Design of Boolean Function from a Great Number of Variables Satisfying Strict Avalanche Criterion

E.G.BARDIS*, N.G.BARDIS*, A.P.MARKOVSKI*, A.K.SPYROPOULOS**

*Department of Computer Science
National Technical University of Ukraine
(Kiev Polytechnic Institute)
Glyfada-Athens
Tainarou 66
16561
HELLAS
**Department of Mathematics
University of Athens
e-mail: bardis@akcecc.kiev.ua

Abstract:-Cryptoresistance of a broad class of cryptographic algorithms is determined by their correspondence to some special criteria of bit transform Boolean functions being implemented in these algorithms. One of such criteria is a strict avalanche criterion (SAC). Obtaining of Boolean functions satisfying this criterion is an important constituent of cryptoresistant algorithm design. The existing methods of SAC-function obtaining which utilize in the explicit or implicit form the truth tables of a function being formed are practically useless for synthesis of SAC-functions from a great number of variables, because they demand memory capacity in proportion to 2^n (n is the number of variables).
This paper presents investigation of Boolean SAC-function properties and suggests a new method for function obtaining without making use of the truth tables. The method deals with the algebraic normal form whose storage demands memory capacity of many orders lower comparing to that for truth table storage. The method is helpful both for obtaining ordinary SAC-functions and for synthesis of high-order SAC-functions.The formalized procedure for construction of zero and higher orders SAC-functions is expounded in details, examples of functions design are given.
Key-Words:- Cryptography, Boolean Functions, SAC functions.
CSCC'99 Proceedings - Pages 3111-3116

1 Introduction

A mathematical problem which is insolvable with analytical methods and whose only practical way to be solved is searching provides the basis for all cryptographic algorithms. So, the problem of large number factoring lies in the foundation of the well-known algorithm RSA. The algorithm ElGamal gets its security from the difficulty of calculating discrete logarithms in a finite field. Security properties of cryptographic schemes based on combination of confusion and diffusion are determined by an analytically intractable problem of finding the roots of a system of nonlinear Boolean functions. Such algorithms as DES, IDEA, SHA and many others widespread in practice belong to that class. In this case "break" of a cryptographic algorithm is equivalent to solution of the corresponding system of nonlinear Boolean equations. The only practical way for nonlinear Boolean equations solution is searching. The search area may be decreased significantly by application of different expedients based on taking into account the special features of Boolean functions constructing the system of nonlinear Boolean equations. There is a certain identity between the methods for decrease of searching at finding the roots of a system of nonlinear Boolean equations and the methods for break of cryptographic algorithms. For example, the known method of cryptanalysis [3] is identical in fact with the method of linear approximation at finding the roots of nonlinear Boolean equation systems. The method makes it possible to diminish the search area if the Boolean functions possess low non-linearity.

If Boolean functions constituting a system equivalent to a cryptographic algorithm satisfy certain properties, the search area at solution of the corresponding Boolean equation systems can not be diminished and, consequently, efficiency of all break methods will also be minimal. Obviously, such

cryptographic algorithms will be of maximal resistance to breaks.

Properties of Boolean functions providing maximal resistance to breaks were determined in the result of S-boxes DES [1,4,6] investigation and formulated as the following criteria: a function must have a high nonlinear order, must be 0/1 balanced, complete and satisfy a strict avalanche criterion.

The Strict Avalanche Criterion (SAC) was introduced by Webster and Travares [6] in connection with study of design of S-boxes. A Boolean function is said to satisfy SAC if complementing a single input bit results in changing the output bit with probability of one half.

2 Problem Formulation

Development of methods for obtain in Boolean functions with properties mentioned above is an essential problem of cryptographic algorithm working-out.

Formally, a Boolean function $f(x_1,\dots,x_n)$ satisfies SAC if the $g(x_1,\dots,x_n)=f(x_1,\dots x_i,\dots,x_n)\oplus f(x_1,\dots,x_i\oplus 1,\dots,x_n)$ is balanced for any $i\in\{1,\dots,n\}$.

Forre [1] extended the concept of SAC by defining higher order strict avalanche criteria. The Boolean function $f(x_1,\dots,x_n)$ is said to satisfy the strict avalanche criterion of k order (designed as SAC(k)) if any function obtained from $f(x_1,\dots,x_n)$ by keeping any *k* input bits constant satisfies SAC.

A number of methods for solution of this problem has been put forward by now. In particular, SAC-functions are supposed to be obtained by application of Walsh transforms [1]. In study [2] SAC-functions are suggested to be obtained from matrix transformations. All these methods imply explicitly or implicitly utilization of the truth tables. This impose a processing restriction on the number of variables. So, a SAC-function from 100 variables can not be obtained by the known methods because it will demand memory capacity of 2^{100} bits that is impossible for implementation in modern computers. Therefore, methods for SAC-function design which do not make use of the truth tables but operate only with the function normal algebraic form should be worked out. Such a method was suggested by B.Preneel [5], however, the functions formed with its application have low non-linarity and the maximal degree of the terms is equal to 2.

From the tend to enhancement of cipherblock capacity the problem of obtaining nonlinear SAC functions of a great number of variables appears to be one of the central to design of cryptoresistant algorithms.

3 Problem Solution

The theoretical basis for the suggested approach to construction of an analytical representation of Boolean SAC-functions is the following theorem.

Any Boolean function $f(x_1,\dots,x_n)$ may always be represented in the form:

$$f(x_1,\dots,x_n)=x_i\cdot\varphi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)\oplus \oplus\psi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n),i=1,\dots,n \quad (1)$$

Theorem: In order a Boolean function $f(x_1,\dots,x_n)$ to correspond to SAC, it is necessary and sufficient that each of n functions $\varphi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)$, $i=1,\dots,n$ be balanced.

Proof: By definition, the Boolean function $f(x_1,\dots,x_n)$ corresponds to SAC if for any $i=1,\dots,n$ the Boolean function $f(x_1,\dots,x_i,\dots,x_n)\oplus f(x_1,\dots,1\oplus x_i,\dots,x_n)$ is balanced and, with making use of (1), it may be transformed into the following form:

$$f(x_1,\dots,x_i,\dots,x_n)\oplus f(x_1,\dots,1\oplus x_i,\dots,x_n)$$
$$=x_i\cdot\varphi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)\oplus$$
$$\oplus\psi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)\oplus$$
$$\oplus(1\oplus x_i)\cdot[\varphi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)]\oplus$$
$$\oplus\psi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)=$$
$$=\varphi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)$$

Thus, the condition of the Boolean function $f(x_1,\dots,x_n)$ satisfaction to SAC is equivalent to the condition of balansedness for all n Boolean functions $\varphi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)$, which is what had to be proved.

Making use of the theorem above, the following practically important corollary may be proved:

Corollary :The Boolean function $f(x_1,\dots,x_n)=\phi(x_1,\dots,x_n)\oplus\delta(x_1,\dots,x_n)$ corresponds to SAC if the function $\phi(x_1,\dots,x_n)$ corresponds to SAC, and the function $\delta(x_1,\dots,x_n)$ is linear or, what is the same: addition of a linear function to a SAC-one does not break correspondence of the latter to SAC.

Proof: If the Boolean function $\phi(x_1,\dots,x_n)$ corresponds to SAC, then *n* representation of the form (1) may be always indicated:

$$\phi(x_1,\dots,x_n)=x_i\varphi_i'(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)\oplus \oplus\psi_i'(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n),$$

in this case each of n functions $\varphi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)$, $i=1,\dots,n$, is balanced on the set of variables $x_1,\dots,x_{i-1},x_{i+1},\dots,x_n$, as it follows from the theorem above. The function $f(x_1,\dots,x_n)$ may also be given by n representations of the form: $f(x_1,\dots,x_n)= =x_i\cdot\varphi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)\oplus \oplus\psi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)$. Since the function $\delta(x_1,\dots,x_n)$ is linear, it does not comprise products and, consequently, its addition to $\phi(x_1,\dots,x_n)$ will not

change any of n functions $\varphi_i'(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)$, $i=1,\dots,n$ (only the function $\psi_i'(x_1,\dots,x_{i-1},\ x_{i+1},\dots,x_n)$ will be changed). Hence it is true for all $i=1,\dots,n$ that $\varphi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)= =\varphi_i'(x_1,\dots,x_{i-1},\ x_{i+1},\dots,x_n)$ and as a result all the functions $\varphi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)$ of representations (1) of the Boolean function $f(x_1,\dots,x_n)$ appear to be balanced in view of the balancedness of the functions $\varphi_i'(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)$. This, in accordance with the theorem proved above, implies that the function $f(x_1,\dots,x_n)$ itself corresponds to SAC.

Starting from the theorem given above, the problem of SAC-function formation is transformed into the problem of finding a system with n balanced Boolean functions of n-1 variables:

$$\sum_{(x_2,\dots,x_n)\in U_i}\varphi_i(x_2,\dots,x_n)=2^{n-2} \qquad (2)$$

where U_i is the set of 2^{n-1} all the possible values of the Boolean variables $x_1,\dots,x_{i-1},x_{i+1},\dots,x_n$. The resulting Boolean function that corresponds to SAC is presented in the form:

$$f(x_1,\dots,x_n)=\bigvee_{i=1,\dots,n} x_i\cdot\varphi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n) \qquad (3)$$

The problem of finding system (2) of the balanced Boolean functions without making use of the truth tables is rather a complicated one taking into account that the Boolean functions compiling them are dependent.

For practical obtaining of the system (2), special cases of balanced Boolean functions should be applied when the fact of balansedness may be proved by analysis of the normal algebraic form of Boolean functions. A number of corresponding methods may be developed for construction of system (2) of balanced Boolean functions and further on of the SAC-function according to an approach utilized for obtaining balanced functions.

For analytical construction of a system of balanced Boolean functions the known concept [5] is applied according to which the Boolean function $\varphi'(x_1,..,x_n)= =\xi(x_1,..,x_k)\oplus\ \oplus\chi(x_{k+1},\dots x_n)$ is balanced if the function $\xi(x_1,\dots,x_k)$ is linear. This concept leads to the other one that the Boolean function $\varphi'(x_1,..,x_n)= \xi(x_1,..,x_k)\oplus\ \oplus\chi(x_{k+1},\dots x_n)$ is balanced if the function $\xi(x_1,..,x_k)$ is balanced on k Boolean variables.

The essence of the method suggested for obtaining system (3) of balanced functions consists in performing the following sequence of actions:

1. The set of variables $\{x_1,\dots x_n\}$ is divided into two subsets $\vartheta=\{x_1,\dots x_k\}$ and $\Omega\{x_{k+1},\dots x_n\}$.

2. A certain relations Δ which assigns each element of the set Ω to one of the elements of the set ϑ : $y = \Delta(z)$, $z\in\Omega$,$y\in\vartheta$ is given arbitrary. Substantially the mentioned relation prescribes the set of element pairs of the Ω and ϑ sets.

3 The number k of Boolean functions $\xi_h(x_1,\dots,x_{h-1},x_{h+1},\dots,x_{t-1},x_{t+},\dots,x_n)$, $1\le h\le k$, $k+1\le t\le n$, is constructed, with $h=\Delta(t)$, in the form:

$$\xi_h(x_1,\dots,x_{h-1},x_{h+1},\dots,x_{t-1},x_{t+},\dots,x_n)= =(\lambda_h(x_1,\dots,x_{h-1},x_{h+1},\dots,x_k)\oplus \oplus\mu_h(x_{k+1},\dots,x_{t-1},x_{t+},\dots,x_n) \qquad (4)$$

functions $\mu_h(x_{k+1},\dots,x_{t-1},x_{t+1},\dots,x_n)$ and $\lambda_h(x_1,\dots,x_{h-1},x_{h+1},\dots,x_k)$ being given arbitrary.

4. The number of n-k of Boolean functions is constructed, with $u=\Delta(q)$, in the form:

$$\xi_q(x_u,x_{k+1},\dots,x_{q-1},x_{q+1},\dots,x_n)= =\delta_q(x_{k+1},\dots,x_{q-1},x_{q+1},\dots,x_n) \qquad (5)$$

where $\delta_q(x_{k+1},\dots,x_{q-1},x_{q+1},\dots,x_n)$ is an arbitrary function determined on the set of variables belonging to the set Ω and independent of x_q.

5. The SAC Boolean function to be found is obtained by combination through OR of all the conjunctions of variables x_i and the partial balanced Boolean functions ξ_i, i = 1,...,n: obtained earlier:

$$\xi_h f(x_1,\dots,x_n)=\bigvee_{h=1,\dots,k} x_h\cdot\xi_h(x_1,\dots,x_{h-1},x_{h+1},\dots,x_{t-1},x_{t+1},\dots,x_n)\bigvee_{q=k+1,\dots,n}\dots$$

$$\dots\bigvee x_q\cdot\xi_q(x_u,x_{u+1},\dots,x_{q-1},x_{q+},\dots,x_n) \qquad (6)$$

Let us show that the Boolean function formed in the manner described above corresponds to SAC, i.e. all the functions $\varphi_i'(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)$, $i = 1,\dots,n$ in the expansion of the Boolean function $f(x_1,\dots,x_n)$ in accordance with formula (1) are balanced. The resulting SAC-function $f(x_1,\dots,x_n)$ by virtue of condition (6) comprises necessary, as an added, the product of the variable $x_j(j =1,\dots,k)$ by the function $\mu_j(x_{k+1},\dots,x_{t-1},x_{t+1},\dots,x_n)$ as well as it comprises the $x_j\cdot x_t$ product determined by equation (5), where $j=\Delta(t)$. Except the mentioned product the terms comprising the product of the variable x_j by the variables belonging to the set Ω are absent in the normal algebraic form of the function $f(x_1,\dots,x_n)$. Reasoning from this, the functions $\varphi_j(x_1,\dots,x_{j-1},x_{j+1},\dots,x_n)$, $j=1,\dots,k$ may always be represented as the sum:

$\varphi_j(x_1,\dots,x_{j-1},x_{j+1},\dots,x_n)=$

$$=\phi_j(x_1,\dots,x_{j-1},x_{j+1},\dots,x_n)\oplus$$
$$\oplus\mu_j(x_{k+1},\dots,x_{t-1},x_{t+1},\dots,x_n)\oplus x_j \quad (7).$$

Because the functions $\phi_j(x_1,\dots,x_{j-1},x_{j+1},\dots,x_k)$ and $\mu_j(x_{k+1},\dots,x_{t-1},x_{t+1},\dots,x_n)$ do not depend on the linear function x_j, the function $\varphi_j'(x_1,\dots,x_{j-1},x_{j+1},\dots,x_n)$ is balanced. In the similar way, in view of the fact that the resulting SAC Boolean function $f(x_1,\dots,x_n)$ contains necessarily the product $x_u\cdot x_q$, with $u=\Delta(q)$, and does not contain the terms in which the mentioned product would be contained (this is the consequence of the fact that the product $x_u\cdot\mu_j(x_{k+1},\dots,x_{q-1},z_{t+1},\dots,x_n)$ comprised in the function $f(x_1,\dots,x_n)$ does not contain the variable x_q, the functions $\varphi_q(x_1,\dots,x_{q-1},x_{q+1},\dots,x_n)$, $q=k+1,\dots,n$, may always be represented as the sum modulo 2 of the variable x_u and a function independent of

$$x_u;\varphi_q'(x_1,\dots,x_{q-1},x_{q+1},\dots,x_n)=x_u\oplus$$
$$\oplus\sigma_q\varphi_q(x_1,\dots,x_{u-1},x_u,\dots,x_{q-1},x_{q+1},\dots,x_n) \quad (8).$$

With respect to the above indication of balancedness, the Boolean function $\varphi_q(x_1,\dots,x_{q-1},x_{q+1},\dots,x_n)$ is balanced. Thus, it is proved that all the functions φ_i, $i=1,\dots,n$, of the function representation $f(x_1,\dots,x_n)$ in the form (3) are balanced, and according to the basis theorem proved, it means the correspondence of the Boolean function $f(x_1,\dots,x_n)$, being the result of the suggested method, to SAC.

The method suggested provides obtaining the SAC-function possessing the maximal degree of the terms equal to max(n-k,k). As it follows from the analysis of the maximal value of the term degree in expressions (4) and (5) for the partial generating functions ξ_i, $i=1,\dots,n$, the maximal degree of the term in the subfunction $\lambda_h(x_1,\dots,x_{h-1},x_{h+1},\dots,x_k)$ is equal to k-1 (for this purpose the shown function must comprise the term $x_1\cdot x_2\cdot\dots\cdot x_{h-1}\cdot x_{h+1}\cdot,\dots,\cdot x_k$), and the maximal degree of the subfunction $\mu_h(x_{k+1},\dots,x_{t-1},x_{t+1},\dots,x_n)$ is equal to n-k-1 basing on the similar reasonings. Accordingly the maximal degree of all the functions ξ_h, $h=1\dots k$, equals max(k-l, n-k-1). The maximal degree of the term for the functions ξ_q, $q=1\dots n-k$, defined by expression (5) is determined by the maximal degree of of the subfunction $\delta_q(x_{k+1},\dots,x_{q-1},x_{q+1},\dots,x_n)$, which is equal to n-k-1. For this purpose, the term which is the product of all the variables $x_{k+1},\dots,x_n$, except x_q, must be comprised in the normal algebraic sum of the mentioned function. Since each of the function ξ_i, $i=1\dots n$, enters into the resulting SAC-function $f(x_1,\dots,x_n)$, according to (6), as the product by the corresponding variable x_i, the degree of its maximal term makes max(n-k-1,k-1)+1=max(n-k,k). Evaluation of non-linearity of the SAC-function being generated by the suggested manner may be carried out in the following way. According to the theorem on the lower bound of Boolean function non-linearity, proved in the study [2], if Boolean function may be represented in the form $f(x_1,\dots,x_n)=\mu(x_1,\dots,x_n)\oplus\tau(x_1,\dots,x_n)$, where $\tau(x_1,\dots,x_n)$ is the term of degree s which is not comprised in the Boolean function $\mu(x_1,\dots,x_n)$, then non-linearity of the Boolean function $N(f)\geq 2^{n-s}$. Consequently, the lower bound of non-linearity of the arisen function makes 2^{n-2} in the case of the SAC-function being generated in the correspondence with (6) because any of the terms $x_p\cdot x_S, 0<p\leq k$, $k+1\leq s\leq n$, where $x_p=\Delta(x_s)$, of the degree s=2, being presented in the normal algebraic form of the function $f(x_1,x_2,\dots,x_n)$ enters into none of other terms of the latter, that has been demonstrated above at proving the correspondence of the function $f(x_1,x_2,\dots,x_n)$ to SAC.

The suggested method for SAC-function generation may be illustrated by the following example of obtaining the SAC-function of 6 variables. Let k=3 and the sets ϑ and Ω be given respectively in the form $\vartheta=\{x_1,x_2,x_3\}$ and $\Omega=\{x_4,x_5,x_6\}$. Let us set the following relation Δ which performs the one-to-one mapping of the set Ω elements into the set ϑ: $x_1=\Delta(x_4)$, the $x_2=\Delta(x_5)$, $x_3=\Delta(x_6)$. Partial generating functions in correspondence with (4) and (5) may be formed in the following manner:

$$\xi_1=x_2\cdot x_3\oplus x_5\cdot x_6$$
$$\xi_2=x_1\oplus x_3\oplus x_4\cdot x_6$$
$$\xi_3=x_1\oplus x_5$$
$$\xi_4=x_1\oplus x_5\cdot x_6$$
$$\xi_5=x_2\oplus x_4$$
$$\xi_6=x_3\oplus x_5$$

The resulting SAC-function $f(x_1,\dots,x_6)$ is formed by combining through OR the conjunctions of the partial generating functions presented above onto the corresponding variables, according to (6), in the form:

$$f(x_1,\dots,x_6)=x_1\cdot x_2\oplus x_2\cdot x_3\oplus x_1\cdot x_3\oplus x_1\cdot x_4\oplus x_2\cdot x_5\oplus x_3\cdot x_5\oplus$$
$$\oplus x_4\cdot x_5\oplus x_3\cdot x_6\oplus x_5\cdot x_6\oplus x_1\cdot x_2\cdot x_3\oplus x_1\cdot x_5\cdot x_6\oplus x_2\cdot x_4\cdot x_6$$

The maximal degree of the terms comprised in the normal algebraic form of the generated function $f(x_1,\dots,x_6)$ is equal to max(n-k,k)=3, while non-linearity $N(f(x_1,\dots,x_6))=24$ and exceeds the lower non-linearity bound determined above, which is equal to $2^{n-2}=16$. It should be pointed out that non-

linearity of the SAC-function obtained is close to the maximal possible non-linearity for 6 variables equal to 28. The proved theorem and evidence of Boolean function balancedness may be used as the basis for high order SAC-function generation method development.

Below the method of m-order SAC-function generation is presented.

The essence of the method consists in performance of the following actions:

1.The set of variables $\{x_1,...,x_n\}$ is being deviated into two subsets: $\vartheta=\{x_1,...,x_k\}$ and $\Omega=\{x_{k+1},...,x_n\}$, the number of variables in each subset must be not less than m+1.

2.The set Θ of pairs of variable which do not belong simultaneously to the set ϑ and Ω $:<x_r,x_q>\in\Theta$, $r\in\{1,...,k\}$, $q\in\{k+1,...,n\}$ is being determined, in so doing the number of pairs into which every variable $x_1,...,x_n$ enters must be not less than m+1 and the pairs themselves which are comprised in the set Θ , must not recur.

3.SAC-function of the m-th order is being formed in the following manner:

$$f(x_1,...,x_n)=$$
$$=\sum_{\forall<x_d,x_p>\subset\Theta} x_d\cdot x_p\oplus\xi(x_1,..,x_k)\oplus\mu(x_{k+1},...,x_n) \qquad (9)$$

where $\xi(x_1,...,x_k)$ and $\mu(x_{k+1},...,x_n)$ are arbitrary Boolean functions determined on the variable sets ϑ and Ω respectively.

Let us show now that the Boolean function $f(x_1,...,x_n)$ formed in such a manner corresponds to SAC of the m-th order. To meet this demand [1], a Boolean function must possess the maximum of conditional entropy, i.e. it must correspond to SAC of the zero order, correspond to SAC of the (m-1) order and any function $\phi_{k1,...,km}(x_1,...,x_{n-m})$ determined on n-m variables, into which the initial function $f(x_1,...,x_n)$ is transformed by setting m variables $x_{k1},x_{k2},...,x_{km},k_1,\ k_2,...,k_m\in\{1,...\ ,n\}$ to zero or one, must also correspond SAC. Let us show first that the initial function $f(x_1,...,x_n)$ corresponds to this criterion. For this purpose let us present $f(x_1,...,x_n)$ in the form (1) and consider the function $\varphi_i(x_1,...,x_{i-1},x_{i+1},...,x_k)$, $i=1,...,k$, which with account for (9) has the form:

$$\varphi_i(x_1,...,x_{i-1},x_{i+1},...,x_n)=$$
$$=\sum_{\forall X_t:<X_i,X_t>\subset\Theta} x_t\oplus\zeta_i(x_1,...,x_{i-1},x_{i+1},...,x_k),\ i=1,,k, \qquad (10)$$

In this case every x_t included in expression (10) does not belong to the set ϑ, i.e. $x_t\in\Omega$ on the strength of the fact that the set Θ comprises only pairs of variables belonging to the different subsets. Thus the function $\varphi_i(x_1,...,x_{i-1},x_{i+1},...,x_n)$ represents the sum modulo 2 of a linear function (that is the sum of variables x_t) determined on the variables of the set Ω and of the function $\zeta_i(x_1,...,x_{i-1},x_{i+1},...,x_k)$, determined on the variables belonging to the set ϑ, i.e. it is balanced. In the similar way the balancedness property of the functions $\varphi_i(x_1,...,x_{i-1},x_{i+1},...,x_n)$ for $i=k+1,...,n$ may be proved. Thus, it is proved that the $f(x_1,...,x_n)$ corresponds to SAC.

Let us show now that any function $\phi_{k1,...ku}(x_1,...,x_{n-u})$ $u\le m$, determined on n-u variables, into which the initial function $f(x_1,...,x_n)$ transforms through setting u variables $x_{k1},x_{k2},...,x_{ku},k_1,k_2,...,k_u\in\{1,...,\ n\}$ to zero or one possesses such a property. To do this, let us represent the function ϕ in the form (3), supposing that s of the excluded variables belong to the set ϑ and v of those belong to the set Ω, $0\le s\le k$, $0\le v\le n-k$, $s+v=u$. The set of the excluded variables is denoted through

$$\Xi:x_{k1}\cup x_{k2}\cup...\cup x_{ku}=\Xi:\phi_{k1,...ku}(x_h,h=1,...,x_h\notin\Xi)=x_j\cdot\varphi_{j,kl,...,ku}(x_h,h=1,...,j-1,j+1,...,n,\ \ ,x_h\notin\Xi)\ \oplus\ \psi_{j,kl,...,ku}(x_h,h=1,...,j-1,\ j+1,...,n,\ x_h\notin\Xi)$$

Let us consider the function $\varphi_{j,kl,...,ku}(x_h,h=1,...,j-1,j+1,...,n,x_h\notin\Xi)$ for $j\le k$, which may be represented in the following form:

$$\varphi_{j,kl,...,ku}(x_h,h=1,...,j-1,j+1,...,n,x_h\notin\Xi)=$$
$$=\sum_{x_e\in W_j} x_e\oplus\rho_{j,kl,...,ku}(x_h,h=1,j-1,j+1,.,k,x_h\notin\Xi)$$

The set W_j comprises the variables satisfying the following conditions:

$x_e\in W_j:<x_e,x_j>\in\Theta$, $x_e\notin\Xi$. It is also apparent that $x_e\in\Omega$. Let us show that the set $W_j\neq\emptyset$. Consider the variable $x':x'\in\Omega$, $<x',x_j>\in\Theta$: if $x'\in\Xi$, two versions are possible: $x'=0$ and $x'=1$, by the first version the term $x'\cdot x_j$ is excluded out of ANF of the ϕ function of being investigated, and by the second version this term is transformed into the term x_j which being a linear one, according to the Corollary proved at the beginning of the article, does not effect the SAC-property of the initial function and may be also omitted. Thus, if $x'\in\Xi$, it need not be taken into account at analysis of the function φ balancedness. Since, according to the condition of subsets ϑ and Ω formation, the number of variables in each of them is more than m, i.e. more than the maximal amount of the variables being excluded, while v is less than the number of variables entering into the set of Ω, so a variable x_e such that $x_e\notin\Xi$ may be always found. Further on, the method under consieration supposes

the existence of the number of the pairs $<x_g,x_j>\in\Theta$, which also more than m+1, so a variable x_e such that $<x_e,x_j>\in\Theta$ and at the same time $x_e\in\Omega$ will be always found. Consequently, the set $W_j \neq\emptyset$. Then the function $\varphi_{j,kl,\ldots,ku}(x_h,\ h=1,\ldots,j-1,j+1,\ldots,n,\ x_h\notin\Xi)$ will always have the linear constituent subfunction determined on the variables of the set Ω and the other constituent subfunction that does not depend on the variables of the mentioned set, and it means that the function being investigated is balanced. The balancedness of the subfunctions corresponding to the variables of the set Ω may be proved by quite the similar reasoning. According to the theorem proved at the beginning of the article, the function ϕ under investigation is a SAC-function. Correspondingly, the Boolean function $f(x_1,\ldots,x_n)$ constructed in conformity with this technique is a SAC-function of the m-th order.

Non-linearity of the function generated in such a manner also exceeds 2^{n-2}, and the maximal degree of non-linearity, providing the optimal choice of the subsets ϑ and Ω, makes 2^{n-m-1}.

To illustrate the considered method intended to obtaining SAC-functions of the m-th order, an example of first order SAC-function formation from 6 variables is given below.

According to the procedure described above, each of the subsets ϑ and Ω into which the set of variables is divided must comprise not less than n+1=2 variables. Let $\vartheta=\{x_1,x_2\}$ and $\Omega=\{x_3,\ldots,x_6\}$. Determine the set of pairs of the variables belonging to the indicted subsets in the following way:
$\Theta=\{<x_1,x_3>,<x_1,x_4>,<x_1,x_5>,<x_1,x_6>,<x_2,x_3>,<x_2,x_4>,<x_2,x_5>,<x_2,x_6>\}$. The function $\xi(x_1,x_2)$ may be taken arbitrary or even omitted. The function $\xi(x_3,\ldots,x_6)$ may be arbitrary set equal to $x_3\cdot x_4\cdot x_5\oplus x_3\cdot x_4\cdot x_6\oplus x_3\cdot x_4\cdot x_5\cdot x_6$. Then the sought-for SAC-function of the first order may be represented in the following form:
$f(x_1,\ldots,x_n)=x_1\cdot x_3\oplus x_1\cdot x_4\oplus x_1\cdot x_5\oplus x_1\cdot x_6\oplus x_2\cdot x_3\oplus x_2\cdot x_4\oplus x_2\cdot x_5\oplus x_2\cdot x_6\oplus x_3\cdot x_4\cdot x_5\oplus x_3\cdot x_4\cdot x_6\oplus x_3\cdot x_4\cdot x_5\cdot x_6$.

The function obtained corresponds to SAC of the zero and first orders and has non-linearity equal to 20, the degree of non-linearity is equal to 4.

4 Conclusion

Removal of the important for practical problems processing restriction on construction of SAC-functions from a great number of variables inherent in the existing methods for synthesis of this class of Boolean functions may be attained by working-out of a problem solution strategy which does not demand storage of the truth tables in the explicit or implicit manner in the memory. This method must operate with the normal algebraic forms of Boolean functions that needs memory capacity of several order lower. The properties of the normal algebraic forms of SAC-functions are studied, it is shown that the problem of such function obtaining may be reduced to construction of a system of balanced Boolean functions. A method for solution of the last-mentioned problem has been suggested that made it possible to work out formalized procedures for obtaining the normal algebraic forms of SAC-functions of the zero and higher orders with high non-linearity.

The investigation carried out extends in effect the principles set forth by B.Preneel [5]. Comparing to the basic concept, the method suggested makes it possible to increase the non-linearity characteristics of the SAC-functions being obtained. The experimental study carried out has shown the practical usefulness of the developed methods for obtaining Boolean functions from 100 and a greater number of variables.

References:

[1]R.Forre. The strict avalanche criterion: spectral properties of Boolean functions and an extend definition, CRYPTO'88, Vol.403, 1990, pp 450-468.

[2]K.Kurosawa and T.Saton. Design of SAC/PC(l) of Order *k* Boolean Functions and Three Other Cryptographic Criteria, EUROCRYPT'97, Vol.1233, Springer 1997, pp.434-449.

[3]M.Matsui. Linear cryptanalysis method for DES-chiper, EUROCRYPT'93, Vol.765, Springer-Verlag.1994, pp. 386-397.

[4]B.Preneel, W.Van Leekwijck, L.Van Linden, R.Govaerts and J.Vandewalle. Propagation characteristics of Boolean functions, EUROCRYPT'90, Vol.473, Springer-Verlag.1991, pp.161-173.

[5]B.Preneel, R.Govaerts and J.Vandewalle. Boolean functions satisfying higher order propagation criteria, EUROCRYPT'91, Vol.547, Springer-Verlag 1991, pp.141-152.

[6]A.F.Webster and S.E.Tavares. On the design of S-boxes. CRYPTO'85, Vol..218, Springer - Verlag.1986,pp.523-534.