Research on Performance Enhancement for Electromagnetic Analysis and Power Analysis in Cryptographic LSI

Hongying LIU

November 2012

Abstract

Nowadays, various cryptographic devices, such as smart cards, RFID tags, voting machines, cryptographic coprocessor, are widely used to provide authentication of users and to store secrete information for secure systems. Recently, not only the mathematical cryptanalysis, but also the side channel analysis which is based on the physical leakage, such as electromagnetic emission and power consumption, is performed. The security of cryptographic devices becomes a significant research topic.

Among the side channel analysis, the electromagnetic analysis (EMA) and power analysis (PA) can be performed with inexpensive equipment, and hence are powerful and play important roles in security fields. They are included in the fundamental research and involved in practical development. Lots of designs are contrived to prevent cryptographic devices from these attacks. In particularly, the security evaluations on the resistance of these designs based on EMA and PA are also actively conducted. However, due to the complexities of the implementation environment of cryptographic algorithms, there are still problems with conventional analysis for the security evaluation shown as follows: **(1)** The Gaussian noise in the side channels has been reduced by Le (IEEE TIFS07). Diversified non-Gaussian noises that especially occur to EM side channel because of the coupling of encryption modules are still unsolved. **(2)** Although the Hamming Distance leakage model proposed by Brier (CHES04) is generally effective, it overlooks the Glitch power consumption and is not device-specific. **(3)**The classical statistical tests-based leakage localization method, which was proposed by Sauvage (ACM TRTS09), is inaccurate for determining the locations of EM emission. All these problems lead to the deficient performance (the correctness of key detection and computational time) of EMA and PA.

The target of this dissertation is to enhance the performance of EMA and PA. Because the performance of EMA and PA depends on the following key factors: less

noisy signals, accurate leakage model and sound statistical test. This dissertation proposes effective methods for them respectively. The contributions of this dissertation also cover three main aspects: **(1)** Addressing the noise issue that occurs to the EM side channel, especially, for correlated noise and simultaneous noise, several effective algorithms are proposed to decrease their influence on EMA. **(2)** Considering the Glitch power consumption, a new Switching Glitch leakage model is proposed to improve the performance of both EMA and PA. **(3)** To improve the localization accuracy, a novel leakage localization method using instant signal variance based on near-field scan is proposed to enhance the performance of EMA. With all these proposed methods for EMA and PA, the correct key is detected with less computational time, and efficient security evaluation is also made possible.

Therefore, this dissertation consists of 6 chapters as follows.

Chapter 1 [Introduction] gives a brief introduction to side channel analysis. The main principles of PA and EMA are explained. The researches on performance enhancement are overviewed. The motivation and contribution of this dissertation are summarized.

Chapter 2 [Correlated Noise Reduction for EMA] proposes three techniques to reduce the correlated noise for EMA. The correlated noise is caused by the interferences of clock network to the cryptographic module and exhibits strong correlation with encryption signal. The Discrete Wavelet Transform (DWT) proposed by Pelletier (NIST 05) cannot separate this noise because the noise has overlapped frequency bands with encryption signal. It is discovered that unlike the encryption signal, the clock signal has a high variance at the signal edges. Based on this property, the first and second techniques: single-sample SVD and multi-sample SVD reduce the correlated noise by extracting the high variance component from encryption signal. And the third technique: averaged subtraction is efficient when background samplings are included. These techniques are validated by the EM emission acquired from the AES (Advanced Encryption Standard) implementation on both ASIC (Application-Specific Integrated Circuit) and FPGA. Compared with existed methods,

the proposed techniques increase the SNR as high as 22.94dB, and the success rates of EMA shows that the data-independent information is retained and the performance of EMA is enhanced.

Chapter 3 [Simultaneous Noise Reduction for EMA] presents the Source Recovery algorithm to reduce simultaneous noise for EMA. The simultaneous noise is introduced in the EM side channel by running multiple encryption modules simultaneously, which is probably used as an effective countermeasure. However, the fourth-order cumulant-based Gaussian noise reduction strategy presented by Le (IEEE TIFS07) fails to deal with this type of noise. The proposed Source Recovery algorithm takes advantage of the FastICA algorithm (Hyvärinen, IEEE TNN99) to separate the single encryption from mixed encryptions, and then by the amplitude recovery follows the correlation judgment to attenuate the noise. The effectiveness is demonstrated through the analyses of multiple AES and Camellia encryption modules on synthesized ASIC. The number of signals needed to detect keys has been dramatically reduced by 47.8% compared with standard EMA, and the performance of EMA is greatly enhanced.

Chapter 4 [The Switching Glitch Leakage Model for EMA and PA] presents a new leakage model for EMA and PA. The conventional leakage model: Hamming Distance model, which was formalized by Brier(CHES 04), is widely used due to its generality. However Glitch (Glitch is the unnecessary signal transition due to the unbalanced path delays to the inputs of a logic gate in a circuit) effects, which account for 20% to 40% of the dynamic switching power in CMOS circuits, are not involved. This leads to a low performance for EMA and PA. The proposed leakage model not only considers the data dependent switching activities but also includes Glitch power consumptions in cryptographic module. Furthermore, the switching factor and Glitch factor are introduced in the model. The estimation of these factors is shown. The advantage of this model is that the factors can be adjusted according to the analyzed devices during evaluation. Compared with Hamming Distance model, the power traces of recovering keys have been decreased by as much as 24.5%, and the EM

traces has been decreased by as much as 17.1%. Namely, the performance of EMA and PA are both enhanced.

Chapter 5 [A novel Leakage Localization Method for EMA] proposes a novel leakage localization method for EMA. Due to the locality of EM emission, namely, secret information leaks from multiple locations around cryptographic devices, it is challenging to determine the exact location before conducting an EMA. Sauvage (ACM TRTS09)'s localization method has limitation in finding all the data-dependent EM emissions. Based on the EM emission acquired from near field scan, the instant signal variance of EM emission is proved as an equivalent statistical test to DoM (Difference-of-Means) test. Thus, it is proposed to identify the locations that have data-dependent EM emission. Additionally a small and low-cost probe is made to verify the proposed EMA on ASIC implementations. The EMA against unprotected AES indicates that the localization accuracy is improved by 48.6% compared with Sauvage (ACM TRTS09)'s method. Moreover, the EMA on AES WDDL (Wave Dynamic Differential Logic) implementation shows that proposed method is also effective to expose the leakage locations in the presence of countermeasure.

Chapter 6 [Conclusion] concludes this dissertation.

Acknowledgement

First of all, I would like to express my sincere gratitude to my advisor, Professor Satoshi Goto at Waseda University, who has constantly guided and supported me during my doctor course study. I got valuable advices from him in each research meeting and seminar. He inspired my interest in research and also taught me a lot in my life.

I would also like to express my appreciation to Professor Takeshi Yoshimura, Professor Nozomu Togawa, at Waseda University for their suggestions, encouragement and insightful comments in completion of this work.

I also thank Dr. Tsunoo, Dr.Yamashita (NEC Knowledge Discovery Research Laboratories) for advising me in cryptography research. Their broad knowledge in cryptography and signal processing helps me to find the right research directions and instruct me how to continue my work. Thank Mr. Kimura (Y.D.K.Corp.), Mr. Nozawa (Y.D.K. Corp.) and Mr. Syouji (Y.D.K. Corp.) for helping me to dealing with SASEBO, Amplifier, etc.

I also thank the graduated students from Security Group in Goto lab: Mr. Guoyu Qian, Ms.Ying Zhou, Ms.Yue Xing, and Mr.Bin Hu for working with me in side-channel attack. Discussion with them gave me great inspirations in my research work.

In addition, I would like to thank all the students in Goto lab. They gave me a lot of help and advices in my research and made my life in Waseda a wonderful memory.

Finally, I thank my parents, Binlong Liu and Xiumei Wang, and my brother, Xinjun Liu, for their kind support over the years.

Contents

List of Tables

List of Figures

1 Introduction

In this chapter, side channel analysis is introduced. The main principles of PA and EMA are explained. The researches on PA and EMA are overviewed. The motivation and contributions of this work are summarized. The experimental platform is described and the outline of this dissertation is provided.

1.1 Side Channel Analysis

In modern society, the information security is a major concern during the acquisition, storage, processing, and transmission of data. It is conventionally considered that the cryptographic devices are secure because they are based on elaborate cryptographic algorithms and authentication mechanisms. However, this is not true. Various attacks have threatened the security of cryptographic devices. According to the ways of access them, these attacks are classified into three main categories [1]: invasive attacks, semi-invasive attacks and non-invasive attacks.

An invasive attack involves depackaging the cryptographic devices to get direct access to the internal components. For example, an attacker may perform reverse engineering to integrated circuits (ICs). He grinds away the IC layer by layer and takes pictures with an electron microscope to reveal the complete hardware and software part of the IC.

A semi-invasive attack is access the device through the authorized surface without damaging it. The fault-induced attacks are such attacks. The attacker may use a laser beam to influence the operation of the processor, and utilize the incorrect output of this device to deduce the internal data state or the instructions that the processor is running.

A non-invasive attack involves close observation or manipulation of the device's operation. Unlike the invasive attack involves depackaging the

cryptographic devices to get direct access to the internal components. This attack only exploits externally available information which is often unintentionally leaked, such as supply voltage and clock signal of the processor.

An important class of non-invasive attacks is side channel attack or side channel analysis (SCA). It exploits the physical information leaked from cryptographic devices during encryption or decryption to infer secrets.

There are several types of SCA developed during the last decades.

The timing attack is one of the pioneer works, which was presented by Kocher [2] on the international conference CRYPTO in 1996. The time of RSA modular exponentiation was observed and used to recover the secret key.

The power attack, was introduced by Kocher et al.[3] in 1999. They perform power analysis on the implementation of DES. And it is demonstrated as a powerful attack for most straightforward implementations of symmetric and public key ciphers.

Electromagnetic Analysis attacks were also studied by many researchers. The electromagnetic emission was investigated as one of the compromising emanations by military in TEMPEST document presented by National Security Agency [4].

The cache attack monitors the cache misses to recover secret information. It usually happens to CPU which has cache. When CPU accesses data that are not stored in the cache, a delay is engendered for loading the data from main memory to cache. And the measurement of this delay maybe utilized to determine the occurrence and frequency of the cache misses. It was exploited as one of the side channel attacks by Kelsey et al.[5]

There are other side channel attacks that exploit the physical leakage. For example, the acoustic attack that explores the correlation between the sound of a processor and its computation was proposed in [6]. The visible light, e.g., average luminosity of a CRT's diffuse reflection off a wall was reconstructed to recover the signal displayed on the CRT in [7].

The characteristics of SCAs are summarized based on the above introduction, shown as follows.

- SCAs are applicable to almost all the cryptographic devices;
- SCAs threaten a cryptographic device even if its cryptographic algorithm is secure;
- SCAs are easily to mount (they seldom require to modify the hardware design);
- SCAs are likely to occur to cryptographic devices in the absence of awareness;

Therefore, SCAs are very powerful. They play an important role in the researches and applications of security of cryptographic devices.

Among the SCAs, the power analysis (PA) and electromagnetic analysis (EMA) can be performed with inexpensive equipment, and hence represent serious threats to cryptographic devices in hostile environments. They have attracted more attention from research community and the industry. In this dissertation, we also concentrate on the researches of PA and EMA.

1.2 Power Analysis (PA)

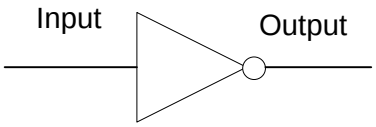
Power analysis is based on analyzing the power consumption of the module while it performs the operation of encryption or decryption.

The principle of PA is introduced below. Various cryptographic devices are essentially based on CMOS technology, and made up of logic gates, such as NOT gate, AND gate, and OR gate, etc.

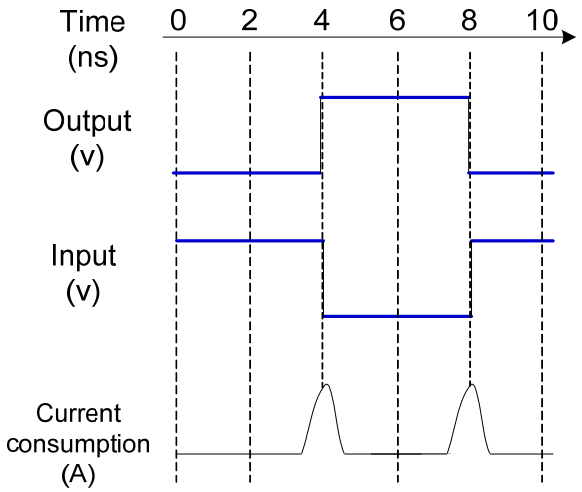
The power consumption of cryptographic module is the sum of its individual gate. For each logic gate, when it works, its output is likely to have 4 types of transitions: $0\rightarrow0$, $1\rightarrow1$, $0\rightarrow1$, $1\rightarrow0$. And the gate draws power from the source. The main power consumption happens in the latter two cases, i.e., from 0 to 1, and from 1 to 0. For these two cases, the gate switches and it is called the

dynamic power consumption occur. And this power consumption is data dependent. A CMOS NOT gate is shown in Fig.1.1 (a), and its power consumption is illustrated in Fig.1.1 (b). When the output switches at 4th ns and 8th ns, the current drawn from the source displays peaks respectively.

For a cryptographic module, all the switches at each logic gate contribute to the power consumption of the circuit. When it encrypts or decrypts, the number of transitions that occur during a certain time interval has correlation with the real power consumption. This is the key idea for power analysis.



(a) CMOS NOT gate



(b) Dynamic power consumption

Fig.1.1 CMOS NOT gate and its dynamic power consumption

The real power consumption is measured across the inserted resistor on the VDD line or GND line. It is shown in Fig.1.2. Typically, an oscilloscope is used to measure the voltage from the resistor. And the current drawn from the source is computed as

$$I = \frac{U}{R} \qquad \text{(Eq.1.1)}$$

where U is the measured voltage, R is the resistance value of the resistor. Since the measured voltage is proportional to the consumed current, it is used to perform power analysis in practice.

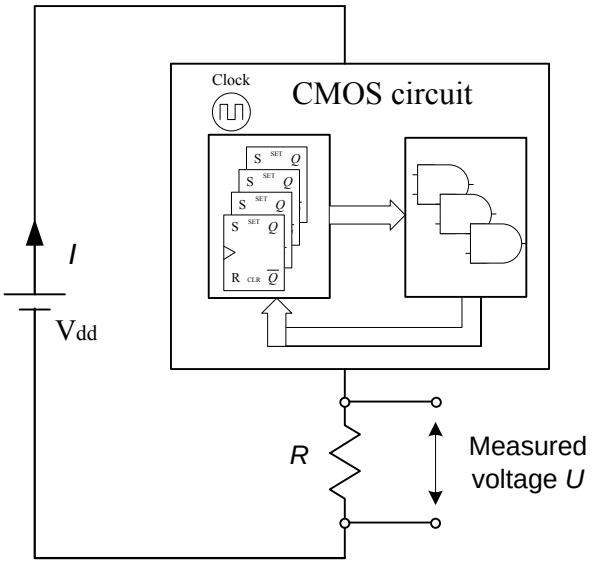


Fig.1.2 The measurement for power analysis

Though the key idea for power analysis is the same, several different types of power analysis were developed.

Simple Power Analysis (SPA) was introduced by Kocher et al. [3] in 1999. It relies on detailed knowledge of the cryptographic algorithm being implemented and visual inspection of the power consumption curves, and tries to extract secret keys. In fact it is quite challenging to recover the keys by only one or very few power traces. The author revealed the sequence of DES instructions executed on a smart card by SPA in [3].

Differential Power Analysis (DPA) was introduced in the same literature with SPA. The author used Hamming Weight model, and the Difference of Means (DoM) test to recover secret key. Compared with SPA, DPA adds statistical techniques to separate signal from noise, and requires less detailed knowledge of the implementation of cryptographic algorithm. The original DPA was single-bit, which based on the fact that the power consumption to switch one bit to 1 is different from the power consumption to switch it to 0. Then it was extended to

multi-bit DPA in [8] and [9], respectively. In general, they are more efficient than single-bit DPA because more bits contribute to the power consumption in most of the implementations.

The Correlation Power Analysis (CPA) was developed by Brier et al.[10] in 2004. The correlation-based power analysis has been suggested in several papers [11, 12, 13]. It was formalized to use the Pearson correlation factor between Hamming Distance and measured power consumption to analyze secret keys in [10].

The Partitioning Power Analysis (PPA) was proposed by Le et al.[14] in 2006 . It is an extension to CPA. The weights are adaptively set for each Hamming Distance in each partition. Though PPA builds a flexible relationship between Hamming Distances and power consumption, the selection of suitable weights for each partition is left open.

Template-based attack is another important type of PA attack, which was initially proposed in [15], and then developed under the name Template Attack [16]. It also exploits the dependency of power consumption on the processed data, but assumes that an identical device is available to the attacked one. It consists of two stages: template building (also named as profiling) stage and template matching stage. In template building stage, a large number of power traces are used to characterize the device, e.g., a multivariate Gaussian distribution is extracted based on the sampled signals, and several templates, which describe the data and corresponding key is built up. In the template matching stage, the secret key is determined by analyzing the given power trace matching the template. The template attack does not try to reduce noise but uses the multivariate-Gaussian noise model to extract information present in a single sample.

The procedure for PA attack is as follows. Firstly, a certain process of the encryption of the algorithm is selected as the target for analysis. Secondly, the real leakage during the execution of encryption is measured. Then statistic analysis is performed based on leakage models, such as Hamming Weight,

Hamming Distances, and using distinguishers such as correlation coefficient, DoM. Finally the secret keys are recovered.

1.3 Electromagnetic Analysis (EMA)

EMA is performed with electromagnetic sensors to extract the secret information from cryptographic devices. The EM sensor is always a magnetic probe, which was suggested in EMC measurement methods [17]. The principle of EMA is briefly introduced below.

An IC which is composed of a number of power lines and signal lines can be simplified as a current loop excited by alternating current source. A time-variant electromagnetic field is generated around the current loop.

The region around a radiant source can be divided into near field and far field in general case. The boundary *d* is given by

$$\boldsymbol{d} = \frac{\lambda}{2\pi} = \frac{\boldsymbol{c}}{2\pi \boldsymbol{f}} \qquad \text{(Eq.1.2)}$$

where λ is the wavelength, *c* is the propagation velocity of EM wave, *f* is the frequency.

The magnetic field generated by the current loop [18] at location P with spherical coordinates $(\boldsymbol{r}, \theta, \phi)$ as shown in Fig.1.3 are given by

$$\begin{aligned} \boldsymbol{H} \simeq & \frac{\boldsymbol{IAe}^{-j\beta r}}{4\pi \boldsymbol{r}^3}[2\cos\theta(1+\boldsymbol{j}\beta\boldsymbol{r})\vec{\boldsymbol{r}} \\ & + \sin\theta(1+\boldsymbol{j}\beta\boldsymbol{r}+\beta^2\boldsymbol{r}^2)\vec{\theta}] \end{aligned} \qquad \text{(Eq.1.3)}$$

where ***A*** denotes the area of the current loop, β denotes a constant of $(2\pi/\lambda)$, $\vec{r}$ and $\vec{\theta}$ represent that the magnetic field ***H*** has two components along the ***r*** and θ direction. $\vec{\phi}$ represents that the electric field ***E*** has component in ϕ direction. μ is the magnetic permeability. In near field, $\boldsymbol{r} \ll \lambda/2\pi$, i.e. $\beta r \ll 1$, the ***H*** field of current loop is given by Eq.1.4. While in far field, $\boldsymbol{r} \gg \lambda/2\pi$, i.e. $\beta r \gg 1$, the ***H*** field of current loop is given by Eq.1.5.

$$\boldsymbol{H} \simeq \frac{\boldsymbol{IAe}^{-j\beta r}}{4\pi \boldsymbol{r}^3}[2\cos\theta\vec{\boldsymbol{r}} + \sin\theta\vec{\theta}] \quad \text{(Eq.1.4)}$$

$$\boldsymbol{H} \simeq -\frac{\boldsymbol{IAe}^{-j\beta r}\beta^2}{4\pi \boldsymbol{r}}\sin\theta\vec{\theta} \quad \text{(Eq.1.5)}$$

The equations indicate that the intensity of measured H-field of a current loop is proportional to its current ***I***.

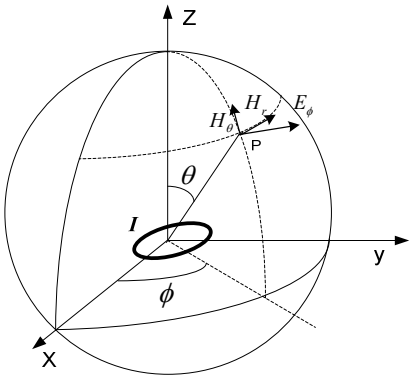


Fig.1.3 The EM field generated by a current loop

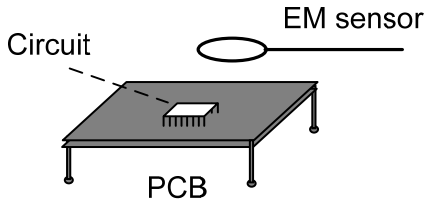


Fig.1.4 Typical measurement for EMA

EMA works because this current ***I*** is caused by the switching activity of the circuit (e.g., from 0 to 1, or vice versa). In other words, this current ***I*** is data-dependent. Thus by monitoring these data-dependent EM emissions, the data handled by the device can be disclosed. Thereby, the more precise the EM field one measures, the more accurate the current can be represented and the switching activity can be revealed, which leads to a faster EMA. A typical measurement is illustrated in Fig.1.4. The measured voltage on EM sensor is proportional to the radiated EM field from data-dependent current in the circuit.

The initial EMA published with experimental results was against smart cards in 2001 by Quisquater et al. [19]. And Simple EMA (SEMA) and differential EMA (DEMA) have been demonstrated by Gandolfi et al.[20] in the same year.

EM leakage has been assessed in [21]. The authors showed that EM emanations not only can be used to attack cryptographic devices where the power side-channel is unavailable, but also can be used to break countermeasures which designed for power analysis. Gebotys et al.[22] showed EMA attacks on a PDA which runs Rijndeael and elliptic curve cryptography.

Because both EMA and PA exploit the power consumption which depends on the cryptographic computation, similar to PA, there are also DEMA, CEMA, etc., classified according to the used distinguishers. Additionally, the procedure for EMA is the same as that for PA, except that the EM emission is acquired not the power leakage.

One important characteristic that discriminates EM side channel from Power side channel is that the EM emission can be measured locally. This may further lead to several positive influences on EMA. (1)With an EM probe of small size and high resolution, the measurements can be carried out over the cryptographic module, which can result in a higher signal-to-noise ratio. The number of measurements is likely to reduce. (2)Because that the EM side channel is measured locally not globally as the case of power side channel, the power consumption which is not related to the processing of encryption, can be attenuated. The “ghost peaks”, which means that the differential curve with the highest peak does not represent the correct subkey, is therefore avoided to some extent. (3) EMA may overcome some of the countermeasures which are implemented against DPA.

1.4 Countermeasures

Countermeasures that protect cryptographic devices from side channel attacks have been actively studied. The PA attack works because the power consumption of cryptographic device depends on intermediate values of the executed

cryptographic algorithm. The goal of every countermeasure is to make the power consumption of a cryptographic device independent of the intermediate values.

They are mainly classified into two categories, the software countermeasure and hardware countermeasure. The software countermeasure tries to implement a cryptographic algorithm in terms of an existing hardware processor, such as a general computer, Advanced RISC Machines (ARM) or digital signal processors. The hardware countermeasure tries to design certain circuits to perform the cryptographic algorithm. There is no clear boundary between these two types of countermeasures. For example, the cryptographic algorithm may be modified and implemented with a customized architecture.

Software countermeasures include the insertion of dummy code, power consumption randomization, etc. Chari et al. [23] proposed to split all intermediate data results using a secret sharing scheme, in which the attacker has to analyze joint distribution functions on multiple points in the power signal. Goubin et al. [24] proposed a similar strategy: duplication method, to protect the DES algorithm from DPA.

Another one of the software countermeasures is selecting the instructions that used for the implementation. Since not every instruction of the hardware platform leaks the same amount of information about its operands, it is possible to choose the instructions that leak the smallest amount of information to reduce the leakage. And the codes which include conditional jumps that depend on the key should be avoided.

The disadvantage of software countermeasures is that they may result in significant memory and execution time overhead. And it is very limited by selecting the instructions.

Hardware countermeasures mainly fall into two categories: masking and hiding. Masking is a method that masks all the intermediate values of circuit by random number. Typical masking is random precharging, and masking buses. In random precharging, random values are sent through the circuit to prcharge all

combinational and sequential cells. The implementation examples are shown in [25] and [26]. The masking buses are used for the implementations that encrypt the data and address buses which connect the processor to memory and cryptographic co-processor. One scheme for bus masking is to exclusive-or random value with the value on the bus. Some publications are [27] and [28].

Hiding conceals the power consumption by inserting dummy operations, dummy cycles or adding power supply filter. And several proposed schemes also tried to affect the clock signal of the cryptographic devices, such as randomly changing the clock frequency, or generating multiple clock signals and randomly switching between them. In [29], a switched-capacitor power filter was demonstrated on an AES core. But it needs additional custom design for its power filter. In addition, a number of hiding countermeasures are implemented with balanced logic cells that try to make the power consumption of each cell constant in every clock cycle for all processed logic values. Examples are the dual-rail precharge (DRP) logic styles. They include Sense Amplifier Based Logic (SABL) which was proposed by Tiri et al.[30], and Wave Dynamic Differential Logic (WDDL) which was proposed by Tiri and Verbauwhede[31].

It is noted that the goals for designing countermeasures are not the same for softwares and hardwares. Goals for software countermeasure are high speed. Goals for hardware countermeasure are maximum throughput and minimum area.

1.5 Overview of PA and EMA

Abundant researches on PA and EMA are carried out, and they have attracted the attention from academic communities and industries. They are mainly summarized into 4 directions:

- Attack methodologies;
- Principles and mechanisms;
- Design of countermeasures;
- Evaluation of implementations;

1.5.1 Attack Methodologies

Plenty of researches are carried out on exploring the advanced attack methodologies.

More distinguishers are developed besides the classical DoM test. They are listed as follows.

- DoM: Presented in DPA [3];
- Pearson Correlation: Proposed in CPA [10];
- Maximum Likelihood: Introduced in Template attacks [32] and used when probability density functions (PDF) can be estimated;
- Covariance: Introduced initially as the multi-bit generalization of the DoM [33];
- Variance: Proposed as an alternative to the mutual information distinguisher [34];
- Spearman rank correlation: Introduced in Rank Correlation Based DPA [35];
- Least square: Introduced in stochastic attacks [36];
- Mutual information [37]: Introduced in Mutual Information Analysis(MIA) from information-theoretic perspective;
- Principal components analysis (PCA): Introduced in differential cluster analysis (DCA) [38], and then presented in [39] which is a typical example of DCA;

There are also other works that propose advanced attack methods mainly against countermeasures, such as second-order DPA attacks that exploit the leakage of two intermediate values related to the same mask.

Against masking. The earliest publication with practical attack is against software masking by Messerges.[40]. Waddle et al.[41] proposed several second order attacks that extend the DPA by preprocessing the signal traces, i.e. computing the difference of the means of squares. The first second-order PA against hardware masking was proposed by Mangard[42]. And the first few work against the asymmetric cryptographic algorithm was the attack on RSA with a secret-sharing scheme in [43].

Against hiding: Clavier et al.[44] proposed the sliding-window DPA to attack the implementations with random process interrupts and noisy power consumption. DPA attack against WDDL was presented in [45].The hidden Markov model was used to attack the asymmetric cryptography with randomized addition-subtraction chains [46].

Moreover, numerous of papers have been published on enhancement of PA and EMA. Agrawal et al.[47] proposed multi-channel attack. It is to combine the acquired power signal and EM emission together for analysis, which leads to enhanced performance of attack. The efficiency of higher-order attacks were studied and improved against masking in [48, 49, 50]. Addressing the misalignment of signal, Homma et al.[51]applied the phase-only correlation in EMA to eliminate the misalignment of the signals on DES implementation. Le et al.[52] proposed the energy-based DPA to overcome the misaligned signal acquire from DES implementation on ASIC. In [53], the WDDL implementation on FPGA was attacked successfully by the EM cartography in frequency domain. In [54], an overview of the application of signal processing techniques to PA and EMA was provided. In [55], the Entropy Power Analysis (EPA) was proposed to attack the protected implementation based on the masking countermeasure. It uses a weighted sum of conditional entropies as a distinguisher. A better success rate is shown when compared with the MIA attack and Variance-based PA attack. In [56], the author proposed an enhanced SEMA, which can find out the demodulation frequencies of the acquired signal. And the effectiveness is presented by the attack against RSA implementation on FPGA.

1.5.2 Principles and Mechanisms

Besides the pioneer works that proposed the DPA, EMA, etc. There are numbers of works that discuss the principles and mechanisms of PA and EMA.

In [57], the weakness of previously known hardware countermeasures was analyzed based on a new model and the secure conditions for the hardware countermeasure were also discussed.

In [58], the classical leakage model was analyzed and a so-called switching distance model was proposed for both PA and EMA. The locality of EM leakage was also indicated.

In [59], the "ghost peaks" problem was analyzed and explained regarding to different power consumption model and various weighting techniques. The properties of S-boxes were studied and the resistance of an s-box against DPA attacks was quantified by introducing the notion of transparency order in [60].

In [61], the mechanism of EMA from remote locations was discussed from the view point of Electromagnetic Compatibility. And the authors claimed that the radiation of cryptographic chip conducts to peripheral circuits was based on ground bounce.

In [62] it is proved that all the distinguishers (e.g. DoM test, correlation factor, maximum likelihood) essentially have the same efficiency given the same leakage model. And the authors also concludes that the correlation factor and information theoretic metric are equally suitable to compare the leakage of devices in unprotected implementations.

1.5.3 Design of Countermeasures

Numerous works have been published on the design of countermeasures for cryptographic devices. Lots of hardware architectures were designed to prevent PA and EMA. The decoupling of the power supply of the cryptographic devices using two capacitors was proposed in [63]. A similar work was presented in [64], which used a three-phase charge pump to supply the power to the devices. Rakers et al. [65] discussed the use of active circuits to reduce the leakage for RFID devices. The non-deterministic processors, which can randomly change the sequence of executed program, was presented by May et al.[66]. In [67], the

authors propose to randomly change the supply voltage and the clock frequency of the circuit.

Some other novel logic styles were proposed besides the SABL and WDDL. The Dual-Spacer Dual-Rail (DSDR) logic style, in which both of the possible precharge values are used alternately, was presented by Bystrov et al.[68]. A similar logic style, named as Three-Phase Dual-Rail Precharge Logic(TDPL) was proposed by Bucci et al.[69]. Moore et al.[70] discussed the use of asynchronous circuits to counteract PA attacks. The Dynamic Current-Mode Logic (DyCML) [71] was proposed as DPA resistant logic style.

Various countermeasures for masking have also been proposed. Messerges[72] presented algorithms for Boolean and arithmetic masking for AES candidates. Akkar and Girand[73] proposed to mask AES S-box. In 1996, the additive or multiplicative masking for RSA was discussed by Kocher in [2]. The similar techniques for ECC were discussed in [74]. The masked AND gate were proposed by Trichina et al. [75]

1.5.4 Evaluation of Implementations

The security evaluations are also carried out on both the unprotected and protected implementations. The DPA and EMA were used to evaluate the security of asynchronous smart-card style device, which were implemented as 16-bit RISC architecture processors by Fournier et al. [76]. The result of assessing the resistance of implementations with WDDL and differential routing against DPA was presented in [45]. The PA was used to evaluate the WDDL and MDPL implementations on FPGA in [77]. The security of Boolean masking for block ciphers was theoretically analyzed in [78]. PA against the DES implementations of WDDL logic style and SecLib (is a logic style that is based on quasi delay insensitive (QDI) asynchronous primitives) logic style, were evaluated by Guilley et al. [79], and the authors concluded that, provided that the back-end of the WDDL module is carefully designed, its vulnerability cannot be

exploited by state-of-the-art attacks. Two analysis methods were proposed based on a preprocessing of the power traces to analyze the DPA-resistant S-box in [80]. A novel analysis method, named as algorithmic collision analysis, was proposed to evaluate the implementations of cryptographic algorithms in [81].

1.6 Contributions of Dissertation

From the procedure for PA and EMA, it indicates that the key reveal depends on side-channel signals, leakage models and statistic test. In general, less noisy signals lead to a faster key detection. An accurate leakage model and sound test can accelerate the key detection.

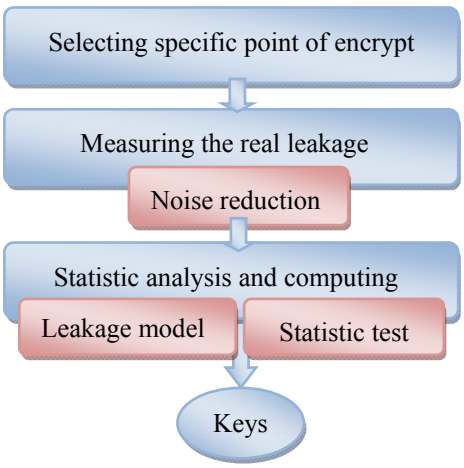


Fig.1.5 Contributions of the dissertation

Therefore, we concentrate on all these three factors, and the corresponding contributions are summarized as follows, shown in Fig.1.5.

- Critical factors for key detections: our contributions
  - less noisy signals : Noise reduction
  - accurate leakage model: Model improvement
  - sound statistic test: Equivalent test methods

1.6.1 Noise Reduction for EMA

In this dissertation, two types of noises that frequently emerge in the EM side channel are thoroughly studied. Consequently new algorithms for the noise reduction are proposed. For the correlated noise, three techniques are proposed, namely: single-sample SVD, multi-sample SVD, and averaged subtraction. For the simultaneous noise, the Source Recovery algorithm is proposed.

The correlated noise is caused by the interferences of clock network in the cryptographic module and exhibits strong correlation with encryption signal. This noise frequently presents in the acquired EM signal in both FPGA and ASIC implementations, which dramatically decreases the performance of EMA. From the observation and simulation, we discovered that unlike the encryption signal, the clock signal has a high variance at the signal edges. Then based on this property, the first and second techniques: single-sample SVD and multi-sample SVD reduce the correlated noise by extracting the high variance component from encryption signal. And the third technique: averaged subtraction is efficient when background samplings are included.

The main characteristics of the proposed techniques are: single-sample SVD can extract the clock signal with only one EM sample. Multi-sample SVD is capable of suppressing the clock signal with short sampling length. The averaged subtraction is suitable for estimation of correlated noise. Furthermore, these techniques are validated by the EM emission acquired from the AES implementation on both ASIC and FPGA. Compared with existed noise reduction methods, the proposed three techniques increase the SNR as high as 22.94dB, and the success rates of EMA shows that the data-independent information is retained and the performance of EMA is enhanced.

The simultaneous noise is introduced in the EM side channel by running multiple encryption modules simultaneously, which is probably used as one

effective countermeasure. However, the conventional noise reduction strategies fail to deal with this type of noise. The proposed Source Recovery algorithm takes advantage of the FastICA algorithm to separate the uncorrelated encryption from mixed encryptions, and then by a difference computation follows the peak judgment to attenuate the noise.

The effectiveness is demonstrated through the analyses of multiple AES and Camellia encryption modules on synthesized application-specific integrated circuit (ASIC). Experiments show that the proposed algorithm recovers the secret key in the presence of the simultaneous noise. The number of signals needed to reveal keys has been dramatically reduced by 47.8%. And the performance of EMA is greatly enhanced. In addition, the results also provide enlightment for the design of countermeasures. It is that the mixed execution of different encryption sources can be bypassed with signal processing techniques, which means it is not an effective countermeasure.

1.6.2 EMA and PA Enhancement Based on A New Leakage Model

In this dissertation, a new leakage model: Switching Glitch leakage model is proposed for EMA and PA.

The conventional leakage model, Hamming Distance model is widely used due to its generality. However in this model, glitch effects, which account for 20% to 40% of the dynamic switching power in CMOS circuits, are not involved. This leads to a low performance for PA and EMA. The Switching Glitch leakage model not only considers the data dependent switching activities but also includes glitch power consumptions in cryptographic module. Furthermore, the switching factor and glitch factor are introduced in the model. And from a theoretical point of view, we show how to estimate these factors. The advantage of this model is that the factors can be adjusted according to the analyzed devices

during evaluation, which makes it device specific and more accurate for the modeling of power consumption.

The PA and EMA on AES implementation validate the proposed model. Compared with conventional Hamming Distance model, the power traces of recovering keys have been decreased by as much as 24.5%, and the EM traces have been reduced as much as 17.1%..

1.6.3 EMA Enhancement Based on A Novel Leakage Localization Method

In this dissertation, a novel leakage localization method that based on the DoM-equivalent statistic test is proposed for EMA.

Due to the locality of EM emission, namely, secret information leaks from multiple locations around cryptographic devices, it is challenging to determine the exact location before conducting an EMA. Based on the EM emission acquired from near field scan, the instant signal variance of EM emission is proved as an equivalent statistical test to DoM test. Thus, it is proposed to identify the information leakage of cryptographic modules. Therefore, by calculating the instant signal variance at each scanning point and computing the higher values, the points that have data-dependent EM emission are disclosed, namely, the leakage locations are found. And the time complexity is also reduced compared with conventional EMA. In addition, a small and low-cost probe is made to verify the proposed EMA on ASIC implementations.

The EMA on AES PPRM1 implementation indicates that misjudgments of the leakage are reduced, and the accuracy is improved by 48.6% compared with existing methods. Moreover, the EMA on AES WDDL implementation shows that proposed method is also effective to expose the leakage locations in the presence of countermeasure.

1.7 Experimental Platform

In this section, the experimental platform for this dissertation is introduced. What is more, a small and low-cost magnetic probe is made for the near-field scan. The making process and verification are also presented in this section. The hardware platforms for PA and EMA are quite different in the research community. Since the ASIC and FPGA are widely applied as cryptographic devices, Side-channel Attack Standard Evaluation Board (SASEBO)-R and SASEBO are used as hardware platforms, which have many cryptographic modules implemented on ASIC and FPGA respectively and provided by National Institute of Advanced Industrial Science and Technology (AIST) [101]. The cryptographic modules used in this dissertation are listed in Table 1.1. The configurations of the two boards are similar. The RS-232 interfaces are connected to the host PC for communication. The analyses are programmed in C language.

Table 1.1 Cryptographic modules on ASIC [101]

Name	Implementation	Gate No.	Area (μm^2)
AES0	Composite field S-box with encryption and decryption	25,483	129,763
AES1	S-box with Look Up Table	20,639	105,097
AES2	* PPRM-based S-box using 1-stage AND-XOR logic	61,801	314,702
AES3	PPRM-based S-box using 3-stage AND-XOR logic	16,541	84,230
AES4	Composite field S-box with encryption only	12,059	61,408
AES11	AES with WDDL (Wave Dynamic Differential Logic)	29,894	152,225
Camellia	Composite field S-box with encryption only	14,416	73,407

*Note: PPRM denotes Positive Prime Reed Muler.

1.7.1 Measurement Setup for PA

A cryptographic LSI and a control FPGA are mounted on PCB of dimension 230 mm x 180 mm x 1.6 mm. The cryptographic cores on SASEBO-R use 0.13μm TSMC standard library of CMOS process technology. DES, AES and some other encryptions are implemented on the LSI. For SASEBO, Two FPGAs of Xilinx Virtex-II pro series are used.

The connection between devices is shown in Fig.1.6. The computer randomly generates 56-bit plaintext for DES (or 128-bit plaintext for AES) in groups, which are transmitted to the FPGA through RS-232 serial ports, and then upon receipt of the plaintext, the FPGA control the LSI to execute DES (or AES) encryption. At the same time, the execution signal on LSI triggers the oscilloscope to start sampling, and thus the oscilloscope acquires and records power signals through coaxial cable, shown in Fig.1.7. The oscilloscope is Agilent MSO 54832D. The sampled data is transmitted to computer through LAN. When the encryption is finished, the ciphertexts are transmitted to computer and used to perform PA.

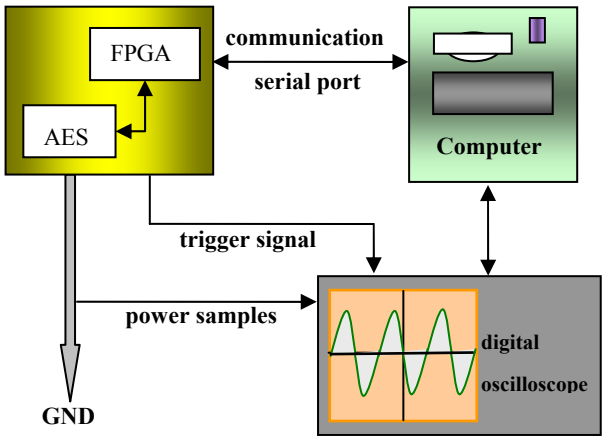


Fig.1.6 Block diagram of experimental environment for power analysis

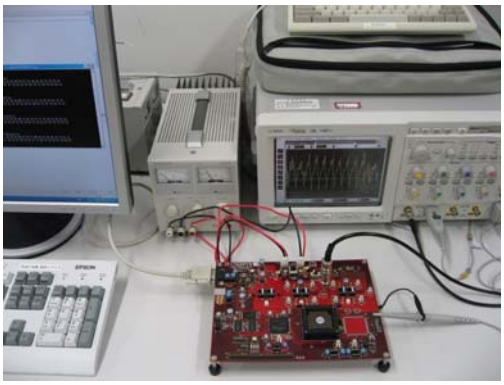


Fig.1.7 Experimental environment for power analysis

1.7.2 Measurement Setup for EMA

For EMA, besides the instruments for PA, a sustentation, a baseplate, an EM probe and a preamplifier are needed.

The connection between instruments for EMA is shown in Fig.1.8. A preamplifier with gain of 51 dB is connected to EM probe through coaxial cable to magnify the weak EM signals before they are sent to the oscilloscope. The experimental environment is shown in Fig.1.9. A 3D-positioning sustentation with scales in three dimensions is used to control the position of the probe above the PCB. A close-up of the scales is shown in Fig.1.10. With this configuration, we can record the specific location of each measurement, which yields repeatable and accurate experiments.

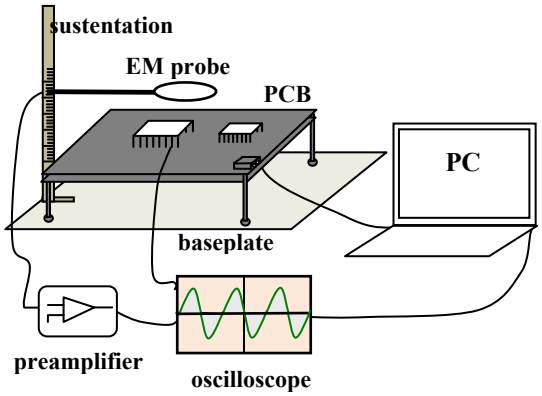


Fig.1.8 Connections between devices for EMA

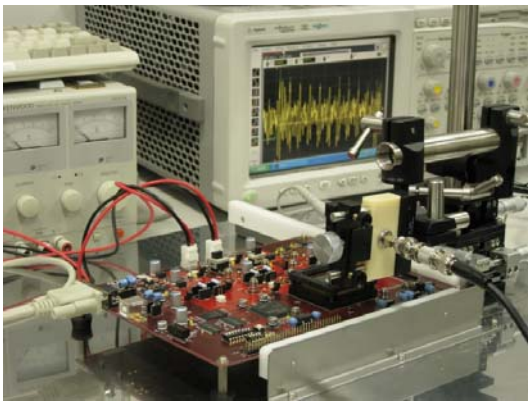


Fig.1.9 Experimental environment for EMA

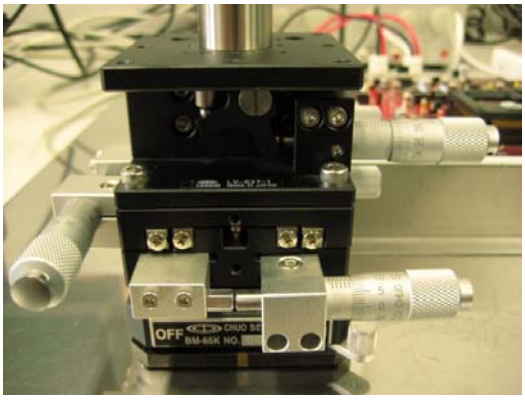


Fig.1.10 Scales on 3D-positioning sustentation

1.7.3 Probe Making

The quality of the obtained EM field signals strongly depends on the utilized field probes. Inductive loop probes are used for magnetic field measurement. Three probes are made out of semi-rigid coaxial cable mounted on a SMA connector. The design process and verification are presented.

A loop probe is sensitive to magnetic field. It outputs a voltage proportional to that field. The voltage V_i induced in the loop probe by an electromagnetic wave [82] is determined from Maxwell's equations and Stoke's theorem, given by

$$\boldsymbol{V_i = -jw\mu H_i NS} \qquad \text{(Eq.1.6)}$$

where $\boldsymbol{H}_i$ is the time variant magnetic field, $\boldsymbol{w}$ is the angular frequency of $\boldsymbol{H}_i$, $\boldsymbol{N}$ is the number of loop turns, and $\boldsymbol{S}$ is the area of the loop.

The distributed capacitance of the probe has no detrimental effect since it is shorted. At low frequencies, the probe acts as a resistance. At high frequencies, this resistance gets negligible against the impedance of the inductance. Thus the frequency response of the probe is flat with a closed circuit.

In addition, the performance of probe is influenced by the number of turns and the area of the loop. In general, a multiple turns of loop has better sensitivity than single loop. Large area loop has better sensitivity than small loop. However, as the loop becomes larger in size, it not only introduces more disturbances in the field being measured, but also reduces the spatial resolution. Therefore, a tradeoff between sensitivity and spatial resolution has to be found.

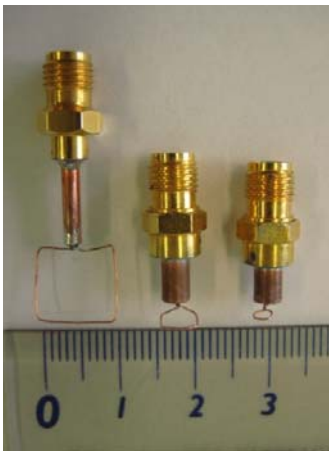


Fig.1.11 Magnetic field probes: MP3, MP2 and MP1

Three single-turn probes: MP1, MP2 and MP3 were made and shown in Fig.1.11 in order to measure the magnetic field near cryptographic LSI. They are in square aperture, which have side length 2mm, 5mm and 10mm respectively, and soldered on the inner conductor of the semi-rigid coaxial,. The diameter of the copper loop is 150 um. Because the aperture of the loop is square and the dimensions of the loop probes are much smaller than the wavelength, the induced electric field is compensated in the loop.

The magnetic field 0.5 mm above a micro-strip line is measured with probes MP1, MP2 and MP3 respectively, shown in Fig.1.12. The micro-strip line terminated with a 50 Ohm load SMA connector is fabricated on a FR-4 substrate

with dielectric constant of 4.5 and a substrate height of 0.8 mm. The length of the micro-strip line is 36 mm, and width is 1.0 mm. The micro-strip line is oriented horizontally along ***X*** axis and the center is at ***Y***=0 mm. Then it is excited by a network analyzer Agilent N3382A with 10 dbm input power at 300 MHz. The probe is controlled by a positioning sustentation and moves along ***Y*** direction at a step of 0.25 mm. The measured amplitudes with three probes are shown in Fig.1.13. The simulation is also plotted.

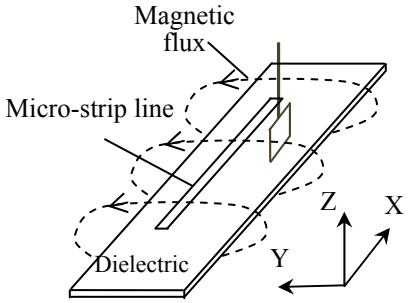


Fig.1.12 Micro-strip line and probe

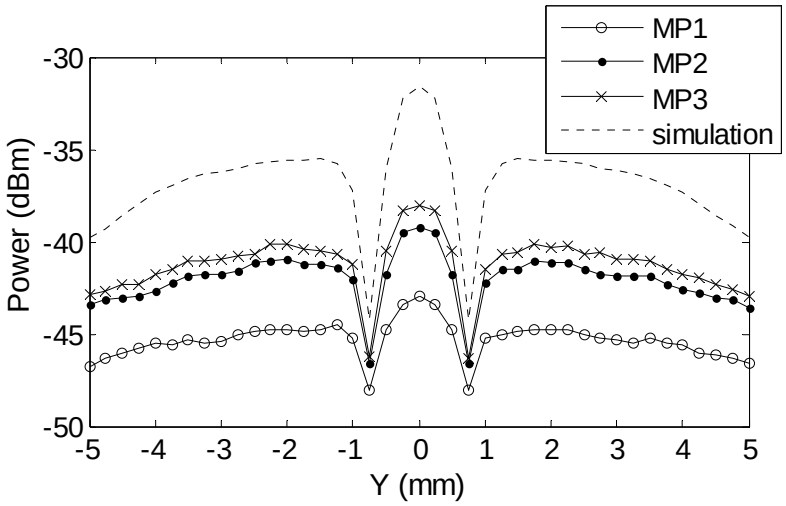


Fig.1.13 Measurement and simulation over micro-strip line

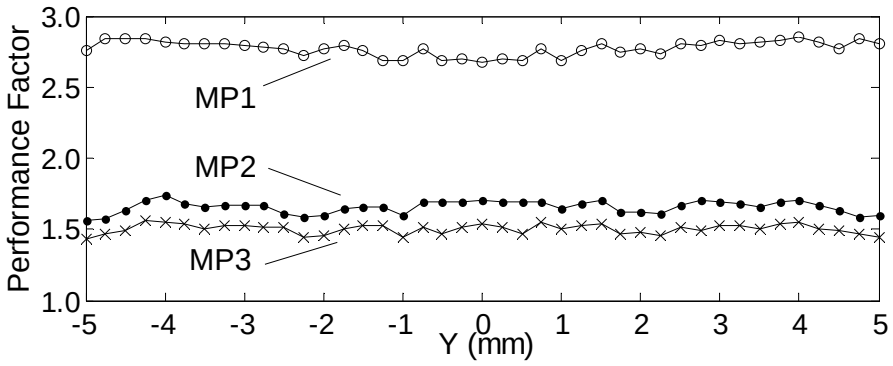


Fig.1.14 Performance factor of 3 probes

Simulations over the micro-strip line are carried out for comparison. An EM simulator based on Finite Element Method (FEM) by Ansoft HFSS is used. The transmitted power is calculated between one port of the line and the port of the probe.

To evaluate the agreement between the measured value and simulated one, performance factor of the 3 probes are calculated and shown in Fig.1.14. Performance factor in time domain is defined as the ratio of the amplitude of the simulated value to the amplitude of the actual measured value. They are expected as straight lines in theory. The standard deviation of MP1, MP2 and MP3 are computed respectively: 0.0233, 0.0331 and 0.0415. The results indicate that measurement from MP1 is more accurate than MP2 and MP3.

1.8 Organization of Dissertation

The remainder of this dissertation is organized as follows. The researches on noise reduction are presented in Chapter 2 and Chapter 3. The three techniques for correlated noise reduction are described in detail in Chapter 2. The algorithm for simultaneous noise reduction is proposed in Chapter 3. The Switching glitch leakage model is proposed, and the performances on both PA and EMA are demonstrated in Chapter 4. The novel leakage localization method is presented,

and EMA on unprotected and protected module is shown in Chapter 5. The conclusion for this dissertation is drawn in Chapter 6.

2 Correlated Noise Reduction for EMA

The issue of correlated noise is studied. Firstly, the background and related works are briefly introduced. Secondly, the problem is defined based on signal model, and the signal characteristics under correlated noise are analyzed. Then the proposed 3 techniques for correlated noise reduction are explained in detail in section 2.3, 2.4, and 2.5 respectively. The experiments are demonstrated in section 2.6. Finally, a summary for this chapter is presented.

2.1 Background and Related Works

The background and related works for correlated noise reduction are introduced in this section.

2.1.1 Noise in Side Channel

Similar to the power side channel, there are several types of noise that influence the EM side channel, such as external noise, which is caused by the environmental interference; and the intrinsic noise results from the physical variation of circuits which is widely occurs to electronic device, as well as the noise introduced by analog-to-digital converter. They are classified as white noise in [83]. Several techniques have been proposed to deal with this noise. The averaging, which was first mentioned by Kocher[3] was used to eliminate the noise which influence the amplitude of the sampled signals. Ryoo et al.[84] proposed the signal companding, which is a non-linear weighting method to minimize the noise presented in the amplitude for DES encryption. The filtering was used to suppress white noise in the work of [21, 85]. Moreover, Le et al.[83] adopted the Fourth-order Cumulant to preprocess the white noise which contains in the acquired EM signal during DES encryption. Charvetd et al. [86] applied

the Discrete Wavelet Transform (DWT) to denoise and smooth the signal. In[89], convolutive noise, which was present in the encryption signal on smartcard, was decreased with the technique of Ceptstrum.

Noises may occur due to the complicated transmission mechanism of EM emissions. For any cryptographic devices, besides the cryptographic module, other modules such as I/O interfaces, Phase-locked loops (PLL) circuit, and clock network, also radiate and may be coupled into the encryption signal during sampling as a result of the superposition of EM fields. They are not directly to encryption activities. Namely, they are data independent.

In [92], Dehbaoui et al. applied the magnitude squared incoherence function to differentiate the EM signals above the DES module and the clock network. The former signals were considered as data dependent and were used to perform EMA. This method uses only two sampling signals in time domain to identify the data dependent signals, which is efficient. However, the noise reduction is left open in [92]. Even if the EM signal is data dependent, it is probably corrupted by correlated noise in EM side channel.

2.1.2 Correlated Noise

The correlated noise in this work is defined as the unwanted signal consisted in the sampled signal, which is correlated by sampling locations. An example of such EM sample is shown in Fig.2.1 (b). This sample is collected over the crypto core of the LSI but close to the clock network. The AES encryption is almost totally immersed in the signals radiated from clock network compared with Fig. 2.1(a). Fig. 2.1(a) is a low-noise AES encryption signal collected above the crypto core, and Fig. 2.1(c) is the clock signal collected above the clock network. The correlation coefficients between the three samples are calculated, and shown in Table 2.1. The EM sample in Fig. 2.1(b) has a strong correlation as high as 0.5676 with clock signal. And in frequency domain, shown in Fig. 2.2, because

that the crypto LSI works at the same frequency with clock signal, their frequency spectrums overlap and covers a wide frequency band.

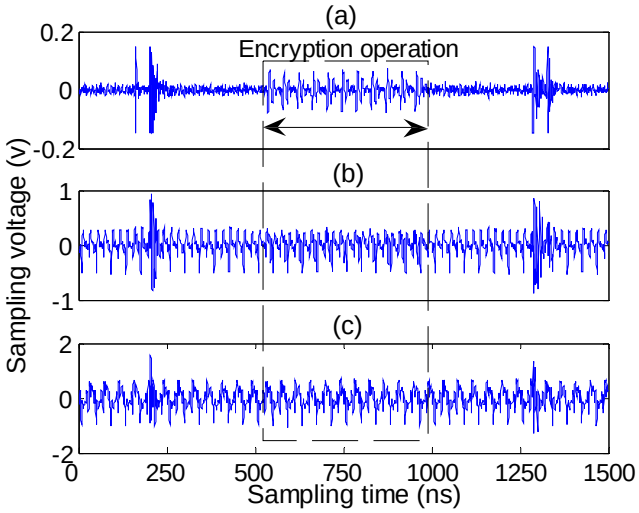


Fig.2.1 Three EM samples in time domain

(a) low-noise encyption signal acquired above the AES crypto core (b) mixed signal (corrupted by clock signal) acquired over the crypto core and close to the clock network (c) clock signal acquired above the clock network.

Table 2.1 Correlation coefficients between EM traces

Corr(*a*,*b*)	Corr(*a*,*c*)	Corr(*b*,*c*)
-0.1780	0.2114	0.5676

a , *b*, and *c* are from the EM samples shown in Fig.1(a),(b),and (c). respectively.

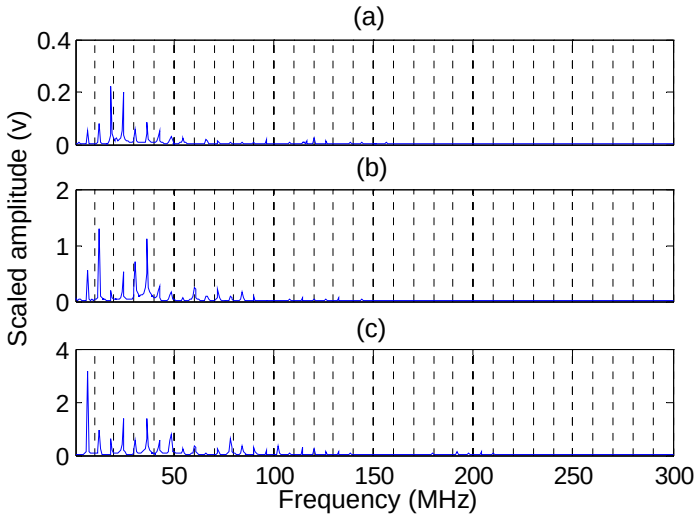


Fig.2.2 The frequency spectrums (0-300MHz) for the three EM samples

The correlated noise in the EM sample is caused by the interference of the clock network, which results from different sampling locations. In fact, if the sampling locations can be properly chosen, e.g., always over the crypto core, such noise can be minimized. However, in general case, during an automatic sampling process, the location of the crypto core can not be determined without preliminary knowledge. The interference from the clock network is unavoidable. The correlated noise occurs frequently in practice.

2.1.3 The Influence on EMA

The correlated signal can be acquired from both the surface of cryptographic ASIC and FPGA. And when EMA performed with such signals, the peak of correlation coefficient for the correct key can not be detected with 10000 plaintexts, which is shown in Fig.2.3.

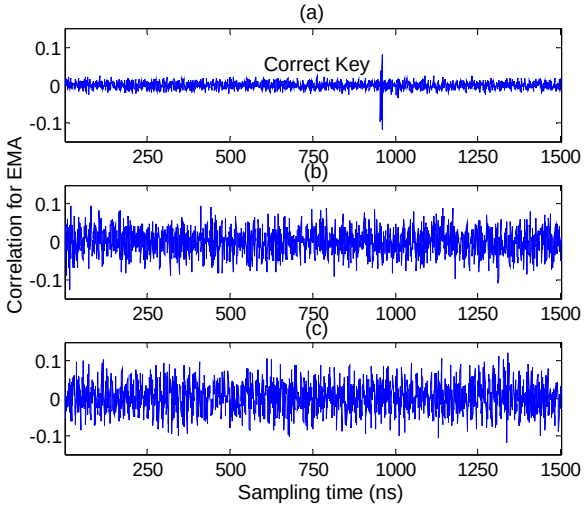


Fig.2.3 Correlation-based EMA of the correct key for signals shown in Fig.2.1(a),(b), and (c)

The conventional averaging can not eliminate the correlated noise. And the filtering, such as the one mentioned in [21, 85] is ineffective due to the severely overlapped frequency spectrum. Because that the mixed signal is not independent, and generally there is only one EM sensor, which does not meet the basic assumptions of the solution such as Difference ICA [90] to separate the mixed

signal. Moreover, since the amount of correlated noise coupled in the sampled signal varies at different sampling locations, it is unable to refer the samples from other locations to reduce the correlated noise.

2.2 Characteristics of EM Traces

In order to solve the noise reduction problem, the signal is modeled and the characteristics of EM traces are studied in this section.

2.2.1 Signal Model

The acquired EM trace during encryption is formulated as follows. For a cryptographic device, generally, multiple modules radiated their own EM emissions during encryption activities of the system. Theses emissions are superposed, and they are picked up by EM probes. The EM probe outputs a voltage proportional to the strength of emissions. And then the emissions are observed and converted to digital signals in time domain by oscilloscope.

Because that devices during this acquisition chain are linear, the measured EM emission W(t) is represented by a combination of the primary source signals: $S_{\mathrm{enc}}(t)$, which is mainly from cryptographic module, such as AES or DES core, contains data dependent information. And $S_{\mathrm{clk}}(t)$ represents the correlated interference that comes from the clock network. η denotes other independent noises, such as white noise etc, expressed by Eq.2.1.

$$W(t) = S_{enc}(t) + S_{clk}(t) + \eta \qquad \text{(Eq.2.1)}$$

where t is sampling time.

The problem is that each source signal can not be measured directly. Only the mixed EM emission is acquired. Owing to the addition of multiple sources, the mixed EM emission covers a wide range of frequency content. Furthermore, the radiated $S_{\mathrm{enc}}(t)$ may contains direct EM emission and modulated emission at certain frequency

due to the coupling of EM field. A formal physical expression is unavailable. Thus, we deduce some characteristics of mixed EM emission by analyzing the characteristic of each source in time domain.

2.2.2 Edge Variance

It is observed that fluctuation of clock signal occurs at the edges and is larger than the fluctuation of encryption signal. We explain this fact by modeling the EM emission.

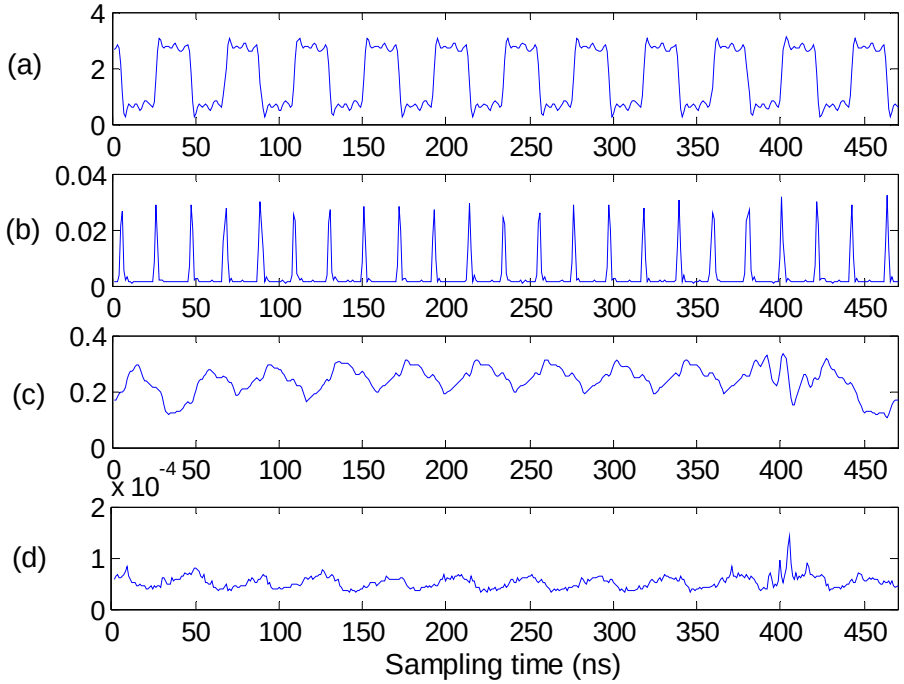


Fig.2.4 Power traces and the computed variances from SASEBO during AES runs

(a) clock signal; (b) variance of this clock signal; (c) encryption signal; (d)variance of this encryption signal. The vertical axis denotes the sampling voltage (in volt). Sampling length is 11 rounds.Sampling time is 470 ns.

(1)The observation of edge variance

The high clock edge variance is observed in the power traces. Fig.2.4 shows the power traces from SASEBO when AES runs. Fig.2.4(a) is the clock signal sampled through the resistor on the board. Fig.2.4(b) is the variance of (a) computed by 5000

traces. Fig.2.4(c) is the encryption signal sampled by the resistor through the power node on board. Fig.2.4(d) is the variance of (c) computed by 5000 traces. It clearly indicates that the clock signal has a much higher variance at the edges.

The high clock edge variance is also observed in the EM traces. Fig.2.5 is the EM traces from SASEBO when AES runs. Fig.2.5(a) is the clock signal sampled by EM probe above the clock oscillator. Fig.2.5(b) is the variance of (a) computed by 5000 traces. Fig.2.5(c) is the encryption signal sampled above the crypto LSI. Fig.2.5 (d) is the variance of (c) computed by 5000 traces. It indicates that for the EM traces the clock signal has a much higher variance at the edges. This is explained by the the fact that the EM signals are radiated by the current in the power/ground network.

Similarly, we also measured the clock signal and the encryption signal from SASEBO-R. The EM traces and their variance are shown in Fig. 2.6.

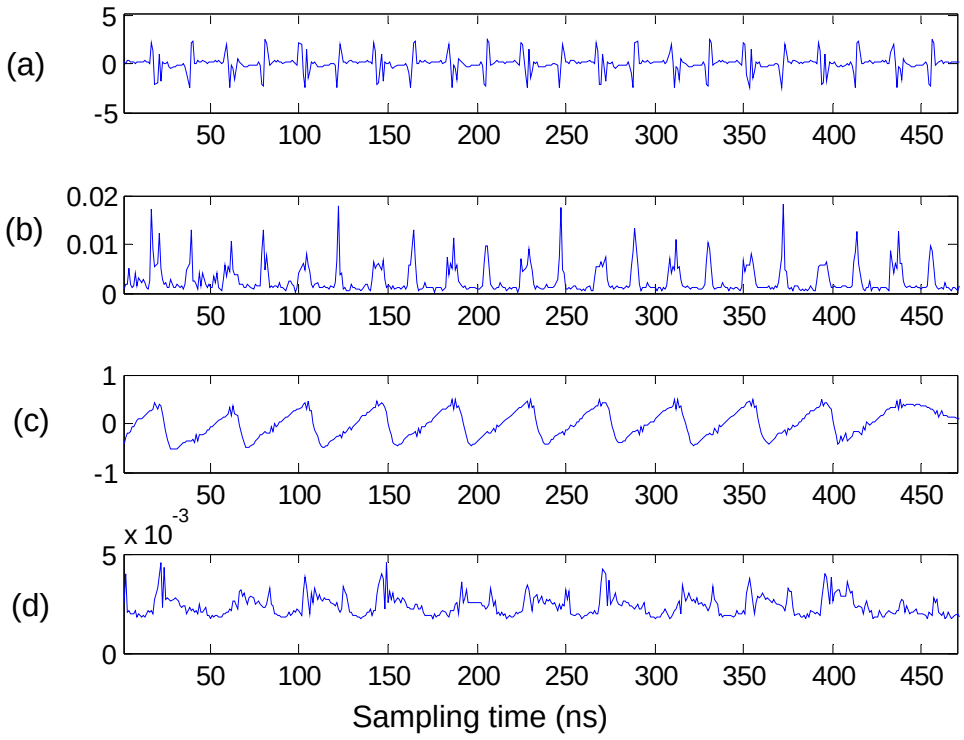


Fig.2.5 EM traces and the computed variances from SASEBO during AES runs

(a) clock signal; (b) variance of this clock signal; (c) encryption signal; (d) variance of this encryption signal. The vertical axis denotes the sampling voltage after the preamplifier (in volt). Sampling length is 11 rounds.Sampling time is 470 ns.

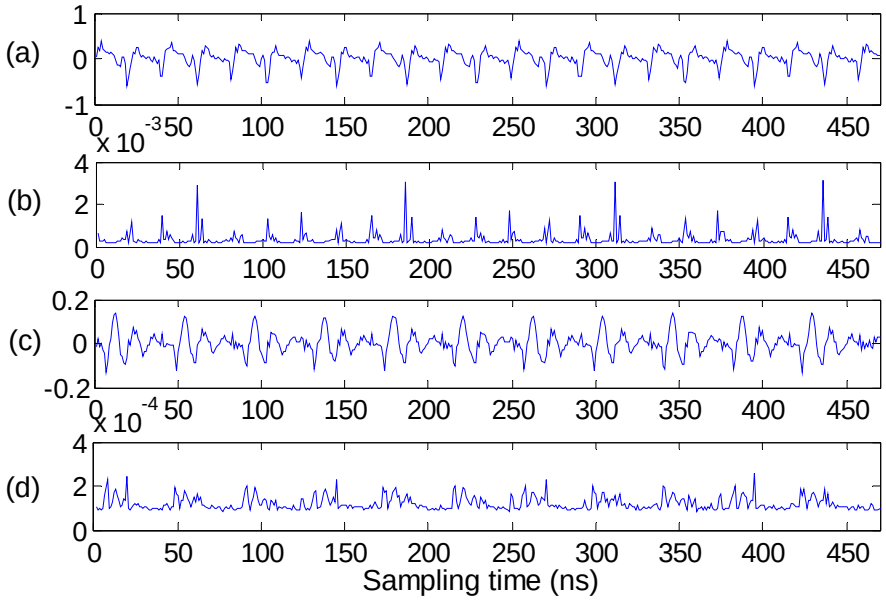


Fig.2.6 EM traces and the computed variances from SASEBO-R during AES runs

(a) clock signal; (b) variance of this clock signal; (c) encryption signal; (d) variance of this encryption signal. The vertical axis denotes the sampling voltage after the preamplifier (in volt). Sampling length is 11 rounds.Sampling time is 470 ns.

(2)The explanation of edge variance

EM emission is radiated by electric current on the circuits during encryption or decryption. The electric current includes not only the one flows on the power/ground network, but also the one flows on the clock network. It is noted that although these two types of electric current originate from similar CMOS primitives, such as logic gates or clock trees. The activities of these primitives are different. The former is data-dependent, and the latter is not. Thus the measured signals from EM field are different though they may both appear spiky due to the wireless transmissions and EM interferences. It is reasonable to model them differently. We describe the characteristics of these emissions according to their source, i.e. the electric current. From the measured current through inserted resistor, it is known that the encryption signal is a slowly changing analogue signal and has a triangle-wave shape, which is different from the clock signal (It has rectangle-wave shape and has shorter rising time and falling time).

Therefore, the encryption signal is modeled by triangle wave and the clock signal is modeled by rectangle wave in the following simulation. They are denoted as Sim_{clk} and Sim_{sig} respectively.

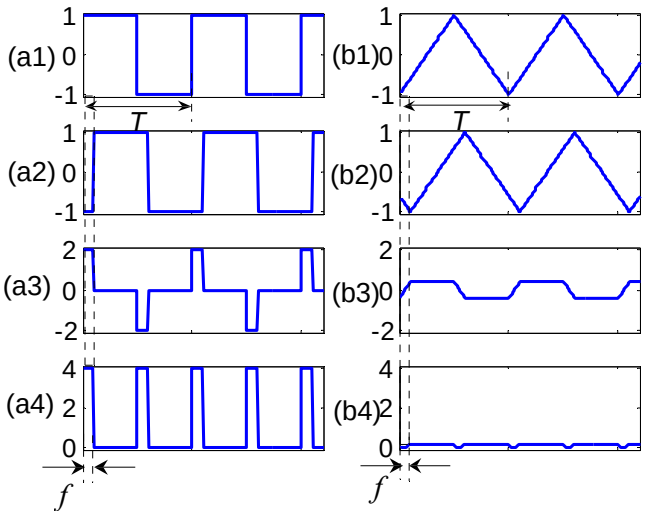


Fig.2.7 Large fluctuations caused by time deviation *f* at clock edges

(a1) and (b1) are clock signal and analogue signal with period *T*, and amplitude 1 respectively; (a2) and (b2) are the expected value of each signal sampled multiple times; (a3) and (b3) are differential curves between each signal and its expected value; (a4) and (b4) are squared deviation of differential curves.

It is generally considered that the clock signal of LSI is constant. This is true except for the short rising and falling time. Due to clock jitter which is the undesired deviation from true periodicity of clock signal and the electromagnetic interference between signals, the slight time difference between each period and fluctuation of the average amplitude of the clock signal probably occur. Because every module of LSI is linked to the clock network, this deviation is dispersed to every module and affects each signal. The fluctuation of the amplitude can be described by the deviation from its expected value, namely variance. Suppose a jitter takes place, which causes a time deviation *f* of the clock signal Sim_{clk} from its expected value. And consequently, the fluctuation is also observed on Sim_{sig}. They are explained by Fig.2.7. The clock signal and the analogue signal have the same period *T*. The time deviation *f* from the expected value is shown in Fig.2.7 (a2) and (b2), respectively. A differential computation is in Fig.2.7 (a3) and (b3). The resulted peaks occur at the edges of clock

signal. Moreover, Fig.2.7 (a4) and (b4) shows the squared deviation. The periodicity of Sim_{clk}, becomes $0.5T$ compared with Sim_{sig}. The resulted amplitude of Sim_{clk} is 400 times larger than Sim_{sig} when $f = 0.025\ T$.

Table 2.2 Simulation results of slight deviation

Time deviation f	Resulted Sim_{clk}	Resulted Sim_{sig}	Resulted Sim_{clk}/Sim_{sig}
2.5×10^{-2}T	4.0	1.0×10^{-2}	$\mathbf{4.0\times10^{2}}$
1.0×10^{-3}T	4.0	1.6×10^{-5}	$\mathbf{2.5\times10^{5}}$

(Note: The original amplitudes of Sim_{clk} and Sim_{sig} are 1, and the periods are T)

In practice, generally, the period of clock jitter is smaller than the rise time of the clock signal according to the datasheet of a typical clock oscillator [104]. For example, for a 50MHz clock oscillator, the period is T=20 ns. The typical period of clock jitter is about 20 ps ($1.0\times10^{-3}T$), while the typical period of the rise time is 2 ns ($0.1T$). Then by simulation with our model, the resulted amplitudes of Sim_{clk} and Sim_{sig} by the small deviation f are listed in Table 2.2.

The second line in Table 2.2 indicates that even if the time deviation is very small, i.e. as short as the period of clock jitter $1.0\times10^{-3}T$, there is a large edge variance for the clock signal compared with encryption signal. Therefore, we have corrected the occurrence of large fluctuation is probably caused by clock jitter or electromagnetic interference between signals.

This result gives hints for the comparison of variance of sampled signals at multiple times. Since the definition of the variance is mean of the squared deviations. It hints that the variance at the clock edges is much larger than an analog signal in the same condition.

2.2.3 Time Delay

Following the explanation of the edge variance, we interpret the timing relation between Sim_{clk} and Sim_{sig} during encryption.

For any cryptographic LSI, all the switching activities of each module are strictly under the control of the clock signal, which is generated from the clock

oscillator and distributed as clock network in the cryptographic LSI. The data begins to change, and the power consumes. In other words, the logic gates switch only when the clock signal arrives at its high level. Consequently, the power consumption signal requires a short time to be observed. In fact, for CMOS clock signal, there is a rising time *RT* before it arrives at the high level and a falling time *FT* preceding the low level. Therefore, there is a time delay Δt between rising point of clock signal and observed power consumption signal (in volt) during encryption, and $\Delta t > RT$, shown in Fig.2.8. *RT* or *FT* is about several nanoseconds for a typical CMOS oscillator, which can be viewed by oscilloscope.

The time delay is further explained in Fig.2.9. Suppose the Sim_{sig} is sampled in correspondence of the rising edges of the clock oscillator. The time delay Δt between (a1) Sim_{clk} and (b1) Sim_{sig} is compared. A jitter causes the both of these signal deviated from their expected value shown in (a2) and (b2). The deviation *f* from the expected value of these signals is shown in (a2) and (b2). Fig.2.9 (a3) and (b3) are the differential curves for each signal. Finally, (a4) and (b4) are the squared deviation for each signal, where P1 denotes the position of the peak value of squared deviation of Sim_{clk}, and P2 denotes the position of the lower value of squared deviation of Sim_{sig}.

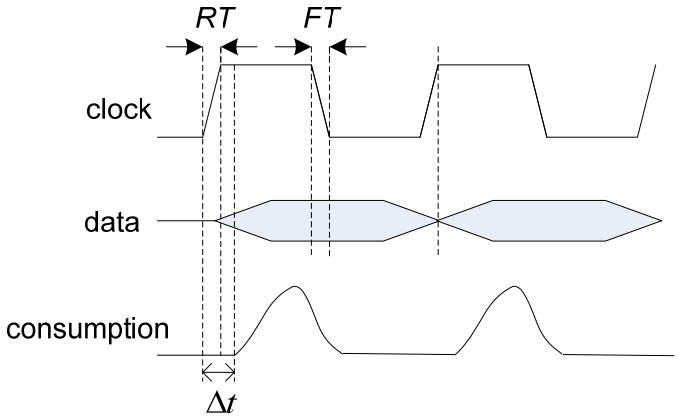


Fig.2.8 Timing relation between the clock signal and power signal during encryption

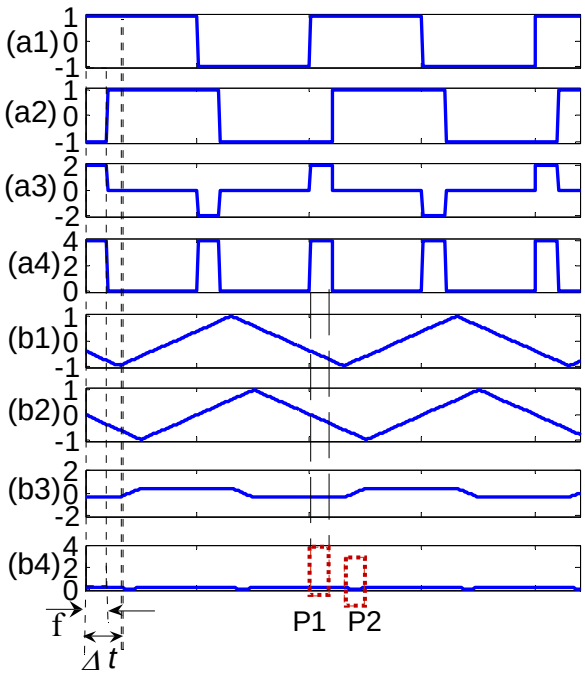


Fig.2.9 The different positions (P1, P2) of peak values of clock signal and analogue signal caused by a time delay

It indicates that there is time delay between the position of peak value of the variance of Sim_{clk} and Sim_{sig}, and the peak value of Sim_{clk} will not be attenuated if the two signals superposed, which is the case for the mixed EM trace. In other words, if the two signals mixed, the variances have a chance to be differentiated due to the time delay.

In the following subsection, we propose three techniques in order to reduce the correlated noise without referring the samples from other locations. Based on the observation that the correlated noise has a high variance at signal edges, the Singular Value Decomposition (SVD)-based techniques which are capable of extracting the signal components with high variance by SVD computation, are proposed to extract the clock noise from the mixed signal. The single-sample SVD takes the advantage of encryption periodicity in neighboring rounds to denoise the high-variance component (correlated noise) in one sample, while the multi-sample SVD extracts the high-variance component using multiple samples

but with short-length. Additionally, the third technique: averaging subtraction is an alternate technique to suppress the correlated noise given sufficient sampling length. Because that the corrupted sample has a strong correlation with clock noise, namely there is a similarity between them, based on this similarity, this technique subtracts clock noise from the sample after averaging to attenuate correlated noise.

The proposed techniques are validated on sampling signals from the AES encryption on both FPGA and ASIC implementations. The SNR gains via the number of samples and length of samples are demonstrated. Their effectiveness for EMA are also compared with existed works: [86][90].

2.3 Proposed Single-sample SVD Algorithm

Suppose that only several rounds of an encryption is available in one sample, which the sampling length L >= 2 rounds. We propose to use single-sample SVD to reduce the correlated noise. By dividing one sample into periodic segments, the SVD can extract the component with largest variance, which is the clock signal. And finally subtract it from the mixture. Therefore, the steps include:

- Period division
- SVD
- Clock subtraction

2.3.1 Period Division

Most block ciphers are designed by iterated computation (each iteration is termed as round), which results in the periodicity of the EM trace. And each round encryption takes one or more clock cycles in practice. Therefore it is possible to make use of the periodicity to divide one EM trace into multiple similar segments.

One acquired EM trace is one signal, which can be represented by a sequence and handled as a row vector.

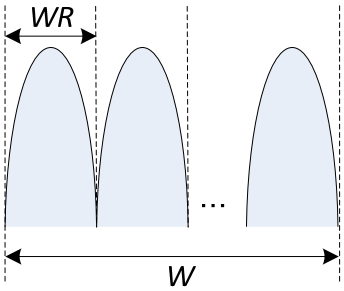


Fig.2.10 The length of EM trace *W*, and the length of its one round *WR*

The cross-correlation function measures the similarity between two signals. We adopt it to divide one EM trace into several segments, and each of them is nearly periodic. The cross-correlation function between two signals *X* and *Y* is defined as

$$Cross((X,Y)_\tau) = \sum_i X_i^* \bullet Y_{i+\tau} \qquad \text{(Eq.2.2)}$$

where signal *X* is denoted as a sequence{x(1), x(2), … , x(n)}, and *n* is the sampling point. *X** denotes the conjugate transpose of *X*, which is defined by taking the vector transpose and then taking the complex conjugate of *X* (for an real-valued signal, its complex conjugate is itself). And *i* is the index for the sampling point, τ denotes an integer offset. Eq.2.2 is a function of offset respect to signal *X*. A peak value of this function means signal *Y* is most similar to *X* at an offset of τ.

Suppose one EM trace is *W* ={w(1), w(2),…, w(n)}, where *n* denotes the sampling point. If we take one-round sample as *WR* ={ w(1), w(2),…, w(k)} with period length *k*, and shown in Fig.2.10, then the offsets that maximize the cross-correlation between *W* and *WR* is given by $H(\tau)$ with length *m*, given by

$$H(\tau) = \underset{\tau=1,2,\ldots,m}{\arg\max}(Cross(W,WR)_\tau) \qquad \text{(Eq.2.3)}$$

Then with these offsets, a single EM trace *W* is divided into a matrix *WP* of consecutive periods of size (*m*+1) x k, given by

$$\boldsymbol{WP} = \begin{bmatrix} \boldsymbol{w}(1) & \boldsymbol{w}(2) & ... & \boldsymbol{w}(\boldsymbol{k}) \\ \boldsymbol{w}(\boldsymbol{H}(1)) & \boldsymbol{w}(\boldsymbol{H}(1)+1) & ... & \boldsymbol{w}(\boldsymbol{H}(1)+\boldsymbol{k}-1) \\ \boldsymbol{w}(\boldsymbol{H}(2)) & \boldsymbol{w}(\boldsymbol{H}(2)+1) & ... & \boldsymbol{w}(\boldsymbol{H}(2)+\boldsymbol{k}-1) \\ ... & ... & ..... & ... \\ \boldsymbol{w}(\boldsymbol{H}(\boldsymbol{m})) & \boldsymbol{w}(\boldsymbol{H}(\boldsymbol{m})+1) & ... & \boldsymbol{w}(\boldsymbol{H}(\boldsymbol{m})+\boldsymbol{k}-1) \end{bmatrix} \quad \text{(Eq.2.4)}$$

Through the period division, the variance of the correlated noise which consists in one encryption round is able to be distinguished by statistical tool.

2.3.2 SVD

Singular value decomposition (SVD) is such a tool that can extract the component of the largest variance from a series of samples. It is widely applied in signal processing, which is a factorization of a real or complex matrix. The SVD of the real-valued WP, is given by

$$WP = U\Sigma V^T \quad \text{(Eq.2.5)}$$

where T denotes the transpose of matrix, U and V are named as the left and right singular vector matrices for $\sum$, respectively. U is in size (m+1)x(m+1), and V is in size k x k. The (m+1)xk matrix $\sum$ is comprised by non-negative singular values $\sigma_1,\sigma_2,\ldots,\sigma_p$, p $\in$ Z, which are ranged in decreasing order: $\sigma_1 \geq \sigma_2 \geq \ldots,\sigma_p \geq 0$ in the diagonal, given by

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 & 0 \\ 0 & 0 & ... & 0 & 0 \\ 0 & 0 & 0 & \sigma_p & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{(Eq.2.6)}$$

In fact, the singular values are the non-negative square roots of the eigenvalues of $(WP)^T(WP)$, and they are the variances of WP. The energy of WP is expressed by

$$\|WP\|_F^2 = \sigma_1^2 + \sigma_2^2 + ... + \sigma_p^2 \quad \text{(Eq.2.7)}$$

where the subscript F denotes Frobenius norm.

According to the characteristics that the variance of the edges of clock signal is much larger than the encryption signal, which means it includes the major energy of

WP and consists in σ_1; therefore the matrix $\sum$ of singular values can be partitioned into three groups, given by Eq.2.8.

$$\Sigma = \begin{bmatrix} \Sigma_{clk} & 0 & 0 \\ 0 & \Sigma_{enc} & 0 \\ 0 & 0 & \Sigma_0 \end{bmatrix} \quad \text{(Eq.2.8)}$$

where $\sum_{clk}$ contains r_{clk} singular values, which corresponding to the clock signal; $\sum_{enc}$ contains r_{enc} singular values, which corresponding to the encryption signal; the $\sum_0$ contains *the* remaining singular values, which represents the other EM interferences and white noise. Thus the contribution from clock signal is :

$$Sclk = \sum_{i=1}^{r_{clk}} u_i \sigma_i v_i^T \quad \text{(Eq.2.9)}$$

where u_{i} is the i_{th} column of *U*, and v_i is the i_{th} column of *V*. *S*clk is a matrix in the same size with *WP*.

2.3.3 Clock Subtraction

The reduction of the correlated noise comes down to the subtraction of the clock signal from *WP*:

$$Wres = WP - Sclk \quad \text{(Eq.2.10)}$$

Then a recovery of this single-sample signal is achieved by assembling each row of the *W*res.

2.4 Proposed Multi-sample SVD Algorithm

If the sampling length is limited L <= 2 rounds, which means there is no sufficient round-signals that can be used to form into periodic matrix WP, then multiple samples are required to compute SVD. The single-sample SVD is extended to the multi-sample SVD.

Suppose *g* samples, each of which has length *c*, with different plaintexts are acquired and have been aligned during sampling by trigger signal.

Theses samples are organized into one matrix W_g of size g x c. Then the following steps are the same as in multi-sample SVD, whereas the resulted Wres are the low noise samples, which can be used to perform EMA, given by

$$W_g = U\Sigma V^T \quad \text{(Eq.2.11)}$$

$$Wres = W_g - Sclk \quad \text{(Eq.2.12)}$$

2.5 Proposed Averaged Subtraction Algorithm

Provided that the sampling length is sufficient, L >= full rounds, but the number of sample is limited. The Averaged subtraction is proposed to get rid of the correlated noise. This technique is based on the following fact: firstly, there is periodicity of the clock signal in one EM trace W. Secondly, clock signals which are present in encryption phase, and non-encryption phase have high similarity. This argument is based on the strong auto-correlation in W. And thirdly, it is shown in section 2.2 that the correlated clock signal is an addictive impact on the EM trace. Therefore, it is possible to extract one averaged, period segment of the clock signal from non-encryption phase, and subtracted from WP. In this way, the impact from the clock signal can be attenuated.

Thus this method includes two steps: clock extraction, and clock subtraction.

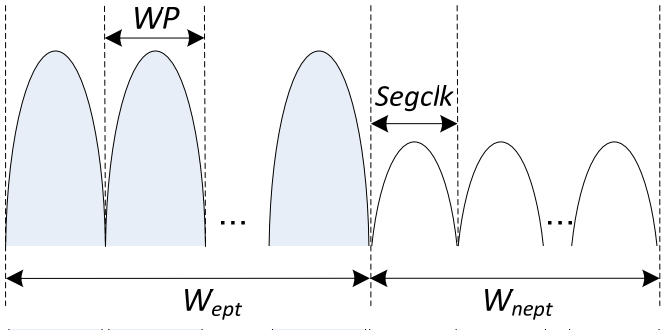


Fig.2.11 The length of encryption phase W_{ept} and non-encryption phase W_{nept}

(WP is the length of one-round encryption, $Segclk$ is one segment of non-encryption signal)

2.5.1 Clock Extraction

It is assumed that the correlated noise only appears in encryption phase W_{ept} for one sample trace W. Referring to Eq.2.4, W_{ept} is expressed as W_{ept}={w(1),w(2),…,w(H(m)+k-1)}. Thus the non-encryption phase is expressed as W_{nept}={w(H(m)+k), w(H(m)+k+1),…,w(n)}, which is considered contains clock signal as well as addictive white noise. Fig.2.11 shows their relation. Owing to the clock synchronization of crypto LSI, the W_{nept} can be divided into J segments *Segclk*(*i*), given by

$$\boldsymbol{Segclk}(\boldsymbol{i}) = \left[\boldsymbol{w}_{\boldsymbol{nept}}((\boldsymbol{i}-1)\boldsymbol{k}+1) \quad \boldsymbol{w}_{\boldsymbol{nept}}((\boldsymbol{i}-1)\boldsymbol{k}+2) \quad \ldots \quad \boldsymbol{w}_{\boldsymbol{nept}}(\boldsymbol{ik}) \right] \qquad \text{(Eq.2.13)}$$

where i=*1,2,...,J,* and J is the number of segments, and k is the length of segment. The remaining sampling points which are insufficient for one period are discarded. Furthermore, to diminish the accidental error which may emerge at each clock signal due to the EM interference, these segments are averaged as $\overline{Segclk}$, given by

$$\overline{Segclk} = \frac{1}{J}\sum_{i=1}^{J} Segclk(i) \qquad \text{(Eq.2.14)}$$

2.5.2 Clock Subtraction

$\overline{Segclk}$ is duplicated to a matrix of (m+1) equivalent rows and subtracted from WP, given by

$$Wres = WP - (\overline{Segclk})_{(m+1)\times k} \qquad \text{(Eq.2.15)}$$

The limitation of this technique is that the subtraction may cause negative amplitudes, which becomes noise. And this technique is applicable to the EM sample, whose length is larger than encryption period.

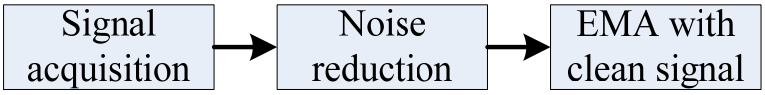


Fig.2.12 Experimental procedures

2.6 EMA Based on Correlated Noise Reduction

The procedures for the experiments are shown in Fig.2.12. Three steps are included.

The signal acquisition is carried out on both FPGA and ASIC implementation of AES, and the platform is SASEBO and SASEBO-R, respectively. The configurations of the devices are similar. The computer randomly generates 128-bit plaintext in groups, which are transmitted to FPGA through RS-232 serial ports, and the execution signal on LSI triggers the oscilloscope for sampling. AES runs at 24MHz, therefore the clock cycle is about 41.6 ns. The sampling rate is 2GSa/s to ensure a sufficient bandwidth. The EM probe is placed 0.5mm over the surface of cryptographic LSI. The EM traces that corrupted by clock signal are captured at the locations close to the clock oscillator, where 10000 traces with different plaintexts are recorded. The key is fixed but randomly chosen in 16-byte (the final round): 28 AF CE 9F 5A FF C8 F1 E0 54 B3 52 B0 CE 43 0E.

Then the proposed three techniques are compared with other two methods for both noise reduction and their effectiveness on EMA. The two methods are DWT which was proposed in[86] and Difference ICA which was presented in [90].

In order to evaluate the performance of proposed techniques, the SNR is defined. Because that we do not have knowledge about the power or amplitude ratio between encryption signal and noise, thus instead of the conventional definition, the SNR is defined as the correlation peak corresponding to the correct key, given by

$$SNR_{e/c}(t) = 20\log_{10}\frac{|A_{peak}|}{|A_{noise}|} \quad \text{(Eq.2.16)}$$

where | | denotes the amplitude of the signal, and A_{peak} is the amplitude of correlation peak, A_{noise} is the average amplitude of the other parts of the curve.

In addition, the effectiveness of proposed techniques is assessed by success rate [18] for EMA. Success rate, which was proposed by Standaert et al., expresses the number of correct subkey guesses among the secret key. Both of these two metrics are

used in the following experiments.

2.6.1 EMA on FPGA Implementation

Noise reductions are carried out with various techniques on the EM traces acquired from FPGA. One corrupted EM trace (unprocessed trace) is shown in first line of Fig.2.13. The sampling points are trimmed to 3000, which involves the full-round AES encryption and background noise.

For the proposed single-sample SVD (S-SVD), only the sampling points of the first 1832 points are used, thus W=[w(1),w(1832)].These points are sufficient to suppress the correlated noise. Meanwhile it avoids introducing the inaccuracy of variances due to the difference of amplitudes. WR is [w(1), w(167)]. SVD is computed on WP, Because that $\sigma_1^2/\sigma_2^2 \approx 784.1$ and it is far larger than 1, which means that the first variance is much larger than the second one and the energy is concentrated in the first component, the first singular value is picked up to project the clock component from the mixture. Finally, the results from clock subtraction is assembled back to a sample and combined to the complete sample.

For the multi-sample SVD (M-SVD), in fact the selection of points of interested is not constrained. We used 3000 sampling points in order to compare with other methods. Similarly, After SVD, σ_1^2/σ_2^2 is computed. Its value is 625.3 which is much larger than 1 and indicates that the clock noise mainly consists in the first singular value. Thus the first singular value is used to extract clock component.

In the technique of averaged subtraction (A.Subt.), 6 clock segments are averaged and subtracted. Additionally, the “Symlet” family of DWT is used to process the mixed signal as in[86]. And a repeated sampling is performed with another EM probe putting close to the clock oscillator in order to obtain two groups of samples which satisfy the condition of Difference ICA[90].

The signals after denoising are shown in Fig.2.13. And the correlation peak of the correct key with 10000 each signal is shown in the right column. The extracted clock noise by proposed S-SVD and M-SVD are shown in Fig.2.14 (a) and (b), respectively.

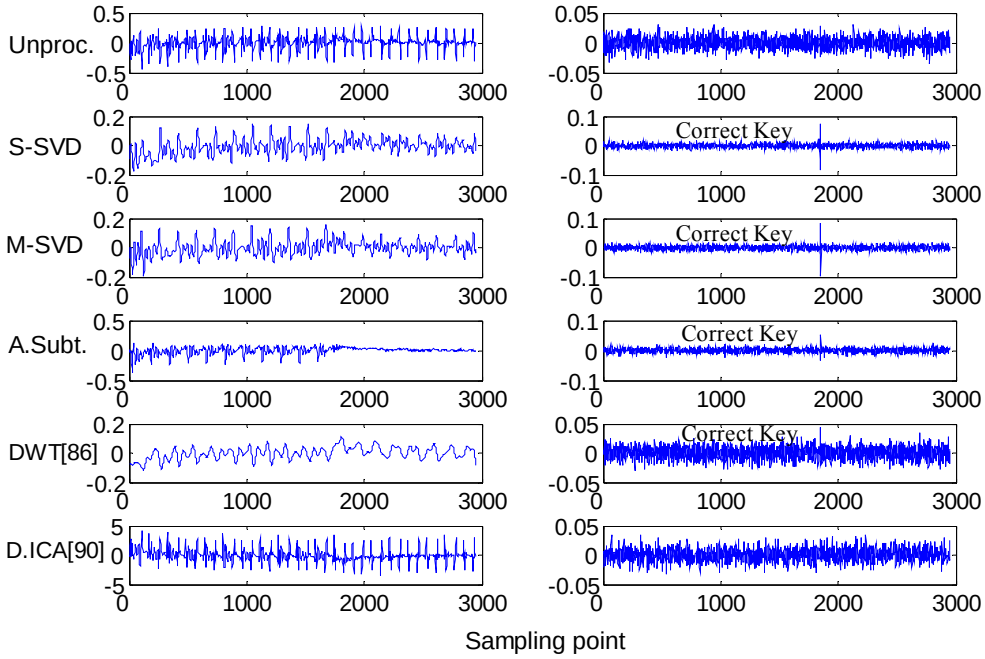


Fig.2.13 EM signals from SASEBO

Left column: The unprocessed EM signal, and signals processed by Single SVD(S-SVD), Multi-SVD (M-SVD), Averaged subtraction (A.Subt.), DWT[86], and Difference ICA[90]. Right column: correlation peak of EMA corresponding to the correct key when 10000 traces are used.

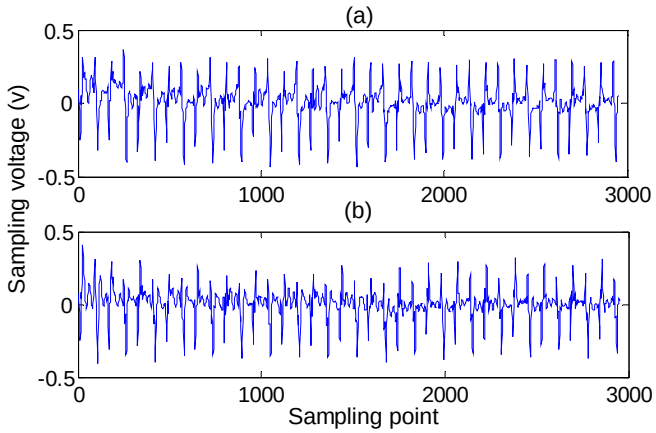


Fig.2.14 Extracted clock noise by proposed SVD-based methods on SASEBO (a) by S-SVD (b) by M-SVD.

Table 2.3 Average SNR comparison of signals on FPGA

	Unproc.	S-SVD	M-SVD	A.Subt.	DWT[86]	D.ICA[90]
SNR(dB)	-0.86	14.49	21.58	8.08	2.45	0.03

(10 groups, each with 10000 different samples)

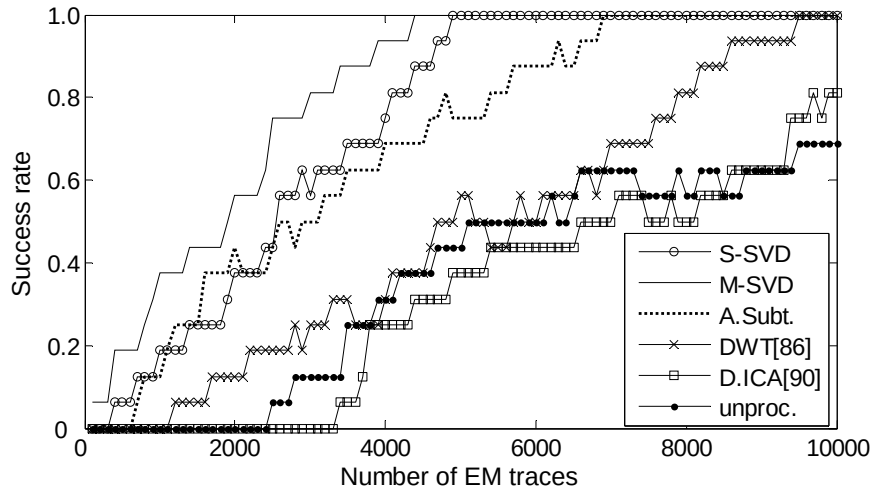


Fig.2.15 Success rates of EMA on SASEBO with unprocessd signal and the noise reduced signals

In order to evaluate the results quantitatively, 10 groups of acquired EM traces with correlated noise, which with 10000 samples are processed with each method. The averaged SNR of each signal are calculated according to Eq.2.16, shown in Table 2.3.

The S-SVD and M-SVD have larger SNR compared with the unprocessed signal and other methods. They are as high as 14.49 dB and 21.58 dB respectively compared with other 3 methods. This is because that the clock component which contains strong edge variances is extracted by SVD method. And M-SVD has higher SNR than S-SVD, because that more samples are used and the singular value is computed more accurate. By contrast, for DWT[86], though by setting the wavelet coefficients of the low-valued details to zero, white noise can be filtered. This does not have effect on the correlated noise, which covers the same frequency bands with encryption signal. And it indicates that the Difference ICA [90] removes the noise only to some limited extent. The algorithm does not decorrelate

the mixed signals adequately, which leads to its low performance. The negative SNR occurs because the correlation for peak of the correct key is smaller than the amplitude of noise.

EMAs are performed with these 6 groups of signals to further confirm their effectiveness. The output of s-box in the final round of AES is chosen as a target to analyze. Then the Pearson correlation is computed to recover all the 16 sub-keys.

The success rates versus number of traces are plotted in Fig.2.15. EMA with M-SVD succeeds fastest, and all the sub-keys are recovered with only 4308 traces. S-SVD is with 4825 traces, and 6861 traces for Averaged subtraction. The DWT requires 9425 traces. And Difference ICA only recovers 81.25% with 10000 traces. This result is in consistent with the SNR. EMA performs better with the proposed methods, since higher SNR are achieved.

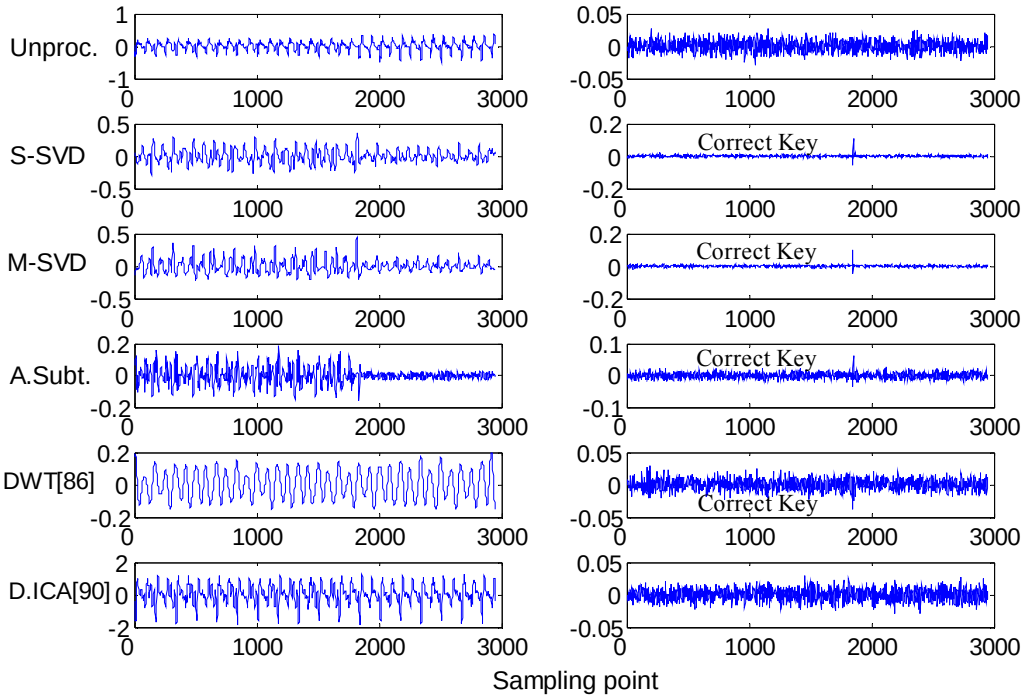


Fig.2.16 EM signals from SASEBO-R

Left column: The unprocessed EM signal, and signals processed by Single SVD(S-SVD), Multi-SVD (M-SVD), Averaged subtraction(A.Subt.), DWT[86], and Difference ICA[90]. Right column: correlation peak of EMA corresponding to the correct key when 10000 traces are used.

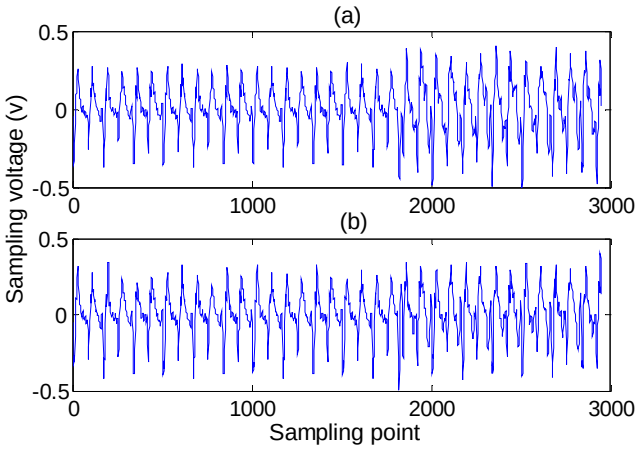


Fig.2.17 Extracted clock noise by proposed SVD-based methods on SASEBO-R(a) by S-SVD (b) by M-SVD.

2.6.2 EMA on ASIC Implementation

The correlated noise is also observed on the implementation of cryptographic LSI on ASIC.

The unprocessed signals and signals after noise reduction with each method are shown in Fig.2.16, respectively. 3000 sampling points for each signal. Similar to the FPGA experiments, σ_1^2/σ_2^2 is calculated. It is 1224.7 and 961.4 for the S-SVD and M-SVD method, respectively. The first singular values are extracted. The extracted clock noise by proposed S-SVD and M-SVD on SASEBO-R are shown in Fig.2.17 (a) and (b), respectively. It shows that after the noise reduction, the peak of correct key is exposed by S-SVD, M-SVD, A.Subt., and DWT. The averaged SNR is listed in Table 2.4. S-SVD performs best, which has an averaged SNR of 23.06 dB. By contrast, the average SNR for DWT is only 4.13 dB. The correct key is immersed in the background noise for the Difference ICA method and the unprocessed signal.

Table 2.4 Average SNR comparison of signals on ASIC

	Unproc.	S-SVD	M-SVD	A.Subt.	DWT[86]	D.ICA[90]
SNR(dB)	-3.87	23.06	20.45	11.14	4.13	-1.86

(10 groups, each with 10000 different samples)

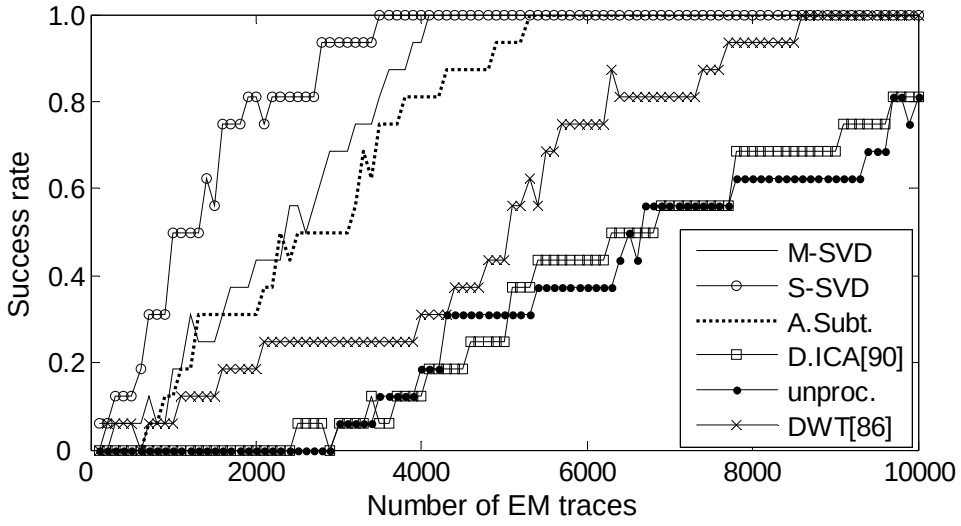


Fig.2.18 Success rates of EMA SASEBO-R with unprocessd signal and the noise reduced signals

The success rates of EMA for recovery all the sub-keys with each method are shown in Fig.2.18. All the sub-keys are revealed by S-SVD, M-SVD, A.Subt., and DWT within 10000 EM traces. Only 3319 traces are required for S-SVD to recover all the sub-keys, 4127 traces for M-SVD, and 5413 traces for A.Subt. While the success rate for D.ICA and unprocessed trace is only 81.25%. These results are similar to the results given by FPGA experiments.

Then the performances of proposed techniques are further investigated. They are related to the variation of number of samples and the length of samples. SNR Gain means that the gain of SNR compared with unprocessed EM traces. The number of samples is the sample used for preprocessing the EM trace. It always refers to the number of samples stored and processed online. It can be very small, such as one, two, etc. This is different from the meaning of number of traces, which are the traces required for key recovery. Since EMA relies on the statistics of EM traces, generally at least several hundred traces are required to reveal the key. Thus this discussion is meaningful for the real time processing of EM samples.

2.6.3 Performance Evaluation

The discussion about the performance of proposed techniques is provided below. The variation of SNR Gain of S-SVD, M-SVD, and A.Subt., along with the number of samples is shown in Fig.2.19. For this result, the length of sample is fixed at 3000 points. When only one sample is involved in the preprocessing, the S-SVD achieves a SNR gain of 14.49 dB by utilizing the periodicity of encryption rounds. And the A.Subt. has a gain of 8.08 dB. By contrast, the gain for M-SVD is around 0, because that this method depends on the variance computation and it is not available for a single sample. When the number of samples increases, the SNR gain of M-SVD rises since the computation of variance become accurate. It arrives as high as 21.58 dB when there are 10000 samples. However, the computational cost for preprocessing these samples also increases. It indicates that the S-SVD can achieve a relative high SNR gain when the number of sample is limited.

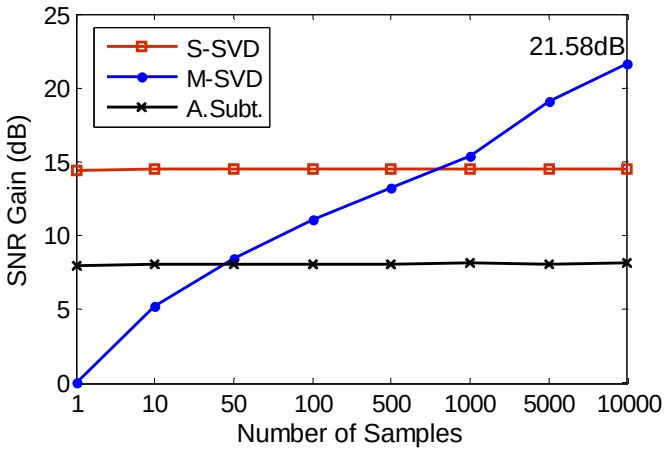


Fig.2.19 The variation of SNR Gain of proposed 3 techniques along with the number of samples

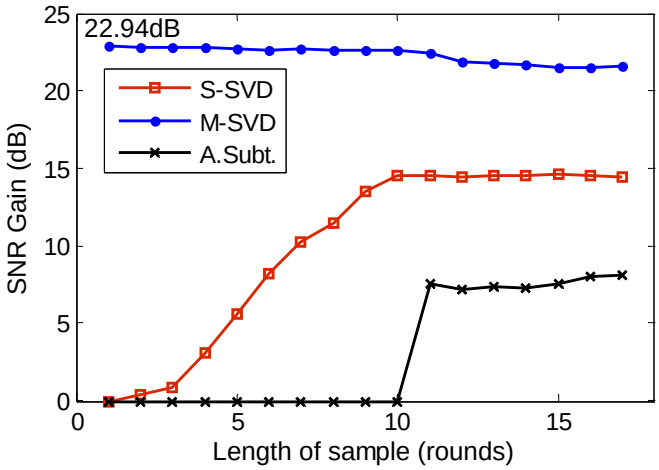


Fig.2.20 The variation of SNR Gain of proposed 3 techniques along with the length of sample

The variation of SNR Gain of S-SVD, M-SVD, and A.Subt., along with the length of samples is summarized in Fig.2.20. This experiment is performed with 10000 samples for each method. When the sample involves only one round of encryption (It is assumed that the interested round, which is used for key recovery, is included), the M-SVD achieves a high SNR gain of 22.94 dB by computing the variances among all the available samples, while S-SVD can not, which leads to 0 gain. And similarly, the A.Subt. can not be performed, since there is no extra length of sample for subtraction. The gain of A.Subt. increases unless the length of sample exceeds the full encryption round, and its performance is bellow S-SVD and M-SVD.

Thereby, it indicates that M-SVD has best performance when the length of sample is limited. And A.Subt. is a coarse estimation if the background noise sampling is available.

The proposed three techniques can be applied to the online pre-processing of EM samples. The online pre-processing is needed especially for the EM cartography or near-field scan based EMA, in which it needs to collect and store many samples. These samples are stored on oscilloscope temporarily. Because the storage capacity M_{on} of oscilloscope is limited, the number of samples N_{on} that can be stored and the sampling length L_{on} have constraints. Generally, it has

$$M_{on} = N_{on} \times L_{on} \quad \text{(Eq.2.17)}$$

In this case, it is necessary to select an appropriate method. Given storage capacity M_{on}, when L_{on} < 2 rounds, it is better to select M-SVD because N_{on} can be larger, which means that multi samples are available. When L_{on} >=2 rounds, it is better to use S-SVD (N_{on} is smaller). And when L_{on} >= full rounds, A.Subt. also can be used (N_{on} is smaller).

In addition, the proposed three techniques can be applied to various scenarios for EMA. It is noted that these three techniques reduce the correlated noise in time domain, whereas the applications are not constrained. For S-SVD, after the correlated noise is attenuated, the clean signal can be transformed to frequency domain for performing EMA. Since the sampling length L >= 2 rounds, the clean signal has a fine frequency resolution (The frequency resolution is computed by Fs/N, where Fs is the sampling rate, and N is the number of sampling points. Given sampling rate, if the sampling length is larger, namely the number of sampling points is larger, then the frequency resolution is finer). This leads to efficient EMA. In addition, since more than one round's signal is available, the comparative EMA between different rounds is enabled. For M-SVD, it is more appropriate for EMA in time domain. Because the sampling length L < 2 rounds (the interested round for key recovery is included), the data-dependent information concentrates in the limited sampling points, the computational time for the comparison of the correlation-based EMA can be saved (This is similar to the key idea of "compression of power traces"). For A.Subt., as the sampling length is larger than the full rounds, the EMA in frequency domain is more efficient. Moreover, since the background noise is included in the sampled signal, the comparative study on the noise characteristics at different locations is possible, and this is helpful in finding the best location for EMA. A summarization is shown in Table 2.5.

Table 2.5 The applications of proposed methods

Sampling length L	Method	Applications
L < 2 rounds	M-SVD	Time domain EMA;
L >= 2 rounds	S-SVD	Frequency domain EMA; Round comparison of EMA;
L >= full rounds	A.Subt.	Frequency domain EMA; Location finding for EMA;

2.7 Summary

In this chapter, the correlated noise consisted in EM emission, which is commonly occurs to cryptographic devices, is studied. The characteristics of such signal are investigated based on observation and analyses. They indicate that the clock signal of the cryptographic system has a high variance at signal edges. Therefore, the single-sample SVD and the multi-sample SVD are proposed to reduce the correlated noise by extracting the clock component from the mixture. The single-sample SVD is suitable for online preprocessing of EM trace since only one sample is needed to suppress the correlated noise. The multi-sample SVD is able to reduce the correlated noise without much constraint of sampling length unless the points of interest are included. Additionally, the third technique: averaged subtraction is effective for a coarse estimation of clock noise.

Furthermore, these techniques are validated by the EM emission acquired from the AES implementation on ASIC and FPGA. Compared with existed noise reduction methods, the proposed 3 techniques increase the SNR as high as 22.94dB, and the success rates of EMA shows that the data-independent information is retained and the performance of EMA is improved

3 Simultaneous Noise Reduction for EMA

In this chapter, the simultaneous noise is studied. Firstly, we give a brief introduction about the noise in side channel. Secondly, the proposed algorithm is presented in detail. Then the EMA with the proposed algorithm is shown and compared with EMA based on bandpass filtering [90]. Finally, a short summary is concluded.

3.1 Background and Related Works

The necessary background for our work is introduced in this section.

3.1.1 Intentional Noise

The noise discussed in subsection 2.1.1 is mainly unintentional. Noise in side channel can be intentional, which means noise is introduced as countermeasure to prevent the EMA or PA. It is named as **noise-countermeasure**. This is another category of typical countermeasures, besides the application of logic styles, hiding, and masking. It makes the PA/EMA difficult or impossible by either adding extra hardware or random processing to change the signature of power consumption of the cipher.

There are mainly 3 of such noise-countermeasures in the literatures: variable clock, random delay insertion, and correlated power-noise generator, shown as follows.

(1) Variable clock

It consists in clocking a chip with an internal oscillator whose parameters (frequency, duty cycle, shape, etc.) vary randomly in time. It was mentioned and by Kafi, et al[105]. These authors also proposed the parametric deconvolution to reduce it. Another solution to this noise is proposed in [86]. In that work, DWT

was used firstly to denoise the power traces, and then “Simulated annealing” algorithm was used to resynchronization the power traces

(2) Random delay insertion (RPI)

The delays can be inserted randomly during the operations of a cipher, e.g. delays are inserted before SubByte during AES computation, as this operation is more vulnerable. Such implementation leads to the misalignment of the collected signal traces. The solutions are actively studied. Clavier et al.[44] restored the original amplitude of the power traces by integrating the RPI-protected signal over several consecutive cycles. Homma et al.[87]proposed phase-based waveform matching method. Gebotys et al.[88]introduced the phase replacement technique.

(3) Power-noise generator

Kamoun et al. [106] proposed a power-noise generator for AES. The generator is composed by the most vulnerable AES functions, and works concurrently with the AES core but using different keys. However, to our best knowledge, there are no solutions discussed in the literatures.

3.1.2 Simultaneous Noise

The simultaneous noise is defined as the unwanted signal, which is mixed with interested signal but independent from interested signal in this work. The simultaneous noise belongs to class (3) Power-noise generator, and it is one of the noise-countermeasure. It was discovered during our experiment. When multiple crypto modules work simultaneously, some of the modules act as noise generator to other interested modules. Its generation and reduction is studied in the following part.

In the literatures, simultaneous switching noise (SSN) [119] is well known as a phenomenon with adverse and severe effects when a large number of high speed chip drivers switch simultaneously causing a large amount of current to be

injected into the power distribution grid. It was actively studied in [120-123]. In the field of side channel analysis, the hardware decoupling has been a technique for reducing the conduction of the SSN current of logic ICs [124]. The simultaneous noise defined in this dissertation is kind of implementation of SSN in cryptographic LSI.

The advantage of such noise generator is that there is not any hardware cost. Compared with the noise generator proposed in [106], the Camellia module acts as a noise generator for AES module, and they use the same key to encrypt in this work. There are no additional gate counts; meanwhile it plays the role of countermeasure to prevent the attacks.

In order to meet demands for the various encryption standards, multiple cryptographic modules are integrated to one LSI. For example, AES, DES, Camellia, and RSA are likely produced on one ASIC to support for both the private and the public cryptosystems, shown in Fig.3.1. Multiple AES modules (AES0, AES1,…, AESn) and Camellia are integrated on one ASIC.

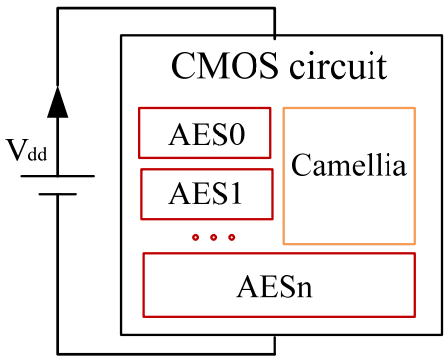


Fig.3.1 Multiple cryptographic modules on circuit

A short description of AES and Camellia is provided below.

Advanced Encryption Standard (AES) was published by the National Institute of Standards and Technology (NIST) of United States in 2001 [107]. It has become one of the most popular symmetrical encryption algorithms. It has a 128-bit block size, with key sizes of 128, 192 and 256 bits. AES is designed to be easy to implement on hardware and software, as well as in restricted

environments and offer good defenses against various attack techniques. And plenty of research works are carried out on the security evaluation of AES.

Camellia [108] is a symmetrical key block cipher developed jointly by Mitsubishi Electric Corporation and Nippon Telegraph and Telephone Public Corporation (NTT in short) in 2000. And it is specified in ISO/IEC 18033-3. Camellia's block size is 16 bytes (128 bits), and can use 128-bit, 192-bit or 256-bit keys. The block cipher was designed to be suitable for both software and hardware implementations, from low-cost smart cards to high-speed network systems. The cipher has security levels and processing abilities comparable to AES.

In general, only one encryption module runs and the corresponding EM signals are measured and collected during an EMA. However, in order to hide the data-dependent information from attackers, multiple encryption modules may run simultaneously. This is considered as simultaneous noise, which is an effective countermeasure that slower the key detection. For instance, AES0, AES1 and Camellia run at the same time, shown in Fig.3.2. The plaintext is input for each of the 3 modules, and the same key is used for encryption, and finally the ciphertexts are output as Ciphertext 1, Ciphertext 1, and Ciphertext 2, respectively, since the algorithms of AES and Camellia differ. The only difference between AES0 and AES1 is that they have different S-box structures: AES0 is Composite-field based S-box, and AES1 is Look-up -table based S-box.

There is reason for their simultaneous run. AES and Camellia have the same block size: 128 bits, the same key size, and work under the system clock, although the structures of the ciphers vary: AES is Substitution-Permutation Network (SPN) implemented in 11 rounds while Camellia is Feistel Network (FN) structure in 23 rounds. Given sufficient sampling time (larger than 23 rounds), a mixed power curves can be collected, which includes the power signature of AES and Camellia.

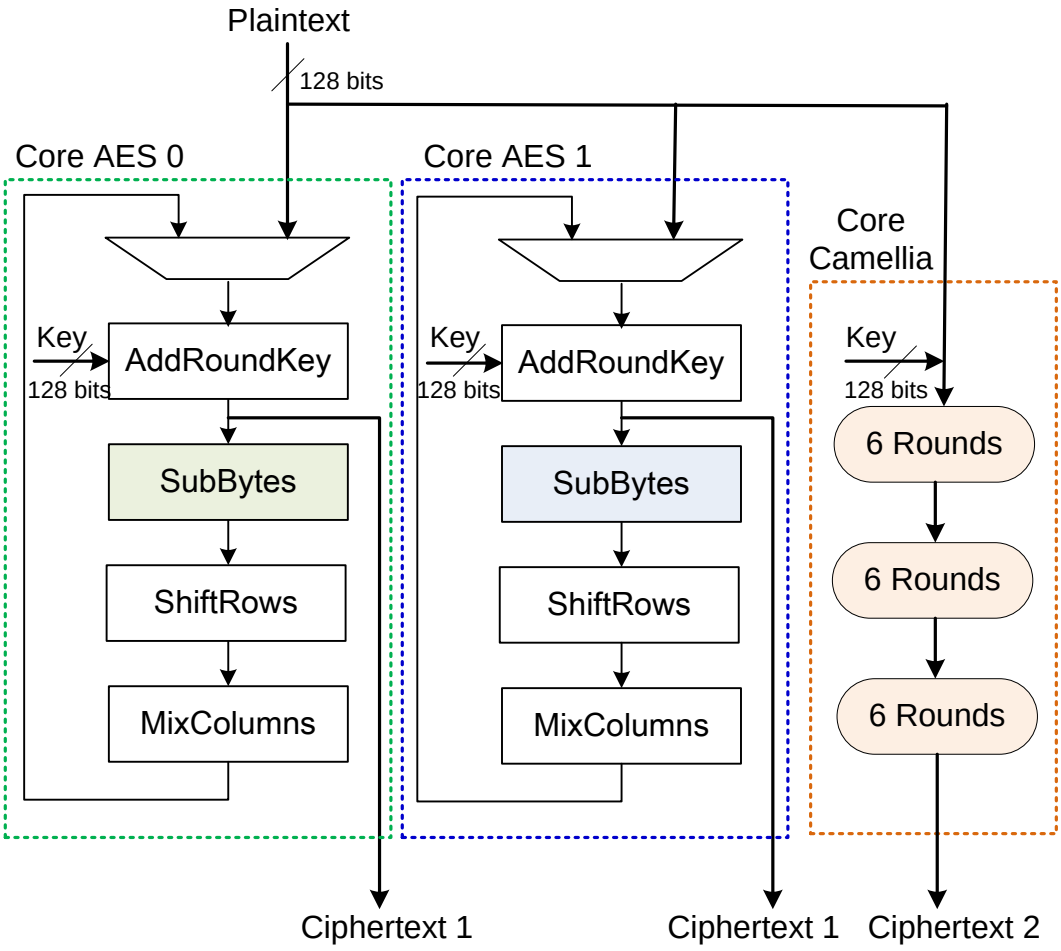


Fig.3.2 The generation of the simultaneous noise: 2 AES modules and Camellia work simultaneously

When this kind of noise generator works, it prevents the EMA on interested cipher. For example, Camellia acts as noise generator to prevent AES key detection. When EMA or PA is performed against AES, a number of EM traces or power traces and ciphertexts are required. The signal traces denote the power consumption of AES. The ciphertexts are used to compute the data-dependent switching activities in Hamming Distance (HD) model for ASIC device. (A detailed introduction about the models is in subsection 4.1.1) It assumes that the power consumption of AES P_{AES} is proportional to the data-dependent switching activities HD_{AES}, given by

$$P_{\mathrm{AES}} \sim HD_{\mathrm{AES}} \qquad \text{(Eq.3.1)}$$

However, when AES module and Camellia work simultaneously, the ciphertext of AES is used to compute HD_{AES}, but the power consumption becomes P_{AES} + $P_{Camellia}$, since the power traces have altered, shown as

$$P_{AES} + P_{Camellia} \not\sim HD_{AES} \quad \text{(Eq.3.2)}$$

namly, the power consumption is no longer proportional to the data-dependent switching activities HD_{AES} . Thus, Camellia alters the power signatures for AES key detection, and it is a noise generator in this case.

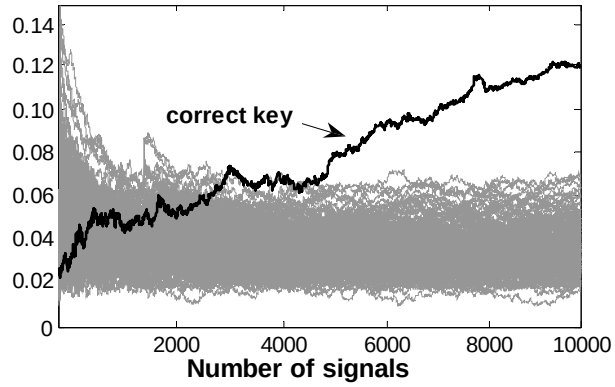
Therefore, the requirements for such noise generators are: (1) it is implemented in the same LSI with interested cipher in order that they work in the same clock cycle. (2) It shares the input channel with the interested cipher. Usually this requires the other cipher is the same class and has the same block size with the interested cipher. (3) Its key size is the same with the interested cipher. There are other noise generators for AES besides Camellia. Here, AES is the interested cipher. For example, the block ciphers: Serpent (block size :128 bits, key size: 128, 192 or 256 bits), Twofish (block size :128 bits , key size: 128, 192 or 256 bits), SEED (block size :128 bits , key size: 128 bits), RC5 (block size :32,64, 128 bits , key size: 0 to 2040 bits). (4) There should be one interface-register that acts as control circuit, for example, when some bits of the register is set to 1, both of these crypto modules could run. Otherwise, only one crypto module runs.

An example of such noise generator is shown in detail below. The key detection process becomes quite difficult.

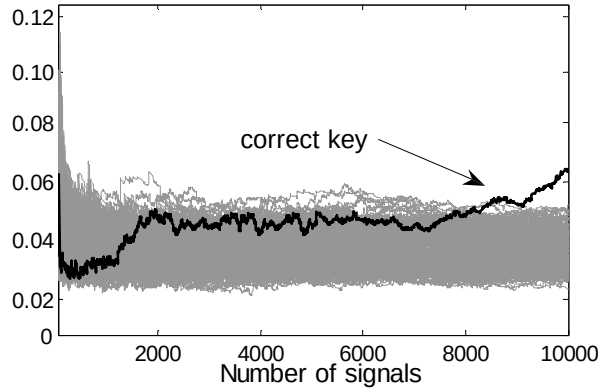
On the ASIC, we activate module AES0-AES4 and Camellia simultaneously, record the mixed signals, and perform EMA. The 16 byte-keys are detected within 8291 signals, the evolution of the second byte key “AF” is shown in Fig.3.3(b). By contrast, only 3614 signals are needed when only AES0 runs, namely without simultaneous noise, shown in Fig.3.3(a).

The simultaneous noise can not be reduced by existed works as mentioned in subsection 2.1.1, because averaging [3]: attenuates both the interested signal and

noise; The Fourth-order Cumulant-based method [83] reduces the wide-band Gaussian noise; Signal compandings [84] only compress/expand the amplitude of signal; And Digital Filtering [21,85] to remove the signal and noise together at certain frequency bands. These methods are not capable of reducing the simultaneous noise, which is from another encryption source, appears as strong as the interested signal, and has the same frequency band with interested signal. Therefore, new algorithms are in need to reduce such noise.



(a)EMA without simultaneous noise



(b) EMA without simultaneous noise

Fig.3.3 The evolution of the second key byte: "AF"

3.1.3 Blind Signal Separation

Aiming at probing into the possible solutions to reduce simultaneous noise, we studied the problem of blind signal separation (BSS). It is explained as follows.

Consider that there are a number of signals emitted by some physical objects or sources, such as the electric signals emitting by different areas of brain, the radio signals emanated by mobile phones or the speech signals etc. Then the sensors receive and record these signals in the form of a mixture of the original source signals. We are interested to find the original source signals from the mixture with little or no knowledge about the source signals.

ICA has been proved as an effective way to solve this problem. And it has been applied to separation of different speech signals, analysis of EEG data, functional magnetic resonance imaging (fMRI) data, and as a model of biological image processing. There are plenty of derivative algorithms of ICA.

According to the computational complexity, the algorithms fall into two categories: the low-computational algorithms, which are based on the first-order or second-order statistics, such as AMUSE[109,110], GED [111], SOBI[112], and the gradient-based algorithm [113]. And the algorithms based on high-order (larger than second-order) statistics of the mixed signals, such as JADE[114], EASI[115], [116], [117], and FastICA[91].

The basic model for these algorithms is shown as follows. Assume that we observe *m* linear mixed signals X of *n* independent source signals (*m*>=*n*)

$$\boldsymbol{X} = \boldsymbol{AS} + \boldsymbol{N} \qquad \text{(Eq.3.3)}$$

where $X=(X_1,X_2,\ldots,X_m)^T$, is *m* mixed signals which are observed, $S=(S_1,S_2,\ldots,S_n)^T$, is the *n* source signals, $N=(N_1,N_2,\ldots,N_m)^T$, denotes the *m* noise vector, superscript T denotes transpose of matrix. All of these signals have sampling length L. Then, after estimating a matrix **W**, the independent component can be obtained by: **S** =**WX**.

FastICA is introduced as follows.

The FastICA is based on a fixed-point iteration scheme to find a direction, i.e. a unit vector **W** such that the projection $\mathbf{W}^T\mathbf{X}$ maximizes nongaussianity. The algorithm is as follows, given by Eq.3.4-3.6.

$$\boldsymbol{Z} = \boldsymbol{QX} \qquad \text{(Eq.3.4)}$$

$$\boldsymbol{W}^{+} \leftarrow \boldsymbol{Z}g(\boldsymbol{Z}^{T}\boldsymbol{W}) - \boldsymbol{W}g'(\boldsymbol{W}^{T}\boldsymbol{Z})\boldsymbol{O}_{Lx1} \quad \text{(Eq.3.5)}$$

$$\boldsymbol{W}^{+} \leftarrow \boldsymbol{W}^{+} / \left\| \boldsymbol{W}^{+} \right\| \quad \text{(Eq.3.6)}$$

where $\boldsymbol{X}$ is centered and whitened to simplify the computation. **Q** is unitary matrix, $E(ZZ^{T})=\mathbf{I}$ is satisfied, where E is the mathematical expectation, g is the non-linear contrast function (namely objective function), Vector $\mathbf{O}_{Lx1}$ has all values of one. And in Eq.3.6, the normalization has been added to improve the stability. g' denotes the mathematical derivative.

The value of **W** iterate until it converges. Namely, the old and new values of **W** point in the same direction, which the contrast function reaches its maxima. Thus the nongaussianity is maximized, and **S** is solved.

Indeterminacy of the Independent Components.

Due to the shortage of the knowledge of the source signal and the transmission of the mixed signal, there are two indeterminacies of the separation algorithms inherent to all the solutions, as mentioned in [118].

(1)The order of the separated signal is undetermined.

Suppose B is a permutation matrix, then it holds

$$X = AB^{-1}BS \quad \text{(Eq.3.7)}$$

where BS is the separated signal, AB^{-1} is the mixing matrix. It is obvious that BS may still equals to S only differs with the order.

(2) The amplitude and the phase of the separated signal are undetermined.

This means the separated source signal may have different amplitude with the real source signal. Suppose that b_j is a non-zero constant, it satisfies

$$X = \sum_{j=1}^{N} (a_j \frac{1}{b_j})(b_j S_j) \quad \text{(Eq.3.8)}$$

where b_jS_j is one of the separated signal. b_j is a scale for the separated signal. It shows that the separated signal may have amplitude as b_jS_j , which is different from

the amplitude of the real source S_j . Since that b_j can be negative, the phase of the separated signal also can not be determined.

3.2 Proposed Source Recovery Algorithm

The Source Recovery (SR) algorithm is proposed to reduce the simultaneous noise, and it is introduced in detail.

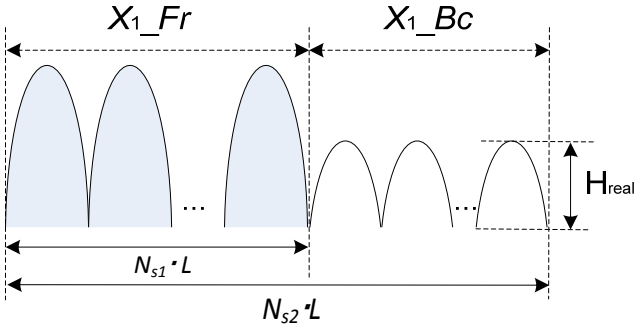
The problem of simultaneous noise fits well with the ICA model. EMA is conducted when the details of encryption module is unknown. One can only measure the leaked mixed signals. The source encryption signals are independent, since the algorithms of different ciphers differ and their power signatures vary (the switching activities are different even if the secret key is the same).

More restrictions are needed in order to meet the requirements of ICA and to solve simultaneous noise. The restrictions are: (1) The encryption for one class of cipher is interested, the signals from other modules act as noise signal. (2) Only the mixed signal can be collected (The interested signal and noise signal can not be collected separately). (3) The implementations of the ciphers are in different cycles (the interested signal and the noise signal have different number of peaks in the collected signals).

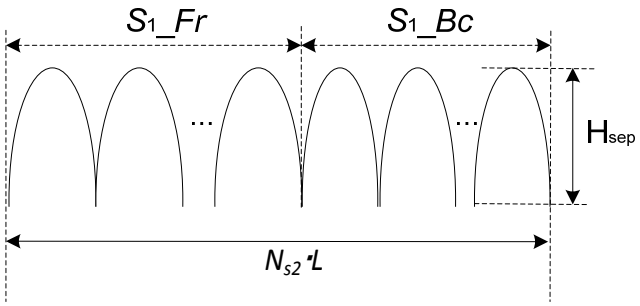
In this case, the classical ICA algorithms can be adopted to separate the noise signal from the collected mixed signal as the first step. However, because there are two indeterminacy of the separated signal as mentioned in subsection 3.1.3, the separated signal can not be simply used for EMA (As the success of EMA depends on the statistics of a large number of signal traces, which is sensitive to the variations of signal amplitudes). More processing techniques are proposed to determine the interested signal and recover the amplitudes. They are described as SR algorithm in detail to reduce the simultaneous noise for EMA.

3.2.1 Overview of the Algorithm

The signal model for the proposed algorithms is similar to the classical ICA model, as shown in Eq.3.3. It is assumed that the measured mixed signal X= (X_1, X_2), contains two different encryption sources, S_1 and S_2. One of it is correlated with the Hamming Distance (HD) of interested cipher, and the other is uncorrelated to this HD. In addition, the number of rounds for the two source ciphers are N_{s1} and N_{s2}, and $N_{s1} < N_{s2}$. The number of sampling point in each round is L, thus the sampling lengths for them are N_{s1}·L and N_{s2}·L respectively, shown in Fig.3.4.



(a) Mixed encryption signal X_1



(b) Separated encryption signal S_1

Fig.3.4 Illustration of the mixed encryption signal and separated encryption signal

In order to obtain the real source signal S_{real1} and S_{real2} from two mixed samples X_1 and X_2, the proposed SR algorithm includes the following three steps:

(1) Source separation

(2) Correlation judgment

(3) Amplitude recovery

The step (1) takes advantage of the classical FastICA to obtain the raw separated signals S_1 and S_2 from the 2 mixed samples X_1 and X_2. It is assumed that the signal is correctly separated but the order of S_1 and S_2 are not determined, and the real amplitudes of them are mismatched. Then step (2) determines the interested signal, namely to determine the order of the separated signal. For EMA, the interested signal is the one correlated to the HD of interested cipher, since the number of encryption rounds of different cipher is different because of the cipher structure. Then by judging the number of peaks of the separated signals, the correlated source can be identified. Finally step (3) computes the amplitude of the real source signal for the interested cipher. The flow chart of the algorithm is shown in Fig.3.5.

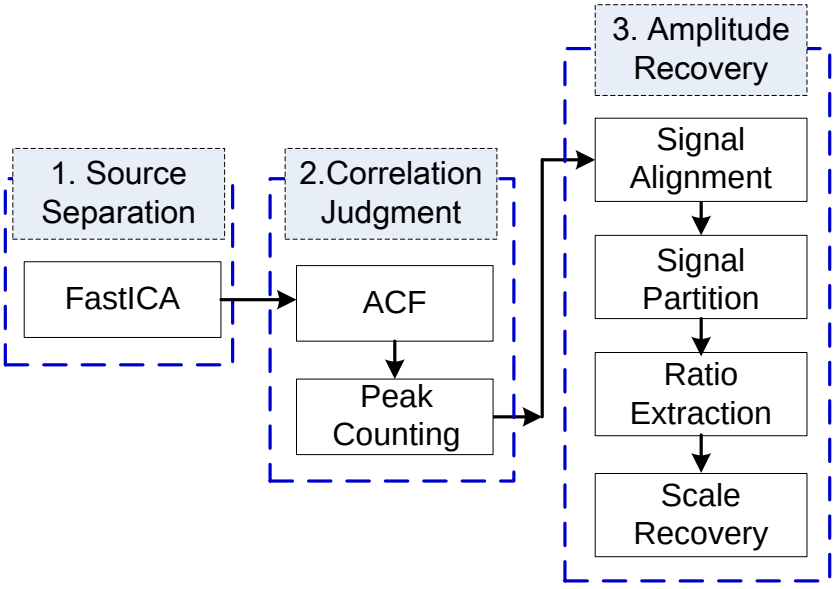


Fig.3.5 Flow chart of SR algorithm

The reasons for adopting FastICA are explained. Among the algorithms for BSS problem, there are low-computational algorithms, AMUSE and GED require that the source signals completely uncorrelated, namely, the correlated matrix for the source signal should be strictly diagonal matrix, because these algorithms obtain the separated source by decomposing the eigenvalue of the mixed matrix.

This requirement is seldom satisfied for the mixed encryption signals since they may slightly correlated due to the clock synchronization in practical cases. For the SOBI algorithm, it requires the source signal is non-whitten as it utilizes the nonzero-lag correlated function. And the gradient-based algorithm demands the source signal is non-stationary. Due to the above limitations, these algorithms are not suitable for separating the mixed encryption signals.

On the other hand, the algorithms, such as JADE[114], EASI[115], and FastICA [91] are based on high-order (larger than second-order)statistics of the mixed signals without the requirements of non-whittening or non-stationary. And another advantage of these algorithms is that they are immune to Gaussian noise in the mixed signal, as the high-order of the Gaussian noise is zero. Among these algorithms, JADE and EASI need very complex matrix or tensorial operations. [116,117] are based on stochastic gradient methods , which has slow convergence and the convergence depends on the correct choice of the learning rate parameters. On the contrary, FastICA [91] has a fast convergence, which means it finds the separated signal in several iterations. It has been one of the most popular algorithms. This is suitable for solving the mixed encryption signals which always need thousands of separations of mixed signals with different plaintexts.

3.2.2 Step 2 of Source Recovery Algorithm

The step 2 of SR algorithm is Correlation Judgment, which includes (1) ACF (Auto-Correlation function), (2) Peak Counting. ACF is to enhance the peak feature of different ciphers, and then the Peak Counting differentiates the separated signals.

(1) ACF

ACF measures the similarity of one sequence (A signal is expressed as one sequence). For the signal X, its ACF is defined as

$$ACF(X) = \frac{1}{N-m} \sum_{i=1}^{N-m-1} X_i X_{i+m}^T \quad \text{(Eq.3.9)}$$

where m = 0, 1, 2,… , (N-1). m denotes an integer offset. N is the length of signal X, i is the index for the sampling point, and T denotes the vector transpose of X.

From the properties of ACF, it is known that the ACF of a periodic signal is still periodic, and the length of the periodicity remains the same with that of the signal. Another property is that the superposition of uncorrelated noise. It states the ACF of one signal contains uncorrelated noise is composed of the individual ACF's of both the signal and noise. And this property is often used to detect signal from noise. Therefore, ACF is used to enhance the peak feature of different ciphers with noise.

(2) Peak Counting

Since the number of peaks of different ciphers varies, the Peak Counting follows the ACF of the signals can differentiate the signals, shown as follows.

Peak Counting (ACF(S_1))
For i = 1 to N
If ACF($S_1(i)$) > Th, Num ++;
End

where Th is the threshold of amplitude of the signal. This parameter depends on the SNR of the ACF resulted signal. In experience, k ·AVR(ACF(S_1)), AVR denotes the average value, and k = 0.1 for the low-noise ACF resulted signal. Num is the number of peaks detected from the signal S_1.

Then after the Peak Counting, the interested signal is identified, since the number of rounds of the interested cipher is known. Here suppose S_1 is the interested signal, which has the correlated power signature and it will be further processed.

3.2.3 Step 3 of Source Recovery Algorithm

The step 3 of SR algorithm is Amplitude Recovery. Its goal is to recover the amplitude of the real source signal. And it includes the (1) Signal Alignment, (2)

Signal Partition, (3) Ratio Extraction, and (4) Scale Recovery. They are explained in detail as follows.

(1) Signal Alignment

The cross-correlation function is proposed to align the signal. It measures the similarity between two signals. It can identify the signal offset between two similar signals. The cross-correlation function between two signals X and Y is defined as

$$Cross((X,Y)_\tau) = \sum_i X_i^* \bullet Y_{i+\tau} \quad \text{(Eq.3.10)}$$

where X^* denotes the conjugate transpose of X, which is defined by taking the vector transpose and then taking the complex conjugate of X (for an real-valued signal, its complex conjugate is itself). And i is the index for the sampling point, τ denotes an integer offset. A peak value of this function means signal Y is most similar to X at an offset of τ.

Cross(X_1, S_1) is computed. With the resulted offset, the signal S_1 and X_1 are aligned to each other.

(2)Signal Partition

Because X_1 is sampled in the strict clock cycle when the ciphers run, it has near periodicity. It can be partitioned into two segments X_1_Fr and X_1_Bc according to the length of the two ciphers, expressed by

$$X_1\_Fr = \begin{bmatrix} X_1(1) & X_1(2) & \dots & X_1(N_{s1} \bullet L) \end{bmatrix} \quad \text{(Eq.3.11)}$$

$$X_1\_Bc = \begin{bmatrix} X_1(N_{s1} \bullet L+1) & X_1(N_{s1} \bullet L+2) & \dots & X_1(N_{s2} \bullet L) \end{bmatrix} \quad \text{(Eq.3.12)}$$

Similarly, S_1 is partitioned into 2 segments S_1_Fr and S_1_Bc, given by

$$S_1\_Fr = \begin{bmatrix} S_1(1) & S_1(2) & \dots & S_1(N_{s1} \bullet L) \end{bmatrix} \quad \text{(Eq.3.13)}$$

$$S_1\_Bc = \begin{bmatrix} S_1(N_{s1} \bullet L+1) & S_1(N_{s1} \bullet L+2) & \dots & S_1(N_{s2} \bullet L) \end{bmatrix} \quad \text{(Eq.3.14)}$$

(3) Ratio Extraction

Then the quotients of the uncorrelated parts of signal X_1 and S_1 is obtained in order to compute the ratios between them, shown as follows.

$$R = X_1\_Bc \ / \ S_1\_Bc \quad \text{(Eq.3.15)}$$

The average scale is the mean value of these ratios, computed by

$$\overline{R} = \frac{1}{(N_{s2} \bullet L - N_{s1} \bullet L)} \sum_{i=1}^{N_{s2} \bullet L - N_{s1} \bullet L} R(i) \qquad \text{(Eq.3.16)}$$

(4) Scale Recovery

Finally the real source signal S_{real1} is recovered by this average scale, given by

$$S_{real1} = S_1 \bullet \overline{R} \qquad \text{(Eq.3.17)}$$

It is noted that the discussion of the algorithm is limited to the 2 classes of sources. It can be extended to more classes of sources. The interested source can be processed as that for S_1 in the above SR algorithm. However, most of the applications involve only 2 classes of sources for one LSI in practice, as the number of sources increase, the power consumption of the LSI also increases dramatically.

3.3 Experimental Results

In the following experiments, SR algorithm is applied to the mixed signal to separate the most uncorrelated source of encryption. And it is compared with bandpass filtering.

3.3.1 Two Simultaneous Encryption Sources

We set the bits in the interface circuits through computer. The AES0 and Camellia on the LSI execute simultaneously. Two mixed signals which are shown in Fig.3.6 (a) (b), with different plaintext and the same key are input to the FastICA algorithm. This leads to two separated signals: one is AES0 signal, and the other is Camellia signal. They are shown by Fig.3.6 (d) (f) respectively. Then the AES0 and Camellia executes individually. These are supposed to be the source signals, which are plotted in Fig.3.6 (c) (e) respectively.

In order to evaluate the effectiveness of the separation, we compute the correlation coefficient between the resulted signals and the source signals. The

correlation coefficients are listed in Table 3.1. The correlation of resulted AES0 and source AES0 is 0.8791, which is much greater than the correlation between resulted AES0 and source Camellia: 0.0236. This is the same case for the resulted Camellia. The strong correlations indicate that the separation is successful.

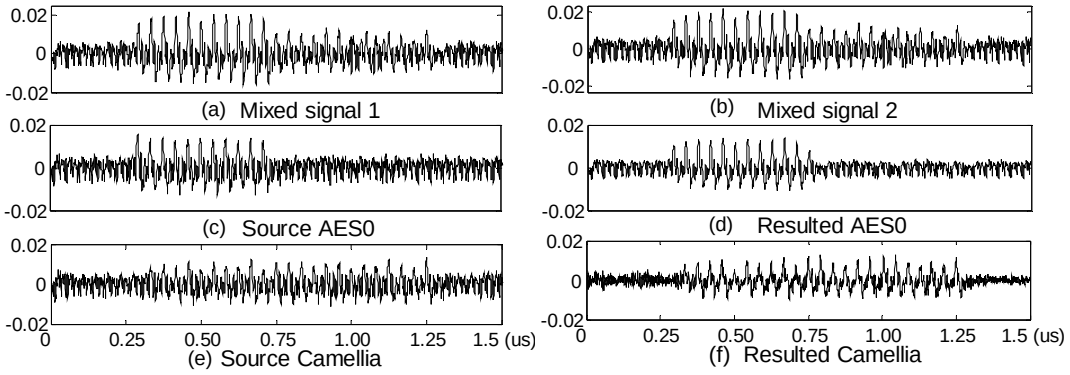


Fig.3.6 The signals of two encryption source: AES0 and Camellia

Table 3.1 Correlation coefficients between the source signal and resulted signal

	S.AES0	S.Camellia
Re.AES0	**0.8791**	0.0236
Re.Camellia	0.0207	**0.8904**

("S."denotes"source"and"Re."denotes "resulted".)

Fig.3.7 shows the 2 mixed signals processed with bandpass filtering, respectively. As suggested in [90], the pass band is [0Hz, 40MHz]. This processing technique only smoothes the signals to some extent. A close-up of the processed signal is shown in Fig.3.8. Although the bandpass filtering suppresses the frequency components upon 40MHz, the filtering does not separate the mixed signals, since AES0 and Camellia work at the same frequency.

During the execution of AES0 and Camellia, 10000 EM signals each with sampling length 2000, are recorded with oscilloscope. Then every two signals are processed with SR algorithm, it yields 10000x2 separated signals in total, which has the same number of AES signals and Camellia signals. The 10000 resulted Camellia signals are subtracted from the mixed signals. EMA is performed with

the differential signals. In order to compare it with the bandpass filtering, the 10000 mixed signals are processed with filtering.

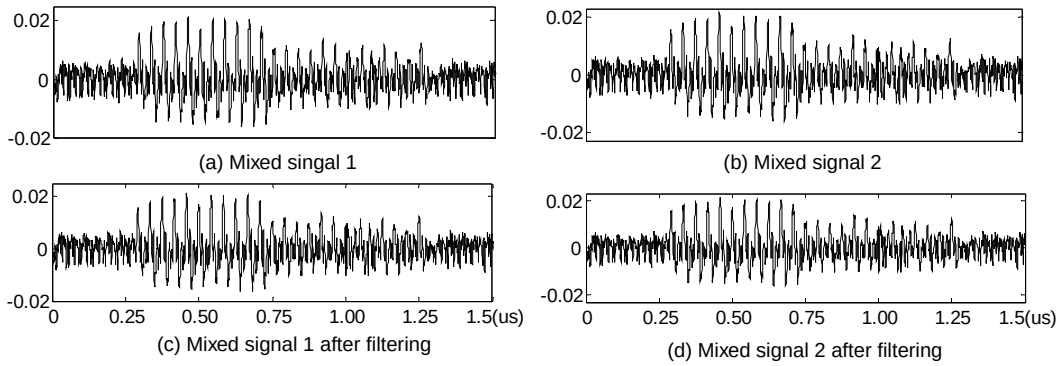


Fig.3.7 Mixed signal processed with bandpass filtering

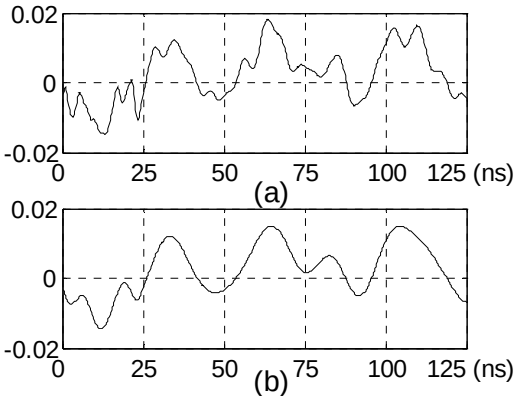


Fig.3.8 A close-up of filtering (a) EM signal without filtering (b) EM signal with filtering

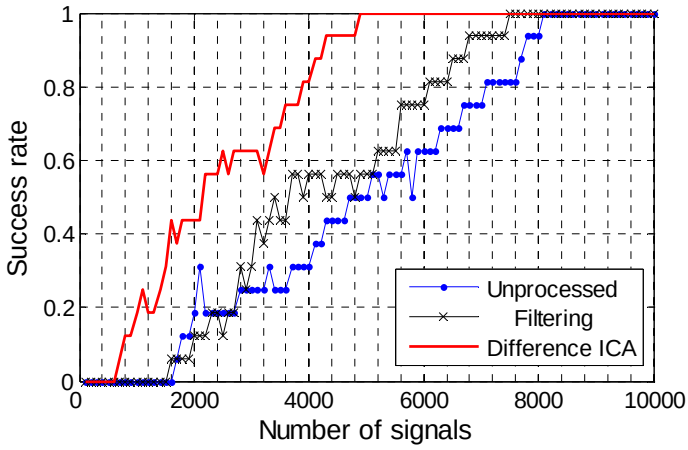


Fig.3.9 Success rates of unprocessed signal (two sources), filtered signal, and SR algorithm resulted signal

The success rates of EMA with unprocessed signals, filtering, and SR algorithm are shown in Fig.3.9, respectively. All the key bytes are revealed within 4926 signals by SR algorithm, while 7482 signals are required for filtering. They are faster than the unprocessed signals, which is with 8133.

3.3.2 Three Simultaneous Encryption Sources

The situation becomes complex when 3 encryption sources are mixed. The encryption signals are recorded when AES1, AES2 and Camellia run simultaneously. Similar to the previous process, we use 3 mixed signals and attempt to obtain the 3 separated signals. However, the resulted signals are not clearly separated. Only one of the resulted signals has a greater correlation with the source Camellia. This indicates the Camellia has been separated successfully.

The explanations for these results are: because any one of the AES executions (AESi, i=0-5) on LSI has a linear relation with Hamming Distance, the relation between different AES is not independent. The independence assumption of ICA is not satisfied. Thus the separation of different AES fails. The resulted mixed signal, namely the mixture of AES1 and AES2, is shown in Fig.3.10 (c). The differential signal is shown in Fig.3.10 (b).Then EMA is conducted with the differential signal.

The success rates are compared with the case of unprocessed signals, and filtered signals, shown in Fig.3.11, respectively. More than 10000 signals are required to reveal all the key bytes for the unprocessed signals. The filtering costs 8793 signals, while only 5371 signals are needed for SR algorithm. The success rate is greatly enhanced. It also suggests that the mixed execution of AES0 and AES1 do not have much influence for the result of EMA.

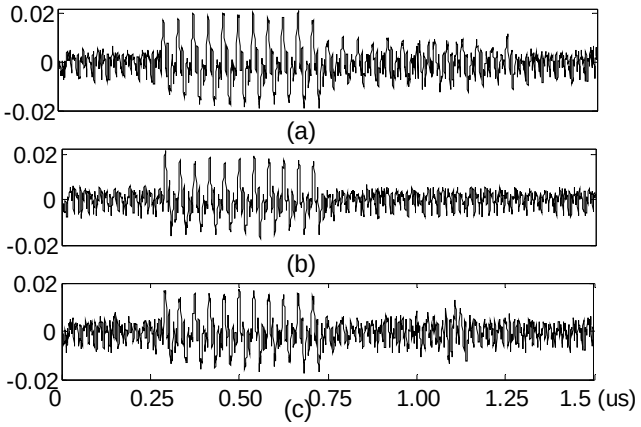


Fig.3.10 (a) The mixed signal of three source; (b) the differential signal; (c) the resulted mixed AES1 and AES2

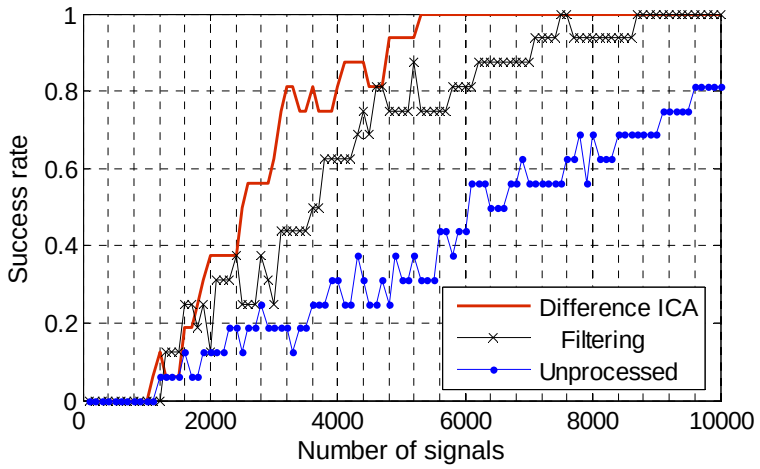


Fig.3.11 Success rates of unprocessed signal (three sources), filtered signal, and SR algorithm resulted signal

Table 3.2 The number of needed signals and correlations for each mixed encryption

Mixed type	Unprocessed		SR algorithm			Filtering		
	No.	Corr.	No.	Corr.	Rate	No.	Corr.	Rate
AES1,2, C	Fails	0.0422	5,371	0.0996	**46.3%**	8793	0.0514	12.1%
AES1,3, C	7,012	0.0627	4,126	0.1098	**41.1%**	6025	0.0682	14.1%
AES1,4, C	6,411	0.0703	3,679	0.1327	**42.6%**	5517	0.0835	13.9%
AES2,3, C	Fails	0.0419	5,301	0.0921	**47.0%**	9102	0.0501	8.9 %
AES2,4, C	9,835	0.0580	5,527	0.0908	**43.8%**	8359	0.0589	15.0%
AES3,4, C	7,164	0.0695	4,175	0.1162	**41.7%**	6120	0.0674	14.6%
AES1-4, C	8,291	0.0613	4,327	0.1204	**47.8%**	7038	0.0636	15.1%

*No.: denotes the number of needed signals.
*Corr.: the maximal correlation coefficients for key revealing.
* AES1,2, C: denotes the mixed type of “AES1, AES2 and Camellia”.
*Fails: the keys can not be revealed within 10000 signals.
*Reduction rate: the number of needed signals compared with unprocessed signal, and “Fails” is computed as 10000 for the rate.

This inference has been further confirmed in the mixed executions of AES1, AES3 and Camellia, AES1, AES4 and Camellia, AES2, AES3 and Camellia etc. The number of signals to reveal all the key bytes and the maximal correlation coefficient are listed in Table 3.2. After the application of SR algorithm, The number of signals has been reduced by 41.1% at least, while the Filtering-based EMA is only 15.1% at most. This result indicates the effectiveness of SR algorithm.

3.3.3 More Than Three Simultaneous Encryption Sources

From the hint of Experiment2, we only need to separate Camellia from the mixed signals of multiple modules of AES executions and Camellia. Five signals, namely AES0-AES4 and Camellia are processed by SR algorithm. We perform EMA with the resulted signal. The number of signals used to reveal all the key bytes has been reduced 47.8%, which is much better than the filtering-based EMA listed in the last line of Table 3.2.

All the above three groups of experiments indicate the successful application of the proposed SR algorithm to EMA.

3.4 Summary

The main contribution of this chapter is that we propose SR algorithm and successfully apply it to reduce the simultaneous noise for EMA. This is confirmed by the experiments of EMA against AES and Camellia implementation on ASIC. Additionally, the proposed algorithm is compared with bandpass filtering. It indicates that the proposed algorithm can reduce the number of EM traces as much as 47.8%, which is much better than bandpass filtering. Several conclusions are elicited. Bandpass filtering is a general processing technique, which can attenuate the inference from multiple frequency components, but is not suitable for simultaneous noise. By contrast, SR algorithm is particularly effective to

separate uncorrelated signals, which is fit for the mixed encryption implementations. With SR algorithm, the countermeasure of simultaneous noise is greatly weakened. The mixed execution of different encryption can be bypassed with signal processing techniques. These results may also provide enlightment for the design of countermeasures.

In the future, more advanced signal processing techniques will be investigated and studied. They will be applied to the evaluation of other countermeasures in order to improve the security of cryptographic devices.

4 The Switching Glitch Leakage Model for EMA and PA

A new switching glitch model is proposed for PA and EMA. Section 4.1 describes the background and related works. Section 4.2 presents the proposed leakage model in detail. Section 4.3 shows the application of the proposed model in PA and EMA. Section 4.4 summarizes this chapter.

By improving the leakage models, it is possible to enhance both PA and EMA. As mentioned by many research works [20], EMA uses the same power models as PA. The measurement from EM field generated by the device is indirect method for power consumption, which is presented in the IEC standard 61967[93].

4.1 Background and Related Works

The related works on leakage model are briefly introduced. The power consumption of CMOS circuits which is the basis of modeling power consumption of cryptographic devices is analyzed.

4.1.1 Leakage Models for EMA and PA

Leakage models are used to establish a hypothesized power consumption of attacked cryptographic device for PA.

Several works have been done on it. Aimed at improving Hamming Weight (HW) leakage model [3] in PA, Brier et al. [10] proposed Hamming Distance (HD)-based correlation power analysis (CPA) attack which utilized the correlation factor between HD and measured power to reveal keys in 2004.

For a n-bit processor, Hamming Weight (HW) expresses the number of bits that are set in the processed data value $\boldsymbol{X}$, given by Eq.4.1, where x_i is the value of (i+1)th bit of the processing data value $\boldsymbol{X}$.

$$\boldsymbol{HW}(\boldsymbol{X}) = \sum_{i=0}^{n-1} \boldsymbol{x}_i \quad , \quad x_i \in \{0,1\} \qquad \text{(Eq.4.1)}$$

In HW model, it is assumed that the power consumption $\boldsymbol{Y}$ is proportional to Hamming Weight, given by Eq.4.2, where $\boldsymbol{a}$ is a scalar gain, $\boldsymbol{b}$ denotes the offsets, time dependent components and noise.

$$\boldsymbol{Y} = \boldsymbol{aHW}(\boldsymbol{X}) + \boldsymbol{b} \qquad \text{(Eq.4.2)}$$

Hamming Distance is computed according to Eq.4.3, where $\boldsymbol{X}$ is an intermediate value during a target implementation, $\boldsymbol{R}$ is a reference state of the running algorithm.

$$\boldsymbol{HD}_{\boldsymbol{X},\boldsymbol{R}} = \boldsymbol{HW}(\boldsymbol{R} \oplus \boldsymbol{X}) \qquad \text{(Eq.4.3)}$$

It assumes that power consumption $\boldsymbol{Y}$ is proportional to the transitions of the intermediate values not the value being processed, given by Eq.4.4.

$$\boldsymbol{Y} = \boldsymbol{aHD}_{\boldsymbol{X},\boldsymbol{R}} + \boldsymbol{b} \qquad \text{(Eq.4.4)}$$

Since there is a linear relationship between the real leakage, namely the measured power consumption $\boldsymbol{P}$ of the cryptographic device and the assumed power model $\boldsymbol{Y}$[10].The more accurate the hypothesized power consumption $\boldsymbol{Y}$, the more legible the relation with the measured power consumption appears.

Thereby several other models contribute to more accurate description of the power model.

The Switching Distance model was suggested by Peeters et al. [58] and they mounted the simulated attacks on an 8-bit PIC16F877 microprocessor against S-box output in 2007. The transition activities of the CMOS circuits, namely from 0 to 1 and from 1 to 0, were differentiated by defining a normalized parameter φ. The advantage of this model is that it is more precise than HD model. The performance of PA and EMA may be both improved. However, the problem is that the method of determining the value of φ is left open. Probably one could not know the exact value of φ without pre-measurement, and this value might varies with the attacked devices.

A so-called Zero-Value model [94] was proposed, which is based on the observation that the input value zero of S-box for a cipher implementation consumes significantly less power than in case of all other input values. Thus the power consumption for the data value zero is set lower than the power

consumption for all other values. However, this model is only effective for the S-box implemented with the composite field arithmetic, because essentially all multiplications in S-box are multiplications by zero in such case.

Therefore it is a trade-off. A more accurate leakage model is possible at the cost of the knowledge of many details of the attacked devices. For instance, if the transistor netlist of the attacked device is known, then the difference equations might be available to characterize the power consumption. In this work, we assume that the attackers do not have knowledge about the details of the cryptographic devices.

4.1.2 Power Consumption of CMOS Circuit

To simplify the problem, the following conditions are satisfied: the cryptographic device is built of CMOS logic gates and edge-triggered flip-flops, and is synchronous. The power supply and ground voltage are fixed on the chip. The simplified structure of a typical CMOS circuit is shown in Fig.4.1.

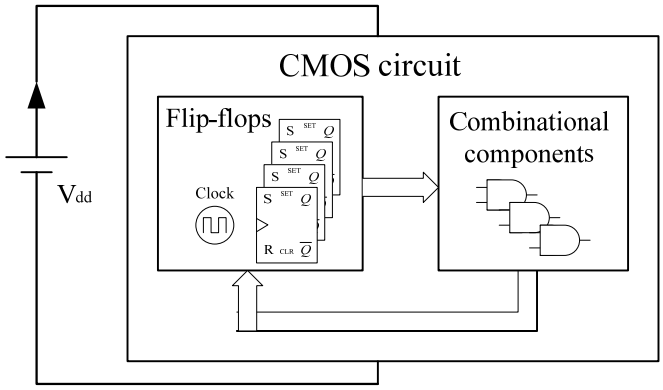


Fig.4.1 Simplified structure of CMOS circuit

The power consumption of CMOS circuits includes two parts: dynamic power and static power [95]. Static power is consumed when there are no transition activities. Dynamic power consumption occurs if the sequential components or/and combinational components transits. This is the dominant part of the dynamic power. Besides, there is a short-circuit power consumption, which is

caused by signal transition at a gate output when the pull-up and pull-down transistors conducts simultaneously for a short period of time. A summarization is shown in Fig.4.2.

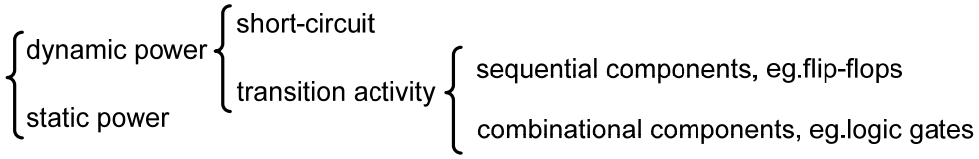


Fig.4.2 A summarization of power consumption of CMOS circuits

Table 4.1 Power consumptions of CMOS circuits and their computations

Dynamic	Switching activity	Flip-flop	$0.5V_{dd}^2 fC_x P(x)$	
		Combinational components	Normal Transition	$0.5V_{dd}^2 f\sum_{i=1}^{n} C_{xi} P(x_i)$
			glitch	$V_{dd}^2 fC_{av} N_{glch}$
	Short-circuit	$\frac{1}{t}\int_0^t P_{sc}(t)\,dt$		
Static	$V_{dd}\, I_{leak}$			

Glitch or hazard is the temporary states of the output of combinational components because of the different arrival times of the input signals. Glitch power is a significant portion of the dynamic power consumption. On account of the indeterminate property, a probabilistic modeling is adopted to characterize the glitch power P_{glch}. The power dissipation due to input glitches [102,103] is shown by

$$P_{glch} = V_{dd}^2 fC_{av} N_{glch} \quad \text{(Eq.4.5)}$$

where C_{av} is the average capacitance at logic gates. N_{glch} expresses the number of generated glitches within a circuit.

The total power consumption of a CMOS circuit is summarized in Table 4.1, where f denotes the clock frequency. C_x is the total capacitance at one flip-flop output. P(x) is the transition probabilities. P_{norm} expresses the power consumption of logic gates when they perform normal signal transitions to finish required logic

functions. C_{xi} denotes the total capacitance at one gate. $P(x_i)$ is the transition probabilities(Transition probability for a logic gate or flip-flop X: The average fraction of clock cycles in which the value of X at the end of the cycle is different from its initial value.). n is the total number of logic gates in the circuit. V_{dd} is the voltage of the power supply. I_{leak} is the leakage current. $P_{sc}(t)$ is the instantaneous short-circuit power consumed by one flip-flop or logic gate.

4.2 Proposed Leakage Model

Our work follows the Hamming Distance model whereas considering a more accurate description to the data dependent power dissipation.

4.2.1 Switching Factor and Glitch Factor

In power analysis, the data dependent and operation dependent power consumption are of the main interest of research. However, in general cases, operation dependent power consumption are more relied on specific cryptographic devices and it is totally black box for attackers, while the short-circuit and static power are negligible [23]. When the logic states change on the arrival of clock cycle, the intermediate value or cipher is stored in flip-flops, and the combinational components perform switching activities. During this period, glitches take place at some logic gates. The dissipated power which is related to data P_{data} is the sum of the power consumed by flip-flops and combinational components, given by

$$P_{data} = P_{flip} + P_{norm} + P_{glch} \qquad \text{(Eq.4.6)}$$

For an encryption circuit, it is costly to compute the exact value of each part of P_{data} by the equation in 1. But quantitatively, from the equations in Table 4.1, we can conclude that the power consumption of flip-flops and combinational components is in a close magnitude. In other words, a number of flip-flops could

consume similar power with the normal combinational components except the little difference of load capacitances.

So we simplify this proportional relation with two factors: switching factor α and glitch factor β to summarize the corresponding power, shown in Table 4.2.

Table 4.2 Switching factor and glitch factor

Switching power	1 N_{01}
	α N_{10}
Glitch power	β

N_{01} and N_{10} are the number of switches performed by the circuit from 0 to 1 and 1 to 0 respectively. α is switching factor, it characterizes the difference of such transitions. The existence of switching factor is based on the fact that these two kinds of transitions consume different power in normal switching activities. While glitch factor β describes the glitch power which is circuit and algorithm specific. For different chips and encryption data paths, the glitch factors vary a lot.

4.2.2 The New Leakage Model

Our work follows the Hamming Distance model whereas considering a more accurate description to the data dependent power dissipation. Since there is a linear relationship between the power consumption Y of the cryptographic device and the assumed power consumption E, given by Eq.4.7-Eq.4.9.

$$Y \sim E \quad \text{(Eq.4.7)}$$

$$E = E_{sf} + \beta \quad \text{(Eq.4.8)}$$

$$E_{sf} = N_{01} + \alpha N_{10} \quad \text{(Eq.4.9)}$$

where E is the estimated energy of encryption, E_{sf} denotes the power of normal switching activities of combinational gates and flip-flops. The more accurate the assumed power, the more legible this relation appears.

Then we show how to estimate the switching factor and glitch factor.

In order to find the optimal switching factor, the "accumulation factor" S_{acu} is defined and given by Eq. 4.15 to quantitatively explore which switching factor is better for PA or EMA,

$$S_{acu}(\alpha) = \sum_{i=1}^{L}(X_{HDi} - X_{SFi}) \qquad \text{(Eq.4.10)}$$

where α is switching factor, for the i_{th} key byte, X_{HDi} denotes the number of power traces required by HD, and X_{SFi} denotes the number of power traces required by switching factor α.

This definition means that for switching factor α, the accumulative numbers of power traces of all the L bytes (eg. For AES encryption, L=16) is the sum of the differences of each byte. It expresses the improved accumulative number of power traces compared with HD. When S_{acu} is positive, that means the number of power traces is decreased. While the negative S_{acu} stands for an increase of power traces. For HD itself, this value is 0. By selecting the largest S_{acu}, the optimal switching factor is determined.

In order to estimate glitch factor β , the division is computed as

$$E_{sf} / \beta = P_{norm} / P_{glch} = 0.5V_{dd}^2 f\sum_{i=1}^{n} C_{xi}P(x_i) \Big/ V_{dd}^2 fC_{av}N_{glch} \qquad \text{(Eq.4.11)}$$

Suppose that the total capacitance at each gate C_{xi} equals to C_{av}, and then the expression is simplified as

$$E_{sf} / \beta = 0.5\sum_{i=1}^{n} P(x_i) \Big/ N_{glch} \qquad \text{(Eq.4.12)}$$

For an encryption circuit, the value of $P(x_i)$ depends on the input data. Furthermore, if we know the relation between the numbers of generated glitches N_{glch} and the logic gates n, then β can be expressed by some expression of E_{sf}. In fact, because of the complexity of the design technologies and detailed processing techniques of CMOS circuits, it seems that this relation is unpredictable without CAD simulation tool. We will make a further reckon and verify it through experiments.

From Eq.4.12, β can be expressed by

$$\beta = E_{sf} N_{glch} / 0.5\sum_{i=1}^{n} P(x_i) \qquad \text{(Eq.4.13)}$$

It is always assumed that the plaintext is randomized when power analysis is performed, the transition probability of the logic gates can be estimated with value 0.5. Then β is derived as

$$\beta = E_{sf} N_{glch} / 0.25 N_{gats} \quad \text{(Eq.4.14)}$$

where N_{gats} denotes the number of logic gates in encryption circuit. For one byte input data, with switching factor α = 1.0, which is the HD model, E_{sf} is in 10 magnitude. Suppose that the average generation rate of glitches at the logic gates is 0.1, and then β is calculated from Eq.4.14 at 1.0 magnitude.

In practice, it is possible to estimate the optimal value for α and β by above method to make them device specific, and thus more accurate. This is further explained by the experiments.

4.3 Side Channel Analysis with SG Model

In this section, the proposed leakage model is verified by both PA and EMA on AES implementations.

4.3.1 PA with SG Model

Firstly the optimal value of switching factor and glitch factor are estimated by power analysis on one AISC, named as IC_a. Then these factors are applied to IC_b, which is produced with the same technology. The cryptographic core uses 0.13μm TSMC standard library of CMOS process technology. Throughout the experiments, the initial encryption 16-byte keys are set as hexadecimal numbers: 12 34 56 78 90 AB CD EF 12 34 56 78 90 AB CD EF. The final round encryption keys are: C3 BE 32 F4 60 A9 B3 4E F7 43 61 57 F2 B9 19 D8.

Table 4.3 The number of power traces used to recover all the 16-byte keys

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**Sacu**
SF3.5	42	62	34	16	29	74	30	37	46	36	39	48	100	22	63	63	-132
SF3.0	42	61	34	16	29	74	30	35	41	35	28	49	90	21	51	50	-75
SF2.5	42	57	16	16	29	74	31	32	38	34	29	48	90	21	51	54	-53
SF2.0	42	54	34	11	31	55	32	35	37	33	22	48	41	21	50	50	13
SF1.8	42	52	17	11	31	55	32	35	38	34	22	48	41	22	50	50	29
SF1.7	43	51	15	11	29	58	36	35	33	34	21	51	40	21	50	48	35
SF1.6	35	51	16	15	31	56	40	24	32	32	28	48	41	21	50	47	42
SF1.5	35	51	16	15	31	56	40	24	28	32	28	47	38	20	50	47	51
SF1.3	35	54	16	15	28	75	55	31	31	29	38	58	38	15	49	28	14
SF1.1	35	65	16	15	46	76	55	21	26	35	23	64	37	15	49	28	3
SF0.9	38	70	15	20	46	93	67	21	38	39	31	66	40	25	50	28	-72
SF0.8	35	70	15	20	46	93	67	21	28	41	34	66	43	23	50	26	-66
HD	35	65	16	15	46	76	55	24	26	35	23	64	37	15	49	28	0

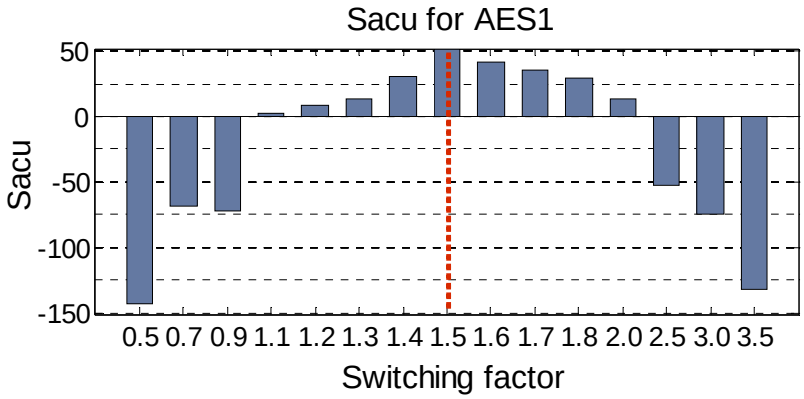


Fig.4.3 Sacu of switching factors for AES1

Table 4.3 shows S_{acu} and the number of power traces used to recover all the 16-byte (128-bit) keys of AES when the value of α is set to range from 0.8 to 3.5. Some identical lines are omitted. For example, when SF is 1.4, the number of power traces is the same as SF1.3. Note that, all the numbers are in unit 100. For instance,

when switching factor is 3.5, the first bytes of AES keys can be recovered at 4200 power traces.

Fig.4.3 gives a more clear vision of S_{acu} at different switching factors. We can see that the S_{acu} is the largest when switching factor is 1.5. Therefore, for AES encryption on IC_a, when switching factor is set to 1.5, the 16-byte keys can be recovered with least power traces. Similarly, the same experiments are repeated when AES2, AES3 and AES4 run. The S_{acu} for AES2, AES3 and AES4 are shown in Fig.4.4. The switching factors for them are 1.6, 1.2, and 1.4, respectively.

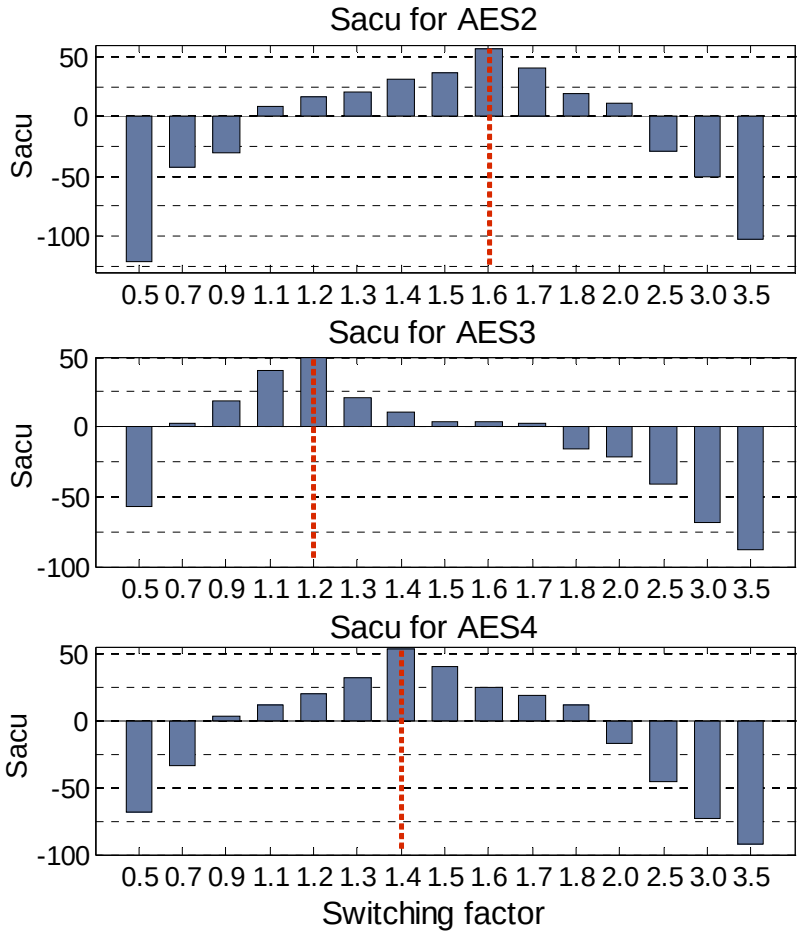


Fig.4.4 Sacu of switching factors for power traces

Fig.4.5 shows the glitch factor for power traces from AES1, AES2, AES3, and AES4. In order to determine glitch factor for each of AES, PA is conducted with different values, and the number of traces is compared with HD model. The glitch factor which has the largest reduction rate is the optimal one. For AES1, we find that

1.0E-1 is the optimal value rather than the theoretical value 1.0. That means when β = 0.1 the number of power traces is least. Therefore, the glitch factors for AES2, AES3, and AES4 are 0.09, 0.09, and 0.20 respectively.

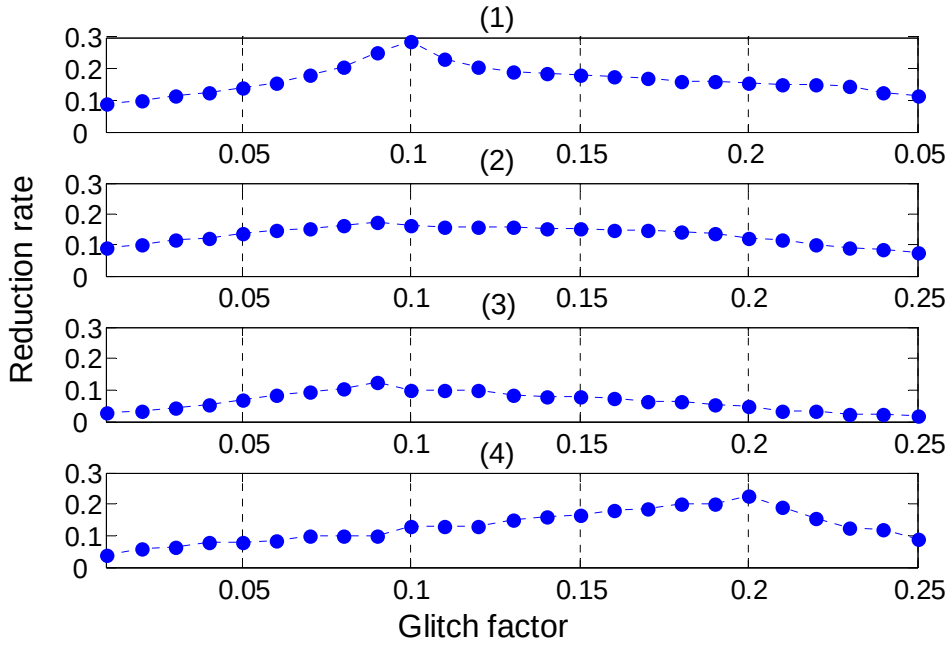


Fig.4.5 Glitch factor for power traces (1)AES1 (2)AES2 (3)AES3 (4)AES4

Table 4.4 Switching factor and glitch factor of PA with SG model on IC_a

Name	Switching factor	Glitch factor	SG model	Reduction rate (%)*
AES1	1.5	0.10	5400	28.9
AES2	1.6	0.09	2900	17.1
AES3	1.2	0.09	2800	12.5
AES4	1.4	0.20	3800	22.4

(* Reduction rate : compared with HD model.)

Table 4.4 lists the switching factor, glitch factor and the least number of power traces used to recover 16-byte keys by SG model as well as the reduction rate for power traces for each of the AES respectively. For example, the first line of this table means: For AES1 implementation on IC_a, the switching factor is 1.5, glitch factor is 0.10. The SG-based CPA recovers 16-byte keys with 5400 power traces. And the power traces have been reduced by 28.9% compared with that with HD model.

Table 4.5 PA with HD model[10], SD model[58] and SG model on IC_b

Name	HD model[10]	SD model[58]	Glitch factor	SG model	Reduction rate (%)*
AES1	6100	5800	0.05	5000	18.0
			0.10	**4600**	**24.5**
			1.00	5100	16.3
AES2	3000	2700	0.05	2700	10.0
			0.09	**2600**	**14.4**
			1.00	2900	3.3
AES3	2300	2300	0.05	2200	4.3
			0.09	**2000**	**13.0**
			1.00	2200	4.3
AES4	3400	3300	0.05	2900	14.7
			0.20	**2700**	**20.6**
			1.00	3000	11.8

(* Reduction rate : compared with HD model.)

The switching factors and glitch factors trained from IC_a are further verified on IC_b. The results are also compared with HD model [10], SD model [58]-based PA, and listed in Table 4.5. According to [58], the value of φ is 0.17. Namely the transition from 0 to 1 is denoted by 1, and from 1 to 0 is denoted by 0.63 (1-0.17). Compared with HD model, the number of power traces is slightly reduced by SD model for AES1, AES2 and AES4. It keeps the same for AES3. The reason for the very limited reduction is likely that the value of φ is optimal for microprocessor PIC16F877 but not for SASEBO. For the SG model, different glitch factors are used for each AES implementation, the reduction rates for different glitch factors vary. The largest reduction rate 24.5% is achieved at glitch factor 1.0 for AES1. This shows that the SG model with glitch factors trained from IC_a has highest reduction rate. Thus the result of SG model is best among the 3 models.

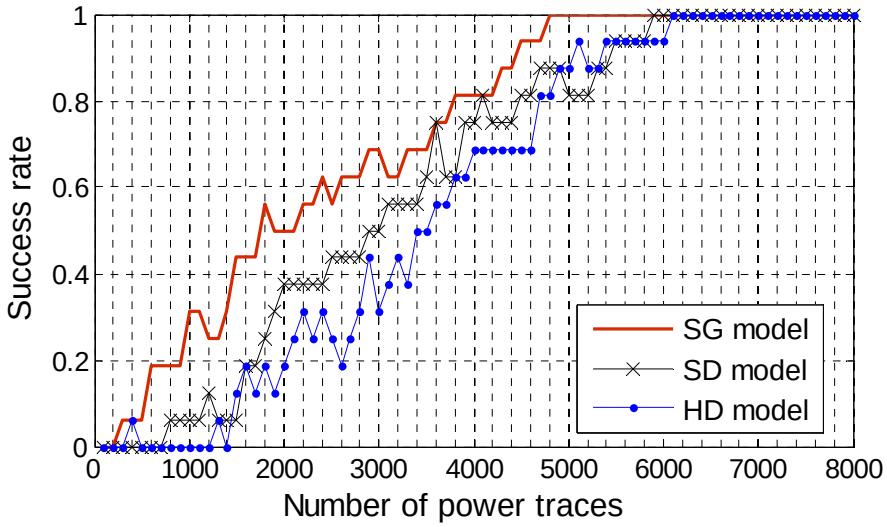


Fig.4.6 Success rates of PA against AES1on IC_b

The success rates are compared with HD model-based PA, and SD model-based PA against AES1 are illustrated in Fig.4.6. Success rates express the number of power traces when all the keys can be recovered. With HD model, 100% appears at 6100 power traces, while for SD model, it cost 5800 power traces. But with SG model, 100% appears at 4600 power traces. The power traces of recovering keys have been reduced by 24.5%.

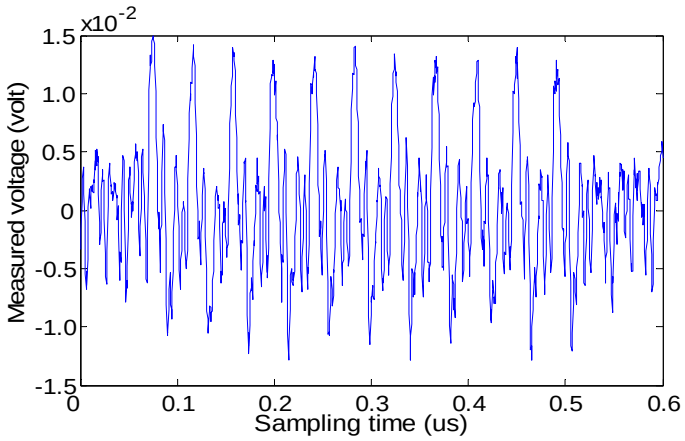


Fig.4.7 An EM trace

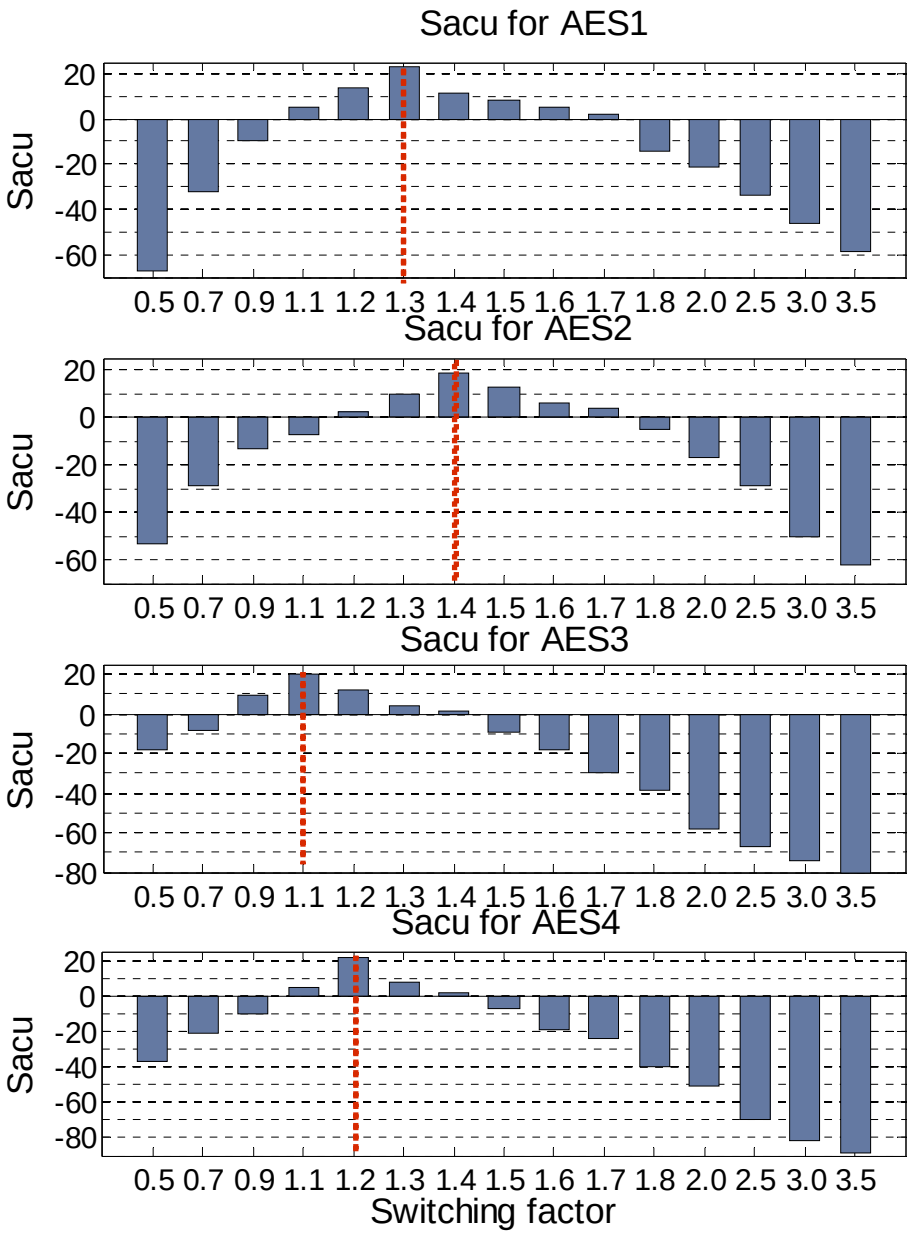


Fig.4.8 Sacu of switching factors for EM traces

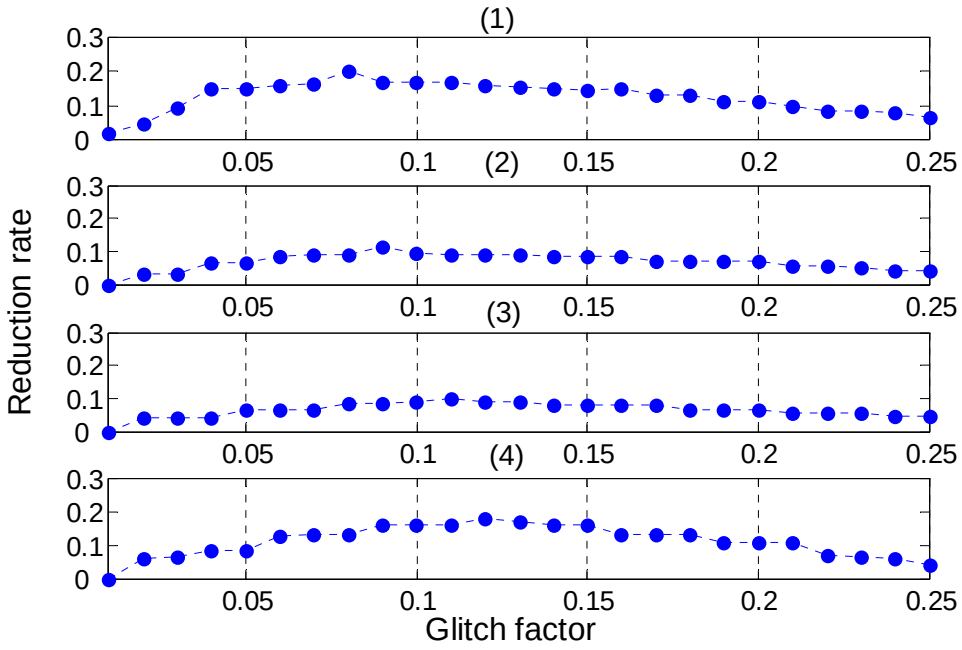


Fig.4.9 Glitch factor for EM traces (1)AES1 (2)AES2 (3)AES3 (4)AES4

4.3.2 EMA with SG Model

The proposed SG model is verified by EMA experiments in this subsection. As it is mentioned by many research works, an effective way to find the best location performing EMA attack with less EM signals is to move the EM probe manually around radiation area. The location where EM signal has higher amplitude might succeed faster. It suggests the upper right corner of LSI has a larger EM leakage. An EM signal during AES1 runs is shown in Fig.4.7.

The switching factors using EM traces for each of the AES on IC_a are shown in Fig.4.8. The value of switching factor is determined by the largest S_{acu}. The glitch factors are calculated by EMA and shown in Fig.4.9.The switching factors and glitch factors for each AES implementation are listed in Table 4.6.

Similarly, we perform EMA on IC_b using different glitch factors for SG model and compare the results of PA with HD model[10], SD model[58]. The results are shown in Table 4.7. For the SD model, φ is 2.0. It shows that the number of EM traces is not reduced for AES2 and AES3, probably because this model deviates from

the real consumption for SASEBO. On the contrary, for SG model, the numbers of EM traces are reduced at different rates: the lowest is 8.7% for AES3, and the highest is 17.1% for AES1. Furthermore, it indicates that the least number of EM traces is achieved at the estimated glitch factor trained from IC_a, which validated that the proposed model is more accurate than HD model [10] and SD model [58].

Table 4.6 Switching factor and glitch factor of EMA with SG model on IC_a

Name	Switching factor	Glitch factor	SG model	Reduction rate (%)*
AES1	1.3	0.08	4500	20.2
AES2	1.4	0.09	2700	11.5
AES3	1.1	0.11	2500	10.0
AES4	1.2	0.12	3700	17.8

(* Reduction rate : compared with HD model)

Table 4.7 EMA with HD model[10], SD model[58] and SG model on IC_b

Name	HD model[10]	SD model[58]	Glitch factor	SG model	Reduction rate (%)*
AES1	5000	4900	0.01	4900	2.0
			0.08	**4100**	**17.1**
			1.00	4800	4.0
AES2	2700	2800	0.01	2700	0.0
			0.09	**2400**	**10.2**
			1.00	2600	3.7
AES3	2400	2600	0.01	2400	0.0
			0.11	**2200**	**8.7**
			1.00	2300	4.2
AES4	3600	3500	0.01	3600	0.0
			0.12	**3200**	**12.0**
			1.00	3500	2.8

(* Reduction rate : compared with HD model)

4.4 Summary

In this chapter, a new power consumption model, namely Switching Glitch model, is proposed, which characterizes the power consumption of cryptographic devices more accurately. The distinguishing of two different dynamic switching

activities and the power consumption of glitches are both included. Compared with Hamming Distance model-based PA, it can reduce the number of power traces by as much as 24.5%, which also outperforms the Switching Distance model. The performance is also enhanced for EMA; the number of EM traces is reduced by as much as 17.1%. In addition, we also show how to estimate the switching factor and glitch factor, and they are experimentally represented from AES implementation on ASIC after theoretical derivations. Therefore, this model is appropriate for the evaluation of information leakage of implementations against both EMA and PA.

5 A Novel Leakage Localization Method for EMA

A novel leakage localization method is proposed for the performance enhancement of EMA. Section 5.1 introduces the background and related works. Section 5.2 presents the proposed leakage localization method in detail. Section 5.3 shows the application of the proposed method in EMA. Section 5.4 summarizes this chapter.

5.1 Background and Related Works

The background and related works for the leakage localization are introduced in this section.

5.1.1 DoM Test

Suppose that at one scanning point, a leakage model is used. We adopt the widely admitted leakage model. It assumes that EM signal $S(t)$ depends on a selection function H, which is an intermediate value of encryption, and related to plaintext and key[3], given by Eq.5.1, where t is sampling time, a represents a scalar gain, b denotes the offset, and time dependent components.

$$S(t) = aH + b \qquad \text{(Eq.5.1)}$$

Then a distinguisher is applied to test the dependence between $S(t)$ and H. Our leakage indicator is from the distinguisher Difference of Means (DoM), which is briefly reviewed here. To determine whether one candidate key K_c is correct or not, DoM uses N random plaintexts C_i (i =1,2,…, N) which yield N sampling signals, $S(t)$ = $S_i(t)$. The selection function $H = H(C_i, \beta, K_c)$ partitions $S_i(t)$ into two sets: $S^1=\{S_i(t) \mid H(C_i, \beta, K_c)=1 \}$ and $S^0 =\{S_i(t) \mid H(C_i, \beta, K_c)=0 \}$ under an examined bit β. For example, H is the Hamming weight of a single-bit output of SubBytes computation for AES, and $H \in \{0,1\}$. β denotes one bit of s-box. Then DoM computes a differential

trace $D_\beta(t)$, which is the difference between the averaged S^1 and S^0, given by

$$D_\beta(t) = \frac{1}{\left|S^1\right|}\sum_{S_i(t)\in S^1} S_i(t) - \frac{1}{\left|S^0\right|}\sum_{S_i(t)\in S^0} S_i(t) \quad \text{(Eq.5.2)}$$

where $|S^1| + |S^0| = N$, and it is simplified as

$$D_\beta(t) = \mathrm{E}[S(t) \mid H = 1] - \mathrm{E}[S(t) \mid H = 0] \quad \text{(Eq.5.3)}$$

$D_\beta(t)$ tends to 0 for wrong key guess because the partitioning is statistically random. $D_\beta(t) \neq 0$ for correct key and this results in a peak. The correct key is identified as the one that yields the highest peak in differential trace at some instant $t = \tau$.

5.1.2 Leakage Localization Methods

Then for conducting EMA attacks, majority of the published works use probes of small size, in the millimeter range or even smaller. The benefit of this probe is that it distinguishes EM emissions from close locations, thus the noise caused by modules not related to cryptographic computation is attenuated. An example of a handmade probe described in literature was 3 mm long [20]. The commercially available tiny magnetic-field probes, which are designed for electromagnetic compatibility (EMC) analysis, such as the one mentioned in [96], were also used for EMA.

In this case, a challenging issue is where the possible locations are before conducting EMA attacks. In general, an attacker lacks the knowledge of the exact locations from which EM signals are emitted by a cryptographic module or communication interface. He may open the package of cryptographic LSI to recognize its different modules with a microscope. Nevertheless this is a semi-invasive approach, which is destructive [97].

Another way is to put the probe blindly, for example, far away from the cryptographic module, which leads to a very slow key detection or even failure. The drawback of this approach which is named “blind placement”, is that the leakage regions are not localized accurately.

Other approaches have been proposed. Quisquater and Samyde[98] exploited EM Cartography, which is an imaging technique, to observe the EM emissions of a smart card. Sauvage et al. [10] applied this technique to reveal the active regions of DES encryption modules on FPGA. In their work, 50 points x 50 points on a region of 2.08 cm x 2 cm over the FPGA were scanned, and the maximum peak-to-peak amplitudes of EM signals in the time domain were extracted to acquire an EM map. Then the most radiating point was identified based on the EM map and used to perform EMA. This approach is named "peak-to-peak amplitude" in this paper. It is feasible and more accurate than blind placement. In fact, the EM image was achieved from a near-field scan using an EM probe of high spatial resolution over the surface of FPGA in their work.

However, it is noted that the maximum peak-to-peak amplitude of EM signals after a subtraction computation between the active and idle phase of the DES module is utilized to draw the EM map, which is not an optimal indicator for revealing the locations of highest leakage and probably causes misjudgment, because the maximum peak-to-peak amplitude only represents the region where EM emissions are highest, but not necessarily the data-dependent EM signals, which are crucial for the success of EMA. Additionally, though the influence of surrounding noise might be reduced by the subtraction computation, numerous other data-independent EM signals, such as signals from communication interfaces, still exist and may prevent correct judgments of information leakage. Furthermore, when countermeasures are applied, the data dependence of encryption is concealed. The peak-to-peak amplitude of EM signals does not expose real leakage locations.

5.2 Proposed Leakage Localization Method

The proposed EMA includes two steps: near-field scan and leakage localization. In near-field scan, EM signals are acquired. In leakage localization, a leakage indicator is used to identify leakage locations. They are explained in detail in this section.

5.2.1 Near-field Scan for EMA

Near-field scan is a technique that is used to specify the radiated source on LSIs or printed circuit boards (PCBs). It has been standardized as International Electro-technical Commission (IEC) 61967-3[17]. The near-field scanning system comprises a magnetic-field probe, a device under test (DUT), a sustentation and positioning instrument which is used to fix and move the probe over the DUT. Moreover, a spectrum analyzer or oscilloscope is required to receive the measured values from the magnetic-field probe. A preamplifier, which magnifies weak signals, is optional.

A typical near-field scanning system is similar for EMA experiment shown in Fig. 1.8. In the context of EMA, an exact computation for the strength of the measured EM field is not necessary because the voltage output from the probe is proportional to the EM field around the cryptographic LSI and it represents the activity of each encryption. In DEMA or CEMA, a differential voltage or correlation coefficient is sufficient to detect the correct key. In addition, although the quality of the obtained EM signals depends on the utilized probes, there is no standard for its size in the application to EMA.

After setting up of a near-field scanning system, it is used to acquire EM signals over the surface of DUT when the encryption algorithm runs. Suppose that at each scanning point, N different random plaintexts are used, during each run i (i =1,2,…, N), an EM signal trace $W_i(t)$ is recorded, which consists of encryption-related signals $S_i(t)$ and independent noise η, expressed by

$$W_i(t) = S_i(t) + \eta \qquad \text{(Eq.5.4)}$$

where t is sampling time. In this paper, we assume that noise is well reduced by preprocessing techniques.

5.2.2 The Equivalence of Instant Signal Variance

To localize hot spots of DUT, i.e., cryptographic LSI, the most accurate method is to perform EMA with signal traces at each scanning point. Then the locations where EMA succeeds faster are hot spots. However, the time computation for such an exhaustive method is quite large. Every key candidate must be examined to test the success of EMA at each location. Moreover, hot spots cannot be exposed unless EMA is conducted. To enable an accurate prediction of the hot spots and reduce the computation, we attempt to devise a leakage indicator, which is an equivalent metric for EMA to localize hot spots and avoid the computation of key searches. Signal variance is such a metric. The derivation and proposition are shown below.

In fact, we do not compute $D_\beta(t)$ to localize hot spots at the scanning point, but attempt to look for a substitute. It is noted that the variance of EM signal $S(t)$ is $Var[S(t)]$, given by

$$\begin{aligned} Var[S(t)] &= \mathrm{E}[(S(t) - \mathrm{E}[S(t)])^2] \\ &= \mathrm{E}[(aH + b - \mathrm{E}[aH + b])^2] \\ &= a^2 Var[H] \end{aligned} \quad \text{(Eq.5.5)}$$

The covariance of $S(t)$ and H is expressed as

$$\begin{aligned} Cov[S(t), H] &= Cov[aH + b, H] \\ &= Cov[aH, H] + Cov[b, H] \\ &= aVar[H] \end{aligned} \quad \text{(Eq.5.6)}$$

where $Cov[b, H] = 0$, since b and selection function H are independent. Then Eq. 5.5 rewrites as

$$Var[S(t)] = a(aVar[H]) = aCov[S(t), H] \quad \text{(Eq.5.7)}$$

The covariance of $S(t)$ and H is calculated as

$$\begin{aligned} Cov[S(t), H] &= \mathrm{E}[(S(t) - \mathrm{E}[S(t)]) \cdot (H - \mathrm{E}[H])] \\ &= \mathrm{E}[S(t) \cdot (H - \mathrm{E}[H]) - \mathrm{E}[S(t)] \cdot H + \mathrm{E}[S(t)] \cdot \mathrm{E}[H]] \\ &= \mathrm{E}[S(t) \cdot (H - \mathrm{E}[H])] - \cancel{\mathrm{E}[S(t)] \cdot \mathrm{E}[H]} + \cancel{\mathrm{E}[S(t)] \cdot \mathrm{E}[H]} \\ &= \mathrm{E}[S(t) \cdot (H - \mathrm{E}[H])] \end{aligned} \quad \text{(Eq.5.8)}$$

Then, according to the definition of mathematical expectation, Eq. 5.8 rewrites as

$$Cov[S(t),H] = \sum_s \sum_h \mathrm{P}[S(t)=s, H=h] \cdot s \cdot (h - \mathrm{E}[H])$$

$$= \sum_s \sum_h \mathrm{P}[S(t)=s \mid H=h] \cdot \mathrm{P}[H=h] \cdot s \cdot (h - \mathrm{E}[H]) \quad \text{(Eq.5.9)}$$

For single-bit selection function H, the probability of its value being 1 and 0 is equal, namely, P[H=1]=P[H=0]=1/2, and H-E[H] $\in$ {-1/2,+1/2}. Thus Eq. 5.9 rewrites as

$$Cov[S(t),H] = \sum_s \mathrm{P}[S(t)=s \mid H=1] \cdot s \cdot (-\frac{1}{2}) \quad + \sum_s \mathrm{P}[S(t)=s \mid H=0] \cdot s \cdot (+\frac{1}{2})$$

$$= \frac{1}{2} \sum_s -\mathrm{P}[S(t)=s \mid H=1] \cdot s + \sum_s \mathrm{P}[S(t)=s \mid H=0] \cdot s$$

$$= \frac{1}{2} \mathrm{E}[S(t) \mid H=1] - \mathrm{E}[S(t) \mid H=0]$$

$$= \frac{1}{2} D_\beta(t) \quad \text{(Eq.5.10)}$$

From Eq. 5.8 and Eq.5.10, we have the relation

$$D_\beta(t) = \frac{2}{a} Var[S(t)] \quad \text{(Eq.5.11)}$$

When the selection function is multi bit, H=H(C_i, ϵ, K_c), where ϵ =$\beta_1\beta_2...\beta_G$. For example, H is the Hamming weight of the 8-bit output of SubBytes computation for AES, and $H \in$ {0,1,2,3,4,5,6,7,8}. ϵ denotes 8 bits of the s-box, and G=8. DoM computes the differential trace $D(t)$ as a sum of each examined bit in the case of single bit, given by

$$D(t) = D_{\beta_1}(t) + D_{\beta_2}(t) + ... + D_{\beta_G}(t) \quad \text{(Eq.5.12)}$$

$$= G \cdot D_\beta(t)$$

$$= \frac{2G}{a} Var[S(t)] \quad \text{(Eq.5.13)}$$

under the assumption that each bit contributes identically to the power dissipation. Indeed, this assumption is true for a number of hardware platforms, such as ASIC. Thus, Eq. 5.13 is obtained, which proves that DoM is equal to the signal variance of EM emissions despite a constant gain of $\frac{2G}{a}$.

Therefore, our proposition is: the signal variance Var[S(t)] at time t for N leakage signals, given by Eq.5.14, is used as an equivalent metric to DoM to test data

dependence.

$$Var[S(t)] = \frac{1}{N}\sum_{i=1}^{N}(S_i(t) - \frac{1}{N}\sum_{i=1}^{N}S_i(t))^2 \quad \text{(Eq.5.14)}$$

5.2.3 EMA with Instant Signal Variance

This proposition means that signal variance is the equivalent metric for evaluating the dependence between EM emissions and data encryption, because DEMA identifies the correct key by DoM test, whereby the dependence between EM emission and data encryption can be evaluated. For a certain encryption implementation, a high signal variance denotes intensive fluctuation of the EM field, which is caused by the dynamic change of instantaneous current in the LSI. This dynamic change is due to the switching activities of its components, i.e., from 0 to 1, or 1 to 0. In other words, a high variance represents strong dependence on input data, and a low variance means that the instantaneous signal remains the same and is independent of input data. This is the reason why signal variance can reveal information leakage. It also indicates that there is no direct relationship between "peak-to-peak amplitude" and the evaluation metric. Although a high "peak-to-peak amplitude" means strong EM emission, this emission is not necessarily data-dependent. Thus it cannot accurately express the data dependence of cryptographic operation, and it may result in misjudgments of hot spots. It is not an optimal indicator for localizing leakage.

The time complexity is reduced by computing signal variance. Suppose that the sampling length is M for each of N signal traces. DoM decides the correctness of only one partitioning at each run. Thus for a DEMA attack that attempts L partitioning, where L is the size of one subkey, e.g., $L=2^8$ for AES, it requires a time complexity of $\theta(NML)$. For a CEMA attack, the correlation coefficient between signal traces and leakage model must be calculated. It is $\theta(2NML)$. With our proposition, the time complexity is $\theta(NM)$, because the partitioning for guessing key is avoided.

Signal variance can be used to disclose the information leakage of implementations with countermeasures. As we know, when countermeasures, either masking or hiding, are applied to cryptographic modules, information leakage becomes difficult to detect by SCAs because of the concealment of data-dependent operations. However, they still exist. In a masked circuit, a logic gate potentially switches more than once during one clock cycle, which results in considerable amount of power dissipation. Thus this dissipation of the gate is still correlated to some unmasked inputs and outputs. The masked implementations are susceptible to DEMA and DPA attacks. For countermeasures that use dual-rail circuits, such as WDDL, MDPL, when input signals have a difference of delay time, the timing of starting the power dissipation varies independent of the signal values during an operation cycle. Then the difference of power dissipation remains detectable by DEMA and DPA. Thereby, the signal variance is still capable of identifying information leakage in these cases. Nevertheless, more signal traces are required to expose hot spots.

The equivalence of signal variance to the DoM test has been presented. For every scanning point, the above proposition holds. Therefore, with the signals acquired for each scanning point from near-field scan, we calculate the signal variance at instant $t = \tau$, which is the time the examined value is handled (Note that this instant is estimated in accordance with the attacked encryption operation in specific implementation). A leakage map can be plotted. Hot spots are those locations with higher values of signal variance. EMA succeeds faster at these locations.

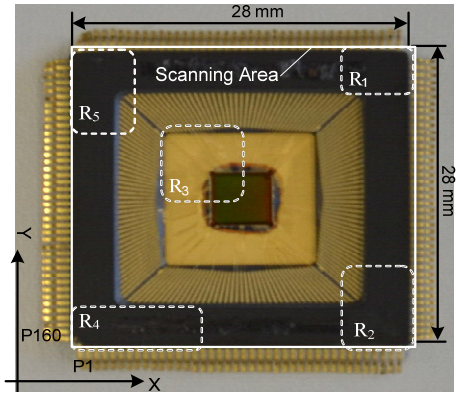
5.3 EMA Based on Proposed Method

The verification of the proposed method on unprotected module and protected module is shown in this section.

5.3.1 EMA on Unprotected Module

A near-field scan over the surface of the LSI when AES PPRM1 implementation runs, is carried out. The origin of the Cartesian coordinate system is set at the corner of pin1 and pin160 of the LSI, which has a package area of 28 mm x 28 mm, shown in Fig. 5.1. The probe plane is kept at 0.5 mm over the packaged surface in order to receive the strong vertical field, and it moves in steps of 1.0 mm from location (1, 1). Thus, there are 784 (28x28) scanning points. Encryption proceeds with 10000 random plaintexts at each point and a fixed but randomly chosen 16-byte key (the final round): 28 AF CE 9F 5A FF C8 F1 E0 54 B3 52 B0 CE 43 0E. The EM signals $W_i(t)$ (i=1,2,…,10000, and t = [1,1000] ns) are acquired and averaged 30 times by the oscilloscope. The sampling rate is 1G Sa/s, and 1000 points are recorded for each sample. This covers the total encryption period. Start timing for trigger signal is the EXEC signal, which is obtained from the pin of encryption execution on LSI. The clock cycle of encryption is 41.6 ns. A signal trace at location (1, 1) is shown in Fig. 5.4. The 10-round encryption and a register access are shown by 11 peaks in 11 clock cycles.

Signal variance in the final round of AES encryption is calculated according to Eq. (13), and normalized to [0,1]. The leakage map is shown in Fig. 5.3(a). Since we are interested in the output of s-box in the final round for analysis, t = 709.4 ns at each point. At this instant, the 10-round encryption is finished, and the ciphertext is to store in the data registers in the area of silicon die. The leakage map indicates several active and inactive regions. Four regions (R_1-R_4), which have hot spots, and region R_5, which has cold spot (note that cold spot is defined as the point with the minimum value on leakage map in this paper), are marked with rounded rectangles in Fig. 5.3(a).



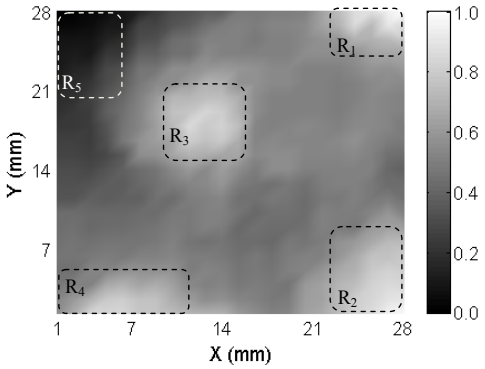
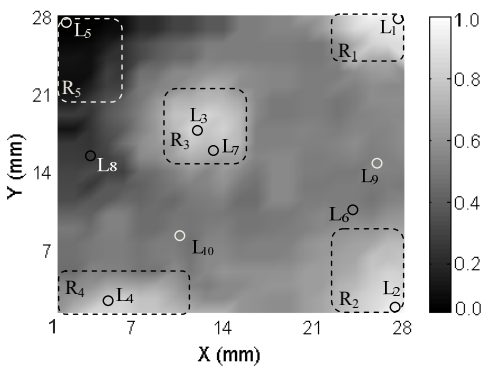


Fig.5.1 Depackaged cryptographic LSI

Fig.5.2 Correlation coefficients of EMA for the scanning area



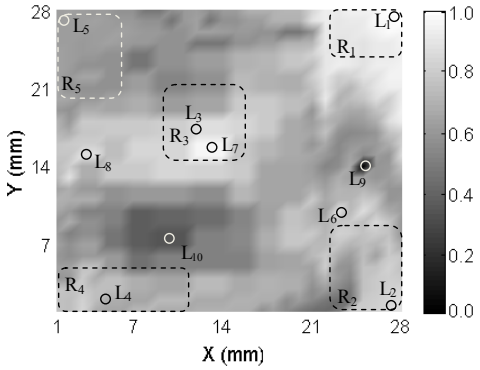


Fig.5.3 (a) Leakage map for AES PPRM1 calculated with proposed method, (b) Leakage map for AES PPRM1 calculated with peak-to-peak amplitude [96]

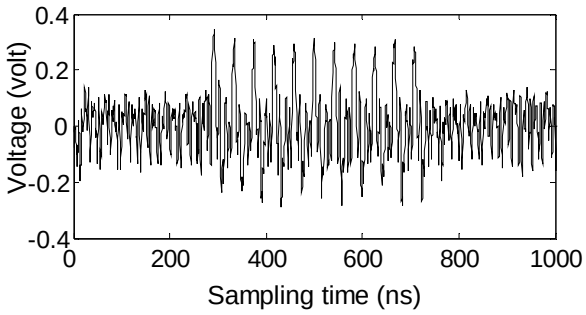


Fig.5.4 Signal trace of AES PPRM1 at (1,1)

The leakage map in Fig. 5.3(a) agrees with present activity of cryptographic LSI according to the manual of LSI [99]. These regions(R_1-R_5) are also marked in Fig.5.3. As we have mentioned, higher signal variance represents strong dependence on data encryption activities of LSI. Region R_3 which is around the silicon die, exhibits hot spots due to the encryption of the cryptographic core.

Region R_1 which is around the pins of data output, and region R_2, which is in the vicinity of the address bus, also have hot spots. Although pins around R_4 and R_5 are not active at this moment, R_4 has hot spots because of the EXEC pin. A summarization of the connected pins, present activity, and range of signal variance of these regions is listed in Table 5.1.

Table 5.1 Summarization of regions R_1-R_5 and calculated signal variance

Region	Connected pins	Present activity	Range of signal variance
R_1	Data output	Active	[0.8712, 1.0000]
R_2	Address bus	Active	[0.7955, 0.9103]
R_3	Silicon die	Active	[0.8339, 0.8821]
R_4	N.C.*, EXEC*	Inactive, active	[0.6818, 0.8014]
R_5	Data input	Inactive	[0.0000, 0.2052]

*N.C.: denotes not connected, *EXEC: execution signal for cryptographic core

It is noted that region R_3 is not at the exact center of the silicon die. This is probably due to the distribution of power/ground grid of the cryptographic LSI. Because of the complexity of this distribution, it is difficult to deduce any characteristic about EM emission. The mechanism behind leakage has been actively studied by researchers, such as Schmidt et al. [100]. It is not discussed further in this paper.

In order to compare the results with conventional methods, maximum peak-to-peak amplitude [96] at the same instant after a subtraction of idle sampling at each point is calculated, normalized and plotted in Fig. 5.3(b).

The locations of hot spots exposed in Fig. 5.3(a) and Fig. 5.3(b) are quite different. They are marked with small circles. Hot spots of Fig. 5.3(a) are L_1, L_2, L_3, and L_4, with L_5 as a cold spot. In Fig.5.3(b), hot spots are L_1, L_6, L_7, and L_8. Cold spots are L_9 and L_{10}. L_6, L_7, and L_8 are not hot spots, and L_9 is not a cold spot in Fig. 5.3(a). But they are identified as hot spots and cold spots, respectively, in Fig. 5.3(b).

To verify whether these hot spots shown in the above two leakage maps are true or not, EMA at 784 locations is performed. Hamming Distance model is used. Correlation coefficients between the signal traces and hypothesized leakage of the output of s-boxes in the final round are calculated to reveal each subkey. In terms of

correlation-based attacks, the correlation coefficient corresponding to the correct key guess represents data dependence and determines MTD. A higher correlation coefficient suggests that the EM emission has a stronger data dependence on encryption, thus secret key is not difficult to detect. In this case, EMA succeeds faster and MTD is small, vice versa. Thereby, the value of correlation coefficient corresponding to the correct key guess at each location is used to represent the performance of EMA. It is normalized to [0,1] and plotted in Fig. 5.2. The correlation coefficients of 10 locations L_1-L_{10}, are sorted and listed in descending order in the first column of Table 5.2. In a similar way, the signal variance obtained by the proposed method and peak-to-peak amplitude in[96] are also listed in descending order in Table 5.2. It is expected that the orders of leakage indicators (signal variance or peak-to-peak amplitude) are consistent with the orders of the results of EMA. In this case, the method is accurate, and the leakage indicator correctly reveals the data dependence at each location.

Table 5.2 Results of two methods at locations L_1-L_{10}

EMA results Corr., MTD, Loc.*	**Proposed** Loc.,Signal Var.*	**Method[96]** Loc.,Peak.Amp.*
0.2113, 3218, **L_1(28,28)**	**L_1**, 1.0000	L_1, 1.0000
0.1989, 3590, **L_2(27,01)**	**L_2**, 0.9103	L_7, 0.9762
0.1896, 4713, **L_3(13,16)**	**L_3**, 0.8821	L_8, 0.9215
0.1783, 5421, **L_4(05,02)**	**L_4**, 0.8014	L_6, 0.8819
0.1715, 5986, **L_7(14,17)**	**L_7**, 0.7380	L_3, 0.8734
0.1290, 6495, **L_{10}(11,08)**	**L_{10}**, 0.4979	L_2, 0.8246
0.1161, 7542, **L_6(24,11)**	**L_6**, 0.3015	L_4, 0.7043
0.1032, 8270, **L_9(26,14)**	**L_9**, 0.2928	L_5, 0.5921
0.0797, 9251, **L_8(04,16)**	**L_8**, 0.1276	L_{10}, 0.2852
0.0634, 9982, **L_5(02,27)**	**L_5**, 0.0000	L_9, 0.0000

* Corr., MTD, Loc.: correlation coefficient, MTD, and location, respectively
* Loc.,Signal Var.: location and signal variance, respectively
* Loc., Peak.Amp.: location and peak-to-peak amplitude, respectively

Table 5.2 is the comparison of the two methods with the results of EMA at 10 locations. The orders determined by the proposed method agree well with the orders of correlation coefficients from EMA. EMA succeeds fastest at L_1, where the maximum correlation coefficient reaches 0.2113, and only 3218 MTD is required to

detect secrete key. Both methods correctly predict that L_1 is a hot spot. However, EMA succeeds slower at L_7, L_6, and L_8 than at L_1, L_2, L_3, and L_4. The method in[96] is unable to reveal this relative relation. On the contrary, this is correctly indicated by the proposed method. In other words, the data dependence is misjudged by method[96], whereas the hot spots indicated by our proposed method are accurate.

To determine which leakage map (Fig. 5.3(a) or Fig.5.3 (b)) better agrees with Fig. 5.2, namely, to quantitatively evaluate the accuracy of proposed method and the method of peak-to-peak amplitude[96], we adopt a "sorting and consistency counting" approach for all the scanning points. Firstly, sorting, the correlation coefficients from EMA are ranged in descending order. The locations identified by the proposed method and method [96] are also sorted in descending orders according to the signal variance and peak-to-peak amplitude respectively. Secondly, consistency counting, for one location, if its order determined by one method matches with its order determined by EMA, then this location is counted as consistent, and that method is considered as accurate. If there is no match, the method is not accurate. Finally, the accuracies of the two methods are evaluated and listed in Table 5.3.

Table 5.3 Accuracy calculations for the two methods at scanning area

EMA results Corr., MTD, Loc.*	**Proposed** Loc., Signal Var.*	**Method[96]** Loc., Peak.Amp.*
0.2113, 3218, **(28,28)**	**(28,28)**, 1.0000	(28,28), 1.0000
0.2082, 3365, **(28,27)**	**(28,27)**, 0.9791	(28,27), 0.9923
0.2016, 3380, **(27,28)**	**(27,28)**, 0.9348	(28,26), 0.9881
0.2001, 3417, **(26,28)**	**(26,28)**, 0.9250	(27,28), 0.9642
0.1998, 3428, **(24,28)**	**(25,28)**, 0.9187	(28,25), 0.9576
…	…	…
0.0672,9680, **(01,25)**	**(01,25)**, 0.0236	(26,12), 0.2454
0.0663,9762, **(01,26)**	**(01,26)**, 0.0187	(25,13), 0.1730
0.0659,9775, **(02,28)**	**(02,28),** 0.0158	(25,14), 0.1326
0.0657,9831, **(01,27)**	**(01,27),** 0.0104	(26,13), 0.0578
0.0634,9982, **(02,27)**	**(02,27),** 0.0000	(26,14), 0.0000
Accuracy	**573/784≈73.1%**	**192/784≈24.5%**
Improved Accuracy	**73.1%-24.5%= 48.6%**	

* Corr., MTD, Loc.: denotes correlation coefficient, MTD, and location, respectively
* Loc.,Signal Var.: denotes location and signal variance, respectively
* Loc., Peak.Amp.: denotes location and peak-to-peak amplitude, respectively

Table 5.3 shows the accuracy calculations for the two methods at the scanning area. The results indicate that the leakage map Fig.5.3(a) calculated by the proposed method better fits Fig. 5.2. Most of the orders determined by the proposed method agree with the orders determined by EMA. The proposed method has an accuracy of 73.1%. By contrast, the orders calculated from method [96] appear to be inconsistent. Only 24.5% locations agree with the order determined by EMA. The accuracy of proposed method is improved by 48.6% compared with that of the method [96].

The above experiments confirm that signal variance accurately reveals the data dependence of encryption that leads to the success of EMA, and peak-to-peak amplitude suffers from misjudgments of data dependence. In addition, it is noted that the accuracies of the two methods are not as high as expected. There are several possible reasons.

The first possible reason is the influence of noise. Signal variance and peak-to-peak amplitude of EM emission are influenced by noise during signal acquisition. To show a naive result of the proposed method, only averaging of signal traces was adopted to attenuate surrounding noise in the above experiments. More sophisticated techniques can be applied to reduce noise during the preprocessing to improve the accuracy of these methods. A detailed discussion of noise sources and reductions can be found in [83, 90]. It is not iterated here.

The second possible reason is the accuracy calculation approach. The approach of “sorting and consistency counting” was used in the experiment to quantitatively compute the accuracy of these two methods. This is a strict evaluation approach. For higher accuracy, it requires a correct relative relation between the points around hot spots. However, it is fair to use it for evaluating these two methods, and it shows that the proposed method is more accurate.

5.3.2 EMA on Protected Module

WDDL proposed by Tiri and Verbauwhede [31], is a countermeasure in the family of Dual-rail with Precharge Logic (DPL) that attempts to make power consumption independent of manipulated data. WDDL includes two stages of the calculation: precharge and evaluation. The differential signals are forced into the same state in the precharge phase. Then, in the evaluation phase, either the true or false signal turns to the opposite state. Thereby the number of transitions is considered as constant when switching from the precharge to the evaluation phase. However, as pointed out by Suzuki and Saeki[77], because of the flaw that there is leakage caused by the difference in delay time between input signals of WDDL gates, it is still vulnerable to SCA.

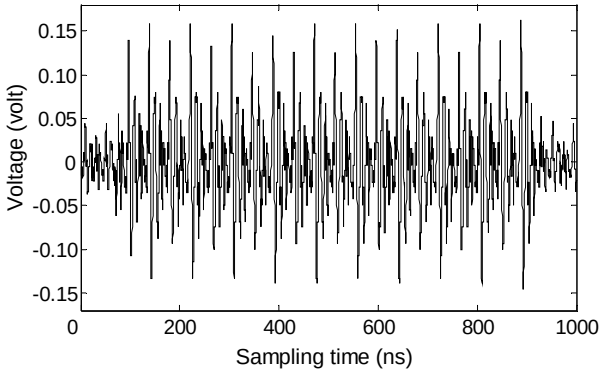


Fig.5.5 Signal trace of AES WDDL at location (1,1)

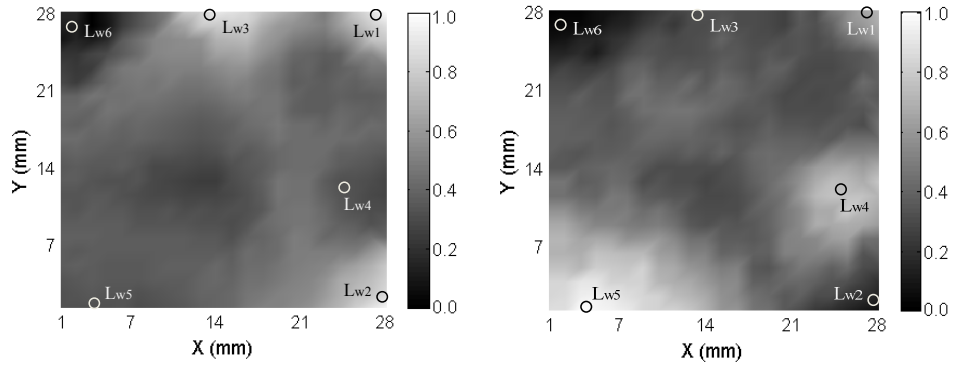


Fig.5.6 (a) Leakage map for WDDL calculated by proposed method (b) Leakage map for WDDL calculated with peak-to-peak amplitude [96]

A near-field scan over the surface of LSI when AES WDDL runs, is conducted. 20000 samplings are acquired at each scanning point. A signal trace at (1, 1) is shown in Fig. 5.5. The 10-round encryption is shown by 20 peaks.

Signal variance at t = 887.1 ns in the final round of encryption for each point is computed and plotted in Fig.5.6(a). A leakage map by computing peak-to-peak amplitude [96], is shown in Fig. 5.6(b). The hot spots in these two leakage maps are totally different. Three hot spots, Lw1, Lw2, and Lw3, are indicated by Fig. 5.6(a), while two other hot spots, Lw4 and Lw5 are revealed in Fig. 5.6(b). Their positions are further away from each other over the surface of LSI. The positions of hot spots Lw4 and Lw5 exhibit rather dark in Fig. 5.6(a). The cold spot Lw6 in Fig. 5.6(a) agrees with that in Fig. 5.6(b).

Table 5.4 EMA results and two leakage indicators for AES WDDL at 6 locations

Loc.	Coordinates	MTD	Corr.	Proposed*	Method[96]*
Lw3	(14,28)	12,057	**0.0713**	**1.0000**	0.5205
Lw2	(28,02)	12,732	**0.0689**	**0.9732**	0.1814
Lw1	(27,28)	13,169	**0.0630**	**0.8874**	0.7829
Lw5	(04,01)	>20,000	**0.0302**	**0.4136**	1.0000
Lw4	(25,12)	>20,000	**0.0281**	**0.2912**	0.9816
Lw6	(02,27)	>20,000	**0.0154**	**0.0000**	0.0083

*Proposed: The signal variance is calculated and normalized

*Method[96]: The peak-to-peak amplitude is calculated and normalized

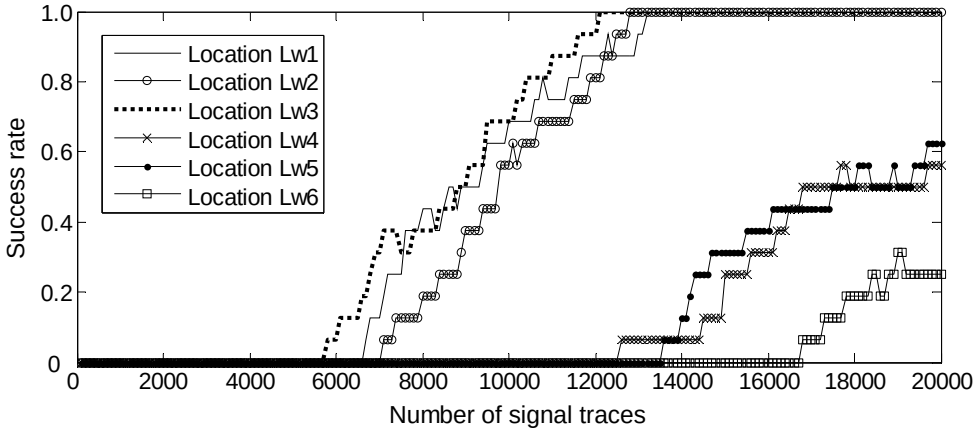


Fig.5.7 Success rates for AES WDDL at 6 locations

Table 5.5 MTD and maximal correlation for each s-box of AES WDDL at Lw3 and Lw5

S-box	S1	S2	S3	S4	S5	S6	S7	S8
Lw3	**12,057**	6,739	8,806	9,503	**5,864**	11,127	9,104	10,274
	0.0713	0.0943	0.0848	0.0796	**0.1022**	0.0720	0.0815	0.0771
Lw5	**x***	14,031	19,436	x	**13,812**	x	19,815	17,523
	0.0302	0.0495	0.0342	0.0336	**0.0517**	0.0311	0.0340	0.0376
S-box	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**	**S16**
Lw3	9278	6,873	11,564	8,721	6,925	10,338	6,121	7,012
	0.0802	0.0921	0.0729	0.0867	0.0927	0.0765	0.0980	0.0913
Lw5	x	14,308	x	16,149	14,797	x	14,176	15,534
	0.0331	0.0429	0.0318	0.0382	0.0418	0.0324	0.0489	0.0407

* x: the subkey is not revealed with 20000 MTD.

Correlation-based EMA at the 6 locations is performed to verify the hot spots. We are more interested in the correctness of the proposed method in the case of countermeasures. Therefore, instead of a strict "sorting and consistency counting" approach at all the scanning points to compare the accuracy, only sorting is used. The locations are shown in descending order according to the values of correlation coefficient, and the signal variance and peak-to-peak amplitude are also listed in Table 5.4. All the subkeys are revealed at Lw3, Lw2, and Lw1 within 20000 signal traces, but not at Lw5 and Lw4. Table 5.4 indicates that the values of signal variance agree well with the correlation coefficients at 6 locations.

Success rates of EMA at the 6 locations are shown in Fig. 5.7. It clearly displays that EMA succeeds faster at Lw3, Lw2, and Lw1. The success rate is 62.5%, namely, only 10/16 subkeys are recovered at Lw5 and 9/16 at Lw4 when signal traces reach 20000. In other words, Lw3, Lw2, and Lw1 are hot spots, but Lw5 and Lw4 are not. This is correctly indicated by the proposed method.

The results of EMA at Lw3 and Lw5 are shown in Table 5.5. The fastest guess for the key is the fifth s-box, where 5864 signal traces are required at Lw3 and 13812 signal traces at Lw5. The slowest guess is for the first s-box. Table 5.5 further demonstrates that EMA succeeds faster at Lw3 than at Lw5. This confirms that the

proposed method correctly reveals this data dependence and predicts the possible leakage locations in the presence of WDDL.

It is noted that 10000 signal traces in section 5.3.1 and 20000 signal traces in section 5.3.2 are acquired for each scanning point. They are sufficient for this experimental configuration. The number of signal traces varies in terms of signal-to-noise ratio of a specific platform. For instance, if stronger countermeasures are applied, then the signal-to-noise ratio decreases, and more signal traces are required to compute signal variance.

Furthermore, a trigger signal is used to align the signal traces during signal acquisition in the experiments presented in section 5.3.1 and section 5.3.2. If other countermeasures, such as the insertion of random delays, are applied in the implementation, additional preprocessing techniques, such as phase-only correlation proposed by Homma et al.[87], are necessary to remove the displacements in signal traces.

5.4 Summary

In this chapter, instant signal variance was proposed as an indicator for localizing hot spots over the surface of cryptographic LSI. It was proved as an equivalent metric to DoM in classical DEMA. Although signal variance does not reveal the specific locations of cryptographic modules by near-field scan, it is capable of identifying data-dependent EM emissions, which leads to the success of EMA. Blind placement is avoided, thus EMA is conducted accurately. Additionally, signal variance is also effective in finding leakage points when countermeasures are applied. Furthermore, a small and low-cost probe was made to verify the proposed method. The experiment of EMA against AES PPRM1 implementation revealed that misjudgments of the leakage are reduced and the accuracy is improved 48.6% compared with the method of peak-to-peak amplitude. The experiment on AES WDDL implementation

demonstrated that a faster EMA is enabled under the guidance of signal variance. The performance of EMA is enhanced.

We have shown the richness of the information disclosed by signal variance based on near-field scan in the time domain, which is an effective tool to explore the secret of cryptographic LSI. In the future, with this tool, more features of EM emissions in the frequency domain will be studied to improve the performance of EMA.

6 Conclusion

In this dissertation, the performance of EMA and PA are enhanced by the proposed algorithms. The algorithms cover the 3 key aspects of side channel analysis, that is: noise reduction, model improvement and equivalent test method.

For the noise reduction, three techniques are proposed to reduce the correlated noise for EMA. From the observation and simulation, we discovered that unlike the encryption signal, the clock signal has a high variance at the signal edges. Then based on this property, the first and second techniques: single-sample SVD and multi-sample SVD reduce the correlated noise by extracting the high variance component from encryption signal. And the third technique: averaged subtraction is efficient when background samplings are included. The main characteristics of the proposed techniques are: single-sample SVD can extract the clock signal with only one EM sample. Multi-sample SVD is capable of suppressing the clock signal with short sampling length. The averaged subtraction is suitable for estimation of correlated noise. Furthermore, these techniques are validated by the EM emission acquired from the AES implementation on both ASIC and FPGA. Compared with existed noise reduction methods, the proposed three techniques increase the SNR as high as 22.94dB, and the success rates of EMA shows that the data-independent information is retained and the performance of EMA is enhanced.

In addition, Source Recovery algorithm is proposed to reduce the simultaneous noise that occurs to the EM side channel. The fourth-order cumulant-based Gaussian noise reduction strategies presented by Le et al. fail to deal with this type of noise. The proposed Source Recovery algorithm takes advantage of the FastICA algorithm to separate the single encryption from mixed encryptions, and then by the amplitude recovery follows the correlation judgment to attenuate the noise. The effectiveness is demonstrated through the analyses of

multiple AES and Camellia encryption modules on synthesized application-specific integrated circuit (ASIC). Experiments show that the proposed algorithm recovers the secret key in the presence of the simultaneous noise. The number of signals needed to reveal keys has been dramatically reduced by 47.8%. And the performance of EMA is greatly enhanced. In addition, the results also provide enlightment for the design of countermeasures. It is that the mixed execution of different encryption sources can be bypassed with signal processing techniques, which means it is not an effective countermeasure.

For the model improvement, a new Switching Glitch leakage model is proposed. It not only considers the data dependent switching activities but also includes glitch power consumptions in cryptographic module. Furthermore, the switching factor and glitch factor are introduced in the model. And from a theoretical point of view, we show how to estimate these factors. The advantage of this model is that the factors can be adjusted according to the analyzed devices during evaluation, which makes it device specific and more accurate for the modeling of power consumption. The EMA on AES implementation validates the proposed model. Compared with conventional Hamming Distance model, the power traces of recovering keys have been decreased by as much as 24.5%.

For the equivalent test method, a novel leakage localization method is proposed for EMA. Based on the EM emission acquired from near field scan, the instant signal variance of EM emission is proved as an equivalent statistical test to DoM test. Thus, it is proposed to identify the information leakage of cryptographic modules. Therefore, by calculating the instant signal variance at each scanning point and computing the higher values, the points that have data-dependent EM emission are disclosed, namely, the leakage locations are found. And the time complexity is also reduced compared with conventional EMA. In addition, a small and low-cost probe is made to verify the proposed EMA on ASIC implementations. The EMA on AES PPRM1 implementation indicates that misjudgments of the leakage are reduced, and the accuracy is

improved by 48.6% compared with existing methods. Moreover, the EMA on AES WDDL implementation shows that proposed method is also effective to expose the leakage locations in the presence of countermeasure

With the above proposed algorithms, the performance of EMA and PA are enhanced. The efficiency of security evaluation is improved, and guidelines for the test practices in industries are also shown.

The future work covers three aspects.

For noise reduction, the uniqueness of the correlated noise will be further studied to detect this noise. Although the frequently occurred noise either in EM or power side channels has been reduced successfully, the noise detection is still an open problem. The noise detection means to differentiate the noise and the leakage signals. To some extent, it is a signal classification problem. The target is to identify what the noise signal is, and how strong it is among the acquired samples. And one needs to differentiate the data-dependent signal from the data-independent signal. This is also quite challenging. Even if the implementation details, such as the cryptographic algorithm, hardware architecture, and the gate counts, are known, the data-dependent signal and the data-independent signal is still as a black box to the analyzer due to the complexity of the generation, transmission, etc. Luckily, the correlated noise has been reduced based on the discovered characteristics in this dissertation. Its uniqueness will be further studied to identify this noise.

For leakage model, the switching glitch leakage model will be extended to analyze the cryptographic LSI with different countermeasures. The proposed leakage model is targeted at analyzing the cryptographic LSI without countermeasure. The possibilities of improving the leakage model based on considering both the glitch and switching activities has been explored. However, when the crypto algorithm implemented with countermeasure, the situation becomes more complex. Both the switching characteristics and the glitch are altered due to the additional hardware circuits. They might be masked or hidden

completely or incompletely. The switching glitch leakage model needs to be extended to consider the new characteristics incurred by the countermeasure.

For leakage localization method, more features in frequency domain will be studied to further improve the localization accuracy. The instant signal variance-based method utilizes the statistical characteristics in time domain. The data-dependent information has been exploited in time domain. As we know, the time-domain signal is straight expression for its periodicity, peak, mean, and variance. Due to the energy equivalence principle, the signal also has its frequency expression, which is described by frequency band, phase, and amplitude. There is likely contains more rich information that can be explored whether it is data-dependent. This will be one of the future works.

References

[1] R. Anderson, M. Bond, J. Clulow, S. Skorobogatov, “Cryptographic Processors -A Survey”, Proceedings of the IEEE, Vol.94, No.2, pp.357-369, 2006.

[2] P. Kocher, “Timing Attacks on Implementations of Diffie-Hellmann, RSA, DSS, and Other Systems”, CRYPTO’96, LNCS 1109, pp.104-113, 1996.

[3] P. Kocher, J. Jaffe, B. Jun, “Differential Power Analysis”, CRYPTO’99,LNCS 1666,pp.388-397,1999.

[4] National Security Agency, “NACSIM 5000 Tempest Fundamentals (U)”, Fort George G. Meade, Maryland, USA. Available from http://cryptome.org/nacsim-5000.htm.

[5] J. Kelsey, B. Schneier, D. Wagner, and C. Hall, “Side Channel Cryptanalysis of Product Ciphers”, Proc. 5th European Symposium on Research in Computer Security, LNCS 1485, pp.97-110, 1998.

[6] A. Shamir, E. Tramer, “Acoustic Cryptanalysis: on Noisy People and Noisy Machines”, EUROCRYPT 2004 rump session, 2004.

[7] M. Kuhn, “Optical Time-Domain Eavesdropping Risks of CRT Displays”, Proc of the 2002 Symposium on Security and Privacy, pp.3-18, 2002.

[8] T. S. Messerges, E. A. Dabbish, and R. H. Sloan, “Investigations of Power Analysis Attacks on Smartcards,” USENIX1999, Jun. 1999. http://www.usenix.org/

[9] R. Bevan and E. Knudsen, “Ways to Enhance DPA”, International Conference on Information Security and Cryptology (ICISC 2002), LNCS 2587, pp.327-342, Springer-Verlag, Dec. 2003.

[10] E. Brier, C. Clavier, and F. Olivier, “Correlation Power Analysis with a Leakage Model”, proceedings of CHES 2004, LNCS 3156, pp. 16-29, 2004.

[11] J.-S. Coron, P. Kocher, and D. Naccache, “Statistics and Secret Leakage”, In Financial Cryptography (FC2000), LNCS 1972, pp. 157-173, Springer-Verlag, 2001.

[12] R. Mayer-Sommer, “Smartly Analyzing the Simplicity and the Power of Simple Power Analysis on Smartcards”, In Cryptographic Hardware and Embedded Systems-CHES 2000. LNCS 1965, pp. 78-92, Springer-Verlag, 2000.

[13] E. Oswald, “On Side-Channel Attacks and the Application of Algorithmic Countermeasures”, PhD Thesis, Faculty of Science of the University of Technology Graz (IAIK-TUG), Austria, May 2003.

[14] T. Le, J. Clédière, C. Canovas, et al, “A Proposition for Correlation Power Analysis Enhancement”, LNCS 4249, pp.174, 2006.

[15] P. Fahn and P. Pearson, “IPA: A New Class of Power Attacks”, Proc. CHES 1999, LNCS 1717, pp. 173-186, Massachusetts, USA, 1999, Springer-Verlag.

[16] A. Chari, J. Rao, and P. Rohatgi, “Template Attacks”, Proc. of CHES 2002, LNCS 2523, pp. 13-28, San Francisco Bay, USA, 2002, Springer-Verlag.

[17] IEC 61967-3, “Integrated Circuits-measurement of Electromagnetic Emissions, 150 kHz to 1 GHz, Part 3: Measurement of Radiated Emissions, Surface Scan Method (10 kHz to 3 GHz)”, 47A/620/NP, New Work Item Proposal, Date of proposal: July 2001.

[18] D. J. Griffiths, “Introduction to Electrodynamics (3rd edition)”, Benjamin/Cummings Pub Co. ISBN 0-13-805326-X., section 6.1, 1998.

[19] J.J. Quisquater, D. Samyde, “Electromagnetic Analysis (EMA): Measures and Counter measures for Smart Cards”, E-smart 2001, LNCS 2140, pp.200-210,2001.

[20] K. Gandolfi, C. Mourtel, and F. Olivier, “Electromagnetic Analysis: Concrete Results”, Proc. Workshop on Cryptographic Hardwar and Embedded Systems (CHES), LNCS, Vol. 2162, pp. 251-261, 2001.

[21] D. Agrawal, B. Archambeault, J. R. Rao, and P. Rohatgi, “The EM Side Channel(s) ”, Proc. CHES 2002, LNCS, Vol. 2523, pp. 29-45, 2002.

[22] C. Gebotys, S. Ho and C. Tiu, “EM Analysis of Rijndael and ECC on A Wireless Java-based PDA”, Proc. CHES 2005, LNCS, Vol. 3659, pp. 250-264, 2005.

[23] Suresh Chari, Charanjit S. Jutla, Josyula R. Rao and Pankaj J. Rohatgi, “Towards Sound Approaches to Counteract Power-Analysis Attacks,” Proc. of Advances in Cryptology-CRYPTO ‘99, Springer-Verlag, 1999, pp. 398-412.

[24] Louis Goubin and Jacques Patarin, “DES and Differential Power Analysis-The Duplication Method,” Proc. CHES 1999, Springer-Verlag, August 1999, pp. 158-172.

[25] M. Bucci, M. Guglielmo, and R. Luzzi, et al., “A Power Consumption Randomization Countermeasure for DPA-resistant Cryptographic Processors”, Integrated Circuit and System Design. Power and Timing Modeling, Optimization and Simulation, pp. 481-490, 2004.

[26] D. May, H. Muller, and N. Smart, “Random Register Renaming to Foil DPA”, Cryptographic Hardware and Embedded Systems-CHES 2001, pp. 28-38, 2001.

[27] J. Golic, “DeKaRT: A New Paradigm for Key-dependent Reversible Circuits”, Cryptographic Hardware and Embedded Systems-CHES 2003, pp. 98-112 ,2003.

[28] R. Elbaz, L.Torres, and G. Sassatelli, et al., “Hardware Engines for Bus Encryption: A Survey of Existing Techniques”, Proc. Design, Automation and Test in Europe conference and Exhibition (DATE), pp. 40-45, 2005.

[29] C. Tokunaga, D. Blaauw, “Secure AES Engine with A Local Switched-capacitor Current Equalizer”, Proc.2009 IEEE international Solid-State Circuits Conference, pp.64-66, 2009.

[30] K.Tiri, M. Akmal, and I. Verbauwhede, “A Dynamic and Differential CMOS Logic with Signal Independent Power Consumption to Withstand Differential Power Analysis on Smart Cards”, Proc. the 28th European Solid-State Circuits Conference, 2002 (ESSCIRC 2002), pp. 403-406, 2002.

[31] K. Tiri, and I. Verbauwhede, “A Logic Level Design Methodology for A Secure DPA Resistant ASIC or FPGA Implementation”, Proc. Design, Automation and Test in Europe conference and Exhibition (DATE), volume 1, pp. 246-251, 2004.

[32] S. Chari, J. R. Rao, and P. Rohatgi, “Template Attacks” , Proc.CHES, LNCS, Vol. 2523. Springer, August 2002, pp. 13-28, San Francisco Bay (Redwood City), USA.

[33] R. Bevan and E. Knudsen, “Ways to Enhance Differential Power Analysis”, in ICISC, ser. Lecture Notes in Computer Science, vol. 2587, pp. 327-342, Springer, November 28-29 2002, Seoul, Korea.

[34] F.-X. Standaert, B. Gierlichs, and I. Verbauwhede, “Partition vs. Comparison Side-Channel Distinguishers: An Empirical Evaluation of Statistical Tests for Univariate Side-Channel Attacks against Two Unprotected CMOS Devices”, in ICISC, LNCS, vol. 5461, pp. 253-267, Springer, December 3-5 2008, Seoul, Korea.

[35] L. Batina, B. Gierlichs, and K. Lemke-Rust, “Comparative Evaluation of Rank Correlation Based DPA on an AES Prototype Chip”, Lecture Notes in Computer Science, vol. 5222. Springer, September 15-18 2008, pp. 341-354, Taipei, Taiwan.

[36] W. Schindler, K. Lemke, and C. Paar, “A Stochastic Model for Differential Side Channel Cryptanalysis”, CHES 2005, LNCS, vol. 3659. Springer, Sept 2005, pp. 30–46, Edinburgh, Scotland, UK.

[37] B.Gierlichs, L. Batina, P. Tuyls, B. Preneel, “Mutual Information Analysis-A Generic Side-Channel Distinguisher”, In: Oswald, E., Rohatgi, P. (eds.) CHES 2008. LNCS, vol. 5154, pp. 426-442. Springer, Heidelberg, 2008.

[38] L. Batina, B. Gierlichs, and K. Lemke-Rust, “Differential Cluster Analysis”, in CHES2009, LNCS, vol. 5747.Lausanne, Switzerland: Springer-Verlag, 2009, pp. 112-127.

[39] Y. Souissi, M. Nassar, S. Guilley, J.-L. Danger, and F. Flament, “First Principal Components Analysis: A New Side Channel Distinguisher”, Proc. ICISC, LNCS. Springer, December 1-3 2010, Seoul, Korea.

[40] T. S. Messerges, “Using Second-Order Power Analysis to Attack DPA Resistant Software”, Cryptographic Hardware and Embedded Systems (CHES 2000), LNCS 1965, pp.238-251, Aug. 2000.

[41] J. Waddle and D. Wagner, “Towards Efficient Second-Order Power Analysis”, Cryptographic Hardware and Embedded Systems (CHES 2004), LNCS 3156, pp. 1-15, Springer-Verlag, 2004.

[42] S Mangard, N. Pramstaller ,and E. Oswald, “ Successfully Attacking Masked AES Hardwar Implementation”, Proc. CHES 2005, pp.157-171, 2005.

[43] F.Muller, and F. Valette, “Higher-order Attacks Against the Exponent Splitting Protection ”, Public Key Cryptography-PKC 2006, pp.315-329, 2006.

[44] C. Clavier, J.S. Coron, and N. Dabbous,” Differential Power Analysis in the Presence of Hardware Countermeasures”, Proc. CHES 2000, pp.13-48, 2000.

[45] K.Tiri, D. Hwang, and A. Hodjat, et al., “Prototype IC with WDDL and Differential Routing-DPA Resistance Assessment”, Proc. CHES 2005, pp. 354-365, 2005.

[46] C. Karlof and D.Wagner, “Hidden Markov Model Cryptanalysis”, Proc. CHES 2003, pp. 17-34, 2003.

[47] D. Agrawal, J.R. Rao and P. Rohatgi, “Multi-channel Attacks” , Proc.CHES 2003, LNCS 2779, Springer-Verlag , Cologne, Germany 2003.

[48] J. Waddle and D. Wagner, “Towards Efficient Second-order Power Analysis”, In Proc.CHES 2004, LNCS 3156, Springer-Verlag, pp. 1-15, Cambridge (Boston), USA 2004.

[49] E. Peeters, F.X. Standaert, N. Donckers, and J.J. Quisquater, “Improved Higher-order Side-channel Attacks with FPGA Experiments”, Cryptographic Hardware and Embedded Systems-CHES 2005,pp.309-323,2005.

[50] T.H. Le, J. Clédière, C. Servière and J.L. Lacoume, “Higher Order Statistics for Side Channel Analysis Enhancement”, Proc. of e-Smart 2006, Sophia Antipolis, France, September 2006.

[51] N. Homma, S. Nagashima, and Y. Imai, et al., “High-resolution Side-Channel Attack Using Phase-Based Waveform Matching”, CHES 2006. LNCS, vol. 4249, pp. 187-200. 2006.

[52] T.H. Le, J. Clédière, C. Servière and J.L. Lacoume, “Efficient Solution for Misalignment of Signal in Side Channel Analysis”, In Proceedings of ICASSP 2007, Honolulu, Hawaii, USA, April 2007.

[53] L.Sauvage, S.Guilley, J.L.Danger, Y.Mathieu and M.Nassar, “Successful Attack on An FPGA-based WDDL DES Cryptoprocessor without Place and Route Constraints”, Proc. DATE, pp. 640-645, 2009.

[54] M. Knežević, L. Batina, and E. Mulder, “Signal Processing for Cryptography and Security Applications”, Handbook of Signal Processing Systems, pp. 161-177. 2010.

[55] H. Maghrebi, S. Guilley, and J.L. Danger, et al, “Entropy-based Power Attack”, Proc. 2010 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST), pp.1-6,2010.

[56] O. Meynard, and D. Real, and F. Flament, et al., “Enhancement of Simple Electro-magnetic Attacks by Pre-characterization in Frequency Domain and Demodulation Techniques”, Proc. Design, Automation & Test in Europe Conference &Exhibition (DATE), pp.1-6, 2011.

[57] D.Suzuki, M.Saeki, T .Ichikawa, “DPA Leakage Models for CMOS Logic Circuits”, In: Rao, J.R., Sunar, B. (eds.) CHES 2005. LNCS, vol. 3659, pp. 366–382. Springer, Heidelberg(2005).

[58] E. Peeters, F. X. Standaert, and J.J.Quisquater, “Power and Electromagnetic Analysis: Improved Model, Consequences and Comparisons”, Integration, the VLSI Journal, Elsevier, vol.40, issue 1, pp.52-60, 2007.

[59] Canovas, C. and Clediere, J., “What Do S-boxes Say in Differential Side Channel Attacks”, IACR e-Print archive, volume 311, 2005.

[60] E. Prouff, “DPA Attacks And S-boxes”, In proceedings of FSE 2005, LNCS 3557, pp.424 - 441, Springer, 2005.

[61] T.Sugawara, Y. Hayashi, N.Homma, et al., “Mechanism Behind Information Leakage in Electromagnetic Analysis of Cryptographic Modules”, Proc. of WISA 2009, LNCS, Vol. 5932, pp. 66-78, 2009.

[62] S. Mangard, E.Oswald, and F.X.Standaert, “One for All-all for One: Unifying Standard Differential Power Analysis Attacks”, IET Information Security, vol.5, no.2, pp.100-110, 2011.

[63] A.Shamir, “Protecting smart cards from passive power analysis with detached power supplies”, Cryptographic Hardware and Embedded Systems- CHES 2000 , pp. 121-132, 2000.

[64] P. Corsonello, S. Perri, and M.Margala, “A New Charge-Pump based Countermeasure against Differential Power Analysis ”, 6th International Conference On ASIC (ASICON 2005),volume 1 , pp. 66-69, 2005.

[65] P. Rakers, L.Connell, and T. Collins, et al., “Secure Contactless Smartcard ASIC with DPA Protection”, IEEE Journal of Solid-State Circuits ,volume 36 ,number3 ,pp. 559-565 ,2001.

[66] D.May, H. Muller, and N. Smart, “Non-deterministic Processors”, Information Security and Privacy (book), pp. 115-129, 2001.

[67] S. Yang, W. Wolf, and N. Vijaykrishnan, et al., “Power Attack Resistant Cryptosystem Design: A Dynamic Voltage and Frequency Switching Approach”, Proc. Design, Automation and Test in Europe (DATE), pp. 64-69, 2005.

[68] A. Bystrov, D.Sokolov, and A.Yakovlev, et al., “Balancing Power Signature in Secure Systems”, Proc. 14th UK Asynchronous Forum, 2003.

[69] M. Bucci, L. Giancane, and R.Luzzi, et al., “Three-phase Dual-rail Pre-charge Logic”, Cryptographic Hardware and Embedded Systems-CHES 2006 , pp. 232-241, 2006.

[70] S. Moore, R. Anderson, and P. Cunningham, et al. , “Improving Smart Card Security Using Self-timed Circuits”, Proc. Eighth International Symposium on Asynchronous Circuits and Systems, pp. 211-218, 2002.

[71] M.W.Allam, and M.I.Elmasry, “Dynamic Current Mode Logic (DyCML): A New Low-power High-performance Logic Style”, IEEE Journal of Solid-State Circuits, volume36, number3, pp. 550-558, 2001.

[72] T. Messerges, “Securing the AES Finalists Against Power Analysis Attacks”, Proc. Fast Software Encryption (FSE), pp. 293-301, 2001.

[73] M.L.Akkar, and C. Giraud, “An Implementation of DES and AES, Secure Against Some Attacks”, Cryptographic Hardware and Embedded Systems-CHES 2001, pp. 309-318, 2001.

[74] J.S.Coron, “Resistance Against Differential Power Analysis for Elliptic Curve Cryptosystems”, Cryptographic Hardware and Embedded Systems-CHES1999, pp. 725-725, 1999.

[75] E. Trichina, T. Korkishko, and K. Lee, “Small Size, Low Power, Side Channel-immune AES Coprocessor: Design and Synthesis Results”, Advanced Encryption Standard-AES, pp. 572-572 ,2005 .

[76] J. Fournier, S. Moore, and H. Li, et al., “Security Evaluation of Asynchronous Circuits”, Cryptographic Hardware and Embedded Systems-CHES 2003, pp.137-151, 2003.

[77] D. Suzuki, and M. Saeki, “Security Evaluation of DPA Countermeasures Using Dual-rail Pre-charge Logic Style”, Cryptographic Hardware and Embedded Systems-CHES 2006, pp.255-269, 2006.

[78] G. Piret, and F.X.Standaert, “Security Analysis of Higher-order Boolean Masking Schemes for Block Ciphers (with conditions of perfect masking)”, IET Information Security, volume 2, number 1, pp.1-11, 2008.

[79] S. Guilley, L. Sauvage, P. Hoogvorst, “Security Evaluation of WDDL and SecLib Countermeasures against Power Attacks”, IEEE Transactions on Computers, Volume 57, Number 11, pp.1482-1497, 2008.

[80] Y. Li, K. Sakiyama, S. Kawamura, “Security Evaluation of a DPA-Resistant S-box Based on the Fourier Transform”, Information and Communications Security, pp. 3-16, 2009.

[81] Q. Luo, and Y. Fei, “Algorithmic Collision Analysis for Evaluating Cryptographic Systems and Side-channel Attacks”, 2011 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST), pp.75-80, 2011.

[82] M. Kanda, “Standard Probes for Electromagnetic Field Measurements”, IEEE Transactions on Antennas and Propagation, Vol.41, Issue 10, pp. 1349-1364, 1993.

[83] T.-H. Le, C. Servière, and J. Cledière, “Noise Reduction in the Side Channel Attack Using Fourth Order Cumulants”, IEEE Trans. Inf. Forensic Security, Vol. 2, No. 4, pp. 710-720, 2007.

[84] J. Ryoo, D.G. Han, S.K.Kim, and S. Lee, “Performance Enhancement of Differential Power Analysis Attacks with Signal Companding Methods”, IEEE signal processing letters, Vol.15, pp. 625-628, 2008.

[85] 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, 51(5), pp.541-552, May 2002.

[86] H. Pelletier, and X. Charvet, “Improving the DPA Attack Using Wavelet Transform”, NIST’s Physical Security Testing Workshop, (2005) Honolulu, Hawai, USA, Website: http://csrc.nist.gov/groups/STM/cmvp/documents/fips140-3/physec/papers/physecpaper14.pdf.

[87] N. Homma, S. Nagashima, and T. Sugawara, et al., “A High-Resolution Phase-Based Waveform Matching and Its Application to Side-Channel Attacks”, IEICE Transactions, 91-A(1):193-202, 2008.

[88] C. H. Gebotys and B. A. White, “EM Analysis of a Wireless Java-Based PDA”, ACM Transactions on Embedded Computing Systems, 7(4), pp.1-28, 2008.

[89] J. Kang, D.G. Lee, and D. Choi, “Convolutive Noise Filtering in Power Analysis on Smartcards Using the Cepstrum”, 4th International Conference on Embedded and Multimedia Computing, (EM-Com 200), pp.1-4, 2009.

[90] H. Liu, Y. Tsunoo, and S. Goto, “Electromagnetic Analysis Enhancement with Signal Processing Techniques”, 16th Australasian Conference on Information Security and Privacy (ACISP 2011, Jul. 11-13, Melbourne, Australia) LNCS, Vol. 6812, pp. 456-461, 2011.

[91] A. Hyvärinen, “Fast and Robust Fixed-point Algorithms for Independent Component Analysis”, IEEE Transactions on Neural Networks, vol.10, no.3, pp.626-634, 1999 .

[92] A. Dehbaoui, V. Lomne, and T. Ordas, et.al, “Enhancing Electromagnetic Analysis Using Magnitude Squared Incoherence”, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, No.99, pp.1-5, 2010.

[93] International Electrotechnical Commission (IEC), “IEC 61967: Integrated Circuits-Measurement of Electromagnetic Emissions, 150 kHz to 1 GHz, http://www.iec.ch, 2003.

[94] S. Mangard, E. Oswald, and T. Popp, “Power Analysis Attacks: Revealing the Secrets of Smart Cards”, pp.151, Springer, NewYork , 2007.

[95] M. Pedram, “Power Minimization in IC Design: Principles and Applications”, ACM Transactions on Design Automation of Electronic Systems, Vol. 1, No. 1, January 1996.

[96] L. Sauvage, S. Guilley, and Y. Mathieu, “Electromagnetic Radiations of FPGAs High Spatial Resolution Cartography and Attack on A Cryptographic Module”, ACM Trans. Reconfigurable Technology and Systems, Vol. 2, No. 1, pp. 4-24, 2009.

[97] S. P. Skorobogatov, “Semi-invasive Attacks-a New Approach to Hardware Security Analysis”, Technical Report UCAM-CL-TR-630, University of Cambridge, Computer Laboratory, April, 2005.

[98] J. J. Quisquater and D. Samyde, “Electromagnetic Analysis (EMA): Measures and Countermeasures for Smart Cards”, Smart Card Programming and Security (E-smart 2001) LNCS, Vol. 1240, pp. 200-210, 2001.

[99] Research Center for Information Security (RCIS) of AIST, “Standard cryptographic LSI specification-with side channel attack countermeasures-Ver.1.0”, http://staff.aist.go.jp/akashi.satoh/SASEBO/en/board/crypto_lsi.html.

[100] J. M. Schmidt, T. Plos, M. Kirschbaum, M. Hutter, M. Medwed, and C. Herbst,

"Side-channel Leakage Across Borders", Proc. the 9th Smart Card Research and Advanced Application IFIP Conference (CARDIS 2010), LNCS 6035, pp. 36-48, 2010.

[101] Research Center for Information Security (RCIS) of AIST, "Side-channel Attack Standard Evaluation Board(SASEBO)". http://www.risec.aist.go.jp/project/sasebo/

[102] A. Raghunathan, S. Dey, and N.K. Jha, "High-level Macro-modeling and Estimation Techniques for Switching Activity and Power Consumption", IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Vol.11, NO.4, pp.538-557, 2003.

[103] S.Mangard, T.Popp, and B. M. Gammel, "Side-channel Leakage of Masked CMOS Gates", LNCS 3376, Springer, pp.351-365, 2005.

[104] EPSON Tokyocom, Data sheet of crystal oscillator XG-1000CA/CB (CMOS), http://www.epsondevice.com/docs/qd/ja/DownloadServlet?id=ID000771.

[105] M. Kafi, S. Guilley, and S.Marcello, et al., "Deconvolving protected signals", Proc. IEEE International Conference on Availability, Reliability and Security, ARES'09, pp. 687-694, 2009.

[106] N. Kamoun, L. Bossuet, and A. Ghazel, "Correlated Power Noise Generator as Low Cost DPA Countermeasures to Secure Hardware AES Cipher", Proc. of the International Conference on Signals, Circuits and Systems (ICCS'09), pp. 1-6, October, 2009.

[107] National Institute of Standards and Technology (NIST) of U.S. Department of Commerce, FIPS 197: Advanced Encryption Standard, Nov. 2001.

[108] K. Aoki, T. Ichikawa, M. Kanda, M. Matsui, S. Moriai, J. Nakajima, and T. Tokita, Specification of Camellia--a 128-bit Block Cipher, http://info.isl.ntt.co.jp/crypt/camellia/dl/01espec.pdf, 2001.

[109] L. Tong, R. Liu, V.C.Soon, and Y.-F.Huang, "Indeterminacy and Identifiability of Blind Identification", IEEE Trans on Circuits and Systems, 38(5), pp.499-509, 1991.

[110] L.Tong, and R.Liu, "Blind Estimation of Correlated Source Signals", Proceedings of ACSSC, (1), pp.258-262, 1990.

[111] C. Chang, Z. Ding, S.F.Yau, and D.H.Y.Chan, "A Matrix-pencil Approach to Blind Separation of Colored Non-stationary Signals", IEEE Signal Processing Letters, 7 (12), pp.348-350, 2000.

[112] A. Belouchrani, K. Abed-Meraim, and J-F.Cardoso, et al "A Blind Source Separation Technique Using Second-order Statistics", IEEE Trans.on Signal Processing, 45(2), pp.434-443, 1997.

[113] S.Choi, A.Cichocki, and S.Amari, "Equivariant Nonstationary Source Separation", Neural Networks, 15, pp.121-130, 2002.

[114] J-F.Cardoso, and A. Souloumiac, "Blind Beamforming for Non-Gaussian Signals", IEE Proceedings-F, 140(6), pp.362-370, 1993.

[115] J-F.Cardoso, B. H. Laheld, "Equivariant Adapative Source Separation", IEEE Trans.on Signal Processing, 44(12), pp.3017-3030, 1996.

[116] S. Amari, A. Cichocki, and H. Yang, "A New Learning Algorithm for Blind Source Separation", Advances in Neural Information Processing (Proc.NIPS'95). Cambridge, MA: MIT Press, pp. 757-763, 1995.

[117] N. Delfosse, and P. Loubaton, "Adaptive Blind Separation of Independent Sources: A Deflation Approach", Signal Processing, 45, pp.59-83, 1995.

[118] E. Bingham, and A. Hyvärinen, "A Fast Fixed-point Algorithm for Independent Component Analysis of Complex Valued Signals", International journal of neural systems, Vol.10, No.1, pp.1-8, 2000.

[119] J. G. Yook, V. Chandramouli, L.P.B. Katehi, K.A. Sakallah, T.R. Arabi, and T.A. Schreyer, "Computation of Switching Noise in Printed Circuit Boards", IEEE Transactions on Components, Packaging, and Manufacturing Technology, Part A, Vol.20, No.1, pp.64-75,1997.

[120] G. A. Katopis, "Deltai Noise Specification for A High-performance Computing Machine", Proc. IEEE, pp. 1405-1415, 1985.

[121] A. J. Rainal, "Computing Inductive Noise of Chip Packages", AT&T Bell Lab. Tech. J., pp. 177-195, Jan. 1984.

[122] R. Senthinathan, and J. L. Prince, "Simultaneous Switching Ground Noise Calculation for Packaged CMOS Devices" , IEEE Journal of Solid-State Circuits, Vol.26 , No.11, pp. 1724-1728, 1991.

[123] R. Senthinathan and J. L. Prince, "Simultaneous Switching Noise of CMOS Devices

and Systems" , Boston, MA: Kluwer, 1994.

[124] K. Iokibe，T. Amano，Y. Toyota, "On-Board Decoupling of Cryptographic FPGA to Improve Tolerance to Side-Channel Attacks," Proc. International Symposium on Electromagnetic Compatibility, pp. 925-930, Aug. 2011.

Publications

Journal Papers

1. Hongying Liu, Xin Jin, Yukiyasu Tsunoo, and Satoshi Goto, "Correlated Noise Reduction for Electromagnetic Analysis", IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Jan. 2013 (Accepted).

2. Hongying Liu, Yukiyasu Tsuoo, Yibo Fan, Bin Hu, and Satoshi Goto, "Electromagnetic Analysis Enhancement Based on Near-Field Scan", Journal of Signal Processing, Vol.16, No.3, pp. 241-250, May 2012.

Conference Papers

International Conference (with review)

1. Hongying Liu, Satoshi Goto, and Yukiyasu Tsunoo, "Correlation Power Analysis with Companding Methods", International Conference on Advances in Control Engineering and Information Science (CEIS), Procedia Engineering, Vol.15, pp.2108-2112, Dali, China, Aug. 2011.

2. Hongying Liu, Yukiyasu Tsunoo, and Satoshi Goto, "Electromagnetic Analysis Enhancement with Signal Processing Techniques", 16th Australasian Conference on Information Security and Privacy (ACISP), Lecture Notes in Computer Science (LNCS), No. 6812, pp. 456-461, Melbourne, Australia, Jul. 2011.(The paper was then published by the following journal)
Hongying Liu, Yibo Fan, and Satoshi Goto, "Secret Recovery from Electromagnetic Emissions", Advanced Science Letters, Vol. 7, pp.182-186, 2012.

3. Hongying Liu, Guoyu Qian, Satoshi Goto, and Yukiyasu Tsunoo, "Correlation Power Analysis based on Switching Glitch Model", 11th International Workshop on Information Security Applications (WISA), Lecture Notes in Computer Science (LNCS), No.6513, pp. 191-205, Jeju Island, Korea, Aug. 2010. (The paper was then published by the following journal)
Hongying Liu, Guoyu Qian, Satoshi Goto, and Yukiyasu Tsunoo, "The Switching Glitch Power Leakage Model", Journal of Software, Vol. 6, No. 9, pp.1787-1794, Sep. 2011.

4. Hongying Liu, Guoyu Qian, Satoshi Goto, and Yukiyasu Tsunoo, "AES Key Recovery based on Switching Distance Model", 3rd International Symposium on Electronic Commerce and Security (ISECS), pp. 218-222, GuangZhou, China, Jul.2010.

5.Ying Zhou, Guoyu Qian, Yueying Xing, Hongying Liu, Satoshi Goto, and Yukiyasu Tsunoo, "An approach of using different positions of double registers to protect AES

hardware structure from DPA", 3rd International Symposium on Electronic Commerce and Security (ISECS), pp. 223-227, GuangZhou, China, Jul.2010.

6. Hongying Liu, Satoshi Goto, and Junhuai Li, "An Indoor Localization System with RFID Passive Tags ", International Symposium on Ubiquitous Computing Systems (UCS), pp.1-7, Beijing, China, Aug. 2009.

7. Hongying Liu, Satoshi Goto, and Junhuai Li; " The Study and Application of Tree-based RFID Complex Event Detection Algorithm ", International Symposium on Web Information Systems and Applications (WISA), pp. 520-524, NanChang, China, May 2009.

Domestic Conference (without review)

8. Hongying Liu, Guoyu Qian, Yukiyasu Tsunoo, and Satoshi Goto, "CPA Attack with Switching Distance Model on AES ASIC Implementation", The 27th Symposium on Cryptography and Information Security(SCIS), Kagawa, Japan, Jan. 2010.

Book Chapter

1. Hongying Liu, Satoshi Goto, and Junhuai Li, "An Efficient Approach for Data Transmission in RFID Middleware", Chapter 18 for the book of "Radio Frequency Identification Fundamentals and Applications Bringing Research to Practice ", invited by publisher of IN-TECH, Vienna, Austria, Feb. 2010.