An Improved Method for SO-DPA Based on the Optimal Distinguish Function

Wuyin Wang, Ning Wu, Jinbao Zhang and Fang Zhou

Abstract—**Higher Order Differential Power Analysis (HO-DPA) is a powerful side-channel attack that allows an attacker to bypass the widely used masking countermeasure, so the research of HO-DPA is of great significance to the cryptographic chip design. In this paper, we propose a new analysis method for Second Order Differential Power Analysis (SO-DPA) by choosing optimized distinguish function of Differential Power Analysis (DPA). With this method, we need not to pre-process the power trace as traditional methods. Therefore, the amount of computation is reduced and the attack efficiency is improved, compared to methods proposed by previous researchers. The experiment results show that we have successfully attacked the masked AES circuit and recover the secret key through the developed platform based synopsis software.**

Index Terms—**HO-DPA, SO-DPA, DPA, Distinguish function**

I. INTRODUCTION

With the rapid development of information technology, information security issues become more and more important. American Institute of Standard and Technology (AIST) proposed the advanced encryption standard (AES) to solve that problem. However, side channel analysis technology, especially power attack poses a serious threat to cryptographic chip and power analysis becomes the most concern for the safety chip designers in recent years. The wildly used method against first-order power analysis is Boolean masking [1-3].

However, by combining power consumptions leaked at multiple times, such protected devices using the masking technology can be defeated by Higher-Order attacks as originally described in [4]. Currently, there are two kinds of higher order attacks, including Higher-Order DPA and Mutual Information Analysis (MIA) [5-6].Various combining function are applied to attack the masked AES implementation, such as absolute difference combining function, product combining function and sum combining function, but finding an appropriate combining function is still a serious issue. T. Messerges et al, used the absolute difference between two leakage samples as a combining function based on the assumption of Hamming weight (HW) model [4], but the complexity is 2^{128} and attacker need to know the exact time of sensitive intermediate variable. Waddle, et al, proposed the so-called zero-offset 2DPA and FFT 2DPA [7], however this method need to select more power trace with the increase of noise. Oswald, et al, proposed the strategy for SO-DPA that pre-process power trace at first and then make a DPA attack on the pre-processed power trace [8], but the value of the correlation coefficient is too small. E. Prouff, et al, conducted an in-depth analysis of HO-DPA and proposed a normalization step that improved product method to the best efficiency [9], however it is inevitable to pre-process the power trace and know the exact time of sensitive intermediate variable. Based on the problem above, we use DPA method by choosing optimize distinguish function, which need not pre-process power trace.

The rest of this paper is organized as follows. In Sections Ⅱ, we introduce the basic principles and survey related work of the SO-DPA. We explain our concept of SO-DPA and the process of power acquisition and data processing in Sections Ⅲ. We show the results of our new attacks on a masked AES implementation in Section Ⅳ and conclude this article in Section Ⅴ.

II. THE PRINCIPLE OF SO-DPA

The main idea of SO-DPA is that combine statistical properties of the power consumption at multiple sample times to eliminate the random of power consumption. The method of this power analysis is successes in attacking masked AES circuit, which is used to against DPA. Generally, there are two processes for SO-DPA, first step preprocess the power traces as figure1, second step attack on the pre-processed power traces [10-11]. In the figure1, *m* denotes the masked random variable and *a* denotes the intermediate variable.

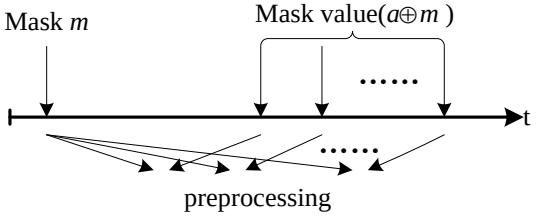

Figure 1 the process of power preprocessing

For example, if the adversary wants to attack the first round of the masked AES encryption system, he may combine the

Manuscript received July 2, 2016, revised July 16, 2016. This work was supported by the National Natural Science Foundation of China(No. 61376025, No. 61106018),the Industry-academic Joint Technological Innovations Fund Project of Jiangsu (BY2013003-11),the Funding of Jiangsu Innovation Program for Graduate Education(No. KYLX_0273).

W. Wang is with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing , 210016, China.(email:wwynuaa@163.com).

N. WU is with the College of Electronic and Information Engineering, NUAA, Nanjing, 210016, China (e-mail: wunee@nuaa.edu.cn).

J. Zhang is with the College of Electronic and Information Engineering, NUAA, Nanjing, 210016,China(e-mail:376106072@qq.com).

F. Zhou is with the College of Electronic and Information Engineering, NUAA, Nanjing, 210016,China(e-mail:zfnuaa@nuaa.edu.cn).

leakage sample L_{SM} of the masked S-box output $S(p \oplus k) \oplus m_1$ at time t_1 and the leakage sample L_M of the mask m1 at time t_2 to approximately construct the leakage of $S(p \oplus k)$ using a combining function $C(L_{SM}, L_M)$. With the assumption that the leakages at all relevant time instants follow the Hamming weight model, the adversary can mount a DPA by evaluating the Pearson's correlation coefficient ρ between the combined result $C(L_{SM}, L_M)$ and the Hamming weight of $S(p \oplus k)$ for all hypotheses k as following:

$$\rho\left(C\left(L_{\mathrm{SM}}, L_{\mathrm{M}}\right), HW\left(S\left(p \oplus k\right)\right)\right) \quad (1)$$

If adversary guesses the right key, the peak will appear in the curve of correlation coefficient. However, there is no peak in the curve of correlation coefficient when guess the wrong key for the Hamming weight value of the predictive sensitive intermediate variable is no relevant to the pre-processed power.

The absolute difference method is primarily introduced based on the relation in (2), S denotes the sensitive intermediate variable and M denotes the random variable, where S, M, $S \oplus M$ are all single bit wide[4].

$$HW(\mathrm{S}) = \left|HW\left(S \oplus M\right) - HW(M)\right| \quad (2)$$

However, an input of S-box is 8 bits wide, so the other 7 bits will be taken as Algorithm noise when still using (2). That is to say, the equality of (2) no longer holds, if adversary needs to attack one byte of the S-box output. The relation between correlation coefficient ρ and the bit wide of variable is shown in figure 2.

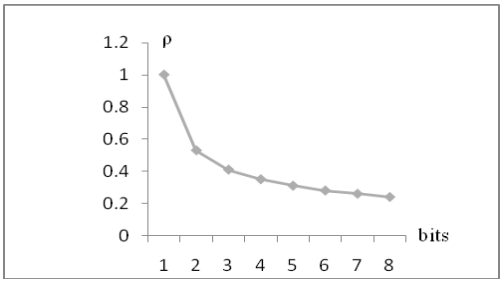

Figure 2 the relation between ρ and bits

In order to get the optimal correlation coefficient, we research the effect for different values of β (denoted as the power of a function) and for attacks on different numbers of bits based on absolute difference function. We calculate the correlation between HW(S) and $|\mathrm{HW(S} \oplus \mathrm{M)\text{-}HW(M)}|^{\beta}$ proposed by Oswald[8]. We calculate the correlation coefficient ρ for different value of β and for attacks on different numbers of bits. The result is shown in TableI.

Table I the values of ρ for different β and bits

Bits \ β	1	2	3	4	5	6
1	0.0871	0.0987	0.0970	0.0879	0.0785	0.0675
2	0.1129	0.1325	0.1285	0.1186	0.1082	0.0974
3	0.1415	0.1662	0.1614	0.1485	0.1361	0.1213
4	0.1743	0.1924	0.1836	0.1684	0.1498	0.1328
5	0.1956	0.2107	0.2013	0.1828	0.1633	0.1455
6	0.2094	0.2296	0.2188	0.1989	0.1769	0.1564
7	0.2288	0.2416	0.2345	0.2135	0.1897	0.1665
8	0.2407	0.2624	0.2511	0.2278	0.2031	0.1785

These questions arise as below based on the above analysis:

1) The variable of equation (2) holds only single bit.
2) We need know the exact time of sensitive intermediate variable executing.
3) The value of correlation coefficient is so small that it is not easy to distinguish from the wrong guessed key.

According to the questions above, we proposed an improved different power analysis method by choosing optimal distinguish function. We guess every possible key and calculate the value of distinguish function, which can be used to divide power traces into two groups. Therefore we can plot the differential power curve and there are some peaks in the curve when the guessed key is right. We detail the process of choosing distinguish function in the next section.

III. IMPROVED METHOD FOR SO-DPA

The object that HO-DPA attacked is masked AES circuits, which make power consumption randomization by add some random variable in algorithm. We will detail how to choose the optimal distinguish function and the process of power analysis as below.

A. The option of distinguish function

The masked AES circuit that we adopt is shown in figure 3 and the secret key of AES that we researched is 128 bits. In order to keep the same masked random for every round, we add the masked random at the input and output of every AES round. The symbols of figure 3 are introduced as follows: M denotes plaintext, K_0 denotes the original key, and K_1 denotes the expanded key for the first round. X_0, X_1, X_2, X_3, X_4, and X_5 denote masked random variable. ARK denotes the exclusive or operation in the AES algorithm. SB denotes the transformation in the AES algorithm that processes the state using a nonlinear byte substitution. NSB denotes inverse transformation in SB , AF denotes affine operation in SB. SR denotes the transformation in the AES algorithm that process the state by cyclically shifting the last three rows of state. MC denotes the transformation that takes all of columns of the state and mixes their data to produce new columns.

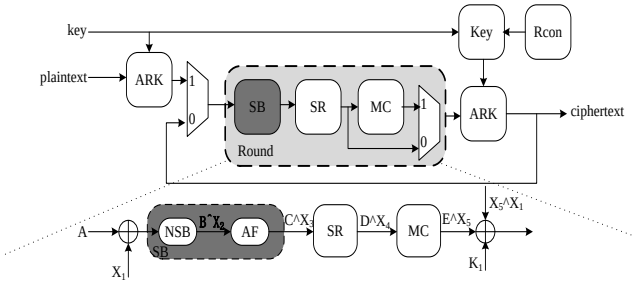

Figure 3 the masked AES circuit

A denotes the output of ARK, B denotes the output of NSB, C denotes the output of AF, D denote the output of SR, E denotes the output of MC, and F denotes the output of AES first round without calculating the masked random.

The relationship is as follows:

$$A = M \wedge K_0;$$

$$I_1 = NSB(A \wedge X_1) = B \wedge X_2, \quad B = NSB(A);$$

$$I_2 = AF\left(I_1\right) = C \wedge X_3, \quad C = AF(B), \quad X_3 = AF(X_2);$$

ISBN: 978-988-14047-1-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$I_3 = SR(I_2) = D \wedge X_4$, $D = SR(C)$, $X_4 = SR(X_3)$;
$I_4 = MC(I_3) = E \wedge X_5$, $E = MC(C)$, $X_5 = MC(X_3)$;
$F = E \wedge K_1$, $(A \wedge X_1) \wedge (F \wedge X_1) = A \wedge E \wedge K_1 = A \wedge F$;

It is easy to find that the input and output of first round ARK operation own the same mask variable, then we choose the sampling point shown in figure 3. We select the optimal distinguish function as equation (3). S denotes the same operation as SB among equation (3), $Z_{0,0}$ denotes one byte of state, $K0_{0,0}$,$K0_{1,1}$,$K0_{2,2}$,$K0_{3,3}$,$K1_{0,0}$ denote involved key and $A_{0,0}$,$A_{1,1}$,$A_{2,2}$,$A_{3,3}$ denote one byte of A. We selected $K0_{0,0}$ as target and the term contained other key are as constant, expressed α. According to one bit value of $Z_{0,0}$, we divide the power trace into two groups and make statistical analysis of power data when guess every possible $K0_{0,0}$ to recover right value of $K0_{0,0}$. Based on the ideas of DPA, we can get differential power wave, which will appear some peak at some times, so that we can easily distinguish the peak and recover the secret key.

$$\begin{aligned} Z_{0,0} &= (A_{0,0} \wedge X_{0,0}) \wedge (E_{0,0} \wedge X_{0,0}) \\ &= \{02\} S(A_{0,0}) \wedge \{03\} S(A_{1,1}) \wedge S(A_{2,2}) \wedge S(A_{3,3}) \wedge K_{1,0} \\ &= \{02\} S(A_{0,0}) \wedge \alpha \end{aligned} \quad (3)$$

The detailed process of power analysis contains two parts: power acquisition module and power data analysis module. In this paper, the power acquisition module adopts the EDA tool to simulate the operation of the Cryptographic chip to gain the power data. Based on the predecessors' research work and resources of lab, we acquire power data through the set of Synopsys IC tool and make the statistical analysis of power data using MATLAB tools. We will introduce the process of power analysis as follows:

B. The platform of power acquisition

The flow of power acquisition show in figure 3, which is consist of five steps.

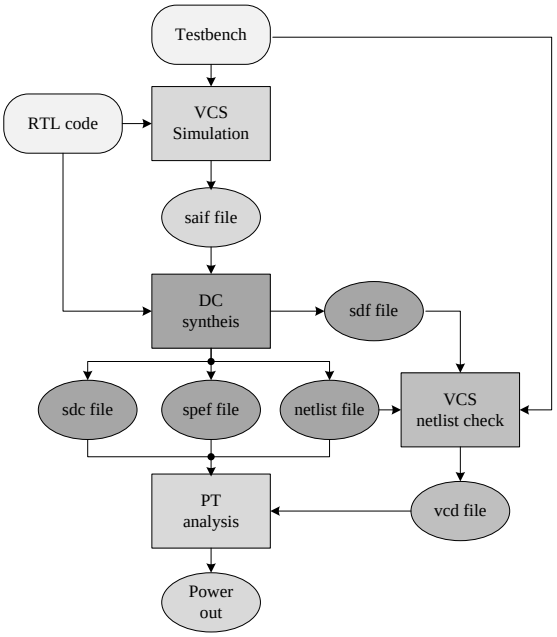

Figure 4 The flow of power acquisition

First step: Write code using Verilog at RTL level to implement the AES algorithm.

Second step: Verify the function of the RTL level code. We write a Testbench as the input of AES and make Verilog compiled simulator (VCS) to verify the RTL level code.

Third step: Design compiler (DC) the Verilog code to generate the netlist file. Design compiler read the saif file which generate from the above step, generate a set report beside netlist file, such as sdf file, sdc file, spef file.

Forth step: Verify the function of netlist file. Verilog compiled simulator (VCS) adopt same testbench as second step and sdf file to simulate the netlist file for verifying the function of netlist file. At the same time, it generate vcd file at this testbench when the netlist file is right.

Final step: Generate power report. We use Prime Time (PT) tool to generate power report based sdc file, spef file and vcd file and save the report in filename called Power.out.

C. The process of power analysis

We analyze the power data through MATLAB tool and the detailed flows of power data analysis as follows: show in figure 5.

We divide the power traces into two groups based on the equation (4), the number of power traces belong to every group, respectively expressed as A0 and A1. P_j denotes the power trace and Z(.) denotes the optimal distinguish function.

$$A0 = \{P_j \mid Z(\cdot) = 0\}, \quad A1 = \{P_j \mid Z(\cdot) = 1\} \quad (4)$$

Obviously, $|A0| + |A1| = N$, N denotes the total number of power trace. We execute the operation as equation (5) to both A0 and A1 and generate the average value of two groups power trace, respectively expressed as S0 and S1.

$$S0 = \frac{1}{|A0|} \sum_{P_j \in A0} P_j, \quad S1 = \frac{1}{|A1|} \sum_{P_j \in A1} P_j \quad (5)$$

We execute the operation as equation (6) and generate differential power curve when the guessed key is k_j, expressed as T_j.

$$T_j = S0 - S1 \quad (6)$$

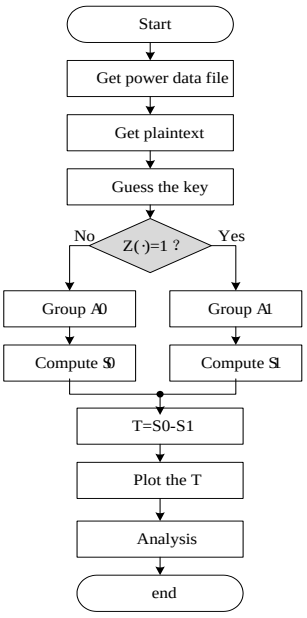

Figure 5 the flow of power data analysis

We can get 256 differential power curves under same operation when guess the every possible $K0_{0,0}$, respectively

expressed as $T=(T_1, T_2, ..., T_{256})$. Compare the value of differential power curve at the time of sensitive intermediate variable and find corresponding differential power curve own max value at that time. Obviously, the guess key of that curve is right key which we want to recover.

IV. THE RESULT OF EXPERIMENT

We have made experiments on the developed power analysis software platform, targeting the high 8 bit of original key, which realistic value of that byte key is 8'h3c. The result of experiment is demonstrated in figure 6 and the peak of differential power curve is corresponding to guessed key with value 8'h3c. Therefore, we successfully recovered the right key through the software platform of power analysis and achieved some conclusion as follows:

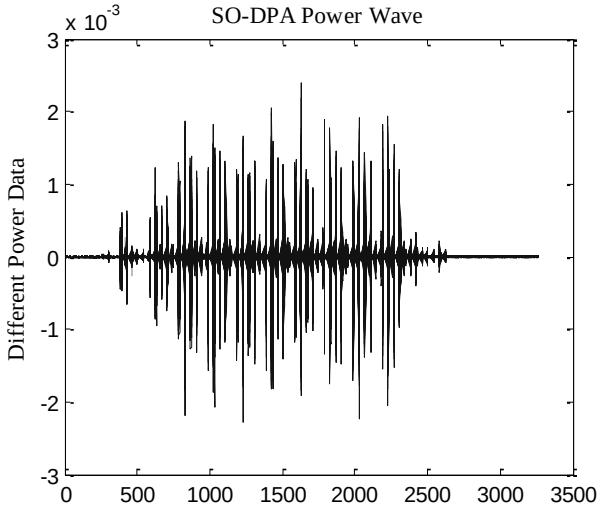

Figure 6 the result of experiment

1) We guess every possible value for one byte key and get the 256 DPA power waves through MATLAB tools shown in figure 6.
2) Obviously, there are ten rounds of transformation seen from the figure 6 and there are a peak in the SO-DPA power waves at the time when the output of first round.
3) We can only need no more than 256 power traces to recover one byte secret key, which obviously improved attack efficiency.

V. CONCLUSION

According to the above process of power analysis, software simulation own much advantages, such as low cost, low noise easy operation and so on. We sum up the conclusions as follows:

1) We need not to preprocess the power traces that reduce the amount of computation through the method proposed by us compared to [8], [12-13].

2) We can recover one byte secret key one time using the method proposed by us, which reduce attack complexity, comparing to [4].

3) The result of our experiment is easier to distinguish the right key than the method using correlation coefficient in [4], [7-8].

4) Besides, we only need less than 256 power traces to recover one byte key when compared to traditional methods, such as [5-8]. Therefore, it improved attack efficiency by the method we proposed. However, we only attacked the masked AES circuit with software simulation and the experiment isn't confirmed by physical equipment, so, we need to do further research in the future.

REFERENCES

[1] M.Akkar, C.Giraud. "An Implementation of DES and AES, Secure against Some Attack"[C], *Cryptology Hardware and Embedded Systems- CHES 2001*, 2001, pp.309-318.
[2] S. Chari, C. S. Jutla, J. R. Rao, and P. Rohatgi, "Towards Sound Approaches to Counteract Power-Analysis Attacks", *Advances in Cryptology-CRYPTO'99*, Spring Berlin Heidelberg, 1999,pp.398-412.
[3] K. Schramm and C. Paar, "Higher Order Masking of the AES" [C], *Topics in Cryptology-CT-RSA 2006*,Springer Berlin Heidelberg, 2006, pp.208-225.
[4] Thomas S. Messerges, "Using Second-Order Power Analysis to Attack DPA Resistant Software" [C], *Cryptology Hardware and Embedded Systems-CHES 2000*, Spinger Berlin Heidelberg,2011,pp.238-251.
[5] B. Gierlichs, L. Batina, P. Tuyls, and B. Preneel, "Mutual Information Analysis: A Generic Side-Channel Distinguisher" [C], *Cryptology Hardware and Embedded Systems-CHESS 2008*, Spring Berlin Heidelberg, 2008, pp.426-442.
[6] L. Batina, B. Gierlichs, E. Prouff, and M. Rivain, "Mutual Information Analysis: a Comprehensive study", *Journal of Cryptology*, Vol 24, No. 2, 2011, pp.269-291.
[7] Jason Waddle and David Wagner, "Towards Efficient Second-Order Power Analysis" [C], *Cryptology Hardware and Embedded Systems-CHES 2004*, Springer Berlin Heidelberg, 2004, pp.1-15.
[8] E. Oswald, S. Mangard, C. Herbst, and S. Tillich, "Practical Second-Order DPA Attacks for Masked Smart Card Implementations of Block Ciphers"[C],*Topics in Cryptology-CT-RSA 2006*,Springer Berlin Heidelberg, 2006, pp.192-207.
[9] E. Prouff, M. Rivain, and R. Bevan, "Statistical Analysis of Second Order Differential Power Analysis"[J],IEEE *Transaction on Computers*, Vol 58, No.6, 2009, pp.799-811.
[10] FX Standaert, N Veyrat-Charvillon, E Oswald and B Gierlichs, ''The World Is Not Enough: Another Look on Second-Order DPA''[C], *Advances in Cryptology-ASAIC-RYPT*, Spring Berlin Heidelberg, 2010,pp.112-129,
[11] Emmanuel Prouff, Thomas Roche, "Attack on Higher-Order Masking of the AES Based on Homographic Functions" [C], *Progress in Cryptology-INDOCCRYPT*,2010,pp.262-281.
[12] Miao Yuan, Bai Guoqiang, "Improving Second-Order DPA Attacks with New Modeled Power Leakages" [C], *International Conference on Computational Intelligence and Security*, IEEE, 2015,pp.394-397.
[13] F Durvaux, FX Standaert, N Veryat-Charvillon, JB Mairy, "Efficient Selection of Time Samples for Higher-Order DPA with Projection Pursuits", *Constructive Side-Channel Analysis and Secure Design*, Springer International Publishing,2015,pp.34-50.

ISBN: 978-988-14047-1-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)