Секция 2

# ДИСКРЕТНЫЕ ФУНКЦИИ

УДК 519.7 DOI 10.17223/2226308X/14/4

## ГИБРИДНЫЙ ПОДХОД К ПОИСКУ БУЛЕВЫХ ФУНКЦИЙ С ВЫСОКОЙ АЛГЕБРАИЧЕСКОЙ ИММУННОСТЬЮ НА ОСНОВЕ ЭВРИСТИЧЕСКИХ МЕТОДОВ[1]

Н. Д. Атутова

Предложен комбинированный подход к поиску булевых функций с высокой алгебраической иммунностью на основе эвристических методов, в частности генетического алгоритма и алгоритма Hill Climbing. Для булевых функций от $n \leqslant 8$ переменных проведены вычислительные эксперименты, продемонстрировавшие эффективность предлагаемого подхода.

**Ключевые слова:** *генетический алгоритм, алгоритм Hill Climbing, алгебраическая иммунность, нелинейность, эвристики.*

Развивающийся интерес к криптоанализу повышает потребность в улучшении стойкости шифров. Для защиты от статистических и аналитических методов криптоанализа для построения компонент шифра необходимо использовать булевы функции, обладающие хорошими криптографическими характеристиками. В 2003 г. N. Courtois и W. Meier в [1] предложили новый метод криптоанализа шифров, названный алгебраическим криптоанализом. Высокая алгебраическая иммунность помогает противостоять такому криптоанализу.

Целью работы является построение булевых функций с максимальной алгебраической иммунностью — характеристикой, повышающей стойкость шифра к алгебраическим атакам.

*Алгебраическая иммунность* булевой функции $f$ ($AI(f)$) — минимальное число $d$, такое, что существует булева функция $g$ степени $d$, не тождественно равная нулю, для которой выполняется равенство $fg = 0$ или $(f \oplus 1)g = 0$, где $f$ и $g$ — функции от равного числа переменных. Известно, что для любой функции $f$ от $n$ переменных справедливо $AI(f) \leqslant \lceil n/2 \rceil$.

Задача полного описания класса булевых функций, обладающих максимальной алгебраической иммунностью, а также получения новых конструкций таких функций является открытой проблемой.

Существует три способа нахождения функций с высокой алгебраической иммунностью: полный перебор, алгебраическое конструирование и эвристики. При росте числа переменных множество булевых функций растёт дважды экспоненциально, что ухудшает эффективность полного перебора. Алгебраическое построение заведомо сужает множество решений. Перспективным является подход, использующий эвристические

[1]Работа выполнена при поддержке Математического центра в Академгородке, соглашение с Министерством науки и высшего образования РФ № 075-15-2019-1613, и лаборатории криптографии JetBrains Research.

методы, в основе которых лежит структурированный перебор с параметрами для достижения желаемого результата.

Предлагается рассмотреть применение эвристических методов, в частности генетического алгоритма и алгоритма Hill Climbing. Специфика применения данных алгоритмов для поиска булевых функций с высокими значениями нелинейности впервые описана в [2]. Получены также теоретические результаты применения этих алгоритмов для функций от 16 переменных. Эффективность эвристических методов продемонстрирована в ряде работ: в [3] исследована возможность применения алгоритма имитации отжига для поиска функций с высокой нелинейностью и низкой автокорреляцией; в [4] реализован генетический алгоритм для поиска бент-функций и сбалансированных функций; в [5] приведены результаты применения гибридного генетического алгоритма для построения сбалансированной булевой функции с оптимальными криптографическими характеристиками.

Для достижения максимального значения алгебраической иммунности реализованы два алгоритма.

**Генетический алгоритм** (ГА) — это метод поиска, аналогичный естественному отбору в природе. Для получения более жизнеспособных потомков к особям из начальной популяции итерационно применяются скрещивание и мутация. Последовательно происходит полное обновление популяции потомками, обладающими наибольшими значениями целевой функцией. В терминах нашей задачи:

— особь — вектор значений булевой функции;
— начальная популяция — случайное множество особей, без ограничений;
— мутация — перестановка двух случайных различных битов вектора значений входной функции;
— целевая функция — алгебраическая иммунность;
— скрещивание — однородный кроссинговер. На вход поступают булевы функции $f$ и $g$, представленные векторами значений $(f_0, f_1, \ldots, f_{2^{n-1}})$ и $(g_0, g_1, \ldots, g_{2^{n-1}})$. На выходе — булева функция $h$, вектор значений $(h_0, h_1, \ldots, h_{2^{n-1}})$ которой определяется следующим образом: если $f_i = g_i$, то $h_i = f_i$; если $f_i \neq g_i$, то $h_i$ принимается равным $f_i$ или $g_i$ с одинаковой вероятностью, $i = 0, 1, \ldots, 2^{n-1}$.

Введём некоторые ограничения на скрещивание. Для этого нам потребуется расстояние Хэмминга.

*Расстоянием Хэмминга* $\mathrm{dist}(f, g)$ между булевыми функциями $f$ и $g$ называется число координат, в которых различаются их векторы значений.

Если $\mathrm{dist}(f, g) > 2^{n-1}$, то вместо вектора значений функции $g$ рассматривается вектор, полученный инверсией всех битов вектора значений функции $g$. В рамках работы вероятность выполнения операции скрещивания принималась равной 0,8.

Далее в таблице представлены экспериментальные результаты применения генетического алгоритма для булевых функций при $n = 4, 6, 8$. Алгоритм повышает значения алгебраической иммунности до максимальной теоретической оценки для всех особей популяции. Подсчитано количество функций с максимальным значением целевой функции, получаемых на каждой итерации при каждом обновлении популяции.

**Hill Climbing** — итерационный алгоритм, который начинается с произвольного решения задачи, а затем пытается найти лучшее путём пошагового изменения одного из элементов решения. В рамках рассматриваемой задачи используется понятие нелинейности — характеристики, повышающей стойкость к линейному криптоанализу [6]. *Преобразованием Уолша — Адамара* булевой функции $f$ называется функция $W_f : \mathbb{F}_2^n \to \mathbb{Z}$,

где

$$W_f(y) = \sum_x (-1)^{f(x) \oplus \langle x,y \rangle}, \quad y \in \mathbb{F}_2^n, \quad \langle x, y \rangle = x_1 y_1 \oplus \ldots \oplus x_n y_n.$$

С помощью преобразования Уолша — Адамара можно вычислить нелинейность функции $f$:

$$N_f = 2^{n-1} - \frac{1}{2} \max_{y \in \mathbb{F}_2^n} |W_f(y)|.$$

При чётном числе переменных $n$ максимально возможное значение нелинейности равно $2^{n-1} - 2^{n/2-1}$. В случае нечётного $n$ точное значение максимальной нелинейности неизвестно.

Алгоритм для повышения нелинейности описан в [2]. На вход поступает вектор значений булевой функции. Алгоритм итеративно пытается улучшить нелинейность, изменяя одну из координат. На каждой итерации коэффициенты Уолша — Адамара разбиваются на множества и последовательно происходит проверка условий повышения значения целевой функции. В настоящей работе Hill Climbing применяется для поддержания высокой нелинейности после мутации потомков на каждой итерации генетического алгоритма.

Известно следующее соотношение, связывающее нелинейность и алгебраическую иммунность булевой функции [7]:

$$N_f \geqslant 2 \sum_{i=0}^{AI(f)-2} \binom{n-1}{i}.$$

Соотношение определяет верхнюю границу алгебраической иммунности: если нелинейность булевой функции достаточно высока, то эта граница увеличивается. В данной работе при поиске булевых функций с максимальной алгебраической иммунностью на каждой итерации поддерживается высокое значение нелинейности для получаемых потомков с помощью алгоритма Hill Climbing. Проведённые эксперименты показали эффективность такого подхода. Результаты экспериментов для $n = 4, 6, 8$ представлены в таблице, где $n$ — число переменных; $P$ — размер популяции; $T$ — число итераций.

**Результаты применения ГА и Hill Climbing**

| $n$ | $P$ | $T$ | min, среднее, max значения $AI(f)$ в исходной популяции | min, среднее, max значения $AI(f)$ после применения ГА | Количество функций с max $AI(f)$ при применении ГА | Количество функций с max $AI(f)$ при применении ГА + Hill Climbing |
|---|---|---|---|---|---|---|
| 4 | 10 | 20 | (0; 1,1; 2) | (2; 2; 2) | 609 | 715 |
| 6 | 10 | 20 | (0; 1,6; 3) | (3; 3; 3) | 649 | 718 |
| 8 | 10 | 20 | (1; 2,5; 4) | (4; 4; 4) | 683 | 703 |
| 8 | 20 | 20 | (1; 2,75; 4) | (4; 4; 4) | 2864 | 2989 |

Полученные булевы функции могут быть использованы при поиске векторных булевых функций с высокой компонентной алгебраической иммунностью. Наличие высокой компонентной алгебраической иммунности S-блоков способствует противостоянию алгебраическому криптоанализу поточных и блочных шифров.

## ЛИТЕРАТУРА

1. *Courtois N. and Meier W.* Algebraic attacks on stream ciphers with linear feedback // LNCS. 2003. V. 2656. P. 345–359.

2. *Millan W., Clark A., and Dawson E.* An effective genetic algorithm for finding highly nonlinear Boolean functions // LNCS. 1997. V. 1334. P. 149–158.
3. *Clark J., Jacob J., Stepney S., et al.* Evolving Boolean functions satisfying multiple criteria // LNCS. 2002. V. 2551. P. 246–259.
4. *Picek S., Jakobovic D., Miller J., et al.* Cryptographic Boolean functions: one output, many design criteria // Appl. Soft Computing. 2016. No. 40. P. 635–653.
5. *Behera P. and Gangopadhyay S.* An improved hybrid genetic algorithm to construct balanced Boolean function with optimal cryptographic properties // Evolutionary Intelligence. 2021. No. 1. P. 1–15.
6. *Matsui M.* Linear cryptanalysis method for DES cipher // LNCS. 1994. V. 765. P. 386–397.
7. *Лобанов М. С.* Точные соотношения между нелинейностью и алгебраической иммунностью // Дискретный анализ и исследование операций. 2008. Т. 15. №6. С. 34–47.

УДК 519.7 DOI 10.17223/2226308X/14/5

# S-БЛОКИ С МАКСИМАЛЬНОЙ КОМПОНЕНТНОЙ АЛГЕБРАИЧЕСКОЙ ИММУННОСТЬЮ ОТ МАЛОГО ЧИСЛА ПЕРЕМЕННЫХ[1]

Д. А. Зюбина, Н. Н. Токарева

Пусть $\pi$ — перестановка $n$ элементов, $f$ — булева функция от $n$ переменных. Рассмотрим векторную булеву функцию $F_\pi : \mathbb{F}_2^n \to \mathbb{F}_2^n$ вида $F_\pi(x) = (f(x), f(\pi(x)), \cdots, f(\pi^{n-1}(x)))$. Изучается компонентная алгебраическая иммунность функции $F_\pi$ в зависимости от булевой функции $f$ и перестановки $\pi$ при $n = 3, 4, 5$. Получены полные множества булевых и частичные векторных булевых функций с максимальной алгебраической иммунностью от малого числа переменных.

**Ключевые слова:** *булева функция, векторная булева функция, алгебраическая иммунность, компонентная алгебраическая иммунность.*

S-блоки играют решающую роль в обеспечении стойкости блочных шифров к различным типам атак. Основная причина этого в том, что в классических и современных блочных шифрах нелинейный слой представлен именно данными блоками. S-блок является отображением множества двоичных векторов длины $n$ в множество двоичных векторов длины $m$. В 2003 г. в [1] был представлен новый вид криптоанализа — алгебраический, основанный на понижении степени системы уравнений, описывающей шифр. Для противостояния такому роду атак необходимо, чтобы S-блок имел максимально возможное значение компонентной алгебраической иммунности.

Будем рассматривать S-блоки определённого вида, а именно $F_\pi(x) = (f(x), f(\pi(x)), \ldots, f(\pi^{n-1}(x)))$, где $F_\pi : \mathbb{F}_2^n \to \mathbb{F}_2^n$; $f$ — булева функция от $n$ переменных; $\pi$ — циклический сдвиг влево на один. Эта конструкция предложена А. Удовенко в решении олимпиадной задачи на NSUCRYPTO-2016 [2]. Он показал, что при таком построении векторной функции можно получить функцию с максимальной алгебраической иммунностью от 3, 4, ..., 10 переменных. В настоящее время остаётся открытым вопрос о существовании векторной булевой функции с максимальной компонентной иммунностью $\lceil n/2 \rceil$ от произвольного числа переменных $n$.

*Алгебраической иммунностью* $\mathrm{AI}(f)$ булевой функции $f$ называется минимальное число $d$, такое, что существует булева функция $g$ степени $d$, не тождественно равная ну-

[1]Работа выполнена в рамках госзадания ИМ СО РАН (проект №FWNF-2022-0018) при поддержке лаборатории криптографии JetBrains Research.