Полная классификация квадратичных APN-функций от 7 переменных

Отчет по работе на вычислительном кластере НГУ за 2020 год

Коллектив:
Токарева Наталья Николаевна, к.ф.-м.н., с.н.с. ИМ СО РАН, доцент каф. комп.систем ФИТ НГУ, каф. теор.киб. ММФ НГУ;
Калгин Константин Викторович, к.ф.-м.н., ст.препр. ФИТ НГУ, лаборатория Криптографии ФИТ НГУ, м.н.с. ИВМиМГ СО РАН;
Идрисова Валерия Александровна, к.ф.-м.н., м.н.с. ИМ СО РАН.
E-mail для связи: kalginkv@gmail.com

1. Определение APN-функций

Через $\mathbb{F}_2^n$ будем обозначать векторное пространство размерности n над полем $\mathbb{F}_2$. Функция F, действующая из $\mathbb{F}_2^n$ в $\mathbb{F}_2^m$, где n и m — некоторые целые числа, называется *векторной булевой функцией*. Каждая векторная булева функция F может быть представлена в виде упорядоченного набора из m *координатных функций* $F = (f_1, \dots, f_m)$, где f_i — булева функция от n переменных. Любая векторная функция F может быть представлена в виде *алгебраической нормальной формы* единственным образом:

$$F(x) = \sum_{I \in \mathcal{P}(N)} a_I \Big(\prod_{i \in I} x_i \Big),$$

где $\mathcal{P}(N)$ — всевозможные сочетания из $N = \{1, \dots, n\}$ и $a_I \in \mathbb{F}_2^m$. Тогда *алгебраической степенью* функции F называют степень ее АНФ: $\deg(F) = \max\{|I| : a_I \neq 0, I \in \mathcal{P}(N)\}$. Если алгебраическая степень функции F равна 2, то F называется *квадратичной*. Далее будем рассматривать только случай $m = n$.

Пусть F — векторная булева функция из $\mathbb{F}_2^n$ в $\mathbb{F}_2^n$. Для векторов $a, b \in \mathbb{F}_2^n$, где $a \neq 0$, рассмотрим величину

$$\delta(a, b) = \big| \{ \, x \in \mathbb{F}_2^n \mid F(x + a) + F(x) = b \} \big|.$$

Обозначим через Δ_F следующую величину:

$$\Delta_F = \max_{a \neq \mathbf{0},\ b \in \mathbb{F}_2^n} \delta(a, b).$$

Тогда функция F называется *дифференциально Δ_F-равномерной* функций. Чем меньше параметр Δ_F, тем лучше стойкость шифра, содержащего функцию F в качестве S-блока, к дифференциальному криптоанализу. Для векторных функций из $\mathbb{F}_2^n$ в $\mathbb{F}_2^n$ минимальное значение Δ_F равно 2. В этом

случае функция F называется ***почти совершенно нелинейной (APN)***. APN функции представляют особый интерес для специалистов в области булевых функций и криптографии - APN функции обеспечивают устойчивость к дифференциальной атаке.

2. Современное состояние проблемы

Несмотря на повышенный интерес к таким функциям, в области остается еще много открытых вопросов. Так, например, неизвестны точное число APN-функций и какие-либо оценки их количества, неизвестны комбинаторные и итеративные конструкции, метрические свойства и т.д. Одним из главных вопросов является вопрос классификации APN-функций, а именно, получения всех функций для конкретной размерности с точностью до CCZ-эквивалентности. Напомним это определение. Две функции F и G называются *CCZ-эквивалентными*, если их графики $\{(x, y) \in \mathbb{F}_2^n \times \mathbb{F}_2^n \mid y = F(x)\}$ и $\{(x, y) \in \mathbb{F}_2^n \times \mathbb{F}_2^n \mid y = G(x)\}$ аффинно эквивалентны, то есть существуют аффинные автоморфизмы $A = (A_1, A_2)$ на $\mathbb{F}_2^n \times \mathbb{F}_2^n$ такие, что $y = F(x) \Leftrightarrow A_2(x, y) = G(A_1(x, y))$.

Полная классификация APN-функций была получена М. Бринкманном и Г. Леандером лишь до $n = 5$. Для $n = 6$ APN-функции классифицированы вплоть до третьей степени. До недавнего времени было известно 487 CCZ-классов квадратичных APN-функций от 7 переменных и 8179 CCZ-классов от 8 переменных. Однако, в совсем недавней работе К. Бейерле и Г. Леандера было найдено 12923 новых квадратичных APN-функций от 8 переменных, 35 новых квадратичных APN-функций от 9 переменных и 5 новых квадратичных APN-функций от 10 переменных. Несмотря на это, полной классификации квадратичных функций от 7 и более переменных к настоящему моменту получено не было.

3. Новая APN-функция от 7 переменных

Квадратичная векторная функция G от n переменных может быть представлена в виде симметричной матрицы $\mathcal{G} = (g_{ij})$, где каждый элемент $g_{ij} \in \mathbb{F}_2^n$ есть вектор коэффициентов, соотвествующих слагаемомуу $x_i x_j$ в АНФ функции G, а диагональные элементы g_{ii} все равны нулю. Через $e_1, \ldots, e_n$ будем обозначать стандартный базис в $\mathbb{F}_2^n$. Для некоторого натурального n рассмотрим следующую матрицу $\mathbb{F}_2^n$:

$$\mathcal{T} = \begin{bmatrix} 0 & e_1 & e_2 & e_3 & \ldots & e_{n-2} & e_{n-1} \\ e_1 & 0 & e_3 & e_4 & \ldots & e_{n-1} & e_n \\ e_2 & e_3 & 0 & e_5 & \ldots & e_n & t_{3,n} \\ e_3 & e_4 & e_5 & 0 & \ldots & t_{4,n-1} & t_{4,n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ e_{n-2} & e_{n-1} & e_n & t_{n-1,4} & \ldots & 0 & t_{n-1,n} \\ e_{n-1} & e_n & t_{n,3} & t_{n,4} & \ldots & t_{n,n-1} & 0 \end{bmatrix},$$

где $t_{i,j} = t_{j,i}$ и $t_{i,j}$ обозначают некоторые неизвестные векторы из $\mathbb{F}_2^n$. Наша задача состоит в том, чтобы означить неизвестные элементы из матрицы так, чтобы функция, соответствующая матрице $\mathcal{T}$ являлась APN-функцией. Перебором значений элементов $t_{i,j}$ на кластере НГУ была найден новая APN-функция при $n = 7$, приводим ее в виде полинома над конечным полем:

$F(x) = a^{100}x+a^{88}x^2+a^{89}x^3+a^{107}x^4+a^{57}x^5+a^{98}x^6+a^{56}x^8+a^9x^9+a^{58}x^{10}+a^{60}x^{12}+a^{109}x^{16}+a^{47}x^{17}+a^{44}x^{18}+a^{27}x^{20}+a^{91}x^{24}+a^{71}x^{32}+a^{96}x^{33}+a^{101}x^{34}+a^7x^{36}+a^{12}x^{40}+a^{34}x^{48}+a^{66}x^{64}+a^4x^{65}+a^4x^{66}+a^{73}x^{68}+a^{73}x^{72}+a^{56}x^{80}+a^{20}x^{96}$,

где a — примитивный элемент, чей минимальный полином над $\mathbb{F}_{2^7}$ равен $x^7 + x + 1$.

4. Классификация квадратичных APN-функций от 7 переменных

Пусть F — квадратичная векторная функция, а $\mathcal{F}$ — соответсвующая ей симметрическая матрица. Нетрудно заметить, что первая строка матрицы $\mathcal{F}$ равняется (0 1 2 4 8 16 32) с точностью до эквивалентности. В работе Y. Yu и др. 2014 года было показано, что если APN-функции F и G эквивалентны, то для их матриц $\mathcal{F}$ и $\mathcal{G}$ выполняется следущее соотношение:

$$\mathcal{G} = L(P\mathcal{F}P^t),$$

где P — обратимая матрица с элементами из $\mathbb{F}_2$, а L — линейная перестановка на $\mathbb{F}_2^n$.

Перебирать все матрицы (квадратичные формы) не представляется возможным, слишком большой перебор. Основная идея сокращения перебора - перебирать лексикографически минимальные матрицы в классе. Использовался метод ветвей и границ - перебор по одной из веток останавливался, если выяснялось, что по данной частично построенной матрице можно построить другую эквивалентную ей, но лексикографически меньшую матрицу.

Опишем кратко алгоритм нахождения лексикографически минимальной матрицы в классе эквивалентности. Наша задача заключается в том, что бы преобразовать первые 2 строки матрицы для того, чтобы получить лексикографически минимальные матрицы, используя только те преобразования, которые сохраняют эквивалентность.

Для всевозможных вариантов первых двух строк матрицы $\mathcal{F}$, удовлетворяющих свойству APN, был реализован поиск всевозможных матриц P вида:

$$\mathcal{P} = \begin{bmatrix} x & x & 0 & 0 & 0 & 0 & 0 \\ x & x & 0 & 0 & 0 & 0 & 0 \\ * & * & x & x & x & x & x \\ * & * & x & x & x & x & x \\ * & * & x & x & x & x & x \\ * & * & x & x & x & x & x \\ * & * & x & x & x & x & x \end{bmatrix},$$

Для каждой такой матрицы P мы:

P1: Перебираем всевозможные L, такие, что первая строка матрицы $\mathcal{G}$ равна (0 1 2 4 8 16 32);

P2: Если пара P и L такова, что $\mathcal{G} < \mathcal{F}$ лексикографически, мы отбрасываем $\mathcal{F}$.

Случаи трёх и четёрых строк описываются аналогично. Эта процедура отбора матриц колоссально сокращает перебор. Но даже после такого сокращения необходимо использование кластера. В полном объёме данная процедура перебора с отсечениями была реализована на кластере. Было

установлено, что существует всего 5 вариантов второй стороки (с точностью до эквивалентности). Ниже перечислены полученные варианты и число функций, которые имеют данную строку в качестве второй строки.

1. Случай (1 0 4 8 16 32 64) содержит 3 квадратичных функции с точностью до эквивалентности;
2. Случай (1 0 4 6 16 32 64) содержит 2 квадратичных функции;
3. Случай (1 0 4 6 16 32 24) не содержит квадратичных функций;
4. Случай (1 0 4 6 16 26 64) содержит 220 квадратичных функций;
5. Случай (1 0 4 6 16 24 64) содержит 263 квадратичных функции.

Результаты

Найден новый 488-й квадратичный класс APN-функций от 7 переменных. Получена полная классификация квадратичных APN-функций от 7 переменных (показано, что других классов нет). Существует всего 488 квадратичных APN-функции с точностью до CCZ-эквивалентности. Начата работа над классификацией квадратичных APN-функций от 8 переменных.

Все вычисления были произведены на вычислительном кластере НГУ в течение 2020 года. Наличие разных очередей (для коротких и долгих запусков программ) позволило эффективно использовать ресурсы и быстро достичь результата.

Грантовая поддержка

Грант РФФИ N 18-07-01394 "Математические методы в современных криптографических приложениях (2018-2020 годы) руководитель - Токарева Наталья Николаевна.

Публикации

1. Kalgin K., Idrisova V. "The classification of quadratic APN functions in 7 variables and combinatorial approaches to search for APN functions"// Cryptography and Communications. Отправлено 2.12.2020 по следам конференции SETA-2020. Рецензирвоание до 1.06.2021. IF=1.291. Quartile Q1/Q2. H-index 13.

2. Kalgin K., Idrisova V. On secondary and cyclic approaches to search for quadratic APN functions // Sequences and Their Applications (Тезисы конференции SETA-2020).

3. K. V. Kalgin, V. A. Idrisova. On a secondary construction of quadratic APN functions. // Прикладная дискретная математика. Приложение, 2020, N 13, с.37–39. (Тезисы SIBECRYPT'20)