ON THE NUMBER OF SUBFUNCTIONS OF RANDOM BOOLEAN FUNCTION WHICH ARE CLOSE TO THE AFFINE FUNCTIONS SET

A.A. Serov
Steklov Mathematical Institute of RAS
Moscow
e-mail: serov@mi.ras.ru

Abstract

We present formulas and inequalities for the mean and the variance of the number of subfunctions of a random Boolean function such that their Hamming distances to the set of affine Boolean functions do not exceed a given value. The critical values of the number of variables of subfunctions having good affine approximations are obtained.

This work was supported by RFBR, grant 11-01-00139.

Let $V_n = (\mathrm{GF}(2))^n$. Denote by $\mathbb{F}_2^{V_n}$ the set of all Boolean functions and by $\mathbb{A}_n$ the set of all affine functions of n Boolean variables. Let $\mathbb{F}_2^{V_n}(r) \subseteq \mathbb{F}_2^{V_n}$ be the set of all Boolean functions with Hamming distance to $\mathbb{A}_n$ not exceeding r.

Let $f = f(x_1, \ldots, x_n) \in \mathbb{F}_2^{V_n}$. For the partition $\{1, \ldots, n\} = I_n^s \cup J_n^{n-s}$, $I_n^s = \{i_1, \ldots, i_s\}$, $J_n^{n-s} = \{j_1, \ldots, j_{n-s}\}$ and for the collection of constants $C_{n-s} = \{c_1, \ldots, c_{n-s} \in \mathbb{F}_2\}$ we define the *subfunction* $g(J_n^{n-s}, C_{n-s}; y_1, \ldots, y_s) \in \mathbb{F}_2^{V_s}$ of f as the function obtained from $f(x_1, \ldots, x_n)$ by the following change of variables: $x_{j_k} = c_k\,(k = 1, \ldots, n-s)$, and $x_{i_m} = y_m\,(m = 1, \ldots, s)$.

Let $\nu(f, s, r)$ be the number of subfunctions $g(J_n^{n-s}, C_{n-s}; y_1, \ldots, y_s) \in \mathbb{F}_2^{V_s}$ with distance from $\mathbb{A}_s$ not exceeding r, i. e. the number of pairs $(J_n^{n-s},\, C_{n-s})$ such that $g(J_n^{n-s},\, C_{n-s}) \in \mathbb{F}_2^{V_s}(r)$.

Theorem 1. *If $\varphi(x_1, \ldots, x_n)$ is a random Boolean function with the uniform distribution on $\mathbb{F}_2^{V_n}$ then for $r < 2^{s-2}$*

$$\mathbf{E}\,\nu(\varphi, s, r) = C_n^s\, 2^{n-2^s+1} \sum_{j=0}^{r} C_{2^s}^j,$$

$$2^{n-2^s+1} C_n^s C_{2^s}^r \leqslant \mathbf{E}\,\nu(\varphi, s, r) \leqslant 2^{n+1}\, C_n^s \left(\tfrac{2}{3}\right)^{2^{s-2}}.$$

Remark 1. *Note that the left hand side of the inequality tends to infinity if $s \leqslant \log_2 n$, $n \to \infty$, and the right hand side tends to 0 if $s \geqslant \log_2 n + 3$, $n \to \infty$; thus a "threshold" value of s belongs to the range from $\log_2 n$ to $\log_2 n + 3$ when $n \to \infty$.*

State 1. a) *The values of s maximizing $\mathbf{E}\,\nu(f, s, 2^{s-2} - 1)$ for the sufficiently large n belongs to the interval*

$$[\log_2 \ln n + 2,5,\ \log_2 \ln n + 3, 5].$$

b) *The inequality* $\mathbf{E}\,\nu(f,s,2^{s-2}-1)<1$ *holds for* $s\geqslant\log_2 n+4$ $(n\geqslant 1)$ *and for* $s\geqslant\log_2 n+3$ $(n\geqslant 70)$.

State 2. *If* f *be a random Boolean function with the uniform distribution on* $\mathbb{F}_2^{V_n}$ *then for* $r<2^{s-3}$

$$\mathbf{D}\nu(f,s,r)<\left(1-2^{s+1-2^s}\sum_{j=0}^{r}C_{2^s}^{j}\right)\mathbf{E}\nu(f,s,r)+$$
$$+\frac{3}{4}\frac{s\cdot C_n^s\cdot 2^{2n+5s-5}}{\pi\sqrt{\pi 2^{s-2}}2^{s-2}}\exp\left\{\frac{-3\cdot 2^{2s-3}-4r(r+1)+2^{s-1}(4r+3)}{2^s+2}\right\}.$$

Theorem 2. *If* f *is a random Boolean function with the uniform distribution on* $\mathbb{F}_2^{V_n}$ *then for* $r<2^{s-2}$

$$\mathbf{E}\nu(f,s,r)^{[2]}=$$
$$=\left(2^{n-s}C_n^{n-s}(2^{n-s}C_n^{n-s}-1)-C_n^{n-s}\sum_{d=\max\{0,2s-n\}}^{s-1}2^{n-d}C_{n-s}^{d+n-2s}C_s^{s-d}\right)\times$$
$$\times 2^{2s+2-2^{s+1}}\left(\sum_{j=0}^{r}C_{2^s}^{j}\right)^2+\frac{C_n^{n-s}}{2^{2^{s+1}}}\sum_{d=\max\{0,2s-n\}}^{s-1}2^{n+2^d-d}C_{n-s}^{d+n-2s}C_s^{s-d}\times$$
$$\times\left(2^{2s+1-d}\left(|Z_1(l_1,l_2)|+|Z_2(l_1,l_2)|-2|Z_3(l_1,l_2)|\right)+2^{2(s+1)}|Z_3(l_1,l_2)|\right),$$

where

$$|Z_1(l_1,l_2)|=\sum_{k_1=0}^{\min\{2^d,r\}}C_{2^d}^{k_1}\left(\sum_{k_2=0}^{r-k_1}C_{2^s-2^d}^{k_2}\right)^2,$$
$$|Z_2(l_1,l_2)|=\sum_{k_1=\max\{2^d-r,0\}}^{\min\{2^d,r\}}C_{2^d}^{k_1}\sum_{k_2=0}^{r-k_1}C_{2^s-2^d}^{k_2}\sum_{k_4=0}^{r-2^d+k_1}C_{2^s-2^d}^{k_4},$$
$$|Z_3(l_1,l_2)|=\sum_{m_0=0}^{\min\{2^{d-1},\,r\}}C_{2^{d-1}}^{m_0}\times$$
$$\times\sum_{m_1=\max\{0,\,m_0+2^{d-1}-r\}}^{\min\{2^{d-1},\,r-m_0\}}C_{2^{d-1}}^{m_1}\sum_{k_2=0}^{r-(m_0+m_1)}C_{2^s-2^d}^{k_2}\sum_{k_4=0}^{r-(m_0+2^{d-1}-m_1)}C_{2^s-2^d}^{k_4}.$$

References

[1] Logachev O. A., Salnikov A. A., Yashchenko V. V. Boolean functions in coding theory and cryptology (in Russian) — M.: MCCME, 2004.