Security Analysis of the Masking-Shuffling based Side Channel Attack Countermeasures

Jong-Won_Cho and Dong-Guk Han

Korea System Assurance Inc., Kookmin University
mr_jo@kosyas.com, christa@kookmin.ac.kr

Abstract

Side Channel Attacks are known to be effective in cracking secret keys utilized in smart cards, electronic passports, and e-ID cards. A combination of masking and shuffling methods has been proposed as a practical countermeasure to such attacks. Using a template attack (TA), S. Tillich recently analyzed an AES using masking and shuffling techniques with a biased-mask attack technique. To apply this method, however, we need to collect the template information on the masking value in advance. Moreover, this method requires knowing the exact time position of the target masking value for a higher probability of success. In this paper, we suggest a new practical method called a Biasing Power Analysis (BPA) to find a secret key used in an AES based on a masking-shuffling method without the use of the time position and template information of the masking value. We conducted an experiment on a BPA attack against a 128-bit AES secret key based on a masking-shuffling method operating on an MSP430 chip and succeeded in finding the entire secret key. The results of this study can be utilized for next-generation ID cards to verify their physical safety.

Keywords: *Side channel Attacks, masking, shuffling, Biasing Power Analysis, AES*

1. Introduction

Side Channel Attacks are known to be effective tools for attacking secret keys by using the calculation time, power consumption, and electromagnetic waves. The known side channel attack method is related to an analysis of the power consumption and has been developed for a simple power analysis (SPA), differential power analysis (DPA), and correlation power analysis (CPA). It is the one of the strongest methods for analyzing side channels [3, 6, 7, 8].

Several countermeasures have been studied to improve the stability of devices with weaknesses regarding a side channel analysis. Masking and shuffling techniques are widely used practical countermeasures. The masking technique uses random masking values to hide information that an attacker may try to analyze. Although the masking technique defends the information from the first-order CPA, security cannot be guaranteed as it cannot protect information from the second-order CPA. In this case, using a shuffling method in which the calculation sequence is changed, we can increase the complexity of a side channel analysis and thereby defend the information from the second-order CPA.

S. Tillich recently suggested the use of a biased-mask attack technique, which uses a template attack (TA) as an analysis method of the AES [4], in which the masking technique and shuffling technique are both used [9]. The biased-mask attack method is an attack technique against the countermeasure to which the masking is applied. It helps to analyze the secret key by the first-order CPA, rather than the second-order CPA. In addition, to attack an algorithm in which a shuffling prevention technique is applied, they used a windowing

technique proposed by C. Clavier and attacked an algorithm by applying both masking and shuffling techniques [1, 2, 9].

However, the biased-mask technique suggested by S. Tillich requires template information on the masking value and its time position, which is difficult to estimate. If the above information is not known, the analysis will not be sufficiently correct. This means that we require the time position of the masking value, and that we must collect the template information for a correct analysis.

In this paper, we propose new biasing-power analysis (BPA), which can crack an AES applied with a masking-shuffling technique without having the information regarding the time position of the masking value and the template information. This method enables us to analyze the information using previously collected power traces only, without the need for the template information of the masking value.

This paper is organized as follows. Section 2 describes the suggested BPA technique and the bias classification method necessary for the use of the BPA technique. Section 3 shows the performance of the BPA analysis and its availability based on a power analysis experiment conducted on a practical MSP430 software board. Finally, Section 4 shows the advantages and disadvantages of the suggested method based on the experimental results.

2. Biased-power Analysis

The biased-power analysis (BPA) described in this paper, which is a power-analysis method, uses the phenomenon in which the collected power wave forms are biased at the same time zone and classified into power values such as {0,1,…,8}, which are the same as the Hamming weight values for a relative power analysis. According to the classification method adopted, the power value of the collected power wave form for each time can be grouped into 9 Hamming weights ranging from 0 to 8. According to this method, even though we do not have the template information appearing at the correct masking time, we can classify the masking values. If we know the power position of the masking, we can analyze the Hemming weight at the point of the masking power section using the classification method, and we therefore do not need to obtain separate template information. In this section, we describe the biased-power analysis at the masking position. Prior to our explanation of these two methods, we also describe the necessary classification method.

2.1. Classification Method using the Power Bias

In this section, we describe the classification function on the Hemming weight in detail. The distribution of power values at the point of analysis, as shown in Figure 1. The classification method shall therefore be designed in such a way that the power value and Hamming weight can be properly classified.

We suggest the use of a classification method in which the power value is classified into N wave forms and the maximum and minimum values are obtained from the power values of the analysis point. The maximum value and minimum values are then divided into 9 same sizes and a number of 0 to 8 is assigned, with the bigger number assigned to a larger power. To classify a value of 0 to 8 at the time of the analysis point, we divide the fluctuation zone into maximum to minimum values based on the 9 classes shown in Figure 2. Using this range, we then classify the power value from 0 to 8, with the largest value being 8 and the least being 0.

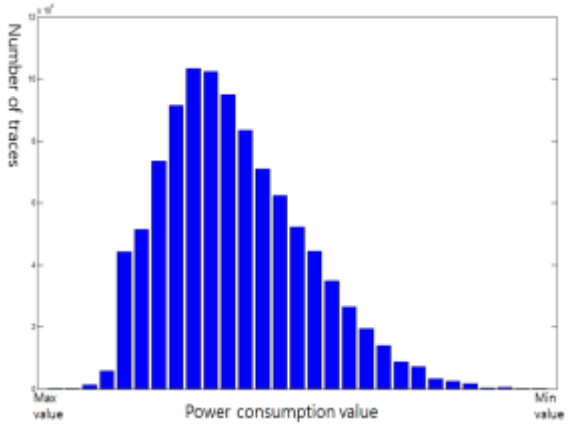

Figure 1. Power Distribution of the Classified Points

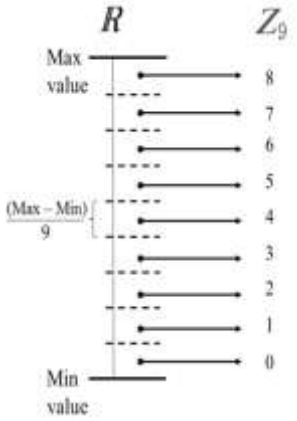

Figure 2. Classification Method used for a Uniform Division of Values

2.2. Biased-power Analysis at the Masking Position

We propose a biased-power analysis (BPA) at the masking position. It analyzes the power in such a way that it limits the masking values used. If the HW(*m*) of the masking value *m* is 8 or 0 only, the analysis is possible by XORing "OxFF" to the classification function to be calculated or leaving it as is. HW(*m*) is 0, the value of (*s* XOR *m*) becomes (*s* XOR 0x00) = *s*, where the out value of *s* is calculated using the masking value of *m*. Here, HW(*x*) means the Hamming weight of the value *x*. If HW(*m*) is 8, *m* becomes ' 0xFF' , making it similar to (*s* XOR 0x00). Therefore, we can determine the median value used by the masking. However, this can only be obtained when we know the Hamming weight of the masking value. Actually, what we can know from the collected information is simply C(*m*), or the power value of masking value *m*. We can therefore limit the masking value to the required Hamming weight only when we can convert C(*m*) to the value having the same phase as the Hamming weight according to the classification method, as described in Figure 2. As a result, if we know the time position of the masking value of *m*, we can analyze it by simply selecting the specific masking value according to the classification method.

We can select the wave form required for the analysis by applying the method used for converting the power value into the Hamming weight to the masking value. The selection of a wave form of 0 or 8 Hamming weight of *m* is fine in terms of the analysis efficiency. However, in such a case, as the number of classified wave forms is reduced to 1/128 of the number of collected wave forms, an analysis failure is more likely to occur. Accordingly, to solve the problem caused by the number of wave forms, we can use an HW(*m*) of 1 or 7 rather than 0 or 8. This analysis is possible since an HW(*m*)of 0 or 1 is related to an HW(*m*) of 0, while an HW(*m*) of 7 or 8 is related to an HW(*m*) of 8.

The independent use of two types of wave forms selected through a classification method may be considered a waste of data, as an analysis is possible using two kinds of wave forms at the same time. Of course, an HW(m) value of 8 indicates that the masking value of m is "0xFF", which has the same effect as when m is removed by calculating XOR of "0xFF" to the classification function. Accordingly, when we analyze the wave form related to an HW(m) of 8, we can XOR calculate "0xFF" to the classification function value. Otherwise, we can conduct an analysis without any additional calculations, and can instead use two kinds of wave forms.

3. Experiment Environment and Analysis Results

3.1. Experiment Environment

For the BPA experiment proposed in this paper, we used an MSP430 chip board. An AES was constructed using the masking-shuffling technique suggested by C. Herbst [2]. The reconstructed AES collected 1,000,000 wave forms of the AES algorithm at an oscilloscope sampling rate of 250MS/s.

Figure 3 shows the collected wave forms which are compressed, particularly the section where 16 1-round Sbox calculations were conducted. The narrow peak section at the front side is the power of six masking values called by the experiment.

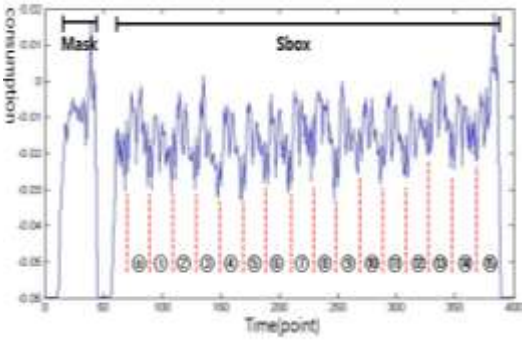

Figure 3. The first round of Sbox of AES with Masking Time

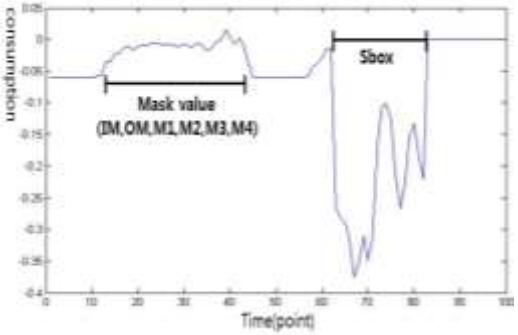

Figure 4. AES, traces to applied Windowing Method

The wave form analyzed in this paper was applied using the masking and shuffling techniques, and therefore, the windowing technique as suggested by C. Clavier was also applied for the analysis [1].

Figure 4 shows the wave form made by combining 16 Sbox calculations into one using the windowing technique. As 16 Sbox calculations are combined into one, the peak formed down the back side is used in the Sbox calculation.

In our experiment, we analyzed the first through fourth Sbox calculations and used four values of m0, m1, m2, and m3 as masking values calculated using the Sbox output value.

To compare the newly suggested BPA analysis method with the other analysis methods after obtaining the BPA experiment results, we compare the results using the second-order CPA analysis, which is usually used for a general analysis [5]. For the second-order CPA analysis, we use the time position of the masking value when the differential method is used, and the time position of the calculation when the pressure and windowing techniques are used. In the second-order CPA analysis, we use point 1 as the time position of the masking value, which has the highest correlation, as in the BPA.

3.2. Experimental Results

We used 1,000,000 wave forms for the analysis. The BPA at the masking position used 110,000 wave forms, or 1/10 of this total number.

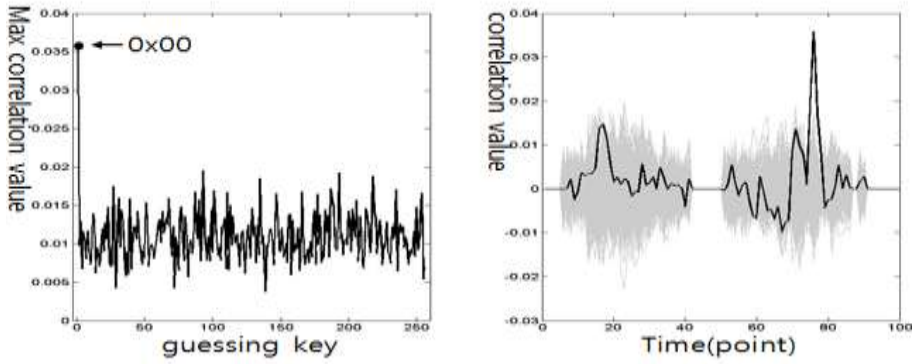

Figure 5. BPA with a mask, the first Sbox, the maximum correlation value for the guessing keys (left), and the correlation trace of the guessing keys (right)

Figure 5 shows the results of the BPA analysis at the masking position. The left-side of the figure shows the maximum correlation coefficient of the 256 guessing keys, while the black lines on the right-side are the CPA analysis wave form for the first key of "0x00", and the grey lines are the CPA analysis wave form for the other guessing 255 wrong keys.

The HW of the masking value was limited to assumed values of 0 and 8, and the classified function is the output value of Sbox. Therefore, we can see the peak formed at the Sbox position, as shown in Figure 5.

3.3 Comparison between the Second CPA Analysis and BPA Analysis

A comparison between the second-order CPA analysis and the BPA analysis was conducted based on the results of 1,000,000 wave forms. As shown in the shaded part of Figure 1, the general second-order CPA analysis found only four key bytes, while the BPA analysis found all 16 keys. Although one result can reveal which analysis is better, we used the full experiment to show that the BPA has the better performance over the second-order CPA power analysis.

Table 1. Comparison of the Second-order CPA with the BPA

Key Bytes	Analysis method		Key Bytes	Analysis method	
	2nd CPA	BPA		2nd CPA	BPA
1st 0x00	0x84	0x00	9th 0x08	0x08	0x08
2nd 0x01	0x23	0x01	10th 0x09	0xEA	0x09
3rd 0x02	0xDC	0x02	11th x0A	0x6C	0x0A
4th 0x03	0xF3	0x03	12th 0x0B	0x0E	0x0B
5th 0x04	0x04	0x04	13th 0x0C	0x0C	0x0C
6th 0x05	0xC6	0x05	14th 0x0D	0x0D	0x0D
7th 0x06	0x32	0x06	15th 0x0E	0x34	0x0E
8th 0x07	0xF7	0x07	16th 0x0F	0x64	0x0F

4. Conclusion

We suggested the use of a BPA analysis method in which the measurements based on the masking-shuffling technique are taken against a side channel attack without the use of a second-order CPA. The conventional method by S. Tillich needs both the template information and time information on the masking values. However, the BPA analysis suggested in this paper can be used when we do not have one or both pieces of information. To verify the analysis method, we conducted an experiment to find the secret key on the AES-128 algorithm using the side channel information obtained from an MSP 430 chip, and succeeded by conducting a million power wave form experiments. However, a MCU chip such as an MSP430 detects the side channel information with less noise, and has an environment where there is no hardware related measure. The BPA is possible using a million wave forms. However, when an IC chip card is used, which is typically the case, the self-hardware action is driven by default and an attack will be less likely to succeed when the analysis is conducted using the suggested BPA. Accordingly, future research will be conducted to develop an enhanced signal pre-treatment, classification criteria, and differential technology allowing a BPA analysis to be conducted even in actual IC chip environment where various kinds of hardware noise exist.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(2012-0007285).

.

References

[1] C. Clavier, J. -S. Coron and N. Dabbous, “Differential Power Analysis in the Presence of Hardware Countermeasures”, CHES 2000, **(2000)**, LNCS 1965, Springer-Verlag, pp. 252–263.

[2] C. Herbst, E. Oswald and S. Mangard, “An AES Smart Card Implementation Resistant to Power Analysis Attacks”, ACNS 2006, **(2006)**, LNCS 3989, Springer-Verlag, pp. 239-252.

[3] E. Brier, C. Clavier and F. Olivier, “Correlation power analysis whit a leakage model”, CHES 2004, **(2004)**, LNCS 3156, Springer-Verlag, pp. 16-29.

[4] E. Oswald and S. Mangard, “Template Attacks on Masking—Resistance is Futile”, CT-RSA 2007, **(2007)**, LNCS 4377, Springer-Verlag, pp. 243–256.

[5] E. Oswald, S. Mangard, C. Herbst, and S. Tillich, “Practical Second-Order DPA Attacks for Masked Smart Card Implementations of Block Ciphers”, CT-RSA 2006, **(2006)**, LNCS 3860, Springer-Verlag, pp. 192–207.

[6] P. Kocher, “Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems”, CRYPTO 1996, **(1996)**, LNCS1109, Springer-Verlag, pp. 104-113.

[7] P. Kocher, J. Jaffe, and B. Jun, “Differential power analysis”, CRYPTO 1999, **(1999)**, LNCS 1666, Springer-Verlag, pp. 388-397.

[8] P. Kocher, J. Jaffe and B. Jun, “Introduction to differential power analysis and related attack”, Cryptography Research, White Paper, **(1998)**.

[9] S. Tillich, C. Herbst and S. Mangard, “Protecting AES Software Implementations on 32-Bit Processors Against Power Analysis”, Computer Science, **(2007)**, LNCS 4521, Springer-Verlag, pp. 141-157.

Authors

Jong-Won Cho

He received his B.S. degree in mathematics from Kookmin University in 2010, and his M.S. degrees in mathematics from Kookmin University in 2012, respectively. He is currently working as a researcher with the Korea System Assurance, Inc., Seoul, Korea. He is interested in Side Channel Analysis.

Dong-Guk Han

He received his B.S. degree in mathematics from Korea University in 1999, and his M.S. degrees in mathematics from Korea University in 2002, respectively. He received Ph.D. of engineering in Information Security from Korea University in 2005. He was a Post.Doc. in Future University-Hakodate, Japan. After finishing the doctor course, he had been an exchange student in Dep. of Computer Science and Communication Engineering in Kyushu University in Japan from Apr. 2004 to Mar. 2005. He was a senior researcher in Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. He is currently working as an associate professor with the Department of Mathematics of Kookmin University, Seoul, Korea. He is a member of KIISC, IEEK, and IACR.