On construction of correlation-immune functions via minimal functions

Karelina E. K. — Infotecs JSC
Alekseev E. K. — CryptoPro LLC
Logachev O. A. — ISI MSU

2017

Examples of the cryptographic properties of Boolean functions

1. Nonlinearity
2. Algebraic immunity
3. Nondegeneracy
4. Correlation immunity

Existing construction methods

1. Brute-force search method

   The search of functions in the sets which a priori possess a positive set of properties:

   - Maiorana-McFarland class
   - $\mathcal{PS}$ class

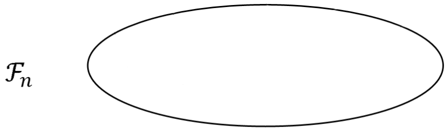
2. Recursive method

$$f_1 \in \mathcal{F}_n \longrightarrow f_2 \in \mathcal{F}_{n+1} \longrightarrow f_3 \in \mathcal{F}_{n+2} \longrightarrow \dots \longrightarrow f_{m-1} \in \mathcal{F}_{n+m-1} \longrightarrow f_m \in \mathcal{F}_{n+m}$$

The main aspects of this paper

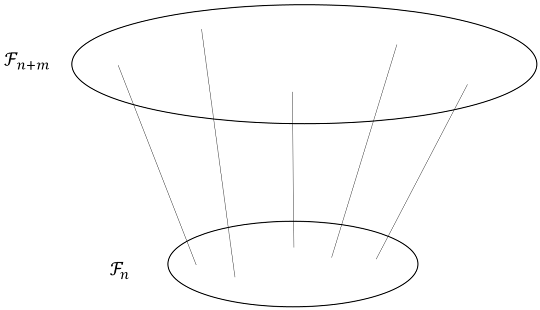
1. Method of functions construction with the specified order of correlation immunity based on a combination of the above-stated approaches

   - The construction of a base set



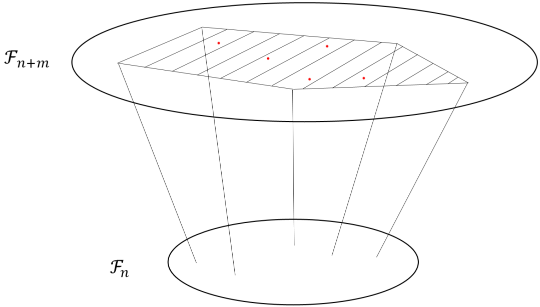
The main aspects of this paper

1. Method of functions construction with the specified order of correlation immunity based on a combination of the above-stated approaches
   - The construction of a base set
   - Recursive method



The main aspects of this paper

1. Method of functions construction with the specified order of correlation immunity based on a combination of the above-stated approaches
   - The construction of a base set
   - Recursive method
   - Search for functions in the target dimension



The main aspects of this paper

2. Study of «neighbourhoods» of the already known functions

Basic concepts and notations

- Let $\mathbb{F}_2$ be the finite field of 2 elements.
  $\forall n \in \mathbb{N}\ V_n = (\mathbb{F}_2 \times \ldots \times \mathbb{F}_2) = \mathbb{F}_2^n$.
  $V_n^* = V_n \setminus \{0^n\}$, where $0^n = (0, \ldots, 0) \in V_n$.

- Boolean function of n variables is the correspondence from V_n into $\mathbb{F}_2$. Constant Boolean functions are denoted as $\mathbf{1}$ and $\mathbf{0}$. The set of all Boolean functions is denoted as $\mathcal{F}_n$.

- The support 1_f of a Boolean function $f \in \mathcal{F}_n$ is a set $1_f = \{\, x \in V_n \mid f(x) = 1 \,\}$. The weight $\mathrm{wt}\,(f)$ of a Boolean function $f \in \mathcal{F}_n$ is a cardinality of the support. The distance $\mathrm{dist}\,(f, g)$ between $f \in \mathcal{F}_n$ and $g \in \mathcal{F}_n$ is value of $\mathrm{wt}\,(f \oplus g)$.

Basic concepts and notations

- The algebraic degree $\deg(f)$ of a Boolean function $f \in \mathcal{F}_n$ of n variables is the number of variables in the longest term ANF (Zhegalkin polynomial).
- For $u \in V_n$ a Boolean function l_u denotes a linear Boolean function $l_u(x) = \langle u, x \rangle$, where $\langle u, x \rangle = \bigoplus_{i=1}^{n} u_i \cdot x_i$ is a scalar product of vectors u and x. The set $\{l_u(x) \oplus b | u \in V_n, b \in \mathbb{F}_2\}$ of affine Boolean functions of n variables is denoted as $\mathcal{A}_n$.
- Nonlinearity $\mathrm{nl}(f)$ of a Boolean function $f \in \mathcal{F}_n$ is the Hamming distance to the set of all affine functions $\mathcal{A}_n$:

$$\mathrm{nl}(f) = \mathrm{dist}(f, \mathcal{A}_n) = \min_{l \in \mathcal{A}_n} \mathrm{dist}(f, l).$$

Basic concepts and notations

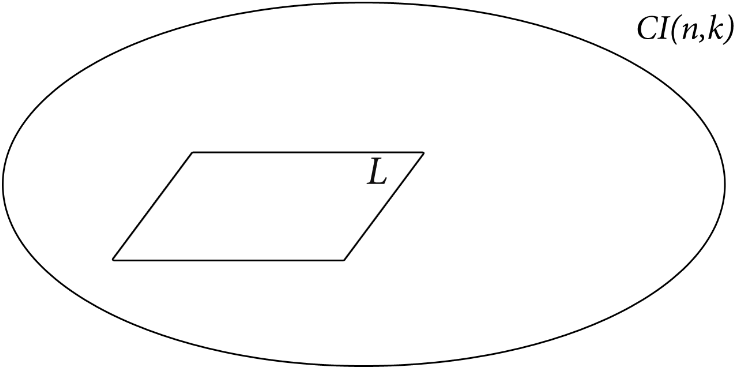
- $f \in \mathcal{F}_n$ is correlation-immune of order m, $1 \leqslant m \leqslant n$ (further CI-function), if the following equality $\mathrm{wt}\left(f^{'}\right) = \frac{\mathrm{wt}(f)}{2^m}$ holds for any subfunction $f^{'}$ of $n-m$ variables.
- $cor(f) = \max\{m \in \mathbb{N} \mid f$ — correlation immune of order $m\}$.
- $\mathrm{CI}(n,k) = \{f \in \mathcal{F}_n | \mathrm{cor}\,(f) \geqslant k\}$
  $\mathrm{CI}(n) = \mathrm{CI}(n,1)$

Basic concepts and notations

- The balanced function $f \in \mathcal{F}_n$ is k-resilient, if $\mathrm{cor}\,(f) \geqslant k$.
- The Walsh-Hadamard transform of a Boolean function $f \in \mathcal{F}_n$ is an integral function $W_f : V_n \to \mathbb{Z}$, $W_f(u) = \sum_{x \in V_n} (-1)^{f(x) \oplus \langle u, x \rangle}$.
- $f \in \mathcal{F}_n$ is correlation-immune function of m order, $0 < m \leqslant n$, $\Leftrightarrow \forall u \in V_n : 1 \leqslant \mathrm{wt}\,(u) \leqslant m$, the equality $W_f(u) = 0$ performs.
- Functions $f, g \in \mathrm{CI}(n, k)$ such that $f \cdot g = \mathbf{0}$ are called *orthogonal*. Let $f, g \in \mathrm{CI}(n, k)$ be *orthogonal* then $f \oplus g \in \mathrm{CI}(n, k)$.

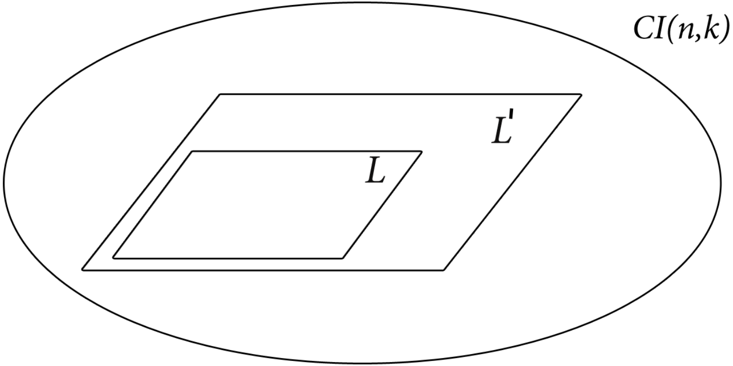
Minimal correlation-immune functions

L — linear space, $L \subset \mathrm{CI}(n, k)$.
$f_1, \ldots, f_r \in \mathrm{CI}(n, k)$ — basis L, which consist of mutually orthogonal functions.



Minimal correlation-immune functions

Could we expand L to L' such that $L \subset L' \subset CI(n, k)$?



Minimal correlation-immune functions

Expand basis:

1. $g = \mathbf{1} \oplus f_1 \oplus \ldots \oplus f_r \in CI(n,k)$, $\forall i \in [1,r]$ $g \cdot f_i = \mathbf{0}$.
2. Let's decompose existing functions f_i into a sum of mutually orthogonal functions $f_i', f_i'' \in \mathrm{CI}(n,k)$ for all $i \in [1,r]$.

Minimal correlation-immune functions

Functions $f \in \mathrm{CI}(n, k)$, which can't be represented as a sum of orthogonal functions $f', f'' \in \mathrm{CI}(n, k)$, will be called *k-minimal correlation-immune functions* (k-minimal for short).

$\mathrm{MCI}(n, k)$ — a set of k-minimal functions of n variables.

The construction of 1-minimal functions with a given number of variables

The truth table of function $f \in \mathcal{F}_n$ is called the matrix T_f of order $\mathrm{wt}\,(f) \times n$, the rows of this matrix are vectors from 1_f lexicographically-ordered.
For example, for function $f(x_1, x_2, x_3) = x_1x_2 \oplus x_3 \in \mathcal{F}_3$

$$T_f = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

The construction of 1-minimal functions with a given number of variables

Let be $\mathcal{F}_n^w = \{f \in \mathcal{F}_n | \mathrm{wt}\,(f) = w\}$. For any $w \in \{1, \ldots, 2^n\}$ define the map $\mathrm{AC}^{(w)}$:

$$\mathrm{AC}^{(w)} : \mathcal{F}_n^w \times V_w \times \{1, \ldots, n+1\} \mapsto \mathcal{F}_{n+1}^w.$$

The function $g = \mathrm{AC}_{v,i}^{(w)}(f) = \mathrm{AC}^{(w)}(f, v, i)$ is defined as follows. The matrix $\mathrm{wt}\,(f) \times (n+1)$ is formed by adding vector v of dimension w in the truth table T_f as i-th column. Whereas i-th and the following columns T_f are shifted to the right. The rows of formed matrix is support of function g. If $i = n+1$, then the column is added to the end of the table.

The construction of 1-minimal functions with a given number of variables

f(x1, x2, x3, x4)		g(x1,x2,x3, x4, x5)
0 0 0 0	→	0 0 0 0 1
0 1 1 1		0 1 1 1 0
1 0 1 0		1 0 1 0 1
1 1 0 1		1 1 0 1 0

$$g(x_1, x_2, x_3, x_4, x_5) = \mathrm{AC}^{(4)}_{v,5}(f), \text{ where } v = (1010)$$

The construction of 1-minimal functions with a given number of variables

Theorem

Let $f \in \mathrm{CI}(n)$ and $w = \mathrm{wt}\,(f)$. Then for any $v \in V_w$, such that $\mathrm{wt}\,(v) = w/2$, and for any $i \in \{1, \ldots, n+1\}$ the following is true $g = \mathrm{AC}^{(w)}_{v,i}(f) \in \mathrm{CI}(n+1)$.

Theorem

Let $f \in \mathrm{MCI}(n, 1)$ and $w = \mathrm{wt}\,(f)$. Then for any $v \in V_w$, such that $\mathrm{wt}\,(v) = w/2$, and for any $i \in \{1, \ldots, n+1\}$ the following is true $g = \mathrm{AC}^{(w)}_{v,i}(f) \in \mathrm{MCI}(n+1, 1)$.

The search of function with a given order of correlation immunity

- $L \subset \mathrm{CI}(n, k)$ — a linear space
- $f_1, \ldots, f_r \in \mathrm{CI}(n, k)$ — basis of mutually orthogonal functions

The search of a $(k + m)$-*resilient function* $g \in L$:

- $\mathrm{cor}\,(g) \geqslant k + m,\ m \geqslant 1$
- $\mathrm{wt}\,(g) = 2^{n-1}$

The search of function with a given order of correlation immunity

For any u, $\mathrm{wt}\,(u) > 0$, and for any

$$g = b_1 \cdot f_1 \oplus \ldots \oplus b_r \cdot f_r, b_1, \ldots, b_r \in \mathbb{F}_2,$$

the following equality holds:

$$W_g(u) = b_1 \cdot W_{f_1}(u) + \ldots + b_r \cdot W_{f_r}(u).$$

The search of function with a given order of correlation immunity

- g — CI-function of $(k+m)$-th order $\Leftrightarrow$ for any u, $1 \leqslant \mathrm{wt}\,(u) \leqslant k+m$, the equality $W_g(u) = 0$ is true.
- $f_i \in \mathrm{CI}(n,k) \Rightarrow W_{f_i}(u) = 0$ for any $u : 1 \leqslant \mathrm{wt}\,(u) \leqslant k$.

So the function g — CI-function with $\mathrm{cor}\,(g) \geqslant k+m \Leftrightarrow$ for all u, $k+1 \leqslant \mathrm{wt}\,(u) \leqslant k+m$, $\binom{n}{k+1} + \ldots + \binom{n}{k+m}$ equations are true:

$$b_1 \cdot W_{f_1}(u) + \ldots + b_r \cdot W_{f_r}(u) = 0$$

The search of function with a given order of correlation immunity

The condition $\mathrm{wt}\,(g) = 2^{n-1}$ is true if the following equality is true:

$$b_1 \cdot \mathrm{wt}\,(f_1) + \ldots + b_r \cdot \mathrm{wt}\,(f_r) = 2^{n-1}.$$

The search of function with a given order of correlation immunity

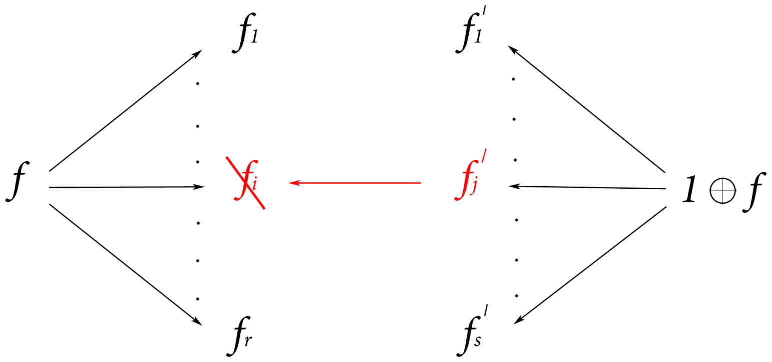
In order to find $(k+m)$-resilient function $g \in L$ it is sufficient to find $(0,1)$-solutions $(b_1, \ldots, b_r)$ of the system of $\binom{n}{k+1} + \ldots + \binom{n}{k+m} + 1$ linear equations

$$\begin{cases} b_1 \cdot W_{f_1}(u) + \ldots + b_r \cdot W_{f_r}(u) = 0, \text{ for } u : k+1 \leqslant \mathrm{wt}\,(u) \leqslant k+m \\ b_1 \cdot \mathrm{wt}\,(f_1) + \ldots + b_r \cdot \mathrm{wt}\,(f_r) = 2^{n-1} \end{cases}$$

Analysis of «neighbourhoods» of known functions

- $f \in \mathrm{CI}(n, k)$,
  $f = f_1 \oplus \ldots \oplus f_r$, where f_i mutually orthogonal functions
- $f \oplus \mathbf{1} \in \mathrm{CI}(n, k)$, $f \oplus \mathbf{1} =$
  $f_1^{'} \oplus \ldots \oplus f_s^{'}$, where $f_i^{'}$ mutually orthogonal functions

Analysis of «neighbourhoods» of known functions



The results of applying the proposed methods

Consider the function $f_T \in \mathcal{F}_{10}$:

$\mathrm{wt}(f_T) = 512$	$\mathrm{cor}(f_T) = 6$	$\deg(f_T) = 3$	$\mathrm{nl}(f_T) = 384$	$\mathrm{nd}(f_T) = 0$

Functions f_T and $f_T \oplus \mathbf{1}$ were decomposed on 128 1-minimal functions with the weight 4.
The main disadvantage of f_T: $\mathrm{nd}(f_T) = 0$.

The new function g_T has the following parameters:

$\mathrm{wt}(g_T) = 512$	$\mathrm{cor}(g_T) = 2$	$\deg(g_T) = 7$	$\mathrm{nl}(g_T) = 360$	$\mathrm{nd}(g_T) = 8$

The results of applying the proposed methods

The filter function f_c is used in stream cipher LILI128. This function of 10 variables has the following parameters:

$\mathrm{wt}(f_c) = 512$	$\mathrm{cor}(f_c) = 3$	$\deg(f_c) = 6$	$\mathrm{nl}(f_c) = 480$	$\mathrm{nd}(f_c) = 80$

The new function g_c is constructed with the following parameters:

$\mathrm{wt}(g_c) = 512$	$\mathrm{cor}(g_c) = 3$	$\deg(g_c) = 6$	$\mathrm{nl}(g_c) = 480$	$\mathrm{nd}(g_c) = 112$

The results of applying the proposed methods

The example of function construction without the use of known «good» functions:

- $f = f_1 \oplus \ldots \oplus f_{256} \in CI(10, 7)$
- $f_i \in MCI(10, 1), \mathrm{wt}(f_i) = 2, i \in [1, 256]$

$\mathrm{wt}(f) = 512$	$\mathrm{cor}(f) = 7$	$\deg(f) = 2$	$\mathrm{nl}(f) = 256$	$\mathrm{nd}(f) = 0$

This function is 7-resilient function and it achieves the upper bound for nonlinearity.

Thanks for your attention!

alekseev@cryptopro.ru
Ekaterina.Karelina@infotecs.ru
logol@iisi.msu.ru

Nondegeneracy

- A — $(n \times k)$-matrix over $\mathbb{F}_2$
- $f \in \mathcal{F}_k$
- $f^A \in \mathcal{F}_n$, $f^A(x) = f(xA)$

1. The order of algebraic degeneracy $\mathrm{AD}(f)$ of $f \in \mathcal{F}_n$ is the maximum possible value of $(n - k)$, where the integer k, $0 \leqslant k \leqslant n$ such that a function $g \in \mathcal{F}_k$ and $(n \times k)$-matrix A over $\mathbb{F}_2$ exist, that there is an equality $f = g^A$.
2. Functions with $\mathrm{AD}(f) > 0$ are algebraically degenerate.
3. The set of all degenerate algebraic functions of n variables is denoted as $\mathrm{DG}(n) = \{f \in \mathcal{F}_n \mid \mathrm{AD}(f) > 0\}$.
4. Nondegeneracy of a function $f \in \mathcal{F}_n$ is the following value:

$$\mathrm{nd}(f) = \mathrm{dist}(f, \mathrm{DG}(n)).$$

Maiorana-McFarland class

- $\pi : V_n \to V_n$ — a substitution on the space V_n
- $\psi \in \mathcal{F}_n$ — a Boolean function of n variables

$$M = \{f(x, y) \in \mathcal{F}_{2n} : f(x, y) =< \pi(y), x > \oplus \psi(y), x, y \in V_n\}$$

— Maiorana-McFarland class.

$\mathcal{PS}$ class

$L_1, \ldots, L_{2^{n-1}}$ — subsets of V_{2n}

- $dim L_i = n, i = 1, \ldots, 2^{n-1}$
- $L_i \cap L_j = \mathbf{0}, i \neq j, i, j = 1, \ldots, 2^{n-1}$

$$\mathcal{PS}^- = \{f(x) = I_{L_1} \oplus \ldots \oplus I_{L_{2^{n-1}}}\}$$

$L_1, \ldots, L_{2^{n-1}+1}$ — subsets of V_{2n}

- $dim L_i = n, i = 1, \ldots, 2^{n-1} + 1$
- $L_i \cap L_j = \mathbf{0}, i \neq j, i, j = 1, \ldots, 2^{n-1} + 1$

$$\mathcal{PS}^+ = \{f(x) = I_{L_1} \oplus \ldots \oplus I_{L_{2^{n-1}+1}}\}$$

$$\mathcal{PS} = \mathcal{PS}^- \cup \mathcal{PS}^+$$

Open problems

1. The search of efficient criteria for approving the k-minimality of this function.
2. The development of a method of the increasing number of variables k-minimal functions, $k > 1$, is as effective as for the $k = 1$.
3. The development of efficient searching method of balanced functions with a given values of nonlinearity/nondegeneracy/algebraic immunity in the space generated by k-minimal functions.