UDC 001.89
IRSTI 81.93.29

STUDY OF A DYNAMIC SUBSTITUTION BOX

N.A. Kapalova[1], A. Haumen[2]

[1,2]Institute of Information and Computational Technologies, Almaty, Kazakhstan
[1]kapalova@ipic.kz, [2]haumen.armanbek@gmail.com
[1]ORCID ID: https://orcid.org/0000-0001-9743-9981
[2]ORCID ID: https://orcid.org/0000-0002-1670-2520

Abstract. The paper examines a previously developed method for generating a dynamic substitutionbox. This method allows generating substitution boxes (S-boxes) that change depending on the values of some parameter obtained from the secret key of the algorithm. The generated S-boxes are resistant to linear and differential cryptanalysis, since they are random, not known in advance, and depend on the secret key. The obtained dynamic substitution boxes were examined for nonlinearity, and the results were compared with S-boxes of well-known symmetric block algorithms such as AES, Kuznyechik, SM4, BelT, and Kalyna.

Keywords: encryption algorithm, substitution box, dynamic S-box, nonlinear transformation.

Introduction

In the era of information technology development, symmetric block encryption algorithms are the main cryptographic means of ensuring confidentiality when processing information in modern information and telecommunication systems. Besides, block ciphers are used to ensure the integrity of information, and also as a basic element in the construction of other cryptographic primitives [1], such as pseudo-random sequence generators (PRNG), stream ciphers, and hash functions. The level of strength and properties of the symmetric block encryption algorithm used in a system to a significant extent determine the strength of the cryptographic protection of information, the security of cryptographic protocols, and the security of the information and telecommunications system as a whole.

What conditions must a strong block cipher satisfy? These conditions were formulated by C. Shannon in a number of his fundamental works on the theory of encryption [2]. A strong cipher must offerthe properties of diffusion and confusion:

1. Diffusion is a property of a cipher, when one character (bit) of the original text affects several characters (bits) of the ciphertext, optimally – all characters within one block. If this condition is met, then when encrypting two data blocks with minimal differences between them, completely different ciphertext blocks should be obtained. Exactly the same result should take place for the dependence of the ciphertext on the key – one character (bit) of the key must affect several characters (bits) of the ciphertext. Diffusion hides the relationship between the ciphertext and the original text.
2. Confusion is a property of a cipher to hide dependencies between characters in the original text and the ciphertext. If the cipher "mixes" the bits of the original text well enough, then the corresponding ciphertext does not contain any statistical and, moreover, functional regularities – again, for an outside observer with only limited computational resources. Confusion hides the relationship between the ciphertext and the key.

To meet the requirements for block ciphers, modern encryption algorithms use various transformations. This paper considers nonlinear transformations used in symmetric block encryption algorithms.

Nonlinear transformation in symmetric block encryption algorithms

A nonlinear transformation is required for every modern encryption algorithm and has been proven to be a strong cryptographic primitive against linear and differential cryptanalysis. Nonlinear

transformations in modern symmetric block algorithms are implemented in the form of substitution boxes (S-boxes) [3].

Considering that most modern block algorithms (Rijndael, Camellia, DES, etc.) use a single linear operation (modulo 2 addition) to introduce round keys and combine inter-round values, S-boxes are the only element that determines the nonlinearity of the encryption transformation and its level. resistance to cryptanalytic attacks. The number of rounds required to ensure the resistance of block ciphers to known types of cryptographic analysis is calculated based on the specified properties of nonlinear substitution nodes [4].

Many stream algorithms, cryptographic hash functions, and pseudo-random sequence generators are based on block ciphers or their structural elements. Thus, the cryptographic strength of most modern symmetric block algorithms largely depends on the properties of the selected S-boxes.

Despite the multi-modulus addition, the main nonlinear elements of modern ciphers are substitutions. Many different S-box criteria determine the strength of an encryption algorithm. However, today there is no unambiguous opinion about the need for most of them. Regardless of many existing solutions in the field of symmetric cryptography, the question of finding substitutions, the use of which in encryption algorithms provides protection against existing and promising types of attacks, remains relevant [4].

Substitutions used in cryptoalgorithmsshould meet the following criteria [5]:

a) Maximizingnonlinearity;
b) Theminimumdegreeis 3;
c) Minimizing the maximum value of the table of differentials.

In this paper, the properties of the nonlinearity of well-known encryption algorithms such as AES, Grasshopper, SM4, BelT, and others are considered, and the nonlinearity values of the S-boxes of these algorithms are calculated using a computer program. The program is implemented in Python 3.

Substitution boxes of known encryption algorithms

AES/Rijndael

Advanced Encryption Standard (AES), also known as a symmetric block cipher algorithm adopted as an encryption standard by the US government through an AES competition.

AES [7] is an encryption algorithm for 128-bit data blocks with keys of 128, 192, and 256 bits. AES is a simplified version of the Rijndael algorithm [8]. The original Rijndael algorithm differs in that it supports a wider set of block lengths.

The AES algorithm is widely used in cryptographic applications. In the AES algorithm, the SubBytes transformation is a nonlinear byte substitution performed independently with each byte [9]. The S-box substitution boxes are a fixed 8x256 bit table (Table 1). The nonlinear substitution operation is the main strength factor of this algorithm.

Table 1.

63, 7c, 77, 7b, f2, 6b, 6f, c5, 30, 01, 67, 2b, fe, d7, ab, 76, ca, 82, c9, 7d, fa, 59, 47, f0, ad, d4, a2, af, 9c, a4, 72, c0, b7, fd, 93, 26, 36, 3f, f7, cc, 34, a5, e5, f1, 71, d8, 31, 15, 04, c7, 23, c3, 18, 96, 05, 9a, 07, 12, 80, e2, eb, 27, b2, 75, 09, 83, 2c, 1a, 1b, 6e, 5a, a0, 52, 3b, d6, b3, 29, e3, 2f, 84, 53, d1, 00, ed, 20, fc, b1, 5b, 6a, cb, be, 39, 4a, 4c, 58, cf, d0, ef, aa, fb, 43, 4d, 33, 85, 45, f9, 02, 7f, 50, 3c, 9f, a8, 51, a3, 40, 8f, 92, 9d, 38, f5, bc, b6, da, 21, 10, ff, f3, d2, cd, 0c, 13, ec, 5f, 97, 44, 17, c4, a7, 7e, 3d, 64, 5d, 19, 73, 60, 81, 4f, dc, 22, 2a, 90, 88, 46, ee, b8, 17, de, 5e, 0b, db, e0, 32, 3a, 0a, 49, 06, 24, 5c, c2, d3, ac, 62, 91, 95, e4, 79, e7, c8, 37, 6d, 8d, d5, 4e, a9, 6c, 56, f4, ea, 65, 7a, ae, 08, ba, 78, 25, 2e, 1c, a6, b4, c6, e8, dd, 74, 1f, 4b, bd, 8b, 8a, 70, 3e, b5, 66, 48, 03, f6, 0e, 61, 35, 57, b9, 86, c1, 1d, 9e, e1, f8, 98, 11, 69, d9, 8e, 94, 9b, 1e, 87, e9, ce, 55, 28, df, 8c, a1, 89, 0d, bf, e6, 42, 68, 41, 99, 2d, 0f, b0, 54, bb, 16

Kuznyechik

Kuznyechik is a symmetric block cipher with a block size of 128 bits and a key length of 256

bits. The number of rounds is 10.

This cipher is approved as the standard GOST R 34.12-2015 RF [10]. The standard came into effect on January 1, 2016. The code was developed by the Center for Information Protection and Special Communications of the FSB of Russia with the participation of InfotecsOJSC.

The nonlinear transformation of the Kuznyechik algorithm is specified by the substitution S, which is a one-dimensional array of 256 bytes.

SM4

In January 2006, the SMS4 block cipher algorithm, developed by Professor Shu-Wang Lu, was declassified. This algorithm is used in China as the national standard for wireless local area networks. After the release of the GM / T 0002-2012 SM4 Block Cipher Algorithm standard on March 21, 2012, the algorithm was officially renamed SM4.

The block size of the SM4 encryption algorithm is 128 bits, the key size is also 128 bits. The number of rounds is 32 [11].

The nonlinear transformation is implemented by substitution $\boldsymbol{\tau}$, it is also specified in the form of a table of 8x256 bits.

BelT

BelT is the state standard for symmetric encryption and integrity control of the Republic of Belarus (STB 34.101.31-2007). Adopted as a preliminary standard in 2007. Introduced as the final standard in 2011.

The cipher has the block length of 128 bits and the key length of 256 bits.

Encryption is performed by eight rounds of transformations applied to the input block [12].

The nonlinear transformation of this algorithm is carried out through the substitution S. The substitution itself is specified in the form of a fixed table with the size of 8x256.

Kalyna

The Kalyna algorithm is a symmetric block cipher. It supports blocks of 128, 256, or 512 bits; the key length is equal to or doubled the block size.

Kalyna was adopted as the national encryption standard of Ukraine in 2015 (standard DSTU 7624: 2014) after the All-Ukrainian competition of cryptographers. The algorithm design is based on the Rijndael Algorithm (AES). The Kalyna algorithm retains all the basic operations of the Rijndael algorithm. The main differences are in the use of different S-boxes (4 different S-boxes are used), generated randomly, instead of identical S-boxes, and in the use of alternating addition with cyclic subkeys modulo 2 and modulo 2^{64} [13].

By means of a developed program, the values of the nonlinearity of the S-boxes of the above-considered encryption algorithms were calculated and analyzed.

Let$S = (f_0, f_1, \dots, f_{m-1})$be some$n \times m$ substitution, wheref_i is a Boolean function of$n$ variables. We denote byg_i the set of all linear combinations off_i. Thenthenonlinearity$S$is equal to [6]:

$$NL(S) = \min\left(NL(g_j)\right), 0 < j < 2^m$$

Table 2 below shows the results of calculating the value of the nonlinearity of the S-boxes.

Table 2.

Algorithms	Nonlinearityvalue
AES	112
Kuznyechik	116
SM4	112
BelT	110
Kalyna – S0	112
Kalyna – S1	110
Kalyna – S2	110
Kalyna – S3	112

Dynamic S-boxes

Recently, various scientific journals have published works on the creation of dynamic S-blocks of the encryption algorithm. In [3, 14, 15], methods of dynamic creation of S-boxes in various ways are considered. In [3], dynamic S-boxes are created based on the S-box of the AES algorithm. In [14], [15] S-boxes are generated dynamically using pseudo-random sequence generators.

In our work, we investigated one of the methods for creating dynamic S-boxes. Consider the developed method for generating S-boxes based on known and proven S-boxes. The idea behind this method is to generate dynamic S-boxes that modify with each change in the secret key. The main advantage of this approach is that S-boxes are random, key-dependent, and unknown in advance since both linear and differential cryptanalysis require known S-boxes.

The well-known S-box of the AES algorithm was chosen for the study. The values of this S-box are presented in Table 1. From the encryption master key, using various transformations, we obtain one byte, for example, by summing all the bytes of the master key modulo 2. The same byte will be used as a constant in the next affine transformation:

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \end{bmatrix} + \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \end{bmatrix}$$

where$a_i$arethebitsofthe S-boxbyte, $c_i$are the constant bits (the byte obtained from the master key), $b_i$are the bits of the new S-box byte

As a result of this affine transformation, we get a new S-box that differs from the original S-box. The resulting S-block will be used in the process of nonlinear transformation of the encryption algorithm

In the course of the study, it was revealed that after such a transformationpreserved the nonlinearity properties of the S-boxes. The experimental results are shown in Table 3.

Table 3.

Algorithm	Initialvalueofnonlinearity	Non-linearity value after transformation
AES	112	112
Kuznyechik	116	116
SM4	112	112
BelT	110	110
Kalyna-S0	112	112
Kalyna-S1	110	110
Kalyna-S2	110	110
Kalyna-S3	112	112

Also, various properties of dynamically generated S-boxes were investigated using special programs developed in our laboratory.

Table 4 shows the results of the study of the dynamic S-box of the AES algorithm, obtained with the value of C = 36, 109, and 221.

Next, we investigated various cryptographic properties of the created S-box. We consider a nonlinear node effective if it provides resistance to currently known cryptographic analysis methods.

In most of the well-known works in the field of analysis and synthesis of nonlinear substitution nodes of modern symmetric block ciphers, the mathematical apparatus of cryptographic Boolean functions is used [16]. Moreover, each S-box is represented by a set of component Boolean functions, the properties of which characterize the efficiency of the nonlinear substitutionnode. The following are used as the main criteria and performance indicators: balance and nonlinearity of component Boolean functions; correlation immunity; propagation criterion; algebraic degree; the value of the autocorrelation function.

Let us introduce the basic concepts and definitions used in what follows to assess the efficiency of nonlinear substitution nodes [16].

A Boolean function f of *n* variables is a function that maps from the field $GF(2^n)$ of all binary vectors $x = (x_1, \ldots, x_n)$ of length *n* into the field $GF(2)$. Boolean functions are usually represented in algebraic normal form. The field $GF(2^n)$ consist of 2^n vectors α_i: $\alpha_0 = (0,\ldots,0,0)$, $\alpha_1 = (0,\ldots,0,1)$,..., $\alpha_{2^n-1} = (1,\ldots,1,1)$, $\alpha_i \in V_n$, where V_n where V_n is a vector space in $GF(2^n)$.

The sequence of a function f is a (1,-1) sequence defined as $\left((-1)^{f(\alpha_0)}, (-1)^{f(\alpha_1)}, \ldots, (-1)^{f(\alpha_{2^n-1})}\right)$.

The truth table of a function f is a (0,1) sequence defined as $(f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1}),)$.

The sequence of a function *f* is *balanced* if its (0,1) sequence ((1,-1) sequence) contains the same number of zeros and ones (ones and minus ones). A function *f* is balanced if its sequence is balanced.

Equivalent definition: A function *f* over $GF(2^n)$ is balanced if its outputs are equiprobable:

$$|\{x|f(x) = 0\}| = |\{x|f(x) = 1\}| = 2^{n-1}.$$

An affine function f is a function of the form $f = a_1x_1 \oplus \ldots \oplus a_nx_n \oplus c$, where a_j, $c \in GF(2)$, $j = 1, 2,\ldots, n$. A function *f* is called linear if $c = 0$.

The Hamming weight of the vector α, denoted as $W(\alpha)$, is the number of ones in the vector (sequence). *The Hamming distance d(f,g)* between the sequences of two functions *f* and *g* is the number of positions in which the sequences of these functions are different.

The nonlinearity of transformation (N_S) is the minimum Hamming distance between the output sequence *S* and all output sequences of affine functions over some field:

$$N_s = min\{d(S, j)\}, \text{ where j is the set of affine functions.}$$

The nonlinearity of a function (N_f) is the minimum Hamming distance N_f between the function *f* and all affine functions over $GF(2n)$:

$$N_f = min\{d(f, j)\}, \text{ where j is the set of affine functions.}$$

For an arbitrary function *f* the nonlinearity N_f over $GF(2^n)$ can reach:

$$N_f \leq 2^{n-1} - 2^{\frac{n}{2}-1}.$$

For a balanced function *f* over $GF(2^n)$ ($n \geq 3$) the nonlinearity N_f can reach:

$$N_f \leq \begin{cases} 2^{n-1} - 2^{\frac{n}{2}-1} - 2, n = 2k \\ \left[\!\left[2^{n-1} - 2^{\frac{n}{2}-1}\right]\!\right], n = 2k = 1 \end{cases}$$

where $\|x\|$ is the maximum even integer less than or equal to x.

A function *f* possesses *correlation immunity* of order *k* if the output sequence of the function $y \in Y$ is statistically independent of any subset of *k* input coordinates:

$$\forall\{x_1, \ldots x_k\} P(y \in Y/\{x_1, \ldots x_k\} \in X) = P(y \in Y).$$

Table 4.

Properties	S-box-1	S-box-2	S-box-3
ParameterCvalues	36	109	221
Hammingweight	128	128	128
Balance	True	True	True
Hammingweight - Minimum	128	128	128
Hammingdistance	128	128	128
Nonlinearity(min)	112	112	112
Nonlinearity(max)	144	144	144

Correlationvalue(min)	-32	-32	-32
Correlationvalue(max)	32	32	32
|AC| min	-32	-32	-32
|AC|max	32	32	32
| SSI |min	133120	133120	133120
| SSI |max	133120	133120	133120
SAC	False	False	False
Propagationcriterion	N/A	N/A	N/A
CI	N/A	N/A	N/A
t-stability	N/A	N/A	N/A

The resulting S-box will be used for encryption. Each encryption process will have its own S-box, the values of which are not known in advance. This property, in turn, complicates both differential and linear cryptanalysis. The decryption uses the inverse substitution box. The inverse S-box is also generated dynamically from the created main S-box. There is no need to store the inverse S-box.

Conclusion

Work in this area continues and the results obtained in the study will be used in the development of an encryption algorithm and research work that is being carried out at the Institute of Information and Computational Technologies of the Committee of Science of the Ministry of Education and Science of the Republic of Kazakhstan. Further, as a continuation of this work, other cryptographic properties of dynamically generated S-boxes will be investigated.

Acknowledgment

The research work was carried out within the framework of the project AP08856426 "Development and study of an encryption algorithm and the creation of a hardware and software complex for its implementation" at the Institute of Information and Computational Technologies.

References

[1] Gorbenko I.D., Dolgov V., Olejnikov R.V., Ruzhencev V.I., Mihajlenko M.S., Gorbenko YU.I., Razrabotka trebovanij i princip proektirovaniya perspektivnogo simmetrichnogo blochnogo algoritma shifrovaniya. *Izvestiya yuzhnogo federal'nogo universiteta. Tekhnicheskie nauki*. 2007. 76(1). 183-189. (in Russian)

[2] Shennon K. Raboty po teorii informacii i kibernetike, Moscow., IL, 1963, P.333-369. (in Russian)

[3] Kazlauskas K., Kazlauskas J., Key-dependent S-box generation in AES block cipher system. *Informatica*. 2009. 20. 23-34

[4] Olejnikov R.V., Kazimirov A.V. Vybor S-blokov dlya simmetrichnyh kriptograficheskih algoritmov na osnove analiza algebraicheskih svojstv. *Vesn. Hark. nac. un-tu. Ser. Matematichne modelyuvannya. informacijni tekhnologii*. Avtomatizovani sistemi upravlinnya, Kharkiv., 2010. 925. 79-86. (in Russian)

[5] Gorbenko I. D., Gorbenko YU. I. Prikladna kriptologiya,. - Kharkiv.: Fort, 2012. – 870 p. (in Russian)

[6] Kazimirov A.I. Metody i sredstva generacii nelinejnyh uzlov zameny dlya simmetrichnyh kriptoalgoritmov. Dissetraciya na soiskanie uchenoj stepeni kanditata tekhnicheskih nauk. Har'kovskij nacional'nyj universitet radioelektroniki. Kharkiv, 2013 (in Russian)

[7] Specification for the Advanced Encryption Standard (AES) [Electronic resource]. Available at: URL: http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf. (Accessed 02.05.2021)

[8] Daemen J., Rijmen V. AES Proposal: Rijndael [Electronic resource]. Available at:– URL: http://csrc.nist.gov/archive/aes/rijndael/Rijndael-ammended.pdf (Accessed 02.05.2021)

[9] Babenko L.K., Ishchukova E.A. Sovremennye algoritmy blochnogo shifrovaniya i metody ih analiza. – Moscow: Gelios APV, 2006. - 376 p. (in Russian)

[10] Kriptograficheskaya zashchita informacii Blochnye shifry -GOST R 34.12-2015 [Electronic resource]. Available at: – URL: https://tc26.ru/standard/gost/GOST_R_3412-2015.pdf (Accessed:

02.05.2021) (in Russian)

[11] SMS4 Encryption Algorithm for Wireless Networks. Translated and typeset by Whit field Diffie of Sun Microsystems and George Ledin of Sonoma State University, 15 May 2008

[12] Agievich S.V., Galinskij V.A., Mikulich N.D., Harin YU.S. Algoritm blochnogo shifrovaniya BelT / Upravlenie zashchitoj informacii, Vol. 6, №4, 2002. - P.407-412. (in Russian)

[13] Kuznecov O.O., Ivanenko D.V., Kolovanova E.P. Modelirovanie perspektivnogo blochnogo shifra «Kalina». *Prikladna radioelektronika: nauk.-tekhn. zhurnal*. 2014. 13(3). 201-207. (in Russian)

[14] Cassal-Quiroga B. B., Campos-Cantin E.. Generation of Dynamical S-Boxes for Block Ciphers via Extended Logistic Map. *Mathematical Problems in Engineering*. 2020. 2. 12. https://doi.org/10.1155/2020/2702653

[15] Jiqiang Lu, Hwajung Seo. A Key Selected S-Box Mechanism and Its Investigation in Modern Block Cipher Design, Security and Communication Networks, vol.2020, Article ID1457419, 26 pages, 2020. https://doi.org/10.1155/2020/1457419

[16] Kuznetsov A.A., Moskovchenko I.V., Prokopovich-Tkachenko D.I., Kuznetsova T.Yu.. Heuristic methods of gradient search for cryptographic Boolean functions. *Kharkiv National University of Radio Electronics, Radio engineering journal,* 2018. 195. 150-164.

Исследование динамической таблицы подстановок

Н.А.Капалова[1], **А.Хаумен**[1]
[1]Институт информационных и вычислительных технологий КН МОН
Алматы, Казахстан
e-mail: kapalova@ipic.kz, haumen.armanbek@gmail.com

Аннотация: В работе исследуется ранее разработанный метод генерации динамической таблицы подстановок. Данный метод позволяет генерировать таблицы подстановок (S-блоки), изменяющиеся в зависимости от значений некоторого параметра, получаемых из секретного ключа алгоритма. Генерируемые S-блоки устойчивы к линейному и дифференциальному криптоанализу, так как они являются случайными, заранее не известны и зависят от секретного ключа. Полученные динамические таблицы подстановок были исследованы на нелинейность, а результаты сравнивались с S-блоками известных симметричных блочных алгоритмов, таких как AES, Кузнечик, SM4, BelT и Калина.

Ключевые слова: алгоритм шифрования, таблица подстановок, динамический S-блок, нелинейное преобразование.

Динамикалық алмастыру кестелерін зерттеу

Н.А.Капалова[1], **А.Хаумен**[1]
[1]Ақпараттық және есептеуіш технологиялар институты, БжҒМ ҒК, Алматы, Қазақстан
e-mail: kapalova@ipic.kz, haumen.armanbek@gmail.com

Аннотация: Аталмыш жұмыста динамикалық алмастыру кестелерін жасаудың бұрынырақ құрастырылған тәсілі зерттелінеді. Бұл тәсіл құпия кілттен алынатын қандай да бір параметрдің мәніне байланысты өзгеріп отыратын динамикалық алмастыру кестелерін (S-блоктар) жасауға мүмкіндік береді. Құрастырылған S-блоктар сызықтық және дифференциалдық криптоталдауларға төтеп бере алады, себебі, алынған S-блоктар кездейсоқ, алдын ала белгісіз және олар құпия кілттерге тәуелді болмақ. Алынған динамикалық алмастыру кестелерінің сызықты емес болу қасиеті зерттелініп, нәтижелері AES, Кузнечик, SM4, BelT және Калина сияқты танымал симметриялы блоктық алгоритмдердің S-блоктарымен салыстырылып қарастырылды.

Түйін сөздер: шифрлеу алгоритмі, алмастыру кестесі, динамикалық S-блок, сызықты емес түрлендіру.