An implementation of SM4 algorithm based on mask

LI Xinchao
QuanZhou Coast Guard . Fujian, Quanzhou, 362000

Abstract: Aiming at the problem of SM4 algorithm resisting power attack, this paper studies the implementation structure and S-box characteristics of SM4 algorithm, constructs the S-box realized by multiplication mask, and designs a semi random mask circuit implementation scheme of SM4 algorithm. In this scheme, the middle values of the first six rounds and the last six rounds of SM4 algorithm are completely random masked, and the intermediate values of the middle 24 rounds are fixed mask, which reduces the consumption of random masks. After security analysis, the scheme can resist the first-order power consumption attack, and has less random number consumption and lower implementation area.

1. Introduction
In 2006, in order to ensure the data security of WLAN products in WAPI Standard, the state commercial password management office issued the SM4 algorithm standard [1]. As the standard of commercial cipher algorithm in China, the research on the security of SM4 algorithm is of great significance to the development of cryptographic algorithm in China.

In 2007, Liu f et al. Analyzed and studied SM4 algorithm, and provided the S-box implementation structure and algebraic expression of SM4, which laid the foundation for the follow-up research of SM4 algorithm [2]. In 2008, Bai X and others successfully carried out differential power attack on SM4 algorithm, and the implementation security of SM4 algorithm faced a major threat [3]. After that, various schemes against power analysis attack of SM4 algorithm were proposed one after another, among which mask defense scheme was commonly used. Referring to the fixed value mask scheme in AES algorithm, X Duan et al. Proposed a fixed value mask scheme for the secure implementation of SM4 algorithm [4]. Hao et al. Through the analysis and study of AES algorithm S-box, the normal basis inversion method in AES algorithm is introduced into SM4 algorithm, and a composite domain mask implementation scheme of SM4 algorithm S-box is proposed [5]. Pei Chao and others proposed a new random mask scheme based on S-box table look-up method and multi-path multiplication mask to conduct random linear transformation on S-box, which achieved the purpose of resisting DPA attack [6].

In this paper, according to the structure characteristics of SM4 algorithm S-box, the multiplication mask S-box of SM4 algorithm is designed and implemented by using multiplication mask method. Combined with the implementation structure of SM4, the semi-random mask circuit scheme of SM4 algorithm is designed and implemented. In this scheme, the intermediate values of the first six rounds and the last six rounds are completely random masked, and the intermediate values of the middle 24 rounds are fixed mask, which reduces the number of random mask The number of machine masks consumed. After security analysis, the scheme can resist the first-order power attack, and has less random number and lower implementation area.
2. Preliminaries

2.1 SMS4 block cipher
In 2006, China's national Password Administration announced the algorithm standard for WLAN products SM4 cipher algorithm, which is the first commercial cipher algorithm published in China. The packet length and key length of SM4 algorithm are 128 bits, and the unbalanced Feistel structure is used to carry out 32 rounds of iterative operation, which has high complexity and security [7].

![SM4 algorithm implementation process](image)

Fig. 1. SM4 algorithm implementation process

The encryption process of SM4 is shown in Figure 1. According to the implementation process of S-box obtained by Liu f et al., the specific implementation framework is as follows:

![S-box structure of SM4 algorithm](image)

Fig. 2. S-box structure of SM4 algorithm

2.2 DPA
Differential power analysis (DPA) is a common power attack method, which has the characteristics of high efficiency and simple implementation. The differential power consumption attack aims at the cryptographic scheme, but measures the power consumption information generated by the cryptographic device when processing the data in the specific encryption and decryption process. By
using the mathematical statistical analysis method, the data is determined according to the power consumption change of the cryptographic device when processing different data, and repeated guessing is carried out. Finally, the encryption used in the algorithm encryption and decryption is obtained Key [2]. Differential power attack only needs to know the encryption and decryption algorithm, and does not need to know the details of the cryptographic device to recover the key.

Differential power attack includes the following four steps:

a) select a certain intermediate value \( D \) produced by cryptographic equipment in the process of encryption and decryption, and construct a differential function \( D(m, k) \), in which, is a small part of the key \( k \), the non constant data that can be measured in the process of encryption and decryption is \( m \);

b) a large number of different data are injected into the cryptographic device to obtain \( m_i \) and the trajectory curve \( t_i \) generated in the process of encryption and decryption;

c) assuming the possible sub key used in encryption and decryption is \( k_i \), the intermediate value \( D \) should be generated by \( k_i \) and the known calculation \( m_i \), the selected model is used to filter and classify \( t_i \), and the average value \( \bar{t} \) of all curves and the value of classified curve are calculated respectively, and the difference operation is performed finally, the formula is as follows

\[
\Delta_{ij}(j) = \frac{\sum_{m} D(m_i, k_i) \cdot f_i(j)}{\sum_{m} D(m_i, k_i)} - \frac{\sum_{m} (1 - D(m_i, k_i)) \cdot f_i(j)}{\sum_{m} (1 - D(m_i, k_i))}
\]

(1)

According to the result of the difference operation, whether there is an independent and obvious peak can be used to judge whether the assumed sub key is correct. If there is an independent obvious peak, the sub key is assumed to be correct and proceed to the next step. Otherwise, repeat step c).

d) All the sub keys can be obtained through the above steps, so as to calculate the initial key of the recovery cipher device.

3 Implementation of masked algorithm

3.1 Mask technology

Masking is an effective anti power attack technology. Its basic principle is to cover all the intermediate values generated in the encryption and decryption process of cryptographic devices with random numbers, and cut off the dependence between the intermediate values and power consumption, resulting in the attacker unable to recover the key according to the intermediate value, so as to achieve the purpose of resisting power consumption attack [9]. In the mask protection scheme of the algorithm, every intermediate value \( x \) is masked by random mask value \( m \), so the masked intermediate value \( x_m \) and unmasked intermediate value \( x \) are independent of each other, and the power consumption of \( x \) and power consumption of \( x_m \) are also independent of each other. Satisfy:

\[
x_m = x \# m
\]

(2)

Where "\#" is defined as an operation between the intermediate value \( x \) and the mask value \( m \). The common functions are Boolean function ("\( \oplus \)"), Moore addition function ("\( + \)"") and Moore multiplication function ("\( \times \")).

If the masked intermediate value \( x_m \), unmasked intermediate value \( x \) and random mask value satisfy the condition of pairwise independence, the scheme can resist first-order differential power attack (DPA). Due to the low cost, easy implementation and good effectiveness of mask technology, it has been widely studied and applied in cryptography [10].

In this paper, the method of combining additive mask and multiplication mask is used to mask the intermediate value.
### 3.2 Multiplication mask S-box

According to section 2.2, the S-box of SM4 algorithm is divided into two affine transformations and one inverse transformation. Let the two affine transformations be expressed as $f_1(x), f_2(x)$ respectively, and satisfy the requirements

$$f_1(x) = f_2(x) = x \cdot a + b$$

(3)

The inverse operation is expressed as $g(x) = x^{-1}$, and the input of S-box is set as $x$, the output is satisfied

$$sbox(x) = f_2(g(f_1(x))) = x^{-1} + ab + b$$

(4)

Using multiplication mask to realize S-box, because there is a $g(x \cdot m) = (x \cdot m)^{-1} = x^{-1} \cdot m^{-1}$, we can get the structure of the inverse part of S-box.

The implementation flow of multiplication mask is shown in Figure 3.

![Fig. 3. Inverse transformation flow of multiplication](image)

Using multiplication mask to realize S-box, because there is a $g(x \cdot m) = (x \cdot m)^{-1} = x^{-1} \cdot m^{-1}$, we can get the structure of the inverse part of S-box.

The implementation flow of multiplication mask is shown in Figure 3.

![Fig. 4. Implementation flow of multiplication mask](image)

The output result of multiplication mask S-box is represented by $sbox'$. It can be seen from Fig. 4 that

$$sbox' = sbox(x + m) = sbox(x) + m'$$

(5)
Here, we have $m' = ma^2$.

3.3 Mask implementation circuit

Since the attacker mainly chooses the first few rounds or the last few rounds of SM4 algorithm implementation for power consumption attack, this scheme focuses on protecting the first six rounds and the last six rounds of algorithm implementation, carries out full random mask, and carries out fixed random mask for the middle 24 rounds, which effectively reduces the consumption of random mask.

The implementation flow of SM4 algorithm mask circuit is shown in Figure 5. Four 32-bit registers reg1, reg2, reg3 and Reg4 are set, with a total of 128 bits, to store the intermediate value results of each round of operation of SM4 algorithm. The S-box substitution is replaced by the multiplication mask S-box, which is called ys box. B1, B2, B3 and B4 are four 32-bit compensation transforms, which are used to ensure that the result after each round of compensation operation is equal to the sum of the exclusive or of the output without mask and the random mask value of the current round.

![Fig. 5. SM4 algorithm mask circuit](image)

The 128 bit plaintext data input by SM4 is XOR with 128 bit random number. The 128 bit plaintext is divided into four words from left to right, with 32 bits in each word, which are marked as $min1, min2, min3, min4$ in turn. Similarly, 128 bit random numbers are also divided into 4 32-bit words, which are marked as $r1, r2, r3, r4$ in turn. The initial input plaintext data is randomly masked and stored in reg1, reg2, reg3, Reg4 four intermediate value registers as the input value of the first round in the next 32 rounds of iterative operation.

First of all, the first six rounds of iterative operation are performed, and the full random mask is used, that is, each round needs to update a set of masks. Let the initial input value of the unmasked $i$ iteration operation be marked as $a_i$, divide it into four 32 bit words from left to right, mark them as $a1_i, a2_i, a3_i, a4_i$, and the corresponding random mask is $r_i$, and marked as $r1_i, r2_i, r3_i, r4_i$ from left to right, here $i \in \{1, 2, 3, 4, 5, 6\}$

As can be seen from Figure 5, the value in register reg2 needs to be preserved in register reg1 through compensation transformation B1. The operation of compensation transformation B1 can be expressed by the following formula

$$B1((a2_i + r2_i), r1_i, r2_i) = (a2_i + r2_i) + r1_i + r2_i$$

(6)
After the compensation transformation $B_1$, the random mask number $C$ is updated to $D$. Similarly, the operation formula of the compensation transformation $B_2$ and $B_3$ is

$$B_2((a_3 + r_3_i), r_{2_i+1}, r_3) = (a_3 + r_3_i) + r_{2_i+1} + r_3$$ (7)

$$B_3((a_4 + r_4_i), r_{3_{i+1}}, r_4) = (a_4 + r_4_i) + r_{3_{i+1}} + r_4$$ (8)

After compensation transformation $B_2$, $B_3$, the random number $r_{2_i}$ is updated to $r_{2_{i+1}}$, and $r_3$ is updated to $r_{3_{i+1}}$.

The input value of $B_4$ goes through round key addition, $y_s$ transformation and $l$ transformation in turn. As can be seen from Figure 5, the input value of $y_s$ transformation is

$$y_s \text{ in} = (a_2_i + r_2_i) + (a_3_i + r_3_i) + (a_4_i + r_4_i) + r_k_i$$ (9)

Where $r_k_i$ is the round key of round $i$.

$$y_s ((a_2_i + r_2_i) + (a_3_i + r_3_i) + (a_4_i + r_4_i) + r_k_i)$$

$$= S \left( a_2_i + a_3_i + a_4_i + r_k_i \right) + S \left( r_{4_{i+1}} \right)$$ (10)

$y_s$ transformation is a multiplication random mask S-box. See Section 2.2 for specific implementation. $y_s$ transformation is expressed by formula (10).

After $y_s$ transformation, the random number $r_{4_i}$ is updated to $r_{4_{i+1}}$. $L$ transform into linear transformation, the process is as follows

$$L \left( S (a_2_i + a_3_i + a_4_i + r_k_i) + S (r_{4_{i+1}}) \right)$$

$$= L \left( S (a_2_i + a_3_i + a_4_i + r_k_i) \right) + L \left( S (r_{4_{i+1}}) \right)$$ (11)

Therefore, the input of compensation transformation $B_3$ is

$L \left( S (a_2_i + a_3_i + a_4_i + r_k_i) + L \left( S (r_{4_{i+1}}) \right) \right) + (a_1_i + a_2_i)$ (12)

According to the definition of compensation transformation, in order to ensure that the result of each round of compensation operation is equal to the sum of the exclusive or of the output without mask and the random mask value of the current round, the compensation transformation $B_4$ is defined as

$$B_4(x, r_{4_{i+1}}, r_2_i) = x + r_{4_{i+1}} + r_2_i + L \left( S (r_{4_{i+1}}) \right)$$ (13)

The output after $B_4$ compensation transformation is the input of register $Reg_4$, which is shown as follows

$$L \left( S (a_2_i + a_3_i + a_4_i + r_k_i) \right) + a_1_i + a_{4_{i+1}}$$ (14)

Secondly, the fixed random mask method is used to execute the middle 24 rounds, i.e. the random masks used in each round are the same, which are uniformly marked as $r_1, r_2, r_3, r_4$. Compared with the previous six rounds, the operation formula of compensation transformation $B_1$, $B_2$, $B_3$ and $B_4$ in the middle 24 rounds is as follows:

$$B_1((a_2_i + r_2_i), r_1_i, r_2_i) = (a_2_i + r_2_i) + r_1_i + r_2_i$$ (15)

$$B_2((a_3_i + r_3_i), r_2_i, r_3_i) = (a_3_i + r_3_i) + r_2_i + r_3_i$$ (16)

$$B_3((a_4_i + r_4_i), r_3_i, r_4_i) = (a_4_i + r_4_i) + r_3_i + r_3_i$$ (17)

$$B_4(x, r_4_i, r_2_i) = x + r_4_i + r_2_i + L \left( S (r_4_i) \right)$$ (18)

After the compensation transformation $B_1$, $B_2$, $B_3$, the random number is no longer updated. The input to register $Reg_4$ is represented as follows

$$L \left( S (a_2_i + a_3_i + a_4_i + r_k_i) \right) + a_1_i + r_4_i$$ (19)
The process of the last six rounds is the same as that of the first six rounds, which is not detailed here.

The algorithm realizes the chip built-in true random number generator. The mask required in the scheme is generated by the random number generator. The above process is the whole implementation process of SM4 algorithm semi random mask circuit. After analysis, a complete SM4 encryption operation needs 13 128 bit random mask.

4. experimental results
Using the FPGA power attack platform designed by ourselves, after collecting the power curve data, according to the power analysis theory of differential power attack technology, the power consumption data curve is processed, and finally presented in a certain way, so as to achieve the purpose of cracking the key.

![Fig. 6. DPA attack results of original SM4 algorithm](image)

4000 power consumption curves are collected for analysis. When attacking the original SM4 algorithm, there is a obvious peak in the picture, as shown in Figure 6, however, When attacking the masked SM4 algorithm, there is no obvious peak in the picture, as shown in Figure 7, which proves the effectiveness of the proposed scheme against power consumption attacks.

![Fig. 7. DPA attack results of masked SM4 algorithm](image)

5. Conclusion
As a block cipher standard in China, SM4 has high theoretical security, but its implementation security is facing a huge threat of power consumption attacks. In order to ensure the security implementation of SM4 algorithm and its hardware scheme, and improve the ability of SM4 algorithm to resist power consumption attacks, this paper designs and implements the SM4 algorithm multiplication mask S-box based on the structure characteristics of SM4 algorithm S-box, and combines with the implementation structure of SM4 algorithm, designs and implements the semi-random mask circuit scheme of SM4.
algorithm, which implements the first six rounds and the last six rounds of SM4 algorithm. The middle value is completely random mask, and the middle 24 round intermediate value is fixed mask, which effectively reduces the consumption of random mask. After security analysis and experimental verification, the scheme can resist the first-order power attack, and has less random number and lower implementation area. In the next step, we will improve the scheme to make it have the ability to defend against second-order and even higher-order power consumption attacks.

References
[1] National commercial password management office. SMS cipher algorithm for WLAN products.http://www.oscca.gov.cn/UpFile/.pdf.
[2] Liu F, Ji W, Hu L, et al. Analysis of the SMS4 Block Cipher[C]// Information Security and Privacy, Australasian Conference, ACISP 2007, Townsville, Australia, July 2-4, 2007, Proceedings. DBLP, 2007:158-170.
[3] Bai X, Guo L, Li T. Differential power analysis attack on SM4 block cipher[C]//Circuits and Systems for Communications, 2008. ICCSC 2008. 4th IEEE International Conference on. IEEE, 2008:613-617.
[4] Duan X, Hu R, Li X Y. Research and Implementation of DPA-resistant SM4 Block Cipher[C]// International Conference on Computational Intelligence & Security. IEEE Computer Society, 2011:1033-1036.
[5] Liang H, Wu L, Zhang X, et al. Design of a Masked S-Box for SM4 Based on Composite Field[C]// Tenth International Conference on Computational Intelligence and Security. IEEE, 2014:387-391.
[6] Pei Chao. A Method of Masking SM4 and Analysis against DPA Attacks[J]. Journal of Cryptologic Research,2016,3(1):79–90.
[7] Lv Shuwang, Su Bozhan, Wang Peng. et al.SMS4 Block Cipher Algorithms [J]. Journal of Information Security Research, 2016,11:995-1007.
[8] Kocher P, Jaffe J, Jun B. Differential Power Analysis[J]. Lecture Notes in Computer Science, 1999, 1666:388-397.
[9] Mangard S, Oswald E, Popp T. Power Analysis Attacks: Revealing the Secrets of Smart Cards[M]. Springer Publishing Company, Incorporated, 2010.
[10] Zhu Nianhao. Research on Power Analysis Attack of Symmetric Encryption Algorithm[D]. Shanghai Jiao Tong University,2014.
[11] Dong Liling. The Optimization and Research for AES Cipher Chips with Power Attack Resistance [D]. Nanjing University of Aeronautics and Astronautics,2016.
[12] Blömer J, Guajardo J, Krummel V. Provably Secure Masking of AES[M]/ Selected Areas in Cryptography. Springer Berlin Heidelberg, 2004:69--83.
[13] Canright D, Batina L. A very compact "Perfectly masked" S-box for AES[C]// International Conference on Applied Cryptography and Network Security. Springer-Verlag, 2008:446-459.Fridrich J, Kodovsky J. Rich Models for Steganalysis of Digital Images[J]. IEEE Transactions on Information Forensics & Security, 2012, 7(3):868-882.