Общероссийский математический портал

S. V. Matveev, Маскирование ГОСТ 28147-89 для защиты от атак по побочным каналам, *Матем. вопр. криптогр.*, 2015, том 6, выпуск 2, 35–43

DOI: <https://doi.org/10.4213/mvk143>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.80.73.141

26 апреля 2023 г., 16:37:09

УДК: 519.719.2

GOST 28147-89 masking against side channel attacks

S. V. Matveev

Technical committee for standardization (TC 26), Moscow

Получено 16.IX.2014

Side-channel attacks exploit leakage from the physical implementation of a cryptographic algorithm to obtain some additional information on its secret parameters. During the last decade we observe the intensive development of various side-channel attacks, that affect security of many popular cryptosystems. In an attempt to reduce the possible damage a general method that masks the intermediate data was proposed. This method was studied for popular cryptographic algorithms such as RSA, DES, AES and several of the AES candidates. In this paper we propose a strategy of masking for Russian cryptographic standard GOST 28147-89 and perform an analysis of its properties.

Key words: GOST 28147-89, side-channel attack

Маскирование ГОСТ 28147-89 для защиты от атак по побочным каналам

С. В. Матвеев

Технический комитет по стандартизации (ТК 26), Москва

Аннотация. Атаки по побочным каналам используют сигналы, порождаемые физической реализацией криптографических алгоритмов, для получения дополнительной информации о секретных параметрах. В последнее десятилетие наблюдается интенсивное развитие различных типов атак по побочным каналам, которые влияют на оценки стойкости ряда распространенных криптосистем. Для противодействия атакам подобного типа был предложен общий метод маскирования промежуточных данных. Этот метод изучался для алгоритмов RSA, DES, AES и некоторых кандидатов конкурса AES. Здесь предлагается способ маскирования для российского криптографического стандарта ГОСТ 28147-89 и проводится анализ его свойств.

Ключевые слова: ГОСТ 28147-89, атаки по побочным каналам

Citation: *Mathematical Aspects of Cryptography*, 2015, vol. 6, no. 2, pp. 35–43 (Russian).

1. Introduction

One of the modern cryptography research areas are attacks on the cryptographic algorithms based on supervision of physical processes in cryptographic device. In literature this research area has received the name of side-channel attacks. Examples of such attacks are: Differential Power Analysis (DPA) [1, 2], Timing Analysis (TA) [3], Correlation Power Analysis (CPA) [4, 5], Mutual Information Attack (MIA) [19]. Various countermeasures were also introduced for the most popular algorithms [6, 7]. For example, in paper [6] DPA attacks on AES candidates was investigated. A popular method to counteract side-channel attacks uses masking for the data being processed. This method uses random masks for results of internal operations of a cryptographic algorithm.

The possibility of using side-channel attacks is determined and depends on the way of realization of the algorithm on a hardware platform, and additional data available by the attacker. Usually the investigations of such possibilities are based on the Hamming weight model [1,2], that is on the number of nonzero bits in a data word. The main countermeasures for protection of block encryption algorithms from side channel attacks are: when implementing the branching of iterative transformations, with equal energy, introduction of randomness in dependence between the secret algorithm parameters and possible intermediate data

In this paper we propose an algorithm for protection against side-channel attacks for the Russian cryptographic standard GOST 28147-89 (GOST). The development of the specified algorithm protection aims to simplify its implementation and to minimize the additional costs of hardware resources. The simplicity of implementation in this case does not assume the choice of special parameters of the algorithm. This algorithm is universal for all hardware platforms.

To substantiate the effectiveness of the proposed method of protection against side-channel attacks we consider the dependence between the success probability of attack, signal-to-noise ratio (SNR) and the number of measurements.

2. Preliminaries

The GOST encryption algorithm is an iterative block cipher based on the Feistel network. 32-bit block of data will be processed with the following operations:

- addition of 32-bit round key modulo 2^{32} ,
- non-linear substitution,
- cyclic left shift,
- bitwise addition of the obtained result to the previously calculated block.

Thus, for side-channel attacks the following types of information may be used: a 32-bit key, outputs of adders modulo 2^{32} , observations at the output of a substitution and the output after cyclic left shift. Therefore, the possibly best way of protection from the attack may be to ensure masking of each possible leakage point. When considering a security methodology we assume that the attacker may observe the information leakage from the part of the 32-bit subkey and from some bits of masks.

3. Masking strategy

3.1. Masking for cryptographic algorithms GOST

Consider the encryption process of the GOST cipher. It could be seen as a recurrent sequence. The initial elements of a sequence of 32-bit words are the plaintexts: $A(0), A(1)$.

The subsequent elements of the sequence are calculated according to the rule:

$$A(i+1) = A(i-2) \oplus R(S(A(i-1) \boxplus X(i-1))), i = 2, \dots, 33,$$

where:

$\boxplus$ — addition modulo 2^{32} ,

$\oplus$ — bitwise addition of 32-bit vectors,

$S(Z)$ — substitution on 32-bit vector, $S : V_{32} \mapsto V_{32}, V = \{0, 1\}$,

R — cyclic left shift of a 32-bit vector on 11 bits,

$$X^0 = \{X(i) : i = 1, \dots, 32\} =$$

$= \{X_0, \dots, X_7, X_0, \dots, X_7, X_0, \dots, X_7, X_7, \dots, X_0\}$ — sequence of 32-bit round keys.

Words $A(32), A(33)$ are the result of encryption.

The masking algorithm consists of the following steps:

- generate a sequence of random bits $m(1), \dots, m(32)$,
- generate a sequence of inverse round keys $X^1 = \overline{X} = \{\overline{X(i)}, i = 1, \dots, 32\}$,
- generate substitutions $S^{m_1, m_2}(Z) = S(Z \oplus m_2 \cdot e) \oplus m_1 \cdot e$, where $m_1, m_2 \in \{0, 1\}$, $e = (1, \dots, 1) \in V_{32}$ is the unit vector.

The calculation of the recurrent sequence is as follows:

$$\begin{aligned}
 B(0) &= A(0) \oplus m(0) \cdot e = A(0), \\
 B(1) &= A(1) \oplus m(0) \cdot e, \\
 m(0) &= m(33) = 0, \\
 B(i) &= B(i-2) \oplus R(S^{m(i) \oplus m(i-2), m(i-1)}(B(i-1) \boxplus X^{m(i-1)}(i-1) \boxplus m(i-1))), \\
 i &= 2, \dots, 33, \\
 A(32) &= B(32) \oplus m(32) \cdot e, A(33) = B(33) \oplus m(33) \cdot e = B(33).
 \end{aligned}$$

It is easy to show that the algorithm is equivalent to the original.

Therefore, each round key X depends on the random bit of the mask and could be observed by an attacker like X or $\bar{X}$. The output of the adder modulo 2^{32} depends on the value of mask bit and observed either directly as Z , either as $\bar{Z} - 1$. The substitution output depends on the values of the combination of random masks and observed directly or inversely.

Additional condition for implementation of the algorithm is the absence of conditional transitions.

It is easy to see that the proposed algorithm will require additional memory for 32-bit mask, a 256 bit key and, in the case of using 8 different 4-bit substitutions, $4 \times 16 \times 8 \times 3 = 1536$ bits. Totally no more than 1824 extra bits. This overhead is required only at the preparatory stage of encryption to generate the inverse keys and substitutions.

3.2. Masking of cryptographic algorithms based on SP Network

The proposed method of masking can be applied to the cryptographic algorithms based on the substitution permutation network.

Let's consider the SPN-based cryptographic algorithm as a sequence of transformations of the n -bit vectors. We will use following transformations: a substitution, a permutation and a bitwise addition of a round key.

Denote:

- $A(0)$ — initial n -bit vector,
- $\oplus$ — bitwise addition of n -bit vectors,
- $S(Z)$ — substitution of the n -bit vector Z ,
- $X^0(i)$ — n -bit round key, $i=1, \dots, t$,

t — number of iterations of the algorithm,

$L_{n \times n}$ — linear transformation of the size $n \times n$ with elements from the set $\{0,1\}$, $L = (l_1^\downarrow, \dots, l_n^\downarrow)$, where $l_i^\downarrow$ — have odd weight for all i ,

$A \cdot L_{n \times n}$ — operation of multiplication of a vector A on a matrix $L_{n \times n}$,

$E_{n \times n}$ — identity matrix of size $n \times n$,

e — unit vector of length n .

The transformation of the initial vector $Y(0)$ is performed as follows:

$A(i) = S(A(i-1) \oplus X(i)) \cdot L$, for $i = 1, \dots, t$. The result is $A(t)$.

The protection algorithm consists in the following.

- generate a sequence of random bits $m(i), i = 1, \dots, t-1$,
- generate a sequence of vectors $\{X^1(i) = X(i) \oplus e, i = 1, \dots, t\}$,
- generate substitutions $S^m(Z) = (S(Z \oplus m \cdot e)) \oplus m \cdot e, m \in \{0, 1\}$.

The transformation of the original vector follows the rule:

$$B(0) = A(0),$$

$$B(i) = (S^{m(1)+\dots+m(i)}(B(i-1) \oplus X^{m(i)}(i))) \cdot L,$$

$$i = 1, \dots, t.$$

Obviously, $A(t) = B(t) \oplus \sum_{j=1}^t m(j) \cdot e$ by the definition of $m(t)$ and the matrix L .

4. Analysis

For the analysis of the proposed approaches for side-channel attacks protection we consider simplified models proposed in [10, 20]. We assume that a Hamming weight of subkey is a sensitive variable. The consider the following leakage model: $Y(t) = HW(X_c) + \xi(t)$. Here X_c is the correct subkey, $HW(X)$ is a Hamming weight of subkey X and $\xi(t) \sim N(0, \sigma^2)$ is the independent Gaussian noise. In our analysis we consider attack on one subkey. The aim of side-channel attack is to distinguish the correct key X_c from all possible key hypotheses X based on T independent realizations of the noisy physical leakage $Y(1), \dots, Y(T)$, where T is the number of measurements.

Using maximum likelihood method, we obtain the estimate Y^* :

$$Y^* = \frac{1}{T} \sum_{t=1}^T Y(t).$$

We take the closest integer to Y^* as an estimate of the unknown value $HW(X_c)$. Under our model the estimate Y^* has a normal distribution $Y^* \sim N\left(HW(X_c), \frac{\sigma^2}{T}\right)$. Therefore, the success probability of attack is

$$SR = \mathbf{P}\left(HW(X_c) - \frac{1}{2} < Y^* \leq HW(X_c) + \frac{1}{2}\right) = 2\Phi\left(\frac{\sqrt{T}}{2\sigma}\right) - 1. \quad (1)$$

Here $\Phi(\cdot)$ is the cumulative distribution function of Gaussian distribution with mean 0 and variance 1.

In the case of the proposed protection method we use the following leakage model: $Y(t) = HW(X_c^{m(t)}) + \xi(t)$, $m^*(t) = m(t) + \eta(t)$. Here $\xi(t)$ and $\eta(t)$ are independent Gaussian noises with mean 0 and variance σ^2 , like above. We suppose that $Y(t)$ and $m^*(t)$ are observations at different points.

Denote by d the maximum count of bits in the observed subkey.

Likelihood function for this case is

$$\begin{aligned} Lik(Y(t), m^*(t)) &= \\ &= \frac{1}{(2\sqrt{2\pi}\sigma)^T} \prod_{t=1}^T \left(\exp\left(-\frac{(m^*(t))^2 + (Y(t) - Y)^2}{2\sigma^2}\right) + \exp\left(-\frac{(m^*(t) - 1)^2 + (Y(t) - d + Y)^2}{2\sigma^2}\right) \right). \end{aligned}$$

Denote

$$\begin{aligned} Z^*(t) &= Y(t) - \frac{d}{2}, \quad Z = Y - \frac{d}{2} \\ U^*(t) &= m^*(t) - \frac{1}{2}, \quad U(t) = m(t) - \frac{1}{2}. \end{aligned}$$

Logarithm of likelihood function (without some constants) is

$$L(Z) = - \sum_{t=1}^T \frac{(Z^*(t))^2 + (U^*(t))^2}{2\sigma^2} - \frac{T}{8\sigma^2} - \frac{TZ^2}{2\sigma^2} + \sum_{i=1}^T \ln \left(2 \operatorname{ch} \left(\frac{U^*(t) - 2Z \cdot Z^*(t)}{2\sigma^2} \right) \right),$$

where $\operatorname{ch}(\cdot)$ is hyperbolic cosine.

We use approximations $\ln \left(\operatorname{ch} \left(\frac{x}{\sigma^2} \right) \right) \approx \frac{x^2}{2\sigma^2}$ and

$$\begin{aligned} L(Z) &\approx - \sum_{t=1}^T \frac{(Z^*(t))^2 + (U^*(t))^2}{2\sigma^2} - \frac{T}{8\sigma^2} - \frac{TZ^2}{2\sigma^2} + T \ln 2 + \\ &\quad + \sum_{t=1}^T \frac{U^*(t)^2 - 4Z \cdot Z^*(t) \cdot U^*(t) + 4(Z \cdot Z^*(t))^2}{8\sigma^4} \end{aligned}$$

for large values σ^2 .

Basing on this approximation we define two statistics

$$T_1(Z^*, U^*) = \sum_{t=1}^T Z^*(t) U^*(t)$$

$$T_2(Z^*, U^*) = \sum_{t=1}^T Z^*(t)^2.$$

We use the statistic T_1 with additional coefficients as an estimate of Y :

$$Y^* = -\frac{2}{T} \sum_{t=1}^T Z^*(t) U^*(t) - \frac{d}{2} = -\frac{2}{T} \sum_{t=1}^T \left[(Y(t) - \frac{d}{2}) \cdot (m^*(t) - \frac{1}{2}) \right] + \frac{d}{2}.$$

We have

$$\mathbf{E} \left[(Y(t) - \frac{d}{2}) \cdot (m^*(t) - \frac{1}{2}) \right] = \frac{1}{2} \left[\frac{d}{2} - HW(X_c) \right],$$

$$\begin{aligned} \mathbf{D} \left[(Y(t) - \frac{d}{2}) \cdot (m^*(t) - \frac{1}{2}) \right] &= \sigma^4 + \sigma^2 \left[(HW(X_c) - \frac{d}{2})^2 + \frac{1}{4} \right] = \\ &= \sigma_1^2 (HW(X_c), d, \sigma). \end{aligned}$$

Further we use the notation $\sigma_1 = \sigma_1(HW(X_c), d, \sigma)$ to simplify formulas.

The statistic Y^* has an asymptotically normal distribution:

$$Y^* \sim N \left(HW(X_c), \frac{4\sigma_1^2}{T} \right) \text{ for } T \rightarrow \infty.$$

As in the previous case, we take the closest integer as an estimate of the target data. Thus, the success probability of attack is

$$SR = \mathbf{P} \left(HW(X_c) - \frac{1}{2} < Y^* \leq HW(X_c) + \frac{1}{2} \right) = 2\Phi \left(\frac{\sqrt{T}}{4\sigma_1} \right) - 1. \quad (2)$$

The mean value of the probability of success in this case is

$$\mathbf{E}(SR) = \sum_{i=0}^d \frac{C_d^i}{2^d} \left[2 \cdot \left[\Phi \left(\frac{\sqrt{T}}{4\sigma_1} \right) - 1 \right] \right]. \quad (3)$$

We define $SNR = \frac{1}{\sigma}$ as the signal-to-noise ratio (SNR) for the side channel attack. From (1) and (3) we can see that for protected GOST the attacker needs more measurements for key recovery with equal SR and SNR. For example, in the Table 1 we show dependence between SNR and the number of measurements

Table 1. Dependence between SNR and the number of measurements with SR=0.95.

SNR	The number of measurements			
	Unprotected GOST	GOST with masking		
		d=4	d=8	d=16
0.1	1536	622335	628543	641055
0.05	6146	$9.88 \cdot 10^6$	$9.88 \cdot 10^6$	$9.93 \cdot 10^6$
0.01	153658	$6.114 \cdot 10^9$	$6.114 \cdot 10^9$	$6.115 \cdot 10^9$

sufficient for the mean of probability of successful attack to be no smaller than SR=0.95 for some values d for unprotected algorithm and for GOST with masking.

In practice we may use the following approximation for the probability of successful attack for scheme with masking:

$$SR = 2\Phi\left(\frac{\sqrt{T}}{4 \cdot \sigma^2}\right) - 1. \quad (4)$$

5. Conclusion

The proposed strategy allows an effective protection of cryptographic algorithm from side-channel attacks with small hardware costs. There exists a possibility of using this strategy for a sufficiently large class of block ciphers.

References

- [1] Kocher P., Jaffe J., Jun B., "Introduction to differential power analysis and related attacks. Tech. Rept", Cryptography Research Inc., 1998, <http://www.cryptography.com/resources/whitepapers/DPA-technical.html>.
- [2] Kocher P., Jaffe J., Jun B., "Differential power analysis", CRYPTO'99, Lect. Notes Comput. Sci., **1666**, Springer-Verlag, 1999, 388–397.
- [3] Kocher P.C., "Timing attacks on implementations of Diffie-Hellman, RSA, DSS and other systems", CRYPTO'96, Lect. Notes Comput. Sci., **1109**, 104–113.
- [4] Coron J.-S., Kocher P., Naccache D., "Statistics and secret leakage", FC 2000, Lect. Notes Comput. Sci., **1972**, 2001, 157–173.
- [5] Brier E., Clavier C., Olivier F., "Correlation power analysis with a leakage model", CHES 2004, Lect. Notes Comput. Sci., **3156**, 2004, 16–29.
- [6] Messerges T. S., "Securing the AES finalists against power analysis attacks", FSE 2001, Lect. Notes Comput. Sci., **1978**, 2001, 150–164.
- [7] Rivain M., Dottax E., Prouff E., "Block ciphers implementations provably secure against second order side channel analysis", FSE 2008, Lect. Notes Comput. Sci., 2008, 127–143.

- [8] Gerard B., Grosso V., Naya-Plasencia M., Standaert F.-X., “Block ciphers that are easier to mask: How far can we go”, CHES 2013, Lect. Notes Comput. Sci., **8086**, 2013., 383–399.
- [9] Fei Y., Luo Q., Ding A. A., “A statistical model for DPA with novel algorithmic confusion analysis”, CHES 2012, **7428** (2012), 233–250.
- [10] Fei Y., Ding A. A., Lao J., Zhang L., “A statistics-based fundamental model for side-channel attack analysis”, <https://eprint.iacr.org/2014/152.pdf>.
- [11] Debraize B., “Efficient and provably secure methods for switching from arithmetic to Boolean masking”, CHES 2012, Lect. Notes Comput. Sci., **7428**, 2012, 107–121.
- [12] Doget J., Prouff E., Rivain M., Standaert F.-X., “Univariate side channel attacks and leakage modelling”, *J. Cryptographic Engineering*, **1:2** (2011), 123–144.
- [13] Mangard S., Oswald E., Standaert F.-X., “One for all – all for one: unifying standard DPA attacks”, *IET Information Security*, **5**, 2011, 100–110.
- [14] Standaert F.-X., Malkin T. G., Yung M., “A unified framework for the analysis of side-channel attacks”, EUROCRYPT 2009, Lect. Notes Comput. Sci., **5479**, 2009, 443–461.
- [15] Whitnall C., Oswald E., “A fair evaluation framework for comparing side-channel distinguishers”, *J. Cryptographic Engineering*, **1:2** (2011), 145–160.
- [16] Duc A., Dziembowski S., Faust S., “Unifying leakage models: from probing attacks to noisy leakage”, <https://eprint.iacr.org/2014/079.pdf>.
- [17] Rivain M., Prouff E., “Provably secure higher-order masking of AES”, <https://eprint.iacr.org/2010/441.pdf>.
- [18] Rivain M., Prouff E., Doget J., “Higher-order masking and shuffling for software implementations of block ciphers”, CHES 2009, Lect. Notes Comput. Sci., **5747**, 2009, 171–188.
- [19] Gierlichs B., Batina L., Tuyls P., Preneel B., “Mutual information analysis”, CHES 2008, **5154** (2008), 426–442.
- [20] Messerges T. S., Dabbish E. A., Sloan R. H., “Examining smart-card security under the threat of power analysis attacks”, *IEEE Trans. on Computers*, **51:5** (2002), 541–552.
- [21] Hajra S., Mukhopadhyay D., “SNR to success rate: reaching the limit of non-profiling DPA”, <https://eprint.iacr.org/2013/865.pdf>.