УДК 519.1, 519.7

НИЖНЯЯ ОЦЕНКА РАССТОЯНИЯ МЕЖДУ БИЮНКТИВНОЙ ФУНКЦИЕЙ И ФУНКЦИЕЙ С ЗАДАННОЙ АЛГЕБРАИЧЕСКОЙ ИММУННОСТЬЮ[1]

А. В. Покровский

Институт проблем информационной безопасности Московского государственного университета им. М. В. Ломоносова, г. Москва, Россия

Получены нижние оценки расстояния Хэмминга (которые могут быть достижимы при выполнении определённых условий) между функцией с заданной алгебраической иммунностью и биюнктивными функциями. Данные оценки позволяют в совокупности оценить устойчивость функции к методу линеаризации, предложенному Н. Куртуа, и возможность её приближения биюнктивными статистическими аналогами.

Ключевые слова: *алгебраическая иммунность, биюнктивные функции, нелинейность, аннуляторы, расстояние между функциями.*

DOI 10.17223/20710410/34/3

A LOWER BOUND FOR THE DISTANCE BETWEEN A BIJUNCTIVE FUNCTION AND A FUNCTION WITH THE FIXED ALGEBRAIC IMMUNITY

A. V. Pokrovskiy

Lomonosov Moscow State University, Moscow, Russia

E-mail: AlexPokrovskiy@yandex.ru

Let $f = f(x_1, \ldots, x_n)$ be a bijunctive Boolean function, that is, the multiplication of some disjunctions of two variables or their negations, $L_f = \{i_1, \ldots, i_{|L_f|}\} \subset \{1, \ldots, n\}$, and, for $\mathbf{y} = (y_1, \ldots, y_{|L_f|}) \in \mathbb{F}_2^{|L_f|}$, the Boolean function $f_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}}$ obtained by substitution of $y_1, \ldots, y_{|L_f|}$ instead of $x_{i_1}, \ldots, x_{i_{|L_f|}}$ respectively into $f(x_1, \ldots, x_n)$ is not const and is equivalent relatively the Jevons group to the function

$$f_{d_{\mathbf{y}},m_{\mathbf{y}}}(\mathbf{x}) = \begin{cases} (x_1 \vee x_2) \cdot \ldots \cdot (x_{2d_{\mathbf{y}}-1} \vee x_{2d_{\mathbf{y}}}) \cdot x_{2d_{\mathbf{y}}+1} \cdot \ldots \cdot x_{2d_{\mathbf{y}}+m_{\mathbf{y}}}, \text{ if } 1 \leqslant d_{\mathbf{y}} \leqslant \lfloor n/2 \rfloor, \\ \qquad\qquad 1 \leqslant m_{\mathbf{y}} \leqslant n - 2d_{\mathbf{y}}; \\ x_1 \cdot \ldots \cdot x_m, \quad \text{if } d_{\mathbf{y}} = 0, 1 \leqslant m_{\mathbf{y}} \leqslant n; \\ (x_1 \vee x_2) \cdot \ldots \cdot (x_{2d_{\mathbf{y}}-1} \vee x_{2d_{\mathbf{y}}}), \quad \text{if } 1 \leqslant d_{\mathbf{y}} \leqslant \lfloor n/2 \rfloor, m_{\mathbf{y}} = 0. \end{cases}$$

Let $f_0 = f_0(x_1, \ldots, x_n)$ be a Boolean function with the algebraic immunity $\mathrm{AI}(f_0)$ satisfying the condition $1 < k = \mathrm{AI}(f_0) - 2|L_f|$, $C = |\{(y_1, \ldots, y_{|L_f|}) \in \mathbb{F}_2^{|L_f|} :$

[1]Работа выполнена при финансовой поддержке Академии криптографии Российской Федерации.

$f^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}} = \mathrm{const}\}|$, and $\mathrm{dist}(f, f_0)$ is the Hamming distance between f and f_0. Then

$$C \sum_{i=0}^{\mathrm{AI}(f_0)-2|L_f|-1} \binom{n-|L_f|}{i} + \sum_{\substack{\mathbf{y}\in\mathbb{F}_2^{|L_f|}:\\ f^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\neq\mathrm{const}}} \Big(\sum_{i=0}^{k-1}\binom{n-|L_f|}{i} + \\ + \sum_{p=0}^{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}} \sum_{j=2d_{\mathbf{y}}+m_{\mathbf{y}}+p-k+1}^{d_{\mathbf{y}}} \Big(2^j\binom{d_{\mathbf{y}}}{j}\binom{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}}{p}\Big) - \\ - \sum_{p=0}^{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}} \sum_{j=0}^{k-1-p} \Big(2^j\binom{d_{\mathbf{y}}}{j}\binom{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}}{p}\Big)\Big) \leqslant \mathrm{dist}\,(f, f_0)\,.$$

In cryptography, functions like f_0 and f are widely used for solving systems of Boolean equations by respectively linearization and statistical approximation methods.

Keywords: *algebraic immunity, bijunctive functions, nonlinearity, annihilator, distance between functions.*

Введение

Решение систем булевых уравнений является одной из актуальных задач дискретной математики. Двумя широко известными подходами к её решению являются метод линеаризации и метод статистических аналогов. При первом подходе система различными путями сводится к линейной системе-следствию. Параметром, характеризующим эффективность одного из вариантов линеаризации, является алгебраическая иммунность булевых функций, задающих левую часть рассматриваемой системы. При втором подходе левая часть каждого уравнения, определяемая некоторой булевой функцией f_i, заменяется на более простую функцию g_i. Правая часть уравнения при этом не меняется. При такой замене не все уравнения остаются верными. Функции g_i выбираются из следующих соображений. Во-первых, вероятность совпадения g_i с f_i при случайном равновероятном выборе аргумента должна быть достаточно велика. Во-вторых, полученная система-следствие должна «легко» решаться. В простейшем случае g_i могут выбираться из класса аффинных функций. В данной работе устанавливается взаимосвязь между алгебраической иммунностью булевой функции и удаленностью её от множества биюнктивных функций, которые могут быть использованы в качестве статистического аналога.

1. Основные обозначения и определения

Пусть $\mathcal{F}_n$ — множество булевых функций от n переменных и $\mathcal{A}_n$ — его подмножество, состоящее из всех аффинных функций. Вес произвольного двоичного вектора $\mathbf{x} \in \mathbb{F}_2^n$ обозначим символом $\mathrm{wt}(\mathbf{x})$. Под *весом* булевой функции $f \in \mathcal{F}_n$ будем понимать вес её вектора значений и записывать его в виде $\mathrm{wt}(f)$. Обозначим D_{i_1,i_2} дизъюнкцию двух булевых переменных

$$D_{i_1,i_2} = x_{i_1} \vee x_{i_2}.$$

С точки зрения метода статистических аналогов одной из основных характеристик булевой функции является её нелинейность. Она определяется через понятие расстояния и характеризует удалённость функции от множества аффинных функций в метрике Хэмминга.

Определение 1. *Расстоянием* между функциями $f, g \in \mathcal{F}_n$ называется

$$\mathrm{dist}(f, g) = \mathrm{wt}(f \oplus g).$$

Определение 2. *Нелинейностью* функции $f \in \mathcal{F}_n$ называется

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

Понятие нелинейности допускает некоторые обобщения. Можно рассматривать минимальное расстояние от функции не только до множества $\mathcal{A}_n$, но и до других практически значимых множеств, например до множества функций степени не выше r. Такой параметр называется нелинейностью r-го порядка. В данной работе будем рассматривать вопрос об удалённости функции от биюнктивных функций.

Определение 3. Булева функция f из $\mathcal{F}_n$ называется *биюнктивной*, если существует представление f в виде 2-КНФ:

$$f(x_1, \ldots, x_n) = \bigwedge_{i=1}^{t} (x_{s_{i1}}^{\alpha_{i1}} \vee x_{s_{i2}}^{\alpha_{i2}}). \tag{1}$$

Множество всех биюнктивных функций обозначим символом $\mathcal{B}\mathfrak{i}_n$. Интерес к биюнктивным функциям объясняется тем, что они порождают классы нелинейных булевых систем, решаемых с полиномиальной трудоемкостью [1, 2].

Определение 4. Для параметров d, m из $\mathbb{N}_0$, одновременно не равных нулю, определим функцию $f_{d,m}$ из множества $\mathcal{B}\mathfrak{i}_n$ как

$$f_{d,m}(\mathbf{x}) = \begin{cases} D_{1,2} \cdot \ldots \cdot D_{2d-1,2d} \cdot x_{2d+1} \cdot \ldots \cdot x_{2d+m}, & \text{если } 1 \leqslant d \leqslant \lfloor n/2 \rfloor,\ 1 \leqslant m \leqslant n - 2d; \\ x_1 \cdot \ldots \cdot x_m, & \text{если } d = 0,\ \ 1 \leqslant m \leqslant n; \\ D_{1,2} \cdot \ldots \cdot D_{2d-1,2d}, & \text{если } 1 \leqslant d \leqslant \lfloor n/2 \rfloor,\ \ m = 0. \end{cases}$$

Важной характеристикой булевых функций является их алгебраическая иммунность. Она определяется через множество аннуляторов.

Определение 5. *Множеством аннуляторов* или просто *аннуляторами* функции $f \in \mathcal{F}_n$ называется множество

$$\mathrm{Ann}(f) = \{g \in \mathcal{F}_n : \forall\, \mathbf{x} \in \mathbb{F}_2^n\, (f(\mathbf{x})g(\mathbf{x}) = 0)\}.$$

Определение 6. *Алгебраической иммунностью* функции f из $\mathcal{F}_n$, обозначаемой символом $\mathrm{AI}(f)$, называется

$$\mathrm{AI}(f) = \min\{\deg(g) : g \in (\mathrm{Ann}(f) \cup \mathrm{Ann}\,(f \oplus 1)) \setminus \{0\}\}.$$

Для булевой функции f определим множество $\mathrm{Ann}_k(f)$.

Определение 7. *Множеством аннуляторов* функции $f \in \mathcal{F}_n$ *степени не выше* k, где $k \in \mathbb{N}$, называется

$$\mathrm{Ann}_k(f) = \{g(\mathbf{x}) \in \mathrm{Ann}(f) : \deg(g) \leqslant k\}.$$

Очевидно, что $(\mathrm{Ann}_k(f), \oplus)$ является пространством над $\mathbb{F}_2$ относительно внешней операции — умножения функции на элемент $\mathbb{F}_2$. Для описания множеств $\mathrm{Ann}_k(f)$ и $\mathrm{Ann}(f)$ удобно использовать матрицы, введённые далее в определении 8. Для их задания используем функции из $\mathcal{F}_n$, полином Жегалкина которых состоит из одного монома. Такие функции в дальнейшем будем обозначать символом **mon**.

Определение 8. Матрица $P_d(f)$ есть матрица над полем $\mathbb{F}_2$, где $d \in \mathbb{N}$, $f \in \mathcal{F}_n$, размера $\sum\limits_{i=0}^{d} \binom{n}{i} \times 2^n$. Её строки состоят из коэффициентов при мономах полиномов Жегалкина функций $\mathbf{mon} \cdot f(\mathbf{x})$, где $\mathbf{mon}$ — ненулевой моном степени не выше d.

Функции $\mathbf{mon} \cdot f(\mathbf{x})$ в дальнейшем будем обозначать $g_{\mathbf{mon}}$, когда ясно, о какой функции f идет речь.

Определение 9. Пусть G — группа преобразований векторов из $\mathbb{F}_2^n$ в $\mathbb{F}_2^n$. Будем называть группу G *подходящей*, если для любых функций f_1, f_2 из $\mathcal{F}_n$ и произвольного преобразования $g \in G$ выполняются равенства

$$\operatorname{dist}(f_1(\mathbf{x}), f_2(\mathbf{x})) = \operatorname{dist}(f_1(g(\mathbf{x})), f_2(g(\mathbf{x}))),$$
$$\operatorname{AI}(f_1(\mathbf{x})) = \operatorname{AI}(f_1(g(\mathbf{x}))),$$

т. е. любое преобразование g из группы сохраняет расстояние между функциями и алгебраическая иммунность инвариантна относительно действия этого преобразования. Если в этой группе существует преобразование g, удовлетворяющее условию

$$f_1(g(\mathbf{x})) = f_2(\mathbf{x}),$$

то функции f_1 и f_2 будем называть *эквивалентными* относительно группы G и обозначать эквивалентность как $f_1 \underset{G}{\sim} f_2$.

В качестве примера подходящей группы можно взять полную аффинную группу.

2. Использование размерности пространства $\operatorname{Ann}_k(f)$ для построения нижних оценок нелинейности функции

Для фиксированной функции $f_{d,m}$ введём на множестве мономов следующую маркировку.

Определение 10. Моном $\mathbf{mon}$ из $\mathcal{F}_n$ относительно фиксированной функции $f_{d,m}$ будем маркировать системой из четырёх множеств

$$\mathbf{mon} = \mathbf{mon}_{V,I,J,L},$$

где

$$\mathbf{mon}_{V,I,J,L} = \prod_{v \in V} (x_{2v-1}x_{2v}) \prod_{i \in I} x_i \prod_{j \in J} x_j \prod_{l \in L} x_l.$$

Первое множество индексов V либо пусто, либо $V = \{v_1, \ldots, v_u\}$ является непустым подмножеством множества $\{1, \ldots, d\}$. Если $V = \{v_1, \ldots, v_u\}$, то оно задаёт номера дизъюнкций в представлении функции $f_{d,m}$, у которых обе переменные входят в моном $\mathbf{mon}$. В случае, когда оно пусто, никакая пара переменных, одновременно входящая в одну дизъюнкцию, в моном не входит, либо $d = 0$.

Второе множество индексов I либо пусто, либо $I = \{i_1, \ldots, i_t\}$ является непустым подмножеством множества $\{1, \ldots, 2d\} \setminus \{2v_1 - 1, 2v_1, \ldots, 2v_u - 1, 2v_u\}$ (в случае, когда $V \neq \varnothing$) или множества $\{1, \ldots, 2d\}$ (в случае $V = \varnothing$). Элементы этого множества задают номера переменных, входящих в дизъюнкцию, отличную от дизъюнкций с номерами из множества V. При этом из рассматриваемой дизъюнкции выбирается лишь одна переменная. Например, во втором множестве не может быть индексов $i_1 = 1$, $i_2 = 2$ одновременно, так как переменные x_1 и x_2 вместе входят в первую дизъюнкцию в представлении $f_{d,m}$. Поэтому во втором множестве возможен либо индекс $i_1 = 1$, либо $i_2 = 2$, но не одновременно оба. Как и в предыдущем случае, если второе множество индексов пусто, то никакие из описанных переменных в моном не входят.

Элементы множества $J = \{j_1, \ldots, j_r\}$ определяют номера переменных, удовлетворяющие условию $2d + 1 \leqslant j_1 < \ldots < j_r \leqslant 2d + m$, которые одновременно входят и в $f_{d,m}$, и в $\mathbf{mon}$. Если таковых нет либо $m = 0$, то множество J полагаем пустым.

Четвёртое множество $L = \{l_1, \ldots, l_s\}$ определяет номера переменных, удовлетворяющих условию $2d + m + 1 \leqslant l_1 < \ldots < l_s \leqslant n$, которые входят в $\mathbf{mon}$. Если таковых нет либо $2d + m = n$, то указанное множество полагаем пустым.

Моном $\mathbf{mon}_{\varnothing,\varnothing,\varnothing,\varnothing}$ по определению будем считать равным 1.

Поясним определение 10 на примере.

Пример 1. Рассмотрим моном $\mathbf{mon}_{\{1,3\},\{3,7\},\{9\},\{12\}}$ и функцию $f_{d,m} = f_{4,2}$ при $n = 12$. Функция $f_{4,2}$ равна

$$f_{4,2} = D_{1,2} \cdot D_{3,4} \cdot D_{5,6} \cdot D_{7,8} \cdot x_9 \cdot x_{10}.$$

Моном $\mathbf{mon}_{\{1,3\},\{3,7\},\{9\},\{12\}}$ в этом случае равен

$$\mathbf{mon}_{\{1,3\},\{3,7\},\{9\},\{12\}} = x_1 \cdot x_2 \cdot x_5 \cdot x_6 \cdot x_3 \cdot x_7 \cdot x_9 \cdot x_{12}.$$

Определим на множестве мономов отношение эквивалентности, которое будем обозначать «$\sim$».

Определение 11. Будем говорить, что два монома $\mathbf{mon}$ и $\mathbf{mon}'$, промаркированные соответствующими множествами индексов относительно фиксированной $f_{d,m}$, удовлетворяют отношению «$\sim$», и обозначать этот факт $\mathbf{mon}_{V,I,J,L} \sim \mathbf{mon}'_{V',I',J',L'}$, если выполняются равенства $V = V'$, $I = I'$, $L = L'$.

В работе [3] установлена связь между нелинейностью булевой функции и её алгебраической иммунностью, а также приведены достижимые нижние оценки нелинейности через алгебраическую иммунность. Эти результаты были получены с помощью утверждения, связывающего величину $\mathrm{dist}(f, f_0)$ с размерностями пространств $\mathrm{Ann}_k(f)$ и $\mathrm{Ann}_k(\overline{f})$. Сформулируем его в том виде, в котором оно приведено в [3].

Утверждение 1. Пусть f и f_0 из $\mathcal{F}_n$, $\mathrm{AI}(f_0) \geqslant k$. Тогда

$$\dim\left(\mathrm{Ann}_{k-1}(f)\right) + \dim\left(\mathrm{Ann}_{k-1}(\overline{f})\right) \leqslant \mathrm{dist}(f, f_0).$$

Дальнейшее направление исследований — получение точного значения размерности пространств $\mathrm{Ann}_{k-1}\left(f_{d,m}\right)$ и $\mathrm{Ann}_{k-1}(\overline{f}_{d,m})$. Эти значения позволяют оценить снизу расстояние между функциями $f_{d,m}$ и функцией f_0 из $\mathcal{F}_n$ с алгебраической иммунностью $\mathrm{AI}(f_0) \geqslant k$. Эти оценки остаются справедливыми и для любых f и f'_0, удовлетворяющих условию $f \underset{G}{\sim} f_{d,m}$ и $f_0 \underset{G}{\sim} f'_0$, где G — подходящая группа. Важным преимуществом данного подхода к построению нижних оценок $\mathrm{dist}(f, f_0)$ является то, что эти оценки при определённых значениях параметров достижимы. Приведём результаты, доказанные в работе [3].

Определение 12. Пусть $f \in \mathcal{F}_n$. Обозначим через $B_k(f)$ линейное пространство функций из $\mathcal{F}_n$ степени не выше k, которые при умножении на h снова дают функции степени не выше k, т. е.

$$B_k(h) = \{g(\mathbf{x}) : \deg(g) \leqslant k,\ \deg(g \cdot f) \leqslant k\}.$$

Следующее утверждение устанавливает связь между размерностями пространств $\mathrm{Ann}_k(f)$, $\mathrm{Ann}_k(\overline{f})$ и $B_k(f)$.

Утверждение 2. Верно равенство

$$\dim(\mathrm{Ann}_k(f)) + \dim(\mathrm{Ann}_k(\overline{f})) = \dim(B_k(f)).$$

Наконец, утверждение 3 даёт ответ на вопрос о достижимости получаемых оценок.

Утверждение 3. Пусть $\deg(f) \leqslant \lceil n/2 \rceil$, $k \leqslant \lceil n/2 \rceil$, $\dim(B_{k-1}(f)) > 0$. Тогда существует функция f_0, такая, что $\mathrm{AI}(f_0) = k$ и $\mathrm{dist}(f, f_0) = \dim(B_{k-1}(f))$.

В качестве функции f из утверждения 3 выберем биюнктивную функцию $f_{d,m}$, удовлетворяющую условию $2d + m \leqslant \lceil n/2 \rceil$, где $n \geqslant 6$ и $3 \leqslant k \leqslant \lceil n/2 \rceil$. Для такой функции $f_{d,m}$ всегда существует квадратичный аннулятор, поэтому $\dim(B_{k-1}(f_{d,m})) > 0$. Выбранная функция удовлетворяет условиям утверждения и для неё существует f_0 с алгебраической иммунностью k, такая, что оценка $\mathrm{dist}(f_{d,m}, f_0)$ достижима.

Теорема 1. Размерность пространства $\mathrm{Ann}_k(\overline{f}_{d,m})$, где $k \in \{1, \ldots, \lceil n/2 \rceil\}$, $0 \leqslant d \leqslant \lfloor n/2 \rfloor$, $m \in \{0, \ldots, n-2d\}$, равна

$$\dim\left(\mathrm{Ann}_k\left(\overline{f}_{d,m}\right)\right) = \sum_{p=0}^{n-2d-m} \sum_{j=2d+m+p-k}^{d} \left(2^j \binom{d}{j} \binom{n-2d-m}{p}\right).$$

Доказательство. В [4] доказано, что любая функция g из $\mathrm{Ann}(f)$, где f — произвольная функция из $\mathcal{F}_n$, представима в виде $g = \overline{f} \cdot h$ для некоторой h из $\mathcal{F}_n$. Следовательно, любая функция из $\mathrm{Ann}(\overline{f}_{d,m})$ представима в виде линейной комбинации функций $g_{\mathbf{mon}} = f_{d,m} \cdot \mathbf{mon}$.

Рассмотрим сначала случай $0 < d$. Промаркируем все мономы относительно функции $f_{d,m}$ согласно определению 10. Из определения 11 следует, что для любых двух эквивалентных мономов $\mathbf{mon}$ и $\mathbf{mon}'$ функции $g_{\mathbf{mon}}$ и $g_{\mathbf{mon}'}$ равны. В связи с этим в дальнейшем будем рассматривать лишь представителей классов эквивалентности, а не сами классы. В качестве представителей выберем мономы вида $\mathbf{mon}_{V,I,\varnothing,L}$.

Докажем рекуррентную формулу для вычисления полинома функции $g_{\mathbf{mon}}$:

$$g_{\mathbf{mon}_{\{v_1,\ldots,v_u\},\{i_1,\ldots,i_t\},\varnothing,L}} = g_{\mathbf{mon}_{\{v_2,\ldots,v_u\},\{2v_1-1,i_1,\ldots,i_t\},\varnothing,L}} \oplus$$
$$\oplus g_{\mathbf{mon}_{\{v_2,\ldots,v_u\},\{2v_1,i_1,\ldots,i_t\},\varnothing,L}} \oplus g_{\mathbf{mon}_{\{v_2,\ldots,v_u\},\{i_1,\ldots,i_t\},\varnothing,L}}.$$

По определению 10 имеем

$$\mathbf{mon}_{\{v_1,\ldots,v_u\},\{i_1,\ldots,i_t\},\varnothing,L} = x_{2v_1-1}x_{2v_1}\,\mathbf{mon}_{\{v_2,\ldots,v_u\},\{i_1,\ldots,i_t\},\varnothing,L} =$$
$$= (D_{2v_1-1,2v_1} \oplus x_{2v_1-1} \oplus x_{2v_1})\,\mathbf{mon}_{\{v_2,\ldots,v_u\},\{i_1,\ldots,i_t\},\varnothing,L}\,.$$

Домножая слева и справа последнее равенство на $f_{d,m}$, получим доказываемую формулу. Применяя её u раз к функции $g_{\mathbf{mon}_{\{v_1,\ldots,v_u\},\{i_1,\ldots,i_t\},\varnothing,L}}$ и возникающим слагаемым, получим в результате, что исходная функция выражается в виде линейной комбинации функций вида $g_{\mathbf{mon}_{\varnothing,\{*,\ldots,*,i_1,\ldots,i_t\},\varnothing,L}}$.

В случае $d = 0$ любой моном имеет маркировку $\mathbf{mon}_{\varnothing,\varnothing,J,L}$. В качестве представителя класса эквивалентности будем рассматривать моном с маркировкой $\mathbf{mon}_{\varnothing,\varnothing,\varnothing,L}$.

Отсюда следует, что в обоих случаях в качестве претендентов на базис пространства $\mathrm{Ann}(\overline{f}_{d,m})$ остаются лишь функции из множества Bas:

$$\mathrm{Bas} = \left\{g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}} : I \text{ и } L \text{ удовлетворяют определению 10}\right\}.$$

Найдём степень функций из Bas, изучим их свойства и вычислим мощность этого множества. Из свойства

$$\begin{aligned} x_i \cdot D_{i,j} &= x_i, \\ x_j \cdot D_{i,j} &= x_j \end{aligned} \tag{2}$$

и определения 10 следует, что $\deg\left(g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}}\right) = 2d + m - |I| + |L|$.

Второе важное свойство рассматриваемых функций заключается в том, что в полином Жегалкина функции $g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}}$ входит единственный моном степени $\deg\left(g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}}\right)$ и для разных функций эти мономы разные. Это непосредственно следует из вида функции $f_{d,m}$ и соотношения (2). Отсюда, в частности, получается, что любая нетривиальная линейная комбинация функций $g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}}$ не равна нулю, а значит, указанные функции линейно независимы.

Найдём мощность множества Bas и размерность $\mathrm{Ann}\left(\overline{f}_{d,m}\right)$. Подсчитаем число мономов вида $\mathbf{mon}_{\varnothing,I,\varnothing,L}$, удовлетворяющих определению 10. Нетрудно видеть, что их количество равно

$$\sum_{p=0}^{n-2d-m} \sum_{j=0}^{d} \left(2^j \binom{d}{j}\right) \binom{n-2d-m}{p} = 3^d 2^{n-2d-m}.$$

Поскольку функции $g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}}$ линейно независимы, то среди них нет одинаковых, т. е. разным мономам $\mathbf{mon}_{\varnothing,I,\varnothing,L}$ соответствуют разные функции, поэтому их число совпадает с числом нужных мономов.

Рассмотрим произвольную нетривиальную линейную комбинацию функций из множества Bas. Покажем, что если в неё входит функция $g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}}$, удовлетворяющая условию $k < \deg(g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}})$, то и степень полученной комбинации больше k. Из второго свойства функции $g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}}$ следует, что она содержит единственный моном максимальной степени, большей k. Для того чтобы при сложении он сократился, необходимо, чтобы в линейной комбинации присутствовала функция с таким же мономом. Рассмотрим функции $g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}}$, входящие в линейную комбинацию и имеющие максимальную степень среди всех слагаемых. Пусть степень таких функций равна $q > k$. Поскольку полином каждой такой функции содержит единственный моном степени q и у разных функций они различны, то такие мономы друг с другом не сократятся. Следовательно, чтобы степень линейной комбинации функций $g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}}$ была не выше k, необходимо, чтобы степень каждой функции $g_{\mathbf{mon}_{\varnothing,I,\varnothing,L}}$, входящей в линейную комбинацию, удовлетворяла условию $\deg(g_{\mathbf{mon}}) \leqslant k$. Это требование эквивалентно выполнению неравенства

$$2d + m - |I| + |L| \leqslant k.$$

Отсюда получаем нижнюю границу $|I|$. Верхняя граница определяется условием $|I| \leqslant d$. Резюмируя все вышесказанное, подсчитаем число таких базисных функций. Нетрудно видеть, что оно вычисляется по формуле

$$\sum_{p=0}^{n-2d-m} \sum_{j=2d+m+p-k}^{d} \left(2^j \binom{d}{j} \binom{n-2d-m}{p}\right),$$

где индекс суммирования p соответствует возможным значениям параметра $|L|$, а индекс j — возможным значениям $|I|$. ■

Перейдём к изучению вопроса о размерности пространства $\mathrm{Ann}_k\left(f_{d,m}\right)$. Значение этой величины может быть вычислено с помощью следующей теоремы.

Теорема 2. Размерность пространства $\mathrm{Ann}_k(f_{d,m})$, где $k \in \{1,\dots,\lceil n/2\rceil\}$, $0 \leqslant d \leqslant \lfloor n/2\rfloor$, $m \in \{0,\dots,n-2d\}$, равна

$$\dim\left(\mathrm{Ann}_k(f_{d,m})\right) = \sum_{i=0}^{k}\binom{n}{i} - \sum_{p=0}^{n-2d-m}\sum_{j=0}^{k-p}\left(2^j\binom{d}{j}\binom{n-2d-m}{p}\right).$$

Доказательство. Для описания пространства $\mathrm{Ann}_k\left(f_{d,m}\right)$ используем матрицу $P_k(f_{d,m})$. Из определения 8 следует, что каждый элемент множества $\mathrm{Ann}_k(f_{d,m})$ соответствует решению системы

$$\mathbf{z}\cdot P_k(f_{d,m}) = 0.$$

Каждая строка матрицы $P_k(f_{d,m})$ построена из коэффициентов функций $g_{\mathbf{mon}}$, полученных с помощью мономов степени не выше k. Следовательно,

$$\dim\left(\mathrm{Ann}_k\left(f_{d,m}\right)\right) = \sum_{i=0}^{k}\binom{n}{i} - \mathrm{rang}\left(P_k(f_{d,m})\right).$$

При доказательстве теоремы 1 получено, что функции из множества Bas, порождаемые мономами $\mathbf{mon}_{\varnothing,I,\varnothing,L}$, линейно независимы. Условие $\deg(\mathbf{mon}) \leqslant k$ эквивалентно неравенству $|I| + |L| \leqslant k$, откуда получаем верхнюю границу на параметр $|I|$:

$$|I| \leqslant k - |L|.$$

Найдём число линейно независимых строк матрицы $P_k(f_{d,m})$, порождённых функциями $g_{\mathbf{mon}}$ из Bas и удовлетворяющих условию $\deg(\mathbf{mon}) \leqslant k$. Учитывая ограничения на параметр $|I|$, получим, что количество таких строк равно

$$\sum_{p=0}^{n-2d-m}\sum_{j=0}^{k-p}\left(2^j\binom{d}{j}\binom{n-2d-m}{p}\right) = \mathrm{rang}\left(P_k(f_{d,m})\right),$$

где индекс суммирования p соответствует возможным значениям параметра $|L|$, а индекс j — возможным значениям $|I|$. ■

Рассмотрим вопрос о расстоянии между функцией f_0 с фиксированной иммунностью и функцией f из множества $\mathcal{B}\mathfrak{i}_n$, у которой сомножители в представлении (1) могут зависеть от пересекающихся множеств переменных или их отрицаний. Обозначим $f_{i_1,\dots,i_k}^{y_1,\dots,y_k}$, где $k \in \mathbb{N}$, сужение функции f из $\mathcal{F}_{n-k}$, получающееся фиксацией переменных с номерами $i_1,\dots,i_k$ константами $y_1,\dots,y_k \in \{0,1\}$ соответственно.

Определение 13. Множество $L_f = \left\{i_1,\dots,i_{|L_f|}\right\} \subset \{1,\dots,n\}$ назовём *разделяющим* множеством функции f из $\mathcal{B}\mathfrak{i}_n$, если для любого вектора $\mathbf{y} = \left(y_1,\dots,y_{|L_f|}\right) \in {\mathbb{F}_2}^{|L_f|}$ функция $f_{i_1,\dots,i_{|L_f|}}^{y_1,\dots,y_{|L_f|}}$ либо константа, либо эквивалентна относительно группы Джевонса функции $f_{d_{\mathbf{y}},m_{\mathbf{y}}}$. Если мощность разделяющего множества равна n или $n-1$, то такое множество будем называть *тривиальным*. Множество L_f минимально возможной мощности будем называть *минимальным*.

Заметим, что поскольку группа Джевонса является подгруппой полной аффинной группы, то она удовлетворяет определению 9. У рассмотренных выше функций $f_{d,m}$ множество $L_{f_{d,m}}$ пустое. В качестве примера разделяющего множества L_f можно взять такие номера, что каждая переменная с номером из L_f или её отрицание входит более чем в один сомножитель в представлении (1). Очевидно, что при любой фиксации

переменной с номером из L_f, входящей в дизъюнкцию, дизъюнкция становится равной либо единице, либо второй входящей в неё переменной, либо её отрицанием.

Сформулируем утверждение [5], устанавливающее связь между весом функции и её алгебраической иммунностью.

Утверждение 4. Если функция $f_0 \in \mathcal{F}_n$ удовлетворяет условию $\mathrm{AI}(f_0) > k$, то

$$\sum_{i=0}^{k} \binom{n}{i} \leqslant \mathrm{wt}(f_0) \leqslant \sum_{i=0}^{n-k-1} \binom{n}{i}.$$

Используя этот результат, продемонстрируем, как могут быть применены теоремы 1 и 2 для оценки расстояния между функциями с заданной алгебраической иммунностью и биюнктивными функциями с непустым разделяющим множеством.

Теорема 3. Пусть $f \in \mathcal{B}\mathfrak{i}_n$, $L_f = \{i_1, \ldots, i_{|L_f|}\} \neq \varnothing$ — разделяющее множество функции f, $f_0 \in \mathcal{F}_n$, $1 < k = \mathrm{AI}(f_0) - 2|L_f|$. Обозначим через C мощность множества

$$C = \left| \left\{ \left(y_1, \ldots, y_{i_{|L_f|}}\right) \in {\mathbb{F}_2}^{|L_f|} : f_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}} = \mathrm{const} \right\} \right|.$$

Пусть $d_{\mathbf{y}}$ и $m_{\mathbf{y}}$, где $\mathbf{y} = (y_1, \ldots, y_{|L_f|}) \in {\mathbb{F}_2}^{|L_f|}$, — параметры функции $f_{d_{\mathbf{y}},m_{\mathbf{y}}} \neq \mathrm{const}$, получаемой после фиксации набора переменных $(x_{i_1}, \ldots, x_{i_{|L_f|}})$ координатами вектора $\mathbf{y}$. Тогда

$$\begin{aligned} C \sum_{i=0}^{\mathrm{AI}(f_0)-2|L_f|-1} \binom{n-|L_f|}{i} + \sum_{\substack{\mathbf{y}\in{\mathbb{F}_2}^{|L_f|}: \\ f_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}} \neq \mathrm{const}}} \left(\sum_{i=0}^{k-1} \binom{n-|L_f|}{i} + \right. \\ + \sum_{p=0}^{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}} \sum_{j=2d_{\mathbf{y}}+m_{\mathbf{y}}+p-k+1}^{d_{\mathbf{y}}} \left(2^j \binom{d_{\mathbf{y}}}{j} \binom{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}}{p} \right) - \\ \left. - \sum_{p=0}^{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}} \sum_{j=0}^{k-1-p} \left(2^j \binom{d_{\mathbf{y}}}{j} \binom{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}}{p} \right) \right) \leqslant \mathrm{dist}\,(f, f_0)\,. \end{aligned}$$

Доказательство. Согласно определению 1, имеет место цепочка равенств

$$\begin{aligned} \mathrm{dist}\,(f, f_0) = \mathrm{wt}(f \oplus f_0) &= \sum_{\mathbf{y}\in{\mathbb{F}_2}^{|L_f|}} \mathrm{wt}\left(f_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}} \oplus (f_0)_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}} \right) = \\ &= \sum_{\mathbf{y}\in{\mathbb{F}_2}^{|L_f|}} \mathrm{dist}\left(f_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}}, (f_0)_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}} \right), \end{aligned} \tag{3}$$

где $f_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}}$ и $(f_0)_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}}$ — сужения функций f и f_0 соответственно, полученные фиксацией переменных с номерами $i_1, \ldots, i_{|L_f|}$ координатами вектора $\mathbf{y} = \left(y_1, \ldots, y_{|L_f|}\right) \in {\mathbb{F}_2}^{|L_f|}$.

Если $f_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}} = \mathrm{const}$, то $\mathrm{dist}\left(f_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}}, (f_0)_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}} \right) = \mathrm{wt}\left((f_0)_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}} \oplus \mathrm{const} \right)$. Поскольку для любой f_0 из $\mathcal{F}_n$ по определению 6 выполняется равенство $\mathrm{AI}(f_0) = \mathrm{AI}(\overline{f}_0)$, то, согласно утверждению 4,

$$\sum_{i=0}^{\mathrm{AI}\left((f_0)_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}} \right)-1} \binom{n-|L_f|}{i} \leqslant \mathrm{wt}\left((f_0)_{i_1,\ldots,i_{|L_f|}}^{y_1,\ldots,y_{|L_f|}} \oplus \mathrm{const} \right). \tag{4}$$

В работе [6] доказано, что для произвольной функции $f_0 \in \mathcal{F}_n$, линейного подпространства $L < \mathbb{F}_2^n$ и вектора $\mathbf{a} \in \mathbb{F}_2^n$ выполняется неравенство

$$\mathrm{AI}(f_0) \leqslant \mathrm{AI}\,|_{L\oplus\mathbf{a}}\,(f_0) + n - \dim(L)\,, \tag{5}$$

где

$$\mathrm{AI}\,|_{L\oplus\mathbf{a}}\,(f_0) = \min\{\deg(g) : \\ g \in \mathcal{F}_n : f(\mathbf{x})g(\mathbf{x}) = 0 \text{ или } (f(\mathbf{x})\oplus 1)\,g(\mathbf{x}) = 0\ \forall\,\mathbf{x} \in L\oplus\mathbf{a},\ \mathrm{supp}(g)\cap L\oplus\mathbf{a} \neq \varnothing\}.$$

Здесь $\mathrm{supp}(g) = \{\mathbf{x} \in \mathbb{F}_2^n : g(\mathbf{x}) = 1\}$. В рассматриваемом случае роль аффинного пространства $L \oplus \mathbf{a}$ выполняет аффинное пространство, состоящее из векторов, у которых координаты с номерами $i_1, \ldots, i_{|L_f|}$ зафиксированы значениями $y_{i_1}, \ldots, y_{i_{|L_f|}}$. Тогда $\dim(L) = n - |L_f|$. Отсюда и из (5) получаем

$$\mathrm{AI}(f_0) - |L_f| \leqslant \mathrm{AI}\,|_{L\oplus\mathbf{a}}\,(f_0)\,. \tag{6}$$

Оценим сверху величину $\mathrm{AI}\,|_{L\oplus\mathbf{a}}\,(f_0)$. Пусть $\mathrm{AI}\left((f_0)^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\right) = \deg(\tau)$, где

$$\tau \in \mathrm{Ann}\left((f_0)^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\right) \cup \mathrm{Ann}\left((f_0\oplus 1)^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\right) \subset \mathcal{F}_{n-|L_f|}.$$

Определим функцию $h = (x_{i_1}\oplus y_1\oplus 1)\cdot\ldots\cdot\left(x_{i_{|L_f|}}\oplus y_{|L_f|}\oplus 1\right)\cdot\tau$. Поскольку функция τ от переменных $x_{i_1}, \ldots, x_{i_{|L_f|}}$ не зависит, то

$$\deg(h) = \deg(\tau) + |L_f| = \mathrm{AI}\left((f_0)^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\right) + |L_f|.$$

Из определения $\mathrm{AI}\,|_{L\oplus\mathbf{a}}\,(f_0)$ следует, что

$$\mathrm{AI}\,|_{L\oplus\mathbf{a}}\,(f_0) \leqslant \mathrm{AI}\left((f_0)^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\right) + |L_f|, \tag{7}$$

поэтому из (6) и (7) получаем неравенство

$$\mathrm{AI}\,(f_0) - 2|L_f| \leqslant \mathrm{AI}\left((f_0)^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\right).$$

Используя последнее соотношение в качестве оценки верхнего индекса суммирования в (4), получаем цепочку неравенств

$$\sum_{i=0}^{\mathrm{AI}(f_0)-2|L_f|-1}\binom{n-|L_f|}{i} \leqslant \sum_{i=0}^{\mathrm{AI}\left((f_0)^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\right)-1}\binom{n-|L_f|}{i} \leqslant \mathrm{wt}\left((f_0)^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\oplus\mathrm{const}\right).$$

Поскольку количество слагаемых в сумме (3), у которых сужение функции f обращается в константу, равно C, их общий вклад в эту сумму снизу можно оценить величиной

$$C\left(\sum_{i=0}^{\mathrm{AI}(f_0)-2|L_f|-1}\binom{n-|L_f|}{i}\right).$$

Для оценки вклада слагаемых, удовлетворяющих условию $f^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}} = f_{d_{\mathbf{y}},m_{\mathbf{y}}} \neq \mathrm{const}$, воспользуемся нижней оценкой алгебраической иммунности сужения функции, теоремами 1, 2 и утверждением 1. Согласно им, если положить $k = \mathrm{AI}(f_0) - 2|L_f|$, выполняется цепочка неравенств

$$\sum_{p=0}^{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}} \sum_{j=2d_{\mathbf{y}}+m_{\mathbf{y}}+p-k+1}^{d_{\mathbf{y}}} \left(2^j \binom{d_{\mathbf{y}}}{j} \binom{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}}{p}\right) +$$

$$+ \sum_{i=0}^{k-1} \binom{n-|L_f|}{i} - \sum_{p=0}^{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}} \sum_{j=0}^{k-1-p} \left(2^j \binom{d_{\mathbf{y}}}{j} \binom{n-|L_f|-2d_{\mathbf{y}}-m_{\mathbf{y}}}{p}\right) =$$

$$= \dim\left(\mathrm{Ann}_{\mathrm{AI}(f_0)-2|L_f|-1}\left(f_{d_{\mathbf{y}},m_{\mathbf{y}}}\right)\right) + \dim\left(\mathrm{Ann}_{\mathrm{AI}(f_0)-2|L_f|-1}\left(\overline{f}_{d_{\mathbf{y}},m_{\mathbf{y}}}\right)\right) \leqslant$$

$$\leqslant \dim\left(\mathrm{Ann}_{\mathrm{AI}\left((f_0)^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\right)-1}\left(f_{d_{\mathbf{y}},m_{\mathbf{y}}}\right)\right) + \dim\left(\mathrm{Ann}_{\mathrm{AI}\left((f_0)^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\right)-1}\left(\overline{f}_{d_{\mathbf{y}},m_{\mathbf{y}}}\right)\right) \leqslant$$

$$\leqslant \mathrm{dist}\left(f_{d_{\mathbf{y}},m_{\mathbf{y}}}, (f_0)^{y_1,\ldots,y_{|L_f|}}_{i_1,\ldots,i_{|L_f|}}\right).$$

Суммируя полученные нижние оценки, получаем утверждение теоремы. ■

ЛИТЕРАТУРА

1. *Горшков С. П.* Применение теории NP-полных задач для оценки сложности решения систем булевых уравнений // Обозрение прикладной и промышленной математики. Сер. Дискретная математика. 1995. Т. 2. Вып. 3. С. 325–398.
2. *Тарасов А. В.* О свойствах функций, представимых в виде 2-КНФ // Дискретная математика. 2001. Т. 13. Вып. 4. С. 99–115.
3. *Лобанов М. С.* О соотношениях между алгебраической иммунностью и нелинейностью булевых функций: дис. ... канд. физ.-мат. наук. М.: МГУ им. М. В. Ломоносова, 2009. 64 с.
4. *Meier W., Pasalic E., and Carlet C.* Algebraic attacks and decomposition of Boolean functions // EUROCRYPT'04. LCNS. 2004. V. 3027. P. 474–491.
5. *Dalai D. K.* On Some Necessary Conditions of Boolean Functions to Resist Algebraic Attacks: PhD Thesis. Kolkata, 2006. 139 p.
6. *Буряков М. Л.* Алгебраические, комбинаторные и криптографические свойства параметров аффинных ограничений булевых функций: дис. ... канд. физ.-мат. наук. М.: МГУ им. М. В. Ломоносова, 2008. 114 с.

REFERENCES

1. *Gorshkov S. P.* Primenenie teorii NP-polnykh zadach dlya otsenki slozhnosti resheniya sistem bulevykh uravneniy [Application of the NP-complete Problems Theory for Estimating the Complexity of Solving Systems of Boolean Equations]. Obozrenie Prikladnoy i Promyshlennoy Matematiki, Ser. Diskr. Mat., 1995, vol. 2, iss. 3, pp. 325–398. (in Russian)
2. *Tarasov A. V.* O svoystvakh funktsiy, predstavimykh v vide 2-KNF [On the properties of functions representable in the form of a 2-CNF]. Diskr. Mat., 2001, vol. 13, iss. 4, pp. 99–115. (in Russian)
3. *Lobanov M. S.* O sootnosheniyakh mezhdu algebraicheskoy immunnost'yu i nelineynost'yu bulevykh funktsiy [On the Relations between the Nonlinearity and Algebraic Immunity of Boolean Functions]. PhD Thesis, Moscow, MSU Publ., 2009. (in Russian)
4. *Meier W., Pasalic E., and Carlet C.* Algebraic attacks and decomposition of Boolean functions. EUROCRYPT'04, LCNS, 2004, vol. 3027, pp. 474–491.
5. *Dalai D. K.* On Some Necessary Conditions of Boolean Functions to Resist Algebraic Attacks: PhD Thesis, Kolkata, 2006. 139 p.

6. *Buryakov M. L.* Algebraicheskie, kombinatornye i kriptograficheskie svoystva parametrov affinnykh ogranicheniy bulevykh funktsiy [Algebraic, Combinatorial, and Cryptographic Properties of Parameters of Boolean functions Affine Restrictions]. PhD Thesis, Moscow, MSU Publ., 2008. (in Russian)