2	1	0	0
WFTF	531	382	268	156	99	63	35	15	3	2	2	0	0	0
BLPF (80MHz)	19	17	16	12	10	8	7	5	2	0	0	0	0	0

Consider the following classification problem. Suppose that all the data points in class +1 are denoted by a matrix $X_1 \in R^{m_1 \times n}$, where the i-th row $X_i \in R^n$ represents a sample. Similarly, the matrix $X_2 \in R^{m_2 \times n}$ represents the data points of class –1.

TWSVM [18-22] seeks a pair of nonparallel hyperplanes

$$f_1(x) = w_1^{\mathrm{T}} x + b_1 \text{ and } f_2(x) = w_2^{\mathrm{T}} x + b_2 \tag{14}$$

In the nonlinear separable case, we use the kernel function $K(x, X)$, then the two hyperplanes of TWSVM are as follows:

$$K(x, X)u^{(1)} + b^{(1)} = 0 \text{ and } K(x, X)u^{(2)} + b^{(2)} = 0 \tag{15}$$

where $X = [X_1 ; X_2]$, and K is an appropriately chosen kernel. The nonlinear classifiers are obtained by solving the following two optimization problems:

$$\begin{cases} \min\limits_{u^{(1)}, b^{(1)}, q} & \frac{1}{2} c_3 (\left\| u^{(1)} \right\|^2 + (b^{(1)})^2) + \frac{1}{2} \left\| K(X_1, X)u^{(1)} + e_1 b^{(1)} \right\|^2 + c_1 e_2^{\mathrm{T}} \xi_2 \\ \text{s.t.} & -(K(X_2, X)u^{(1)} + e_2 b_1) + \xi_2 \geq e_2 \quad \xi_2 \geq 0 \end{cases} \tag{16}$$

$$\begin{cases} \min\limits_{w_1, b_1, \xi_2} & \frac{1}{2} c_4 (\left\| u_2 \right\|^2 + (b^{(2)})^2) + \frac{1}{2} \left\| K(X_2, X)u^{(2)} + e_2 b^{(2)} \right\|^2 + c_2 e_1^{\mathrm{T}} \xi_1 \\ \text{s.t.} & (K(X_1, X)u^{(2)} + e_1 b_2) + \xi_1 \geq e_1 \quad \xi_1 \geq 0 \end{cases} \tag{17}$$

From the KKT conditions, the augmented vector $z^{(1)} = [(u^{(1)})^{\mathrm{T}} \; b^{(1)}]^{\mathrm{T}}$ and $z^{(2)} = [(u^{(2)})^{\mathrm{T}} \; b^{(2)}]^{\mathrm{T}}$ given by

$$z^{(1)} = -(S^{\mathrm{T}} S + c_3 I)^{-1} R^{\mathrm{T}} \alpha \tag{18}$$

$$z^{(2)} = -(R^{\mathrm{T}} R + c_4 I)^{-1} S^{\mathrm{T}} \gamma \tag{19}$$

A new data sample $x \in R^n$ is then assigned to class r $(r = 1, 2)$, depending on which of the two planes given by Formula (15) it lies closest to. Thus

$$\text{Class}(x) = \arg \min_{r=1,2} (d_r(x)) \tag{20}$$

where $d_r(x) = \left(\frac{\left| K(x, C)u^{(r)} + b^{(r)} \right|}{\left\| u^{(r)} \right\|} \right)$, and $\left\| u \right\|$ is the L_2 norm of vector.

4.2.2. Comparison of Experiment Results with Different Filters: To evaluate the performance of the proposed denoising method, different filters have been used to preprocess the traces, then vectors with 120 dimensions are selected by Pearson correlation approach. The train traces cover $\{900, 1350, 1800, 2250, 2700, 3150, 3600, 4050, 4500\}$ and other 900 traces for test.

In this experiment, Gaussian kernel function $K(x_i, x_j) = \exp(-\left\| x_i - x_j \right\|^2 / 2\sigma^2)$ are selected to evaluate the performance of the multi-class classifier. The predictive results of these classifiers will be affected by the choices of parameters, so Grid Search approach is used for the optimal parameters selection. There are four penalty parameters c_1, c_2, c_3, c_4 and one kernel parameter σ. To reduce the computational complexity, we set $c_1 = c_2 = c_3 = c_4$. The optimal values for penalty parameters and kernel parameter are selected from the following range: $c_i \in \{10^{-7}, \cdots, 10^3\}$, $\sigma \in \{2^{-4}, \cdots, 2^6\}$. Figure 6, is the test results of different preprocessing methods based multi-class TWSVM classifier.

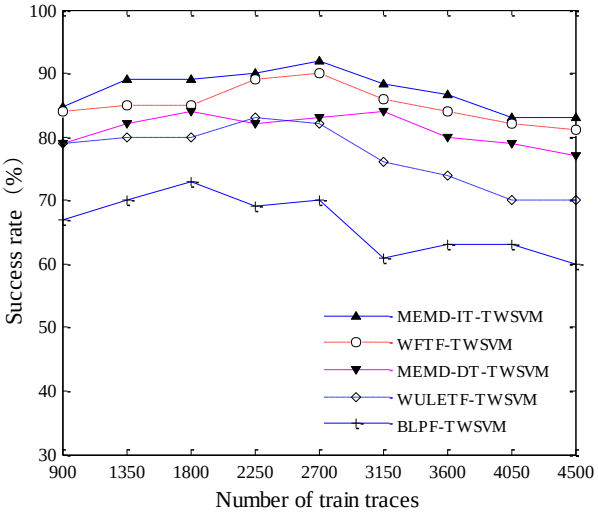

Figure 6. Success Rate of Different Preprocessing

As shown in Figure 6, under the same sample dimension of the feature vectors, the predictive success rate will increase with the increase of the training sample size, then decrease if the sample size is too large. Using the MEMD-IT preprocessing method, the average attack success rate for ETA is nearly 90%. The highest predictive accuracy is more than 92% when the training sample is about 2700. Followed by the wavelet fixed threshold filter, it is about 89%. But it is susceptible to wavelet function and threshold value. The highest predictive accuracy achieve 84%, 83% and 73% for MEMD-DT-TWSVM, WULETF-TWSVM and BLPF-TWSVM, respectively. As the noise is relevant to the encryption chip and the software implementation of the encryption algorithm, the cutoff frequency is a key issue for the Butterworth filter. So the Butterworth filter is dependent on the priori knowledge of side-channel analysis.

5. Conclusion

The electromagnetic radiations preprocessing is the key technique for side-channel template analysis. After analyzing the electromagnetic radiations noise properties of attack targets, MEMD-IT filter is proposed for signal preprocessing. Feature vectors extracted by Pearson correlation approach are used to train TWSVM classifier. Quantities of comparison experiments proved that the MEMD-IT filter is superior to the traditional de-noising methods for signals preprocessing of electromagnetic analysis. The predictive accuracy for 9 Hamming Weights of the key reaches 92% with appropriate train sample size and hyper-parameters, and that can reduce the cipher key search space successfully. The experimental results proved that the proposed electromagnetic analysis scheme combined EMD interval threshold preprocessing with TWSVM classifier model has higher efficiency. It should urge the designers of encryption chip to strongly consider the possibility of physical attacks for trusted embedded devices.

Acknowledgments

This work has been supported by the National Natural Science Foundation of China (61571063, 61202399, 61171051).The authors thank all the partners and the participants in the experiment for their help.

References

[1] L. Lerman, G. Bontempi and O. Markowitch, "Power analysis attack: an approach based on machine learning", International Journal of Applied Cryptography, vol. 3, no. 2, **(2014)**, pp. 97-115.
[2] G. Hospodar, B. Gierlichs, E. D. Mulder, I. Verbauwhede and J. Vandewalle, "Machine learning in side-channel analysis: a first study", Journal of Cryptographic Engineering, vol. 1, no. 4, **(2011)**, pp. 293-302.
[3] Z. Martinasek and V. Zeman, "Innovative Method of the Power Analysis", Radioengineering, vol. 2, no. 22, **(2013)**, pp. 17-20.
[4] Z. Zhong, G. Dawu and L. Junrong, "An improved Side-Channel Attack Based on Support Vector Machine", The 10th international Conference of Computational intelligence and Security, Beijing,China, **(2014)**, September 20-22.
[5] T. H. Le, J. Clediere, C. Serviere and J. L. Lacoume, "Noise Reduction in side Channel Attack Using Fourth-Order Cumulant", IEEE Transactions on Information Forensics and Security, vol. 2, no. 4, **(2007)**, pp. 710-720.
[6] L. Peizhi, "The Research on Template Attack for Block Cipher Chips", Master Dissertation，PLA Information Engineering University, **(2013)**.
[7] S. Chari, J. R. Rao and P. Rohatgi, "Template Attacks", Cryptographic Hardware and Embedded Systems-CHES 2002, Lecture Notes in Computer Science, vol. 2523, **(2003)**, pp. 13-28.
[8] T. S. Messerges, E. A. Dabbish and R. H. Sloan, "Examining smart-card security under the threat of power analysis attacks", IEEE Transactions on Computers, vol. 51, no. 5, **(2002)**, pp. 541-552.
[9] P. Flandrin, G. Rilling and P. Goncalves, "Emperical Mode Decomposition as a filter bank", IEEE Signal Processing Letters, vol. 11, no. 2, **(2004)**, pp. 112-114.
[10] G. Rilling, P. Flandrin and P. Goncalves, "Empirical mode decomposition, fractional Gaussian noise and hurst exponent estimation", Proceedings of the IEEE International Conference on Acoustics, Speech,& Signal Processing, vol. 4, **(2005)**, pp. 489-492.
[11] P. W. Shan and M. Li, "An EMD based simulation of fractional Gaussian noise. Proceedings of the 7th WSEAS International Conference on Instrumentation", Measurement, Circuits and Systems, Hangzhou,China, **(2008)** April 6-8.
[12] B. G. Jeong, B. C. Kim, Y. H. Moon and I. K. Eom, "Simplified noise model parameter estimation for signal-dependent noise", Signal Process, vol. 96, no. 5, **(2014)**, pp. 266-273.
[13] A. O. Boudraa and J. C. Cexus, "Denoising via empirical mode decomposition", Proceeding of IEEE Inter-national Symposium on Control Communications and Signal Processing(ISCCSP), Marrakech, Morocco, **(2006)** March 4-8.
[14] Jayadeva, R. Khemchandani and S. Chandra, "Twin support vector machines for pattern classification", IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 29, no. 5, **(2007)**, pp. 905-910.
[15] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method", Advances in Adaptive Data Analysis, vol. 1, no. 01, **(2009)**, pp. 1-41.
[16] J. He, Z. Qinghua, S. Guoxi, Y. J. Cheng and X. Jianbin, "A Vibration Signal Analysis Method Based on Enforced De-noising and Modified EMD", International Journal of Signal Processing,Image Processing and Pattern Recognition, vol. 8, no. 1, **(2015)**, pp. 87-98.
[17] A. V. D. Ziel, "Noise in Solid State Devices and Circuits", Wiley, New York, **(1986)**.
[18] Y. H. Shao, C. H. Zhang, X. B. Wang and N. Y. Deng, "Improvements on twin support vector machines", Neural Networks, vol. 22, no. 6, **(2011)**, pp. 962-968.
[19] Z. Qi, Y. Tian and Y. Shi, "Robust twin support vector machine for pattern classification", Pattern Recognition, vol. 46, no. 1, **(2013)**, pp. 305-316.
[20] X. Peng, L. Kong and D. Chen, "Improvements on twin parametric-margin support vector machine", Neurocomputing, vol. 151, **(2015)**, pp. 857-863.
[21] D. Tomar and S. Agarwal, "A comparison on multi-class classification methods based on least squares twin support vector machine", Knowledge-Based Systems, vol. 81, **(2015)**, pp.131-147.
[22] J. A. Nasiri, N. M. Charkari and S. Jalili, "Least squares twin multi-class classification support vector machine", vol. 48, no. 3, **(2015)**, pp. 984-992.

Authors

Duan Li, received her Bachelor's degree in Electronic information engineering from PLA Information Engineering University, the Master's degree from Henan Polytechnic University, in 2001 and in 2007 respectively. Currently working on her Ph.D in school of electronic engineering, Beijing University of Post and Teleco- mmunication. She is an assistant professor in Henan Polytechnic University. Her research interests

are digital signal processing, machine learning and Information interception.

Hongxin Zhang, Ph.D., professor and supervisor of Ph.D candidate in Beijing University of Posts and Telecommunications. Besides; Director of communication and microwave engineering laboratory. The peer review expert of the national natural science fund project. His interests mainly in environmental electro-magnetic compatibility, Signal processing in mobile communi-cations and pattern recognition. In important international journals and conferences, he has published over 180 papers including about 100 papers are indexed by SCI or EI.

Qiang Li, received the B.S. degree in communication engineering and the M.S. degree in Signal and Information Processing from the Tianjin University of Technology and Education, Tianjin, China, in 2005 and in 2012, respectively. Currently he is working for the Ph.D. degree at the Beijing University of Posts and Telecommunications. His current research include electromagnetic compatibility, signal Processing and computer security.

Xinjie Zhao, received his B.E. degree and M.E. degree in Department of Computer Engineering, Ordnance Engineering College, Shijiazhuang, China, in 2006 and 2009, respectively. He is currently a Ph.D. student in Department of Information Engineering, Ordnance Engineering College. His main research interest includes side channel analysis, fault analysis and combined analysis of block ciphers. He won the best paper in Darmstadt-the 3rd International Work-shop on Constructive Side-Channel Analysis and Secure Design (COSADE 2012).

