Construction of balanced functions with high nonlinearity and other cryptographic properties

Alexandr Shaporenko
shaporenko.alexandr@gmail.com

Novosibirsk State University

CTCrypt 2024

Определения и обозначения

A function $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ is called **Boolean function in n variables**.

A function f is **balanced** if $|\{x \in \mathbb{Z}_2^n : f(x) = 0\}| = |\{x \in \mathbb{Z}_2^n : f(x) = 1\}|$.

Every Boolean function f can be uniquely represented y represented by its **algebraic normal form (ANF)**:

$$f(x_1, \ldots, x_n) = (\bigoplus_{k=1}^{n} \bigoplus_{i_1,\ldots,i_k} a_{i_1,\ldots,i_k} x_{i_1} \cdot \ldots \cdot x_{i_k}) \oplus a_0, \text{ где } a_{i_1}, \ldots, a_{i_k}, a_0 \in \mathbb{Z}_2.$$

A Boolean function h of the form $h(x_1, \ldots, x_n) = h_1(x_1, \ldots, x_{n-1}) \oplus x_n$ is said to **depend linearly on variable** x_n.

The variable x_n is said to be **non-essential variable** of a Boolean function h in n variables if h has the form $h(x_1, \ldots, x_n) = h_1(x_1, \ldots, x_{n-1})$.

Определения и обозначения

A Boolean function f is **affine** if it can be represented as $\ell_{a,b}(x) = \langle a, x\rangle \oplus b$, where $a \in \mathbb{Z}_2^n$ and $b \in \mathbb{Z}_2$.

The Boolean function $D_y f(x) = f(x) \oplus f(x \oplus y)$ is called a **derivative** of a Boolean function f in n variables in **the direction** y

The Walsh-Hadamard transform of a Boolean function f in n variables: $W_f(y) = \sum\limits_{x \in \mathbb{Z}_2^n} (-1)^{f(x) \oplus \langle x, y\rangle}$, for every $y \in \mathbb{Z}_2^n$.

The Hamming distance between two Boolean functions f and g is given by $dist(f, g) = |\{x \in \mathbb{Z}_2^n : f(x) \neq g(x)\}|$.

The nonlinearity N_f of a Boolean function f is the Hamming distance from f to the set of all affine functions, i.e., $N_f = \min\limits_{a \in \mathbb{Z}_2^n, b \in \mathbb{Z}_2} dist(f, \ell_{a,b})$.

A **bent function** is a Boolean function in an even number of variables that has the maximal nonlinearity, i.e., $N_f = 2^{n-1} - 2^{n/2-1}$.

Denote by $\mathcal{F}_n$, $\mathcal{A}_n$ and $\mathcal{B}_n$ the sets of all Boolean function, affine functions and bent functions in n variables, respectively.

Since bent functions are not balanced, a known problem is to find balanced functions with high nonlinearity.

Works [1-6] are devoted to the search and methods for constructing balanced functions with high nonlinearity.

---

[1] Gini A., Meaux P.: Weightwise Perfectly Balanced Functions and Nonlinearity // Codes Cryptol. Inform. Secur. C2SI 2023 — 2023 — P. 338–359.

[2] Carlet C., Djurasevic M., Jakobovic D., Mariot L., Picek S.: Evolving Constructions for Balanced, Highly Nonlinear Boolean Functions — available at https://arxiv.org/abs/2202.08743.

[3] Hu X., Yang B., Huang M.: A construction of highly nonlinear Boolean functions with optimal algebraic immunity and low hardware implementation cost // Discr. Appl. Math. — 2020 — V. 285 — P. 407–422.

[4] Picek S., Sisejkovic D., Jakobovic D.: Immunological algorithms paradigm for construction of boolean functions with good cryptographic properties // Eng. Appl. AI — 2017 — V. 62 — 320–330.

A Boolean function f in n variables has a **linear structure** if there is a nonzero vector $y \in \mathbb{Z}_2^n$ such that $D_y f \equiv const$.

A Boolean function f in n variables satisfies **the Strict Avalanche Criterion (SAC)** if for any $y \in \mathbb{Z}_2^n$ such that $\mathrm{wt}(y) = 1$ the derivative $D_y(f)$ is balanced, where $wt(y) = |\{y_i \in \mathbb{Z}_2 : y_i = 1\}|$.

A Boolean function $f \in \mathcal{F}_n$ is called **correlation immune of order** r, $1 \leqslant r \leqslant n$, if for any subfunction $g = f_{i_1,\ldots,i_r}^{a_1,\ldots,a_r}$, which obtained from f by substituting constants $a_1, \ldots, a_r$ instead of variables $x_{i_1}, \ldots, x_{i_r}$, it holds $wt(g) = \frac{wt(f)}{2^r}$, where $wt(f) = |\{x \in \mathbb{Z}_2^n : f(x) = 1\}|$.

A detailed description of these properties can be found in monograph [1].

[1] Pankratova I. A.: Boolean functions in cryptography // Tomsk: Tomsk State University — 2014.

We propose the following iterative construction of Boolean functions.

Construction 1

Let $g_1, g_2 \in \mathcal{F}_n$, $h \in \mathcal{F}_{n+2}$ and $y \in \mathbb{Z}_2^n$. We will construct the function $f \in \mathcal{F}_{n+2}$ in the following way:

$$f(x, x_{n+1}, x_{n+2}) = \big((D_y g_1(x) \oplus 1)h(x, x_{n+1}, x_{n+2}) \oplus D_y g_2(x)\big)x_{n+1} \oplus g_1(x)h(x, x_{n+1}, x_{n+2}) \oplus g_2(x),$$

where $x \in \mathbb{Z}_2^n, x_{n+1}, x_{n+2} \in \mathbb{Z}_2$.
The functions g_1, g_2, h and vector y are parameters of the construction.

Remark

Let $h \in \mathcal{F}_{n+2}$, $y \in \mathbb{Z}_2^n$, $y_{n+2} \in \mathbb{Z}_2$. If h depends linaerly on variable x_{n+2} and $D_{(y,1,y_{n+2})}h(x, x_{n+1}, x_{n+2}) \equiv 0$, then the construction characterizes all functions $f \in \mathcal{F}_{n+2}$ that have h as their derivative in the direction $(y, 1, y_{n+2})$.

The following statement describes the parameters of Construction 1, under which the functions will have cryptographic properties.

Proposition 1 (Shaporenko A., 2024)

Let g_1, $g_2 \in \mathcal{F}_n$, $y \in \mathbb{Z}_2^n$ and $h(x, x_{n+1}, x_{n+2}) = \langle b, x \rangle \oplus c \oplus x_{n+2}$ for every $x \in \mathbb{Z}_2^n$, where $b \in \mathbb{Z}_2^n$ and $c \in \mathbb{Z}_2$. Then for the function $f \in \mathcal{F}_{n+2}$ from Construction 1 it holds

1. *f have h as its derivative in the direction $(y, 1, \langle b, y \rangle)$;*
2. *f is balanced if and only if g_2 is balanced;*
3. *$N_f = 2^{n+1} - \max\limits_{a \in \mathbb{Z}_2^n, g \in \{g_2, g_1 \oplus g_2\}} |W_g(a)|$;*
4. *if g_2 and $g_1 \oplus g_2$ are correlation immune functions of order r, then f is correlation immune of order r.*
5. *if g_2 and $g_1 \oplus g_2$ are balanced function and bent function, respectively, then f is balanced without linear structures.*

[1] Shaporenko A. S.: Construction of balanced functions with high nonlinearity and other cryptographic properties // Prikl. Diskr. Mat. — 2024 — V. 63 — P.8 – 23.

Corollary 1

Let $n \geqslant 2$ be an even integer, $g_1, g_2 \in \mathcal{F}_n$, $y, b \in \mathbb{Z}_2^n$ and $h(x, x_{n+1}, x_{n+2}) = \langle b, x \rangle \oplus c \oplus x_{n+2}$, where $c \in \mathbb{Z}_2$. If $g_1 \oplus g_2 \in \mathcal{B}_n$ then for $f \in \mathcal{F}_{n+2}$ from Construction 1 it holds $N_f = 2^n + 2 \cdot N_{g_2}$.

Let us present sufficient conditions for the functions constructed by using Construction 1 to satisfy the strict avalanche criterion.

Proposition 2

Let $n \geqslant 2$ be an even integer, $g_1, g_2 \in \mathcal{F}_n$, $(0, \ldots, 0) = y \in \mathbb{Z}_2^n$ and $h(x, x_{n+1}, x_{n+2}) = \ell_1(x) \oplus x_{n+2}$ for any $x \in \mathbb{Z}_2^n$, where $\ell_1(x) = x_1 \oplus \ldots \oplus x_n$. Then $f \in \mathcal{F}_{n+2}$ constructed by using Construction 1 satisfies the strict avalanche criterion.

It was shown in [1] that there are balanced functions in 16 variables with a nonlinearity of 32 638.

Next, we present an iterative method for constructing balanced functions in an even $n \geqslant 18$ number of variables.

Method 1

We will use Construction 1 with the following parameters:

- *if* $n = 18$ *then* g_2 *is balanced function in* 16 *variables from [1];*
- *if* $n \geqslant 20$ *then* g_2 *is the function* f *from Construction 1, obtained using Method 1 in the previous step;*
- $h(x, x_{n+1}, x_{n+2}) = \ell_1(x) \oplus x_{n+2}$ *for all* $x \in \mathbb{Z}_2^n$, *where* $\ell_1(x) = x_1 \oplus \ldots \oplus x_n$;
- g_1 *such that* $g_1 \oplus g_2$ *is bent;*
- $(0, \ldots, 0) = y \in \mathbb{Z}_2^n$.

[1] Picek S., Sisejkovic D., Jakobovic D.: Immunological algorithms paradigm for construction of boolean functions with good cryptographic properties // Eng. Appl. AI — 2017 — V. 62 — 320–330.

Note that for a bent function $g_1 \oplus g_2$ of n variables we can take, for example, a bent function from the Maiorana-McFarland class: $\langle \pi(x), y \rangle \oplus g(x)$, where π is a one-to-one mapping on $\mathbb{Z}_2^{n/2}$, function g is an arbitrary Boolean function of $n/2$ variables and $x, y \in \mathbb{Z}_2^{n/2}$. Then $g_1 = g_1 \oplus g_2 \oplus g_2$.

Theorem 1

The functions in $n \geqslant 18$ variables obtained by Method 1 are balanced functions whithout linear structures with nonlinearity $2^{n-1} - (2^{\frac{n}{2}-1} + 2^{\frac{n}{2}-7})$ that satisfy the strict avalanch criterion.

In [1] for nonlinearity of balanced functions in an even number n of variables the following upper bound holds $2^{n-1} - 2^{\frac{n}{2}-1} - 2$.

n	$2^{n-1} - 2^{\frac{n}{2}-1} - 2$ [1]	N_f из Theorem 1	N_f [2]	N_f [3]
18	130 814	130 812	130 504	130 688
20	523 774	523 768	523 154	not presented
22	2 096 126	2 096 112	2 094 980	not presented
24	8 386 558	8 386 528	8 384 490	not presented
26	33 550 334	33 550 272	33 545992	not presented
28	134 209 534	134 209 408	134 201 460	not presented

[1] Seberry J., Zhang X-M., Zheng Y.: Nonlinearly Balanced Boolean Functions and Their Propagation Characteristics // Advances in Cryptology - CRYPTO'93 — P. 49–60.

[2] Hu X., Yang B., Huang M.: A construction of highly nonlinear Boolean functions with optimal algebraic immunity and low hardware implementation cost // Discr. Appl. Math. — 2020 — V. 285 — P. 407–422.

[3] Carlet C., Djurasevic M., Jakobovic D., Mariot L., Picek S.: Evolving Constructions for Balanced, Highly Nonlinear Boolean Functions — available at https://arxiv.org/abs/2202.08743.

Thanks for attention!