ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Октябрь–декабрь 2022. Т. 29, № 4. С. 38–58

УДК 519.7 DOI 10.33048/daio.2022.29.730

ПРИМЕНЕНИЕ SAT-РЕШАТЕЛЕЙ К ЗАДАЧЕ ПОИСКА ВЕКТОРНЫХ БУЛЕВЫХ ФУНКЦИЙ С ТРЕБУЕМЫМИ КРИПТОГРАФИЧЕСКИМИ СВОЙСТВАМИ

А. Е. Доронин[1, a], *К. В. Калгин*[2, 3, b]

[1] Новосибирский гос. университет,
ул. Пирогова, 2, 630090 Новосибирск, Россия

[2] Институт математики им. С. Л. Соболева,
пр. Акад. Коптюга, 4, 630090 Новосибирск, Россия

[3] Институт вычислительной математики и математической геофизики,
пр. Акад. Лаврентьева, 6, 630090 Новосибирск, Россия

E-mail: [a] artem96dor@gmail.com, [b] kalginkv@gmail.com

Аннотация. Представлен подход к решению задачи поиска почти совершенно нелинейной (APN) функции, основанный на её сведении к классической задаче выполнимости и использовании SAT-решателей. Описано построение формул, определяющих APN-функцию. Введены два представления функции: разреженное и плотное, в которых описана задача поиска взаимно однозначной векторной булевой функции и APN-функции. Также в работе представлен новый подход к решению задачи построения векторных булевых APN-функций, обладающих дополнительными свойствами. В основе подхода лежит идея представления неизвестной векторной булевой функции в виде суммы известной APN-функции и двух неизвестных булевых функций: $\mathbf{G} = \mathbf{F} \oplus \mathbf{c} \cdot g_1 \oplus \mathbf{d} \cdot g_2$, где $\mathbf{F}$ — известная APN-функция. Показано, что для функций от $n = 6, 7$ переменных такой подход имеет большую эффективность в сравнении с прямым построением APN-функции при помощи SAT. Как итог, описанным в работе методом удалось показать отсутствие кубических APN-функций от 7 переменных, представимых в виде описанной выше суммы. Табл. 3, библиогр. 21.

Ключевые слова: SAT-решатель, криптография, булева функция, APN-функция.

Исследование выполнено в рамках государственного задания Института математики им. С. Л. Соболева (проект № FWNF–2022–0018).

Введение

В данной работе представлено несколько подходов к решению задачи поиска таблично заданных векторных булевых функций с некоторыми криптографическими свойствами, такими как взаимная однозначность и дифференциальная равномерность, основанных на сведении исходной задачи к классической задаче выполнимости и использовании SAT-решателей.

Напомним некоторые основные определения.

Векторным пространством $\mathbb{Z}_2^n$ называется набор всех векторов размера n над полем $\mathbb{Z}_2$. *Булевой функцией* от n переменных называется отображение $f\colon \mathbb{Z}_2^n \to \mathbb{Z}_2$. *Векторной булевой функцией* от n переменных называется отображение $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^m$.

Определение 1. Векторная булева функция $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ называется *дифференциальной δ-равномерной* [1], если при любом ненулевом векторе $a \in \mathbb{Z}_2^n$ и произвольном векторе $b \in \mathbb{Z}_2^n$ уравнение $F(x) \oplus F(x \oplus a) = b$ имеет не более δ решений.

Если $\delta = 2$, то функция называется APN-*функцией*.

APN-функции представляют большой интерес для криптографии, поскольку являются наиболее дифференциально равномерными, т. е. их использование в качестве нелинейных компонент предельно повышает стойкость шифра к дифференциальному криптоанализу. В области APN-функций много открытых проблем. Среди них — построение и классификация APN-функций. Они детально освещаются, например, в таких публикациях, как [2–8]. Наша работа посвящена методам построения APN-функций.

Первый подход заключается в формулировке SAT-задачи для поиска произвольной APN-функции. Во втором подходе строятся APN-функции определённого вида, а именно

$$\mathbf{G}(x) = \mathbf{F}(x) \oplus \mathbf{c} \cdot g_1(x) \oplus \mathbf{d} \cdot g_2(x),$$

где $\mathbf{F}$ — известная APN-функция от n переменных, $\mathbf{c}$ и $\mathbf{d}$ — известные векторы, а g_1 и g_2 — неизвестные булевы функции, которые нужно найти с помощью SAT-решателя или доказать, что подходящих функций не существует.

Напомним постановку задачи выполнимости для булевой формулы.

Задача SAT (задача выполнимости)**.** *Можно ли присвоить переменным некоторой булевой* КНФ*-формулы значения «истина» или «ложь» таким образом, чтобы формула стала истинной.*

В общем случае данная задача NP-полна [9]. Напомним, что NP-полнота задачи означает, что к ней можно свести все задачи из класса

NP за полиномиальное время, и если найдётся полиномиальный алгоритм решения NP-полной задачи, то и все остальные задачи будут решены за полиномиальное время. В настоящее время задаче выполнимости посвящено большое число исследований, а программы, решающие данную задачу, используются во многих областях: проверка моделей, теория расписаний, искусственный интеллект, криптография и многие другие. В связи с этим ежегодно проводятся конкурсы программ, так называемые соревнования SAT-решателей (SAT competitions) [10].

SAT-решатель — это программа, которая проверяет выполнимость формулы, записанной в конъюнктивной нормальной форме (КНФ-формулы). Первые SAT-решатели основывались на алгоритме DPLL (алгоритм Дэвиса — Патнема — Логемана — Лавленда) [11] — полном алгоритме поиска с возвратом для решения задачи выполнимости булевых КНФ-формул. Алгоритм произвольным образом выбирает переменную, присваивает ей значение «истина», упрощает формулу и рекурсивным образом проверяет её на выполнимость. Если алгоритм сталкивается с конфликтом (подмножеством означиваний переменных, при котором возникают ложные дизъюнкции), то он выполняет возврат к переменной и присваивает ей значение «ложь». Основное отличие алгоритма DPLL от перебора в глубину — это наличие вывода значений других переменных после присвоения значения переменной на текущем шаге (constant propagation). Данный алгоритм показал высокую эффективность для ряда практических задач. Непосредственной модификацией алгоритма DPLL является алгоритм CDCL (conflict driven clause learning) [12], который лежит в основе современных SAT-решателей. Аналогично алгоритму DPLL в алгоритме CDCL переменная выбирается и означивается в соответствии с реализованным эвристическим алгоритмом. Основное отличие алгоритма — сохранение значений части переменных в виде новых дизъюнктов, которые не ведут к решению. Это позволяет эффективно отсекать подпространства наборов значений переменных, приводящих к конфликтам. Разработчики алгоритма CDCL создали SAT-решатель GRASP [12]. Алгоритм CDCL используется практически во всех современных решателях.

SAT-решатели используются для решения многих криптографических задач, например для проведения криптоанализа шифрсистем. В [13] рассматривается задача факторизации целых чисел, на сложности которой основана криптосистема RSA. Суть работы заключается в построении и использовании методов, которые находят приближённое решение SAT-задачи. Полученным алгоритмом удалось факторизовать число размерностью до 417 бит в двоичной записи.

В [14] представлена гомоморфная криптосистема с открытым ключом, основанная на задаче выполнимости булевых формул. В данной

системе открытым ключом является сама булева формула, секретным ключом — означивание этой формулы, при которой она станет истинной. Напомним, что гомоморфным шифрованием называется способ шифрования данных, позволяющий проводить некоторые вычислительные операции над зашифрованными данными так, что результат после расшифрования совпадает с результатом операций над открытыми данными. Это понятие впервые было введено в [15].

Работа [16] посвящена проверке обратимости векторных булевых функций. Показано, что эта задача coNP-полна. Предлагается два подхода: с использованием SAT-решателей и бинарных диаграмм решений (BDD).

Приведём структуру данной статьи. В разд. 1 описывается два способа записи задачи поиска криптографических функций. В разд. 2 и 3 описано построение КНФ-формул для задач поиска взаимно однозначной векторной булевой функции и APN-функции соответственно. Подход, описанный в разд. 3, имеет хорошие теоретические оценки, однако он показал свою непрактичность для поиска APN-функций от 6 переменных и более. В разд. 4 представлен новый подход в поиске APN-функций специального вида, в основе которого лежит метод сдвига. Также в этом разделе представлено улучшение данного метода — метод двойного сдвига, в котором APN-функция представляется в виде суммы известной векторной APN-функции и двух неизвестных булевых функций. Этот метод сводится к решению системы квадратичных уравнений. По этой причине используются SAT-решатели.

1. Задание функции

В работе представлены способы записи криптографических свойств с помощью двух представлений — разреженного и плотного.

1.1. Разреженное представление. Для задания векторной булевой функции $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ введём следующие базовые булевы переменные:

$$f_{x,y} = 1 \Leftrightarrow F(x) = y, \quad x, y \in \mathbb{Z}_2^n.$$

Такое представление булевой функции будем называть *разреженным* (по аналогии с разреженными матрицами). Число булевых переменных $f_{x,y}$ равно 2^{2n}.

Теорема 1. *Множество булевых переменных $f_{x,y}$ кодирует функцию $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ тогда и только тогда, когда следующая формула является истинной:*

$$\mathbf{F}^{\mathbf{S}}(f) = \bigwedge_{x\in\mathbb{Z}_2^n} \bigwedge_{\substack{y',y''\in\mathbb{Z}_2^n,\\ y'<y''}} (\overline{f}_{x,y'} \vee \overline{f}_{x,y''}) \wedge \bigwedge_{x\in\mathbb{Z}_2^n} \Big(\bigvee_{y\in\mathbb{Z}_2^n} f_{x,y} \Big). \tag{1}$$

Доказательство. Чтобы переменные $f_{x,y}$ задавали функцию, необходимо, чтобы для каждого элемента из множества прообразов существовал единственный образ, т. е. $\forall\, x \in \mathbb{Z}_2^n\ \exists!\, y \in \mathbb{Z}_2^n\colon f_{x,y} = 1$.

В КНФ это условие записывается в два этапа.

1) Существование образа для каждого прообраза:

$$\forall\, x\ \exists\, y\colon f_{x,y} = 1 \text{ или } \bigwedge_{x \in \mathbb{Z}_2^n} \Big(\bigvee_{y \in \mathbb{Z}_2^n} f_{x,y} \Big).$$

2) Существование не более чем одного образа для каждого прообраза:

$$\forall\, x\ \forall\, y' < y''\ (f_{x,y'} \to \overline{f}_{x,y''}) \text{ или } \bigwedge_{x \in \mathbb{Z}_2^n} \bigwedge_{\substack{y',y'' \in \mathbb{Z}_2^n,\\ y' < y''}} (\overline{f}_{x,y'} \vee \overline{f}_{x,y''}).$$

Объединяя эти формулы при помощи конъюнкции, получаем формулу (1). Теорема 1 доказана.

Таким образом, формула $\mathbf{F}^{\mathbf{S}}(f)$ состоит из $2^{3n-1} - 2^{2n-1}$ дизъюнкций длины 2 и 2^n дизъюнкций длины 2^n.

1.2. Плотное представление. Представим векторную булеву функцию $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ следующим образом:

$$F(x) = (F_0(x), F_1(x), \ldots, F_{n-1}(x)),$$

где $F_i\colon \mathbb{Z}_2^n \to \mathbb{Z}_2$, $i = 0, \ldots, n-1$, — координатные булевы функции. Таким образом, в плотном представлении удобно ввести следующие базовые булевы переменные:

$$fb_{x,k} = 1 \Leftrightarrow F_k(x) = 1, \quad k = 0, \ldots, n-1,\ x \in \mathbb{Z}_2^n.$$

Такое представление булевой функции будем называть *плотным*. Буква b в обозначении fb взята от слова binary ввиду записи в данном представлении значений $F(x)$ в двоичной записи. Число булевых переменных $fb_{x,k}$ равно $n \cdot 2^n$.

В данном представлении каждому прообразу $x \in \mathbb{Z}_2^n$ ставится в соответствие некоторый образ $y \in \mathbb{Z}_2^n$. Таким образом, наложения дополнительных условий на булевы переменные $fb_{x,k}$, как это было сделано для разреженного представления в теореме 1, не требуется.

2. Взаимная однозначность

Определение 2. Векторная булева функция $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ называется *взаимно однозначной* [17], если она инъективна и сюръективна, т. е. одновременно выполняются следующие условия:

(1) $\forall\, x' \in \mathbb{Z}_2^n\ \forall\, x'' \in \mathbb{Z}_2^n\ (x' \neq x'' \to F(x') \neq F(x''))$,

(2) $\forall\, y \in \mathbb{Z}_2^n\ \exists\, x \in \mathbb{Z}_2^n\colon F(x) = y$.

Иными словами, функция F взаимно однозначная, если у каждого образа $y \in \mathbb{Z}_2^n$ существует прообраз $x \in \mathbb{Z}_2^n$, причём различным прообразам соответствуют различные образы.

Утверждение 1. *Векторная булева функция $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ взаимно однозначна, если выполняется хотя бы одно из условий определения* 2.

Доказательство. По определению функции каждый вектор $x \in \mathbb{Z}_2^n$ отображается в некоторый единственный образ $F(x) \in \mathbb{Z}_2^n$.

Поскольку множества образов и прообразов совпадают, выполнение условия инъективности или сюръективности, т. е. первого или второго условия определения 2 соответственно, очевидным образом гарантирует взаимную однозначность функции F. Утверждение 1 доказано.

Взаимная однозначность векторной булевой функции необходима, например, для использования функции в качестве S-блока в различных шифрах и криптосистемах.

2.1. Разреженное представление.

Теорема 2. *Множество булевых переменных $f_{x,y}$ кодирует взаимно однозначную функцию $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ тогда и только тогда, когда истинна следующая формула:*

$$\mathbf{P^S}(f) = \bigwedge_{y \in \mathbb{Z}_2^n} \bigwedge_{\substack{x', x'' \in \mathbb{Z}_2^n, \\ x' < x''}} (\overline{f}_{x',y} \vee \overline{f}_{x'',y}) \wedge \mathbf{F^S}(f). \tag{2}$$

Доказательство. В разреженном представлении для кодирования векторной взаимно однозначной булевой функции помимо условия, задающего функцию, должно выполняться ещё одно условие, показывающее, что у каждого образа существует только один прообраз. В результате оба условия можно записать следующим образом:

$$\begin{cases} \forall\, x\ \exists!\, y\colon f_{x,y} = 1 \text{ — условие для прообразов,} \\ \forall\, y\ \exists!\, x\colon f_{x,y} = 1 \text{ — условие для образов.} \end{cases}$$

Условие для прообразов представлено в теореме 1. Процесс получения условия для образов в КНФ разбивается на два этапа.

(1) Существование хотя бы одного прообраза у каждого образа гарантирует формула $\mathbf{F^S}(f)$ из теоремы 1.

(2) Существование не более чем одного прообраза у каждого образа гарантирует следующая формула:

$$\forall\, y, x', x'' \in \mathbb{Z}_2^n\ \left(f_{x',y} \to \overline{f}_{x'',y}\right) \text{ или } \bigwedge_{y \in \mathbb{Z}_2^n} \bigwedge_{\substack{x', x'' \in \mathbb{Z}_2^n, \\ x' < x''}} \left(\overline{f}_{x',y} \vee \overline{f}_{x'',y}\right).$$

Поскольку эти условия должны выполняться одновременно, запишем их через конъюнкцию. Таким образом получаем формулу (2). Теорема 2 доказана.

Тем самым в формуле $\mathbf{P^S}(f)$ теоремы 2 содержится $2^{3n} - 2^{2n}$ дизъюнкций длины 2 и 2^n дизъюнкций длины 2^n.

2.2. Плотное представление. Запишем условие, задающее взаимную однозначность, в плотном представлении. Это можно сделать двумя способами: воспользовавшись соответственно первым или вторым условием из определения 2.

Представление через инъективность. Первое условие определения 2 означает, что все образы попарно различны, т. е. любые два вектора вида $\mathbf{fb}_i = (fb_{i,0}, \ldots, fb_{i,n-1})$ отличаются хотя бы в одной компоненте. Это условие задаётся следующим образом:

$$\forall\, i, j \neq i\ \exists\, k\colon (fb_{i,k} \neq fb_{j,k}),$$

где $i, j \in \mathbb{Z}_2^n$, $k = 0, \ldots, n-1$.

Также данное условие можно записать в виде следующей формулы:

$$\mathbf{P^D_{sum}}(fb) = \bigwedge_{\substack{i,j\in\mathbb{Z}_2^n,\\ i\neq j}} \left(\bigvee_{k=0}^{n-1} (fb_{i,k} \oplus fb_{j,k}) \right).$$

Чтобы записать эту формулу в КНФ, воспользуемся преобразованием Цейтина [18]. Для этого введём дополнительные булевы переменные:

$$fbq_{i,j,k} = 1 \Leftrightarrow fb_{i,k} \oplus fb_{j,k} = 1, \tag{3}$$

где $i, j \in \mathbb{Z}_2^n$, $i \neq j$, $k = 0, \ldots, n-1$. Число булевых переменных $fbq_{i,j,k}$ равно $n \cdot (2^{2n} - 2^n)$.

Формула $\mathbf{SoP^D}$(sum of pairs), которая связывает булевы переменные $fbq_{i,j,k}$ и $fb_{i,k}$ для любых $i, j \in \mathbb{Z}_2^n$, $k = 0, \ldots, n-1$, получается по таблице истинности для выражения (3):

$$\mathbf{SoP^D}(fb, fbq) = \bigwedge_{i,j,k} (fbq_{i,j,k} \vee fb_{i,k} \vee \overline{fb}_{j,k}) \wedge (fbq_{i,j,k} \vee \overline{fb}_{i,k} \vee fb_{j,k}) \\ \wedge (\overline{fbq}_{i,j,k} \vee fb_{i,k} \vee fb_{j,k}) \wedge (\overline{fbq}_{i,j,k} \vee \overline{fb}_{i,k} \vee \overline{fb}_{j,k}),$$

где конъюнкция берётся по всем $i, j \in \mathbb{Z}_2^n$, $k = 0, \ldots, n-1$.

Теорема 3. *Булевы переменные $fbq_{i,j,k}$ и $fb_{i,k}$ кодируют взаимно однозначную векторную булеву функцию $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ тогда и только тогда, когда истинна следующая формула:*

$$\mathbf{P}^{\mathbf{D}}_{\mathbf{sumCNF}}(fb, fbq) = \bigwedge_{\substack{i,j\in\mathbb{Z}_2^n,\\ i\neq j}} \left(\bigvee_{k=0}^{n-1} fbq_{i,j,k} \right) \wedge \mathbf{SoP}^{\mathbf{D}}(fb, fbq). \qquad (4)$$

Доказательство. Напомним, что у любой пары прообразов образы должны различаться хотя бы в одной координате. Тогда в переменных $fbq_{i,j,k}$ условие взаимной однозначности представляется в виде

$$\bigwedge_{\substack{i,j\in\mathbb{Z}_2^n,\\ i\neq j}} \left(\bigvee_{k=0}^{n-1} fbq_{i,j,k} \right).$$

Поскольку формула $\mathbf{SoP}^{\mathbf{D}}(fb, fbq)$ и полученные условия должны выполняться одновременно, они должны быть записаны через конъюнкцию. Таким образом получаем формулу (4). Теорема 3 доказана.

Тем самым в формуле $\mathbf{P}^{\mathbf{D}}_{\mathbf{sum}}(fb, fbq)$ содержится $4n(2^{2n} - 2^n)$ дизъюнкций длины 3 и $2^{2n} - 2^n$ дизъюнкций длины n.

Представление через сюръективность. Второе условие определения 2 означает, что каждому образу сопоставляется единственный прообраз, т. е. для любого целого числа $y < 2^n$ существует единственное число x такое, что $\mathbf{fb}_x = \mathbf{y}$, где $\mathbf{fb}_x$ — вектор двоичных переменных, т. е. $\mathbf{fb}_x = (fb_{x,0}, fb_{x,1}, \ldots, fb_{x,n-1})$, и $\mathbf{y}$ — двоичное представление числа y, т. е. $\mathbf{y} = (y_0, y_1, \ldots, y_{n-1})$.

Запишем данное условие при помощи логических операций:

$$\bigwedge_y \bigvee_x \left(fb_{x,0}^{y_0} \wedge fb_{x,1}^{y_1} \wedge \cdots \wedge fb_{x,n-1}^{y_{n-1}} \right),$$

где

$$x^b = \begin{cases} x, & \text{если } b = 1, \\ \overline{x}, & \text{если } b = 0. \end{cases}$$

Для преобразования данной формулы в КНФ воспользуемся булевыми переменными из разреженного представления $f_{x,y}$. Можно заметить, что

$$f_{x,y} = 1 \Leftrightarrow fb_{x,0}^{y_0} \wedge fb_{x,1}^{y_1} \wedge \cdots \wedge fb_{x,n-1}^{y_{n-1}} = 1.$$

Получим выражение в КНФ, связывающее булевы переменные $fb_{x,k}$ и $f_{x,y}$. Для этого преобразуем полученную выше формулу с помощью эквивалентных преобразований. В итоге имеем

$$\mathbf{SpDen}(f, fb) = \bigwedge_{x,y,k} (\overline{f}_{x,y} \vee fb_{x,k}) \wedge \left(f_{x,y} \vee fb_{x,0}^{\overline{y}_0} \vee \cdots \vee fb_{x,n-1}^{\overline{y}_{n-1}} \right),$$

где конъюнкция берётся по всем $x, y \in \mathbb{Z}_2^n$, $k = 0, \ldots, n-1$.

Теорема 4. *Булевы переменные $f_{x,y}$ и $fb_{x,k}$ кодируют взаимно однозначную векторную булеву функцию $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ тогда и только тогда, когда истинна следующая формула:*

$$\mathbf{P}^{\mathbf{D}}_{\mathbf{sparse}}(f, fb) = \bigwedge_{y\in\mathbb{Z}_2^n} \Big(\bigvee_{x\in\mathbb{Z}_2^n} f_{x,y}\Big) \wedge \mathbf{SpDen}(f, fb). \tag{5}$$

Доказательство. Используя новые переменные $f_{x,y}$, второе условие определения 2 записываем следующим образом: $\bigwedge\limits_{y} \big(\bigvee\limits_{x} f_{x,y}\big)$.

Поскольку формула выше и формула $\mathbf{SpDen}(f, fb)$ должны выполняться одновременно, они должны быть записаны через конъюнкцию. Таким образом получаем формулу (5). Теорема 4 доказана.

Итого в формуле $\mathbf{P}^{\mathbf{D}}_{\mathbf{sparse}}(f, fb)$ содержится по $n(2^{2n}-2^n)$ дизъюнкций длины 2 и n, а также 2^n дизъюнкций длины 2^n.

3. Дифференциальная равномерность

Дифференциальная равномерность препятствует проведению дифференциального криптоанализа блочных шифров [1]. В данной работе представлена запись дифференциальной равномерности для $\delta = 2$ (APN-функция). Поиск APN-функций является открытой проблемой для чётных $n > 6$. Для $n = 6$ была найдена одна взаимно однозначная APN-функция или APN-перестановка. [19]

3.1. Разреженное представление. В разреженном представлении дифференциальную равномерность можно записать, вводя следующие булевы переменные:

$$d_{x,a,b} = 1 \Leftrightarrow F(x) \oplus F(x \oplus a) = b, \quad x, a, b \in \mathbb{Z}_2^n. \tag{6}$$

Заметим, что число булевых переменных $d_{x,a,b}$ равно 2^{3n}.

Из определения дифференциальной δ-равномерной функции следует, что если x — корень уравнения $F(x)\oplus F(x\oplus a) = b$, то $x\oplus a$ также является корнем уравнения, поэтому будем считать, что x — наименьший из этих двух корней в лексикографическом порядке.

Для записи условия (6), используя только что введённые булевы переменные $d_{x,a,b}$ и основные переменные $f_{x,y}$, необходимо для каждых фиксированных b, a и x просматривать всевозможные значения $F(x)$ и $F(x \oplus a)$ и проверять, равна ли их сумма значению b. Для этого рассмотрим следующие высказывания и определим, являются ли они истинными.

(1°) x — не корень уравнения из условия (6), $F(x) \neq z$ и $F(x \oplus a) \neq z \oplus b$ истинно по определению дифференциальной равномерности.

(2°) x — не корень уравнения из условия (6), $F(x) = z$ и $F(x \oplus a) \neq z \oplus b$ истинно, аналогично высказыванию (1°).

(3°) x — не корень уравнения из условия (6), $F(x) \neq z$ и $F(x \oplus a) = z \oplus b$ истинно, аналогично высказыванию (1°).

(4°) x — не корень уравнения из условия (6), $F(x) = z$ и $F(x \oplus a) = z \oplus b$ ложно, поскольку $F(x) \oplus F(x \oplus a) = b$, что противоречит условию «x — не корень уравнения из условия (6)».

(5°) x — корень уравнения из условия (6), $F(x) \neq z$ и $F(x \oplus a) \neq z \oplus b$ истинно, поскольку для аргумента x существует другое значение z, при котором $F(x) = z$ и $F(x \oplus a) = z \oplus b$.

(6°) x — корень уравнения из условия (6), $F(x) = z$ и $F(x \oplus a) \neq z \oplus b$ ложно, поскольку $F(x) \oplus F(x \oplus a) \neq b$, что противоречит условию «x — корень уравнения из условия (6)».

(7°) x — корень уравнения из условия (6), $F(x) \neq z$ и $F(x \oplus a) = z \oplus b$ ложно, поскольку x — корень, а значит, $x \oplus a$ также является корнем уравнения из условия (6). Дальнейшие рассуждения аналогичны предыдущему высказыванию.

(8°) x — корень уравнения из условия (6), $F(x) = z$ и $F(x \oplus a) = z \oplus b$ истинно по определению дифференциальной равномерности.

Эти высказывания можно записать при помощи следующей формулы:

$$d_{x,a,b} = 1 \Leftrightarrow \exists\, z\colon (f_{x,z} \wedge f_{x\oplus a, z\oplus b}).$$

Построим по этой формуле таблицу истинности, связывающую булевы переменные $f_{x,y}$ и $d_{x,a,b}$.

Таблица 1

Связь переменных $f_{x,y}$ и $d_{x,a,b}$

$d_{x,a,b}$	$f_{x,z}$	$f_{x\oplus a, z\oplus b}$	Значение
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	1

Согласно алгоритму построения КНФ по табл. 1 выпишем формулу $\mathbf{Der^S}(f, d)$ через конъюнкцию дизъюнкций с нулевыми значениями:

$$\mathbf{Der^S}(f,d) = \bigwedge_{x,z,a,b} (f_{x,z} \vee \overline{f}_{x\oplus a,z\oplus b} \vee \overline{d}_{x,a,b}) \wedge (\overline{f}_{x,z} \vee f_{x\oplus a,z\oplus b} \vee \overline{d}_{x,a,b}) \wedge (\overline{f}_{x,z} \vee \overline{f}_{x\oplus a,z\oplus b} \vee d_{x,a,b}),$$

где конъюнкция берётся по всем $x, z, a, b \in \mathbb{Z}_2^n$, $a \neq 0$.

Теорема 5. *Отображение $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ является* APN*-функцией тогда и только тогда, когда выполняются условия теоремы* 1 *и истинна следующая формула:*

$$\mathbf{APN^S}(f,d) = \mathbf{Der^S}(f,d) \wedge \bigwedge_{\substack{a,b,x,y\in\mathbb{Z}_2^n,\\ y\neq x,\, y\neq x\oplus a,\, a\neq 0}} (\overline{d}_{x,a,b} \vee \overline{d}_{y,a,b}). \qquad (7)$$

Доказательство. Надо показать, что уравнение $F(x)\oplus F(x\oplus a) = b$ имеет два решения, либо не имеет решений. Данное требование записывается при помощи следующей формулы:

$$\bigwedge_{\substack{a,b,x,y\in\mathbb{Z}_2^n,\\ y\neq x,\, y\neq x\oplus a,\, a\neq 0}} (\overline{d}_{x,a,b} \vee \overline{d}_{y,a,b}).$$

Покажем корректность этой формулы «от противного». Предположим, что уравнение $F(x) \oplus F(x \oplus a) = b$ имеет хотя бы четыре решения x', $x' \oplus a$, x'' и $x'' \oplus a$ таких, что $x' \neq x''$ и $x' \neq x'' \oplus a$ для некоторых ненулевых a и b. Тогда дизъюнкция $\overline{d}_{x',a,b} \vee \overline{d}_{x'',a,b}$ равна нулю, а значит, и вся формула равна нулю, что приводит к противоречию.

Поскольку формула, которая ограничивает число решений уравнения $F(x) \oplus F(x \oplus a) = b$, и формула $\mathbf{Der^S}(f,d)$ должны выполняться одновременно, они должны быть записаны через конъюнкцию. Тем самым получаем формулу (7). Теорема 5 доказана.

Таким образом, в формуле $\mathbf{APN^S}(f,d)$ в общей сложности $2^{4n} - 2^{3n}$ дизъюнкций длины 3 и $2^{4n-1} - 2^{3n} - 2^{3n-1} + 2^{2n}$ дизъюнкций длины 2.

3.2. Плотное представление. Для записи дифференциальной равномерности в плотном представлении воспользуемся булевыми переменными $fbq_{x,y,k}$. Действительно, $fbq_{x,x\oplus a,k} = 1 \Leftrightarrow fb_{x,k} \neq fb_{x\oplus a,k}$ или, что то же самое, $fbq_{x,y,k} = 1 \Leftrightarrow fb_{x,k} \oplus fb_{x\oplus a,k} = 1$.

Необходимо показать, что уравнение $F(x) \oplus F(x \oplus a) = b$ имеет одну пару решений, либо не имеет решений. Воспользуемся тем, что для любой пары векторов $\mathbf{fbq}_{x,x\oplus a} = (fbq_{x,x\oplus a,0}, fbq_{x,x\oplus a,1}, \ldots, fbq_{x,x\oplus a,n-1})$ должно быть различие хотя бы в одной координате. Это можно записать следующим образом:

$$\forall\, a \neq 0\ \forall\, x, y\colon y \neq x,\ y \neq x \oplus a\ \exists\, k\colon (fbq_{x,x\oplus a,k} \neq fbq_{y,y\oplus a,k}),$$

где $k = 0, \ldots, n-1$, $x, y, a \in \mathbb{Z}_2^n$.

Запишем данный факт в виде формулы

$$\bigwedge_{\substack{a,x,y\in\mathbb{Z}_2^n,\\ y\neq x,\, y\neq x\oplus a}} \left(\bigvee_{k=0}^{n-1} (fbq_{x,x\oplus a,0} \oplus fbq_{y,y\oplus a,0}) \right).$$

Воспользуемся преобразованием Цейтина [18] для записи данной формулы в КНФ. Введём вспомогательные булевы переменные:

$$dbq_{x,y,a,k} = 1 \Leftrightarrow fbq_{x,x\oplus a,k} \oplus fbq_{y,y\oplus a,k},$$

где $k \in 0,\ldots,n-1,\ x,y,a \in \mathbb{Z}_2^n,\ x \neq y,\ x \neq y \oplus a$. Число таких булевых переменных равно $n \cdot 2^{3n-2}$.

Получим выражение, которое для фиксированных x, y, a, k связывает булевы переменные $dbq_{x,y,a,k}$ и $fbq_{x,x\oplus a,k}$. Для этого в соответствии с определением переменных $dbq_{x,y,a,k}$ построим таблицу истинности.

Таблица 2

Связь переменных $dbq_{x,y,a,k}$ и $fbq_{x,x\oplus a,k}$

$dbq_{x,y,a,k}$	$fbq_{x,x\oplus a,k}$	$fbq_{y,y\oplus a,k}$	Значение
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

Согласно алгоритму построения КНФ по табл. 2 выпишем конъюнкцию дизъюнкций с нулевыми значениями:

$$\mathbf{SoPEq^D}(fbq, dbq) = \bigwedge_{x,y,a,k} (dbq_{x,y,a,k} \vee fbq_{x,x\oplus a,k} \vee \overline{fbq}_{y,y\oplus a,k})$$
$$\wedge (dbq_{x,y,a,k} \vee \overline{fbq}_{x,x\oplus a,k} \vee fbq_{y,y\oplus a,k}) \wedge (\overline{dbq}_{x,y,a,k} \vee fbq_{x,x\oplus a,k} \vee fbq_{y,y\oplus a,k})$$
$$\wedge (\overline{dbq}_{x,y,a,k} \vee \overline{fbq}_{x,x\oplus a,k} \vee \overline{fbq}_{y,y\oplus a,k}),$$

где конъюнкция берётся по всем $x, y, a \in \mathbb{Z}_2^n,\ a \neq 0,\ y \neq x,\ y \neq x \oplus a,\ k = 0,\ldots,n-1$.

Теорема 6. *Булевы переменные $fb_{x,k}$, $fbq_{x,x\oplus a,k}$ и $dbq_{x,y,a,k}$ кодируют* APN*-функцию $F\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ тогда и только тогда, когда истинна следующая формула:*

$$\mathbf{APN^D}(fb, fbq, dbq) = \mathbf{SoPEq^D}(fbq, dbq) \wedge \mathbf{SoP^D}(fb, fbq) \wedge \bigwedge_{\substack{a,x,y\in\mathbb{Z}_2^n,\\ y\neq x,\, y\neq x\oplus a}} \left(\bigvee_{k=0}^{n-1} dbq_{x,y,a,k} \right). \quad (8)$$

Доказательство. Напомним, что у любой пары векторов $fbq_{x,x\oplus a}$ должны быть различия хотя бы в одной координате. Тогда в булевых переменных $dbq_{x,y,a,k}$ это условие представляется в таком виде:

$$\bigwedge_{\substack{x,y,a\in\mathbb{Z}_2^n,\\ y\neq x,\, y\neq x\oplus a,\, a\neq 0}} \left(\bigvee_{k=0}^{n-1} dbq_{x,y,a,k} \right).$$

Поскольку формула, которая ограничивает число решений уравнения $F(x)\oplus F(x\oplus a) = b$, а также формулы $\mathbf{SoPEq^D}(fbq, dbq)$ и $\mathbf{SoP^D}(fb, fbq)$ должны выполняться одновременно, запишем их через конъюнкцию. Таким образом получаем формулу (8). Теорема 6 доказана.

Итого в формуле $\mathbf{APN^D}(fb, fbq, dbq)$ имеются $4n(2^{3n-1} - 2^{2n-1})$ дизъюнкций длины 3 и $2^{3n-1} - 2^{2n} - 2^{2n-1} + 2^n$ дизъюнкций длины n.

4. Метод сдвига

Несмотря на хорошие теоретические оценки и сравнительно быструю работу при малом числе переменных $n \leqslant 4$, описанный выше поиск неизвестных векторных булевых функций с помощью SAT-решателей работает долго для больших $n \geqslant 5$. Для существенного ускорения работы SAT-решателя в поиске криптографических булевых функций можно использовать полностью известные функции с некоторыми свойствами.

Далее обозначения векторов и векторных булевых функций будут выделены полужирным шрифтом. Опишем метод построения новых APN-функций из уже известных векторных функций (так называемый метод сдвига). Пусть $\mathbf{F}\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ — полностью известная APN-функция или дифференциальная 4-равномерная функция. Необходимо подобрать булеву функцию $g\colon \mathbb{Z}_2^n \to \mathbb{Z}_2$, $g \neq 0$, таким образом, чтобы векторная булева функция $\mathbf{G}\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ была APN-функцией:

$$\mathbf{G}(x) = \mathbf{F}(x) \oplus \mathbf{c} \cdot g(x),$$

где $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{Z}_2^n$ — параметр метода сдвига, $\mathbf{c} \cdot g(x) = (c_1 \cdot g(x), \ldots, c_n \cdot g(x))$. Данный метод подробно описан в работе [20]. Отметим,

что в [21] доказано, что если функция $\mathbf{G}$, полученная методом сдвига, является APN-функцией, то $\mathbf{F}$ — APN-функция или дифференциальная 4-равномерная функция.

Для дальнейшего описания воспользуемся обозначением двойной производной по направлению. Напомним, что производная по направлению a равна $D_a f(x) = f(x) \oplus f(x \oplus a)$. Известно, что поиск APN-функций при помощи метода сдвига сводится к системе линейных уравнений.

Утверждение 2 [1]. *Для того чтобы функция* $\mathbf{G}$ *обладала свойством* APN, *необходимо и достаточно, чтобы выполнялось следующее утверждение:*

$$\forall\, a \neq 0 \;\; \forall\, b \neq a \;\; \forall\, x \in \mathbb{Z}_2^n \;\; D_a D_b \mathbf{G}(x) \neq 0,$$

где $D_a D_b \mathbf{G}(x) = \mathbf{G}(x) \oplus \mathbf{G}(x \oplus a) \oplus \mathbf{G}(x \oplus b) \oplus \mathbf{G}(x \oplus a \oplus b)$ — *двойная производная по направлениям* a *и* b.

Проведём замену $\mathbf{G}(x) = \mathbf{F}(x) \oplus \mathbf{c} \cdot g(x)$ и перепишем утверждение выше:

$$D_a D_b \mathbf{F}(x) \oplus \mathbf{c} \cdot D_a D_b g(x) \neq 0.$$

Тогда если $D_a D_b \mathbf{F}(x) = c$, то для сохранения свойства APN требуется $D_a D_b g(x) = 0$. Если $D_a D_b \mathbf{F}(x) = 0$, то $D_a D_b g(x) = 1$. Таким образом, если левая часть равна c или 0, то правая часть равна нулю или единице соответственно, и из данных условий можно получить систему уравнений.

Тем самым проблему поиска функции g можно свести к проблеме линейных алгебраических уравнений над полем $GF(2)$ и решить последнюю с помощью метода Гаусса, поэтому в данном случае в использовании SAT-решателей нет необходимости.

4.1. Метод двойного сдвига. В настоящее время при помощи метода сдвига удалось получить единственную известную на данный момент кубическую APN-функцию от 6 переменных, не эквивалентную ни мономиальным функциям (функциям вида x^d над конечным полем), ни квадратичным функциям. Известно, что получена полная классификация квадратичных APN-функций от 7 переменных [21]. Интерес будут представлять кубические классы APN-функций от 7 переменных, которые не удалось получить с помощью метода сдвига. Для этого предлагается обобщение метода сдвига, основанное на прибавлении двух булевых функций. Назовём его *методом двойного сдвига*. На наш взгляд, он позволяет более гибко перемещаться между различными классами APN-функций.

Идея метода двойного сдвига схожа с классической. Однако теперь необходимо найти $g_1, g_2 \colon \mathbb{Z}_2^n \to \mathbb{Z}_2$, $g_1, g_2 \neq 0$. Векторная булева функция $\mathbf{G} \colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$, которая должна быть APN-функцией, теперь строится

следующим образом:

$$\mathbf{G}(x) = \mathbf{F}(x) \oplus \mathbf{c} \cdot g_1(x) \oplus \mathbf{d} \cdot g_2(x),$$

где $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{Z}_2^n$, $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{Z}_2^n$, $\mathbf{c} \neq \mathbf{d}$ — ненулевые параметры метода двойного сдвига, $\mathbf{c} \cdot g(x) = (c_1 \cdot g(x), \ldots, c_n \cdot g(x))$.

Будем требовать наличие у функций g_1 и g_2 следующих свойств.

• Функция g_1 — кубическая булева функция, т. е. имеет в своём алгебраическом представлении моном третьей степени. Функция g_2 — ненулевая произвольная булева функция. Так, в перспективе можно искать новые неквадратичные и, в частности, кубические классы APN-функций, поскольку для поиска квадратичных классов существует множество других методов.

• Функции g_1 и g_2 не должны быть равны друг другу, иначе метод двойного сдвига сводится к обычному.

• Аффинные части функций g_1 и g_2 равны нулю. Иначе будем находить функции в одном и том же классе EA-эквивалентности, т. е. функция $\mathbf{G}$ представима в виде $\mathbf{G} = A_1 \circ \mathbf{F} \circ A_2 + A$, где A_1 и A_2 — взаимно однозначные аффинные векторные булевы функции над $\mathbb{Z}_2^n$ и A — аффинная функция.

4.2. Запись метода сдвига в виде SAT-задачи. Покажем, что поиск APN-функций с помощью метода двойного сдвига сводится к решению системы нелинейных уравнений. Если воспользоваться утверждением 2 и заменой $\mathbf{G}(x) = \mathbf{F}(x) \oplus \mathbf{c} \cdot g_1(x) \oplus \mathbf{d} \cdot g_2(x)$, то получим выражение

$$D_a D_b \mathbf{F}(x) \oplus \mathbf{c} \cdot D_a D_b g_1(x) \oplus \mathbf{d} \cdot D_a D_b g_2(x) \neq 0.$$

Тогда для сохранения свойств APN-функции рассмотрим следующие четыре случая.

1) Если $D_a D_b \mathbf{F}(x) = \mathbf{c}$, то $(D_a D_b g_1(x) = 1) \to (D_a D_b g_2(x) = 1)$.
2) Если $D_a D_b \mathbf{F}(x) = \mathbf{d}$, то $(D_a D_b g_2(x) = 1) \to (D_a D_b g_1(x) = 1)$.
3) Если $D_a D_b \mathbf{F}(x) = \mathbf{c} \oplus \mathbf{d}$, то $(D_a D_b g_1(x) = 0) \vee (D_a D_b g_2(x) = 0)$.
4) Если $D_a D_b \mathbf{F}(x) = \mathbf{0}$, то $(D_a D_b g_1(x) = 1) \vee (D_a D_b g_2(x) = 1)$.

С помощью эквивалентных преобразований булевых формул можно прийти к следующим уравнениям:

1) $D_a D_b g_1(x) \cdot (D_a D_b g_2(x) \oplus 1) = 0$;
2) $(D_a D_b g_1(x) \oplus 1) \cdot D_a D_b g_2(x) = 0$;
3) $D_a D_b g_1(x) \cdot D_a D_b g_2(x) = 0$;
4) $(D_a D_b g_1(x) \oplus 1) \cdot (D_a D_b g_2(x) \oplus 1) = 0$.

Данные уравнения квадратичные, поэтому для решения системы уравнений уместно использовать SAT-решатели.

5. Результаты вычислительных экспериментов

Задача построения функций g_1 и g_2 решалась для известной функции $\mathbf{F}$, заданной таблицей значений, и для каждого возможного значения параметров $\mathbf{c}$ и $\mathbf{d}$. Для каждого такого случая строилась соответствующая SAT-задача, для решения которой использовался SAT-решатель CryptoMiniSat5. В табл. 3 приведены результаты экспериментов. В столбце SAT указано среднее время работы SAT-решателя для конкретной функции $\mathbf{F}$ и конкретных значений параметров $\mathbf{c}$ и $\mathbf{d}$, когда решение g_1 и g_2 существует, в UNSAT — когда решения не существует. Также приведены данные о размере файлов с КНФ-формулой.

Таблица 3

Результаты построения APN-функций

n	Размер файла	SAT	UNSAT
6	1,5 МБ	0,1 с	1,2 с
7	6 МБ	не сущ. $\mathbf{F}$, $\mathbf{c}$, $\mathbf{d}$	20 с

Метод двойного сдвига был применён ко всем 488 квадратичным APN-функциям от 7 переменных. Затраченное время работы поиска кубической APN-функции составило 20 дней на 72 ядрах суперкомпьютера НГУ. В результате было получено

Утверждение 3. *Пусть* $\mathbf{F}$ *— квадратичная* APN*-функция от 7 переменных и* $\mathbf{G}(x) = \mathbf{F}(x) \oplus \mathbf{c} \cdot g_1(x) \oplus \mathbf{d} \cdot g_2(x)$ *—* APN*-функция от 7 переменных. Тогда* $\mathbf{G}$ *может быть только квадратичной для любых* $\mathbf{c}, \mathbf{d}$.

Заключение

В данной работе представлен набор формул для поиска криптографических функций при помощи SAT-решателей. Описано построение формул для поиска взаимной однозначности и дифференциальной равномерности векторных булевых функций (в случае $\delta = 2$). Данный подход показывает свою эффективность уже на аналитическом этапе исследования, поскольку поиск функций с помощью SAT-решателей существенно быстрее перебора всех $2^{n \cdot 2^n}$ векторных булевых функций от n переменных.

Также описан метод сдвига для поиска векторных булевых функций с заданными свойствами. Реализован метод двойного сдвига, который показал свою эффективность на задаче поиска APN-функции от n переменных в сравнении с прямым поиском функций. С его помощью исследованы квадратичные классы APN-функций от 7 переменных. Показано, что не существует кубических APN-функций от 7 переменных специального вида.

ЛИТЕРАТУРА

1. **Nyberg K.** Differentially uniform mappings for cryptography // Advances in Cryptology — EUROCRYPT'93. Proc. Workshop Theory Appl. Cryptogr. Techniques (Lofthus, Norway, May 23–27, 1993). Heidelberg: Springer, 1994. P. 55–64. (Lect. Notes Comput. Sci.; V. 765).
2. **Тужилин М. Э.** Почти совершенные нелинейные функции // Прикл. дискрет. математика. 2009. № 3. С. 14–20.
3. **Brinkmann M., Leander G.** On the classification of APN functions up to dimension five // Des. Codes Cryptogr. 2008. V. 49. P. 273–288.
4. **Carlet C.** Open questions on nonlinearity and on APN functions // Arithmetic of Finite Fields. Rev. Sel. Pap. 5th Int. Workshop. (Gebze, Turkey, Sept. 27–28, 2014). Cham: Springer, 2015. P. 83–107. (Lect. Notes Comput. Sci.; V. 9061).
5. **Calderini M., Budaghyan L., Carlet C.** On known constructions of APN and AB functions and their relation to each other. San Diego: Univ. California, 2020. (Cryptol. ePrint Archive; Pap. 2020/1444). Available at `eprint.iacr.org/2020/1444` (accessed July 21, 2022).
6. **Yu Y., Wan M., Li Y.** A matrix approach for constructing quadratic APN functions // Des. Codes Cryptogr. 2014. V. 73. P. 587–600.
7. **Beierle C., Leander G.** New instances of quadratic APN functions // IEEE Trans. Inf. Theory. 2022. V. 68. P. 670–678.
8. **Городилова А. А.** Характеризация почти совершенно нелинейных функций через подфункции // Дискрет. математика. 2015. Т. 27, № 3. С. 3–16.
9. **Гэри М., Джонсон Д.** Вычислительные машины и труднорешаемые задачи. М.: Мир, 1982. 420 с.
10. The international SAT competition web page. Paderborn: Satisfiability: Application and Theory, 2022. Available at `www.satcompetition.org`.
11. **Davis M., Logemann G., Loveland D.** A machine program for theorem-proving // Commun. ACM. 1962. V. 5, No. 7. P. 394–397.
12. **Marques Silva J. P., Sakallah K. A.** GRASP — A new search algorithm for satisfiability // Proc. 1996 IEEE/ACM Int. Conf. Computer-Aided Design (San Jose, USA, Nov. 10–14, 1996). Washington: IEEE Comput. Soc., 1996. P. 220–227.
13. **Огородников Ю. Ю.** Комбинированная атака на алгоритм RSA с использованием SAT-подхода // Динамика систем, механизмов и машин. 2016. Т. 2, № 1. С. 276–284.
14. **Schmittner S. E.** A SAT-based public key cryptography scheme. Ithaca, NY: Cornell Univ., 2015. (Cornell Univ. Libr. e-Print Archive; arXiv:1507.08094).
15. **Rivest R. L., Adleman L., Dertouzos M. L.** On data banks and privacy homomorphisms // Foundations of Secure Computation. New York: Academic Press, 1978. P. 169–179.
16. **Wille R., Lye A., Niemann P.** Checking reversibility of Boolean functions // Reversible Computation. Proc. 8th Int. Conf. (Bologna, Italy, July 7–8, 2016). Cham: Springer, 2016. P. 322–337. (Lect. Notes Comput. Sci.; V. 9720).

17. Верещагин Н. К., Шень А. Лекции по математической логике и теории алгоритмов. Часть 1. Начала теории множеств. М.: МЦНМО, 2012. 112 с.
18. Цейтин Г. С. О сложности вывода в исчислении высказываний // Исследования по конструктивной математике и математической логике. II. Зап. науч. сем. ЛОМИ. Т. 8. Л.: Наука, 1968. С. 234–259.
19. Browning K. A., Dillon J. F., McQuistan M. T., Wolfe A. J. An APN permutation in dimension six // Finite Fields: Theory and Applications. Proc. 9th Int. Conf. (Dublin, Ireland, July 13–17, 2009). Providence, RI: AMS, 2010. P. 33–42. (Contemp. Math.; V. 518).
20. Edel Y., Pott A. A new almost perfect nonlinear function which is not quadratic // Adv. Math. Commun. 2015. V. 3. P. 59–81.
21. Kalgin K. V., Idrisova V. A. The classification of quadratic APN functions in 7 variables. San Diego: Univ. California, 2020. (Cryptol. ePrint Archive; Pap. 2020/1515). Available at `eprint.iacr.org/2020/1515` (accessed July 21, 2022).

Доронин Артемий Евгеньевич
Калгин Константин Викторович

Статья поступила
30 декабря 2021 г.
После доработки —
11 апреля 2022 г.
Принята к публикации
15 апреля 2022 г.

DISKRETNYI ANALIZ I ISSLEDOVANIE OPERATSII
/DISCRETE ANALYSIS AND OPERATIONS RESEARCH/
October–December 2022. Vol. 29, No. 4. P. 38–58

UDC 519.7 DOI 10.33048/daio.2022.29.730

APPLICATION OF SAT SOLVERS TO THE PROBLEM OF FINDING VECTOR BOOLEAN FUNCTIONS WITH REQUIRED CRYPTOGRAPHIC PROPERTIES

A. E. Doronin[1, a] and *K. V. Kalgin*[2, 3, b]

[1] Novosibirsk State University,
2 Pirogova Street, 630090 Novosibirsk, Russia

[2] Sobolev Institute of Mathematics,
4 Acad. Koptyug Avenue, 630090 Novosibirsk, Russia

[3] Institute of Computational Mathematics and Mathematical Geophysics,
6 Acad. Lavrentyev Avenue, 630090 Novosibirsk, Russia

E-mail: [a] artem96dor@gmail.com, [b] kalginkv@gmail.com

Abstract. We propose a method for finding an almost perfect nonlinear (APN) function. It is based on translation into SAT-problem and using SAT-solvers. We construct several formulas defining the conditions for finding an APN-function and introduce two representations of the function: Sparse and dense, which are used to describe the problem of finding one-to-one vectorial Boolean functions and APN-functions. We also propose a new method for finding a vectorial APN-function with additional properties. It is based on the idea of representing an unknown vectorial Boolean function as a sum of known APN-functions and two unknown Boolean functions: $\mathbf{G} = \mathbf{F} \oplus \mathbf{c} \cdot g_1 \oplus \mathbf{d} \cdot g_2$, where $\mathbf{F}$ is a known APN-function. It is shown that this method is more efficient than the direct construction of APN-function using SAT for dimensions 6 and 7. As a result, the method described in the work can prove the absence of cubic APN-functions in dimension 7 representable in the form of the sum described above. Tab. 3, bibliogr. 21.

Keywords: SAT-solver, cryptography, Boolean function, APN-function.

This research is carried out within the framework of the state contract of the Sobolev Institute of Mathematics (Project FWNF–2022–0018).

English version: Journal of Applied and Industrial Mathematics **16** (4) (2022).

REFERENCES

1. **K. Nyberg,** Differentially uniform mappings for cryptography, in *Advances in Cryptology — EUROCRYPT'93* (Proc. Workshop Theory Appl. Cryptogr. Techniques, Lofthus, Norway, May 23–27, 1993) (Springer, Heidelberg, 1994), pp. 55–64 (Lect. Notes Comput. Sci., Vol. 765).
2. **M. É. Tuzhilin,** APN functions, *Prikl. Diskretn. Mat.,* No. 3, 14–20 (2009) [Russian].
3. **M. Brinkmann** and **G. Leander,** On the classification of APN functions up to dimension five, *Des. Codes Cryptogr.* **49**, 273–288 (2008).
4. **C. Carlet,** Open questions on nonlinearity and on APN functions, in *Arithmetic of Finite Fields* (Rev. Sel. Pap. 5th Int. Workshop., Gebze, Turkey, Sept. 27–28, 2014) (Springer, Cham, 2015), pp. 83–107 (Lect. Notes Comput. Sci., Vol. 9061).
5. **M. Calderini, L. Budaghyan,** and **C. Carlet,** On known constructions of APN and AB functions and their relation to each other (Univ. California, San Diego, 2020) (Cryptol. ePrint Archive, Pap. 2020/1444). Available at `eprint.iacr.org/2020/1444` (accessed July 21, 2022).
6. **Y. Yu, M. Wan,** and **Y. Li,** A matrix approach for constructing quadratic APN functions, *Des. Codes Cryptogr.* **73**, 587–600 (2014).
7. **C. Beierle** and **G. Leander,** New instances of quadratic APN functions, *IEEE Trans. Inf. Theory* **68**, 670–678 (2022).
8. **A. A. Gorodilova,** Characterization of almost perfect nonlinear functions in terms of subfunctions *Diskretn. Mat.* **27** (3), 3–16 (2015) [Russian] [*Discrete Math. Appl.* **26** (4), 193–202 (2016)].
9. **M. R. Garey** and **D. S. Johnson,** *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, 1979; Mir, Moscow, 1982 [Russian]).
10. The international SAT competition web page (Satisfiability: Application and Theory, Paderborn, 2022). Available at `www.satcompetition.org`.
11. **M. Davis, G. Logemann,** and **D. Loveland,** A machine program for theorem-proving, *Commun. ACM* **5** (7), 394–397 (1962).
12. **J. P. Marques Silva** and **K. A. Sakallah,** GRASP — A new search algorithm for satisfiability, in *Proc. 1996 IEEE/ACM Int. Conf. Computer-Aided Design, San Jose, USA, Nov. 10–14, 1996* (IEEE Comput. Soc., Washington, 1996), pp. 220–227.
13. **Yu. Yu. Ogorodnikov,** A combined attack on the RSA algorithm using the SAT approach, *Din. Sist. Mekh. Mash.* **2** (1), 276–284 (2016) [Russian].
14. **S. E. Schmittner,** A SAT-based public key cryptography scheme (Cornell Univ., Ithaca, NY, 2015) (Cornell Univ. Libr. e-Print Archive; arXiv:1507.08094).
15. **R. L. Rivest, L. Adleman,** and **M. L. Dertouzos,** On data banks and privacy homomorphisms, in *Foundations of Secure Computation* (Academic Press, New York, 1978), pp. 169–179.

16. R. Wille, A. Lye, and **P. Niemann,** Checking reversibility of Boolean functions, in *Reversible Computation* (Proc. 8th Int. Conf., Bologna, Italy, July 7–8, 2016) (Springer, Cham, 2016), pp. 322–337 (Lect. Notes Comput. Sci., Vol. 9720).

17. N. K. Vereshchagin and **A. Shen',** *Lectures on Mathematical Logic and Algorithm Theory. Part 1. The Beginnings of Set Theory* (MTsNMO, Moscow, 2012) [Russian].

18. G. S. Tseitin, On the complexity of proof in prepositional calculus in *Studies in Constructive Mathematics and Mathematical Logic. II* (Zap. Nauchn. Semin. LOMI, Vol. 8) (Nauka, Leningrad, 1968), pp. 234–259 [Russian].

19. K. A. Browning, J. F. Dillon, M. T. McQuistan, and **A. J. Wolfe,** An APN permutation in dimension six, in *Finite Fields: Theory and Applications* (Proc. 9th Int. Conf., Dublin, Ireland, July 13–17, 2009) (AMS, Providence, RI, 2010), pp. 33–42 (Contemp. Math., Vol. 518).

20. Y. Edel and **A. Pott,** A new almost perfect nonlinear function which is not quadratic, *Adv. Math. Commun.* **3**, 59–81 (2015).

21. K. V. Kalgin and **V. A. Idrisova,** The classification of quadratic APN functions in 7 variables (Univ. California, San Diego, 2020) (Cryptol. ePrint Archive; Pap. 2020/1515). Available at `eprint.iacr.org/2020/1515` (accessed July 21, 2022).

Artemy E. Doronin
Konstantin V. Kalgin

Received December 30, 2021
Revised April 11, 2022
Accepted April 15, 2022