ATLANTIS PRESS

Advances in Engineering Research, volume 158

International Conference on Aviamechanical Engineering and Transport (AviaENT 2018)

Analysis of data representation algorithms used in automated control systems in railway transport

N.A. Gainulin
Irkutsk State University
Irkutsk, Russia

O.V. Kuzmin
Irkutsk State University
Irkutsk, Russia
E-mail:quzminov@mail.ru

***Abstract*—In this paper, the authors consider the use of data representation algorithms in automated control systems in railway transport. Examples of several algorithms illustrate the approaches that are used when creating such solutions. The issue of the stability of data representation algorithms in the methods of linear analysis is considered separately. In particular, the authors pay special attention to such characteristic of these algorithms as non-linearity. The article considers the methods for calculating the nonlinearity of a vector Boolean function. It presents an algorithm for calculating nonlinearity based on the properties of the Sylvester-Hadamard matrices. The authors give estimates of computational complexity for algorithms, as well as estimation of the memory size of the algorithm necessary for the operation.**

Keywords—Automated control systems; Lightweight cryptography; Block ciphers; Nonlinearity; Boolean functions

I. INTRODUCTION

In the modern world, the role of automated control systems (ACS) increases every year. The use of automated management systems allows increasing the efficiency of process control by increasing the efficiency of management and accelerating the performance of individual data collection and processing operations. The railway transport industry was not an exception.

The ACS used in the railway sector is assigned the following tasks:

• adjusting the processes of moving the rolling stock along the railroad track;

• creating a two-way communication between the dispatcher and the driver;

• control over the processes of loading wagons;

• tracking of passenger traffic;

• immediate processing of current information on the state of rolling stock, railway tracks;

• adjustment of the train traffic.

An example of the scenario of the application of the automated control system is the receiving and processing of information that is read by a sensor installed near the tracks, with RFID tags attached to the cars of the passing train. The received data can contain information about the state of the car and cargo.

Modern automated control systems are complex hardware-software complexes consisting of many components. The effectiveness of such a complex depends on the correctness and integrity of the transmitted information between the elements of the system. In some cases, it is required to ensure that information cannot be obtained by a third party. To protect information from unauthorized access or changes to the automated management system, solutions based on secial algorithms are available that can work in conditions of limited resources [1, 2].

In the last decade there have been many works devoted to this area. Such attention is due to the development of the idea of the "Internet of things", which is a network connecting objects of various types. Examples of such objects are household appliances, vehicles, smart sensors and radio frequency identification tags. Since a lightweight cryptographic algorithm should work with devices with low processing power and a small amount of memory, but at the same time show a speed that is acceptable from a practical point of view, the cipher developer faces the difficult task of choosing the optimal solution in terms of performance, cost, and stability [3, 4].

At the heart of many information security systems lie algorithms that are called block algorithms. Such algorithms break open text into blocks and convert each block into cipher text. Next, several algorithms of data representation will be considered.

II. PRESENT

The specification of the block cipher PRESENT was published in 2007 [5]. This algorithm works with blocks of 64 bits long and supports keys of 80 or 128 bits long. The number

Copyright © 2018, the Authors. Published by Atlantis Press.
This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

of rounds is 31. This algorithm was developed for use in RFID tags and measurement sensors.

This algorithm refers to algorithms of the type of a substitution-permutation network. A key feature of this algorithm is the use of bit permutations. These operations take a relatively large amount of CPU time for software implementation, but are very simple in hardware. Figure 1 shows a scheme of two rounds of PRESENT algorithm.

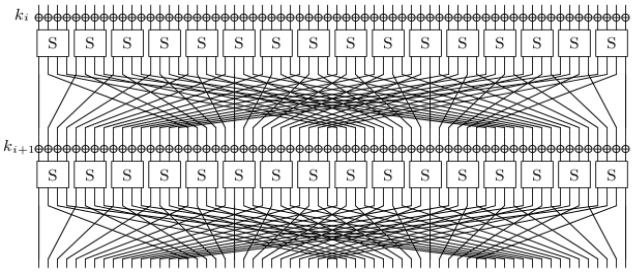

Fig. 1. The schematics of the PRESENT algorithm round.

The PRESENT scheme served as the basis for many other algorithms, such as the block algorithm LED and hash functions PHOTON, DM-PRESENT and SPONGENT.

III. SIMON AND SPECK

These two block algorithms are developed by the US National Security Agency and submitted to the public in 2013 [6]. Both are Feistel networks and belong to the family of so-called ARX-algorithms, whose round functions contain logical AND operations, cyclic shift and modulo addition.

SIMON is designed for hardware implementation. Table 1 shows the characteristics of various SIMON cipher variants.

TABLE I. PARAMETERS OF SIMON BLOCK CIPHER

Parameters of SIMON block cipher		
Block size	***Key length***	***Number of rounds***
32	64	32
48	72 / 96	36
64	96 / 128	42 / 44
96	96 / 144	52 / 54
128	128 / 192 / 256	68 / 69 / 72

Figure 2 shows the schematic of the SIMON cipher round.

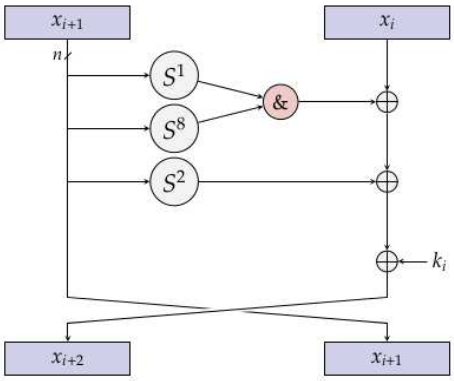

Fig. 2. The schematics of the SIMON algorithm round.

The SPECK code is optimized for application in the form of software implementation. Table 2 gives the characteristics of various variations of the SPECK algorithms.

TABLE II. PARAMETERS OF SPECK BLOCK CIPHER

Parameters of SPECK block cipher		
Block size	***Key length***	***Number of rounds***
32	64	22
48	72 / 96	22 / 23
64	96 / 128	26 / 27
96	96 / 144	28 / 29
128	128 / 192 / 256	32 / 33 / 34

Figure 3 shows the schematics of the SPECK algorithm round.

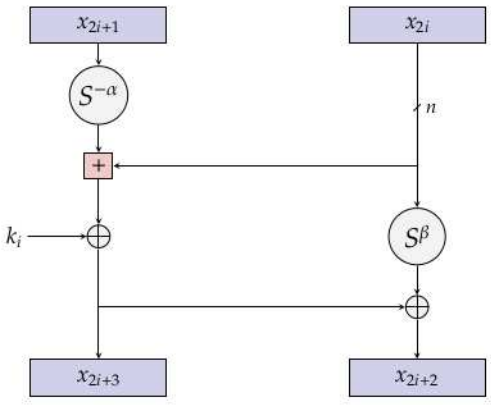

Fig. 3. The schematics of the SPECK algorithm round.

IV. SKINNY

The SKINNY algorithm is introduced in 2016 [7] as an alternative to the SIMON and SPECK algorithm. This code represents a substitution-permutation network. A key feature of this algorithm is that the round key depends on both the master key and some internal state of the algorithm that is generated during the data representation process.

Table 3 shows the characteristics of different variations of the SKINNY algorithms.

TABLE III. PARAMETERS OF SKINNY BLOCK CIPHER

Parameters of SKINNY block cipher		
Block size	**Key length**	**Number of rounds**
64	64 /128 /192	32 / 36 / 40
128	128 / 256 /384	40 / 48 / 56

Figure 4 shows the schematic of the SKINNY round.

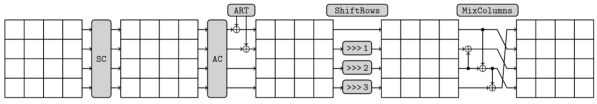

Fig. 4. The schematics of the SKINNY algorithm round.

V. METHODS FOR CALCULATING THE NONLINEARITY OF ROUND FUNCTIONS OF DATA REPRESENTATION ALGORITHMS

The large variety of designs used in the construction of data representation algorithms results in the fact that researchers need objective criteria for evaluating and comparing the properties of these algorithms. The main characteristic for any of these algorithms is the resistance to linear analysis.

The mathematical Boolean transformation model is the vector Boolean function. In turn, the value of the nonlinearity of the Boolean function reflects how well this function is approximated by linear functions. The high value of the round transformation nonlinearity characterizes the stability of the entire algorithms to the methods of linear analysis. The book [8] describes the problems of applying Boolean functions with a high nonlinearity value.

Let $E=\{0,1\}$ be a field of characteristic 2, Z be a field of integers, and E_2^n be a vector space. let us suppose that $u, v \in E_2^n$ is the scalar product of vectors u and v is a quantity $\langle u, v\rangle = u_1 v_1 \oplus \ldots \oplus u_n v_n$, where $\oplus$ is the modulo 2 addition.

Let us consider the Walsh-Hadamard transform, which is defined by formula

$$W_f(u)=\sum_{v\in E_2^n}(-1)^{f(v)\oplus\langle u,v\rangle}. \quad (1)$$

The values of $W_f(u)$ are called the spectral coefficients of the function f in the literature.

Let f and g be Boolean functions depending on the same number of variables, and the Hamming distance between Boolean functions is the quantity. Let be the set of linear Boolean functions that depend on variables. The nonlinearity of a function is given by

$$nl(f)=\min_{g\in A_n} dist(f,g). \quad (2)$$

It is known that the nonlinearity (2) and the spectral coefficients (3) of the Boolean function are related by

$$nl(f)=2^{n-1}-\frac{1}{2}\max_{u\in E_2^n}|W_f(u)|. \quad (3)$$

An ordered set m of Boolean functions of n variables $F(x)=(f_1(x),\ldots f_m(x))$ is called a vector Boolean function, and functions $f_i(x)$ are called coordinate functions. The nonlinearity of a vector Boolean function $F(x)$ is given by

$$nl(F)=\min_{c\in E_2^m\setminus\{0\}}(c_1 f_1(x)\oplus\ldots\oplus c_m f_m(x)). \quad (4)$$

Taking into account expressions (1), (3) and (4), we can conclude that to calculate the nonlinearity of a vector Boolean function $F(x)$, it is required to calculate the nonlinearity of 2^m-1 Boolean functions and to find the least one among them. The calculation of the nonlinearity of a Boolean function implies finding the 2^n spectral coefficients, each of which requires the calculation of 2^n summands. Each summand in (1) contains the scalar product of two vectors from E_2^n which requires $2n-1$ operations of addition and multiplication to calculate it. Thus, the complexity of calculating the nonlinearity of a vector Boolean function $F(x)$ can be estimated as $O(2^{m+2n}n)$.

Let us consider the matrices

$$H_1=\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix},\quad H_n=\begin{pmatrix}H_{n-1} & H_{n-1}\\ H_{n-1} & -H_{n-1}\end{pmatrix}. \quad (5)$$

Matrices of the form (5) are called Sylvester-Hadamard matrices. Their peculiarity is that if you order vectors from E_2^n in lexicographic order, then the element at the intersection of the i-th row and the j-th column of the matrix H_n is equal to $(-1)^{\langle u_i, u_j\rangle}$, where $u_i, u_j \in E_2^n$. Having constructed the matrix H_n of the necessary order, it is possible to obtain values of summands of the sum from (1),

without directly calculating scalar products of all possible pairs of vectors from E_2^n.

Applying the operation of elementwise modulo 2 addition to the truth table of the function f depending on the n variables and the row of the matrix H_n corresponding to the vector u, we get a set of length 2^n consisting of the values $f(v_i) \oplus \langle u, v_i \rangle$, where $v_i \in E_2^n$. It should be noted that the execution time of the elementwise addition modulo 2 operation in modern processors is comparable with the time of other elementary operations, such as addition and multiplication. We introduce the function $S_k(x): E_2^k \to \mathbf{Z}$, where $k < 2^n$:

$$S_k(x) = \sum_{i=1}^{k} (-1)^{x_i} . \quad (6)$$

If you save a set of function $S_k(x)$ values for all vectors $x \in E_2^k$ in memory, you can significantly simplify the calculation of the spectral coefficient of the function f. We divide the set of values of the function f and the matrix H_n row (that is of interest to us) into disjoint groups by 2^k elements, we obtain 2^{n-k} pairs of groups. For each pair we perform an elementwise addition modulo 2 operation and find in memory the value of the function $S_{2^k}(x)$ for the result obtained. Summing the found values of the function $S_{2^k}(x)$, we obtain the required spectral coefficient. The complexity of this algorithm can be estimated as $O(2^{m+2n-k+1})$.

A number of properties (0, 1) of matrices are obtained in [9, 10].

When implementing this algorithm, you should consider the amount of memory that is required to store the necessary structures. The minimum amount of memory that the matrix H_n occupies is equal to 2^{2n} bits. An array with function $S_k(x)$ values requires a minimum 2^k of memory locations. The implementation of the algorithm described above is considered in [11].

VI. CONCLUSIONS

The use of algorithms of data representation in automated railway management systems contributes to ensuring the protection of information transmitted between the elements of the system, as well as reducing the cost of implementing this system. However, one should pay special attention to what algorithms are used in the system and how resistant they are to linear analysis.

References

[1] O.V. Kuzmin, K.P. Orkina, “Creation of the codes correcting errors by means of the Pascal type triangle”, Bull. of Buryat State Univ., vol. 13, pp. 32-39, 2006.

[2] O.V. Kuzmin, B.A. Starkov, “Binary matrices constructed using Pascal's triangle, and noise-immune coding”, Modern Techn. Syst. Anal. Model., vol. 1, pp. 112-117, 2016.

[3] N.A. Gainulin, O.V. Kuzmin, “The use of vector Boolean functions with a high value of nonlinearity in cryptography”, Transp. Infrastruct. of the Siber. Reg., vol. 1, pp. 281-285, 2016. (7th Internat. Sci. and Pract. Conf. Transport Infrastructure of the Siberian Region, p. 672, 2016).

[4] N.A. Gainulin, “Algorithm for calculating the nonlinearity of a vector Boolean function”, Inform. technol. and problems of math. modeling of complex syst., vol. 14, pp. 27-32, 2015.

[5] A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. B. Robshaw et al., “PRESENT: An Ultra-Lightweight Block Cipher” , Crypt. Hardw. and Embed. Syst. - CHES, pp. 450-466, September 2007 (9th International Workshop, 468 p.)

[6] R. Beaulieu, D. Shors, J. Smith, S. Treatman-Clark, B. Weeks, L. Wingers, “The Simon and Speck Families of the lightweight block ciphers”. Cryptology ePrint Archive. URL: https://eprint.iacr.org/2013/404.pdf. (Accessed: Jun 08, 2018)

[7] C. Beierle, J. Jean, S. K. Colbl, G. Leander, A. Moradi, T. Peyrin et al., “The SKINN Family of Block Ciphers and its Low-Latency Variant MANTIS”. Advanc. in Crypt. - CRYPTO 2016, Springer, 2016, pp. 123-153.

[8] O. A. Logachev, A. A. Salnikov, and V. V. Yashchenko, Boolean Functions in Coding Theory and Cryptography. Providence, Rhode Island: American Mathematical Society, 2011, p. 334.

[9] O. V. Kuzmin, L.G. Evsevleeva, “About depth of symmetrical (0, 1)-matrices”, Modern Techn. Syst. Anal. Model., vol. 2, pp. 57-59, 2010.

[10] O. V. Kuzmin, B. A. Starkov, “Binary matrixes based on Pascal’s triangle’s arithmetics and char sequences”, Bull. of Irkutsk State Univ. Ser. Mathematics, vol. 18, pp. 38-47, 2016.

[11] O.V. Kuzmi, N.A. Gainulin, “Methods of computer simulation of Boolean functions with the maximum value of nonlinearity”, Modern Techn. Syst. Anal. Model., vol. 1, pp. 123-127, 2017.