МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ПРИКЛАДНАЯ ДИСКРЕТНАЯ МАТЕМАТИКА

Приложение

№ 16 **Сентябрь 2023**

Зарегистрирован в Федеральной службе по надзору
в сфере связи, информационных технологий и массовых коммуникаций
(Роскомнадзор)

Свидетельство о регистрации ПИ № ФС 77-50702 от 17 июля 2012 г.

ТРУДЫ
XXII Международной конференции
«Сибирская научная школа-семинар
“Компьютерная безопасность и криптография” — SIBECRYPT’23»
имени Г. П. Агибалова
(Барнаул, 4–9 сентября 2023 г.)

УЧРЕДИТЕЛЬ
Томский государственный университет

Адрес редакции и издателя: 634050, г. Томск, пр. Ленина, 36
E-mail: pank@mail.tsu.ru

XXII Международная конференция «Сибирская научная школа-семинар "Компьютерная безопасность и криптография" — SIBECRYPT'23» имени Г. П. Агибалова проведена Томским государственным университетом, Новосибирским государственным университетом и Международным математическим центром в Академгородке в сотрудничестве с Академией криптографии РФ с 4 по 9 сентября 2023 г. в г. Барнауле при финансовой поддержке Международного математического центра в Академгородке (соглашение с Министерством науки и высшего образования РФ № 075-15-2022-282) и Северо-Западного центра математических исследований имени Софьи Ковалевской (соглашение с Министерством науки и высшего образования РФ № 075-02-2023-934).

Редактор *Н. И. Шидловская*
Редактор-переводчик *Т. В. Бутузова*
Верстка *И. А. Панкратовой*

Подписано к печати 18.08.2023. Формат $60 \times 84\frac{1}{8}$. Усл. п. л. 18,25. Тираж 300 экз.
Заказ № 5549. Цена свободная. Дата выхода в свет 23.08.2023.

Отпечатано на оборудовании
Издательства Томского государственного университета
634050, г. Томск, пр. Ленина, 36
Тел.: 8(3822)53-15-28, 52-98-49

СОДЕРЖАНИЕ

Секция 1

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПРИКЛАДНОЙ ДИСКРЕТНОЙ МАТЕМАТИКИ

Секция 2

ДИСКРЕТНЫЕ ФУНКЦИИ

Секция 3

МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

Секция 4

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ И ПРОГРАММИРОВАНИЯ

Секция 5

ПРИКЛАДНАЯ ТЕОРИЯ КОДИРОВАНИЯ И АВТОМАТОВ

Секция 1

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПРИКЛАДНОЙ ДИСКРЕТНОЙ МАТЕМАТИКИ

УДК 519.115.4 DOI 10.17223/2226308X/16/1

ОЦЕНКИ ЧИСЛА РАЗБИЕНИЙ ВЕКТОРНОГО ПРОСТРАНСТВА НАД КОНЕЧНЫМ ПОЛЕМ НА АФФИННЫЕ ПОДПРОСТРАНСТВА ОДИНАКОВОЙ РАЗМЕРНОСТИ

И. П. Баксова, Ю. В. Таранников

Получены нижние и верхние оценки на число упорядоченных и неупорядоченных разбиений пространства $\mathbb{F}_q^m$ на аффинные подпространства одинаковой размерности. В частности, установлена асимптотика логарифма числа неупорядоченных разбиений пространства $\mathbb{F}_3^m$ на одномерные аффинные подпространства.

Ключевые слова: *аффинные подпространства, разбиения пространства, оценки, бент-функции.*

Пусть q — степень простого числа. Обозначим через $N_m^k(q)$ число разбиений пространства $\mathbb{F}_q^m$ на q^{m-k} упорядоченных аффинных подпространств размерности k каждое, а через $\widetilde{N}_m^k(q)$ — число разбиений пространства $\mathbb{F}_q^m$ на q^{m-k} неупорядоченных аффинных подпространств размерности k каждое. Очевидно, что величины $N_m^k(q)$ и $\widetilde{N}_m^k(q)$ связаны соотношением

$$N_m^k(q) = q^{m-k}! \cdot \widetilde{N}_m^k(q). \tag{1}$$

В [1, 2] предложены (в разных формулировках, но по сути эквивалентные) конструкции бент-функций от n переменных, основанные на разбиении векторного пространства над $\mathbb{F}_2$ на аффинные подпространства одинаковой размерности, при этом большой вклад в оценку числа бент-функций, порождаемых предложенными конструкциями, вносят оценки величин $N_n^k(2)$. В [1] получены точные формулы для $N_n^{n-1}(q)$ и $N_n^{n-2}(2)$. В [2] приведена (без доказательства) верхняя оценка величины $N_m^k(2)$.

В [3] уточнены оценки на величины $N_m^k(2)$, а в работе [4] установлен вид асимптотики величины $N_m^k(q)$ при $m-k=\text{const}$, $m\to\infty$. В [5] получены совпадающие в главном члене нижняя и верхняя асимптотические оценки на логарифм величины $\widetilde{N}_m^2(2)$:

$$\frac{m}{2}\cdot 2^m + c_1\cdot 2^m + o(2^m) \leqslant \log_2 \widetilde{N}_m^2(2) \leqslant \frac{m}{2}\cdot 2^m + c_2\cdot 2^m + o(2^m),$$

где $c_1 = -1 - \dfrac{3}{4}\log_2 e \approx -2{,}08$; $c_2 = \dfrac{7}{16} - \dfrac{11}{16}\log_2 3 \approx -0{,}65$, а также асимптотика логарифма числа W_m неупорядоченных разбиений пространства $\mathbb{F}_2^m$, m чётно, на линейные подпространства размерности 2:

$$\log_2 W_m = \frac{m}{3}\cdot 2^m + o(m2^m).$$

В настоящей работе рассматриваются оценки на величины $N_m^k(q)$ и $\widetilde{N}_m^k(q)$ для различных q, при этом основное внимание уделяется случаю $(m-k)\to\infty$.

Очевидно, что $N_m^0(q)=q^m!$, так как, разбивая пространство размерности m на аффинные подпространства размерности 0, получаем $q^m!$ упорядоченных разбиений.

Рассмотрим $N_m^1(q)$. При $q=2$ разбиение $\mathbb{F}_q^m$ на аффинные подпространства размерности 1 представляет собой упорядоченное разбиение на пары неупорядоченных наборов, откуда

$$N_m^1(2)=\frac{2^m(2^m-1)}{2}\,\frac{(2^m-2)(2^m-3)}{2}\cdot\ldots=\frac{2^m!}{2^{2^{m-1}}}.$$

При $q>2$ оценки величины $N_m^1(q)$ представляют собой нетривиальную задачу. В настоящей работе мы устанавливаем асимптотику логарифма величины $N_m^1(3)$.

Теорема 1. Число неупорядоченных разбиений $\widetilde{N}_m^1(3)$ пространства $\mathbb{F}_3^m$ на аффинные подпространства размерности 1 удовлетворяет следующему неравенству:

$$\frac{m}{3}\cdot 3^m+c_1\cdot 3^m+o\left(3^m\right)\leqslant\log_3\widetilde{N}_m^1(3)\leqslant\frac{m}{3}\cdot 3^m+c_2\cdot 3^m+o\left(3^m\right),\tag{2}$$

где $c_1=-\dfrac{1}{3}-\dfrac{2}{3}\log_3 e\approx-0{,}94$; $c_2=-\dfrac{1}{3}\log_3 e-\dfrac{1}{3}\log_3 2\approx-0{,}51$.

Доказательство нижней оценки теоремы 1 основано на подходе из [5].

Пусть Q_m — латинский квадрат порядка 3^m, являющийся таблицей Кэли группы $\mathbb{Z}_3^m$; T_m — число трансверсалей в Q_m.

Утверждение 1. $\widetilde{N}_m^1(3)\geqslant T_{m-1}$.

Асимптотика величины T_m получена в [6]:

$$T_m=\left(e^{-1/2}+o(1)\right)\frac{3^m!^2}{3^{m(3^m-1)}}\quad\text{при}\quad m\to\infty.$$

В [5] для получения нижней оценки величины W_m использована асимптотика из [6] для числа трансверсалей в таблице Кэли группы $\mathbb{Z}_2^m$, а для нижней оценки величины $\widetilde{N}_m^2(2)$ — асимптотика из [6] для числа трансверсалей в трёхмерной таблице Кэли группы $\mathbb{Z}_2^m$.

Верхняя оценка теоремы 1 вытекает из теоремы 5 (см. далее).

Теперь переходим к рекуррентным оценкам.

Лемма 1. Имеет место неравенство

$$N_{m+1}^{k+1}(q)\geqslant N_m^k(q).\tag{3}$$

Теорема 2. Имеет место рекуррентное неравенство

$$N_{m+1}^k(q)\geqslant\left(N_m^k(q)\right)^q\binom{q^{m-k+1}}{q^{m-k},q^{m-k},\ldots,q^{m-k}}.$$

Справедливы следующие нижние оценки.

Утверждение 2. Имеют место неравенства

$$N_m^k(q)\geqslant q^{m-k}!,$$
$$\log_q N_m^k(q)\gtrsim(m-k)q^{m-k},\quad m-k\to\infty.$$

Доказательство. Из оценки (3) и соотношения (1) вытекает справедливость цепочки неравенств

$$q^{m-k}! = N_{m-k}^{0}(q) \leqslant N_{m-k+1}^{1}(q) \leqslant N_{m-k+2}^{2}(q) \ldots \leqslant N_{m-1}^{k-1}(q) \leqslant N_{m}^{k}(q).$$

Утверждение доказано. ■

При $q = 3$ оценку утверждения 2 можно усилить с помощью теоремы 1.

Теорема 3. При $k \geqslant 1$, $m - k + 1 \to \infty$ и $q = 3$ имеет место асимптотическое неравенство

$$\log_3 N_m^k(3) \gtrsim 2(m-k)\,3^{m-k}.$$

Доказательство. Из оценки (3) и соотношения (1) вытекает справедливость цепочки неравенств

$$3^{m-k}! \cdot \widetilde{N}_{m-k+1}^{1}(3) = N_{m-k+1}^{1}(3) \leqslant N_{m-k+2}^{2}(3) \ldots \leqslant N_{m}^{k}(3),$$

отсюда, логарифмируя, используя (2) и формулу Стирлинга, получаем

$$\log_3 N_m^k(3) \gtrsim \log_3 3^{m-k}! + \log_3 \widetilde{N}_{m-k+1}^{1}(3) \gtrsim (m-k)3^{m-k} + \frac{m-k+1}{3}\,3^{m-k+1} \sim$$
$$\sim 2(m-k)\,3^{m-k}.$$

Теорема доказана. ■

Заметим, что для $q = 2$ выполнено [3]

$$\log_2 N_m^k(2) \gtrsim 3(m-k)\,2^{m-k}. \tag{4}$$

Следующая нижняя оценка величины $\widetilde{N}_m^k$ является обобщением результата из [3].

Теорема 4. Имеют место неравенства

$$\widetilde{N}_m^k(q) \geqslant \left(q^{k+1}-1\right)^{q^{m-k-1}},$$
$$\log_q \widetilde{N}_m^k(q) \geqslant k \cdot q^{m-k-1}, \quad k \to \infty,$$
$$\log_q N_m^k(q) \gtrsim \left(m - \frac{q-1}{q}\,k\right) q^{m-k}, \quad k \to \infty, \quad m-k \to \infty.$$

Замечание 1. Легко видеть, что асимптотическая оценка теоремы 4 всегда лучше, чем оценка утверждения 2. А вот при $q = 3$ оценка теоремы 3 лучше, если $k < 3m/4$, а при $k > 3m/4$ лучше оценка теоремы 4. Напомним, что при $q = 2$ оценка (4) [3] лучше, если $k < 4m/5$, а если $k > 4m/5$, то лучше оценка теоремы 4.

В заключение приведём верхние оценки величины N_m^k, являющиеся обобщениями результатов из [3].

Теорема 5. Имеет место неравенство

$$N_m^k(q) \leqslant \prod_{i=1}^{q^{m-k}} C_i, \quad \text{где } C_i = \frac{q^m - (i-1)\,q^k}{q^k} \prod_{j=1}^{k} \frac{\left(q^m - (i-1)\,q^k - q^{j-1}\right)}{q^k - q^{j-1}}.$$

Следствие 1. Справедливо неравенство $N_m^k(q) \leqslant (k+1)(m-k)\,q^{m-k}$.

Теорема 6. При $(m-k) \to \infty$ справедливо неравенство

$$\log_q N_m^k(q) \leqslant (k+1)(m-k-\log_q e)q^{m-k} + O(q^{m-k}) + O(k(m-k)).$$

ЛИТЕРАТУРА

1. *Agievich S.* Bent rectangles // NATO Science for Peace and Security Series — D: Information and Communication Security. 2008. V. 18. P. 3–22.
2. *Баксова И. П., Таранников Ю. В.* Об одной конструкции бент-функций // Обозрение прикл. и промышл. матем. 2020. № 27 (1). С. 64–66.
3. *Баксова И. П., Таранников Ю. В.* Оценки числа разбиений пространства $\mathbb{F}_2^m$ на аффинные подпространства размерности k // Вестник Моск. ун-та. Сер. 1. Математика, механика. 2022. № 3. С. 21–25.
4. *Таранников Ю. В.* О существовании разбиений, примитивных по Агиевичу // Дискретный анализ и исследование операций. 2022. Т. 29. № 4. С. 125–144.
5. *Potapov V. N., Taranenko A. A., and Tarannikov Yu. V.* Asymptotic Bounds on Numbers of Bent Functions and Partitions of the Boolean Hypercube into Linear and Affine Subspaces. `https://arxiv.org/abs/2108.00232`. 2021.
6. *Eberhard S.* More on Additive Triples of Bijections. `https://arxiv.org/abs/1704.02407`. 2017.

УДК 519.7 DOI 10.17223/2226308X/16/2

МУЛЬТИПОДСТАНОВКИ И СОВЕРШЕННАЯ РАССЕИВАЕМОСТЬ РАЗБИЕНИЙ

Б. А. Погорелов, М. А. Пудовкина

Концепция мультиподстановочности является одной из первых, позволяющих формализовать «совершенное» рассеивание в алгоритмах блочного шифрования. Для конечной абелевой группы X рассматривается класс преобразований H группы X^2, предложенный С. Ваденау для реализации этой концепции. Каждое биективное преобразование из H является мультиподстановкой. Установлено соответствие между мультиподстановками из H и ортоморфизмами. Рассматриваются разбиения, задаваемые множеством смежных классов $W_0, \dots, W_{r-1}$ по подгруппе $W_0 \leqslant X$, $W = \{W_0, \dots, W_{r-1}\}$. Описаны множества мультиподстановок из H, совершенно рассеивающих разбиения вида W^2 и $X \times W$. Доказана совершенная рассеиваемость таких разбиений 8- и 16-битными преобразованиями алгоритма блочного шифрования CS.

Ключевые слова: *мультиподстановка, ортоморфизм, квазиадамарово преобразование, совершенное рассеивание, алгоритм блочного шифрования CS.*

Традиционно при синтезе алгоритмов блочного шифрования следуют неформально сформулированным К. Шенноном принципам рассеивания и перемешивания [1]. Под рассеивающим преобразованием К. Шеннон понимает некоторое отображение векторного пространства на себя, при котором каждая или почти каждая его «компактная» область в образе распределяется в большую, «некомпактную» с точки зрения метрики, область. Концепция мультиподстановочности является одной из первых, позволяющих формализовать «совершенное» рассеивание в алгоритмах блочного шифрования, и относится к координатному рассеиванию. Реализующие данную концепцию отображения введены в [2, 3] и названы мультиподстановками, а также совершенно рассеивающими (perfect diffusion).

В [3] (n, d)-мультиподстановкой над алфавитом X называется такое отображение $h\colon X^n \to X^d$, что два различных вектора $(\alpha, h(\alpha)), (\beta, h(\beta)) \in X^n \times X^d$ не могут совпадать ни в каких n координатах. Мультиподстановочность тесно связана с МДР-кодами,

латинскими квадратами, квазигруппами и ортогональными таблицами. Так, существует однозначное соответствие между (n,d)-мультиподстановками и $(|X|^n, n+d, |X|, n)$-ортогональными таблицами [3].

Определение (2,2)-мультиподстановки [2] равносильно следующему. Отображение $h\colon X^2 \to X^2$ называется (2,2)-мультиподстановкой, если частичные функции

$$h_\alpha^{(i,1)}(x) = h^{(i)}(\alpha, x),\ h_\alpha^{(i,2)}(x) = h^{(i)}(x, \alpha)$$

есть подстановки на X для каждых $\alpha \in X$, $i \in \{1,2\}$. Отметим, что каждая (2,2)-мультиподстановка есть подстановка на X^2. Однако не каждая подстановка является (2,2)-мультиподстановкой. В данной работе будем под мультиподстановкой понимать (2,2)-мультиподстановку.

В настоящее время на X^2 (кроме мультиподстановок) для улучшения рассеивающих свойств также рассматриваются псевдоадамаровы [4], квазиадамаровы преобразования [5] и их обобщения [6, 7]. Так, в [8] найден коэффициент (координатного) рассеивания квазиадамаровых преобразований, описаны их свойства относительно линейного и разностного методов.

Пусть $(X,+)$ — конечная абелева группа с бинарной операцией $+$, $S(X)$ — симметрическая группа на X, $\alpha^b = b(\alpha)$ для каждых $\alpha \in X$, $b\colon X \to X$, $\mathrm{Aut}(X)$ — группа автоморфизмов на X, $V_n(2)$ — n-мерное векторное пространство над полем $\mathrm{GF}(2)$, $W_{w,r}(P)$ — множество всех разбиений множества P с r блоками мощности w каждый, $|P| = w \cdot r$. Отметим, что в алгоритмах блочного шифрования чаще всего X — аддитивная группа кольца вычетов или векторного пространства $V_n(2)$, например, $X = \mathbb{Z}_{2^8}$ в алгоритме Safer [4].

В данной работе для произвольной конечной абелевой группы $(X,+)$ рассматривается класс отображений вида $h_{\bar\lambda,\bar\psi,\bar\nu}\colon X^2 \to X^2$,

$$h_{\bar\lambda,\bar\psi,\bar\nu}\colon (\alpha_1,\alpha_2) \mapsto \left((\alpha_1^{\lambda_1} + \alpha_2^{\psi_1})^{\nu_1}, (\alpha_1^{\lambda_2} + \alpha_2^{\psi_2})^{\nu_2}\right),$$

задаваемый наборами $\bar\lambda, \bar\psi, \bar\nu \in S(X)^2$, где $\bar\lambda = (\lambda_1, \lambda_2)$, $\bar\psi = (\psi_1, \psi_2)$, $\bar\nu = (\nu_1, \nu_2)$, предложенный в [3]. Из условия биективности отображения $h_{\bar\lambda,\bar\psi,\bar\nu}$ вытекает его мультиподстановочность.

Подстановка $b \in S(X)$ является ортоморфизмом, если преобразование $x \mapsto x^b - x$ есть подстановка на X [9].

Лемма 1. Пусть $(X,+)$ — произвольная абелева группа, $\bar\lambda, \bar\psi, \bar\nu \in S(X)^2$ и $\lambda_1^{-1}\lambda_2 \in \mathrm{Aut}(X)$. Тогда для мультиподстановочности $h_{\bar\lambda,\bar\psi,\bar\nu}$ на X достаточно ортоморфности $\lambda_2^{-1}\lambda_1\psi_1^{-1}\psi_2$.

Хорошо известно [10], что если силовская 2-подгруппа группы X является нетривиальной и циклической, то не существует ортоморфизмов (полных преобразований) на X, а поэтому не существует мультиподстановок типа $h_{\bar\lambda,\bar\psi,\bar\nu}$. Так как аддитивная группа кольца $\mathbb{Z}_{2^m}$ имеет нетривиальную циклическую силовскую 2-подгруппу, то не существует мультиподстановок типа $h_{\bar\lambda,\bar\psi,\bar\nu}$ на $\mathbb{Z}_{2^m}$, но они существуют на $\mathbb{Z}_{2^m}^2$.

В [11] для характеризации рассеивания подстановкой $g \in S(X)$ разбиения $W = \{W_0, \ldots, W_{r-1}\}$ множества X введена матрица $\mathbf{c}^{(W)}(g) = \left\| c_{i,j}^{(W)}(g) \right\|$, где $c_{i,j}^{(W)}(g) = |W_i^s \cap W_j|$ для $i,j \in \{0,\ldots,r-1\}$, $w = |W_0| = \ldots = |W_{r-1}|$. С помощью элементов матрицы также оценивается расстояние (Хемминга) от подстановки g до группы IG_W,

состоящей из всех подстановок из $S(X)$, которые сохраняют разбиение W. Выделяются подстановки, максимально далёкие (относительно расстояния Хемминга) от группы IG_W. Они имеют наилучшие рассеивающие свойства разбиения W и названы совершенно рассеивающими разбиение W. В [11] показано, что для каждой подстановки $g \in S(X)$, совершенно рассеивающей разбиение W, элементы $c_{i,j}^{(W)}(g)$ матрицы $\mathbf{c}^{(W)}(g)$ удовлетворяют неравенству

$$c_{i,j}^{(W)}(g) \leqslant \lceil w \cdot r^{-1} \rceil$$

для всех $i, j \in \{0, \ldots, r-1\}$.

В настоящей работе получены условия на $\bar{\lambda}, \bar{\psi}, \bar{\nu} \in S(X)^2$, при выполнении которых мультиподстановка $h_{\bar{\lambda},\bar{\psi},\bar{\nu}}$ совершенно рассеивает разбиение

$$W \times W = \{W_i \times W_j : i, j \in \{0, \ldots, r-1\}\}.$$

Утверждение 1. Пусть $w, r \in \mathbb{N}$, $|X| = w \cdot r$, $\{0\} \subset W_0 \subset X$, $w = |W_0|$, $W \in W_{w,r}(X)$, ν_1, ν_2 совершенно рассеивают разбиение W, $\bar{\lambda}, \bar{\psi} \in \mathrm{IG}_W \times \mathrm{IG}_W$. Тогда мультиподстановка $h_{\bar{\lambda},\bar{\psi},\bar{\nu}}$ на X совершенно рассеивает разбиение $W \times W$, причём

$$c_{(i_1,i_2),(t_1,t_2)}^{(W\times W)}\left(h_{\bar{\lambda},\bar{\psi},\bar{\nu}}\right) \leqslant \begin{cases} 1, & \text{если } w \leqslant r, \\ \lceil w^2/r^2 \rceil, & \text{если } w > r. \end{cases}$$

Пусть 0_n — n-мерный вектор с нулевыми координатами.

Утверждение 2. Пусть $d \in \mathbb{N}$, $n = 2d$, $W_0 = V_d(2) \times \{0_d\}$, W — фактор-пространство пространства $V_n(2)$ по подпространству W_0, $w = |W_0|$, $r = |W|$. Тогда существует класс мультиподстановок $h_{\bar{\lambda},\bar{\psi},\bar{\nu}}$ на $V_n(2)$, сохраняющих разбиение $W \times W$, но с преобразованиями $\nu_i, \lambda_i, \psi_i \in S(V_n(2))$, совершенно рассеивающими разбиение W при $i = 1, 2$.

Пусть преобразования $g_v^{(1)}, g_v^{(2)}$ на $V_n(2)^2$ с функцией усложнения $v\colon V_n(2) \to V_n(2)$ заданы соответственно для всех $(\alpha_1, \alpha_2) \in V_n(2)^2$ условиями

$$g_v^{(1)}\colon (\alpha_1, \alpha_2) \mapsto (\alpha_1 + \alpha_2^v, \alpha_2),\ g_v^{(2)}\colon (\alpha_1, \alpha_2) \mapsto (\alpha_1, \alpha_2 + \alpha_1^v).$$

Заметим, что $g_v^{(1)}$, $g_v^{(2)}$ — Фейстель-подобные инволютивные преобразования.

Для $b\colon V_n(2) \to V_n(2)$, $s \in S(V_n(2))$ рассмотрим две подстановки $u_{b,s}^{(1)} = g_b^{(1)} g_s^{(2)}$, $u_{b,s}^{(2)} = g_b^{(1)} g_s^{(2)} g_b^{(1)}$ на $V_n(2)^2$. Заметим также, что 8-битная подстановка $u_{b,s}^{(2)}$ $(n = 8)$ использовалась в алгоритме блочного шифрования CS [12].

Утверждение 3. Пусть U_0 — подпространство пространства $V_n(2)$, $W_0 = V_n(2) \times$ $\times U_0$, $U = \{U_0, \ldots, U_{d-1}\}$ — множество всех смежных классов аддитивной группы $(V_n(2), +)$ по подгруппе U_0, $W^{(U)} = V_n(2) \times U$. Тогда:

1) для каждого преобразования $b : V_n(2) \to V_n(2)$ и каждой подстановки $s \in$ $\in S(V_n(2))$ преобразования $u_{b,s}^{(1)}, u_{b,s}^{(2)}$ совершенно рассеивают разбиение $W^{(U)}$;
2) для каждых таких подстановок $\nu_1, \nu_2, \lambda_1, \lambda_2 \in S(V_n(2))$, что $\lambda_1 \lambda_2^{-1}$ — ортоморфизм на $V_n(2)$, мультиподстановка

$$h_{(\lambda_1,\lambda_2),(e,e),(\nu_1,\nu_2)} : (\alpha_1, \alpha_2) \mapsto \left((\alpha_1^{\lambda_1} + \alpha_2)^{\nu_1}, (\alpha_1^{\lambda_2} + \alpha_2)^{\nu_2}\right)$$

совершенно рассеивает разбиение $W^{(U)}$.

Пусть $n \equiv 0 \pmod 4$, подстановки $\lambda_1, \lambda_2 \in S(V_n(2))$ заданы условиями

$$\lambda_1 \colon (\alpha_1, \alpha_2, \ldots, \alpha_n) \mapsto (\alpha_1, \alpha_2 + \alpha_3, \alpha_3, \alpha_4 + \alpha_5, \alpha_5, \ldots, \alpha_{n-2} + \alpha_{n-1}, \alpha_{n-1}, \alpha_n + \alpha_1),$$
$$\lambda_2 \colon (\alpha_1, \alpha_2, \ldots, \alpha_n) \mapsto (\alpha_2, \ldots, \alpha_n, \alpha_1),$$

где λ_2 — преобразование левого циклического сдвига. Несложно убедиться, что $\lambda_1\lambda_2^{-1}$ — ортоморфизм на $V_n(2)$ и $\lambda_1\lambda_2^{-1} \in \mathrm{Aut}(V_n(2))$.

При $n = 8$ и $\nu_1 = \nu_2 = u_{b,s}^{(2)}$ в раундовой функции $f \colon V_{64}(2)^2 \to V_{64}(2)$ алгоритма CS [12] компонентами являются четыре 16-битные мультиподстановки $h_{(\lambda_1,\lambda_2),(e,e),(u_{b,s}^{(2)},u_{b,s}^{(2)})}$, обеспечивающие реализацию свойств перемешивания и рассеивания К. Шеннона. Из утверждения 3 следует, что мультиподстановка $h_{(\lambda_1,\lambda_2),(e,e),(u_{b,s}^{(2)},u_{b,s}^{(2)})}$ алгоритма CS совершенно рассеивает разбиения вида $W^{(U)} = V_n(2) \times U$.

ЛИТЕРАТУРА

1. Словарь криптографических терминов / под ред. Б. А. Погорелова и В. Н. Сачкова. М.: МЦНМО, 2006. 94 с.
2. *Schnorr C.-P. and Vaudenay S.* Black box cryptanalysis of hash networks based on multipermutations // LNCS. 1995. V. 950. P. 47–57.
3. *Vaudenay S.* On the need for multipermutations: cryptanalysis of MD4 and SAFER // LNCS. 1995. V. 1008. P. 286–297.
4. *Massey J. L.* SAFER K-64: a byte-oriented block-ciphering algorithm // LNCS. 1994. V. 1267. P. 1–17.
5. *Lipmaa H.* On differential properties of pseudo-Hadamard transform and related mappings // LNCS. 2002. V. 2551. P. 48–61.
6. *Погорелов Б. А., Пудовкина М. А.* Вариации ортоморфизмов и псевдоадамаровых преобразований на неабелевой группе // Прикладная дискретная математика. Приложение. 2019. Т. 12. С. 24–27.
7. *Погорелов Б. А., Пудовкина М. А.* Обобщенные квазиадамаровы преобразования на конечных группах // Матем. вопр. криптографии. 2022. Т. 13. № 4. С. 97–124.
8. *St Denis T.* Fast Pseudo-Hadamard Transforms. Cryptology ePrint Archive. Report 2004/010. 2004. `https://eprint.iacr.org/2004/010.pdf`.
9. *Evans A. B.* Applications of complete mappings and orthomorphisms of finite groups // Quasigroups and Related Systems. 2015. V. 23. P. 5–30.
10. *Hall M. and Paige L. J.* Complete mappings of finite groups // Pacific J. Math. 1955. V. 5. P. 541–549.
11. *Погорелов Б. А., Пудовкина М. А.* О расстояниях от подстановок до импримитивных групп при фиксированной системе импримитивности // Дискретная математика. 2013. Т. 25. № 3. С. 78–95.
12. *Stern J. and Vaudenay S.* CS-Cipher // LNCS. 1998. V. 1372. P. 189–204.

Секция 2

ДИСКРЕТНЫЕ ФУНКЦИИ

УДК 511.32 DOI 10.17223/2226308X/16/3

СВОЙСТВА КЛАССОВ БУЛЕВЫХ ФУНКЦИЙ, ПОСТРОЕННЫХ ИЗ НЕСКОЛЬКИХ ЛИНЕЙНЫХ РЕКУРРЕНТ НАД КОЛЬЦОМ ВЫЧЕТОВ $\mathbb{Z}_{2^n}$

А. Д. Бугров

Определён класс булевых функций, построенных из старших разрядных последовательностей линейных рекуррент над кольцом $\mathbb{Z}_{2^n}$. Для выделения старших разрядных последовательностей используются различные координатные множества. Показано, что указанный класс состоит из функций, значительно удалённых от класса всех афинных функций.

Ключевые слова: *линейные рекуррентные последовательности, разрядные последовательности, булевы функции, нелинейность булевых функций.*

Введение

Пусть n — натуральное число, $R = \mathbb{Z}_{2^n} = \{0, 1, \ldots, 2^n - 1\}$ — кольцо вычетов по модулю 2^n, $P = \mathbb{Z}_2 = \{0, 1\}$ — поле из двух элементов. Операцию сложения в кольце R обозначим $+$, а операцию сложения в поле P через $\oplus$. Пусть $F(x)$ — унитарный неприводимый многочлен степени m над полем P. Всюду в дальнейшем будем считать, что $F(x) \neq x$, т. е. $F(x)$ является реверсивным неприводимым многочленом над полем P. Период $T(F)$ такого многочлена равен

$$T(F) = (2^m - 1)/d,$$

где d — некоторый делитель числа $2^m - 1$.

Согласно [1], найдётся единственный унитарный многочлен $G(x) \in R[x]$, такой, что $\bar{G}(x) = F(x)$ и $T(G) = T(\bar{G})$, где $\bar{G}(x)$ — многочлен над полем P, полученный из $G(x)$ приведением всех его коэффициентов по модулю 2. Такой многочлен $G(x)$ называется отмеченным многочленом над кольцом R. Кроме того, как показано в [1], существует простой рекурсивный способ построения многочлена $G(x)$ по многочлену $F(x)$.

Обозначим через $L_R(G)$ множество всех линейных рекуррентных последовательностей (ЛРП) v над кольцом R с характеристическим многочленом $G(x)$, а через $L_R(G)^*$ — его подмножество, состоящее из всех ЛРП, содержащих в своём начальном векторе $(v(0), \ldots, v(m-1))$ хотя бы один обратимый элемент кольца R.

Зададим на множестве $L_P(F)$ всех ЛРП над полем P с характеристическим многочленом $F(x)$ бинарное отношение $\sim$, положив для последовательностей u, $u' \in L_P(F)$ $u \sim u'$ тогда и только тогда, когда найдётся $t \in \mathbb{N}$, такое, что $u' = x^t u$, т. е. $u'(i) = u(i+t)$ при всех $i \geqslant 0$. В силу реверсивности многочлена $F(x)$ это отношение является эквивалентностью на множестве $L_P(F)$.

Выберем и зафиксируем ненулевые попарно неэквивалентные ЛРП:

$$u_0, \ldots, u_{d-1} \in L_P(F).$$

В силу того, что $T(u_0) = \ldots = T(u_{d-1}) = T(F)$ и последовательности выбраны попарно неэквивалентными, векторы

$$(u_0(i), \ldots, u_0(i+m-1)),\ \ldots,\ (u_{d-1}(i), \ldots, u_{d-1}(i+m-1)),\ i = 0, 1, \ldots, T(F)-1,$$

пробегают множество всех ненулевых векторов из P^m.

Для каждого элемента $a \in R$ обозначим через $\bar{a}$ его остаток при делении на 2. Разрядным множеством кольца R называется подмножество $K = \{k_0, k_1\} \subset R$, такое, что $\bar{k}_0 = 0$ и $\bar{k}_1 = 1$. Каждый элемент $a \in R$ однозначно представим в виде

$$a = a_0 + 2a_1 + 2^2 a_2 + \ldots + 2^{n-1} a_{n-1},$$

где $a_i \in K$, $i = 0, 1, \ldots, n-1$. Рассмотрим отображение $\varkappa_{n-1}^K : R \to P$, действующее по правилу $\varkappa_{n-1}^K(a) = \bar{a}_{n-1}$. Зафиксируем разрядные множества K_0, …, K_{d-1} кольца R.

Выберем произвольные ЛРП v_0, …, $v_{d-1} \in L_R(G)^*$ и построим булеву функцию $f : P^m \to P$ от m переменных по следующим правилам:

$$f(0, \ldots, 0) = 0$$

и для всех $i = 0, 1, \ldots, T(F) - 1$

$$\begin{gathered} f(u_0(i), \ldots, u_0(i+m-1)) = \varkappa_{n-1}^{K_0}(v_0(i)), \\ f(u_1(i), \ldots, u_1(i+m-1)) = \varkappa_{n-1}^{K_1}(v_1(i)), \\ \vdots \\ f(u_{d-1}(i), \ldots, u_{d-1}(i+m-1)) = \varkappa_{n-1}^{K_{d-1}}(v_{d-1}(i)). \end{gathered}$$

ЛРП u_0, …, u_{d-1}, а также разрядные множества K_0, …, K_{d-1} будем считать фиксированными. В связи с этим будем использовать обозначение

$$f(\mathbf{x}) = f_{v_0, \ldots, v_{d-1}}(\mathbf{x})$$

и изучать класс булевых функций

$$D_m(F) = \{f_{v_0, \ldots, v_{d-1}}(\mathbf{x}) : v_0, \ldots, v_{d-1} \in L_R(G)^*\}.$$

Для случая $d = 1$ аналогичные функции исследовались в работах [2–5]. Доказано, что данные функции достаточно удалены от класса всех аффинных булевых функций от m переменных. Представляет интерес исследование случая, когда d — произвольный делитель числа $2^m - 1$.

Данная работа продолжает аналогичные исследования случая $d > 1$, начатые в работе [6], где рассмотрена ситуация, когда $K_0 = \ldots = K_{d-1} = \{0, 1\}$. Получена оценка нелинейности для функций из класса $D_m(F)$.

Приведем некоторые вспомогательные результаты. Пусть $\chi : R \to \mathbb{C}^*$ — аддитивный характер кольца R, определённый равенством

$$\chi(x) = e^{2\pi i x / 2^n}, \quad x \in R.$$

Для каждого $s = 0, 1, \ldots, d-1$ рассмотрим отображение $\mu_s : R \to \mathbb{C}^*$, действующее по правилу

$$\mu_s(a) = (-1)^{\varkappa_{n-1}^{K_s}(a)}, \quad a \in R.$$

Для каждой последовательности v над кольцом R обозначим через $\bar{v}$ последовательность над полем P, полученную заменой каждого элемента $v(i)$, $i \geqslant 0$, на его остаток при делении на 2.

Нелинейность функций

Нелинейность $\mathrm{nl}(f)$ функций f из класса $D_m(F)$ определяется как расстояние Хэмминга от функции f до класса всех аффинных булевых функций от m переменных. Для нахождения нелинейности удобно воспользоваться формулой

$$\mathrm{nl}(f) = 2^{m-1} - \frac{1}{2}\max_{\mathbf{a}\in P^m}|W_f(\mathbf{a})|,$$

где $W_f(\mathbf{a})$ — коэффициент Уолша — Адамара функции f, определяемый для каждого вектора $\mathbf{a} = (a_1, \ldots, a_m) \in P^m$ по правилу

$$W_f(\mathbf{a}) = \sum_{\mathbf{b}=(b_1,\ldots,b_m)\in P^m} (-1)^{f(\mathbf{b})\oplus a_1b_1\oplus\ldots\oplus a_mb_m}.$$

Теорема 1. Пусть $R = \mathbb{Z}_{2^n}$, $n \geqslant 2$, $G(x) \in R[x]$ — отмеченный многочлен периода $T(G) = (2^m-1)/d$, $\bar{G}(x) = F(x)$, $\mu_s(0) = 1$ для всех $s = 0, 1, \ldots, d-1$. Тогда для каждой функции $f \in D_m(F)$ выполнено неравенство

$$\mathrm{nl}(f) \geqslant 2^{m-1} - \left(\frac{2(n-1)}{\pi}\ln 2 + 1\right)(d2^{n-1} - 1)2^{m/2-1}.$$

ЛИТЕРАТУРА

1. *Нечаев А. А.* Цикловые типы линейных подстановок над конечными коммутативными кольцами // Матем. сборник. 1993. Т. 184. № 1. С. 21–56.
2. *Былков Д. Н., Камловский О. В.* Параметры булевых функций, построенных с использованием старших координатных последовательностей линейных рекуррент // Матем. вопр. криптографии. 2012. Т. 3. № 4. С. 25–53.
3. *Былков Д. Н.* Об одном классе булевых функций, построенных с использованием старших разрядных последовательностей линейных рекуррент // Прикладная дискретная математика. Приложение. 2014. № 7. С. 59–60.
4. *Камловский О. В.* Нелинейность одного класса булевых функций, построенных с использованием двоичных разрядных последовательностей линейных рекуррент над кольцом $\mathbb{Z}_{2^n}$ // Матем. вопр. криптографии. 2016. Т. 7. № 3. С. 29–46.
5. *Бугров А. Д., Камловский О. В.* Параметры одного класса функций, заданных на конечном поле // Матем. вопр. криптографии. 2018. Т. 9. № 4. С. 31–52.
6. *Груба А. А.* Булевы функции, построенные с использованием разрядных последовательностей линейных рекуррент // Дискретная математика. 2023. Т. 35. № 1. С. 54–61.

УДК 519.7 DOI 10.17223/2226308X/16/4

О ДОСТИЖИМОСТИ НИЖНЕЙ ОЦЕНКИ ЧИСЛА БЕНТ-ФУНКЦИЙ НА МИНИМАЛЬНОМ РАССТОЯНИИ ОТ БЕНТ-ФУНКЦИИ ИЗ КЛАССА МЭЙОРАНА — МАКФАРЛАНДА[1]

Д. А. Быков

Исследуется нижняя оценка $2^{2n+1} - 2^n$ числа бент-функций на минимально возможном расстоянии 2^n от некоторой исходной бент-функции из класса Мэйорана — МакФарланда $\mathcal{M}_{2n}$ от $2n$ переменных. Сформулирован критерий её достижимости для функций в алгебраическом представлении. Конструктивно доказано, что в случае $n = p^k$ для простого $p \neq 2, 3$ и натурального k оценка точна.

[1]Работа выполнена при поддержке Математического центра в Академгородке, соглашение с Министерством науки и высшего образования Российской Федерации № 075–15–2022–282.

Показано, что необходимым условием достижимости оценки является построение функции из $\mathcal{M}_{2n}$ по APN-перестановке, множество значений которой на любом аффинном подпространстве размерности 3 не является аффинным подпространством.

Ключевые слова: *бент-функция, булева функция, минимальное расстояние, класс Мэйорана — МакФарланда, нижняя оценка.*

Введение

Пусть $\mathbb{F}_{2^n}$ — конечное поле из 2^n элементов. Далее функции рассматриваются над полями. Отметим, что для перехода к более привычному представлению над векторным пространством $\mathbb{F}_2^n$ достаточно зафиксировать в поле $\mathbb{F}_{2^n}$ некоторый базис над $\mathbb{F}_2$. Функция $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$ называется *булевой функцией* от n переменных. *Расстоянием Хэмминга* между двумя булевыми функциями f, g называется число аргументов, на которых их значения различаются. *Аффинной* называется булева функция, степень которой не превосходит 1. Булева функция от чётного числа переменных $2n$ называется *бент-функцией*, если расстояние от неё до ближайшей аффинной функции максимально и равно $2^{2n-1} - 2^{n-1}$. В дальнейшем будем рассматривать бент-функции от $2n$ переменных. Пусть $m|n$, функция $\mathrm{tr}_m^n : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$, определённая как $\mathrm{tr}_m^n(x) = x^{2^0} + x^{2^m} + \cdots + x^{2^{n-m}}$, называется *частичным следом* из поля $\mathbb{F}_{2^n}$ в его подполе $\mathbb{F}_{2^m}$. Класс бент-функций Мэйорана — МакФарланда $\mathcal{M}_{2n}$ состоит из функций вида

$$\mathrm{tr}_1^n(x\pi(y)) + \varphi(y),$$

где $\pi : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ взаимно однозначна и $\varphi : \mathbb{F}_{2^n} \to \mathbb{F}_2$ — произвольная булева функция.

Минимальные функции изучались в 1960-х гг. в СССР В. А. Елисеевым и О. П. Степченковым [1], а термин «бент-функции» для них введён О. Ротхаусом в [2]. Отметим, что Елисеевым исследовался и класс Мэйорана — МакФарланда. Бент-функции интересны как своими экстремальными значениями нелинейности, так и приложениями в криптографии, теории кодирования, теории символьных последовательностей. В работе исследуются ближайшие бент-функции к функциям из класса Мэйорана — МакФарланда. Известно, что расстояние между двумя различными бент-функциями от $2n$ переменных не менее 2^n, функции на данном расстоянии изучались в [3–7]. В [6] получен явный вид всех функций из $\mathcal{M}_{2n}$ на минимальном расстоянии от исходной функции из $\mathcal{M}_{2n}$ и дана следующая нижняя оценка числа всех бент-функций на этом расстоянии:

Утверждение 1 (Н. А. Коломеец, 2017). Имеется не менее $2^{2n+1} - 2^n$ бент-функций на минимальном расстоянии от функции из $\mathcal{M}_{2n}$. Все учтённые здесь бент-функции также лежат в классе $\mathcal{M}_{2n}$.

В [7] уточнён критерий принадлежности функции на расстоянии 2^n от функции из $\mathcal{M}_{2n}$ к классу бент-функций и доказано, что нижняя оценка $2^{2n+1} - 2^n$ достижима в случае простого $n \geqslant 5$. Необходимым условием достижимости оценки для функции $\mathrm{tr}_1^n(x\pi(y)) + \varphi(y) \in \mathcal{M}_{2n}$ является использование в качестве π взаимно однозначной APN-функции.

В настоящей работе показано, что оценка достигается и в случае $n = p^k$ для простого p и натурального k, а также то, что для её достижения необходимо, чтобы перестановка в бент-функции из $\mathcal{M}_{2n}$ была APN-функцией, множество значений которой на любом аффинном подпространстве размерности 3 не является аффинным подпространством.

1. Достижимость оценки в случае $n = p^k$

Введём необходимые определения. Любая булева функция f от n переменных представима в виде $f(x) = \mathrm{tr}_1^n\left(\sum\limits_{k=0}^{2^n-1} c_k x^k\right)$, где $c_k \in \mathbb{F}_{2^n}$. Такое представление не единственно, однако его можно сделать таким [8]. Функция $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, определённая как $F(x) = \sum\limits_{k=0}^{n-1} \alpha_k x^{2^k} + \alpha_n$, где $\alpha_0, \ldots, \alpha_n$, называется *аффинной*. Под *линейным подпространством поля* $\mathbb{F}_{2^n}$ будем иметь в виду аддитивные подгруппы $\mathbb{F}_{2^n}$. Соответственно *аффинным подпространством* U *поля* $\mathbb{F}_{2^n}$ назовём $U = a + L$, где $a \in \mathbb{F}_{2^n}$ и L — линейное подпространство $\mathbb{F}_{2^n}$. Функция называется *аффинной на подпространстве*, если её сужение на подпространство U совпадает с сужением некоторой аффинной функции на U. Определим множества $F(S) = \{F(x) : x \in S\}$ и $qS = \{qy : y \in S\}$ для любых $S \subseteq \mathbb{F}_{2^n}$, $q \in \mathbb{F}_{2^n}$ и $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$. Отметим, что в поле $\mathbb{F}_{2^n}$ функция обращения элемента $F(x) = x^{2^n-2}$ является взаимно однозначной.

Основываясь на критерии из работы [7], сформулируем критерий достижимости нижней оценки.

Утверждение 2. На расстоянии 2^n от бент-функции $f(x, y) = \mathrm{tr}_1^n(x\pi(y)) + \varphi(y)$ из $\mathcal{M}_{2n}$ лежит в точности $2^{2n+1} - 2^n$ бент-функций тогда и только тогда, когда для любого аффинного подпространства $E \subseteq \mathbb{F}_{2^n}$ размерности k, $2 \leqslant k \leqslant n$, выполнено одно из следующих условий:

1) $\pi(E)$ не является аффинным подпространством $\mathbb{F}_{2^n}$;
2) $\mathrm{tr}_1^n(H(y)\pi(y)) + \varphi(y)$ не аффинна на E для любой аффинной функции $H : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$.

Теперь можем обобщить достаточные условия достижимости нижней оценки для функции из $\mathcal{M}_{2n}$ от случая простого $n \geqslant 5$ до случая $n = p^k$ для простого $p \neq 2, 3$ и натурального k. При этом, как и прежде, в качестве перестановки используется функция обращения элементов в поле.

Теорема 1. Пусть $n = p^k$, где $p \neq 2, 3$ — простое, k — натуральное и функция $\delta : \mathbb{F}_{2^n}^* \to \mathbb{F}_{2^p}^*$ такая, что $\delta(x) = \delta(y)$ для любых $x, y \in \mathbb{F}_{2^n}^*$, удовлетворяющих $xy^{-1} \in \mathbb{F}_{2^p}$. Пусть также $f(x, y) = \mathrm{tr}_1^n\left(xy^{2^n-2}\right) + \varphi(y) \in \mathcal{M}_{2n}$ такая, что $\varphi(qz) = \varphi'(\delta(q)z)$, где $z \in \mathbb{F}_{2^p}$, $q \in \mathbb{F}_{2^n}^*$ и булева функция $\varphi' : \mathbb{F}_{2^p} \to \mathbb{F}_2$ не является функцией вида

$$\vartheta(z) = c_0 + \mathrm{tr}_1^p\left(c_1 z^{2^1-1} + \cdots + c_{p-1} z^{2^{p-1}-1}\right) + c_p z^{2^p-1}, \tag{1}$$

$z \in \mathbb{F}_{2^p}$, $c_0, c_p \in \mathbb{F}_2$, $c_1, c_2, \ldots, c_{n-1} \in \mathbb{F}_{2^p}$. Тогда для f существует ровно $2^{2n+1} - 2^n$ бент-функций на минимальном расстоянии от неё, т. е. рассматриваемая оценка достигается.

Замечание 1.

1) Пространство $\mathbb{F}_{2^n}$ можно разбить на линейные подпространства $q\mathbb{F}_{2^p}$, $q \in \mathbb{F}_{2^n}^*$, пересекающиеся только в нуле, причём одно и то же подпространство можно описать с помощью любого ненулевого q из этого подпространства. При этом функция δ обеспечивает, что все ненулевые элементы каждого подпространства $q\mathbb{F}_{2^p}$ отображены в один и тот же элемент $\mathbb{F}_{2^p}^*$ независимо от выбора $q \in \mathbb{F}_{2^n}^*$ для $q\mathbb{F}_{2^p}$. Поэтому определение функции φ корректно.
2) Для функции φ' верны результаты, доказанные в [7]. В частности, можно показать, что функций, не удовлетворяющих условию (1), в точности $2^{2^p} - 2^{p^2-p+2}$, и при любом простом $p \geqslant 5$ можно использовать $\varphi'(y) = \mathrm{tr}_1^p(y^5)$ и $\varphi'(y) = \mathrm{tr}_1^p(y^{11})$.

3) Поскольку существуют функции δ и φ', то существуют и функции φ, удовлетворяющие условию.

Таким образом, путём перебора δ и φ' теорема 1 позволяет явно конструировать функции, для которых оценка достигается.

2. Усиление необходимого условия достижимости оценки

Приведём одно из определений *APN-функции* — это функция вида $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, такая, что ни для какого аффинного подпространства $U \subseteq \mathbb{F}_{2^n}$ размерности 2 образ $F(U)$ не является аффинным подпространством $\mathbb{F}_{2^n}$.

В работе [7] показано, что необходимым условием достижимости нижней оценки является использование APN-перестановки в исходной бент-функции из $\mathcal{M}_{2n}$. Для усиления этого условия нам потребуется следующее свойство APN-функций.

Теорема 2. Пусть $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ — APN-функция, $\varphi : \mathbb{F}_{2^n} \to \mathbb{F}_2$ — булева функция, U — аффинное подпространство $\mathbb{F}_{2^n}$ размерности 3 и $\sum\limits_{x \in U} F(x) = 0$. Тогда существует аффинная функция $H : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, такая, что $\varphi(x) + \mathrm{tr}_1^n\left(F(x)H(x)\right)$ является аффинной функцией на U.

Отметим, что в этой формулировке вместо условия $\sum\limits_{x \in U} F(x) = 0$ можно использовать более сильное условие: $F(U)$ — аффинное подпространство $\mathbb{F}_{2^n}$. С помощью этого свойства сужения APN-функций на подпространства размерности 3 можно усилить необходимое условие достижимости нижней оценки.

Теорема 3. Пусть $f(x, y) = \mathrm{tr}_1^n\left(x\pi(y)\right) + \varphi(y) \in \mathcal{M}_{2n}$, $x, y \in \mathbb{F}_{2^n}$ и для некоторого аффинного подпространства $U \subseteq \mathbb{F}_{2^n}$ размерности 2 или 3 верно, что $\pi(U)$ — также аффинное подпространство $\mathbb{F}_{2^n}$. Тогда число бент-функций, лежащих на минимальном расстоянии от f, строго больше нижней оценки $2^{2n+1} - 2^n$.

Таким образом, необходимым условием является то, что π не переводит никакие аффинные подпространства размерности 2 и 3 в аффинные подпространства. Другими словами, π должна быть APN-перестановкой, которая не переводит никакие аффинные подпространства размерности 3 в аффинные подпространства.

Отсюда можно получить также недостижимость оценки при $n = 2, 3$.

Следствие 1. Для $n = 2, 3$ не существует бент-функции из $\mathcal{M}_{2n}$, на минимальном расстоянии от которой лежит в точности $2^{2n+1} - 2^n$ бент-функций.

Отметим, что вопрос существования APN-перестановок при чётном $n \geqslant 8$ является открытым и известен как «The Big APN Problem». В то же время мы полагаем, что рассматриваемая оценка достигается при нечётных $n \geqslant 5$.

Гипотеза 1. Для любого нечётного $n \geqslant 5$ существует бент-функция из $\mathcal{M}_{2n}$, для которой количество бент-функций, лежащих на минимальном расстоянии от неё, равно $2^{2n+1} - 2^n$.

ЛИТЕРАТУРА

1. *Tokareva N.* Bent Functions: Results and Applications to Cryptography. Academic Press, 2015. 220 p.
2. *Rothaus O.* On bent functions // J. Comb. Theory. Ser. A. 1976. V. 20. No. 3. P. 300–305.
3. *Коломеец Н. А., Павлов А. В.* Свойства бент-функций, находящихся на минимальном расстоянии друг от друга // Прикладная дискретная математика. 2009. № 4(6). С. 5–20.

4. *Коломеец Н. А.* Перечисление бент-функций на минимальном расстоянии от квадратичной бент-функции // Дискретн. анализ и исслед. опер. 2012. Т. 19. Вып. 1. С. 41–58.
5. *Коломеец Н. А.* Верхняя оценка числа бент-функций на расстоянии 2^k от произвольной бент-функции от $2k$ переменных // Прикладная дискретная математика. 2014. № 3(25). С. 28–39.
6. *Kolomeec N.* The graph of minimal distances of bent functions and its properties // Des. Codes Cryptogr. 2017. V. 85. No. 3. P. 395–410.
7. *Быков Д. А.* О нижней оценке числа бент-функций на минимальном расстоянии от бент-функций из класса Мэйорана — МакФарланда // Прикладная дискретная математика. Приложение. 2022. № 15. С. 22–25.
8. *Логачев О. А., Сальников А. А., Смышляев С. В., Ященко В. В.* Булевы функции в теории кодирования и криптологии. М.: МЦНМО, 2012. 584 с.

УДК 511.336+519.113.6 DOI 10.17223/2226308X/16/5

НЕКОТОРЫЕ КЛАССЫ УСТОЙЧИВЫХ ФУНКЦИЙ НАД КОЛЬЦАМИ ГАЛУА И ИХ ЛИНЕЙНЫЕ ХАРАКТЕРИСТИКИ

О. В. Камловский, К. Н. Панков

Определяется линейная характеристика функций, заданных на кольце Галуа, которая задаёт «близость» рассматриваемых функций к классу всех аффинных функций данного кольца. Строятся некоторые классы устойчивых функций над кольцами Галуа и оцениваются их линейные характеристики.

Ключевые слова: *дискретные функции, устойчивые функции, кольца Галуа, линейная характеристика функций.*

Введение

Важной задачей криптографии является построение устойчивых дискретных функций, достаточно «удалённых» от класса всех аффинных функций. Устойчивые двоичные функции используются в протоколах квантового распределения ключей [1]. Такие функции зачастую участвуют при построении криптографических примитивов современных симметричных алгоритмов шифрования и представляют интерес при создании постквантовых механизмов защиты информации.

Тематика, связанная с построением подобных функций, значительно проработана для случая функций, заданных на конечных полях [2–5]. Устойчивые функции над кольцами Галуа, которые включают в себя все конечные поля и примарные кольца вычетов, изучены значительно меньше [6]. В работе [7] получено описание корреляционно-иммунных и устойчивых функций, заданных на произвольных конечных алфавитах в терминах спектральных коэффициентов функций. Кроме того, в этой работе построены классы устойчивых функций над конечными полями, имеющих максимально возможную алгебраическую степень.

Данная работа посвящена вопросам построения устойчивых функций над кольцами Галуа, достаточно удалённых от аффинных функций. В качестве меры приближения используется линейная характеристика, которая для функций над конечными полями предложена в [8]. Линейная характеристика, как и функция «близость» из [9] и функция «согласие» из [10], основана на похожих свойствах, однако линейная характеристика выглядит более естественно и проще. Она совпадает с максимальным по модулю коэффициентом корреляции [11] между столбцом значений дискретной функции и столбцами значений всех аффинных функций.

1. Линейная характеристика функций и ее свойства

Пусть $R = \mathrm{GR}(q^l, p^l)$ — произвольное кольцо Галуа из q^l элементов, имеющее характеристику p^l, где $q = p^t$; p — простое число; t, l — натуральные числа [12, 13]. Рассмотрим функцию $f : R^n \to R$ от n переменных, заданную на кольце R. Будем использовать обозначение $f = f(x_1, \ldots, x_n) = f(\mathbf{x})$, где $\mathbf{x} = (x_1, \ldots, x_n)$.

Группа всех аддитивных характеров кольца R (гомоморфизмов группы $(R, +)$ в мультипликативную группу поля комплексных чисел) состоит из гомоморфизмов [14, 15]

$$\chi_a(x) = \exp\left\{2\pi i \frac{\mathrm{tr}_{R_0}^R(ax)}{p^l}\right\}, \quad x \in R,$$

где $a \in R$; $R_0 = \{0, e, 2e, \ldots, (p^l - 1)e\}$ — подкольцо кольца R, порождённое единицей e; $\mathrm{tr}_{R_0}^R$ — функция «след» из кольца R в подкольцо R_0. В дальнейшем будем использовать обозначение $\chi = \chi_e$ для канонического аддитивного характера.

Коэффициентом кросс-корреляции между функциями $f(\mathbf{x})$ и $g(\mathbf{x})$, соответствующим элементу $a \in R$, называют комплексное число

$$C_a(f, g) = \sum_{x_1, \ldots, x_n \in R} \chi(af(\mathbf{x}) - g(\mathbf{x})).$$

Модуль коэффициента $C_a(f, g)$ характеризует «близость» между функциями $af(\mathbf{x})$ и $g(\mathbf{x})$. Чем меньше величина $|C_a(f, g)|$, тем больше отличаются друг от друга рассматриваемые функции [11].

Обозначим через $A_n(R)$ множество всех аффинных функций $g(\mathbf{x})$ от n переменных над кольцом R, т. е. функций вида

$$g(\mathbf{x}) = a_0 + a_1x_1 + \ldots + a_nx_n = a_0 + \langle \mathbf{a}, \mathbf{x} \rangle,$$

где $a_0 \in R$; $\mathbf{a} = (a_1, \ldots, a_n) \in R^n$.

Линейной характеристикой функции f назовём число

$$C(f) = \max_{a \in R \setminus \{0\}} \max_{g \in A_n(R)} |C_a(f, g)|.$$

Рассмотрим, как ведёт себя параметр $C(f)$ в частном случае $R = \mathrm{GR}(2, 2) = \mathrm{GF}(2) = \{0, e\}$. В этой ситуации $l = 1$, $a = e$,

$$\exp\left\{2\pi i \frac{Tr_{R_0}^R(af(\mathbf{x}) - a_0 - a_1x_1 - \ldots - a_nx_n)}{p^l}\right\} = (-1)^{f(\mathbf{x}) \oplus a_0 \oplus a_1x_1 \oplus \ldots \oplus a_nx_n},$$

модуль коэффициента кросс-корреляции $C_e(f, a_0 \oplus \langle \mathbf{a}, \mathbf{x} \rangle)$ равен модулю коэффициента Уолша — Адамара

$$W_f(\mathbf{a}) = \sum_{x_1, \ldots, x_n \in \mathrm{GF}(2)} (-1)^{f(\mathbf{x}) \oplus a_1x_1 \oplus \ldots \oplus a_nx_n}$$

булевой функции f и справедливо равенство

$$C(f) = \max_{\mathbf{a} \in \mathrm{GF}(2)^n} |W_f(\mathbf{a})|.$$

Отметим, что в двоичном случае для измерения удалённости функции от класса всех аффинных функций часто используют нелинейность $\mathrm{nl}(f)$ булевой функции f.

Она равна расстоянию Хэмминга между столбцом значений функции f и столбцами значений всех аффинных двоичных функций от n переменных. Известно [16], что

$$\mathrm{nl}(f) = 2^{n-1} - \frac{1}{2}\max_{\mathbf{a}\in \mathrm{GF}(2)^n} |W_f(\mathbf{a})| = 2^{n-1} - C(f)/2.$$

Укажем некоторые свойства линейной характеристики функции. Обозначим через $N(f, g, b)$ число всех векторов $\mathbf{x} \in R^n$, таких, что $f(\mathbf{x}) - g(\mathbf{x}) = b$.

Утверждение 1. Для всех $g \in A_n(R)$, $b \in R$ справедлива оценка

$$|N(f, g, b) - q^{l(n-1)}| \leqslant \frac{q^l - 1}{q^l} C(f).$$

Таким образом, чем меньше линейная характеристика $C(f)$, тем ближе величина $N(f, g, b)$ к своему естественному «среднему» значению $q^{l(n-1)}$.

Приведём нижнюю оценку для линейной характеристики.

Утверждение 2. Верна оценка $C(f) \geqslant q^{nl/2}$.

Дадим описание бент-функций в терминах линейной характеристики. Функцию f назовём бент-функцией [9, 10, 17], если $|C_a(f, g)| = q^{nl/2}$ для всех $a \in R \setminus \{0\}$ и $g(\mathbf{x}) \in$ $\in A_n(R)$.

Утверждение 3. Функция f является бент-функцией тогда и только тогда, когда $C(f) = q^{nl/2}$, т. е. неравенство из утверждения 2 обращается в равенство.

2. Корреляционно-иммунные и устойчивые функции

Пусть k — натуральное число. Для любых элементов $a_1, \ldots, a_k \in R$ и различных чисел $i_1, \ldots, i_k \in \{1, 2, \ldots, n\}$ обозначим через $f^{a_1,\ldots,a_k}_{i_1,\ldots,i_k}$ функцию, полученную из $f(x_1, \ldots, x_n)$ фиксацией переменных $x_{i_1}, \ldots, x_{i_k}$ значениями $a_1, \ldots, a_k$ соответственно. Назовем функцию f корреляционно-иммунной порядка k, если для всех $a_1, \ldots, a_k \in R$, $i_1, \ldots, i_k$, таких, что $1 \leqslant i_1 < \ldots < i_k \leqslant n$, и всех $z \in R$ для прообразов элемента z при действии отображений $f^{a_1,\ldots,a_k}_{i_1,\ldots,i_k}$ и f верно равенство

$$\left|\left(f^{a_1,\ldots,a_k}_{i_1,\ldots,i_k}\right)^{-1}(z)\right| = \frac{|f^{-1}(z)|}{q^{lk}}.$$

Назовём функцию f сбалансированной, если для всех $z \in R$ выполнено соотношение $|f^{-1}(z)| = q^{l(n-1)}$. Корреляционно-иммунную порядка k функцию f, которая является сбалансированной, называют ещё k-устойчивой функцией.

Нетрудно заметить, что если функция является корреляционно-иммунной порядка k, то она является корреляционно-иммунной порядка $k - 1$. Если функция f является 1-устойчивой, то f сбалансирована. Сбалансированные функции считают 0-устойчивыми. Обозначим через $\|\mathbf{a}\|$ число ненулевых координат вектора $\mathbf{a}$.

Приведём несколько известных фактов, сформулированных в терминах коэффициентов кросс-корреляции функций.

Теорема 1 [7]. Функция $f : R^n \to R$ является корреляционно-иммунной порядка k тогда и только тогда, когда для каждого $\mathbf{a} \in R^n$, такого, что $1 \leqslant \|\mathbf{a}\| \leqslant k$, при всех $a \in R \setminus \{0\}$ имеет место равенство $C_a(f, \langle \mathbf{a}, \mathbf{x}\rangle) = 0$.

Утверждение 4 [7]. Функция f является сбалансированной тогда и только тогда, когда $C_a(f, 0) = 0$ для всех $a \in R \setminus \{0\}$.

Таким образом, справедлив следующий критерий k-устойчивости функции.

Следствие 1. [7] Функция $f : R^n \to R$ является k-устойчивой тогда и только тогда, когда для каждого $\mathbf{a} \in R^n$, такого, что $0 \leqslant \|\mathbf{a}\| \leqslant k$, при всех $a \in R \setminus \{0\}$ имеет место равенство $C_a(f, \langle \mathbf{a}, \mathbf{x} \rangle) = 0$.

Пусть $g_1, \ldots, g_{n+1}$ — произвольные подстановки на множестве R,

$$f(x_1, \ldots, x_n) = g_{n+1}(g_1(x_1) + \ldots + g_n(x_n)).$$

Несложно показать, что такая функция f является $(n-1)$-устойчивой и класс k-устойчивых функций не пуст при каждом $k < n$.

3. Некоторые конструкции k-устойчивых функций

Пусть $R = \mathrm{GR}(q^l, p^l) = \{r_1, \ldots, r_{q^l}\}$, $f_{r_i} : R^{n-1} \to R$ — функции от $n-1$ переменных, где $i = 1, \ldots, q^l$. Зададим функцию $f : R^n \to R$ по правилу

$$f(x_1, x_2, \ldots, x_n) = \begin{cases} f_{r_1}(x_2, \ldots, x_n), & \text{если } x_1 = r_1, \\ \vdots & \\ f_{r_{q^l}}(x_2, \ldots, x_n), & \text{если } x_1 = r_{q^l}. \end{cases} \tag{1}$$

Функцию f будем называть разветвлением функций $f_{r_1}, \ldots, f_{r_{q^l}}$. Аналогичная конструкция для булевых функций рассмотрена в работе [2].

Теорема 2. Пусть функция f построена по правилу (1). Тогда:

1) функция f сбалансирована тогда и только тогда, когда для всех $a \in R \setminus \{0\}$

$$C_a(f_{r_1}, 0) + \ldots + C_a(f_{r_{q^l}}, 0) = 0;$$

2) если функции $f_{r_1}, \ldots, f_{r_{q^l}}$ являются корреляционно-иммунными порядка k и $C_a(f_{r_1}, 0) = \ldots = C_a(f_{r_{q^l}}, 0)$ для всех $a \in R \setminus \{0\}$, то функция f является корреляционно-иммунной порядка k;
3) если функции $f_{r_1}, \ldots, f_{r_{q^l}}$ являются k-устойчивыми, то и функция f является k-устойчивой;
4) линейные характеристики функций связаны соотношением

$$C(f) \leqslant C(f_{r_1}) + \ldots + C(f_{r_{q^l}}).$$

Следствие 2. Пусть функция f построена по правилу (1), где $f_{r_1}, \ldots, f_{r_{q^l}}$ — бент-функции, тогда линейная характеристика функции f удовлетворяет неравенству

$$C(f) \leqslant q^{(n+1)l/2}.$$

Применим конструкцию Майорана — МакФарланда для построения устойчивых функций. Пусть $R = \mathrm{GR}(q^l, p^l)$, $n = 2k$, $\varphi : R^k \to R^k$ — преобразование на множестве R^k с координатными функциями $\varphi_1, \ldots, \varphi_k$, т. е.

$$\varphi(\mathbf{x}) = (\varphi_1(\mathbf{x}), \ldots, \varphi_k(\mathbf{x})), \quad \mathbf{x} \in R^k.$$

Для произвольной функции $h : R^k \to R$ и всех $\mathbf{x}, \mathbf{y} \in R^k$ определим функцию $f : R^n \to R$ равенствами

$$f(\mathbf{x}, \mathbf{y}) = \langle \varphi(\mathbf{x}), \mathbf{y} \rangle + h(\mathbf{x}) = \varphi_1(\mathbf{x}) y_1 + \ldots + \varphi_k(\mathbf{x}) y_k + h(\mathbf{x}). \tag{2}$$

В [6] показано, что если $\varphi(R^k) \subset (R^*)^k$, где R^* — мультипликативная группа кольца R, то функция f является $(k-1)$-устойчивой.

Теорема 3. Пусть функция f определена равенством (2). Тогда если $|\varphi^{-1}(\mathbf{c})| \leqslant t$ для всех $\mathbf{c} \in R^k$, то линейная характеристика функции f удовлетворяет условию

$$C(f) \leqslant tq^{k(2l-1)}.$$

Оценки из теоремы 3 являются наиболее точными в случае, когда $l = 1$, т. е. $R = \mathrm{GF}(q)$ — поле из q элементов. Если при этом $t = 1$ (φ — подстановка на R), то получится известный класс бент-функций Майорана — МакФарланда [18, 19].

ЛИТЕРАТУРА

1. *Панков К. Н.* Оценки мощности классов отображений, применяемых в протоколах квантового распределения ключей // Наукоемкие технологии в космических исследованиях Земли. 2022. № 4. С. 4–18.
2. *Camion P., Carlet C., Charpin P., and Sendrier N.* On correlation-immune functions // LNCS. 1992. V. 576. P. 86–100.
3. *Dobbertin H.* Construction of bent functions and balanced boolean functions with high nonlinearity // LNCS. 1995. V. 1008. P. 61–74.
4. *Xiao G-Z. and Massey J. L.* A spectral characterization on correlation-immune combining functions // IEEE Trans. Inform. Theory. 1988. No. 3. P. 569–571.
5. *Камловский О. В., Панков К. Н.* Классы сбалансированных функций над конечными полями, обладающих малым значением линейной характеристики // Проблемы передачи информации. 2022. № 4. С. 103–117.
6. *Carlet C.* More correlation-immune and resilient functions over Galois fields and Galois rings // LNCS. 1997. V. 1233. P. 422–433.
7. *Camion P. and Canteaut A.* Correlation-immune and resilient function over finite alphabets and their applications in cryptography // Des. Codes Cryptogr. 1999. V. 16. P. 121–149.
8. *Бугров А. Д.* Кусочно-аффинные подстановки конечных полей // Прикладная дискретная математика. 2015. № 4(30). С. 5–23.
9. *Солодовников В. И.* Бент-функции из конечной абелевой группы в конечную абелеву группу // Дискретная математика. 2002. № 1. С. 99–113.
10. *Кузьмин А. С., Марков В. Т., Нечаев А. А. и др.* Бент-функции и гипербент-функции над полем из 2^l элементов // Проблемы передачи информации. 2008. № 1. С. 15–37.
11. *Golomb S. W. and Gong G.* Signal Design for Good Correlation. Cambridge: Cambridge University Press, 2005.
12. *Нечаев А. А.* Код Кердока в циклической форме // Дискретная математика. 1989. № 4. С. 123–139.
13. *Нечаев А. А.* Цикловые типы линейных подстановок над конечными коммутативными кольцами // Математический сборник. 1993. № 3. С. 21–56.
14. *Лидл Р., Нидеррайтер Г.* Конечные поля. М.: Мир, 1988.
15. *Камловский О. В.* Частотные характеристики линейных рекуррентных последовательностей над кольцами Галуа // Математический сборник. 2009. № 4. С. 31–52.
16. *Логачев О. А., Сальников А. А., Смышляев С. В., Ященко В. В.* Булевы функции в теории кодирования и криптологии. М.: МЦНМО, 2012.
17. *Амбросимов А. С.* Свойства бент-функций q-значной логики над конечными полями // Дискретная математика. 1994. № 3. С. 50–60.
18. *McFarland R. L.* A family of noncyclic difference sets // J. Combin. Theory. Ser. A. 1973. No. 15. P. 1–10.
19. *Токарева Н. Н.* Обобщения бент-функций. Обзор работ // Дискретный анализ и исследование операций. 2010. № 1. С. 34–64.

УДК 519.7 DOI 10.17223/2226308X/16/6

О СОХРАНЕНИИ СТРУКТУРЫ ПОДПРОСТРАНСТВ ВЕКТОРНЫМИ БУЛЕВЫМИ ФУНКЦИЯМИ[1]

Н. А. Коломеец

Рассматривается сохранение функцией $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ структуры аффинного подпространства $U \subseteq \mathbb{F}_2^n$, т. е. случаи, когда $F(U) = \{F(x) : x \in U\}$ является аффинным подпространством $\mathbb{F}_2^m$. Приводится связь данного свойства с наличием у F компонентных функций, ограничения которых на рассматриваемое подпространство являются постоянными, а также с оценками нелинейности и порядка дифференциальной равномерности F. Доказано, что множество размерностей аффинных подпространств, структуру которых сохраняет функция обращения элементов поля $\mathbb{F}_{2^n}$, является наименьшим среди всех взаимно однозначных мономиальных функций.

Ключевые слова: *аффинные подпространства, инвариантные подпространства, нелинейность, дифференциальная равномерность, APN-функции, мономиальные функции.*

Введение

Будем называть *подпространством* аффинное подпространство $U \subseteq \mathbb{F}_2^n$, т. е. $U = a \oplus L = \{a \oplus x : x \in L\}$, где L — линейное подпространство $\mathbb{F}_2^n$ и $a \in \mathbb{F}_2^n$. Будем говорить, что векторная булева функция $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ *сохраняет структуру* U, если $F(U) = \{F(x) : x \in U\}$ — подпространство $\mathbb{F}_2^m$. Аналогично, F *разрушает структуру* U, если $F(U)$ не является подпространством $\mathbb{F}_2^m$. Назовём подпространства $\mathbb{F}_2^n$ размерности $0, 1$ и n *тривиальными*, поскольку любые подмножества $\mathbb{F}_2^n$ мощности $1, 2$ и 2^n являются подпространствами. Обозначим через $\langle S \rangle$ минимальное по включению подпространство $\mathbb{F}_2^n$, содержащее множество $S \subseteq \mathbb{F}_2^n$.

Интерес к сохранению/разрушению структуры подпространства векторной булевой функцией обусловлен связью с инвариантными подпространствами отображений. Зная все подпространства, структуру которых сохраняет взаимно однозначная F, можно определить все инвариантные подпространства как самой функции, так и всех функций, аффинно эквивалентных ей. Напомним, что множество $S \subseteq \mathbb{F}_2^n$ называется *инвариантным* относительно F, если $F(S) \subseteq S$. В [1] предложена атака, использующая наличие у раундовой функции G некоторой SP-сети с nk-битным блоком линейного подпространства $L \subseteq \mathbb{F}_2^{nk}$ и $a, b \in \mathbb{F}_2^{nk}$, таких, что

$$G(a \oplus L) = b \oplus L.$$

Если G — композиция слоя S-блоков $S_1, \ldots, S_k$ вида $\mathbb{F}_2^n \to \mathbb{F}_2^n$ и линейного преобразования на $\mathbb{F}_2^{nk}$, то необходимым условием существования такого $a \oplus L$ является сохранение каждым S-блоком S_i структуры подпространства

$$\{(x_{1+n(i-1)} \ldots x_{n+n(i-1)}) : x \in a \oplus L\}$$

для всех $i \in \{1, \ldots, k\}$. Эти подпространства являются проекциями $a \oplus L$ на вход соответствующего S-блока. В этом контексте могут быть интересны S-блоки, сохраняющие структуру минимального числа подпространств и тем самым упрощающие проверку существования таких инвариантных $a \oplus L$.

[1] Работа выполнена в рамках госзадания ИМ СО РАН (проект № FWNF–2022–0018).

Информация об инвариантных подпространствах S-блоков также полезна в силу большего удобства работы с ними. Например, вопрос построения L для упомянутой атаки, начиная с инвариантных подпространств S-блоков, рассматривается в [2]. Есть обобщение данной атаки [3], построение инвариантных множеств для которой можно также начинать с S-блоков [4].

Отметим и прямую связь с APN-функциями [5]: $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ является APN-функцией, если и только если F разрушает структуру любого подпространства $\mathbb{F}_2^n$ размерности 2. Множество открытых вопросов, связанных с этим классом функций, можно найти в [6].

В данной работе продолжаются исследования сохранения структуры подпространств отображениями, начатые в [7, 8]. Приводится связь этого свойства с оценками на нелинейность и порядок дифференциальной равномерности функции, предложены примеры функций, гарантированно разрушающих структуру подпространств больших размерностей, а также показано наличие таких подпространств у взаимно однозначных мономиальных функций.

1. Нелинейность и порядок дифференциальной равномерности

Компонентной функцией для $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ называется функция вида $x \mapsto \langle a, F(x)\rangle$ для всех $a \neq 0$ из $\mathbb{F}_2^m$. Здесь $\langle a, y\rangle = a_1y_1 \oplus \ldots \oplus a_my_m$ при $y \in \mathbb{F}_2^m$.

Утверждение 1. Пусть $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ и U — подпространство $\mathbb{F}_2^n$, тогда как минимум $2^{m-\dim\langle F(U)\rangle} - 1$ компонентных функций F постоянны на U.

Более того, для взаимно однозначных функций справедлив следующий критерий:

Утверждение 2. Пусть $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ взаимно однозначна и U — подпространство $\mathbb{F}_2^n$. Тогда F сохраняет структуру U, если и только если среди компонентных функций F ровно $2^{n-\dim U} - 1$ постоянны на U.

Отметим, что булевы функции, постоянные на некотором подпространстве размерности k, называются k-нормальными [9].

Порядком дифференциальной равномерности $\delta(G)$ функции $G : u \oplus L \to u' \oplus L'$, где L и L' — линейные подпространства $\mathbb{F}_2^n$ и $\mathbb{F}_2^m$ соответственно, $u \in \mathbb{F}_2^n$ и $u' \in \mathbb{F}_2^m$, называется минимальное t, такое, что при любых параметрах $a \in L \setminus \{0\}$ и $b \in L'$ уравнение $G(x) \oplus G(x \oplus a) = b$ имеет не более t решений относительно $x \in u \oplus L$. Мы рассматриваем подпространства исключительно для корректного использования параметра $\delta(F|_U)$ для ограничения функции, все его свойства полностью соответствуют свойствам $\delta(\cdot)$ для функций вида $\mathbb{F}_2^{\dim L} \to \mathbb{F}_2^{\dim L'}$. Если $|L| = |L'|$ и $\delta(G) = 2$, то G называется APN-функцией.

Утверждение 3. Пусть $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ сохраняет структуру подпространства $U \subseteq \mathbb{F}_2^n$. Тогда $\delta(F) \geqslant \delta(F|_U)$.

В общем случае сохранение структуры подпространств определённых размерностей не ограничивает порядок дифференциальной равномерности функции, примером чего является функция обращения элементов конечного поля (см. следующий пункт). Однако утверждение 3 может сработать, например, в следующем случае.

Пример 1. Пусть $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ сохранят структуру подпространства $U \subseteq \mathbb{F}_2^n$ размерности 4, причём $\dim F(U) = 4$. Тогда F не является APN-функцией.

В дополнение к подпространствам размерности 2, взаимно однозначные APN-функции не могут сохранять структуру также подпространств размерности $n - 1$.

Утверждение 4. Пусть $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ взаимно однозначна и сохраняет структуру некоторого подпространства $\mathbb{F}_2^n$ размерности $n-1$. Тогда F не является APN-функцией.

Отметим, что сохранение структуры подпространств размерности $n-1$ можно рассматривать в контексте свойств подфункций APN-функции [10].

Используя результаты [11] (см. также [12]), можно оценить и *нелинейность* функции, которая определяется как расстояние Хэмминга между ближайшими друг к другу её компонентной функцией и аффинной функцией от того же числа переменных.

Утверждение 5. Пусть $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ сохраняет структуру подпространства $U \subseteq \mathbb{F}_2^n$. Тогда $N_F \leqslant 2^{n-1} - 2^{\dim U - 1}$.

Таким образом, функции с высокой нелинейностью разрушают структуру подпространств больших размерностей. Например:

— если n нечётное, то AB-функции (например, функции Голда и Касами) разрушают структуру подпространств $\mathbb{F}_2^n$, размерность которых больше $n/2$;
— если n чётное, те же функции Голда и Касами разрушают структуру подпространств, размерность которых больше $n/2+1$.

2. Мономиальные функции

Функция $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, где через $\mathbb{F}_{2^n}$ обозначено конечное поле, состоящее из 2^n элементов, называется *мономиальной*, если для всех $x \in \mathbb{F}_{2^n}$ справедливо

$$F(x) = \alpha x^k, \text{ где } \alpha \in \mathbb{F}_{2^n} \text{ и } k \in \mathbb{N}.$$

Напомним, что такие функции можно рассматривать и как функции вида $\mathbb{F}_2^n \to \mathbb{F}_2^n$, зафиксировав в поле $\mathbb{F}_{2^n}$ некоторый базис.

В [7] (см. также [8]) доказано, что функция обращения элементов поля $\mathbb{F}_{2^n}$, в мономиальном виде записывающаяся как x^{2^n-2}, сохраняет структуру только определённых подпространств $\mathbb{F}_{2^n}$ размерностей k для любого $k \mid n$. При нечётном n она является APN-функцией. Таким образом, APN-функции могут сохранять структуру множества подпространств разной размерности. В то же время следующая теорема говорит о том, что функцию обращения элементов конечного поля можно считать одной из «лучших» среди всех мономиальных функций.

Теорема 1. Пусть $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ взаимно однозначна и построена по некоторой мономиальной функции. Тогда для всех $k \mid n$ функция F сохраняет структуру некоторого подпространства $\mathbb{F}_2^n$ размерности k.

Таким образом, функция обращения элементов конечного поля разрушает структуру подпространств максимального количества размерностей. Отметим также, что мономиальные функции — одна из самых распространённых конструкций APN-функций. При этом среди них нельзя найти взаимно однозначные функции, которые разрушают структуру всех нетривиальных подпространств при составных n.

ЛИТЕРАТУРА

1. *Leander G., Abdelraheem M. A., AlKhzaimi H., and Zenner E.* A cryptanalysis of PRINTcipher: The invariant subspace attack // LNCS. 2011. V. 6841. P. 206–221.
2. *Трифонов Д. И., Фомин Д. Б.* Об инвариантных подпространствах в XSL-шифрах // Прикладная дискретная математика. 2021. № 54. С. 58–76.
3. *Todo Y., Leander G., and Sasaki Y.* Nonlinear invariant attack: practical attack on full SCREAM, iSCREAM, and Midori64 // LNCS. 2016. V. 10032. P. 3–33.

4. *Буров Д. А.* О существовании нелинейных инвариантов специального вида для раундовых преобразований XSL-алгоритмов // Дискретная математика. 2021. Т. 33. № 2. С. 31–45.
5. *Nyberg K.* Differentially uniform mappings for cryptography // LNCS. 1994. V. 765. P. 245–265.
6. *Carlet C.* Open questions on nonlinearity and on APN functions // LNCS. 2015. V. 9061. P. 83–107.
7. *Kolomeec N. and Bykov D.* On the Image of an Affine Subspace under the Inverse Function within a Finite Field. arXiv preprint arXiv:2206.14980. `https://arxiv.org/abs/2206.14980`. 2022.
8. *Коломеец Н. А., Быков Д. А.* Об инвариантных подпространствах функций, аффинно эквивалентных обращению элементов конечного поля // Прикладная дискретная математика. Приложение. 2022. № 15. С. 5–8.
9. *Charpin P.* Normal Boolean functions // J. Complexity. 2004. V. 20. No. 2–3. P. 245–265.
10. *Городилова А. А.* Характеризация почти совершенно нелинейных функций через подфункции // Дискретная математика. 2015. Т. 27. № 3. С. 3–16.
11. *Canteaut A., Carlet C., Charpin P., and Fontaine C.* On cryptographic properties of the cosets of R(1, m) // IEEE Trans. Inform. Theory. 2001. V. 47. P. 1494–1513.
12. *Carlet C. and Feukoua S.* Three parameters of Boolean functions related to their constancy on affine spaces // Adv. Math. Commun. 2020. V. 14. No. 4. P. 651–676.

УДК 519.7 DOI 10.17223/2226308X/16/7

МАТРИЦЫ ГРАМА БЕНТ-ФУНКЦИЙ И СВОЙСТВА ПОДФУНКЦИЙ КВАДРАТИЧНЫХ САМОДУАЛЬНЫХ БЕНТ-ФУНКЦИЙ[1]

А. В. Куценко

Булева фунция от чётного числа переменных n называется бент-функцией, если она имеет спектр Уолша — Адамара, состоящий из чисел $\pm 2^{n/2}$. Бент-функция называется самодуальной, если она совпадает со своей дуальной бент-функцией. Ранее автором было сформулировано достаточное условие того, что подфункции от $n-2$ переменных самодуальной бент-функции от n переменных, полученные фиксацией первых двух переменных, являются бент-функциями. В настоящей работе доказано, что для квадратичных самодуальных бент-функций данное условие при $n \geqslant 6$ не является необходимым. Введено понятие «матрица Грама бент-функции», установлен общий вид матрицы Грама бент-функции и дуальной к ней функции. Доказано, что если матрица Грама бент-функции от n переменной является необратимой, её подфункции от $n-2$ переменных, полученные фиксацией первых двух переменных, являются бент-функциями. Установлено, что в этом случае подфункции дуальной к ней функции также являются бент-функциями.

Ключевые слова: *самодуальная бент-функция, подфункция, матрица Грама, квадратичная бент-функция, конкатенация бент-функций.*

Через $\mathbb{F}_2^n$ обозначим линейное пространство всех двоичных векторов длины n над полем $\mathbb{F}_2$. *Булевой функцией* от n переменных называется отображение вида $\mathbb{F}_2^n \to \mathbb{F}_2$. Множество всех булевых функций от n переменных обозначается через $\mathcal{F}_n$. *Характеристическим вектором* (характеристической последовательностью) булевой функ-

[1]Работа выполнена в рамках госзадания ИМ СО РАН (проект № FWNF-2022-0018).

ции $f \in \mathcal{F}_n$ называется вектор

$$F \equiv (-1)^f = \left((-1)^{f(0)}, (-1)^{f(1)}, \ldots, (-1)^{f(2^n-1)}\right) \in \{\pm 1\}^{2^n},$$

где $(f(0), f(1), \ldots, f(2^n-1)) \in \mathbb{F}_2^{2^n}$ — вектор значений функции f. Каждая булева функция от n переменных может быть единственным образом представлена в виде многочлена над полем $\mathbb{F}_2$:

$$f(x_1, x_2, \ldots, x_n) = \bigoplus_{i_1, i_2, \ldots i_n \in \mathbb{F}_2} a_{i_1 i_2 \ldots i_n} x_1^{i_1} x_2^{i_2} \cdot \ldots \cdot x_n^{i_n}.$$

Здесь $a_z \in \mathbb{F}_2$ для всех $z \in \mathbb{F}_2^n$ (с соглашением $0^0 = 1$). Данное представление называется *многочленом Жегалкина* булевой функции f. Степенью $\deg(f)$ функции f называется максимальная из степеней слагаемых, входящих в многочлен Жегалкина с ненулевыми коэффициентами. Если $\deg(f) = 2$, функция называется *квадратичной*.

Для каждой пары $x, y \in \mathbb{F}_2^n$ через $\langle x, y \rangle$ обозначим значение $\bigoplus_{i=1}^{n} x_i y_i$. *Преобразованием Уолша — Адамара* булевой функции f от n переменных называется целочисленная функция $W_f : \mathbb{F}_2^n \to \mathbb{Z}$, заданная равенством

$$W_f(y) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) \oplus \langle x, y \rangle}, \quad y \in \mathbb{F}_2^n.$$

Булева функция f от чётного числа переменных n называется *бент-функцией*, если $|W_f(y)| = 2^{n/2}$ для каждого $y \in \mathbb{F}_2^n$ [1]. Для множества бент-функций от n переменных используется обозначение $\mathcal{B}_n$. Для каждой $f \in \mathcal{B}_n$ из соотношения $W_f(y) = (-1)^{\widetilde{f}(y)} 2^{n/2}$ однозначным образом определяется *дуальная* к ней бент-функция $\widetilde{f} \in \mathcal{B}_n$. Бент-функция f называется *самодуальной* (*антисамодуальной*), если $f = \widetilde{f}$ (соответственно $f = \widetilde{f} \oplus 1$).

Изучению данного подкласса бент-функций посвящено множество работ. В частности, в [2–4] исследован вопрос аффинной классификации самодуальных бент-функций от $n \leqslant 8$ переменных, а также квадратичных самодуальных бент-функций относительно преобразований, сохраняющих (анти-)самодуальность. Конструкции самодуальных бент-функций представлены в работах [5–7]. Обзор известных метрических свойств приведён в [8].

Известно, что все подфункции от $n-2$ переменных бент-функции от n переменных имеют одинаковые спектры Уолша — Адамара [9]. Следовательно, либо все подфункции являются бент-функциями, либо $W_f(y) \in \left\{0, \pm 2^{(n+2)/2}\right\}$ для каждого $y \in \mathbb{F}_2^n$ (то есть все подфункции — почти бент-функции), либо их спектры Уолша — Адамара состоят из чисел 0, $\pm 2^{(n-2)/2}$, $\pm 2^{n/2}$.

Далее для булевой функции f от n переменных через (f_0, f_1, f_2, f_3) будем обозначать разложение её вектора значений на четыре подвектора, являющихся векторами значений её подфункций от $n-2$ переменных, полученных фиксацией первых двух переменных. Случай, когда данные подфункции являются бент-функциями, ведёт, в свою очередь, к итеративной конструкции бент-функции, вектор значений которой есть (f_0, f_1, f_2, f_3). В [10] найдены необходимые и достаточные условия, накладываемые на подфункции f_i, $i = 0, \ldots, 3$. В работах [11, 12] данные подфункции рассмотрены для случая, когда f является самодуальной бент-функцией.

1. Линейная независимость характеристических векторов подфункций квадратичной самодуальной бент-функции

В работе [12] доказано:

Теорема 1 [12]. Если характеристические векторы подфункций f_0, f_1, f_2, f_3 самодуальной бент-функции f линейно зависимы, то данные подфункции являются бент-функциями.

Этот результат описывает достаточное условие того, что все подфункции самодуальной бент-функции, полученные фиксацией первых двух переменных, являются бент-функциями. При этом для случая $n = 4$ данное условие также является необходимым. Хорошо известно, что все (самодуальные) бент-функции от 4 переменных являются квадратичными, что позволило обозначить следующий вопрос: является ли линейная зависимость характеристических векторов необходимым условием для *квадратичных* самодуальных функций?

Ответ на данный вопрос даёт следующее

Утверждение 1. Для каждого чётного $n \geqslant 6$ существуют квадратичные самодуальные бент-функции от n переменных, подфункции которых образуют линейно независимые множества характеристических векторов.

Таким образом, обращение теоремы 1 не имеет места при $n \geqslant 6$ и для квадратичных самодуальных бент-функций, то есть линейная зависимость характеристических векторов не является необходимым условием и, как и в случае без ограничения на степень, обеспечивает лишь достаточное условие того, что подфункции f_0, f_1, f_2, f_3 являются бент-функциями.

2. Матрица Грама произвольной бент-функции

Пусть $f \in \mathcal{B}_n$. *Матрицей Грама* $\mathrm{Gram}(f) = \left(g_{ij}\right)$ функции f назовём квадратную матрицу размера 4×4, элементами которой являются числа

$$g_{ij} = \sum_{x \in \mathbb{F}_2^{n-2}} (-1)^{f_i(x) \oplus f_j(x)}, \quad i, j = 0, 1, 2, 3,$$

которые являются скалярными произведениями характеристических векторов её подфункций.

Общий вид матриц Грама бент-функции и дуальной к ней описывает следующая

Теорема 2. Матрицы Грама бент-функции f от n переменных и дуальной к ней функции $\widetilde{f}$ имеют вид

$$\mathrm{Gram}(f) = \begin{pmatrix} 2^{n-2} & b & c & -a \\ b & 2^{n-2} & a & -c \\ c & a & 2^{n-2} & -b \\ -a & -c & -b & 2^{n-2} \end{pmatrix}, \quad \mathrm{Gram}\left(\widetilde{f}\right) = \begin{pmatrix} 2^{n-2} & c & b & -a \\ c & 2^{n-2} & a & -b \\ b & a & 2^{n-2} & -c \\ -a & -b & -c & 2^{n-2} \end{pmatrix}$$

для некоторых целых чисел a, b, c, таких, что

$$-2^{n-2} + |b + c| \leqslant a \leqslant 2^{n-2} - |b - c|.$$

Определители данных матриц совпадают, в частности, для f определитель имеет вид

$$\mathrm{Gramian}(f) = \left(2^{n-2} - a + b - c\right)\left(2^{n-2} - a - b + c\right)\left(2^{n-2} + a - b - c\right)\left(2^{n-2} + a + b + c\right).$$

Теорема 1 в терминах матриц Грама означает, что если матрица Грама самодуальной бент-функции являтся необратимой, то подфункции f_0, f_1, f_2, f_3 являются бент-функциями. Другими словами, для самодуальных бент-функций равенство $\mathrm{Gramian}(f) = 0$ влечёт тот факт, что указанные подфункции являются бент-функциями. Данный результат можно обобщить так:

Теорема 3. Если характеристические векторы подфункций f_0, f_1, f_2, f_3 бент-функции f линейно зависимы, то данные подфункции являются бент-функциями. Бент-функциями являются также подфункции дуальной функции $\widetilde{f}$.

Таким образом, данное утверждение позволяет получить достаточное условие того, что подфункции рассматриваемой бент-функции также являются бент-функциями и, кроме того, отображение дуальности сохраняет их максимальную нелинейность.

ЛИТЕРАТУРА

1. *Rothaus O.* On bent functions // J. Combin. Theory. Ser. A. 1976. V. 20. No. 3. P. 300–305.
2. *Carlet C., Danielsen L. E., Parker M. G., and Solé P.* Self-dual bent functions // Int. J. Inform. Coding Theory. 2010. V. 1. P. 384–399.
3. *Hou X.-D.* Classification of self dual quadratic bent functions // Des. Codes Cryptogr. 2012. V. 63. No. 2. P. 183–198.
4. *Feulner T., Sok L., Solé P., and Wassermann A.* Towards the classification of self-dual bent functions in eight variables // Des. Codes Cryptogr. 2013. V. 68. No. 1. P. 395–406.
5. *Luo G., Cao X., and Mesnager S.* Several new classes of self-dual bent functions derived from involutions // Cryptogr. Commun. 2019. V. 11. No. 6. P. 1261–1273.
6. *Li Y., Kan H., Mesnager S., et al.* Generic constructions of (Boolean and vectorial) bent functions and their consequences // IEEE Trans. Inform. Theory. 2022. V. 68. No. 4. P. 2735–2751.
7. *Su S. and Guo X.* A further study on the construction methods of bent functions and self-dual bent functions based on Rothaus's bent function // Des. Codes Cryptogr. 2023. V. 91. No. 4. P. 1559–1580.
8. *Kutsenko A. V. and Tokareva N. N.* Metrical properties of the set of bent functions in view of duality // Прикладная дискретная математика. 2020. № 49. С. 18–34.
9. *Canteaut A. and Charpin P.* Decomposing bent functions // IEEE Trans. Inf. Theory. 2003. V. 49. No. 8. P. 2004–2019.
10. *Preneel B., Van Leekwijck W., Van Linden L., et al.* Propagation characteristics of Boolean functions // LNCS. 1990. V. 473. P. 161–173.
11. *Kutsenko A.* Metrical properties of self-dual bent functions // Des. Codes Cryptogr. 2020. V. 88. No. 1. P. 201–222.
12. *Куценко А. В.* Свойства подфункций самодуальных бент-функций // Прикладная дискретная математика. Приложение. 2022. № 15. С. 26–30.

УДК 519.7 DOI 10.17223/2226308X/16/8

ПОСТРОЕНИЕ ПОДСТАНОВКИ НА $\mathbb{F}_2^n$ НА ОСНОВЕ ОДНОЙ БУЛЕВОЙ ФУНКЦИИ

И. А. Панкратова, А. А. Медведев

Приведены некоторые необходимые условия того, что векторная булева функция, координаты которой получены из одной булевой функции с помощью перестановок переменных, является подстановкой.

Ключевые слова: *подстановки, векторные булевы функции.*

Подстановки на $\mathbb{F}_2^n$ (обратимые векторные булевы функции $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$) используются во многих криптосистемах, в частности в криптосистемах с функциональными ключами [1, 2]. Одно из требований (эксплуатационных) к подстановке как ключу криптосистемы — это возможность её компактного задания: например, можно строить координатные функции подстановки на основе преобразований одной булевой функции.

В работе [3] предложена следующая конструкция векторной булевой функции:

$$F(x) = \big(f(x), f(\pi(x)), f(\pi^2(x)), \ldots, f(\pi^{n-1}(x))\big), \tag{1}$$

где $f(x)$ — булева функция от n переменных; π — циклический сдвиг вектора переменных влево на 1. Например, при $n = 3$ получаем

$$F(x_1, x_2, x_3) = \big(f(x_1, x_2, x_3), f(x_2, x_3, x_1), f(x_3, x_1, x_2)\big).$$

В данной работе рассмотрим обобщение этой конструкции: пусть $\pi \in \mathbb{S}_n$ — любая подстановка степени n; для $x = (x_1, \ldots, x_n) \in \mathbb{F}_2^n$ обозначим $\pi(x) = (x_{\pi(1)}, \ldots, x_{\pi(n)})$. Подстановка $\pi \in \mathbb{S}_n$ индуцирует подстановку π' на $\mathbb{F}_2^n$ по правилу

$$\pi'(a_1 \ldots a_n) = (a_{\pi(1)} \ldots a_{\pi(n)}), \quad a_1 \ldots a_n \in \mathbb{F}_2^n. \tag{2}$$

Например, для $n = 3$ и $\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ получим

$$\pi' = \begin{pmatrix} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\ 000 & 010 & 100 & 110 & 001 & 011 & 101 & 111 \end{pmatrix}.$$

Утверждение 1. Подстановки π и π' имеют одинаковый порядок:

$$\operatorname{ord}(\pi') = \operatorname{ord}(\pi).$$

Доказательство. Неравенство $\operatorname{ord}(\pi') \leqslant \operatorname{ord}(\pi)$ следует из (2).

Для доказательства обратного неравенства построим вектор $a = (a_1 \ldots a_n) \in \mathbb{F}_2^n$ так: для каждого цикла $(i_1, \ldots, i_s)$ подстановки π положим $a_{i_1} = 1$ и $a_{i_j} = 0$ для всех $j = 2, \ldots, s$. Тогда наименьшее значение k, при котором $(\pi')^k(a) = a$, равно $\operatorname{ord}(\pi)$. ■

Сформулируем некоторые необходимые условия того, что функция F, полученная по формуле (1) для некоторой $\pi \in \mathbb{S}_n$, является подстановкой на $\mathbb{F}_2^n$. Через c^n, $c \in \{0, 1\}$, будем обозначать булев вектор длины n, все компоненты которого равны c.

Утверждение 2. Пусть $f \in P_2(n)$, $\pi \in \mathbb{S}_n$ и функция

$$F(x) = \big(f(x), f(\pi(x)), f(\pi^2(x)), \ldots, f(\pi^{n-1}(x))\big)$$

является подстановкой на $\mathbb{F}_2^n$. Тогда:

1) функция f уравновешенная;
2) индуцированная подстановка π' на $\mathbb{F}_2^n$ имеет только две неподвижные точки — 0^n и 1^n; $f(0^n) \neq f(1^n)$; $\{F(0^n), F(1^n)\} = \{0^n, 1^n\}$;
3) π — полноцикловая подстановка;
4) $f \neq \text{const}$ ни на одном цикле подстановки π' (кроме неподвижных точек).

Доказательство.

1) Функция F является подстановкой, если и только если все её компоненты (ненулевые линейные комбинации координатных функций) уравновешены [4]. Отсюда следует уравновешенность f.
2) Тот факт, что $\pi'(0^n) = 0^n$ и $\pi'(1^n) = 1^n$, следует непосредственно из определения (2). По формуле (1) получаем: $F(0^n) = (f(0^n))^n$, $F(1^n) = (f(1^n))^n$. Поскольку F — подстановка, то $F(0^n) \neq F(1^n)$, значит, $f(0^n) \neq f(1^n)$ и $\{F(0^n), F(1^n)\} = \{0^n, 1^n\}$.
   Предположим, что подстановка π' имеет ещё одну неподвижную точку — $\pi'(a) = a$. Но тогда $F(a) = (f(a))^n \in \{0^n, 1^n\} = \{F(0^n), F(1^n)\}$, что противоречит тому, что F — подстановка.
3) Предположим, что π раскладывается в произведение нескольких циклов. Построим вектор $a = (a_1 \dots a_n) \in \mathbb{F}_2^n$ так: для каждого цикла $(i_1, \dots, i_s)$ подстановки π положим $a_{i_1} = \dots = a_{i_s}$ и $a \notin \{0^n, 1^n\}$ (это можно сделать, так как циклов больше одного). Но тогда $\pi'(a) = a$ — получили третью неподвижную точку, что противоречит п. 2.
4) Предположим, что $f = c \in \mathbb{F}_2$ на всех элементах цикла $(i_1, \dots, i_k)$ подстановки π', где $i_1 \notin \{0^n, 1^n\}$. Тогда по формуле (1) получим $F(i_1) = c^n \in \{F(0^n), F(1^n)\}$ — противоречие с тем, что F — подстановка.

Утверждение 2 доказано. ■

Замечание 1. В соответствии с п. 3 утверждения 2, если F — подстановка, то π — полноцикловая. Поскольку переименование (перенумерация) переменных не влияет на криптографические свойства функции, для их изучения можно ограничиться рассмотрением подстановки частного вида, предложенной в [3] (циклический сдвиг влево): $\pi(i) = i \bmod n + 1$, где n — количество переменных.

Замечание 2. Полноцикловость подстановки π является необходимым условием того, что функция F имеет максимально возможную компонентную алгебраическую иммунность [3].

В дальнейшем планируется рассмотреть вопрос выбора (построения) такой булевой функции f для данной π (или, что эквивалентно, для π, являющейся циклическим сдвигом влево на 1, см. замечание 1), чтобы функция F в (1) была биективной.

ЛИТЕРАТУРА

1. *Agibalov G. P.* Substitution block ciphers with functional keys // Прикладная дискретная математика. 2017. № 38. С. 57–65.
2. *Agibalov G. P. and Pankratova I. A.* Asymmetric cryptosystems on Boolean functions // Прикладная дискретная математика. 2018. № 40. С. 23–33.
3. *Зюбина Д. А., Токарева Н. Н.* S-блоки с максимальной компонентной алгебраической иммунностью от малого числа переменных // Прикладная дискретная математика. Приложение. 2021. № 14. С. 40–42.
4. *Carlet C.* Vectorial Boolean Functions for Cryptography. Cambridge: Cambridge University Press, 2010. 93 p.

Секция 3

МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

УДК 519.24 DOI 10.17223/2226308X/16/9

АТАКА РАЗЛИЧЕНИЯ НА ЧЕТЫРЕ РАУНДА ШИФРА ЛЮБИ — РАКОФФ ПО РАЗНОСТЯМ ДВУБЛОЧНЫХ ТЕКСТОВ[1]

О. В. Денисов

Установлено, что шифр Люби — Ракофф является марковским, найдены его матрицы переходных вероятностей разностей за 1, 2 и 4 раунда. На основе статистики логарифма отношения правдоподобий построена последовательная атака различения на четыре раунда шифра по разностям заданного типа независимых двублочных текстов. Получены оценки среднего числа используемых текстов, проведены эксперименты на шифрах с длинами блоков от 12 до 44 бит.

Ключевые слова: *марковский шифр, сеть Фейстеля, шифр Люби — Ракофф, дивергенция Кульбака, последовательная разностная атака.*

Рассмотрим R-раундовую сеть Фейстеля с алфавитом полублоков $(\mathbb{Z}_2^m, \oplus)$, $\oplus$ — покоординатное сложение по модулю 2, и раундовыми преобразованиями

$$(x^{r-1}, x^r) \to (x^r, x^{r-1} \oplus f^r(x^r)), \quad 1 \leqslant r \leqslant R. \tag{1}$$

В 1988 г. американские криптографы M. Luby и C. Rackoff ввели [1] вероятностную модель сети (1), в которой раундовые функции усложнения выбираются независимо равновероятно из множества всех двоичных вектор-функций от m переменных:

$$f^1, \ldots, f^R \sim U(\mathscr{F}_m), \ \mathscr{F}_m = \{f \mid f : \mathbb{Z}_2^m \to \mathbb{Z}_2^m\}. \tag{2}$$

Для такой случайной сети при $R = 3, 4$ в разных моделях запросов получен ряд нижних оценок стойкости к *атакам различения*, т. е. алгоритмам d проверки гипотез о случайной подстановке F на множестве $\mathbb{X} = \mathbb{Z}_2^{2m}$:

$H_1 : F$ выбрана равновероятно из множества всех подстановок,

$H_2 : F$ получена в R-раундовой модели Люби — Ракофф.

Запросами алгоритма (различителя) d называются такие последовательно подаваемые аргументы $x_1, \ldots, x_q \in \mathbb{X}$, что ему становятся доступны значения $y_t = F(x_t)$, $1 \leqslant t \leqslant q$, т. е. набор $Y = (y_1, \ldots, y_q)$. *Преимуществом* различителя d называется

$$\mathrm{Adv}(d) = \big|\,\mathsf{P}_1[d(Y) = 1] - \mathsf{P}_2[d(Y) = 1]\big| = |1 - \alpha_1(d) - \alpha_2(d)|.$$

Здесь и далее через P_i и $\alpha_i(d)$ обозначаем соответственно вероятностное распределение и вероятности ошибок критерия d при гипотезе H_i, $i = 1, 2$.

Один из главных результатов в этой области следующий [2, с. 50]: при $R = 4$ для любого различителя d в модели запросов “adaptive Chosen-plaintext and Ciphertext

Attack", когда d может формировать запрос x_t с учётом значений $y_1, \ldots, y_{t-1}$, $2 \leqslant t \leqslant q$, выполнено $\mathrm{Adv}^{\mathrm{CCA}}(d) \leqslant q^2/2^m$. Отсюда следует, что если $q(M) = o(\sqrt{M})$ при $M = 2^m \to \infty$, то гипотезы асимптотически неразличимы, т. е. сумма вероятностей ошибок любого критерия стремится к 1.

Рассмотрим модель наблюдений последовательности подстановок $F_1, F_2, \ldots$, полученных независимо друг от друга одним фиксированным способом из двух, но запросов всего два (двублочный текст), и эти пары входных блоков выбираются из $\mathbb{X}$ с ограничением на их разность. При таких условиях наблюдений построим критерий Вальда проверки гипотез по парам входная/выходная разность

$$\Delta X_t = X_{1t} \oplus X_{2t},\ \Delta Y_t = Y_{1t} \oplus Y_{2t},\ t \geqslant 1,$$

с вероятностями ошибок, близкими к заданным значениям α, β.

Далее $\mathbf{e}_j$ ($e_j^{\downarrow}$) обозначает j-й вектор-строку (столбец) стандартного базиса, $j \geqslant 0$. Через $\mathbf{a}$ и $a^{\downarrow}$ обозначаем вектор-строку и вектор-столбец (размерности, определяемой контекстом), все компоненты которых равны константе a, $a \in \{0, 1\}$.

Теорема 1. Если в модели Люби — Ракофф (1) и (2) начальная пара блоков (X^0, X^{0*}) выбирается независимо от $f^1, \ldots, f^R$, то раундовые разности $\Delta X^r = X^r \oplus \oplus X^{r*}$, $0 \leqslant r \leqslant R$, образуют однородную цепь Маркова с матрицей переходных вероятностей разностей (МПВР) за 1 шаг, равной

$$\mathbb{P}(M) = \frac{1}{M} \begin{pmatrix} P_1 & Q_2 & \ldots & Q_M \\ P_2 & Q_2 & \ldots & Q_M \\ \vdots & \vdots & \ddots & \vdots \\ P_M & Q_2 & \ldots & Q_M \end{pmatrix},$$

где $P_j = M e_1^{\downarrow} \mathbf{e}_j$, $Q_k = e_k^{\downarrow} \mathbf{1}$ — клетки ранга 1 размера M для всех $1 \leqslant j \leqslant M$, $2 \leqslant k \leqslant M$, всего $M + (M-1) = 2M - 1$ видов клеток.

Например, $\mathbb{P}(2) = \dfrac{1}{2} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$.

Из теоремы о блочном умножении [3, с. 21] следует, что все степени МПВР сохраняют такое разбиение на M^2 клеток размера M. С помощью этой теоремы описаны все клетки целочисленной матрицы

$$B(M) = (M\mathbb{P}(M))^2 = \begin{pmatrix} B_{11} & B_{12} & \ldots & B_{12} \\ B_{21} & B_{22} & \ldots & B_{12} \\ \vdots & \vdots & \ddots & \vdots \\ B_{21} & B_{12} & \ldots & B_{22} \end{pmatrix},$$

где диагональные клетки $B_{11} = M\,\mathrm{diag}(\mathbf{v}_M)$, $\mathbf{v}_M = (M, 1, \ldots, 1)$ и $B_{21} = M\,\mathrm{diag}(\mathbf{v}_0)$, $\mathbf{v}_0 = (0, 1, \ldots, 1)$ составляют левую горизонтальную полосу, а остальные клетки $B_{12} = v_0^{\downarrow}\mathbf{1}$, $B_{22} = v_M^{\downarrow}\mathbf{1}$ одноранговые.

Аналогично, как суммы произведений клеток матрицы B, найдены пять видов клеток C_{ij} матрицы $C(M) = (M\mathbb{P}(M))^4 = B(M)^2$. Установлено, что строки матрицы, полученной из $C(M)$ удалением верхней строки и левого столбца, могут быть разбиты

на два типа по своему составу, т. е. частотам значений элементов строки:
если

$$(\Delta x_0 = 0 \wedge \Delta x_1 \neq 0) \vee (\Delta x_0 \neq 0 \wedge \Delta x_1 \neq 0) \tag{3}$$

(это строки верхней полосы, а также все строки, кроме верхних, из остальных горизонтальных полос), то имеем

Значение	$2M^2 - M$	$M^2 - M$	$M^2 + 1$
Частота	1	$M - 2$	$M^2 - M$

;

если

$$\Delta x_0 \neq 0,\ \Delta x_1 = 0, \tag{4}$$

(это верхние строки всех горизонтальных полос, кроме верхней), то имеем

Значение	M^2	$2M^2 - M$	$M^2 - M$
Частота	$M - 1$	M	$M^2 - 2M$

.

Отсюда, в частности, следует, что матрицы переходных вероятностей ненулевых разностей за четыре раунда положительны и такие шифры дважды транзитивны.

При гипотезе H_1 выходные разности имеют равномерное распределение на множестве $\mathbb{X}' = \mathbb{Z}_2^{2m} \setminus \{\mathbf{0}\}$ ненулевых блоков. Обозначим эти вероятности через $p_1(\Delta y) = (M^2 - 1)^{-1}$. Через $p_2(\Delta y)$ обозначим вероятности распределения 4-раундовой выходной разности при фиксированной входной разности Δx и рассмотрим дивергенции Кульбака между этими распределениями: $K(i : 3 - i) = \sum\limits_{\Delta y \in \mathbb{X}'} p_i(\Delta y) \ln \dfrac{p_i(\Delta y)}{p_{3-i}(\Delta y)}$, $i = 1, 2$. Вероятностный смысл этих величин следующий: математическое ожидание слагаемого статистики логарифма отношения правдоподобий равно $K(2 : 1) > 0$ при гипотезе H_2 и $-K(1 : 2) < 0$ при гипотезе H_1.

Теорема 2. При входной разности $\Delta x = (\Delta x_0, \Delta x_1) \in \mathbb{X}'$ и $M \to \infty$ для дивергенций Кульбака между распределениями p_2 и p_1 выполнено:

1) при условии (3)

$$K(2:1) = \frac{2M^2 - M}{M^4} \ln \frac{(2M^2 - M)(M^2 - 1)}{M^4} + (M - 2)\frac{M^2 - M}{M^4} \ln \frac{(M^2 - M)(M^2 - 1)}{M^4} +$$
$$+ (M^2 - M)\frac{M^2 + 1}{M^4} \ln \frac{(M^2 + 1)(M^2 - 1)}{M^4} \sim (2 \ln 2 - 1) M^{-2},$$
$$K(1:2) = \frac{1}{M^2 - 1} \ln \frac{M^4}{(2M^2 - M)(M^2 - 1)} + \frac{M - 2}{M^2 - 1} \ln \frac{M^4}{(M^2 - M)(M^2 - 1)} +$$
$$+ \frac{M^2 - M}{M^2 - 1} \ln \frac{M^4}{(M^4 - 1)} \sim (1 - \ln 2) M^{-2};$$

2) при условии (4)

$$K(2:1) = (M - 1)\frac{1}{M^2} \ln \frac{M^2 - 1}{M^2} + M \frac{2M - 1}{M^3} \ln \frac{(2M - 1)(M^2 - 1)}{M^3} +$$
$$+ (M^2 - M)\frac{M - 1}{M^3} \ln \frac{(M - 1)(M^2 - 1)}{M^3} \sim (2 \ln 2 - 1) M^{-1},$$
$$K(1:2) = (M - 1)\frac{1}{M^2 - 1} \ln \frac{M^2}{M^2 - 1} + M \frac{1}{M^2 - 1} \ln \frac{M^3}{(2M - 1)(M^2 - 1)} +$$
$$+ (M^2 - M)\frac{1}{M^2 - 1} \ln \frac{M^3}{(M - 1)(M^2 - 1)} \sim (1 - \ln 2) M^{-1}.$$

Оба асимптотических значения дивергенции в п. 1 ровно в M раз больше, чем соответствующие значения в п. 2. Поэтому атаки различения будут значительно эффективнее при выборе типа (4) входных разностей.

Проведём статистические эксперименты (атаки различения на четыре раунда шифра Люби — Ракофф) по входным разностям (4) второго типа, используя критерии Вальда (как более экономичные в смысле объёма материала) для проверки следующих гипотез о последовательности случайных независимых одинаково распределённых подстановок $F^1, F^2, \ldots$ на множестве $\mathbb{Z}_2^{2m}$:

$$\begin{aligned} &H_1 : F_t \text{ выбраны равновероятно из множества всех подстановок,} \\ &H_2 : F_t \text{ — подстановки 4-раундовой модели Люби — Ракофф, } t \geqslant 1. \end{aligned} \tag{5}$$

Для корректной работы критерия требуется независимость двублочных текстов (запросов с разностью второго типа между блоками), выбираемых для каждой подстановки. Модель независимых двублочных текстов введена в работе автора [4]. В этой модели проведены атаки различения, основанные на статистике отношения правдоподобий при фиксированной входной разности, на марковские модели шифрсистем SmallPresent с длиной блока до 28 бит.

Критерий обеспечивает вероятности ошибок первого и второго родов, близкие к $\alpha = \beta = 0{,}1$, при выборе таких границ A, B, что [5, с. 150]

$$-\ln A = \ln B = \ln \frac{1-\alpha}{\alpha} = \ln 9 = 2{,}197.$$

Среднее число испытаний (пар открытый/шифрованный двублочный текст) до принятия решения при гипотезе H_i близко к $(1-2\alpha)\ln\dfrac{1-\alpha}{\alpha}\big(K(i:3-i)\big)^{-1}$ [5, с. 152], что при больших M, согласно теореме 2, близко к $T_i(M)$, $i \in \{1,2\}$, где

$$T_1(M) = \frac{0{,}8\ln 9}{1-\ln 2}M = 5{,}72M,\ T_2(M) = \frac{0{,}8\ln 9}{2\ln 2 - 1}M = 4{,}55M.$$

В таблице представлены результаты 20 атак различения гипотез (5) с расчётными значениями ошибок $\alpha = \beta = 0{,}1$ при длине полублока m бит: оценки T_i и эмпирические значения $\hat{\mathsf{E}}_i\tau$ числа τ использованных критерием Вальда двублочных текстов при гипотезе H_i, эмпирические вероятности ошибок $\hat{\alpha}_i$. Эмпирические вероятности ошибок, а также средние длины продолжительности экспериментов показывают хорошее соответствие теории и практики.

m	T_1	$\hat{\mathsf{E}}_1\tau$	$\hat{\alpha}_1$	T_2	$\hat{\mathsf{E}}_2\tau$	$\hat{\alpha}_2$
6	366,6	436,5	0	291,2	348,8	0,05
7	733,2	631,6	0	582,4	696,7	0,1
8	1,4 E3	1,4 E3	0,05	1,1 E3	1,2 E3	0,1
9	2,9 E3	3,7 E3	0,1	2,3 E3	2,7 E3	0,05
10	5,8 E3	5,9 E3	0	4,6 E3	7,5 E3	0,1
11	1,1 E4	1,0 E4	0	9,3 E3	7,6 E3	0,2
12	2,34 E4	2,9 E5	0,1	1,8 E4	1,6 E4	0,05
13	4,6 E4	5,8 E4	0	3,7 E4	4,9 E4	0,05
14	9,3 E4	9,7 E4	0,05	7,4 E4	1,1 E5	0,2
15	1,8 E5	2,1 E5	0,2	1,4 E5	1,6 E5	0,1
16	3,7 E5	4,2 E5	0	2,9 E5	3,5 E5	0,1
17	7,5 E5	7,2 E5	0,05	5,9 E5	6,3 E5	0,1
18	1,5 E6	1,3 E6	0,1	1,1 E6	1,3 E6	0,05
19	3,0 E6	3,4 E6	0,05	2,3 E6	1,7 E6	0,05
20	6,0 E6	6,3 E6	0,1	4,7 E6	4,4 E6	0
21	1,2 E7	1,0 E7	0,05	9,5 E6	1,1 E7	0,1
22	2,4 E7	2,6 E7	0,15	1,9 E7	2,1 E7	0,1

ЛИТЕРАТУРА

1. *Luby M. and Rackoff C.* How to construct pseudorandom permutations from pseudorandom functions // SIAM J. Comput. 1988. V. 17. P. 373–386.
2. *Nachef V., Patarin J., and Volte E.* Feistel Ciphers: Security Proofs and Cryptanalysis. Springer, 2017.
3. *Ланкастер П.* Теория матриц. М.: Наука, 1978.
4. *Денисов О. В.* Атаки различения на блочные шифрсистемы по разностям двублочных текстов // Прикладная дискретная математика. 2020. № 48. С. 43–62.
5. *Ивченко Г. И., Медведев Ю. И.* Математическая статистика. М.: Высшая школа, 1984.

УДК 004.8 DOI 10.17223/2226308X/16/10

ОБРАЩЕНИЕ 29-ШАГОВОЙ ФУНКЦИИ СЖАТИЯ MD5 ПРИ ПОМОЩИ АЛГОРИТМОВ РЕШЕНИЯ ПРОБЛЕМЫ БУЛЕВОЙ ВЫПОЛНИМОСТИ[1]

О. С. Заикин

Криптографическая хеш-функция MD5 предложена в 1992 г. Ключевым компонентом MD5 является 64-шаговая функция сжатия. До сих пор не представляется возможным обратить функцию сжатия MD5 за реальное время, поэтому зачастую в данном контексте анализируются версии с сокращённым количеством шагов. В 2007 г. с помощью алгоритмов решения проблемы булевой выполнимости (SAT) была обращена 26-шаговая функции сжатия MD5. В 2012 г. с помощью SAT были обращены 27- и 28-шаговые версии. В настоящем исследовании предлагается подход к формированию 32 промежуточных задач обращения между парой последовательных шагов функции сжатия MD5. С помощью этого подхода построены промежуточные задачи обращения между 28 и 29 шагами. Несколько простых задач использованы для параметризации современного SAT-решателя, в результате чего впервые обращена 29-шаговая функция сжатия MD5.

Ключевые слова: *криптографическая хеш-функция, MD5, алгебраический криптоанализ, логический криптоанализ, проблема булевой выполнимости.*

[1]Работа выполнена за счёт субсидии Минобрнауки России в рамках проекта № 121041300065-9.

Введение

Принимая на вход сообщение произвольной длины, бесключевые криптографические хеш-функции генерируют хеш фиксированной длины [1]. Такие функции имеют несколько обязательных свойств, среди которых легкость вычисления. Из потенциальных свойств отметим устойчивость к поиску коллизий и к поиску прообраза. Последнее свойство означает, что для произвольного хеша поиск его прообраза (т. е. обращение соответствующей криптографической хеш-функции) не представляется возможным осуществить за реальное время. В 1992 г. была предложена бесключевая криптографическая хеш-функция MD5, которая генерирует хеш длиной 128 бит [2]. Главным компонентом MD5 является функция сжатия, которая смешивает 512-битный блок сообщения со специальным 128-битным регистром в течение 4 раундов, каждый из которых состоит из 16 шагов. С 2005 г. MD5 не является стойкой к поиску коллизий [3]. Несмотря на это, данная функция всё ещё широко используется для хеширования паролей и проверки целостности данных. Одна из причин состоит в том, что MD5 ещё не скомпрометирована в контексте поиска прообраза.

В последние годы популярным направлением исследований является обращение неполнораундовых версий функции сжатия MD5. При этом чаще всего используется логический криптоанализ, т. е. исходная задача сводится к экземпляру проблемы булевой выполнимости, для решения которого применяется SAT-решатель [4]. Отметим, что логический криптоанализ является подвидом алгебраического криптоанализа [5]. В 2007 г. была впервые обращена 26-шаговая функции сжатия MD5 [6]. При этом использовался SAT-решатель, основанный на алгоритме Conflict-Driven Clause Learning (CDCL) [7]. Здесь и далее имеются в виду первые шаги функции сжатия, т. е. в [6] представлен результат обращения сокращённой функции сжатия MD5, состоящей из 16 шагов первого раунда и 10 шагов второго раунда. В 2012 г. были впервые обращены 27- и 28-шаговые версии функции сжатия MD5 [8]. Для этого также был применён CDCL-решатель. С тех пор никому не удалось обратить 29-шаговую функцию сжатия MD5. Настоящее исследование нацелено на решение именно этой задачи.

В 2022 г. с помощью метода Cube-and-Conquer были впервые обращены 40-, 41-, 42- и 43-шаговые версии функции сжатия криптографической хеш-функции MD4 [9]. Согласно этому методу, предназначенному для решения трудных экземпляров SAT, сначала с помощью lookahead-решателя задача разбивается на более простые подзадачи, а затем на них запускается CDCL-решатель [10]. Кроме Cube-and-Conquer, в [9] использованы условия Доббертина [11], которые ранее были успешно применены для обращения 39-шаговой функции сжатия MD4 [6, 12]. К сожалению, согласно предварительным исследованиям, условия Доббертина не эффективны в контексте обращения неполнораундовой функции сжатия MD5. Исходя из этого, для решения данной задачи требуется другой подход.

1. Промежуточные задачи обращения функции сжатия MD5

Рассмотрим пару последовательных шагов функции сжатия MD5 с номерами $i-1$ и i, $2 \leqslant i \leqslant 64$. Идея состоит в построении промежуточных задач обращения путём постепенного ослабления i-го шага, при том что первые $i-1$ шагов работают как обычно. Шаг с номером i может быть представлен в виде следующего псевдокода:

$$a \leftarrow b + ((a + Func(b, c, d) + M[j] + K[i]) <<< s).$$

Здесь a, b, c и d — значения четырёх специальных регистров; $Func$ — нелинейная функция; $M[j]$ — 32-битное слово, которое является j-й частью 512-битного сообщения M;

$K[i]$ — известная константа; $<<<$ — циклический сдвиг влево. Отметим, что на каждом из четырёх раундов для смешивания используется собственная нелинейная функция. В ходе предварительных экспериментов для разных значений i были сделаны попытки ослабить шаг путём отбрасывания одной или нескольких операций. Оказалось, что даже если отбросить использование нелинейной функции, для CDCL-решателей задача не становится существенно проще. При этом если отбросить слагаемое $M[j]$, то для CDCL-решателя такая задача становится сопоставимой по сложности с задачей обращения $i-1$ шагов.

Предлагается формировать 32 промежуточные задачи обращения между шагами $i-1$ и i постепенной заменой на i-м шаге $M[j]$ на слово, частично состоящее из нулевых битов. В первой промежуточной задаче $M[j]$ заменяется на 32 нулевых бита. Следует подчеркнуть, что здесь и далее имеется в виду замена $M[j]$ только на шаге i, а не на всех шагах функции сжатия. Во второй промежуточной задаче замена производится на слово, в котором 31 старший бит равен нулю, а оставшийся бит — соответствующему (неизвестному) младшему биту в $M[j]$. Наконец, в 32-й промежуточной задаче на i-м шаге $M[j]$ заменяется на слово, в котором старший бит равен 0, а оставшиеся биты равны соответствующим битам $M[j]$. Таким образом, псевдокод модифицированного i-го шага для p-й промежуточной задачи обращения выглядит следующим образом:

$$a \leftarrow b + ((a + Func(b,c,d) + ((M[j] << (32-p+1)) >> (32-p+1)) + K[i]) <<< s).$$

Предполагается, что для современных CDCL-решателей первая промежуточная задача обращения сравнима по сложности с задачей обращения $i-1$ шагов функции сжатия MD5, а 32-я задача сравнима с обращением i шагов.

2. Обращение 29-шаговой функции сжатия MD5

На первом этапе были построены две КНФ, кодирующие 28 шагов функции сжатия MD5. Для их построения использованы программные комплексы Transalg [13] и CBMC [14]. Было сгенерировано несколько случайных 128-битных слов и путём подстановки в каждую из двух КНФ 128 соответствующих однолитеральных дизъюнктов были получены КНФ, кодирующие задачи обращения для конкретных хешей. Во всех экспериментах использовался CDCL-решатель Kissat [15] версии 3.0, так как он показал отличные результаты на соревнованиях SAT-решателей в 2020–2022 гг. Оказалось, что на каждой КНФ Kissat находит решение на персональном компьютере примерно за 1–3 часа, но при этом среднее время решения было ниже на КНФ, сгенерированных с помощью CBMC. Отметим, что на задачах обращения неполнораундовой функции сжатия MD4 ситуация обратная — там Transalg лучше, чем CBMC [9].

На втором и последующих этапах рассматривались задачи обращения только единичного хеша, т. е. 128-битного слова, состоящего из единиц. Были сделаны 32 КНФ, кодирующие промежуточные задачи обращения между 28 и 29 шагами. Kissat был запущен на персональном компьютере с лимитом времени 48 часов на каждой из промежуточных задач. Только первые восемь из них были решены. После этого была проведена параметризация Kissat на первых четырёх промежуточных задачах. В результате были найдены новые значения параметров Kissat, которые позволяют быстрее решать задачи из этого класса. Параметризованная версия Kissat решила 13 промежуточных задач с тем же лимитом времени.

На третьем этапе для решения остальных промежуточных задач применялся метод Cube-and-Conquer [10], при этом на второй его стадии работал тот же параметризованный Kissat, который использовался на персональном компьютере. Эксперименты

проведены на вычислительном кластере «Академик В. М. Матросов» [16]. Каждой задаче выделялось 7 дней на 5 узлах кластера. В результате из оставшихся 19 промежуточных задач обращения были решены 13, включая 31-ю задачу, в которой на 29-м шаге $M[13]$ заменяется на $(M[13] \ll 2) \gg 2$.

На четвертом этапе с помощью CBMC была сгенерирована КНФ, кодирующая задачу обращения 29-шаговой функции сжатия MD5. На ней в тех же условиях на кластере был запущен Cube-and-Conquer. Решение найдено не было. Напомним, что в самой сложной из решённых промежуточных задач на 29-м шаге $M[13]$ заменяется на слово, в котором два старших бита равны 00, а остальные 30 бит равны соответствующим битам $M[13]$. При этом найденное решение могло также оказаться прообразом 29-шаговой функции сжатия, но у $M[13]$ были другие значения старших битов. Были сделаны ещё 3 КНФ, в которых двум старшим битам на 29-м шаге присваиваются значения не 00, как в 31-й промежуточной задаче, а 01, 10 и 11 соответственно. На всех этих КНФ также был запущен Cube-and-Conquer на кластере, и в результате только на варианте 10 было найдено решение. На этот раз у $M[13]$ значения двух старших битов были равны 10. Таким образом, был найден прообраз единичного хеша, который сгенерирован 29-шаговой функцией сжатия MD5. Корректность данного прообраза проверена на реализации MD5 из работы [2].

ЛИТЕРАТУРА

1. *Алферов А. П., Зубов А. Ю., Кузьмин А. С., Черемушкин А. В.* Основы криптографии. 2-е изд. М.: Гелиос АРВ, 2002.
2. *Rivest R.* The MD5 message-digest algorithm. RFC 1321. 1992. `https://www.ietf.org/rfc/rfc1321.txt`.
3. *Wang X. and Yu H.* How to break MD5 and other hash functions // LNCS. 2005. V. 3494. P. 19–35.
4. *Massacci F. and Marraro L.* Logical cryptanalysis as a SAT problem // J. Automated Reasoning. 2000. V. 1. P. 165–203.
5. *Bard G.* Algebraic Cryptanalysis. N.Y.: Springer, 2009.
6. *De D., Kumarasubramanian A., and Venkatesan R* Inversion attacks on secure hash functions using SAT solvers // LNCS. 2007. V. 4501. P. 377–382.
7. *Marques-Silva J. P. and Sakallah K. A.* GRASP: a search algorithm for propositional satisfiability // IEEE Trans. Computers. 1999. V. 48 (5). P. 506–521.
8. *Legendre F., Dequen G., and Krajecki M.* Encoding hash functions as a SAT problem // Proc. ICTAI. Athens, Greece, 2012. P. 916–921.
9. *Zaikin O.* Inverting 43-step MD4 via cube-and-conquer // Proc. IJCAI-ECAI 2022. P. 1894–1900.
10. *Heule M. J. H., Kullmann O., Wieringa S., and Biere A.* Cube and Conquer: guiding CDCL SAT solvers by lookaheads // LNCS. 2012. V. 7261. P. 50–65.
11. *Dobbertin H.* The first two rounds of MD4 are not one-way // LNCS. 1998. V. 1372. P. 284–292.
12. *Грибанова И. А.* Новый алгоритм порождения ослабляющих ограничений в задаче обращения хеш-функции MD4-39 // Прикладная дискретная математика. Приложение. 2018. № 11. С. 139–141.
13. *Semenov A. A., Otpuschennikov I. V., Gribanova I. A., et al.* Translation of algorithmic descriptions of discrete functions to SAT with application to cryptanalysis problems // Logical Methods in Computer Science. 2020. V. 16. Iss. 1. P. 29:1–29:42.

14. *Clarke E., Kroening D., and Lerda F.* A tool for checking ANSI-C programs // LNCS. 2004. V. 2988. P. 168–176.
15. *Biere A. and Fleury M.* Gimsatul, IsaSAT and Kissat entering the SAT Competition 2022 // Proc. SAT Competition 2022. Solver and Benchmark Descriptions. P. 10–11.
16. Иркутский суперкомпьютерный центр СО РАН. `http://hpc.icc.ru`.

УДК 004.4 DOI 10.17223/2226308X/16/11

СРАВНИТЕЛЬНЫЙ АНАЛИЗ КАЧЕСТВА ПРЕОБРАЗОВАНИЯ ГРАФИЧЕСКОЙ ИНФОРМАЦИИ С ПОМОЩЬЮ БЛОЧНЫХ ШИФРОВ

Е. А. Ищукова, Р. Р. Борлаков

Представлены результаты практических экспериментов по влиянию различных алгоритмов шифрования (DES, AES, Магма) в режиме электронной кодовой книги (ECB) на качество преобразования графической информации в зависимости от её свойств. В ходе эксперимента проверена гипотеза, согласно которой качество шифрования зависит не только от алгоритма шифрования и его режима, но и от свойств самой преобразуемой информации. Экспериментально продемонстрировано, что на качество преобразования информации влияют такие параметры, как количество цветов в цветовой палитре, количество крупных и мелких объектов на рисунке, количество пикселей, наличие или отсутствие фона, а также другие.

Ключевые слова: *шифрование, блочный шифр, режим шифрования, электронная кодовая книга, графическая информация.*

Введение

Известно, что режим шифрования ECB, как правило, используется не для шифрования файлов, а как составная часть других режимов шифрования, которые обеспечивают связь блоков между собой и равномерное перемешивание информации. Тем не менее исследования надёжности шифра всегда начинаются с изучения его работы в режиме ECB. Отдельно обсуждается вопрос о том, как влияет режим шифрования на преобразование разных типов данных. Так, для алгоритма AES было показано, что шифрование графической информации в режиме ECB приводит к сохранению структуры рисунка так, что можно определить содержимое этого рисунка [1, 2]. Рядом исследователей предпринимаются попытки создать новые алгоритмы шифрования, которые бы позволили качественно преобразовывать любую графическую информацию [3, 4]. В настоящей работе мы провели сравнительный анализ влияния различных алгоритмов шифрования, используемых в режиме ECB, на качество преобразования графических изображений в зависимости от свойств самого изображения. В качестве алгоритмов шифрования для исследования были выбраны бывший стандарт шифрования США DES, действующий стандарт шифрования США AES и отечественный алгоритм шифрования Магма.

1. Методы исследования

Алгоритм DES [5] активно использовался до начала 2000 годов, в настоящее время является устаревшим, хотя до сих пор в реализациях некоторых протоколов применяются его модифицированные версии. Например, в стандартах ANSI X9.17 и ISO 8732 используется тройной DES с двумя ключами (Tripple-DES). Алгоритм DES представляет собой симметричный блочный шифр, построенный по схеме Фейстеля. На его вход

поступает 64-битный текст, который обрабатывается под воздействием 56-битного секретного ключа шифрования.

Стандарт AES был принят на смену стандарту DES. В его основе лежит алгоритм шифрования Rijndael, который представляет собой симметричный блочный шифр, построенный по принципу подстановочно-перестановочной сети. Шифр обрабатывает блоки размерностью 128 бит под воздействием секретного ключа, который может принимать разные размеры: 128, 192 или 256 бит [5]. В зависимости от размера ключа изменяется количество раундов шифра. В настоящей работе все эксперименты проводились с использованием версии шифра AES-128.

Третий рассмотренный алгоритм — это отечественный алгоритм шифрования Магма, который является одним из двух шифров, составляющих основу стандарта ГОСТ Р 34.12-2015, и представляет собой симметричный блочный шифр, построенный по схеме Фейстеля. Блок данных у него составляет 64 бита, а секретный ключ — 256 бит [5]. Алгоритм Магма является основой стандарта ГОСТ 28147-89. Основное отличием Магмы от ГОСТ 28147-89 — в ГОСТ 28147-89 блоки замены не зафиксированы и могут использоваться в любых комбинациях; для алгоритма Магма блоки зафиксированы однозначно [5].

При шифровании больших объёмов данных с помощью симметричных блочных шифров рекомендовано использовать различные режимы, направленные на сцепление блоков и перемешивание информации таким образом, чтобы шифрование каждого последующего блока зависело от того, как был зашифрован предыдущий блок. К таким режимам относятся: режим сцепления блоков (CBC), различные режимы гаммирования, режим обратной связи по шифртексту (CFB), режим обратной связи по выходу (OFB) и другие. Основой для всех режимов является режим электронной кодовой книги (ECB), при котором все блоки шифруются независимо друг от друга. Для отечественных алгоритмов режимы шифрования определяются в соответствии с ГОСТ Р34.12-2015.

Целью настоящей работы является исследование качества преобразования графической информации в зависимости от исходных характеристик изображения и от используемых алгоритмов шифрования. Для работы над изображениями был выбран формат .ppm, так как он отличается удобством, простотой и хорошо подходит для экспериментальных исследований. Перед началом преобразования графического изображения необходимо произвести операцию отделения его заголовка от тела (рис. 1), так как при шифровании заголовка его сигнатура будет нарушена и мы больше не сможем просмотреть файл в формате изображения. Поэтому шифруется только тело файла, после чего оно обратно соединяется с заголовком. Здесь следует отметить, что стандартный заголовок для формата .ppm составляет 16 байт, то есть 128 бит. Это укладывается в два блока данных для тех алгоритмов, у которых идет обработка 64-битных блоков (DES и Магма), и в один блок для алгоритма AES, у которого обрабатываемый блок кратен 128 битам. В связи с этим нет большой разницы: сначала убрать заголовок, зашифровать файл и затем соединить его со стандартным заголовком (чтобы графический редактор смог отобразить информацию) или сначала зашифровать файл целиком, а затем заменить зашифрованный заголовок на стандартный нешифрованный.

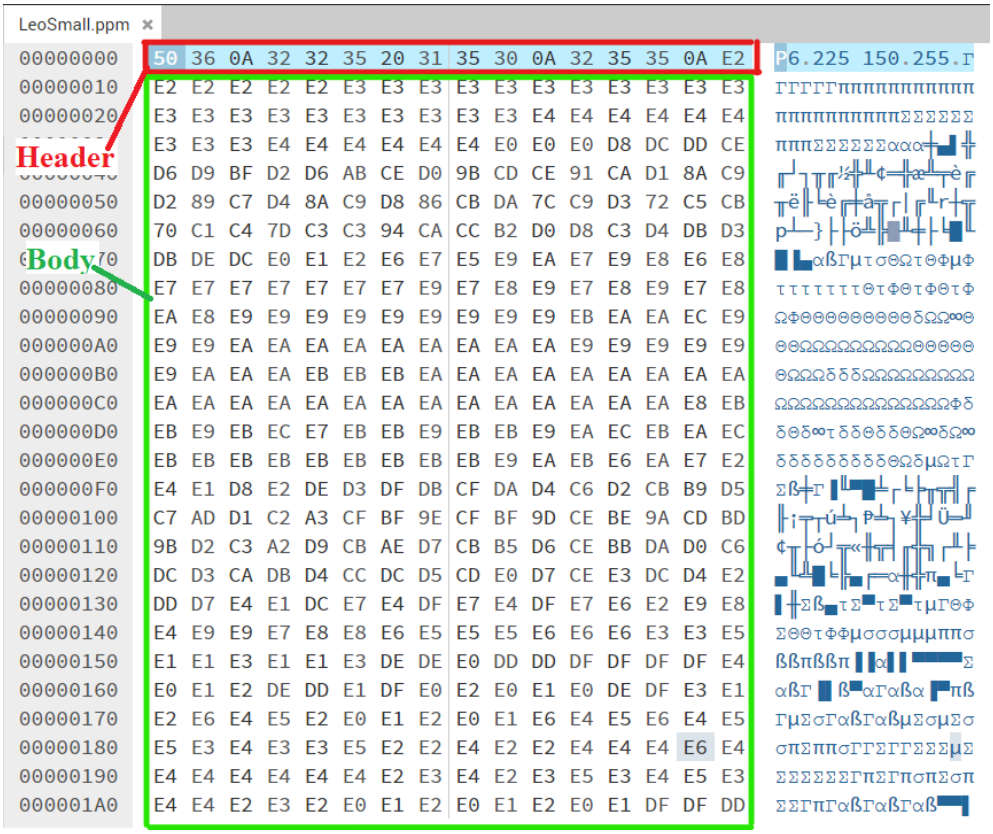

Рис. 1. Header и Body формата .ppm

2. Результаты экспериментов

Эксперимент 1. Сравнение результатов шифрования для фотографий и рисунков. Для данного эксперимента были выбраны две картинки с крупным изображением, первая — фотоизображение, вторая — нарисованная. У обеих картинок задний фон заменён белым цветом (рис. 2).

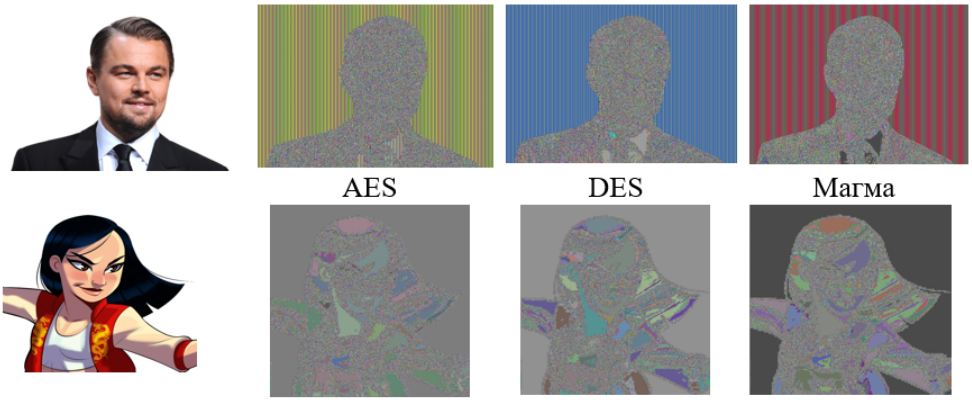

Рис. 2. Эксперимент 1

На рис. 2 можно отчетливо увидеть очертания зашифрованных изображений, соответствующие оригинальному. Можно заметить, что очертания зашифрованного изображения для рисованной картинки более явные по сравнению с фотоизображением.

Связано это с тем, что первое изображение снято на камеру в режиме реального времени, где на конечный цвет пикселя влияют такие факторы, как свет, тень и т. д., а второе изображение нарисовано с использованием инструментов векторной графики, где сложно сымитировать разнородность пикселей. Получается, что чем больше неодинаковых пикселей содержит изображение, тем лучше происходит преобразование.

Эксперимент 2. Влияние фона на качество преобразования. Для данного эксперимента использована одна и та же фотография с отсутствующим задним фоном и с неизменённым фоном (рис. 3).

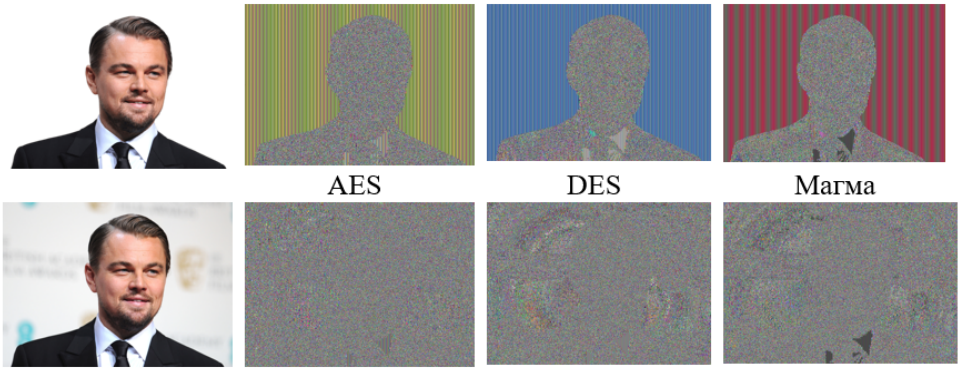

Рис. 3. Эксперимент 2

Из рис. 3 видно, что наличие фона обеспечивает более равномерное распределение зашифрованных пикселей. Фон практически сливается с самим изображением, и оно становится неразличимым. Видно, что наилучшее перемешивание достигается для алгоритма AES.

Эксперимент 3. Влияние разрешения исходной картинки на качество преобразования графической информации. Были использованы изображения из эксперимента 2 с фоном и без фона в двух разных разрешениях: 1200×798 и 300×200. Результаты представлены на рис. 4. Видно, что на восприятие контуров графического изображения расширение оказало влияние только для изображения с фоном, которое и без того обеспечивает достаточно равномерное перемешивание информации. Для изображения без фона изменение разрешения практически не оказало влияния на результат шифрования.

Эксперимент 4. Влияние размеров и количества деталей на рисунке на качество преобразования. Использованы два изображения с множеством предметов. Первое изображение нарисовано графически, второе — фотография. Результаты эксперимента представлены на рис. 5. Видно, что большое количество мелких деталей на исходном изображении приводит к хорошему преобразованию данных даже в режиме ECB. При этом для графического изображения преобразование выполняется качественнее по сравнению с фотографией.

Эксперимент 5. Качество преобразования текстовой информации, представленной в виде изображения (скан-копии). На рис. 6 видно, что для всех алгоритмов шифрования можно проследить структуру текста. Даже если нельзя распознать текст, то можно выделить его отличительные элементы (шапку, заголовок, подпись и т. д.).

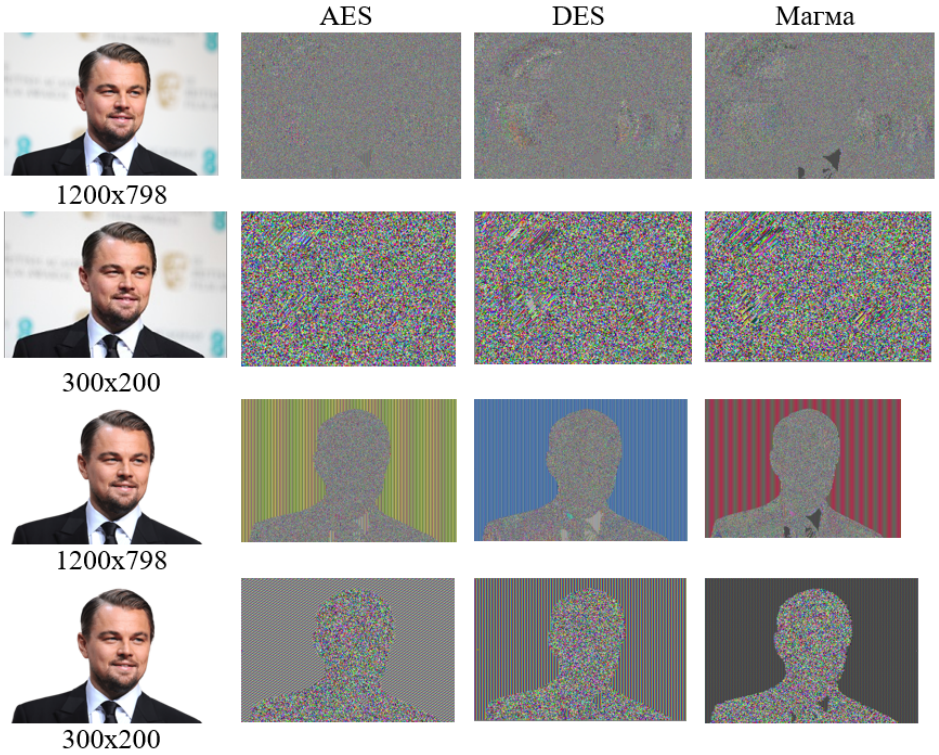

Рис. 4. Эксперимент 3

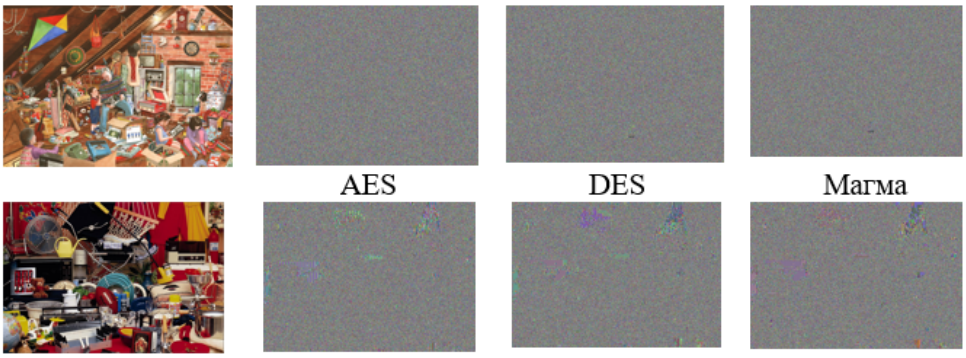

Рис. 5. Эксперимент 4

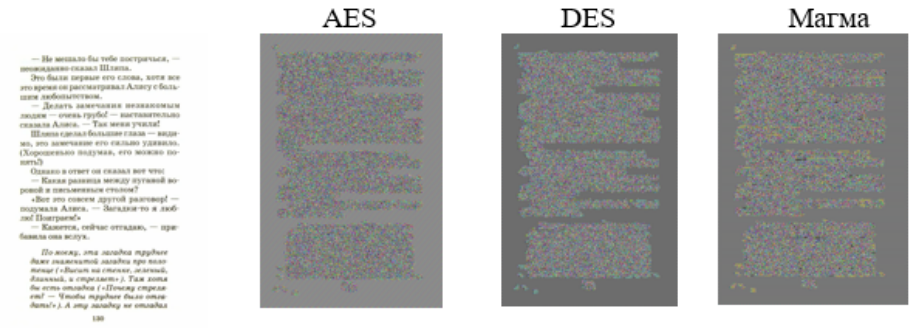

Рис. 6. Эксперимент 5

Эксперимент 6. Качество преобразования текстовой информации, если она содержит стегоконтейнер. Для данного эксперимента было произведено преобразование простой графической информации, после чего в изображение был встроен текст с использованием стегоконтейнера (подойдёт любой онлайн-сервис, например Steganography Online). Результаты шифрования представлены на рис. 7. Видно, что наличие стегоконтейнера внутри изображения сильно искажает результат преобразования графической информации. По результату шифрования можно предположить, встроено ли стегоизображение в картинку с простым рисунком.

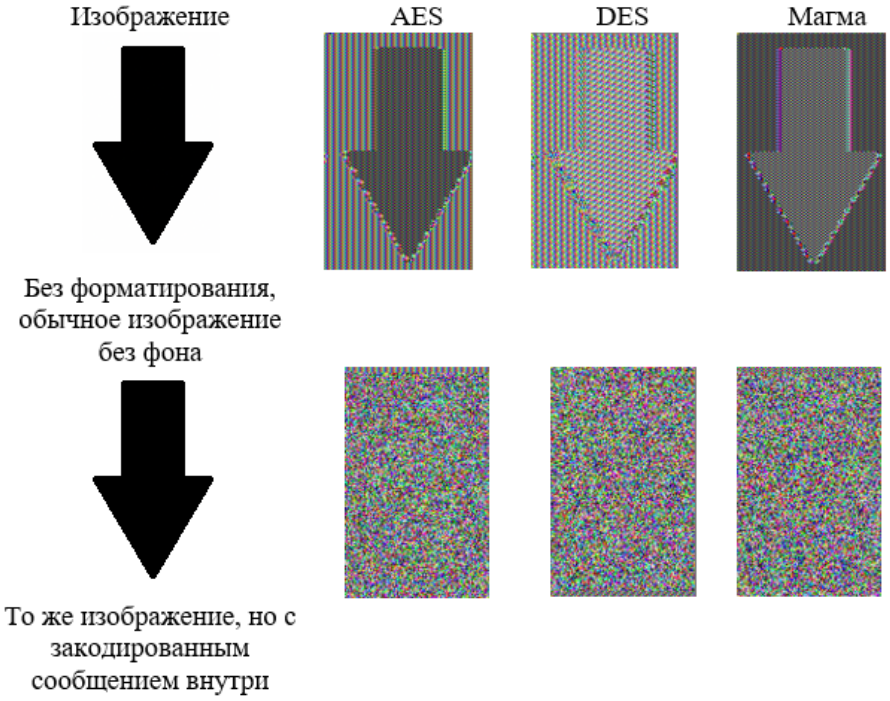

Рис. 7. Эксперимент 6

Эксперимент 7. Качество преобразования информации в зависимости от используемого режима шифрования. Для данного эксперимента одно и то же изображение было зашифровано с использованием различных алгоритмов шифрования как в режиме ECB, так и в режиме CBC (рис. 8). Видно, что независимо от алгоритма шифрования применение режима CBC решает проблему неравномерного перемешивания и обеспечивает хорошее шифрование графической информации.

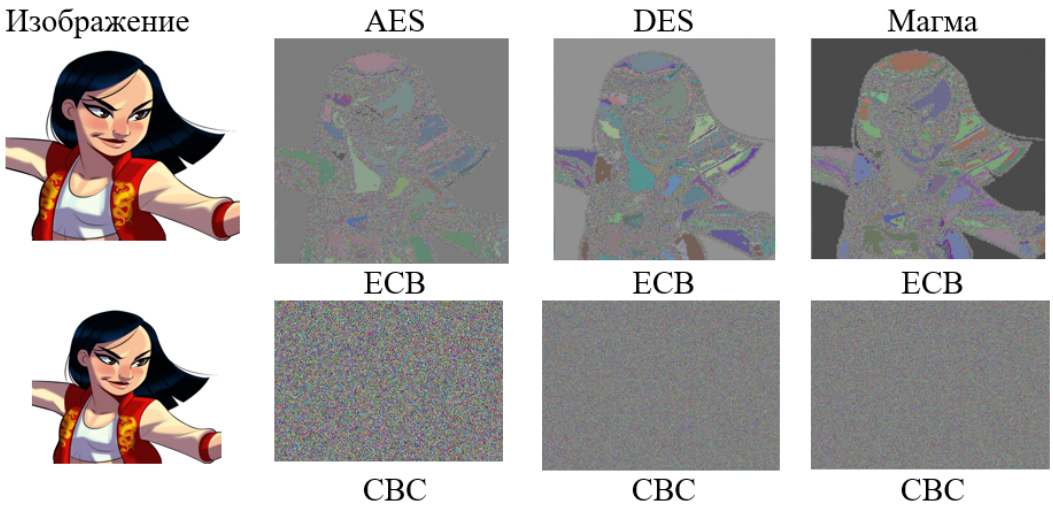

Рис. 8. Эксперимент 7

Заключение

В работе проведён ряд экспериментов по шифрованию графической информации с использованием различных алгоритмов шифрования (DES, AES, Магма) преимущественно в режиме электронной кодовой книги (ECB). Экспериментально показано, что на качество преобразования информации могут оказывать влияние следующие факторы: размер изображения, количество крупных и мелких деталей, задний фон. При этом разрешение изображения практически не оказывает влияния на качество преобразования. Получены интересные экспериментальные данные в части шифрования изображений со встроенными стегоконтейнерами, которые могут быть использованы для детектирования стегоконтейнеров. Предполагается, что данная работа будет продолжена в экспериментальной части. Планируется дополнить эксперименты для всех представленных изображений с использованием алгоритма шифрования Кузнечик, который является частью отечественного стандарта шифрования данных ГОСТ Р 34.12-2015. Также планируется провести ряд экспериментов по использованию монохромных изображений и изображений с разным объёмом встроенной стеганографической информации.

ЛИТЕРАТУРА

1. *Huang C.-W„ Tu Y.-H., Yeh H.-C., et al.* Image observation on the modified ECB operations in Advanced Encryption Standard // Intern. Conf. i-Society. London, UK, 2011. P. 264–269.
2. *Graham R.* ECB Penguin Demonstration. `https://github.com/robertdavidgraham/ecb-penguin`.
3. *He P., Sun K., and Zhu C.* A novel image encryption algorithm based on the delayed maps and permutation-confusion-diffusion architecture // Security Communication Networks. 2021. V. 2021. Article ID 6679288. `https://doi.org/10.1155/2021/6679288`.
4. *Naseer Y., Shah T., Shah D., and Hussain S.* A novel algorithm of constructing highly nonlinear S-p-boxes // Cryptography. 2019. V. 3. No. 1. `https://doi.org/10.3390/cryptography3010006`.
5. *Бабенко Л. К., Ищукова Е. А.* Криптографическая защита информации: симметричное шифрование. Учеб. пособие М.: Изд-во Юрайт, 2019. 220 с.

УДК 519.7 DOI 10.17223/2226308X/16/12

О КОЛИЧЕСТВЕ НЕВОЗМОЖНЫХ РАЗНОСТЕЙ ПО МОДУЛЮ 2^n ДЛЯ ARX-ПРЕОБРАЗОВАНИЯ[1]

Н. А. Коломеец

Исследуется разностная характеристика по модулю 2^n для преобразования $(x \oplus y) \lll r$, где $x, y \in \mathbb{Z}_2^n$ и $1 \leqslant r < n$. Она представляет интерес в контексте разностных атак на шифры, схемы которых состоят из сложений по модулю 2^n, побитовых XOR и циклических сдвигов битов на r позиций. Подсчитано число невозможных разностей, т. е. разностей с вероятностью 0, при всех возможных значениях r и n. Найдена также их доля при $r, n-r \to \infty$ и приведено сравнение с количеством невозможных разностей для преобразования $x \oplus y$ без циклического сдвига.

Ключевые слова: *ARX, разностные характеристики, XOR, сложение по модулю, циклический сдвиг битов, невозможные разности.*

Введение

Современные криптографические примитивы, такие, как CHAM [1], Sparkle [2], Speck [3], Salsa20 [4], ChaCha [5], представленные в виде схем, состоящих только из сложений по модулю 2^n ($\boxplus$), побитовых «исключающих или» (XOR, $\oplus$) и циклических сдвигов битов на r позиций в сторону старших разрядов ($\lll r$), называются ARX-шифрами. Построение оценок их стойкости к разностным атакам [6–8] приводит к необходимости анализа разностных характеристик ARX-преобразований. Далее изучаются разности относительно операции $\boxplus$, в общем же они могут быть выбраны и иными способами [9–12]. Соответствующие разностные характеристики базовых операций исследовались, например, в [13–16].

Будем рассматривать аргументы всех преобразований как элементы пространства $\mathbb{Z}_2^n$ двоичных векторов размерности n, ассоциируя с вектором $x = (x_1, \ldots, x_n) \in \mathbb{Z}_2^n$ целое число

$$2^{n-1}x_1 + 2^{n-2}x_2 + \cdots + 2^0 x_n.$$

Через $x \boxplus y$, где $x, y \in \mathbb{Z}_2^n$, обозначим сложение по модулю 2^n ассоциированных с x и y чисел. Значения разностной характеристики adp^f по модулю 2^n для преобразования $f : (\mathbb{Z}_2^n)^k \to \mathbb{Z}_2^n$ определяются следующим образом:

$$\mathrm{adp}^f(\alpha^1, \ldots, \alpha^k \to \alpha^{k+1}) =$$
$$= 2^{-nk} \#\{x^1, \ldots, x^k \in \mathbb{Z}_2^n : f(x^1 \boxplus \alpha^1, \ldots, x^k \boxplus \alpha^k) = f(x^1, \ldots, x^k) \boxplus \alpha^{k+1}\}.$$

Через $\mathrm{adp}^{\mathrm{XR}}$ обозначим разностную характеристику для преобразования $(x \oplus y) \lll r$, где $x, y \in \mathbb{Z}_2^n$ и $1 \leqslant r < n$, а через $\mathcal{N}_n^r$ — количество её невозможных разностей. Другими словами,

$$\mathcal{N}_n^r = \#\{\alpha, \beta, \gamma \in \mathbb{Z}_2^n : \mathrm{adp}^{\mathrm{XR}}(\alpha, \beta \xrightarrow{r} \gamma) = 0\}.$$

Отметим, что свойства $\mathrm{adp}^{\mathrm{XR}}$ исследовались в [17, 18]. При этом в [18] установлено, что $\mathcal{N}_n^1 = \dfrac{5}{14}8^n - \dfrac{6}{7}$ и $\mathcal{N}_n^r < \mathcal{N}_n^1$ при $r > 1$.

[1]Работа выполнена при поддержке Математического центра в Академгородке, соглашение с Министерством науки и высшего образования Российской Федерации № 075–15–2022–282.

Данная работа продолжает исследования в этом направлении. Мы подсчитаем точное значение $\mathcal{N}_n^r$ и покажем, что в большинстве случаев упомянутая оценка является грубой.

Количество невозможных разностей

В следующей теореме приводится точное значение $\mathcal{N}_n^r$ для всех возможных n и r.

Теорема 1. Пусть $n \geqslant 2$. Тогда для любого r, $2 \leqslant r \leqslant n-3$, справедливо:

$$\mathcal{N}_n^r = \frac{1}{7}8^n + \frac{3}{7}4^r 8^{n-r} - \frac{1}{7}8^r + \frac{1}{14}4^r + \\ + \left(\frac{6}{35}8^{n-r} + \frac{10}{3}4^{n-r} - \frac{8}{5}3^{n-r} + \frac{2}{21}\right)\left(\frac{1}{14}8^r - \frac{1}{12}4^r - \frac{5}{21}\right).$$

Крайние значения выражаются следующим образом:

$$\mathcal{N}_n^{n-2} = \frac{11}{56}8^n + \frac{35}{24}4^n - \frac{250}{21} \text{ при } n \geqslant 3 \text{ и } \mathcal{N}_n^{n-1} = \frac{3}{14}8^n + \frac{2}{3}4^n - \frac{50}{21}.$$

В табл. 1 представлены значения $\mathcal{N}_n^r$ при небольших n, а также доли невозможных разностей по отношению ко всем разностям. В табл. 2 приведены аналогичные доли при самом востребованном значении $n = 32$.

Таблица 1

Значения $\mathcal{N}_n^r$ при $3 \leqslant n \leqslant 10$

n	r	$\mathcal{N}_n^r$	$\mathcal{N}_n^r/8^n \cdot 100\%$	n	r	$\mathcal{N}_n^r$	$\mathcal{N}_n^r/8^n \cdot 100\%$
3	1	182	35,5469%	8	3	3563358	21,2393%
	2	150	29,2969%		4	3232014	19,2643%
4	1	1462	35,6934%		5	3198622	19,0653%
	2	1166	28,4668%		6	3391086	20,2124%
	3	1046	25,5371%		7	3638806	21,6890%
5	1	11702	35,7117%	9	1	47934902	35,7143%
	2	8958	27,3376%		2	34786302	35,7143%
	3	7918	24,1638%		3	28144334	20,9692%
	4	7702	23,5046%		4	25110494	18,7088%
6	1	93622	35,7140%		5	24182542	18,0174%
	2	69806	26,6289%		6	24778398	18,4613%
	3	59422	22,6677%		7	26746478	19,9277%
	4	57454	21,9170%		8	28935702	21,5588%
	5	58902	22,4693%	10	1	383479222	35,7143%
7	1	748982	35,7142%		2	277687598	25,8617%
	2	550206	26,2359%		3	223642910	20,8284%
	3	456078	21,7475%		4	197714510	18,4136%
	4	425118	20,2712%		5	187255262	17,4395%
	5	435822	20,7816%		6	186758926	17,3933%
	6	460310	21,9493%		7	194983582	18,1593%
8	1	5991862	35,7143%		8	212442734	19,7853%
	2	4366574	26,0268%		9	230786582	21,4937%

Большую часть $\mathcal{N}_n^r$ можно оценить сверху не только через $\mathcal{N}_n^1$, но и через $\mathcal{N}_n^{n-1}$.

Следствие 1. Пусть $r \geqslant 5$ и $n - r \geqslant 5$. Тогда $\mathcal{N}_n^r < \mathcal{N}_n^{n-1}$.

Более того, экспериментальные данные показывают, что $\mathcal{N}_n^r < \mathcal{N}_n^{n-1}$ при $2 < r < n-1$, начиная с $n = 7$.

В случае далёких от крайних значений r нетрудно вычислить, к чему стремится доля невозможных разностей.

Таблица 2

Доля $\mathcal{N}_{32}^r$ к 8^{32} при $2 \leqslant r < 32$

r	$\mathcal{N}_{32}^r/8^{32} \cdot 100\%$	r	$\mathcal{N}_{32}^r/8^{32} \cdot 100\%$	r	$\mathcal{N}_{32}^r/8^{32} \cdot 100\%$
2	25,8036%	12	15,5203%	22	15,5328%
3	20,6808%	13	15,5153%	23	15,5550%
4	18,0985%	14	15,5128%	24	15,5987%
5	16,8047%	15	15,5116%	25	15,6843%
6	16,1575%	16	15,5112%	26	15,8504%
7	15,8339%	17	15,5112%	27	16,1691%
8	15,6720%	18	15,5118%	28	16,7690%
9	15,5911%	19	15,5132%	29	17,8571%
10	15,5507%	20	15,5160%	30	19,6429%
11	15,5304%	21	15,5216%	31	21,4286%

Следствие 2. Справедливо

$$\lim_{\substack{r\to\infty \\ n-r\to\infty}} \frac{\mathcal{N}_n^r}{8^n} = \frac{38}{245}.$$

Отметим, что $\dfrac{38}{245} \approx 15{,}5\,\%$. В то же время доля невозможных разностей по модулю 2^n для преобразования $x \oplus y$ равна $4/7$ [14], что составляет примерно $57\,\%$. Таким образом, добавление циклического сдвига к преобразованию $x \oplus y$ существенно снижает число невозможных разностей, особенно при значениях сдвига, далёких от крайних.

Полученные результаты можно применить также к преобразованиям $(x \lll r) \oplus y$, $((x \boxplus y) \lll r) \oplus z$ и т. д., поскольку их разностные характеристики $\mathrm{adp}^{\mathrm{RX}}$ и $\mathrm{adp}^{\mathrm{ARX}}$ выражаются через $\mathrm{adp}^{\mathrm{XR}}$ следующим образом:

$$\mathrm{adp}^{\mathrm{RX}}(\alpha, \beta \xrightarrow{r} \gamma) = \mathrm{adp}^{\mathrm{XR}}(\gamma, \beta \xrightarrow{n-r} \alpha),$$
$$\mathrm{adp}^{\mathrm{ARX}}(\alpha, \beta, \gamma \xrightarrow{r} \delta) = \mathrm{adp}^{\mathrm{RX}}(\alpha \boxplus \beta, \gamma \xrightarrow{r} \delta),$$

где $\alpha, \beta, \gamma, \delta \in \mathbb{Z}_2^n$.

ЛИТЕРАТУРА

1. *Roh D., Koo B., Jung Y., et al.* Revised version of block cipher CHAM // LNCS. 2020. V. 11975. P. 1–19.
2. *Beierle C., Biryukov A., dos Santos L. C., et al.* Lightweight AEAD and hashing using the Sparkle permutation family // IACR Trans. Symmetric Cryptology. 2020. No. S1. P. 208–261.
3. *Beaulieu R., Shors D., Smith J., et al.* The SIMON and SPECK Families of Lightweight Block Ciphers. Cryptology Eprint Archive. 2013. `https://eprint.iacr.org/2013/404`.
4. *Bernstein D. J.* Salsa20 Specification. eSTREAM Project Algorithm Description. `http://www.ecrypt.eu.org/stream/salsa20pf.html`. 2005.
5. *Bernstein D. J.* ChaCha, a variant of Salsa20 // Workshop Record of SASC. 2008. V. 8. No. 1. P. 3–5.
6. *Biham E. and Shamir A.* Differential cryptanalysis of DES-like cryptosystems // J. Cryptology. 1991. V. 4. No. 1. P. 3–72.
7. *Knudsen L.* DEAL —- A 128-bit Block Cipher. Tech. Rep., Department of Informatics, University of Bergen, Bergen, Norway, February 1998.
8. *Biham E., Biryukov A., and Shamir A.* Cryptanalysis of skipjack reduced to 31 rounds using impossible differentials // LNCS. 1999. V. 1592. P. 12–23.

9. *Biryukov A. and Velichkov V.* Automatic search for differential trails in ARX ciphers // LNCS. 2014. V. 8366. P. 227–250.
10. *Leurent G.* Analysis of differential attacks in ARX constructions // LNCS. 2012. V. 7658. P. 226–243.
11. *Малышев Ф. М.* Вероятностные характеристики разностных и линейных соотношений для неоднородной линейной среды // Математические вопросы криптографии. 2019. Т. 10. № 1. С. 41–72.
12. *Малышев Ф. М.* Разностные характеристики основных операций ARX-шифров // Математические вопросы криптографии. 2020. Т. 11. № 4. С. 97–105.
13. *Daum M.* Cryptanalysis of Hash Functions of the MD4-Family. PhD Thesis. Ruhr-Universitat Bochum, May 2005.
14. *Lipmaa H., Wallén J., and Dumas P.* On the additive differential probability of exclusive-or // LNCS. 2004. V. 3017. P. 317–331.
15. *Mouha N., Velichkov V., De Cannière C., and Preneel B.* The differential analysis of S-functions // LNCS. 2010. V. 6544. P. 36–56.
16. *Mouha N., Kolomeec N., Akhtiamov D., et al.* Maximums of the additive differential probability of Exclusive-Or // IACR Trans. Symmetric Cryptology. 2021. No. 2. P. 292–313.
17. *Velichkov V., Mouha N., De Cannière C., and Preneel B.* The additive differential probability of ARX // LNCS. 2011. V. 6733. P. 342–358.
18. *Kolomeec N., Sutormin I., Bykov D., et al.* On Additive Differential Probabilities of the Composition of Bitwise Exclusive-OR and a Bit Rotation. arXiv preprint. 2023. `https://arxiv.org/abs/2303.04097`.

УДК 004.75 DOI 10.17223/2226308X/16/13

АНАЛИЗ ЭФФЕКТИВНОСТИ КРИПТОГРАФИЧЕСКИХ АЛГОРИТМОВ ДЛЯ ПРИМЕНЕНИЯ В ZK-SNARK[1]

Д. О. Кондырев

Проведён сравнительный анализ эффективности различных криптографических алгоритмов с точки зрения применимости в системах на основе zk-SNARK. Разработана инфраструктура на основе ZoKrates для проведения экспериментов с замером параметров. Определены границы практической применимости алгоритмов в распределённых реестрах.

Ключевые слова: *распределённые реестры, доказательство с нулевым разглашением, zk-SNARK, R1CS, эффективность алгоритмов.*

Одним из факторов, существенно ограничивающих применение распределённых реестров в промышленных программных системах, является безопасность информации, которая должна обеспечиваться такими системами.

Для решения проблемы приватности транзакций и смарт-контрактов применяются неинтерактивные протоколы доказательства с нулевым разглашением. Наибольшее распространение в распределённых реестрах получил протокол zk-SNARK [1]. Преимущество zk-SNARK над другими протоколами доказательства с нулевым разглашением заключается в гарантиях эффективности: длина доказательства зависит только от параметра безопасности, а время проверки не зависит от размера схемы и секретного параметра.

[1]Работа выполнена при поддержке Математического центра в Академгородке, соглашение с Министерством науки и высшего образования Российской Федерации № 075-15-2022-282.

Теоретически zk-SNARK позволяет задавать произвольные ограничения [1], но на практике возникают ограничения по времени работы и по памяти. При этом традиционные алгоритмы, оптимизированные для эффективного программного вычисления или реализации в виде аппаратного обеспечения, оказываются не настолько эффективными для применения в zk-SNARK. В связи с этим встают задачи разработки специальных алгоритмов, оптимизированных для применения в zk-SNARK, и оптимизации существующих алгоритмов.

Для доказательства определённого утверждения с помощью zk-SNARK оно должно быть сформулировано в форме, удобной для дальнейшей эффективной обработки. Одним из таких представлений являются системы ограничений ранга 1 (Rank-1 Constraint Systems, R1CS). Благодаря краткости и лаконичности R1CS превратились в де-факто стандарт для реализаций zk-SNARK. Поскольку в R1CS используется стандартная линейная алгебра, они также хорошо подходят для описания криптографических протоколов и теоретического анализа [2].

Эффективность алгоритмов установочной фазы и генерации доказательства zk-SNARK по времени работы и по памяти напрямую зависит от количества ограничений в R1CS-представлении. Таким образом, оптимизация, которая уменьшает количество ограничений в R1CS без изменения её выполнимости, имеет решающее значение для эффективности.

В контексте реальных приложений zk-SNARK эффективность имеет решающее значение для обеспечения практической применимости. Особенно это актуально в распределённых реестрах с их дополнительными ограничениями по времени работы и по памяти. Поэтому для конкретных алгоритмов и задач (систем ограничений) важно понимать, насколько они эффективны с точки зрения количества ограничений в R1CS-представлении.

Цель данной работы — сделать полный сравнительный анализ производительности наиболее часто применяемых в распределённых реестрах криптографических протоколов и определить для них границы практической применимости.

Для проведения вычислительных экспериментов использовался ZoKrates — набор программных инструментов для создания доказательств знания с нулевым разглашением, использующий zk-SNARK в качестве системы проверки [3]. На его основе была разработана программная инфраструктура для автоматизации генерации схем и замера параметров.

Рассмотрены наиболее часто применяемые в схемах zk-SNARK в распределённых реестрах алгоритмы, реализация которых взята из стандартной библиотеки ZoKrates. Проведены эксперименты с алгоритмами хеширования sha256, sha3, blake2, poseidon, mimc.

Протокол zk-SNARK гарантирует, что размер доказательства и время его проверки не зависят от сложности системы ограничений. Однако другие параметры могут отличаться, по ним можно определить эффективность того или иного алгоритма и возможность его применения для определённой задачи. В экспериментах измерялись следующие параметры сравниваемых алгоритмов:

— длина ключа доказательства и ключа верификации;
— время установочной фазы протокола (алгоритма генерации ключей);
— количество ограничений в R1CS-представлении алгоритма;
— время генерации доказательства.

Для каждого алгоритма проводилась серия замеров с различными входными данными. Чтобы повысить точность, каждый эксперимент с замером времени повторялся $N = 10$ раз, после чего полученные значения времени усреднялись. В результате экспериментов определены зависимости измеряемых параметров от размера входных данных.

Полученные данные позволяют сделать выводы об эффективности конкретных алгоритмов и понять границы их применимости при использовании в схемах zk-SNARK. Показано, что классические хеш-функции, такие, как sha256, плохо оптимизированы для zk-SNARK и могут ограниченно применяться только при относительно небольших размерах входных данных (порядка нескольких килобайт). Лучшие результаты показывают алгоритмы, изначально разработанные с целью оптимизации количества ограничений в R1CS-представлении (например, хеш-функция poseidon).

ЛИТЕРАТУРА

1. *Ben-Sasson E., Chiesa A., Genkin D., et al.* SNARKs for C: Verifying program executions succinctly and in zero knowledge // LNCS. 2013. V. 8043. P. 90–108.
2. *Eberhardt J.* Scalable and Privacy-preserving Off-chain Computations. Thesis: Ph.D., Technical University of Berlin, Faculty IV — Electrical Engineering and Computer Science, 2021. 284 p.
3. *Eberhardt J. and Tai S.* ZoKrates — scalable privacy-preserving off-chain computations // IEEE Intern. Conf. Blockchain. Halifax, Canada, 2018. P. 1084–1091.

УДК 003.26 DOI 10.17223/2226308X/16/14

ОБ ОДНОМ РЕЖИМЕ РАБОТЫ БЛОЧНЫХ ШИФРОВ ДЛЯ ЗАЩИТЫ СИСТЕМНЫХ НОСИТЕЛЕЙ С БЛОЧНО-ОРИЕНТИРОВАННОЙ СТРУКТУРОЙ

А. М. Коренева, Г. В. Фирсов

В 2022 г. приняты рекомендации по стандартизации, определяющие режим работы блочных шифров DEC, используемый для защиты носителей информации с блочно-ориентированной структурой. Режим DEC имеет эксплуатационные особенности, усложняющие его использование для шифрования системных дисков, из-за чего востребован синтез альтернативных режимов для полнодискового шифрования. В большинстве существующего ПО для шифрования системных дисков используется режим XTS, но он имеет особенности, ухудшающие его криптографические качества. Предлагается модификация режима XTS — режим XEH (Xor-Encrypt-Hash), для которого получена нижняя оценка уровня информационной безопасности и исследованы некоторые эксплуатационные характеристики.

Ключевые слова: *полнодисковое шифрование, режим работы блочных шифров, симметричная криптография, криптографическая защита информации, системные носители информации.*

Введение

Известно, что для классов систем криптографической защиты информации (СКЗИ), начиная с КС2 и выше, требуется обеспечить защиту от действий нарушителя, который, находясь внутри контролируемой зоны, может реализовать угрозу кражи носителя информации. В связи с этим требуется обеспечить конфиденциальность хранимых данных при наличии у нарушителя непосредственного доступа к диску. Для этого в СКЗИ используется криптографическая подсистема с функци-

ей полнодискового шифрования (ПДШ), то есть полного шифрования всех данных, расположенных на носителе. В существующих технических решениях, используемых для ПДШ, чаще всего применяются симметричные блочные шифры, функционирующие в специально разработанных режимах, которые учитывают эксплуатационные ограничения, возникающие при шифровании носителей информации.

В 2021 г. в ТК 26 (технический комитет по стандартизации «Криптографическая защита информации») завершилась разработка режима работы блочных шифров для защиты носителей информации с блочно-ориентированной структурой — Disk Encryption with Counter (DEC) [1]. Режим DEC требует, помимо шифртекста, хранить дополнительно счётчики, необходимые для его функционирования [2, 3], что ухудшает эксплуатационные качества. В связи с этим востребована разработка альтернативных решений для защиты информации на носителях. Во многих существующих решениях для ПДШ используется режим XEX-based Tweaked-codebook mode with ciphertext Stealing (XTS), описанный в NIST SP 800-38E «Recommendation for Block Cipher Modes of Operation: The XTS-AES Mode for Confidentiality on Storage Devices». При этом режим XTS обладает слабостями, лежащими в основе известных атак [4, 5]. Однако широта распространения режима XTS в прикладном ПО для ПДШ стала основанием для его выбора в качестве базового при построении предлагаемого режима XEH.

1. Целевые характеристики предлагаемого режима XEH

Целевые эксплуатационные характеристики формулируются с учётом повышенных требований к производительности подсистемы ПДШ, а также ограниченности доступной памяти раздела EFI системного носителя. К данным характеристикам относятся:

— минимально возможное количество обращений к примитиву блочного шифра;
— отсутствие дополнительных данных, необходимых для работы режима и хранимых на диске.

Целевые криптографические характеристики формулируются на основании модели и возможностей нарушителя, рассмотренных в [5]. В этой же работе проанализированы некоторые атаки на режим XTS. К целевым криптографическим характеристикам предлагаемого режима относятся:

— стойкость в модели RND-fdeCPA-sector [5] (формализация обеспечения конфиденциальности);
— невозможность построения атак, к которым уязвим режим XTS.

2. Определение режима XEH и оценка его уровня информационной безопасности

Обозначим через l длину блока в битах, n — количество блоков в секторе ($0 < < n < 2^l$). Здесь и далее V_l — множество битовых строк длины l, $\mathbb{F} = \mathrm{GF}(2)[x]/p(x)$, где $p(x) = x^{128} + x^7 + x^2 + x + 1$ для $l = 128$; $p(x) = x^{64} + x^4 + x^3 + x + 1$ для $l = 64$.

В предлагаемом режиме XEH применяется функция вида $g : \mathbb{F}^2 \times \mathbb{F}^n \to \mathbb{F}^n$, которая при фиксированных первых двух аргументах (их будем называть подключами) является обратимой. Обозначим через $g_{\tau_2,\tau_3}(\mathbf{y})$ значение функции $g((\tau_2,\tau_3),\mathbf{y})$ для $\tau_2,\tau_3 \in \mathbb{F}$, $\mathbf{y} \in \mathbb{F}^n$. Определим $g_{\tau_2,\tau_3}(\mathbf{y})$:

$$g_{\tau_2,\tau_3}(\mathbf{y}) = \left(y_1 + Y_{\tau_3} + \tau_2,\ y_2 + Y_{\tau_3} + \tau_2 \cdot \alpha,\ \ldots,\ y_{n-1} + Y_{\tau_3} + \tau_2 \cdot \alpha^{n-2},\ Y_{\tau_3} + \tau_2 \cdot \alpha^{n-1}\right),$$

$$Y_{\tau_3} = \left(\sum_{j=1}^{n} y_j \cdot \tau_3^{n-j}\right) + \left(\sum_{j=1}^{n-1} y_j \cdot \mathfrak{F}_l(j)\right),$$

где $\alpha = x$ — примитивный элемент поля $\mathbb{F}$; $+$ — сложение в $\mathbb{F}$; $\cdot$ — умножение в $\mathbb{F}$; $\mathfrak{F}_l : \mathbb{Z}_{2^l} \to \mathbb{F}$ — отображение, сопоставляющее элементу $r = \sum\limits_{i=0}^{l-1} a_i 2^i$ кольца $\mathbb{Z}_{2^l}$ элемент $\tilde{r} = \sum\limits_{i=0}^{l-1} a_i x^i$ поля $\mathbb{F}$, $a_i \in \{0,1\}$, $i = 0, \ldots, l-1$.

Помимо g, в предлагаемом режиме используется функция $\varphi : \mathbb{F} \times \mathbb{F}^n \to \mathbb{F}^n$, являющаяся обратимой при фиксированном первом аргументе. Обозначим через $\varphi_{\tau_3}(\mathbf{y})$ значение $\varphi(\tau_3, \mathbf{y})$ для $\tau_3 \in \mathbb{F}$ и $\mathbf{y} \in \mathbb{F}^n$. Определим $\varphi_{\tau_3}(\mathbf{y})$ следующим образом:

$$\varphi_{\tau_3}(\mathbf{y}) = (Z_{\tau_3},\ y_2 + Z_{\tau_3},\ \ldots,\ y_n + Z_{\tau_3}),$$
$$Z_{\tau_3} = \sum_{j=1}^{n} y_j \cdot \tau_3^{j-1}.$$

Пусть $\mathcal{E}$ — симметричный блочный шифр; $\mathcal{M}, \mathcal{C}, \mathcal{K}$ — множества открытых текстов, шифртекстов и ключей соответственно; $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$ — функция зашифрования шифра $\mathcal{E}$ (будем обозначать $E(K, m)$ через $E_K(m)$ для $K \in \mathcal{K}$, $m \in \mathcal{M}$); $\Delta_l : V_l \to \mathbb{F}$ — отображение, сопоставляющее строке $a = (a_0, \ldots, a_{l-1})$ из V_l элемент $\tilde{a} = \sum\limits_{i=0}^{l-1} a_i x^i$ поля $\mathbb{F}$, $a_i \in \{0,1\}$, $i = 0, \ldots, l-1$; $\nabla_l : \mathbb{F} \to V_l$ — отображение, обратное к Δ_l.

Как и режим XTS, XEH использует два независимых ключа $K, K' \in \mathcal{K}$. Из номера сектора SN вырабатываются три подключа:

$$\tau_1 = \Delta_l(E_K(SN)),\ \tau_2 = \Delta_l(E_{K'}(\nabla_l(\tau_1))),\ \tau_3 = \Delta_l(E_{K'}(SN)).$$

Процедура зашифрования в режиме XEH с использованием блочного шифра $\mathcal{E}$ состоит в следующем:

$$(w_1, \ldots, w_n) = \varphi_{\tau_3}(\Delta_l(m_1), \ldots, \Delta_l(m_n)),$$
$$y_j = \Delta_l(E_K(\nabla_l(w_j + \tau_1 \cdot \alpha^{j-1}))),\ j = 1, \ldots, n,$$
$$(z_1, \ldots, z_n) = g^{-1}_{\tau_2, \tau_3}(y_1, \ldots, y_n),$$
$$\mathbf{c} = (\nabla_l(z_1), \ldots, \nabla_l(z_n)),$$

где $\mathbf{c} = (c_1, \ldots, c_n)$ — шифртекст; $\mathbf{m} = (m_1, \ldots, m_n)$ — открытый текст; $m_j, c_j \in V_l$, $j = 1, \ldots, n$.

При оценке уровня информационной безопасности использовалась методика, описанная в работах [6, 7], в которых также введены основные термины, обозначения и определения.

Теорема 1. Пусть π — случайная подстановка множества V_l. При фиксированных натуральных числах l, n, q верна следующая нижняя оценка уровня информационной безопасности режима XEH:

$$\mathbf{Adv}^{\text{RND-fdeCPA-sector}}_{\text{XEH}^\pi}(q) \leqslant \frac{2(n+1)^2 q^2}{2^l},$$

где q — количество запросов к экспериментатору в эксперименте RND-fdeCPA-sector [5].

3. Сравнение с существующими режимами работы блочных шифров для ПДШ

Из определения режима XEH следует, что для его работы не требуется сохранять какие-либо данные на носителе информации. Для сравнения, режим DEC требует хранения отдельного счётчика на каждый сектор и раздел (набор последовательных секторов одного диска, объединённых логически) [3]. Таким образом, при шифровании раздела объёмом 32 Гбайт с использованием блочного шифра с размером блока 128 бит потребуется хранить 512 Мбайт дополнительных данных при размере сектора 512 байт либо 64 Мбайт при размере сектора 4096 байт. Для защиты системных носителей информации это может быть затруднительно в связи с ограниченным объёмом системного раздела EFI (100 Мбайт).

Предложенный режим XEH реализован на языке C с использованием встроенных в компилятор функций, соответствующих инструкции умножения без переноса (`PCLMULQDQ`), а также набору инструкций SSE2, и впоследствии интегрирован в модифицированную сборку ПО VeraCrypt. Проведено сравнение производительности зашифрования и расшифрования данных одного сектора в режиме XEH с режимами XTS, CMC (CBC-Mask-CBC) и HEH (Hash-ECB-Hash). Данные режимы выбраны для сравнения, так как не требуют хранения дополнительных данных, кроме номера сектора, то есть удовлетворяют принятым эксплуатационным ограничениям. С режимом DEC сравнение производительности не проводилось по причине невозможности обеспечить одинаковые условия функционирования режимов в проводимом эксперименте. Необходимость хранения счётчиков режима DEC не позволяет смоделировать принятые условия шифрования системного носителя информации. Результаты представлены в таблице, за единицу принято время работы режима XTS. Использовался блочный шифр Кузнечик из ГОСТ 34.12-2018. Эксперимент проводился на ПЭВМ с процессором Intel(R) Core(TM) i7-9750H с постоянной тактовой частотой 2,6 ГГц, ОЗУ типа DDR4 объёмом 8 Гбайт, 64-битной операционной системой macOS 13.1.

Относительное время зашифрования и расшифрования секторов

Режим	Зашифрование, 512 байт	Зашифрование, 4096 байт	Расшифрование, 512 байт	Расшифрование, 4096 байт
XTS	1	1	1	1
XEH	1,087	1,046	1,092	1,069
CMC	1,351	1,409	1,533	1,431
HEH	0,670	0,812	1,022	0,994

Результаты эксперимента показывают, что предложенный режим XEH уступает в производительности режиму XTS не более 10 %, что связано с увеличением количества выполняемых операций. В то же время предложенный режим не позволяет провести некоторые вычисления единожды заранее, в отличие от режима HEH, из-за чего последний оказывается более производительным. Предложенный режим совершает $n+3$ вызова блочного шифра против $2n+1$ для режима CMC, что даёт преимущество более 35 %.

Полученные результаты могут быть востребованы при совершенствовании и разработке подсистем полнодискового шифрования в составе средств криптографической защиты информации.

ЛИТЕРАТУРА

1. Рекомендации по стандартизации Р 1323565.1.042–2022 «Информационная технология. Криптографическая защита информации. Режим работы блочных шифров, предназначенный для защиты носителей информации с блочно-ориентированной структурой». М.: Стандартинформ, 2022.
2. *Bodganov D. and Nozdrunov V.* Some properties of one mode of operation of block ciphers // 10th Workshop CTCrypt 2021. Pre-proceedings. 2021. P. 12–17.
3. *Богданов Д., Ноздрунов В.* Шифрование носителей информации. Режим DEC // Рус-Крипто'2021. `https://www.ruscrypto.ru/resource/archive/rc2021/files/02_bogdanov_nozdrunov.pdf`.
4. *Isobe T. and Minematsu K.* Plaintext recovery attacks against XTS beyond collisions // LNCS. 2020. V. 11959. P. 103–123.
5. *Firsov G. and Koreneva A.* On one block cipher mode of operation used to protect data on block-oriented storage devices // Modern Inform. Technologies and IT-Education. 2022. No. 18(3) P. 691–701.
6. *Смышляев С. В.* Математические методы обоснования оценок уровня информационной безопасности программных средств защиты информации, функционирующих в слабодоверенном окружении: дис. ... докт. физ.-мат. наук. М., 2022.
7. *Ахметзянова Л. Р.* Комбинаторные свойства схем обеспечения конфиденциальности и целостности информации: дис. ... канд. физ.-мат. наук. М., 2022.

УДК 519.719.2 DOI 10.17223/2226308X/16/15

ПОСТРОЕНИЕ РАЗНОСТНОГО СООТНОШЕНИЯ ДЛЯ АЛГОРИТМА КБ-256

А. В. Курочкин, А. Б. Чухно, Д. А. Бобровский

Построено разностное соотношение для алгоритма шифрования КБ-256. Вероятность его выполнения для 15 из 16 раундов не меньше чем 2^{-134}.

Ключевые слова: *разностный и линейный методы криптографического анализа, КБ-256.*

В данной работе построено разностное соотношение для блочного алгоритма шифрования КБ 256-3 на основе обобщённой сети Фейстеля [1].

Применение разностного метода [2] к алгоритму блочного шифрования состоит из двух этапов. На первом, подготовительном, этапе для схемы строятся разностные соотношения. На втором этапе по имеющемуся материалу проверяется гипотеза о вероятности выполнения соотношения — отличие её от равновероятной при опробовании ключа шифрования.

Пусть $h : V_n \to V_m$ — преобразование, $a \in V_n$, $b \in V_m$ — фиксированные векторы. Тогда пара (a, b) называется разностным соотношением, если существуют $x \in V_n$, такие, что верно равенство

$$h(x \oplus a) \oplus h(x) = b.$$

Вероятность разностного соотношения равна

$$p_{a,b} = \mathsf{P}[h(x \oplus a) \oplus h(x) = b] = |\{x : h(x \oplus a) \oplus h(x) = b\}|/2^n.$$

После выделения нелинейных преобразований и оценки их разностных характеристик строится последовательность согласованных соотношений, позволяющая оценить вероятность выполнения разностного соотношения [2].

Для 15 раундов алгоритма КБ-256 построен разностный путь в предположении, что каждая тройка раундовых ключей выработана случайно и равновероятно. Для полораундового алгоритма шифрования КБ-256 построены согласованные локальные разностные соотношения. Для каждого раунда вычислена вероятность возникновения локальных разностных соотношений и для каждого такого соотношения посчитана доля ключей, для которых оно выполнимо. Данные приведены в таблице; входные и выходные разности представлены в ней следующим образом: для $A, B \in V_{256}$ записываем $A \oplus B$ как вектор $(c_0, c_1, \ldots, c_7)$, где $c_i \in V_{32}$ задаётся списком номеров единичных битов $[i_1, i_2, \ldots]$; нулевой вектор c_i будем обозначать как $\varnothing$, $i = 0, \ldots, 7$.

№ раунда	a	b	$p_{a,b}$
1	$(\varnothing, [31], \varnothing, \varnothing, \varnothing, \varnothing, [31], \varnothing)$	$([31], \varnothing, \varnothing, \varnothing, \varnothing, [31], \varnothing, \varnothing)$	1
2	$([31], \varnothing, \varnothing, \varnothing, \varnothing, [31], \varnothing, \varnothing)$	$(\varnothing, \varnothing, \varnothing, \varnothing, [31], \varnothing, \varnothing, [31])$	1
3	$(\varnothing, \varnothing, \varnothing, \varnothing, [31], \varnothing, \varnothing, [31])$	$(\varnothing, \varnothing, \varnothing, [31], \varnothing, \varnothing, [31], \varnothing)$	1
4	$(\varnothing, \varnothing, [31], \varnothing, \varnothing, [31], \varnothing, \varnothing)$	$(\varnothing, \varnothing, [31], \varnothing, \varnothing, [31], \varnothing, \varnothing)$	1
5	$(\varnothing, \varnothing, [31], \varnothing, \varnothing, [31], \varnothing, \varnothing)$	$(\varnothing, [31], \varnothing, \varnothing, [31], \varnothing, \varnothing, \varnothing)$	1
6	$(\varnothing, [31], \varnothing, \varnothing, [31], \varnothing, \varnothing, \varnothing)$	$([31], \varnothing, \varnothing, [31], \varnothing, \varnothing, \varnothing, \varnothing)$	1
7	$([31], \varnothing, \varnothing, [31], \varnothing, \varnothing, \varnothing, \varnothing)$	$(\varnothing, [17, 15], [31], \varnothing, [17], \varnothing, \varnothing, [31, 15])$	2^{-9}
8	$(\varnothing, [17, 15], [31], \varnothing, [17], \varnothing, \varnothing, [31, 15])$	$([17, 15], [31, 17, 15], \varnothing, [17], [17, 15], \varnothing, [31, 15], \varnothing)$	2^{-11}
9	$([17, 15], [31, 17, 15], \varnothing, [17], [17, 15], \varnothing, [31, 15], \varnothing)$	$([31, 17, 15], \varnothing, [17], [17, 15], \varnothing, [31, 15], [17, 15], [17, 15])$	2^{-4}
10	$([31, 17, 15], \varnothing, [17], [17, 15], \varnothing, [31, 15], [17, 15], [17, 15])$	$(\varnothing, [17, 2], [17, 15], \varnothing, [31, 15, 2], [17, 15], [17, 15], [31, 17, 15, 2])$	$2^{-10,42}$
11	$(\varnothing, [17, 2], [17, 15], \varnothing, [31, 15, 2], [17, 15], [17, 15], [31, 17, 15, 2])$	$([17, 2], [19, 17, 15, 2], \varnothing, [31, 15, 2], [19, 17, 15, 2], [17, 15], [31, 17, 15, 2], [19, 2])$	$2^{-20,62}$
12	$([17, 2], [19, 17, 15, 2], \varnothing, [31, 15, 2], [19, 17, 15, 2], [17, 15], [31, 17, 15, 2], [19, 2])$	$([19, 17, 15, 2], [19, 3], [31, 15, 2], [19, 17, 15, 2], [19, 17, 15, 3], [31, 17, 15, 2], [19, 2], [19, 17, 3, 2])$	$2^{-23,51}$
13	$([19, 17, 15, 2], [19, 3], [31, 15, 2], [19, 17, 15, 2], [19, 17, 15, 3], [31, 17, 15, 2], [19, 2], [19, 17, 3, 2])$	$([19, 3], [31, 19, 15, 3, 2], [19, 17, 15, 2], [19, 17, 15, 3], [31, 19, 17, 15, 3, 2], [19, 2], [19, 17, 3, 2], [17, 15, 3, 2])$	$2^{-23,1}$
14	$([19, 3], [31, 19, 15, 3, 2], [19, 17, 15, 2], [19, 17, 15, 3], [31, 19, 17, 15, 3, 2], [19, 2], [19, 17, 3, 2], [17, 15, 3, 2])$	$([31, 19, 15, 3, 2], [19, 17, 15, 2], [19, 17, 15, 3], [31, 19, 17, 15, 3, 2], [19, 2], [19, 17, 3, 2], [17, 15, 3, 2], [19, 3])$	$2^{-6,98}$
15	$([31, 19, 15, 3, 2], [19, 17, 15, 2], [19, 17, 15, 3], [31, 19, 17, 15, 3, 2], [19, 2] [19, 17, 3, 2], [17, 15, 3, 2], [19, 3])$	$([31, 22, 15, 3, 2, 0], [19, 2], [31, 22, 20, 19, 15], [22, 19, 3], \varnothing, [22, 3, 1, 0], [20, 15], [31, 20, 15])$	$2^{-24,55}$

Заключение

Построено разностное соотношение для полораундового алгоритма КБ-256. Итоговая вероятность выполнения разностного соотношения для раунда с номером $i \in \in \{1, \ldots, 15\}$ равна произведению вероятностей выполнения разностных соотношений для раундов с номерами $1, 2, \ldots, i-1$, в частности для 15 раундов она равна $2^{-133,2}$.

ЛИТЕРАТУРА

1. *Fomichev V. and Koreneva A.* Encryption performance and security of certain wide block ciphers // J. Computer Virology Hacking Techniques. 2020. V. 16. P. 197–216.
2. *Malyshev F. M. and Trishin A. E.* Linear and differential cryptanalysis: Another viewpoint // Математические вопросы криптографии. 2020. Т. 11. № 2. С. 83–98.

УДК 519.7 DOI 10.17223/2226308X/16/16

ОСНОВНЫЕ ПОДХОДЫ К ПОСТРОЕНИЮ ПОСТКВАНТОВЫХ КРИПТОСИСТЕМ: ОПИСАНИЕ, СРАВНИТЕЛЬНАЯ ХАРАКТЕРИСТИКА[1]

Е. С. Малыгина, А. В. Куценко, С. А. Новоселов, Н. С. Колесников, А. О. Бахарев, И. С. Хильчук, А. С. Шапоренко, Н. Н. Токарева

Постквантовая криптография является областью теоретических и прикладных исследований, включающей разработку и анализ методов криптографической защиты информации, актуальных в условиях широкого использования квантовых вычислений. В настоящее время наибольший интерес представляют направления, в рамках которых предлагаются криптосистемы, стойкость которых основывается на вычислительной трудности ряда задач из теории решёток, изогений и кодов, исправляющих ошибки. Данная работа является обзорной, она включает краткое изложение двух новых работ, подготовленных авторами и посвящённых описанию основных подходов к построению постквантовых криптографических систем. Рассмотрены вычислительно трудные задачи из данных направлений, проанализированы известные результаты о стойкости и быстродействии соответствующих криптосистем.

Ключевые слова: *постквантовая криптография, теория решёток, линейные коды, изогении эллиптических кривых, квантовый компьютер.*

1. Постквантовая криптография

Термин «постквантовая криптография» появился в середине 2000-х годов. Он используется для обозначения области криптографии, которая охватывает исследование методов построения криптосистем, которые останутся актуальными и после появления квантового компьютера, достаточно мощного для реализации алгоритмов квантового криптоанализа. В 2009 г. был опубликован сборник работ «Post-Quantum Cryptography», явившийся первой попыткой собрать, проанализировать и структурировать информацию о данном направлении криптографии [1]. Растущий интерес к постквантовой криптографии обусловен стремительным развитием квантовых вычислений и увеличением числа логических кубитов, с которыми может работать универсальный квантовый компьютер.

Оценка возможностей квантовых вычислений является актуальной открытой проблемой. Важно отметить, что на данный момент не известен полиномиальный квантовый алгоритм решения какой-либо NP-полной или NP-трудной задачи, что допускает возможность существования постквантовых криптографических систем, основанных, в частности, на различных вариантах таких задач. Таким образом, постквантовая криптография подразумевает развитие именно классической криптографии, при этом с точки зрения стойкости постквантовые криптосистемы должны обладать устойчивостью уже к *квантово-классическому* криптоанализу.

В 2016 г. Национальный институт стандартов и технологий США опубликовал отчёт «NISTIR 8105: Report on Post-Quantum Cryptography» [2], в котором проанализи-

[1]Работа второго, пятого, шестого, седьмого и восьмого авторов выполнена при поддержке Математического центра в Академгородке, соглашение с Министерством науки и высшего образования РФ № 075-15-2022-282. Работа первого, третьего и четвёртого авторов выполнена при поддержке Северо-Западного центра математических исследований имени С. Ковалевской, БФУ им. И. Канта, соглашение с Министерством науки и высшего образования РФ № 075-02-2023-934.

ровано влияние квантовых вычислений на тот момент (таблица) и описаны основные подходы к построению постквантовых криптосистем.

Влияние квантовых вычислений на некоторые известные криптосистемы, используемые в настоящее время [2]

Название	Тип	Предназначение	Влияние квантовых вычислений
AES	Симметричная	Шифрование	Требуется бóльшая длина ключа
SHA-2, SHA-3	Хеш-функция	Хеширование	Требуется бóльшая длина выходной последовательности
RSA	Асимметричная	Подпись, формирование общего ключа	Не безопасен
ECDSA, ECDH (Криптография на эллиптических кривых)	Асимметричная	Подпись, обмен ключами	Не безопасен
DSA (Криптография над конечными полями)	Асимметричная	Подпись, обмен ключами	Не безопасен

В конце 2017 г. Национальным институтом стандартов и технологий США был окончен приём заявок на участие в первом раунде конкурса, по итогам которого должен быть выбран стандарт постквантового асимметричного криптографического механизма для решения задач шифрования, формирования общего секретного ключа и электронной цифровой подписи [3].

В 2022 г. был завершён третий раунд конкурса [4], по итогам которого для шифрования и формирования общего ключа был выбран алгоритм CRYSTALS-Kyber, основанный на теории решёток, а для электронной подписи — алгоритмы CRYSTALS-Dilithium, FALCON и SPHINCS+, базирующиеся на решётках и использовании хеш-функций. В заявку на чётвертый раунд конкурса вошли альтернативные кандидаты, среди которых присутствуют криптосистемы, основанные на использовании кодов, исправляющих ошибки.

В работе приведён анализ текущей ситуации в области постквантовой криптографии, описание основных направлений, актуальных на сегодняшний день, их сравнение с точки зрения стойкости и быстродействия.

2. Решётки

Решётки использовались в математике, по меньшей мере, с XVIII в., однако лишь в 1990-е годы нашли применение не только для взлома уже существующих криптографических систем, но и для создания новых. Начало этому положила работа М. Айтаи [5], в которой показана связь между сложностью в среднем и сложностью в худшем случае для некоторых задач из теории решёток, что позволило использовать эти задачи в приложениях криптографии.

Пусть векторы $v_1, \ldots, v_n \in \mathbb{R}^m$ линейно независимы. *Решёткой, порождённой векторами* $v_1, \ldots, v_n$, называется набор линейных комбинаций векторов $v_1, \ldots, v_n$ с коэф-

фициентами из $\mathbb{Z}$:

$$L = \{a_1 v_1 + a_2 v_2 + \ldots + a_n v_n : a_1, a_2, \ldots, a_n \in \mathbb{Z}\}.$$

Базисом для L является любой линейно независимый набор векторов, порождающий L. Любые два базиса L связаны преобразованием с целочисленной матрицей, определитель которой равен ± 1. *Ранг решётки* L — число векторов в базисе L, *размерность решётки* L равна m; в случае, когда $n = m$, решётка L называется решёткой полного ранга. *Определителем* решётки L называется модуль определителя матрицы, столбцы которой являются базисом.

Минимальным расстоянием λ_1 в решётке L называется евклидова норма кратчайшего ненулевого вектора в L:

$$\lambda_1(L) = \min_{v \in L} ||v||, \ v \neq 0.$$

Вообще, $\lambda_i(L)$ — наименьший радиус r, такой, что в L существуют i линейно независимых векторов, норма которых не превосходит r. Известны следующие оценки минимального расстояния:

$$\lambda_1(L) \leqslant \sqrt{n}(\det L)^{1/n} \text{ (следствие из теоремы Минковского).}$$

Классические трудновычислимые задачи в теории решёток — поиск ненулевого вектора наименьшей длины (**SVP**) и поиск вектора в решётке, ближайшего к заданному вектору (**CVP**):

The Shortest Vector Problem (SVP). Найти ненулевой вектор наименьшей длины в решётке L, то есть найти ненулевой вектор $v \in L$, для которого выполняется $||v|| = \lambda_1(L)$.

The Closest Vector Problem (CVP). При заданном векторе $w \in \mathbb{R}^m$, не принадлежащем решётке L, найти вектор $v \in L$, ближайший к w, то есть найти вектор $v \in L$, который минимизирует евклидову норму $||w - v||$.

В приложениях криптографии используются приближённые версии указанных задач и некоторых других. Точность приближения регулируется параметром $\gamma \geqslant 1$, который, как правило, является некоторой функцией $\gamma(L)$ от решётки L.

Approximated Shortest Vector Problem ($\mathbf{SVP}_\gamma$). При заданном параметре γ найти ненулевой вектор $v \in L$, для которого выполняется $||v|| \leqslant \gamma \cdot \lambda_1(L)$.

Криптосистемы, основанные на задачах из теории решёток, являются популярными кандидатами на роль постквантовых криптосистем. Это обусловлено тем, что такие криптосистемы, например NTRU [6], исследуются уже много лет и до сих пор считаются защищёнными относительно классического и квантового криптоанализа. Кроме того, они легко реализуются, эффективны и хорошо параллелизуются.

Стойкость NTRU основывается на сложности задачи факторизации полинома в кольце полиномов, приведённых по модулю $x^n - 1$, на два полинома с маленькими коэффициентами. Одной из атак на криптосистему NTRU является решение SVP_γ в решётке размерности $2n$.

Существует множество алгоритмов, решающих SVP_γ для решёток размерности n. Примером такого алгоритма с полиномиальным временем работы и экспоненциальной точностью является алгоритм LLL [7]. Известен алгоритм перечисления [8], который для решения SVP задействует полиномиальную память, но его временна́я сложность имеет порядок $2^{\mathcal{O}(n \log n)}$. Алгоритмы BKZ и BKZ 2.0 [9] являются объединением алгоритмов LLL и перечисления. Данные алгоритмы имеют временну́ю сложность $2^{\mathcal{O}(\beta \log \beta)}$, и точность полученного решения составляет $\beta^{\mathcal{O}(n/\beta)}$, где β — параметр

из множества $\{2, \ldots, n\}$. Также активно развиваются алгоритмы просеивания, имеющие экспоненциальную по времени и памяти сложность относительно решения SVP. На данный момент можно выделить два известных алгоритма: Spherical LSF [10], являющийся оптимальным по временно́й сложности, и Triple sieve with NNS [11], являющийся оптимальным по используемой памяти.

Схемы CRYSTALS-Kyber, Saber и NTRU, представленные в третьем раунде конкурса NIST [4], имеют очень высокую скорость инкапсуляции и декапсуляции ключей. Объём требуемой памяти при использования Saber ниже, чем у Kyber, за счёт меньших размеров открытого ключа и шифртекста, однако эта разница незначительна. В свою очередь, затраты на генерацию ключа и, следовательно, общие затраты на использование схемы NTRU значительно выше, чем аналогичные затраты для Kyber. Схема NTRU также имеет наибольшие размеры открытых ключей и шифртекстов среди указанных трёх схем.

Полную версию обзора постквантовых криптосистем на решётках можно найти в [12].

3. Коды, исправляющие ошибки

Коды, исправляющие ошибки, используются для передачи информации в каналах связи, в которых информация искажается. Определённая избыточность в передаваемых кодовых словах (блоках) позволяет обнаруживать ошибки в принятых словах (блоках) и исправлять их, выбирая ближайшие кодовые слова. Особое распространение получили линейные коды в силу более эффективных алгоритмов кодирования и декодирования.

Пусть $\mathbb{F}_q$ — конечное поле, состоящее из q элементов, $\mathbb{F}_q^n$ — векторное пространство над $\mathbb{F}_q$. *Линейным* $[n, k]$*-кодом* $\mathcal{C}$ называется k-мерное векторное подпространство в $\mathbb{F}_q^n$. При этом вектор $(c_1, c_2, \ldots, c_n) \in \mathcal{C}$ называется *кодовым словом* $\mathcal{C}$.

Важным качеством кода является возможность исправления приобретённых в ходе передачи информации по зашумлённому каналу ошибок. Для определения кодового расстояния введём метрику на векторном пространстве $\mathbb{F}_q^n$. *Расстоянием Хэмминга* между векторами $x, y \in \mathbb{F}_q^n$ называется число координат, в которых векторы различаются: $\mathrm{d}_H(x, y) = |\{i : x_i \neq y_i\}|$. *Весом Хэмминга* вектора $x \in \mathbb{F}_q^n$ называется число его ненулевых координат: $\mathrm{wt}_H(x) = |\{i : x_i \neq 0\}| = \mathrm{d}_H(x, 0)$. *Кодовым расстоянием* кода $\mathcal{C}$ называется минимальное расстояние Хэмминга между его различными кодовыми словами: $d = \min\{\mathrm{d}_H(x, y) : x, y \in \mathcal{C}, x \neq y\}$. В случае линейных кодов имеем $d = \min\{\mathrm{wt}_H(x) : x \in \mathcal{C}, x \neq 0\}$. При этом число исправляемых кодом $\mathcal{C}$ ошибок равно $t = \lfloor (d-1)/2 \rfloor$.

Декодированием кода $\mathcal{C}$ называется отображение $D_{\mathcal{C}} : \mathbb{F}_q^n \to \mathcal{C}$. Код *исправляет* t *ошибок*, если для всех $e \in \mathbb{F}_q^n$ и всех $c \in \mathcal{C}$, таких, что $\mathrm{wt}_H(e) \leqslant t$, имеет место $D_{\mathcal{C}}(c + e) = c$.

Первой кодовой криптосистемой была схема шифрования с открытым ключом, предложенная в 1978 г. Р. Мак-Элисом [13]. Однако практически все асимметричные модификации на базе кодов, предложенные позже, имеют общий недостаток — большие требования к памяти. Согласно [14], существуют два основных предположения относительно безопасности схемы Мак-Элиса:

1) сложность задачи декодирования общего неизвестного кода, которая является NP-трудной [15];
2) сложность атак, восстанавливающих структуру базового кода.

Одной из основных задач, обеспечивающих безопасность кодовых криптосистем, является сокрытие структуры используемого кода.

Исходя из размеров открытого и закрытого ключей, классическая схема Мак-Элиса существенно проигрывает и схеме BIKE [16], и схеме HQC [17]. Несмотря на то, что схема HQC обеспечивает достаточный уровень безопасности, а также разумную частоту отказов при расшифровании, она проигрывает BIKE относительно размеров открытого ключа и зашифрованного текста. А потому схема BIKE показала себя наиболее конкурентоспособной.

Полную версию обзора постквантовых криптосистем на кодах можно найти в [18].

4. Изогении

Эллиптической кривой над полем $\mathbb{F}$ называется гладкая кривая E, заданная уравнением

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

где $a_1, a_2, a_3, a_4, a_6 \in \mathbb{F}$. Условие гладкости означает, что кривая не имеет сингулярных точек, т. е. точек, в которых обе частные производные функции $y^2 + a_1xy + a_3y - (x^3 + a_2x^2 + a_4x + a_6)$ равны нулю. В случае, если характеристика поля $\mathbb{F}$ не равна 2 и 3, то кривую можно привести (изоморфным преобразованием) к более простой форме

$$y^2 = x^3 + ax + b,$$

которая называется *краткой формой Вейерштрасса*.

Множество точек кривой E вместе с бесконечно удалённой точкой $\mathcal{O}$ обозначается как $E(\mathbb{F})$. Заметим, что кривая может быть задана над одним полем $\mathbb{F}$ (т. е. коэффициенты a_1, a_2, a_3, a_4, a_6 или a, b лежат в $\mathbb{F}$), но при этом точки кривой мы можем брать над некоторым его расширением, например алгебраическим замыканием $\overline{\mathbb{F}}$. Такие точки получаются из решений уравнения кривой над алгебраическим замыканием поля или, в общем случае, над расширением поля. Множество точек кривой E (заданной над $\mathbb{F}$), которые имеют координаты из поля $\overline{\mathbb{F}}$, обозначается как $E(\overline{\mathbb{F}})$.

Эллиптическая кривая E, заданная над конечным полем $\mathbb{F}_q$, где $q = p^n$ — степень простого числа, называется *суперсингулярной*, если $|E(\mathbb{F}_q)| \equiv 1 \pmod{p}$.

Изогенией двух эллиптических кривых E_1, E_2 над одним и тем же полем $\mathbb{F}$ называется ненулевой гомоморфизм эллиптических кривых, задаваемый рациональными отображениями. Явно изогении задаются в виде рациональных функций

$$\varphi : (x, y) \mapsto \left(\frac{f_1(x, y)}{f_2(x, y)}, \frac{g_1(x, y)}{g_2(x, y)} \right)$$

для некоторых $f_1, f_2, g_1, g_2 \in \mathbb{F}[x, y]$. Используя замену $y^2 \mapsto x^3 + ax + b$, изогению можно привести к виду

$$\varphi : (x, y) \mapsto \left(\frac{u(x)}{v(x)}, \frac{s(x)y}{t(x)} \right),$$

где $u, v, s, t \in \mathbb{F}[x]$. Такая форма изогении называется стандартной. *Степень* изогении определяется как $\deg \varphi = \max(\deg u, \deg v)$.

Пусть X — некоторое множество, а G — группа. Будем говорить, что G действует на множестве X, если задано отображение $* : G \times X \to X$, такое, что для любых $g_1, g_2 \in G$ и $x \in X$ выполняется $g_1 * (g_2 * x) = (g_1 g_2) * x$.

В терминах действия группы на множестве можно описать многие схемы шифрования с открытым ключом, протоколы распределения и инкапсуляции ключа (Key

Encapsulation Mechanism — KEM). При этом действие группы должно обладать некоторым криптографическим («трудновычислимым») свойством, например:

— группа G действует как односторонняя функция, т. е. для любых $x_1, x_2 \in X$ нахождение элемента $g \in G$, такого, что $x_1 = g * x_2$, является трудной задачей (даже полагая, что такой элемент существует);
— действие группы обладает свойством псевдослучайного генератора, т. е. для случайно выбранного элемента $g \in G$ злоумышленник не может отличить множество принятых векторов $\{(x_i, g * x_i)\}_i$ от множества векторов вида $\{(x_i, u_i)\}_i$, где u_i — равномерно распределённые на X случайные величины.

Одной из самых популярных схем на изогениях является схема SIDH (Supersingular Isogeny Diffie — Hellman), предложенная в 2011 г. Л. де Фео, Д. Яо и Дж. Плутом [19]. Она представляет собой протокол обмена ключами, аналогичный протоколу Диффи — Хеллмана, где в качестве X используется множество суперсингулярных эллиптических кривых над конечным полем, а элементы $\varphi_a, \varphi_b \in G$ — изогении суперсингулярных кривых. Однако в 2022 г. В. Кастрик и Т. Декру опубликовали препринт работы [20], в котором описывается полиномиальная атака на криптосистему SIKE (версию SIDH) — кандидата на стандартизацию NIST.

Схема CSIDH (Commutative Supersingular Isogeny Diffie — Hellman) — ещё один протокол обмена ключами, безопасность которого основана на сложности нахождения изогении между двумя суперсингулярными кривыми. Данная схема является устойчивой к атаке Кастрика — Декру. Конструкция схемы основана на криптосистеме Ростовцева — Столбунова, однако вместо обычных эллиптических кривых применяются суперсингулярные эллиптические кривые. Кроме того, в отличие от схемы SIDH, в CSIDH используется действие коммутативной группы. Впервые протокол CSIDH описан в 2018 г. В. Кастриком, Т. Ланге и др. [21], впоследствии к нему опубликовано множество технических оптимизаций [22].

В криптосистемах на изогениях, несмотря на малый размер ключа, главной проблемой остаётся медленная скорость работы схем. Атака Кастрика — Декру вывела из рассмотрения наиболее перспективную с точки зрения практики схему SIDH/SIKE и все схемы, для оптимизации которых использовались значения изогении в точках кручения. Поэтому наиболее важным направлением в области изогений является исследование вопросов оптимизации имеющихся схем.

Полную версию обзора постквантовых криптосистем на изогениях можно найти в [18].

Заключение

На основе проведённого анализа можно сделать вывод, что с точки зрения быстродействия самыми перспективными кандидатами на роль стандарта постквантовой криптографии являются криптосистемы на основе решёток. Они имеют сравнительно небольшой размер ключей, в то время как криптосистемы, основанные на кодах, имеют больший размер ключей и худшую производительность. Схемы на изогениях, несмотря на наименьший среди представленных подходов размер ключей, имеют наихудшую производительность. При этом стойкость актуальных постквантовых криптосистем основана на вычислительной сложности ряда частных случаев NP-трудных задач, и на данный момент не известны атаки, которые бы понижали её серьёзным образом.

ЛИТЕРАТУРА

1. *Bernstein D. J.* Introduction to post-quantum cryptography // Bernstein D. J., Buchmann J., and Dahmen E. (eds). Post-Quantum Cryptography. Berlin; Heidelberg: Springer, 2009. P. 1–14.
2. *Chen L., Jordan S., Liu Y.-K., et al.* NISTIR 8105: Report on Post-Quantum Cryptography. `https://csrc.nist.gov/publications/detail/nistir/8105/final`. 2016.
3. National Institute of Standards and Technology. Post-Quantum Cryptography project. `https://csrc.nist.gov/projects/post-quantum-cryptography`.
4. *Alagic G., Apon D., Cooper D., et al.* Status Report on the Third Round of the NIST Post-Quantum Cryptography Standardization Process. US Department of Commerce, NIST, 2022.
5. *Ajtai M.* Generating hard instances of lattice problems // Proc. 28th Ann. ACM Symp. STOC'96. 1996. P. 99–108.
6. *Hoffstein J., Pipher J., and Silverman J. H.* NTRU: A ring-based public key cryptosystem // LNCS. 1998. V. 1423. P. 267–288.
7. *Lenstra A. K., Lenstra H. W., and Lovász L.* Factoring polynomials with rational coefficients // Math. Ann. 1982. V. 261. No. 4. P. 515–534.
8. *Gama N., Nguyen P. Q., and Regev O.* Lattice enumeration using extreme pruning // LNCS. 2010. V. 6110. P. 257–278.
9. *Chen Y. and Nguyen P. Q.* BKZ 2.0: Better lattice security estimates // LNCS. 2011. V. 7073. P. 1–20.
10. *Becker A., Ducas L., Gama G., and Laarhoven T.* New directions in nearest neighbor searching with applications to lattice sieving // Proc. 27th Ann. ACM-SIAM Symp. on Discrete Algorithms. SIAM, 2016. P. 10–24.
11. *Herold G., Kirshanova E., and Laarhoven T.* Speed-ups and time-memory trade-offs for tuple lattice sieving // LNCS. 2018. V. 10769. P. 407–436.
12. *Малыгина Е. С., Куценко А. В., Новоселов С. А. и др.* Постквантовые криптосистемы: открытые вопросы и существующие решения. Криптосистемы на решётках // Дискретный анализ и исследование операций. 2023. (в печати)
13. *McEliece R. J.* A public key cryptosystem based on algebraic coding theory // DSN Progress Report. 1978. V. 44. P. 114–116.
14. *Minder L. and Shokrollahi A.* Cryptanalysis of the Sidelnikov cryptosystem // LNCS. 2007. V. 4515. P. 347–360.
15. *Berlekamp E., McEliece R., and van Tilborg H.* On the inherent intractability of certain coding problems // IEEE Trans. Inform. Theory. 1978. V. 24. No. 3. P. 384–386.
16. *Misoczki R., Tillich J P., Sendrier N., and Barreto P. S. L. M.* MDPC-McEliece: New McEliece variants from moderate density parity-check codes // IEEE Intern. Symp. Inform. Theory. Istanbul, Turkey, 2013. P. 2069–2073.
17. *Aguilar-Melchor C., Blazy O., Deneuville J. C., et al.* Efficient encryption from random quasi-cyclic codes // IEEE Trans. Inform. Theory. 2018. V. 64. No. 5. P. 3927–3943.
18. *Малыгина Е. С., Куценко А. В., Новоселов С. А. и др.* Постквантовые криптосистемы: открытые вопросы и существующие решения. Криптосистемы на изогениях и кодах, исправляющих ошибки // Дискретный анализ и исследование операций. 2023 (в печати).
19. *De Feo L., Jao D., and Plût J.* Towards Quantum-Resistant Cryptosystems from Supersingular Elliptic Curve Isogenies. Cryptology ePrint Archive. Paper 2011/506. `https://eprint.iacr.org/2011/506`.
20. *Castryck W. and Decru T.* An Efficient Key Recovery Attack on SIDH (preliminary version). Cryptology ePrint Archive. Paper 2022/975. `https://eprint.iacr.org/2022/975`.

21. *Castryck W., Lange T., Martindale C., et al.* CSIDH: An Efficient Post-Quantum Commutative Group Action. // Cryptology ePrint Archive. Paper 2018/383. https://eprint.iacr.org/2018/383.

22. *Chi-Domínguez J.-J. and Rodríguez-Henríquez F.* Optimal strategies for CSIDH // Adv. Math. Commun. 2022. V. 16. No. 2. P. 383–411.

УДК 519.7 DOI 10.17223/2226308X/16/17

АЛГЕБРАИЧЕСКИЙ КРИПТОАНАЛИЗ 9 РАУНДОВ НИЗКОРЕСУРСНОГО БЛОЧНОГО ШИФРА SIMON32/64[1]

Е. А. Маро, О. С. Заикин

Рассматривается низкоресурсный блочный шифр Simon32/64 из семейства Simon. Полная версия этого шифра состоит из 32 раундов. Задачи криптоанализа для 8 раундов Simon32/64 были неоднократно решены с помощью SAT-подхода, т. е. путём сведения к проблеме булевой выполнимости и использования SAT-решателей. Для 9 раундов задача все ещё является сложной для SAT-подхода. Построена SAT-кодировка криптоанализа 9-раундовой версии Simon32/64. Сформированы три класса тестов в зависимости от способа выбора открытого текста. С помощью параллельного SAT-решателя во всех случаях удалось успешно решить задачи криптоанализа при условии, что 16 из 64 битов секретного ключа известны.

Ключевые слова: *низкоресурсный блочный шифр, семейство шифров Simon, алгебраический криптоанализ, SAT-решатель.*

Введение

Низкоресурсные шифры получают всё большее распространение в малоресурсных IoT- устройствах, таких, как RFID-считыватели, сенсоры, индикаторы, датчики, контроллеры. Задача обеспечения безопасности хранения и передачи данных для подобных устройств может быть решена применением специальных криптографических алгоритмов, рассчитанных на эффективную реализацию в условиях ограниченных вычислительных ресурсов. Одним из алгоритмов низкоресурсного шифрования, принятым в качестве международного стандарта для систем связи по радиоинтерфейсу, является семейство симметричных блочных шифров Simon (ISO/IEC 29167-21:2018) [1]. Проектирование низкоресурсных шифров с условием их применимости в малоресурсных системах обуславливает необходимость сокращения типов используемых преобразований и, как следствие, приводит к низкой мультипликативной сложности (низкой нелинейности), что обосновывает важность исследования стойкости низкоресурсных алгоритмов шифрования к алгебраическим методам криптоанализа. В алгебраических методах анализа используется представление криптографического преобразования в виде системы нелинейных уравнений, связывающих наборы известных данных и секретный ключ шифрования, и применяются различные алгоритмы поиска решений системы уравнений. В качестве распространённых подходов к решению описывающих шифры систем уравнений можно выделить:

— метод линеаризации и его расширенные варианты (eXtended Linearization (XL), eXtended Sparse Linearization (XSL), Fix eXtended Linearization (FXL), ElimLin);

[1]Заикин О. С. выполнил свою часть работы за счёт субсидии Минобрнауки России в рамках проекта № 121041300065-9.

— метод сведения к задаче выполнимости булевых формул (SAT-задаче) с применением эффективных SAT-решателей.

Анализ научных исследований, посвящённых алгебраическому криптоанализу низкоресурсных шифров семейства Simon, показал потенциальную перспективность подходов, основанных на сведении к задаче булевой выполнимости формул и последующем использовании SAT-решателей [2–5].

1. Семейство низкоресурсных шифров Simon-$2n/nm$

Низкоресурсный блочный шифр Simon предложен в [6] и представляет собой семейство низкоресурсных шифров LRX-архитектуры с выбираемыми длинами блока текста и ключа шифрования, обеспечивающими возможность оптимизации параметров под конкретную аппаратную реализацию. Шифр Simon (рис. 1) основан на схеме сети Фейстеля, в раундовой функции которой используются операции циклического сдвига влево на 1, 2 и 8 бит, побитового умножения (AND) и побитового сложения (XOR).

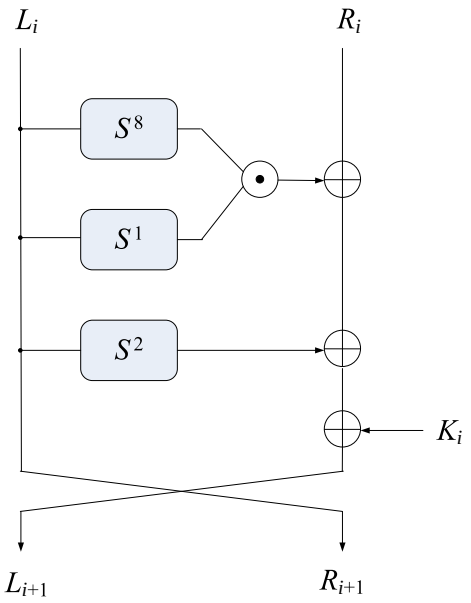

Рис. 1. Структура раундовой функции шифра Simon

Шифр Simon с n-битовым словом обозначается Simon-$2n$, где n может принимать значения 16, 24, 32, 48 и 64 бита. Для задания используемой длины ключа m шифра Simon применяется запись Simon-$2n/mn$. Соотношения длины блока и длины ключа шифрования для семейства Simon приведены в табл. 1.

Раундовое преобразование шифра Simon имеет вид

$$F(L_{i-1}) = (L_{i-1} \lll 1) \wedge (L_{i-1} \lll 8) \oplus (L_{i-1} \lll 2).$$

Таблица 1

Соотношение длины блока и длины ключа шифрования для семейства Simon

Длина блока $2n$	Длина ключа mn	Размер слова n	Число слов в ключе m	Значение константы z	Число раундов T
32	64	16	4	z_0	32
48	72	24	3	z_0	36
	96		4	z_1	36
64	96	32	3	z_2	42
	128		4	z_3	44
96	96	48	2	z_2	52
	144		3	z_3	54
128	128	64	2	z_2	68
	192		3	z_3	69
	256		4	z_4	72

Левая и правая половины входного вектора $L_i||R_i$ i-го раунда шифра Simon могут быть представлены в виде формул

$$L_i = R_{i-1} \oplus F(L_{i-1}) \oplus K_{i-1};$$
$$R_i = L_{i-1}.$$

Первые m раундовых ключей шифра Simon соответствуют секретному ключу шифрования K_0, K_1, …, K_{m-1}, а последующие раундовые ключи в зависимости от параметра m вырабатываются по следующей формуле:

$$K_{i+m} = \begin{cases} c \oplus (z_j)_i \oplus K_i \oplus (I \oplus S^{-1}) S^{-3} K_{i+1}, & m = 2, \\ c \oplus (z_j)_i \oplus K_i \oplus (I \oplus S^{-1}) S^{-3} K_{i+2}, & m = 3, \\ c \oplus (z_j)_i \oplus K_i \oplus (I \oplus S^{-1}) (S^{-3} K_{i+3} \oplus K_{i+1}), & m = 4. \end{cases}$$

2. SAT-кодировка шифра Simon 32/64

На основе результатов алгебраического анализа 8-раундового шифра Simon 32/64 [2] в качестве исследуемого низкоресурсного блочного шифра выбран 9-раундовый Simon 32/64. Описание T раундов шифра Simon $2n/nm$ содержит nT квадратичных уравнений с числом неизвестных $n(T-2)+nm$. В ходе описания 9-раундовой версии Simon 32/64 сформировано 144 квадратичных уравнения со 176 переменными. Структура системы уравнений шифра Simon $2n/nm$ при алгебраическом анализе подробно представлена в [7]. Полученная система может быть задана формулой

$$\begin{cases} k_{0,0} \oplus x_{2,0} \oplus 1 = 0, \\ k_{0,1} \oplus x_{2,1} = 0, \\ k_{0,2} \oplus x_{2,2} \oplus 1 = 0, \\ \dots \\ k_{1,0} \oplus x_{2,2} \oplus x_{2,1} * x_{2,8} \oplus x_{3,0} = 0, \\ \dots \\ k_{0,3} \oplus k_{1,2} \oplus k_{1,4} \oplus k_{1,5} \oplus k_{1,6} \oplus k_{1,14} \oplus k_{1,15} \oplus k_{2,2} \oplus k_{2,6} \oplus k_{2,7} \oplus k_{2,9} \oplus k_{2,13} \oplus \\ k_{2,15} \oplus k_{3,0} \oplus k_{3,6} \oplus k_{3,7} \oplus k_{3,8} \oplus k_{3,9} \oplus k_{3,11} \oplus k_{3,12} \oplus k_{3,15} \oplus x_{8,15} = 0, \end{cases}$$

где $k_{i,j}$ соответствует j-му биту i-го раундового ключа.

Для проведения экспериментов были составлены системы нелинейных уравнений, описывающих шифрование 9-раундового Simon32/64 на трёх различных значениях секретного ключа: $Key_1 = (\texttt{0xb2fe,0x7c97,0xa734,0x8a7f})$; $Key_2 = (\texttt{0xe5e1,0x3e5c,0xfe34,0x7a47})$; $Key_3 = (\texttt{0xd8a6,0x28f0,0x4c35,0xac81})$.

Рассмотрены следующие сценарии генерации системы, в каждом из которых шифруется 64 бита открытого текста:

— сценарий «rand» — система уравнений формируется для двух случайных значений открытого текста (использовались три пары текстов: $P_1 = \texttt{0x65656877}$ и $P_2 = \texttt{0x34895467}$; $P_1 = \texttt{0x30d7f42b}$ и $P_2 = \texttt{0x591a97c5}$; $P_1 = \texttt{0xd19a13c4}$ и $P_2 = \texttt{0x0b597351}$);

— сценарий «nulldist» — система уравнений формируется для двух выбранных значений открытого текста: первый текст принимает нулевое значение ($P_1 = \texttt{0x0}$), второй отличается только в младшем бите ($P_2 = \texttt{0x1}$);

— сценарий «randdist» — система уравнений формируется для двух выбранных значений открытого текста: первый текст выбирается произвольно, второй текст отличается только в младшем бите (использовались три пары текстов: $P_1 = \texttt{0x65656877}$ и $P_2 = \texttt{0x65656876}$; $P_1 = \texttt{0x30d7f42b}$ и $P_2 = \texttt{0x30d7f42a}$; $P_1 = \texttt{0xd19a13c4}$ и $P_2 = \texttt{0xd19a13c5}$).

Конвертация системы нелинейных уравнений в задачу выполнимости булевых формул в конъюнктивной нормальной форме (КНФ) осуществлялась с среде SageMath с использованием инструмента ANF2CNFConverter. Итоговые SAT-задачи для всех девяти описаний шифрования 9-раундового Simon32/64 содержат 1 216 литералов и 20 512 дизъюнктов.

3. Криптоанализ 9 раундов Simon32/64 с помощью параллельного SAT-решателя

Основным полным алгоритмом решения SAT-задач является Conflict-Driven Clause Learning (CDCL) [8]. На первом этапе экспериментов на каждой из девяти КНФ, построенных в п. 2, был запущен современный CDCL-решатель Kissat [9] версии 3.0 с лимитом времени 24 ч на персональном компьютере, оснащённом процессором AMD Ryzen 3900x. В результате решатель не смог решить ни одну из задач в установленный лимит времени.

На втором этапе был применён подход Cube-and-Conquer, согласно которому на первой стадии задача разбивается на более простые подзадачи с помощью lookahead-решателя, а затем на них запускается CDCL-решатель [10]. Этот подход обычно применяется для решения задач комбинаторики и вычислительной геометрии, но в [11] с его помощью впервые решены задачи криптоанализа. В [11] предложен также алгоритм построения прогнозов времени решения SAT-задач. В этом алгоритме варьируется значение основного параметра первой стадии Cube-and-Conquer и для каждого значения строится прогноз на основе решения выборки подзадач. С помощью этого алгоритма для каждой из девяти рассматриваемых SAT-задач построен прогноз времени решения, в качестве CDCL-решателя для второй фазы Cube-and-Conquer использован Kissat. Согласно прогнозам, для решения этих задач данным подходом требуется нереально много времени. Поэтому для каждой задачи был рассмотрен упрощённый вариант, в котором известны правильные значения первых 16 (из 64) переменных, кодирующих секретный ключ.

Эксперименты проведены на вычислительном кластере «Академик В. М. Матросов» [12]. Использовались 5 узлов, в состав каждого из которых входит два 18-ядер-

ных процессора Intel Xeon E5-2695 v4. Итого в каждом запуске были задействованы 180 ядер. В результате для каждой из девяти КНФ был найден выполняющий набор, а соответствующие 48 значений переменных совпали со значениями искомого секретного ключа. В табл. 2 для каждой КНФ указано время решения, процент решённых подзадач на момент нахождения выполняющего набора, а также прогноз времени, необходимого на решение всех подзадач в худшем случае.

Таблица 2

Время решения задач криптоанализа Simon32/64, в которых известны 16 из 64 битов секретного ключа

Задача	Время	Решено подзадач, %	Прогноз на 100 %
rand-Key_1	10 ч 37 мин	7,03	151 ч 6 мин
rand-Key_2	29 ч 33 мин	31,57	93 ч 36 мин
rand-Key_3	12 ч 30 мин	17,17	72 ч 50 мин
nulldist-Key_1	33 ч 52 мин	91	37 ч 13 мин
nulldist-Key_2	4 ч 10 мин	54,35	7 ч 40 мин
nulldist-Key_3	1 ч	10,82	9 ч 18 мин
randdist-Key_1	2 ч 54 мин	75,21	3 ч 51 мин
randdist-Key_2	1 ч 18 мин	46,08	2 ч 49 мин
randdist-Key_3	34 мин	3,6	16 ч 4 мин

Согласно полученным результатам, задачи с выбранным открытым текстом существенно проще для Cube-and-Conquer, чем задачи со случайными открытым текстом.

Заключение

Впервые осуществлён алгебраический криптоанализ 9-раундового блочного шифра Simon32/64 с 64 битами шифртекста и открытого текста в предположении, что известны первые 16 из 64 битов секретного ключа. На суперкомпьютере время решения варьируется от получаса до 34 ч в зависимости от того, как выбран открытый текст. На основе представленных результатов можно сделать вывод, что, развивая используемый подход, возможно осуществить криптоанализ 9 раундов при полностью неизвестном секретном ключе.

ЛИТЕРАТУРА

1. ISO/IEC 29167-21:2018. Information technology — Automatic identification and data capture techniques. P. 21: Crypto suite SIMON security services for air interface communications. `https://www.iso.org/standard/70388.html`.
2. *Courtois N., Mourouzis T., Song G., et al.* Combined algebraic and truncated differential cryptanalysis on reduced-round Simon // Proc. 11th Intern. Conf. SECRYPT. Vienna, Austria, 2014. P. 399–404.
3. *Raddum H.* Algebraic analysis of the Simon block cipher family // LNCS. 2015. V. 9230. P. 157–169
4. *Choo D., Soos M., Chai K. M. A., and Meel K. S.* Bosphorus: Bridging ANF and CNF solvers // Proc. DATE 2019. P. 468–473.
5. *Yeo S. L., Le D., and Khoo K.* Improved algebraic attacks on lightweight block ciphers // J. Cryptogr. Eng. 2011. V. 11. P. 1–19.
6. *Beaulieu R., Shors D., Smith J., et al.* The SIMON and SPECK lightweight block ciphers // Proc. DAC 2015. P. 175:1–175:6.

7. *Куценко А. В., Атутова Н. Д., Зюбина Д. А. и др.* Алгебраический криптоанализ низкоресурсных шифров Simon и Speck // Прикладная дискретная математика. Приложение. 2021. № 14. С. 84–91.
8. *Marques-Silva J. P. and Sakallah K. A.* GRASP: a search algorithm for propositional satisfiability // IEEE Trans. Computers. 1999. V. 48(5). P. 506–521.
9. *Biere A. and Fleury M.* Gimsatul, IsaSAT and Kissat entering the SAT Competition 2022 // Proc. SAT Competition 2022 — Solver and Benchmark Descriptions. P. 10–11.
10. *Heule M. J. H., Kullmann O., Wieringa S., and Biere A.* Cube and Conquer: guiding CDCL SAT solvers by lookaheads // LNCS. 2012. V. 7261. P. 50–65.
11. *Zaikin O.* Inverting 43-step MD4 via cube-and-conquer // Proc. IJCAI-ECAI. Vienna, Austria, 2022. P. 1894–1900.
12. Иркутский суперкомпьютерный центр СО РАН. `http://hpc.icc.ru`.

УДК 519.7 DOI 10.17223/2226308X/16/18

О РАЗНОСТЯХ ПО МОДУЛЮ 2^n, С ВЫСОКОЙ ВЕРОЯТНОСТЬЮ ПРОХОДЯЩИХ ЧЕРЕЗ ARX-ПРЕОБРАЗОВАНИЕ[1]

А. С. Мокроусов, Н. А. Коломеец

Исследуются высокие значения разностной характеристики по модулю 2^n для преобразования $(x \oplus y) \lll r$, где $x, y \in \mathbb{Z}_2^n$ и $1 \leqslant r < n$. Они интересны в контексте разностного криптоанализа шифров, использующих сложение по модулю 2^n, побитовый XOR и циклический сдвиг битов на r позиций как базовые операции. Описаны все разности с вероятностью больше $1/4$ с точностью до симметрий аргументов. Возможными значениями вероятности при этом условии являются $1/3 + 4^{2-i}/6$ для всех $i \in \{1, \ldots, n\}$, что совпадает с аналогичными вероятностями для преобразования $x \oplus y$. Описаны разности, на которых достигается каждое из значений, и подсчитано их количество. Установлено, что общее число разностей с приведёнными вероятностями равно $48n - 68$ при $n \geqslant 2$.

Ключевые слова: *ARX, разностные характеристики, XOR, сложение по модулю, циклический сдвиг битов.*

Введение

ARX — класс криптографических примитивов, базовыми операциями в которых являются сложение по модулю 2^n ($\boxplus$), побитовое «исключающее или» (XOR, $\oplus$) и циклический сдвиг битов на r позиций в сторону старших разрядов ($\lll r$) — см., например, FEAL [1], Speck [2], Salsa20 [3] и ChaCha [4].

Важно, чтобы шифр был стойким к разностному криптоанализу [5], основа которого — изучение преобразования разностей алгоритмом шифрования. Ключевым шагом при реализации метода является вычисление разностных характеристик и их максимальных значений с учётом фиксаций аргументов. Разности могут рассматриваться, например, относительно $\oplus$, $\boxplus$, какой-либо другой операции или вообще быть смешанными [6–9].

Будем рассматривать аргументы всех операций как элементы $\mathbb{Z}_2^n$, т. е. пространства двоичных векторов размерности n. С вектором $x = (x_1, \ldots, x_n) \in \mathbb{Z}_2^n$ мы ассоциируем целое число

$$2^{n-1}x_1 + 2^{n-2}x_2 + \cdots + 2^0 x_n,$$

[1]Работа выполнена при поддержке Математического центра в Академгородке, соглашение с Министерством науки и высшего образования Российской Федерации № 075–15–2022–282.

т. е. биты слева соответствуют старшим разрядам числа. Под $x \boxplus y$, где y также из $\mathbb{Z}_2^n$, подразумеваем сложение по модулю 2^n ассоциированных с ними чисел, $-x$ является обратным к ассоциированному с x числу по модулю 2^n. Как правило, записывать элементы $\mathbb{Z}_2^n$ будем как целые числа, например 0, 2^{n-1} и т. д.

Для преобразований $x \oplus y$ и $(x \oplus y) \lll r$, где $1 \leqslant r < n$, разностные характеристики относительно сложения по модулю 2^n определяются следующим образом:

$$\mathrm{adp}^{\oplus}(\alpha, \beta \to \gamma) = \frac{1}{4^n} \#\{x, y \in \mathbb{Z}_2^n \,:\, (x \boxplus \alpha) \oplus (y \boxplus \beta) = \gamma \boxplus (x \oplus y)\},$$

$$\mathrm{adp}^{\mathrm{XR}}(\alpha, \beta \xrightarrow{r} \gamma) = \frac{1}{4^n} \#\{x, y \in \mathbb{Z}_2^n \,:\, ((x \boxplus \alpha) \oplus (y \boxplus \beta)) \lll r = \gamma \boxplus ((x \oplus y) \lll r)\}.$$

Отметим, что свойства $\mathrm{adp}^{\oplus}$ исследованы в [10–12], а $\mathrm{adp}^{\mathrm{XR}}$ — в [13, 14]. Определим следующие преобразования на тройках $(\alpha, \beta, \gamma) \in \mathbb{Z}_2^n \times \mathbb{Z}_2^n \times \mathbb{Z}_2^n$:

— $\mathcal{T}_{\alpha\beta} : (\alpha, \beta, \gamma) \mapsto (\beta, \alpha, \gamma)$, аналогично определяются $\mathcal{T}_{\alpha\gamma}$ и $\mathcal{T}_{\beta\gamma}$;
— $\mathcal{I} : (\alpha, \beta, \gamma) \mapsto (\alpha \oplus 2^{n-1}, \beta \oplus 2^{n-1}, \gamma)$;
— $\mathcal{S}_\alpha : (\alpha, \beta, \gamma) \mapsto (-\alpha, \beta, \gamma)$, аналогично определяются $\mathcal{S}_\beta$ и $\mathcal{S}_\gamma$.

Обозначим через $\mathcal{E}$ и $\mathcal{E}'$ группы преобразований, порождённые множествами

$$\{\mathcal{T}_{\alpha\beta}, \mathcal{T}_{\beta\gamma}, \mathcal{T}_{\alpha\gamma}, \mathcal{I}, \mathcal{S}_\alpha, \mathcal{S}_\beta, \mathcal{S}_\gamma\} \quad \text{и} \quad \{\mathcal{T}_{\alpha\beta}, \mathcal{I}, \mathcal{S}_\alpha, \mathcal{S}_\beta, \mathcal{S}_\gamma\}$$

замыканием относительно композиции. Они интересны тем, что сохраняют значения $\mathrm{adp}^{\oplus}$ и $\mathrm{adp}^{\mathrm{XR}}$ соответственно [12, 14].

Утверждение 1 [12]. Для любого $\xi \in \mathcal{E}$ и любых $\alpha, \beta, \gamma \in \mathbb{Z}_2^n$ справедливо $\mathrm{adp}^{\oplus}(\alpha, \beta \to \gamma) = \mathrm{adp}^{\oplus}(\alpha', \beta' \to \gamma')$, где $(\alpha', \beta', \gamma') = \xi(\alpha, \beta, \gamma)$.

Утверждение 2 [14]. Для любого $\xi \in \mathcal{E}'$ и любых $\alpha, \beta, \gamma \in \mathbb{Z}_2^n$ справедливо $\mathrm{adp}^{\mathrm{XR}}(\alpha, \beta \xrightarrow{r} \gamma) = \mathrm{adp}^{\mathrm{XR}}(\alpha', \beta' \xrightarrow{r} \gamma')$, где $(\alpha', \beta', \gamma') = \xi(\alpha, \beta, \gamma)$.

Для $\alpha, \beta, \gamma \in \mathbb{Z}_2^n$ введём также $\mathcal{E}(\alpha, \beta, \gamma) = \{\xi(\alpha, \beta, \gamma) : \xi \in \mathcal{E}\}$. Аналогичное обозначение будем использовать и для $\mathcal{E}'$.

Определим бесконечную последовательность рациональных чисел $p_1, p_2, p_3, \ldots$ следующим образом:

$$p_i = \frac{1}{3} + \frac{4^{2-i}}{6}, \text{ где } i \geqslant 1.$$

Её также можно задать рекуррентным соотношением $p_{i+1} = p_i - 2 \cdot 4^{-i}$.

В работе [15] для разностной характеристики $\mathrm{adp}^{\oplus}$ описаны все значения, превышающие 1/4. Приведём соответствующую теорему, используя уже введённые обозначения:

Теорема 1 [15]. Пусть $\alpha, \beta, \gamma \in \mathbb{Z}_2^n$ и $P = \mathrm{adp}^{\oplus}(\alpha, \beta \to \gamma) > 1/4$. Тогда $P \in \{p_1, \ldots, p_n\}$. Более того, $P = p_i \iff (\alpha, \beta, \gamma) \in \mathcal{E}(2^{n-i}, 2^{n-i}, 0)$, где $1 \leqslant i \leqslant n$.

Известны и количества разностей, на которых достигается каждое из значений.

Следствие 1 [15]. Обозначим через C_i количество троек $(\alpha, \beta, \gamma) \in \mathbb{Z}_2^n \times \mathbb{Z}_2^n \times \mathbb{Z}_2^n$, таких, что $\mathrm{adp}^{\oplus}(\alpha, \beta \to \gamma) = p_i$, где $i \in \{1, \ldots, n\}$. Тогда:

— $C_1 = 4$, $C_2 = 24$;
— $C_3 = C_4 = \ldots = C_n = 48$.

Нетрудно подсчитать общее количество таких троек.

Следствие 2. Пусть $n \geqslant 2$. Тогда имеется ровно $48n - 68$ троек $(\alpha, \beta, \gamma) \in \mathbb{Z}_2^n \times \times \mathbb{Z}_2^n \times \mathbb{Z}_2^n$, таких, что $\mathrm{adp}^{\oplus}(\alpha, \beta \to \gamma) > 1/4$.

В данной работе описаны все тройки, значение $\mathrm{adp}^{\mathrm{XR}}$ на которых больше $1/4$. Полученные результаты представлены аналогично теореме 1 и её следствиям. Другими словами, мы рассматриваем, как добавление циклического сдвига к операции $x \oplus y$ влияет на высокие значения её разностной характеристики по модулю 2^n.

Значения $\mathrm{adp}^{\mathrm{XR}}$, превышающие $1/4$

Следующая теорема доказана с использованием формулы для $\mathrm{adp}^{\mathrm{XR}}$, предложенной в [14].

Теорема 2. Пусть $\alpha, \beta, \gamma \in \mathbb{Z}_2^n$ и $P = \mathrm{adp}^{\mathrm{XR}}(\alpha, \beta \xrightarrow{r} \gamma) > 1/4$, где $1 \leqslant r < n$. Тогда $P \in \{p_1, \ldots, p_n\}$. Более того, $P = p_i$, если и только если тройку (α, β, γ) можно привести преобразованиями из $\mathcal{E}'$ к одной из следующих:

1) $(2^{n-i}, 2^{n-i}, 0)$ при $1 \leqslant i \leqslant n$;
2) $(0, 2^{n-i+1}, 2^{r-i+1})$ при $2 \leqslant i \leqslant r+1$;
3) $(2^{n-r-i+1}, 0, 2^{n-i+1})$ при $2 \leqslant i \leqslant n-r+1$.

Замечание 1. Тройки 1, 2 и 3 из условия теоремы с учётом ограничений на i не приводятся друг к другу преобразованиями из $\mathcal{E}'$. Более того, они не приводятся друг к другу и преобразованиями из $\mathcal{E}$, за исключением случая приведения тройки 2 к тройке 3 при $n = 2r$.

Отметим, что $i = 2$ удовлетворяет ограничениям на все тройки при $n \geqslant 2$. Таким образом, значение $p_2 = 1/2$ гарантированно достигается на трёх тройках, неприводимых друг к другу преобразованиями из $\mathcal{E}'$. Перейдём к количеству троек.

Следствие 3. Для троек вида 1, 2 и 3 из условия теоремы справедливо:

— $|\mathcal{E}'(2^{n-i}, 2^{n-i}, 0)| = 8$ при $2 < i \leqslant n$, а также равно 2 и 4 при $i = 1$ и $i = 2$ соответственно;
— $|\mathcal{E}'(0, 2^{n-i+1}, 2^{r-i+1})| = 16$ при $2 < i \leqslant r+1$, а также равно 4 при $i = 2$;
— $|\mathcal{E}'(2^{n-r-i+1}, 0, 2^{n-i+1})| = 16$ при $2 < i \leqslant n-r+1$, а также равно 8 при $i = 2$.

Тогда для каждого из рассматриваемых значений вероятности количества троек будут следующими:

Следствие 4. Пусть C_i — количество троек $(\alpha, \beta, \gamma) \in \mathbb{Z}_2^n \times \mathbb{Z}_2^n \times \mathbb{Z}_2^n$, таких, что $\mathrm{adp}^{\mathrm{XR}}(\alpha, \beta \xrightarrow{r} \gamma) = p_i$, где $i \in \{1, \ldots, n\}$. Тогда C_i равно:

— 2, если $i = 1$;
— 16, если $i = 2$;
— 40, если $2 < i \leqslant \min\{r+1,\ n-r+1\}$;
— 24, если $\min\{r+1,\ n-r+1\} < i \leqslant \max\{r+1,\ n-r+1\}$;
— 8, если $\max\{r+1,\ n-r+1\} < i \leqslant n$.

В частности, $C_3 = C_4 = \cdots = C_n = 24$ при $r = 1$ и $r = n-1$.

Нетрудно подсчитать и общее количество троек.

Следствие 5. Пусть $n \geqslant 2$. Тогда имеется ровно $24n - 30$ троек $(\alpha, \beta, \gamma) \in \mathbb{Z}_2^n \times \times \mathbb{Z}_2^n \times \mathbb{Z}_2^n$, таких, что $\mathrm{adp}^{\mathrm{XR}}(\alpha, \beta \xrightarrow{r} \gamma) > 1/4$.

Заключение

Таким образом, значения $\mathrm{adp}^{\oplus}$ и $\mathrm{adp}^{\mathrm{XR}}$, превышающие $1/4$, совпадают. Однако в случае функции $(x \oplus y) \lll r$ разности, на которых достигаются данные значения, нельзя описать одним набором аргументов (с точностью до симметрий). Кроме того, суммарное количество троек, вероятность которых больше $1/4$, почти в 2 раза меньше при наличии циклического сдвига. При этом, аналогично результатам из [14], минимальные различия получаются при циклическом сдвиге на одну позицию влево или вправо.

Полученные результаты можно применить к более широкому классу преобразований, например к функциям $(x \lll r) \oplus y$, $((x \boxplus y) \lll r) \oplus z$ и т. д., поскольку их разностные характеристики прямо выражаются через значения $\mathrm{adp}^{\mathrm{XR}}$.

ЛИТЕРАТУРА

1. *Shimizu A. and Miyaguch S.* Fast data encipherment algorithm FEAL // LNCS. 1988. V. 304. P. 267–278.
2. *Beaulieu R., Shors D., Smith J., et al.* The SIMON and SPECK Families of Lightweight Block Ciphers. Cryptology Eprint Archive. 2013. `https://eprint.iacr.org/2013/404`.
3. *Bernstein D. J.* Salsa20 Specification. eSTREAM Project Algorithm Description. `http://www.ecrypt.eu.org/stream/salsa20pf.html`. 2005.
4. *Bernstein D. J.* ChaCha, a variant of Salsa20 // Workshop Record of SASC. 2008. V. 8. No. 1. P. 3–5.
5. *Biham E. and Shamir A.* Differential cryptanalysis of DES-like cryptosystems // J. Cryptology. 1991. V. 4. No. 1. P. 3–72.
6. *Biryukov A. and Velichkov V.* Automatic search for differential trails in ARX ciphers // LNCS. 2014. V. 8366. P. 227–250.
7. *Leurent G.* Analysis of differential attacks in ARX constructions // LNCS. 2012. V. 7658. P. 226–243.
8. *Малышев Ф. М.* Вероятностные характеристики разностных и линейных соотношений для неоднородной линейной среды // Математические вопросы криптографии. 2019. Т. 10. № 1. С. 41–72.
9. *Малышев Ф. М.* Разностные характеристики основных операций ARX-шифров // Математические вопросы криптографии. 2020. Т. 11. № 4. С. 97–105.
10. *Lipmaa H. and Moriai S.* Efficient algorithms for computing differential properties of addition // LNCS. 2001. V. 2355. P. 336–350.
11. *Mouha N., Velichkov V., De Cannière C., and Preneel B.* The differential analysis of S-functions // LNCS. 2010. V. 6544. P. 36–56.
12. *Mouha N., Kolomeec N., Akhtiamov D., et al.* Maximums of the additive differential probability of Exclusive-Or // IACR Trans. Symmetric Cryptology. 2021. No. 2. P. 292–313.
13. *Velichkov V., Mouha N., De Cannière C., and Preneel B.* The additive differential probability of ARX // LNCS. 2011. V. 6733. P. 342–358.
14. *Kolomeec N., Sutormin I., Bykov D., et al.* On Additive Differential Probabilities of the Composition of Bitwise Exclusive-OR and a Bit Rotation. arXiv preprint. 2023. `https://arxiv.org/abs/2303.04097`.
15. *Мокроусов А. С.* Вычисление разностных характеристик для сложения k чисел по модулю 2^n // Прикладная дискретная математика. Приложение. 2022. № 15. С. 54–57.

УДК 003.26 DOI 10.17223/2226308X/16/19

НИЗКОРЕСУРСНАЯ СИММЕТРИЧНАЯ КРИПТОГРАФИЯ: ПРИНЦИПЫ, ПОДХОДЫ И КОМПРОМИССЫ

С. П. Панасенко

Настоящая работа является обзорной и посвящена вопросам разработки и стандартизации низкоресурсных симметричных криптографических алгоритмов. Кратко описаны причины выделения низкоресурсных криптоалгоритмов в отдельный класс алгоритмов, рассчитанных на применение в устройствах с ограниченными ресурсами. Перечислены основные направления стандартизации низкоресурсных криптоалгоритмов как на уровне общемировых, так и национальных стандартов ряда стран. Описаны основные методы разработки низкоресурсных криптоалгоритмов как на основе существующих алгоритмов общего назначения, так и оригинальных. Сформулированы основные тенденции предполагаемого дальнейшего развития направления низкоресурсной симметричной криптографии: возможный уход от универсальных низкоресурсных криптостандартов в сторону специализированных, усиление требований к низкоресурсным криптографическим алгоритмам в части их криптостойкости и появление новых применений подобных алгоритмов, что может повлиять на методы их разработки и предъявляемые к ним требования.

Ключевые слова: *низкоресурсная криптография, симметричное шифрование, хеширование.*

Введение

Появление и широкое распространение различных видов вычислительных устройств, обладающих небольшими ресурсами (это, в частности, метки радиочастотной идентификации (РЧИ), имплантируемые медицинские приборы, сенсоры и актуаторы различного назначения), привело к выделению класса низкоресурсных криптографических алгоритмов. В качестве примеров вычислителей с крайне ограниченными ресурсами в [1] приводятся 4-битные микроконтроллеры, а также микроконтроллеры с крайне малым объёмом оперативной памяти — от 16 байт.

Несмотря на отсутствие значительных ресурсов у подобных устройств, во многих из областей их применения требуется обеспечение безопасного обмена данными, поэтому применение криптографических алгоритмов в таких устройствах остается востребованным. Однако, как сказано в [2], «в большинстве обычных криптографических стандартов компромиссы между безопасностью, производительностью и требованиями к ресурсам решены путём оптимизации для десктопных и серверных применений, что делает стандарты сложными или невозможными для применения в устройствах с ограниченными ресурсами». Даже в случаях, когда реализация существующих криптостандартов в подобных устройствах возможна, их быстродействие может быть недопустимо низким (или затраты энергии — слишком большими для устройства). И то и другое, в частности, особенно критично для пассивных РЧИ-меток, работающих на энергии электромагнитного поля, генерируемого считывателем, при этом метка должна успеть сформировать и передать ответ за (возможно, короткое) время нахождения в зоне действия считывателя.

Совокупность изложенных факторов и стала причиной выделения низкоресурсных криптоалгоритмов в отдельный класс, а также обоснованием необходимости разработки низкоресурсных криптографических стандартов, описывающих криптоалгоритмы,

специально предназначенные для использования в устройствах с ограниченными ресурсами.

1. Стандартизация низкоресурсных криптоалгоритмов

Как и разработка обычных криптографических стандартов, стандартизация низкоресурсных криптоалгоритмов преследует задачи выбора хорошо изученных и исследованных алгоритмов с определённым уровнем криптографической стойкости и достаточными показателями быстродействия, а также для обеспечения совместимости различных реализаций, основанных на стандартных криптоалгоритмах.

На текущий момент существует семейство международных стандартов ISO/IEC 29192 (Information technology — Security techniques — Lightweight cryptography), описывающих низкоресурсные криптографические алгоритмы и протоколы следующих категорий: алгоритмы симметричного шифрования (блочные и поточные), асимметричные криптоалгоритмы, функции хеширования, коды аутентификации сообщений и протоколы аутентификации. В Евросоюзе и в ряде стран на национальном уровне также ведутся работы по стандартизации низкоресурсных криптоалгоритмов, в частности:

— завершившийся в 2023 г. конкурс Национального института стандартов и технологий США (NIST — National Institute of Standards and Technology) по выбору алгоритмов хеширования и блочного шифрования в режиме аутентифицированного шифрования с использованием ассоциированных данных (AEAD — Authenticated Encryption with Associated Data);
— прошедший с 2013 по 2019 г. более широкий по целям европейский проект CAESAR (Competition for Authenticated Encryption: Security, Applicability, and Robustness) по выбору портфолио алгоритмов аутентифицированного шифрования; одно из трёх направлений конкурса было посвящено выбору низкоресурсных алгоритмов;
— в рамках работы организации по выбору криптостандартов Японии CRYPTREC (Cryptography Research and Evaluation Committees) также проводится оценка низкоресурсных криптоалгоритмов с целью выработки рекомендаций по их использованию в государственных организациях Японии.

Отметим, что перечисленные проекты имеют сильно различающуюся значимость: по результатам конкурса, проведенного NIST, ожидается выход соответствующего стандарта США на основе выбранного семейства алгоритмов Ascon [3] (аналогичные предыдущие стандарты относительно быстро становились де-факто мировыми), тогда как выводы проекта CAESAR являются только рекомендательными.

2. Основные методы разработки низкоресурсных криптоалгоритмов

При разработке низкоресурсных алгоритмов в качестве «строительных блоков» используются как элементы криптоалгоритмов общего назначения, так и специфические для низкоресурсной криптографии преобразования. Можно выделить несколько основных методов разработки симметричных низкоресурсных криптоалгоритмов:

1) Экстенсивное «упрощение» существующих алгоритмов общего назначения — снижение их ресурсоёмкости за счёт уменьшения размерности основных параметров или упрощения используемых в алгоритмах операций, в частности:
    — уменьшение размеров ключей, блоков данных, выходных значений и внутреннего состояния;
    — уменьшение разрядности и количества таблиц замен, количества раундов преобразований.

В качестве примера приведём известный алгоритм DESL — низкоресурсный вариант алгоритма DES, использующий единственную таблицу замен вместо восьми, за счёт чего достигается значительная экономия ресурсов, требуемых для хранения таблиц [4].

2) Упрощение или полное отсутствие процедур расширения ключей для алгоритмов симметричного шифрования; в ряде случаев — ориентация на «прошитые» ключи, т. е. загруженные в устройство изначально и не подлежащие замене. Прошитые ключи, в частности, используются в алгоритме KTANTAN [5].
3) Использование известных (хорошо проверенных с точки зрения криптостойкости при относительно малом потреблении ресурсов) конструкций в качестве внутренних преобразований, в частности:
   — sponge-структура, свойственная алгоритмам хеширования и лежащая в основе, в частности, стандарта хеширования SHA-3 [6]; из низкоресурсных криптоалгоритмов с подобной структурой можно привести в пример алгоритм Xoodyak [7];
   — AES-подобные преобразования (в ряде случаев использование раундов AES целиком); пример: ряд вариантов алгоритма COMET [8];
   — ARX-конструкции (Add-Rotate-XOR), подразумевающие применение простых раундов, включающих операции сложения по модулю, соответствующему размерности операндов, циклического сдвига и XOR; пример: семейство алгоритмов Sparkle [9].
4) Изначальная ориентация на небольшие размеры основных параметров, а также использование внутренних преобразований с незначительными требованиями к ресурсам, в частности:
   — сдвиговых регистров с линейной обратной связью (LFSR);
   — фиксированных или управляемых битовых перестановок.

   В качестве примера такого алгоритма можно привести алгоритм симметричного шифрования Hummingbird с 16-битным размером блока, основанный на применении LFSR [10].

3. Перспективы дальнейшего развития

В результате анализа текущих требований к низкоресурсным криптоалгоритмам, подходов по их созданию и усилий по стандартизации криптоалгоритмов данного класса, а также происходящих в последние годы изменений в данной области можно сделать ряд предположений о дальнейшем развитии низкоресурсной криптографии.

Во-первых, в течение последних лет мы могли наблюдать смещение спектра устройств, в которых применяются низкоресурсные криптоалгоритмы, в сторону всё более широкого их использования в киберфизических системах различного назначения, т. е. расширяется область применения криптоалгоритмов в устройствах, находящихся в нижней части спектра устройств с точки зрения наличия у них вычислительных ресурсов. Это даёт основание предполагать появление следующих трендов в развитии низкоресурсной криптографии, связанных с возможным уходом от универсальных низкоресурсных криптостандартов в сторону специализированных:

— возможное выделение подкласса алгоритмов с крайне малыми требованиями к ресурсам для применения в минимально оснащённых устройствах; показательным является изначальное выделение подобных криптоалгоритмов в отдельный профиль в рамках конкурса NIST [11];

— возможное разделение низкоресурсных криптоалгоритмов на предназначенные для программной или для аппаратной реализации; здесь также уместно вспомнить про один из профилей алгоритмов конкурса NIST, подразумевавший требования только по аппаратной реализации алгоритмов — участников конкурса [11]; фундаментальные различия между направлениями минимизации ресурсоёмкости при программных или аппаратных реализациях криптоалгоритмов были описаны ещё в 2007 г. в работе [12].

Во-вторых, по сравнению с требованиями по криптостойкости, предъявляемыми ранее к низкоресурсным криптоалгоритмам и подразумевавшими наличие компромисса между уровнем криптостойкости и быстродействием при низких требованиях к ресурсам, современные требования к низкоресурсным криптоалгоритмам подразумевают достаточно высокий уровень криптостойкости (но ниже предъявляемых к криптоалгоритмам общего назначения — в частности, одно из требований конкурса NIST подразумевало минимум 2^{112} операций для успешной атаки на 256-битный алгоритм хеширования или на алгоритм шифрования со 128-битным ключом) при сохранении высоких требований к скорости обработки данных. Кроме того, если раньше экспертами могли высказываться предположения (например, в [13]), что атаки с использованием утечек данных по побочным каналам не обязаны приниматься во внимание при оценке низкоресурсного криптоалгоритма, то требования того же конкурса NIST включают в себя наличие у будущих низкоресурсных стандартов встроенных механизмов противодействия подобным атакам.

В-третьих, с появлением новых технологий сфера применения низкоресурсных криптоалгоритмов расширяется, что может в дальнейшем повлиять на требования к таким алгоритмам. В качестве примера относительно нового применения низкоресурсных криптоалгоритмов приведём низкоресурсный блокчейн (lightweight blockchain), одним (но не единственным) из методов построения которого является применение низкоресурсных криптоалгоритмов [14].

ЛИТЕРАТУРА

1. NISTIR 8114. Report on Lightweight Cryptography. `https://doi.org/10.6028/NIST.IR.8114`. 2017.
2. Announcing Request for Nominations for Lightweight Cryptographic Algorithms. Federal Register. 2018/08/27.
3. *Dobraunig C., Eichlseder M., Mendel F., and Schläffer M.* Ascon v1.2. Submission to NIST. `https://ascon.iaik.tugraz.at`. 2021.
4. *Leander G., Paar C., Poschmann A., and Schramm K.* New lightweight DES variants // LNCS. 2007. V. 4593. P. 196–210.
5. *De Cannière C., Dunkelman O., and Knežević M.* KATAN & KTANTAN — A family of small and efficient hardware-oriented block ciphers // LNCS. 2009. V. 5747. P. 272–288.
6. FIPS PUB 202. SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions. `https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.202.pdf`. 2015.
7. *Daemen J., Hoffert S., Peeters M., et al.* Xoodyak, a lightweight cryptographic scheme // IACR Trans. Symmetric Cryptology. 2020. No. S1. P. 60–87.
8. *Gueron S., Jha A., and Nandi M.* COMET: COunter Mode Encryption with authentication Tag. `https://www.isical.ac.in/\~lightweight/comet/cometv1_spec.pdf`. 2019.
9. *Beierle C., Biryukov A., dos Santos L. C., et al.* Lightweight AEAD and hashing using the Sparkle permutation family // IACR Trans. Symmetric Cryptology. 2020. No. S1. P. 208–261.

10. *Engels D., Fan X., Gong G., et al.* Hummingbird: Ultra-lightweight cryptography for resource-constrained devices // LNCS. 2010. V. 6054. P. 3–18.
11. Profiles for the Lightweight Cryptography Standardization Process. Draft White Paper. https://csrc.nist.gov/CSRC/media/Publications/white-paper/2017/04/26/profiles-for-lightweight-cryptography-standardization-process/draft/documents/profiles-lwc-std-proc-draft.pdf. 2017.
12. *Poschmann A.* Lightweight Cryptography from an Engineers Perspective. Workshop ECC. https://maths.ucd.ie/~gmg/ECC2007Talks/poschmann_LWC.pdf. 2007.
13. *Bogdanov A., Knudsen L. R., Leander G., et al.* PRESENT: An ultra-lightweight block cipher // LNCS. 2007. V. 4727. P. 450–466.
14. *Stefanescu D., Montalvillo L., Galán-García P., et al.* A systematic literature review of lightweight blockchain for IoT // IEEE Access. 2022. No. 10. P. 123138–123159.

УДК 519.7 + 004.056.55 DOI 10.17223/2226308X/16/20

ДОПОЛНИТЕЛЬНАЯ ОПТИМИЗАЦИЯ АЛГОРИТМА ПОИСКА ГАРАНТИРОВАННОГО ЧИСЛА АКТИВАЦИЙ В КРИПТОГРАФИЧЕСКИХ XS-СХЕМАХ[1]

Д. Р. Парфенов, А. О. Бахарев

Предложена дополнительная оптимизация алгоритма вычисления гарантированного числа активаций, предполагающая замену вычисления ранга матрицы соответствующей XS-схемы на проверку префикса пути в дереве перебора. Алгоритм был реализован и дал двукратный прирост производительности по сравнению с предыдущим вариантом. С использованием оптимизированной версии алгоритма проведено несколько вычислительных экспериментов, направленных на перебор XS-схем размерности меньше 8 и найдены их гарантированные числа активаций. На основе полученных данных предложена конструкция XS-схем переменной размерности, обладающая оптимальными числами активации.

Ключевые слова: *гарантированное число активаций,* XS*-схемы, разностный криптоанализ.*

При построении блочных шифров используются простые операции, имеющие однобуквенные обозначения, такие, как:

1) R — циклический сдвиг (rotation);
2) X — побитовое исключающее ИЛИ (XOR);
3) A — сложение слов как целых чисел по модулю 2^m (add);
4) L — побитовые логические операции И или ИЛИ;
5) M — умножение слов как элементов поля порядка 2^m;
6) S — взаимно-однозначные подстановки (S-блоки, substitution).

Конструкции блочных шифров, основанные на операциях X и S, называются XS-схемами и покрывают достаточно широкий спектр блочных шифров, включая SM4, Skipjack, сеть Фейстеля. В данной работе рассматривается задача оптимизации алгоритма поиска гарантированного числа активаций [1], позволяющего получить оценку эффективности разностного криптоанализа [2] XS-схем.

Отметим, что алгоритм связан с исследованием линейного кода определённого вида, который строится на основе рассматриваемого шифра. Для каждого шифра из

[1]Работа выполнена при поддержке Математического центра в Академгородке, соглашение с Министерством науки и высшего образования Российской Федерации № 075-15-2022-282.

класса XS-схем (схема находится в первой канонической форме и однозначно задаётся двоичными векторами a и b длины n, где n — размерность XS-схемы) строится матрица $G = G(n, a, b, t)$ размера $(t + n) \times 2t$ (подробнее см. в [3]).

Одним из подходов к поиску гарантированного числа активаций является алгоритм GNA [3]. Основная идея алгоритма — полный перебор разбиений матрицы G на G_0 и G_1 так, чтобы:

1) матрица G_0 содержала $k + 1$ (k изначально известно) пар столбцов из G;
2) $\mathrm{rank}(G_0) < t + n - 1$, где t — число раундов; n — размерность векторов a и b;
3) в матрице G_1 не было ни одного столбца, линейно зависимого от столбцов в G_0.

В работе [4] предложен способ оптимизации разбиения матрицы G, который существенно уменьшает время работы алгоритма. Его основная идея строится на подходе, основанном на методе ветвей и границ, который является развитием метода полного перебора с отсечением подмножеств допустимых решений, заведомо не содержащих оптимальных решений. Однако данный алгоритм может быть улучшен с помощью замены проверки линейной независимости подмножеств столбцов на более простую операцию.

Рассмотрим двоичное дерево, в котором каждый лист задаёт некоторое разбиение пар столбцов матрицы G на матрицы G_0 и G_1. Корень дерева соответствует полностью неопределённому разбиению. При переходе с i-го уровня на $(i + 1)$-й переход к правому потомку соответствует добавлению $(i+1)$-й пары столбцов в матрицу G_0, а переход влево — в G_1. Таким образом, каждый узел дерева соответствует некоторому разбиению подмножества пар столбцов матрицы G на матрицы G_0 и G_1.

Предлагается следующая оптимизация, учитывающая структуру дерева перебора. Будем рассматривать путь в дереве как двоичную строку, в которой на i-й позиции стоит 0, если после i-го узла был сделан шаг вправо, и 1 — если был сделан шаг влево.

Определение 1. Будем называть разбиение пар столбцов матрицы G на подматрицы G_0 и G_1 «*неувеличивающим*», если в G_1 содержится столбец, линейно зависящий от столбцов G_0.

Утверждение 1. Пусть p — путь до узла дерева, который соответствует «неувеличивающему» разбиению. Тогда все пути с суффиксом p также соответствуют «неувеличивающим» разбиениям.

Данный факт является в достаточной мере очевидным. По определению разбиение является «неувеличивающим», если в подматрице G_1 существует столбец, линейно зависимый от столбцов G_0. А поскольку пары столбцов G отличаются друг от друга только битовым сдвигом, то подмножество столбцов, из-за которого «неувеличивающим» является разбиение, соответствующее p, то «неувеличивающим» будет и разбиение, путь до которого содержит p. В силу последовательности обхода дерева вместо проверки вхождения p в текущий путь в качестве подстроки достаточно проверять, что p не является суффиксом пути.

Прирост производительности от данной оптимизации наиболее заметен при вычислении гарантированных чисел активации XS-схем малых размерностей ($n = 2, 3, 4$), поскольку для них невелика и мощность множества пар столбцов, дающих линейную зависимость, благодаря чему практически все вычисления рангов можно заменить на проверку суффикса пути, что намного быстрее.

Прирост производительности для размерности $n = 8$ растёт с увеличением числа раундов и составляет от 5 до 30 %. В то же время для размерности $n = 2$ подход

с подсчётом суффиксов даёт двукратный прирост производительности (табл. 1). Как видно по данным табл. 2, для размерности $n = 2$ вычисление ранга происходит лишь единожды, а все прочие проверки заменяются на проверку суффикса пути, в то время как для размерности $n = 8$ сохраняется достаточно большое количество операций вычисления ранга.

Таблица 1

Время вычисления гарантированного числа активаций для beltWBL-2 ($n = 2$) и beltWBL-8 ($n = 8$) (в секундах)

Кол-во раундов	beltWBL-2		beltWBL-8	
	Сущ. алг.	Предл. алг.	Сущ. алг.	Предл. алг.
30	7,9	4,0	4,5	3,7
31	3,4	1,7	6,3	5,1
32	8,4	4,1	16,7	13,9
33	20,5	10,1	18,3	15,1
34	8,8	4,3	44,9	34,7
35	21,6	10,7	39,3	30,8
36	51,4	25,3	136,2	106,8
37	22,5	11,0	120,2	95,6
38	53,5	26,8	104,8	83,1
39	127,5	63,0	91,6	72,2
40	56,5	27,8	326,0	252,3

Таблица 2

Количество и виды проверок при вычислении гарантированного числа активаций для beltWBL-2 ($n = 2$) и beltWBL-8 ($n = 8$)

Кол-во раундов	beltWBL-2		beltWBL-8	
	Ранг	Суффикс	Ранг	Суффикс
30	1	1022	328	535
31	1	510	431	510
32	1	1022	929	835
33	1	2046	861	1446
34	1	1022	1283	3942
35	1	2046	1187	3478
36	1	4094	2954	9783
37	1	2046	2720	8563
38	1	4094	2470	7373
39	1	8190	2176	6184
40	1	4094	5677	18305

С помощью усовершенствованной версии алгоритма вычисления гарантированного числа активаций осуществлён полный перебор представителей классов регулярных XS-схем для размерностей $2 \leqslant n \leqslant 8$ и вычислены максимальные и минимальные гарантированные числа активаций для различного числа раундов. Полученные данные позволили проверить ранее предложенные преобразования, которые могли бы быть использованы для построения XS-схем, обладающих одинаковыми гарантированными числами активаций. Одна из таких конструкций полностью соответствует экспериментальным данным и, возможно, может быть применена и для больших размерностей.

Гипотеза 1. XS-схемы размерности n, для которых выполняется $\mathrm{wt}(a + b) = n$, имеют одинаковое гарантированное число активаций.

Кроме того, на основе имеющихся данных возможно предложить следующую конструкцию XS-схем переменной размерности. Для размерности $n = 2$ она задаётся векторами $a = (11)$ и $b = (00)$. Далее, с увеличением размерности n, вектор b остаётся нулевым, а на центральную позицию вектора a дописывается 0, если $\lceil n/2 \rceil$ чётное, и 1 — в противном случае. Так, $a = (101)$ для $n = 3$, $a = (1001)$ для $n = 4$, $a = (10101)$ для $n = 5$, $a = (101101)$ для $n = 6$, $a = (1010101)$ для $n = 7$, $a = (10100101)$ для $n = 8$ и т. д. Данная конструкция обладает большими гарантированными числами активации для своих размерностей и превосходит многие известные схемы (например, SMS-4 или SkipjackG-4). В дальнейшем планируется исследовать свойства построенных на её основе шифров и сравнить их с другими конструкциями переменной размерности (например, BeltWBL).

ЛИТЕРАТУРА

1. *Агиевич С. В.* XS-circuits in block ciphers // Матем. вопр. криптогр. 2019. Т. 10. № 2. С. 7–30.
2. *Biham E. and Shamir A.* Differential cryptanalysis of DES-like cryptosystems // J. Cryptology. 1991. V. 4. No. 1. P. 3–72.
3. *Агиевич С. В.* On the guaranteed number of activations in XS-circuits // Матем. вопр. криптогр. 2021. Т. 12. № 2. С. 7–20.
4. *Парфенов Д. Р., Бахарев А. О., Куценко А. В. и др.* Свойства XS-схем, связанные с гарантированным числом активаций // Прикладная дискретная математика. Приложение. 2022. № 15. С. 62–66.

УДК 519.7 DOI 10.17223/2226308X/16/21

АНАЛИЗ МЕТОДОМ БУМЕРАНГА 4-РАУНДОВОГО АЛГОРИТМА ШИФРОВАНИЯ LILLIPUT-TBC-II-256

М. А. Пудовкина, А. М. Смирнов

Алгоритм аутентифицированного шифрования LILLIPUT-AE — участник конкурса Национального института стандартов и технологий США на стандарт низкоресурсного алгоритма шифрования. Основу его составляет блочная шифрсистема LILLIPUT-TBC в режиме OCB с длиной блока 128 бит и длинами ключей шифрования 128, 192, 256 бит, у которой есть две версии LILLIPUT-TBC-I и LILLIPUT-TBC-II. В работе предложена атака на 4-раундовый алгоритм шифрования LILLIPUT-TBC-II-256 с ключом длины 256 бит методом бумеранга. Для построения различителя применяется подход «йо-йо», первоначально использованный в атаке на алгоритм AES. Основной результат получен благодаря одновременному использованию метода бумеранга и свойства матрицы рассеивающего преобразования алгоритма LILLIPUT-TBC-II. Для реализации атаки на 4-раундовый алгоритм LILLIPUT-TBC-II-256 требуется 2^{24} текстов, $2^{24,3}$ бит памяти. Её трудоёмкость равна 2^{180} зашифрований.

Ключевые слова: *низкоресурсный алгоритм шифрования, подход «йо-йо», аутентифицированное шифрование, линейное преобразование, S-бокс, режим OFB, метод бумеранга.*

Алгоритм блочного шифрования LILLIPUT [1], разработанный в 2015 г., основан на расширенном обобщённом преобразовании Фейстеля (Generalized Feistel Networks (EGFN)) [2]. Длина ключа шифрования алгоритма LILLIPUT равна 80 бит, длина блока — 64 бит, число раундов — 30. Подстановки S-бокса являются 4-битными. Реду-

цированный алгоритм LILLIPUT анализировался интегральным методом [3], методом невозможных разностей [4] и разностным методом [5]. Обнаруженные в ходе анализа слабости привели к синтезу семейства алгоритмов аутентифицированного шифрования LILLIPUT-AE [6] для участия в конкурсе NIST на стандарт низкоресурсного алгоритма шифрования США. Основой семейства LILLIPUT-AE являются блочная шифрсистема LILLIPUT-TBC в режиме OCB [7] (Offset CodeBook) с длиной блока 128 бит и длинами ключей шифрования 128, 192 и 256 бит, где буквы TBC означают модифицированные алгоритм шифрования LILLIPUT с 8-битной подстановкой S-бокса и алгоритм развертывания ключа (Tweakable Block Cipher). В [8] на поданную на конкурс первую версию семейства LILLIPUT-AE приведена практическая атака, позволяющая подделывать сообщения из-за выявленных слабостей модифицированного алгоритма развёртывания ключа. В результате авторами семейства LILLIPUT-AE предложена вторая версия блочной шифрсистемы LILLIPUT-TBC, которая рассматривается далее.

Настоящая работа посвящена анализу 4-раундового алгоритма шифрования LILLIPUT-TBC-II с ключом длины 256 бит (LILLIPUT-TBC-II-256) методом бумеранга. Для построения различителя применяется подход «йо-йо» [9]. Он использовался для атаки на алгоритм блочного шифрования AES, представленной на конференции ASIACRYPT'2017. Основной результат настоящей работы получен благодаря одновременному использованию метода бумеранга и выявленного свойства матрицы рассеивающего преобразования алгоритма LILLIPUT-TBC-II.

Пусть $V_n(2^m)$ — n-мерное векторное пространство над полем $\mathbb{F}_{2^m}$, где $n, m \in \mathbb{N}$; $\oplus$ — операция сложения в $V_n(2^m)$; 0_n — n-мерный вектор, у которого все координаты равны нулю; $I(A)$ — индикатор выполнения условия A; $M^{(4)}_{\mathbb{F}_{2^8}}$ — множество всех (4×4)-матриц над полем $\mathbb{F}_{2^8}$; $g\colon V_{16}(2^8) \times V_{16}(2^8) \to V_{16}(2^8)$ — раундовая функция алгоритма LILLIPUT-TBC-II-256.

Для произвольного $\alpha = (\alpha_{0,0}, \alpha_{0,1}, \ldots, \alpha_{3,3}) \in V_{16}(2^8)$ рассмотрим отображение $\varphi\colon V_{16}(2^8) \to M^{(4)}_{\mathbb{F}_{2^8}}$, заданное условием

$$\hat{\alpha} = \varphi(\alpha) = \begin{pmatrix} \alpha_{0,0} & \alpha_{0,1} & \alpha_{0,2} & \alpha_{0,3} \\ \alpha_{1,0} & \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,0} & \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,0} & \alpha_{3,2} & \alpha_{3,2} & \alpha_{3,3} \end{pmatrix}.$$

Пусть h — $(0,1)$-матрица порядка 16 над полем $\mathbb{F}_{2^8}$ линейного слоя раундовой функции g, а π — 8-битная подстановка S-бокса

$$s = (s_{0,0}, \ldots, s_{0,3}, \ldots, s_{3,0}, \ldots, s_{3,3}) \in S(V_{16}(2^8))$$

нелинейного слоя раундовой функции g, заданного равенством

$$s\colon (\alpha_{0,0}, \ldots, \alpha_{3,3}) \mapsto (\pi(\alpha_{0,0}) \oplus \gamma_{0,0}, \ldots, \pi(\alpha_{3,3}) \oplus \gamma_{3,3}),$$

где $s_{i,j}(\alpha_{i,j}) = \pi(\alpha_{i,j}) \oplus \gamma_{i,j}$; $\gamma_{i,j}$ — фиксированная константа, $\gamma_{i,j} \in \mathbb{F}_{2^8}$, $i, j \in \{0, \ldots, 3\}$. Тогда раундовая функция g задаётся равенством

$$g(\alpha, k) = g_k(\alpha) = h(s(\alpha \oplus k))^{\mathrm{T}}$$

для каждых $(\alpha, k) \in V_{16}(2^8) \times V_{16}(2^8)$, где T — знак транспонирования.

Для произвольных векторов

$$\alpha = (\alpha_0, \ldots, \alpha_{15}) \in V_{16}(2^8),\ \beta = (\beta_0, \ldots, \beta_{15}) \in V_{16}(2^8),\ \theta = (\theta_0, \ldots, \theta_{15}) \in V_{16}(2)$$

и каждого $i \in \{0, \ldots, 15\}$ определим отображения $\rho_\theta^{(i)} \colon V_{16}(2^8)^2 \to \mathbb{F}_2$, $w^{(i)} \colon V_{16}(2^8) \to \mathbb{F}_2$ условиями

$$\rho_\theta^{(i)}(\alpha, \beta) = \begin{cases} \beta_i, & \text{если } \theta_i = 1, \\ \alpha_i, & \text{если } \theta_i = 0, \end{cases}$$
$$w^{(i)}(\alpha) = I(\alpha_i \neq 0).$$

Положим

$$\rho_\theta(\alpha, \beta) = (\rho_\theta^{(0)}(\alpha, \beta), \rho_\theta^{(1)}(\alpha, \beta), \ldots, \rho_\theta^{(15)}(\alpha, \beta)),$$
$$w(\alpha) = (w^{(0)}(\alpha), \ldots, w^{(15)}(\alpha)).$$

Построение различителя для атаки методом бумеранга на 4-раундовый алгоритм LILLIPUT-TBC-II-256 основано на следующей теореме:

Теорема 1. Пусть $\alpha_0, \alpha_1, k_1, k_2, k_3$ — произвольные элементы векторного пространства $V_{16}(2^8)$, $\beta_i = g_{k_3} g_{k_2} g_{k_1}(\alpha_i)$, $i = 1, 2$. Тогда для каждого $\theta \in V_{16}(2)$ справедливо равенство

$$w(g_{0_{128}}(\rho_\theta(\alpha_0, \alpha_1)) \oplus g_{0_{128}}(\rho_\theta(\alpha_1, \alpha_0))) = w(g_{0_{128}}^{-2}(\rho_\theta(\beta_0, \beta_1)) \oplus g_{0_{128}}^{-2}(\rho_\theta(\beta_1, \beta_0))). \quad (1)$$

Заметим, что для построения различителя на основе игрового подхода «йо-йо» [9] для атаки на алгоритм AES применялось равенство, схожее с (1).

Матрице h поставим в соответствие блочную (4×4)-матрицу $\hat{h}$ с блоками-подматрицами $h_{i,j}$ порядка 4, $i, j \in \{0, 1, 2, 3\}$, где

$$\hat{h} = \begin{pmatrix} h_{0,0} & h_{0,1} & h_{0,2} & h_{0,3} \\ h_{1,0} & h_{1,1} & h_{1,2} & h_{1,3} \\ h_{2,0} & h_{2,1} & h_{2,2} & h_{2,3} \\ h_{3,0} & h_{3,2} & h_{3,2} & h_{3,3} \end{pmatrix}.$$

Раундовый ключ $k = (k_{0,0}, \ldots, k_{0,3}, \ldots, k_{3,0}, \ldots, k_{3,3}) \in V_{16}(2^8)$ также представим в матричном виде:

$$\hat{k} = \varphi(k) = \begin{pmatrix} k_{0,0} & k_{0,1} & k_{0,2} & k_{0,3} \\ k_{1,0} & k_{1,1} & k_{1,2} & k_{1,3} \\ k_{2,0} & k_{2,1} & k_{2,2} & k_{2,3} \\ k_{3,0} & k_{3,1} & k_{3,2} & k_{3,3} \end{pmatrix}.$$

Теорема 2. Пусть существуют такие $i, j_1, j_2, r, t \in \{0, \ldots, 3\}$, $\gamma_{j_1,j_2,r} \in \mathbb{F}_{2^8}$, что элементы подматриц h_{i,j_1}, h_{i,j_2} матрицы $\hat{h}$ и подстановки $s_{j_1,r}, s_{j_2,r}$ алгоритма LILLIPUT-TBC-II-256 удовлетворяют условиям

$$(h_{i,j_1})_{t,r} = (h_{i,j_2})_{t,r}, (c_{i,j_1})_{t,r} \neq 0,$$
$$s_{j_1,r}(\beta) = s_{j_2,r}(\beta) \oplus \gamma_{j_1,j_2,r} \text{ для всех } \beta \in \mathbb{F}_{2^8}.$$

Тогда для каждых $\delta \in \mathbb{F}_{2^8}$, $k \in V_{16}(2^8)$ существует вектор

$$\omega \in \langle \alpha_{j_1,r} \oplus \alpha_{j_2,r} \oplus k_{j_1,r} \oplus k_{j_2,r} \rangle \oplus \delta,$$

удовлетворяющий равенству

$$s_{j_1,r}(\alpha_{j_1,r} \oplus k_{j_1,r} \oplus \delta) \oplus s_{j_1,r}(\alpha_{j_1,r} \oplus k_{j_1,r}) \oplus$$
$$\oplus s_{j_2,r}(\alpha_{j_2,r} \oplus k_{j_2,r} \oplus \omega \oplus \delta) \oplus s_{j_2,r}(\alpha_{j_2,r} \oplus k_{j_2,r} \oplus \omega) = 0.$$

На основании теорем 1 и 2 предложена атака методом бумеранга на 4-раундовый LILLIPUT-TBC-II-256.

Для восстановления первого столбца раундового ключа использовалось равенство $(h_{3,0})_{3,1} = (h_{3,1})_{3,1} = 1$, второго столбца — $(h_{3,0})_{3,2} = (h_{3,1})_{3,2} = 1$, а третьего столбца — равенство $(h_{3,0})_{3,3} = (h_{3,3})_{3,3} = 1$. Заметим, что не существует аналогичного равенства для восстановления нулевого столбца ключа k.

Показано, что для атаки требуется 2^{24} текстов, $2^{24,3}$ бит памяти. Трудоёмкость нахождения 256-битного ключа равна 2^{180} операций зашифрования, вероятность ошибки первого рода равна 0.

ЛИТЕРАТУРА

1. *Berger T. P., Francq J., Minier M., and Thomas G.,* Extended generalized Feistel networks using matrix representation to propose a new lightweight block cipher: Lilliput // IEEE Trans. Computers. 2016. V. 65. No. 7. P. 2074–2089.
2. *Berger T. P., Minier M., and Thomas G.* Extended generalized feistel networks using matrix representation // LNCS. 2014. V. 8282. P. 289–305.
3. *Sasaki Y. and Todo Y.* New differential bounds and division property of LILLIPUT: Block cipher with extended generalized Feistel network // LNCS. 2016. V. 10532. P. 264–283.
4. *Sasaki Y. and Todo Y.* New impossible differential search tool from design and cryptanalysis aspects revealing structural properties of several ciphers // LNCS. 2017. V. 10212. No. 3. P. 185–215.
5. *Marriere N., Nachef V., and Volte E.* Differential attacks on reduced round LILLIPUT // LNCS. 2018. V. 10946. P. 188–206.
6. *Adomnicai A., Berger T. P., Clavier C., et al.* Lilliput-AE: a new lightweight tweakable block cipher for authenticated encryption with associated data // NIST Lightweight Cryptography Standardization Process. 2019. `https://csrc.nist.gov/Projects/Lightweight-Cryptography`.
7. *Rogaway P., Bellare M., Black J., and Krovetz T.* OCB: a block-cipher mode of operation for efficient authenticated encryption // Proc. 8th ACM Conf. CCS 2001. Philadelphia, Pennsylvania, USA, 2001. P. 196–205.
8. *Dunkelman O., Keller N., Lambooij E., and Sasaki Y.* A Practical Forgery Attack on Lilliput-AE // 2019. `https://eprint.iacr.org/2019/867.`
9. *Ronjom S., Bardeh N. G., and Helleseth T.* Yoyo tricks with AES // LNCS. 2017. V. 10624. No. 1. P. 217–243.

УДК 004.056.55 DOI 10.17223/2226308X/16/22

РЕАЛИЗАЦИЯ ШИФРАТОРА SD-КАРТ НА ПЛИС С ИСПОЛЬЗОВАНИЕМ ШИФРА МАГМА В РЕЖИМЕ ГАММИРОВАНИЯ

С. И. Разенков

Представлена реализация аппаратного шифратора SD-карт объёмом до 2 Гбайт с использованием шифра Магма в режиме гаммирования. Полученные на ПЛИС разной архитектуры и с использованием разных САПР реализации имеют схожую ресурсоёмкость по логическим элементам. Генерация гаммы не зависит от содержимого открытого текста или шифртекста, что позволяет генератору гаммы работать на собственной, более высокой тактовой частоте, чем остальное устройство. Показано, что таким образом можно значительно сократить время работы устройства.

Ключевые слова: *ПЛИС, шифр Магма, режим гаммирования.*

Задача данного исследования — создать аппаратный шифратор Secure Digital (SD) карт [1], осуществляющий зашифрование и расшифрование всего адресного пространства SD-карты. Аналогичных устройств, с которыми можно было бы провести сравнение, обнаружить не удалось. Несмотря на наличие в SD-картах механизма обеспечения конфиденциальности данных, предоставляющего доступ к содержимому карты по паролю, нет уверенности в отсутствии закладок, позволяющих обходить этот механизм, поскольку в открытом доступе представлена лишь сокращённая спецификация, описывающая поведение SD-карт.

В качестве криптосистемы выбран блочный шифр Магма [2], который имеет небольшую длину блока (64 бит), что позволяет сделать очень компактную с точки зрения аппаратных ресурсов реализацию. Режим гаммирования [3] взят в качестве режима работы блочного шифра в силу своей простоты и идентичности операций зашифрования и расшифрования, что позволяет избежать усложнения логики работы устройства. Кроме того, гамма шифра может независимо генерироваться в процессе чтения блоков данных с SD-карты, после окончания которого можно немедленно приступать к записи данных, покомпонентно сложенных с гаммой. Такой подход, при котором работа шифра не зависит от блоков данных, позволяет избавиться от лишней задержки в работе устройства (рис. 1).

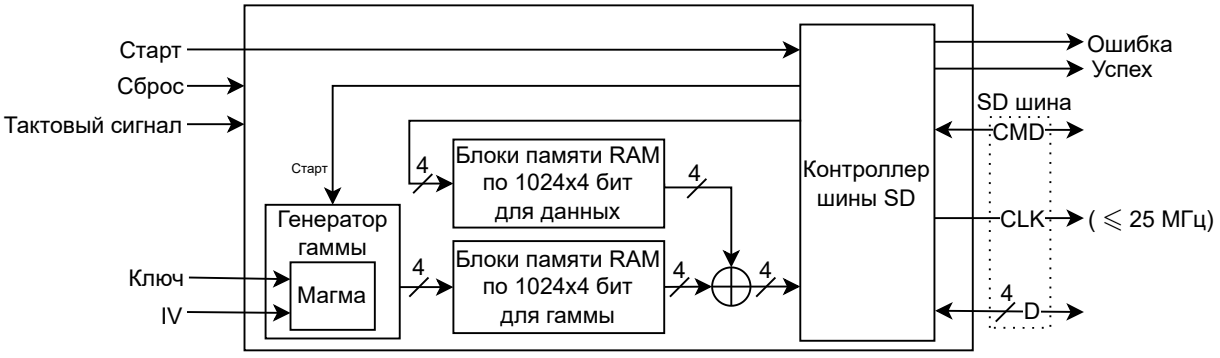

Рис. 1. Структурная схема шифратора SD-карт

Описание шифратора выполнено на языке описания аппаратуры Verilog. Взаимодействие с SD-картой осуществляется с помощью набора взаимосвязанных конечных автоматов. Реализация осуществлялась с расчётом на SD-карты малого объёма (до 2 Гбайт включительно) — SD Standard Capacity (SDSC) карты, чтобы сократить время отладки шифратора. В дальнейшем возможно добавление поддержки SD-карт большего объёма. Согласно стандарту [1], по умолчанию тактовая частота для шины SD составляет не более 25 МГц.

Данные с SD-карты, согласно стандарту, по умолчанию считываются блоками по 512 байт. Кроме того, чтение и запись нескольких подряд идущих блоков осуществляются SD-картой быстрее, чем одиночных. Ввиду ограниченного количества блоков памяти RAM в ПЛИС, использованных для реализации устройства, при испытаниях устройство за транзакцию считывало 8 и 16 последовательно идущих блоков, однако число блоков настраивается и может быть равно произвольной степени двойки, ограниченной лишь имеющимися ресурсами ПЛИС. Параллельно с этим на основе фиксированных ключа и синхропосылки генерируется гамма того же объёма. После чтения данных они покомпонентно складываются с гаммой и записываются обратно на карту по соответствующим адресам.

Симуляция проекта осуществлена с помощью открытого ПО: симулятора Icarus Verilog и библиотеки cocotb языка программирования Python. Шифратор реализован на ПЛИС двух разных семейств: iCE40 фирмы «Lattice Semiconductor» и GW2A фирмы «Gowin». В первом случае в качестве САПР выступили исключительно программы с открытым исходным кодом — Yosys и nextpr, а во втором случае — проприетарное ПО: Gowin EDA Education Edition. Количество использованных ресурсов ПЛИС приведено в табл. 1.

Таблица 1

Сравнение количества использованных ресурсов ПЛИС

Семейство ПЛИС	Блоков в транзакции	Логические элементы	D-триггеры	Блоки RAM
iCE40	8	1111	513	16
GW2A	8	1030	472	16
	16	1119	472	32

Реализации шифратора на обоих семействах ПЛИС протестированы на SD-картах объёмом 1 и 2 Гбайт при разном количестве блоков данных в танзакции, а также разной частоте работы генератора гаммы (табл. 2 и 3). Реализация на ПЛИС семейства iCE40 взаимодействует с SD-картой на частоте 18 МГц, а GW2A — 25 МГц.

Таблица 2

Сравнение времени выполнения зашифрования/расшифрования для реализации на iCE40

Тактовая частота генератора гаммы	Блоков в транзакции	1 Гбайт	2 Гбайт
18 МГц	8	25 м 50 с	52 м 2 с
36 МГц		17 м 25 с	35 м 7 с

Таблица 3

Сравнение времени выполнения зашифрования/расшифрования для реализации на GW2A

Тактовая частота генератора гаммы	Блоков в транзакции	1 Гбайт	2 Гбайт
25 МГц	8	19 м 38 с	40 м 2 с
	16	17 м 20 с	33 м 27 с
100 МГц	8	12 м 34 с	25 м 53 с
	16	8 м 48 с	16 м 24 с

Показано, что увеличение тактовой частоты генератора гаммы позволяет значительно сократить время работы шифратора.

ЛИТЕРАТУРА

1. SD Specifications. Part 1. Physical Layer Simplified Specification Version 9.00. SD Association, 2022.
2. ГОСТ Р 34.12-2015. Информационная технология. Криптографическая защита информации. Блочные шифры. М.: Стандартинформ, 2015.
3. ГОСТ Р 34.13-2015. Информационная технология. Криптографическая защита информации. Режимы работы блочных шифров. М.: Стандартинформ, 2015.

УДК 519.7 DOI 10.17223/2226308X/16/23

ОЦЕНКИ ТРУДНОСТИ ДОКАЗАТЕЛЬСТВ И КРИПТОГРАФИЧЕСКИХ АТАК, ОСНОВАННЫХ НА ЛАЗЕЙКАХ[1]

А. А. Семёнов

Рассматривается задача построения древовидных сертификатов доказательств невыполнимости булевых формул в предположении, что такое доказательство генерируется SAT-решателем, основанным на алгоритме CDCL. Такого сорта древовидные представления удобны, когда для конкретной формулы требуется оценить, насколько трудно доказать её невыполнимость, либо требуется оценить трудоёмкость некоторой криптографической атаки, осуществляемой при помощи SAT-решателя. Предложены древовидные описания сценариев работы CDCL в применении как к невыполнимым формулам, возникающим, например, в задачах символьной верификации, так и к выполнимым формулам, кодирующим задачи обращения дискретных (в том числе криптографических) функций. Доказан ряд свойств введённых древовидных структур. В частности, на языке таких структур сформулировано базовое свойство класса криптографических атак, основанного на инверсных лазейках.

Ключевые слова: *проблема булевой выполнимости (SAT), система доказательства, лазейка, алгоритм CDCL.*

1. Системы доказательства, способы представления доказательств, лазейки

Пусть U — некоторый полный алгоритм, решающий проблему булевой выполнимости, например, U — метод резолюций [1], алгоритм DPLL [2] или алгоритм CDCL [3].

[1]Исследование выполнено в рамках Госзадания Минобрнауки России по проекту «Теоретические основы, методы и высокопроизводительные алгоритмы непрерывной и дискретной оптимизации для поддержки междисциплинарных научных исследований», номер гос. регистрации 121041300065-9.

В соответствии с [4] будем говорить, что алгоритм U задаёт систему доказательства, если существует всюду определённая функция вида

$$f_U : \{0,1\}^* \times \{0,1\}^* \to \{0,1\},$$

заданная некоторым полиномиальным алгоритмом, такая, что $f_U(x,y) = 1$, если y — двоичное описание невыполнимой булевой формулы в КНФ, а x — двоичное описание работы алгоритма U на входе y с момента старта до момента завершения; в противном случае $f_U(x,y) = 0$. Таким образом, можно сказать, что алгоритм, задающий f_U, не только распознаёт по паре (x,y) невыполнимость формулы, представленной словом y, но также проверяет сертификат её невыполнимости, представленный словом x, и тот факт, что это доказательство построено алгоритмом U.

Вопросу выбора способа представления работы алгоритма U на входной формуле может не придаваться особого значения, если речь идёт об оценках сложности доказательств на бесконечных семействах формул, длина которых растёт в зависимости от некоторого натурального параметра (например, формулы Дирихле PHP_n^{n+1} [4]). В таких ситуациях можно использовать различные способы представления доказательств — лишь бы они были полиномиально эквивалентными, то есть большая длина должна быть ограничена полиномом от меньшей длины. Однако если нас интересуют оценки трудности вполне конкретных конечных формул, на основании которых можно сравнивать эффективность различных модификаций базового алгоритма U (например, в результате его дополнения какими-либо новыми эвристиками), то выбор способа описания работы алгоритма U приобретает важное значение. Везде далее будем использовать для представления доказательства x специальные древововидные структуры. Идея представлять доказательства при помощи деревьев не нова и широко используется во многих статьях по сложности доказательств [5–8]. В данной работе мы, однако, детально описываем новый класс древовидных структур, которые хорошо подходят для представления работы современных SAT-решателей, основанных на алгоритме CDCL. Именно эти решатели сейчас доминируют в решении сложных практических задач из широкого спектра областей, включающих верификацию дискретных управляющих систем и криптоанализ.

Ещё один элемент новизны предлагаемого подхода состоит в том, что для оценивания трудности формул с точки зрения размера доказательства мы используем специальные структуры, известные как лазейки (Backdoors). «Классические» лазейки (конкретно, «сильные лазейки» (Strong Backdoors)) описаны в [9]. Как отмечено в работе [8], любая сильная лазейка для невыполнимой булевой формулы автоматически даёт некоторую верхнюю оценку трудности этой формулы. Сильные лазейки, размер которых мал относительно общего числа переменных в формуле, часто могут быть найдены на основе информации о булевой схеме, по которой данная формула построена. Такие ситуации типичны для задач верификации и криптоанализа. Однако зачастую даже такие сильные лазейки слишком велики для того, чтобы построить на их основе нетривиальную оценку трудности рассматриваемой формулы. Другой подход состоит в том, что множество переменных может не быть строгой лазейкой в смысле [9], но тем не менее использование этого множества может существенным образом сокращать время доказательства невыполнимости формулы каким-либо полным алгоритмом решения SAT (например, алгоритмом CDCL). Такого сорта лазейки исследованы в [10]. Наконец, для задач обращения дискретных функций из криптографических приложений могут быть использованы так называемые инверсные лазейки, введённые в [11]. Довольно понятным является тот факт, что произвольную лазейку

(как множество, составленное из некоторых переменных рассматриваемой формулы) можно представить в виде дерева. И, таким образом, процесс доказательства невыполнимости формулы либо криптографическая атака, как некоторый частный случай SAT, могут быть описаны древовидными структурами, в которых представление работы алгоритма решения SAT комбинируется с представлением лазейки, используемой для упрощения рассматриваемой задачи. Далее описан общий класс таких древовидных структур и исследованы их базовые свойства.

2. Древовидные структуры, представляющие работу алгоритмов DPLL и CDCL

Пусть C — некоторая невыполнимая КНФ над множеством переменных $X = \{x_1, \ldots, x_K\}$. Везде далее в роли полного алгоритма, доказывающего невыполнимость C, мы используем CDCL [12] и его слабую версию DPLL как основы систем доказательств, которые по мнению целого ряда исследователей имеют относительно хорошую «автоматизируемость» (то есть относительно эффективно находят относительно короткие доказательства). Для представления процесса доказательства невыполнимости C алгоритмами DPLL и CDCL будем использовать помеченные двоичные деревья следующих двух типов. Дерево первого типа, связанное с DPLL, является полным (full, то есть каждая его нетерминальная вершина имеет два прямых потомка). Обход такого дерева соответствует стандартному поиску в глубину. Рёбра, пройденные в направлении от корня к терминальной вершине, помечаются как пройденные. Если найдена терминальная вершина, то возврат происходит к ближайшей нетерминальной вершине, у которой имеется непройденное ребро.

Тот факт, что сценарий работы алгоритма DPLL на произвольной КНФ C может быть представлен деревом данного типа, хорошо известен. Тем не менее приведём некоторые пояснения, поскольку аналогичные рассуждения потребуются далее. Итак, пусть A — алгоритм DPLL и C — произвольная КНФ над переменными X. Построим дерево, которое обозначим T_A. Любой нетерминальной вершине v данного дерева приписана булева переменная $x_v \in X$. Каждому ребру из нетерминальной вершины v соответствует литерал x_v или $\neg x_v$. Пусть r — корень T_A и x_r — переменная, приписанная r. Пусть v, $v \neq r$, — некоторая нетерминальная вершина и v' — её прямой потомок. Тогда некоторое множество литералов $L_{v'}$ над переменными $\{x_r, \ldots, x_v\}$ соответствует пути из r в v'. Находясь в v', применим правило единичного дизъюнкта (Unit Propagation, UP, [12]) к C и литералам из $L_{v'}$ до тех пор, пока не будет порождён конфликт или не будет получена КНФ с неопределённым статусом, в последнем случае перейдём из v' к следующей вершине. Если переход по ребру (v, v') даёт конфликт, то v' становится терминальной вершиной и помечается символом $\perp$. Если поиск находится в терминальной вершине $\perp$ некоторого пути π, то на следующем шаге делается возврат к ближайшей к $\perp$ нетерминальной вершине v, имеющей инцидентное ей ребро, которое не было пройдено. Если такого ребра не существует, то обход T_A завершается.

Зафиксируем на дереве T_A некоторый порядок его обхода в соответствии с алгоритмом поиска в глубину (DFS) и обозначим полученное дерево через T_A^τ. Заметим, что дерево T_A^τ может быть использовано как сертификат невыполнимости C: будем обходить T_A^τ алгоритмом DFS, применять правило UP к соответствующим литералам и формуле C. Легко понять, что C невыполнима тогда и только тогда, когда каждый путь в T_A^τ оканчивается терминальной вершиной $\perp$. Суммируя все сказанное, имеем следующий факт.

Утверждение 1. Пусть C — невыполнимая КНФ. Тогда существует такое дерево T_A^τ, используя которое, можно верифицировать невыполнимость C детерминированным алгоритмом за время $O(|C| \cdot |T_A^\tau|)$. Все листья дерева T_A^τ в таком случае помечены символом $\perp$.

Второй тип деревьев, рассматриваемых далее, — это двоичные деревья, которые в общем случае не обязаны быть полными, то есть допускается, что некоторая нетерминальная вершина может иметь только одного прямого потомка.

Обозначим через $\tilde{A}$ SAT-решатель на основе алгоритма CDCL, дополнительно полагая, что $\tilde{A}$ не совершает рестартов в процессе опровержения C. Заметим, что $\tilde{A}$ — полный, то есть $\tilde{A}$ опровергает C за конечное время. Свяжем с процессом работы $\tilde{A}$ специальное дерево $T_{\tilde{A}}$ второго типа. Далее опишем структуру $T_{\tilde{A}}$. Ниже использованы понятия, специфичные для CDCL, объяснение которых может быть найдено в [12].

Пусть v — произвольная нетерминальная вершина в $T_{\tilde{A}}$. Данной вершине приписана некоторая переменная из X, которую будем обозначать x_v. Переменная x_v является «переменной выбора» (decision variable), чьё значение (decision literal) выбирается на некотором уровне выбора (decision level), для обозначения которого используется запись $@k$ ($k \geqslant 1$) [3, 12]. Для любой нетерминальной вершины v имеет место: $1 \leqslant \deg(v) \leqslant 2$ ($\deg(v)$ означает выходную степень вершины v).

Только одно ребро, выходящее из нетерминальной вершины v, может соответствовать значению переменной x_v, выбранному на некотором уровне выбора $@k$ ($k \geqslant 1$). Будем называть такое ребро левым (соответствующим образом располагая его на рисунках). Если вершина v имеет только одно исходящее ребро (левое), то это ребро обязательно ведёт в нетерминальную вершину v', соответствующую уровню выбора с номером $@(k+1)$. Такая ситуация соответствует тому факту, что результатом некоторого конфликта является возврат на уровень $@j$, $j \leqslant k-1$ («нехронологический бэктрекинг» [3]). Разберем теперь ситуацию, когда $\deg(v) = 2$. В этом случае, помимо левого исходящего ребра, v имеет ещё одно исходящее ребро, называемое правым. Правое ребро соответствует ситуации, когда в результате анализа конфликта на уровне $@k$ был построен «выученный дизъюнкт» (learnt clause) D_k, содержащий единственный литерал, который был выбран (а не выведен) на уровне $@k$ (полагаем, что для синтеза D_k используется UIP (Unit Implication Point) алгоритм [12]). В этом случае правое ребро, исходящее из v, соответствует ситуации срабатывания процедуры FDA (Failure Driven Assertion [3]) на дизъюнкте D_k. Пусть v^* — корень дерева $T_{\tilde{A}}$. Не ограничивая общности, положим, что v^* имеет два исходящих ребра: левое и правое. Этим двум рёбрам соответствуют два поддерева $T_{\tilde{A}}$, которые обозначим через $T_{\tilde{A}^l}$ и $T_{\tilde{A}^r}$ соответственно. Построение дерева останавливается, если в процессе построения $T_{\tilde{A}^r}$ в результате анализа конфликта порождается выученный дизъюнкт, состоящий из единственного литерала $l(x_{v^*})@1$, где $l(x_{v^*})$ — это литерал, выбранный на уровне $@1$. Фрагмент дерева описанной конфигурации приведён на рис. 1.

Утверждение 2. Произвольному опровержению невыполнимой формулы C посредством алгоритма CDCL без рестартов можно поставить в соответствие некоторое дерево $T_{\tilde{A}}$ описанной структуры.

Заметим, что дерево $T_{\tilde{A}}$ соответствует конкретному сценарию работы алгоритма CDCL без рестартов — в процессе его выполнения дерево $T_{\tilde{A}}$ обходится алгоритмом DFS. Зафиксируем порядок вершин $T_{\tilde{A}}$, пройденных DFS, и обозначим дерево с данным порядком $T_{\tilde{A}}^\tau$. Отметим, что в дереве $T_{\tilde{A}}^\tau$ присутствуют только символы переменных, приписанные вершинам, символ $\perp$, приписанный терминальным вершинам,

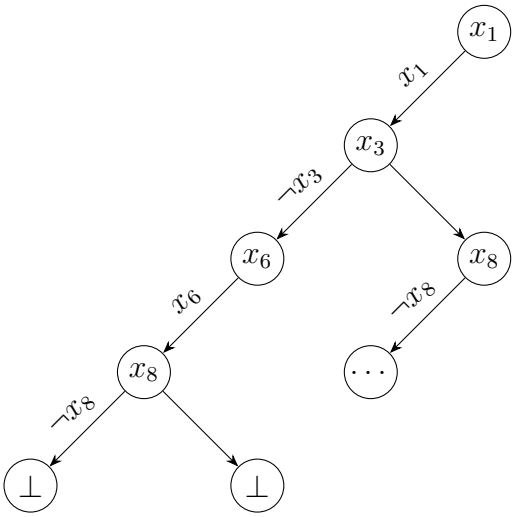

Рис. 1. Фрагмент дерева

символы выбранных литералов, которые приписаны левым рёбрам нетерминальных вершин, и числа, задающие порядок обхода вершин.

Утверждение 3. Пусть C — невыполнимая КНФ и $\tilde{A}$ — CDCL SAT-решатель, не выполняющий рестартов. Тогда существует дерево $T^{\tau}_{\tilde{A}}$, используя которое, можно верифицировать невыполнимость C детерминированным алгоритмом за время $O(|C|\cdot|T^{\tau}_{\tilde{A}}|)$. Все листья $T^{\tau}_{\tilde{A}}$ в таком случае помечены символом $\bot$.

Теперь предположим, что решатель $\tilde{A}$ на основе CDCL совершает в процессе опровержения C рестарты, а также стандартные процедуры, такие, как чистка конфликтных баз.

Сопоставим каждому фрагменту работы $\tilde{A}$ между рестартами с номерами $d-1$ и d дерево описанного типа, обозначаемое через $\left(T^{\tau}_{\tilde{A}}\right)_d$ (считаем, что началу работы $\tilde{A}$ предшествовал рестарт с номером $d=0$). Пусть d^* — номер рестарта, после которого невыполнимость C была доказана. Тем самым мы связываем с процессом опровержения C решателем $\tilde{A}$ некоторый лес $F_{\tilde{A}} = \left\{\left(T^{\tau}_{\tilde{A}}\right)_1, \ldots, \left(T^{\tau}_{\tilde{A}}\right)_{d^*}\right\}$.

Отметим, что лес $F_{\tilde{A}}$, так же как и деревья T^{τ}_{A}, $T^{\tau}_{\tilde{A}}$, можно рассматривать как сертификат невыполнимости C, поскольку $F_{\tilde{A}}$ задаёт конкретный сценарий работы $\tilde{A}$ на формуле C. Очевидно, что размер возникающей в этом случае дополнительной информации (генерируемые $\tilde{A}$ выученные дизъюнкты) ограничен полиномом от $|C|\cdot|F_{\tilde{A}}|$, где $|F_{\tilde{A}}|$ — суммарное число листьев по всем деревьям в $F_{\tilde{A}}$. Таким образом, $F_{\tilde{A}}$ лишь задаёт некоторую фиксированную последовательность уровней решения, всю остальную информацию верифицирующий алгоритм воспроизводит на основе $F_{\tilde{A}}$, подавая алгоритму $\tilde{A}$ уровни решения в соответствии с порядком τ. Следующий факт напрямую следует из анализа структуры леса $F_{\tilde{A}}$.

Теорема 1. Величина $|F_{\tilde{A}}|$ равна числу конфликтов, которые порождает алгоритм $\tilde{A}$, опровергая невыполнимую КНФ C.

3. Древовидные сертификаты доказательств невыполнимости на основе лазеек

Снова рассмотрим проблему доказательства невыполнимости КНФ C над множеством переменных X. Напомним определение сильной лазейки [9]. В его основе лежит

понятие вспомогательного полиномиального алгоритма (sub-solver в терминологии [9]). Такой алгоритм, обозначаемый через P, имеет полиномиальную сложность и должен по произвольной КНФ C либо выдать корректный ответ о выполнимости/невыполнимости C, либо отвергнуть C (неопределённый ответ). Простейшим примером такого алгоритма является правило единичного дизъюнкта.

Произвольное множество $B \subseteq X$ называется сильной лазейкой для C относительно полиномиального алгоритма P, если для любого набора β значений переменных из B ($\beta \in \{0,1\}^{|B|}$) алгоритм P корректно решает проблему выполнимости КНФ $C[\beta/B]$, полученную из C в результате подстановки в неё набора β (обозначаем данный факт через $C[\beta/B] \in S(P)$).

Следующее наблюдение [8] выглядит весьма очевидным. Если B — некоторая сильная лазейка, то имеет место верхняя оценка на время решения SAT в отношении C: $\mathrm{poly}(|C|)\,2^{|B|}$. Действительно, за такое время SAT в отношении C может быть решена за счёт перебора всех $\beta \in \{0,1\}^{|B|}$ и проверки условия $C[\beta/B] \in S(P)$. Таким образом, сильную лазейку B можно считать сертификатом доказательства.

Во многих задачах, связанных со схемами, известны сильные лазейки, размер которых может составлять доли процента от общего числа переменных в формуле. Рассмотрим, например, проблему доказательства эквивалентности двух булевых схем [13] (Logical Equivalence Checking, LEC). В данной задаче рассматриваются две схемы из функциональных элементов некоторого полного базиса (булевы схемы) S_f и S_g, задающие функции $f : \{0,1\}^n \to \{0,1\}^m$ и $g : \{0,1\}^n \to \{0,1\}^m$; требуется доказать или опровергнуть предположение, что f и g на самом деле являются одной и той же функцией ($f \cong g$). Известно, что данную задачу можно за линейное от числа узлов в схемах S_f и S_g время свести к SAT в отношении специальной КНФ $C_{f\cong g}$: схемы эквивалентны (то есть задают одну и ту же функцию) тогда и только тогда, когда $C_{f\cong g}$ невыполнима. КНФ $C_{f\cong g}$ строится по схемам S_f, S_g (выходы которых объединяются специальным функциональным блоком — «майтером» (miter) [13]) при помощи преобразований Цейтина [14]. Основываясь на свойствах этих преобразований, несложно показать, что $C_{f\cong g}$ имеет сильную лазейку относительно правила единичного дизъюнкта, состоящую из n переменных, — такая лазейка образована переменными, приписанными входам схем S_f, S_g. Например, для задачи проверки эквивалентности двух схем, реализующих различные алгоритмы умножения пары 16-битных натуральных чисел, данная лазейка состоит из 32 переменных, тогда как общее число переменных в $C_{f\cong g}$ составляет 9861, если рассматривается задача эквивалентности алгоритмов «Столбик» и «декомпозиция Карацубы» на произвольных парах 16-битных чисел. Таким образом, невыполнимость $C_{f\cong g}$ может быть доказана за 2^{32} применений полиномиального алгоритма P (в роли которого выступает UP) к формулам вида $C\,[\beta/B]$, $\beta \in \{0,1\}^{32}$.

Заметим, что булев гиперкуб $\{0,1\}^n$ может быть представлен в виде полного завершённого бинарного дерева T^n, состоящего из 2^n путей — каждый путь соответствует конкретному вектору из $\{0,1\}^n$. Корню и внутренним узлам такого дерева можно сопоставить булевы переменные, а листьям — значения произвольной булевой функции, определённой всюду на $\{0,1\}^n$. Пусть $X^{\mathrm{in}} = \{x_1, \ldots, x_n\}$ — переменные, приписанные входам схем S_f, S_g. Сопоставим X^{in} дерево T^n, представляющее гиперкуб $\{0,1\}^n$. Семантику терминальных вершин (листьев) T^n определим следующим образом. Каждому пути из корня в лист T^n соответствует конкретный набор значений $\alpha = (\alpha_1, \ldots, \alpha_n)$ переменных из X^{in}, который естественным образом задает набор литералов $x_1^{\alpha_1}, \ldots, x_n^{\alpha_n}$ (здесь и далее $x^\sigma = x$ при $\sigma = 1$ и $x^\sigma = \neg x$ при $\sigma = 0$ [15]). К данному набору литералов и КНФ $C_{f\cong g}$ применяется правило единичного дизъюнкта (UP). Из свойств преобра-

зований Цейтина следует, что схемы S_f и S_g эквивалентны тогда и только тогда, когда для каждой ветви дерева T^n применение UP к соответствующему набору литералов и КНФ $C_{f\cong g}$ даёт конфликт. Соответствующий лист дерева в этом случае помечается символом $\bot$. Таким образом, дерево T^n с листьями, помеченными символом $\bot$, можно рассматривать как сертификат невыполнимости КНФ $C_{f\cong g}$, аналогичный в концептуальном смысле рассмотренным деревьям типа F_A (A — алгоритм DPLL).

Пусть C — произвольная невыполнимая КНФ над переменными X. Рассмотрим произвольное множество $B \subset X$ и представим гиперкуб $\{0,1\}^{|B|}$ в виде дерева T^s, $s = |B|$, как описано выше. Пусть $\beta = (\beta_1, \ldots, \beta_s)$ — произвольный набор из $\{0,1\}^s$ и $x_1^{\beta_1}, \ldots, x_s^{\beta_s}$ — множество литералов над B, соответствующее набору β. Если применение правила UP к КНФ $x_1^{\beta_1} \wedge \cdots \wedge x_s^{\beta_s} \wedge C$ порождает конфликт, то лист дерева T^s, соответствующий пути β, становится терминальной вершиной и получает метку $\bot$. Если же применение UP к КНФ $x_1^{\beta_1} \wedge \cdots \wedge x_s^{\beta_s} \wedge C$ не выводит конфликт, то применяем к $x_1^{\beta_1} \wedge \cdots \wedge x_s^{\beta_s} \wedge C$ алгоритм CDCL, сценарий работы которого представляется лесом $F_{\tilde{A}}$. Таким образом, можно сказать, что к каждой вершине дерева T^s, не помеченной символом $\bot$, «приклеивается» (при помощи вспомогательных непомеченных рёбер) лес вида $F_{\tilde{A}}$, описывающий работу алгоритма CDCL $\tilde{A}$ на соответствующей КНФ вида $x_1^{\beta_1} \wedge \cdots \wedge x_s^{\beta_s} \wedge C$. Полученное в результате дерево (вообще говоря, не бинарное) обозначим через $F_{B,\tilde{A}}$. Если для C имеется некоторое доказательство на основе сильной лазейки, представленное деревом с 2^n листьями (для некоторого n) — например, если C кодирует некоторую LEC задачу и $n = |X^{\text{in}}|$, то очевидный интерес вызывает вопрос существования дерева вида $F_{B,\tilde{A}}$, число листьев которого существенно меньше чем 2^n. Если такое дерево существует, то можно говорить, что C имеет доказательство невыполнимости, которое эффективнее метода грубой силы (brute force). В этом случае также можно считать B, обеспечивающее данное свойство, некоторым вариантом лазейки — такие лазейки исследовались в [10, 16].

4. Древовидные представления криптографических атак на основе инверсных лазеек

Наконец, перенесём идею оценивать трудность доказательств через размер деревьев описанного вида на оценки трудности атак на криптографические функции с использованием SAT-решателей. За отправную точку возьмём понятие инверсной лазейки (Inverse Backdoor Set, IBS) [11]. Напомним контекст проблемы. Рассматривается задача обращения функции $f : \{0,1\}^n \to \{0,1\}^m$, заданной некоторым известным быстрым алгоритмом A_f (например, генератором ключевого потока): известен $\gamma \in \operatorname{Range} f$, требуется найти такой $\alpha \in \{0,1\}^n$, что $f(\alpha) = \gamma$. Одним из следствий теоремы Кука — Левина [17] является тот факт, что по тексту A_f и числу n можно эффективно построить схему S_f из функциональных элементов над произвольным полным базисом, задающую функцию f. По данной схеме при помощи преобразований Цейтина эффективно строится КНФ C_f, называемая далее шаблонной КНФ функции f [18]. Если в C_f подставить (в смысле [19]) набор $\gamma \in \operatorname{Range} f$, то полученная КНФ $C_f(\gamma)$ будет выполнимой, и если удастся найти выполняющий её набор, то будет найден и $\alpha \in \{0,1\}^n : f(\alpha) = \gamma$. С другой стороны, применение UP к произвольной КНФ вида $x_1^{\alpha_1} \wedge \cdots \wedge x_n^{\alpha_n} \wedge C_f$ ($X^{\text{in}} = \{x_1, \ldots, x_n\}$ — переменные, приписанные входу схемы S_f, $\alpha = (\alpha_1, \ldots, \alpha_n)$ — произвольное слово из $\{0,1\}^n$) даёт бесконфликтный вывод всех переменных в C_f, в том числе вывод набора $\gamma : f(\alpha) = \gamma$. Пусть X — множество переменных в КНФ C, рассмотрим произвольное $B \subset X$. Для фиксированного $\alpha \in \{0,1\}^n$

обозначим через β_α набор значений переменных из B, полученный в результате применения UP к КНФ $x_1^{\alpha_1} \wedge \cdots \wedge x_n^{\alpha_n} \wedge C_f$. Назовём β_α набором, индуцированным входом α.

В [11] предложен следующий сценарий поиска прообразов для произвольной функции f, заданной схемой S_f. Зададим на $\{0,1\}^n$ равномерное распределение и пусть α — произвольный вектор, выбранный из $\{0,1\}^n$ в соответствии с этим распределением. Для конкретного $B \subset X$ рассмотрим индуцированный α набор β_α значений переменных из B и набор $\gamma_\alpha = f(\alpha)$. Подставим γ_α и β_α в C_f, обозначим полученную КНФ через $C_f(\gamma_\alpha, \beta_\alpha)$ и применим к $C_f(\gamma_\alpha, \beta_\alpha)$ SAT-решатель $\tilde{A}$ на основе алгоритма CDCL. Очевидно, что КНФ $C_f(\gamma_\alpha, \beta_\alpha)$ выполнима. Зададим некоторое $t > 0$ и определим величину $\xi_{B,t} : \{0,1\}^n \to \{0,1\}$, которая равна 1, если $\tilde{A}$ нашёл набор, выполняющий $C_f(\gamma_\alpha, \beta_\alpha)$ за время $\leqslant t$, в противном случае положим $\xi_{B,t}(\alpha) = 0$. Тогда долю таких $\alpha \in \{0,1\}^n$, для которых $\xi_{B,t}(\alpha) = 1$, можно рассматривать как вероятность успеха в эксперименте Бернулли, обозначаемую через ρ, и оценивать данную вероятность с любой наперёд заданной точностью, используя метод Монте-Карло. Множество B с конкретным значением вероятности ρ называется инверсной лазейкой (IBS) с параметрами $s = |B|, \rho, t$. В [11] описана атака на f, которая анализирует r выходов $\gamma^1, \ldots, \gamma^r$ функции f, построенных по выбираемым случайно и независимо входам $\alpha^1, \ldots, \alpha^r$. Если потребовать, чтобы вероятность обратить хотя бы один из r этих выходов была больше 95 %, то время выполнения такой атаки составит $2^s \cdot t \cdot \lceil 3/\rho \rceil$. Время t в [11] измерялось в секундах, что, вообще говоря, не вполне удачно, если рассматривать выполнение атаки на различных по производительности вычислительных платформах. Представим атаки на основе IBS древовидными структурами описанного выше вида.

Рассматриваем задачу обращения функции $f : \{0,1\}^n \to \{0,1\}^m$. Пусть C_f — шаблонная КНФ функции f; X — множество переменных в C_f; γ — произвольное значение f. Рассмотрим произвольное $B \subset X$ (положим $B = \{x_1, \ldots, x_s\}$). Представим гиперкуб $\{0,1\}^{|B|}$ в виде дерева T^s, $s = |B|$. С каждой ветвью дерева T^s, которая соответствует конкретному набору $\beta \in \{0,1\}^s$, $\beta = (\beta_1, \ldots, \beta_s)$, свяжем лес $F_{\tilde{A}}$, интерпретирующий работу CDCL-решателя $\tilde{A}$ на формуле $x_1^{\beta_1} \wedge \cdots \wedge x_s^{\beta_s} \wedge C_f(\gamma)$. Алгоритм $\tilde{A}$ останавливает работу, как только доказывает невыполнимость соответствующей формулы вида $x_1^{\beta_1} \wedge \cdots \wedge x_s^{\beta_s} \wedge C_f(\gamma)$ либо находит выполняющий её набор. Ограничим теперь множество ветвей леса $F_{\tilde{A}}$ для каждой формулы $x_1^{\beta_1} \wedge \cdots \wedge x_s^{\beta_s} \wedge C_f(\gamma)$ первыми t ветвями. Полученное дерево обозначим через $T_{B,\tilde{A},t}$. Пусть π — произвольный путь в дереве $T_{B,\tilde{A},t}$. Из описания структуры $T_{B,\tilde{A},t}$ следует, что π задает некоторое множество переменных $X^\pi = \{x_1^\pi, \ldots, x_q^\pi\}$ и множество литералов над X^π, обозначаемое $L^\pi = l_1^\pi, \ldots, l_q^\pi$, такое, что применение UP к $l_1^\pi \wedge \cdots \wedge l_q^\pi \wedge C_f(\gamma)$ либо выводит конфликт (в этом случае лист, завершающий π, получает метку $\perp$), либо даёт вывод набора, выполняющего $C_f(\gamma)$. Последний факт соответствует решению задачи обращения рассматриваемого $\gamma \in \operatorname{Range} f$, и листу, завершающему такой путь, присвоим метку $+$.

Следующее утверждение является, по сути, переформулировкой основного теоретического результата [11] на языке деревьев, введенных в настоящей работе.

Теорема 2. Пусть B — инверсная лазейка с параметрами ρ, s, t и $\gamma \in \operatorname{Range} f$ — значение f, полученное для входа α, выбранного из $\{0,1\}^n$ в соответствии с равномерным распределением. Тогда вероятность того, что дерево $T_{B,\tilde{A},t}$, построенное для КНФ $C_f(\gamma)$, содержит путь π, лист которого имеет метку $+$, равна ρ.

Заметим, что предшествующие теореме 2 построения, а также теорема 1 означают, в частности, что время t в атаках на основе инверсных лазеек удобно измерять в числе конфликтов, которые порождает основанный на CDCL решатель $\tilde{A}$.

ЛИТЕРАТУРА

1. *Robinson J. A.* Machine-oriented logic based on resolution principle // J. ACM. 1965. V. 12. Iss. 1. P. 23–41.
2. *Davis M., Logemann G., and Loveland D.* A machine program for theorem-proving // Commun. ACM. 1962. V. 5. No. 7. P. 394–397.
3. *Marques-Silva J. and Sakallah K.* GRASP: A new search algorithm for satisfiability // Proc. ICCAD. San Jose, CA, USA, 1996. P. 220–227.
4. *Cook S. and Reckhow R.* The relative efficiency of propositional proof systems // J. Symbolic Logic. 1979. V. 44. No. 1. P. 36–50.
5. *Ben-Sasson E. and Wigderson A.* Short proofs are narrow — resolution made simple // J. ACM. 2001. V. 48(2). P. 149–169.
6. *Bonet M. L., Esteban J. L., Galesi N., and Johannsen J.* Exponential separations between restricted resolution and cutting planes proof systems // Proc. 39th Ann. Symp. FOCS. Palo Alto, CA, USA, 1998. P. 638–647.
7. *Beame P. and Pitassi T.* Simplified and improved resolution lower bounds // Proc. 37th Conf. FOCS. Burlington, VT, USA, 1996. P. 274–282.
8. *Ansotegui C., Bonet M. L., Levy J., and Manya F.* Measuring the hardness of SAT instances // Proc. AAAI Conf. Chicago, Illinois, USA, July 13–17, 2008. P. 222–228.
9. *Williams R., Gomes C., and Selman B.* Backdoors to typical case complexity // Proc. 18th Int. Conf. IJCAI. Acapulco, Mexico, 2003. P. 1173–1178.
10. *Semenov A., Chivilikhin D., Pavlenko A., et al.* Evaluating the hardness of SAT instances using evolutionary algorithms // Proc. 27th Int. Conf. CP 2021. V. 210. Article No. 47. P. 1–18.
11. *Semenov A., Zaikin O., Otpuschennikov I., et al.* On cryptographic attacks using backdoors for SAT // Proc. AAAI. New Orleans, Louisiana, USA, 2018. P. 6641–6648.
12. *Marques-Silva J., Lynce I., and Malik S.* Conflict-driven clause learning SAT solvers // Handbook of Satisfiability. Second Ed. IOS Press, 2021. P. 133–182.
13. *Molitor P. and Mohnke J.* Equivalence Checking of Digital Circuits: Fundamentals, Principles, Methods. Kluwer Academic Publ., 2004.
14. *Цейтин Г. С.* О сложности вывода в исчислении высказываний // Записки научн. семинаров ЛОМИ. 1968. Т. 8. С. 234–259.
15. *Яблонский С. В.* Введение в дискретную математику. М.: Наука, 1986.
16. *Chivilikhin D., Pavlenko A., and Semenov A.* Decomposing hard SAT instances with metaheuristic optimization // Intern. J. Artificial Intelligence. 2023. V. 21. No. 2. P. 61–92.
17. *Goldreich O.* Computational Complexity: A Conceptual Perspective. Cambridge University Press, 2008.
18. *Semenov A., Otpuschennikov I., Gribanova I., et al.* Translation of algorithmic descriptions of discrete functions to SAT with applications to cryptanalysis problems // Logical Methods in Computer Sci. 2020. V. 16. Iss. 1. Article No. 29. P. 1–42.
19. *Chang C. and Lee R.* Symbolic Logic and Mechanical Theorem Proving. Elsevier, 1973.

УДК 519.7 DOI 10.17223/2226308X/16/24

О СТОЙКОСТИ КЛЮЧЕВЫХ ХЕШ-ФУНКЦИЙ, ОСНОВАННЫХ НА ГОСТ 34.11-2018 («СТРИБОГ»), К АТАКАМ НА КЛЮЧ

А. М. Сергеев, В. А. Кирюхин

Бесключевая хеш-функция ГОСТ 34.11-2018 («Стрибог») является основой многих ключевых криптоалгоритмов, включая HMAC-Стрибог и Стрибог-K. С использованием доказательного подхода к обоснованию стойкости для последних получены оценки сверху на вероятность восстановления секретного ключа. Предложен способ преобразования хеш-функции «Стрибог» в ключевой криптоалгоритм по схеме «сэндвич» (условно называемый Стрибог-C) без внесения изменений в саму хеш-функцию. Стрибог-C является стойкой псевдослучайной функцией и стойким алгоритмом имитозащиты. В отличие от алгоритмов HMAC-Стрибог и Стрибог-K, при любом объёме обрабатываемого материала единственным методом определения секретного ключа алгоритма Стрибог-C будет тотальное опробование в предположении, что аналогичное утверждение верно для функции сжатия, итеративно применяемой внутри хеш-функции.

Ключевые слова: *Стрибог, HMAC, доказуемая стойкость.*

Хеш-функция ГОСТ 34.11-2018 («Стрибог») построена с использованием модифицированной схемы Меркла — Дамгарда. Хешируемое сообщение M дополняется строкой $\mathtt{10..0}$ и разбивается на l блоков $m_1||\ldots||m_l$ по $n = 512$ бит. Исходное состояние хеш-функции равно n-битной константе IV. Функция сжатия g итеративно применяется к блоку сообщения и предыдущему состоянию. На завершающем этапе обработки g применяется ещё два раза — к состоянию «подмешиваются» битовая длина сообщения L и контрольная сумма $\mathsf{sum}(M||\mathtt{10..0}) = (m_1 \boxplus \ldots \boxplus m_l)$ всех блоков по модулю 2^n.

Хеш-функция «Стрибог» служит основой для ряда ключевых криптоалгоритмов, среди которых HMAC-Стрибог [1] и Стрибог-K [2]. Эти алгоритмы не вносят каких-либо изменений в хеш-функцию H, а только подготавливают входные данные:

$$\text{HMAC-Стрибог}(K, M) = \mathsf{H}\left((\overline{K} \oplus opad)||\mathsf{H}(\overline{K} \oplus ipad||M)\right), \ ipad \neq opad,$$
$$\text{Стрибог-K}(K, M) = \mathsf{H}(\overline{K}||M),$$

k-битный ключ K дополняется при необходимости нулями до n бит: $\overline{K} = (K||0\ldots0)$.

В [2] с использованием доказательного подхода [3] к обоснованию стойкости за счёт сведéния к некоторым свойствам функции сжатия показано, что Стрибог-K и HMAC-Стрибог являются стойкими псевдослучайными функциями (PRF) и, следовательно, стойкими схемами имитозащиты. Подобные утверждения о стойкости верны лишь в условиях ограничений на объём материала, обрабатываемого на одном ключе.

Превышение этих ограничений не позволяет говорить об отсутствии атак, обладающих высокой вероятностью успеха. Атаки, направленные на вскрытие секретного ключа, являются при этом наиболее опасными. Для упомянутых ключевых хеш-функций такие атаки существуют [4], их сложность составляет порядка $2^{4n/5}$ операций при сопоставимом объёме материала, что является верхней оценкой стойкости.

В настоящей работе получена нижняя оценка — при надёжности, близкой к единице, и произвольном объёме материала не существует методов восстановления ключа со сложностью менее $2^{k/2}$ операций, $k \leqslant n$.

Результат получен за счёт следующего наблюдения. «Подмешивание» контрольной суммы в процессе финализации приводит к появлению так называемых «связанных

ключей». Так, у алгоритма Стрибог-K при первом вызове g фактически обрабатывается сам ключ $\overline{K}$, а при последнем — $(\overline{K} \boxplus m_1 \boxplus \ldots \boxplus m_l)$, где $m_1, \ldots, m_l$ выбираются противником. Из-за двойного хеширования в схеме HMAC связанные ключи у HMAC-Стрибог возникают соответственно четыре раза, при этом связь задаётся композицией операций «$\oplus$» и «$\boxplus$». Указанные особенности мотивируют потребовать от g стойкости к атакам со связанными ключами, что является достаточным условием стойкости ключевой хеш-функции.

Определение 1. Количественной характеристикой успешности противника $\mathcal{A}$ в модели $KR\text{-}RKA_*$ (Key Recovery under Related Key Attack) для ключевого криптоалгоритма $\mathsf{F} : \mathbf{K} \times \mathbf{X} \to \mathbf{Y}$ назовём вероятность восстановления истинного ключа K

$$\mathrm{Adv}_{\mathsf{F}}^{KR\text{-}RKA_*}(\mathcal{A}) = \mathsf{P}[K \xleftarrow{\mathrm{R}} \mathbf{K};\ \mathcal{A}^{\mathsf{F}_{K*\cdot}(\cdot)} \Rightarrow K';\ K = K'],$$

где «$*$» — некоторая u-арная операция над $\mathbf{K}$. Запрос противника состоит из значения $x \in \mathbf{X}$ и связи по ключу $\varkappa \in \mathbf{K}^{u-1}$. Ресурсы противника: t вычислительных операций; q запросов к оракулу; не более l блоков по n бит в тексте x.

Модель KR (Key Recovery) определяется как модель $KR\text{-}RKA_*$ при унарной тождественной операции в качестве «$*$».

В отсутствие специфических уязвимостей $\mathrm{Adv}_{\mathsf{g}}^{KR\text{-}RKA_*}(t,q)$ можно оценить эвристически значением $t \cdot q \cdot 2^{-k}$, которое соответствует вероятности угадывания хотя бы одного из q связанных ключей за t операций. В то же время $\mathrm{Adv}_{\mathsf{g}}^{KR}(t,q) \lessapprox t \cdot 2^{-k}$.

Теорема 1. Вероятность успеха противника, восстанавливающего секретный ключ криптоалгоритма Стрибог-K (или HMAC-Стрибог), ограничена значениями

$$\mathrm{Adv}_{\text{Стрибог-K}}^{KR}(t,q,l) \leqslant \mathrm{Adv}_{\mathsf{g}}^{KR\text{-}RKA_{\boxplus}}(t', q+1) \lessapprox \frac{q\,t}{2^k},$$

$$\mathrm{Adv}_{\text{HMAC-Стрибог}}^{KR}(t,q,l) \leqslant \mathrm{Adv}_{\mathsf{g}}^{KR\text{-}RKA_{\boxplus\circ\oplus}}(t', 2+2q) \lessapprox \frac{2q\,t}{2^k},\ t' = t + O(q\,l).$$

Интерес представляет построение ключевой хеш-функции, для которой лучшим методом определения ключа было бы тотальное опробование, т. е. обеспечивалась бы k-битная стойкость. Внесение изменений непосредственно в алгоритм хеширования при этом представляется нецелесообразным в силу того, что последний широко распространён и зафиксирован соответствующим государственным стандартом.

Предлагается присоединять к хешируемому тексту специальный блок Σ, обнуляющий в контрольной сумме значение $(m_1 \boxplus \ldots \boxplus m_l)$. Тогда при последнем вызове g обрабатывается только сам ключ $\overline{K}$. Получившаяся конструкция схожа со схемой «сэндвич» [5] (ключ в начале и ключ в конце), условно назовём её «Стрибог-C»:

$$\text{Стрибог-C}(K, M) = \mathsf{H}(\overline{K}||M||\Sigma),$$
$$\Sigma = \Sigma_1||\Sigma_2 = \mathsf{lsb}_p(\widetilde{\Sigma})||\mathsf{msb}_{n-p}(\widetilde{\Sigma}),$$
$$\widetilde{\Sigma} \boxplus \mathsf{sum}(M||10..0) = 0.$$

Здесь p — длина строки $10..0$ ($0 < p \leqslant n$), а $\mathsf{lsb}_x(\widetilde{\Sigma})$ (соответственно $\mathsf{msb}_x(\widetilde{\Sigma})$) означают x младших (старших) бит в блоке $\widetilde{\Sigma}$. Случай, когда длина дополнения равна длине блока ($p = n$), также описывается представленной формулой, $\Sigma = \mathsf{lsb}_n(\tilde{\Sigma}) = \tilde{\Sigma}$.

Из PRF-стойкости алгоритма Стрибог-K непосредственно следует, что Стрибог-C также является стойкой псевдослучайной функцией и стойким алгоритмом имитозащиты:

$$\mathrm{Adv}_{\text{Стрибог-C}}^{\mathrm{PRF}}(t,q,l) \leqslant \mathrm{Adv}_{\text{Стрибог-K}}^{\mathrm{PRF}}(t', q, l+1).$$

Теорема 2. Вероятность успеха противника, восстанавливающего секретный ключ криптоалгоритма Стрибог-С, ограничена значением

$$\mathrm{Adv}^{KR}_{\text{Стрибог-С}}(t,q,l) \leqslant \mathrm{Adv}^{KR}_{\mathtt{g}}(t',q+1) \lessapprox \frac{t}{2^k}.$$

Для обработки сообщения с длиной менее n бит алгоритмам Стрибог-К, Стрибог-С, HMAC-Стрибог-256 и HMAC-Стрибог-512 потребуется соответственно 4, 5, 8 и 9 вызовов функции $\mathtt{g}$, что говорит о сравнительно высокой вычислительной эффективности предложенного криптоалгоритма.

Программные реализации анализируемых криптоалгоритмов представлены в [6].

ЛИТЕРАТУРА

1. Р 50.1.113 2016 — Информационная технология. Криптографическая защита информации. Криптографические алгоритмы, сопутствующие применению алгоритмов электронной цифровой подписи и функции хеширования. М.: Стандартинформ, 2016.
2. *Кирюхин В. А.* Keyed Streebog is a Secure PRF and MAC. CTCrypt 2022. June 6–9, Novosibirsk, Russia. `https://eprint.iacr.org/2022/972`.
3. *Bellare M. and Rogaway P.* Introduction to Modern Cryptography. University of California at Davis, 2005. `https://web.cs.ucdavis.edu/~rogaway/classes/227/spring05/book/main.pdf`.
4. *Dinur I. and Leurent G.* Improved generic attacks against hash-based MACs and HAIFA // LNCS. 2014. V. 8616. P. 149–168.
5. *Yasuda K.* "Sandwich" is indeed secure: How to authenticate a message with just one hashing // LNCS. 2007. V. 4586. P. 355–369.
6. Репозиторий «Ключевой Стрибог». `https://gitflic.ru/project/vkir/streebog`.

УДК 519.6 DOI 10.17223/2226308X/16/25

О СТОЙКОСТИ ГОМОМОРФНОЙ КРИПТОСИСТЕМЫ ДОМИНГО-ФЕРРЕРА ПРОТИВ АТАКИ ТОЛЬКО ПО ШИФРТЕКСТАМ[1]

А. В. Трепачева

Предлагается анализ криптостойкости гомоморфной криптосистемы Доминго-Феррера против атаки только по шифртекстам. Эта криптосистема даёт хороший контрпример к гипотезе об эквивалентности атаки только по шифртекстам и атаки с известными открытыми текстами на криптосистемы, гомоморфные над кольцом вычетов по модулю труднофакторизуемого числа.

Ключевые слова: *гомоморфное шифрование, атака на основе только шифртекста, криптоанализ, факторизация, криптосистема Доминго-Феррера.*

Введение

Изучение гомоморфных криптосистем актуально в связи с приложениями в облачных вычислениях, в которых необходимо обеспечить сложение и умножение шифртекстов так, чтобы расшифрование было гомоморфизмом шифртекстов на кольцо вычетов [1].

[1]Работа выполнена при финансовой поддержке программы «Гранты ИБ МТУСИ» на 2022 г., проект № 49/21-к по договору 40469-49/2021-к.

Интересным направлением в этой области является построение гомоморфных криптосистем, стойкость которых опирается на задачу факторизации чисел [2]. Предложено несколько таких криптосистем, в которых множеством открытых текстов является $\mathbb{Z}_n$, где n — труднофакторизуемое число (RSA-модуль). Многие из упомянутых криптосистем оказались нестойки к атаке с известными открытыми текстами, в связи с чем возник вопрос: может ли какая-то из этих криптосистем, тем не менее, быть стойкой к атаке только по шифртекстам (англ. COA) и в каком случае? Или, другими словами, эквивалентна ли атака по шифртекстам атаке с известными открытыми текстами для упомянутых криптосистем? В случае отрицательного ответа на этот вопрос достаточно одного примера криптосистемы такого типа, для которой эти атаки не эквивалентны [3].

В работе [4] представлена симметричная алгебраически гомоморфная криптосистема, которую будем называть DF96. Стойкость этой криптосистемы проанализирована для случая атаки с известными открытыми текстами [5, 6], однако анализа её стойкости против атаки только по шифртекстам проведено не было.

1. Описание криптосистемы DF96

Будем обозначать как $s \xleftarrow{\$} S$ случайный выбор элемента s из множества S по равномерному распределению. Пусть $n = pq$, где p и q — простые числа, $p < q$. Пространство открытых текстов криптосистемы DF96 — $\mathbb{Z}_n$, пространство шифртекстов — $\mathbb{Z}_p[x] \times \mathbb{Z}_q[x]$, а пространство ключей $\mathcal{K} — \mathbb{Z}_p^* \times \mathbb{Z}_q^*$. Алгоритмы генерации ключей, шифрования и расшифрования приведены на рис. 1.

$\mathsf{KeyGen}_{\mathsf{DF96}}(n)$

1: $r_p \xleftarrow{\$} \mathbb{Z}_p^*$
2: $r_q \xleftarrow{\$} \mathbb{Z}_q^*$
3: **Вернуть** (r_p, r_q)

$\mathsf{Encrypt}_{\mathsf{DF96},d}(m,\ (r_p, r_q))$

1: **Для всех** $i = 2, ..., d-1$
2: $\quad a_i \xleftarrow{\$} \mathbb{Z}_n$
3: $a_d \xleftarrow{\$} \mathbb{Z}_n \backslash \{0\}$
4: $a_1 = (m - \sum_{i=2}^{d} a_i) \pmod n$
5: $a(x) = a_d \cdot x^d + ... + a_1 \cdot x \in \mathbb{Z}_n[x]$
6: $\pi(x) = [a(r_p \cdot x)] \pmod p$
7: $\rho(x) = [a(r_q \cdot x)] \pmod q$
8: **Вернуть** $(\pi(x), \rho(x))$

$\mathsf{Decrypt}_{\mathsf{DF96},d}((\pi(x), \rho(x)),\ (r_p, r_q))$

1: $a_p(x) = \pi(r_p^{-1} \cdot x)$
2: $a_q(x) = \rho(r_q^{-1} \cdot x)$
3: $m_p = [a_p(1)] \pmod p$
4: $m_q = [a_q(1)] \pmod q$
5: **Вернуть** $m_p q(q^{-1} \pmod p) + m_q p(p^{-1} \pmod q)$

Рис. 1. Алгоритмы криптосистемы DF96

Сложение и умножение шифртекстов осуществляются покоординатно:

$$c_\circ = c_1 \circ c_2 = (\pi_1 \circ \pi_2, \rho_1 \circ \rho_2).$$

При умножении размер шифртекстов растет экспоненциально.

2. Основные результаты

Определение 1 (задача $\mathrm{COA}_{\mathsf{DF96},d}(c_1,\ldots,c_l)$). Пусть противник $\mathcal{A}$ владеет шифртекстами $c_1=(\pi_1(x),\rho_1(x)),\ldots,c_l=(\pi_l(x),\rho_l(x))$ криптосистемы DF96, созданными на ключе (r_p,r_q) при параметре d. Задача $\mathrm{COA}_{\mathsf{DF96},d}(c_1,\ldots,c_l)$ состоит в том, чтобы раскрыть $p,(r_p,r_q)$.

Лемма 1. Для любого $f(x)\in\mathbb{Z}_n[x]$ существует некоторый шифртекст криптосистемы DF96, такой, что $f(x)$ является первой (или второй) координатой этого шифртекста.

Доказательство. Для заданного $f(x)$ выберем случайный $r_p\in\mathbb{Z}_n^*$. Считаем, что r_p — первая координата ключа криптосистемы DF96, шифрующая $m\in\mathbb{Z}_n$, который по модулю p равен $f(r_p^{-1})$. При зашифровании коэффициенты полинома приводятся по модулю p, однако при проведении гомоморфных операций приведение по p не происходит, от чего шифртекст не перестаёт быть корректным (т. е. коэффициенты полинома не обязательно должны быть меньше p, чтобы быть первой координатой корректного шифртекста). Аналогично для $r_q\in\mathbb{Z}_n^*$. ■

Теорема 1. Если существует алгоритм $\mathcal{A}$, решающий задачу $\mathrm{COA}_{\mathsf{DF96},d}(c_1,\ldots,c_l)$ за μ ($\mu\leqslant n$) операций с шифртекстами DF96 и с вероятностью ε, то существует алгоритм $\mathcal{B}$, который, имея открытый доступ к $\mathcal{A}$, находит сомножитель n за μ операций с шифртекстами с вероятностью ε.

Доказательство. Рассмотрим алгоритм 1 (алгоритм $\mathcal{B}$ факторизации n).

Алгоритм 1. $\mathcal{B}(n,d,l)$

1: **Для всех** $i=1,\ldots,l$:
2: **Для всех** $j=1,\ldots,d-1$:
3: $\gamma_j\xleftarrow{\$}\mathbb{Z}_n;\quad\delta_j\xleftarrow{\$}\mathbb{Z}_n.$
4: $\gamma_d\xleftarrow{\$}\mathbb{Z}_n\backslash\{0\};\quad\delta_d\xleftarrow{\$}\mathbb{Z}_n\backslash\{0\};$
5: $\pi_i(x):=\gamma_dx^d+\ldots+\gamma_1x;\quad\rho_i(x):=\delta_dx^d+\ldots+\delta_1x.$
6: $(p,(r_p^{-1},r_q^{-1})):=\mathcal{A}((\pi_1(x),\rho_1(x)),\ldots,(\pi_l(x),\rho_l(x))).$
7: **Вернуть** p, n/p.

Если $\mathcal{A}$ с преимуществом ε за μ операций находит секретный ключ DF96, то $\mathcal{B}$ раскрывает факторизацию n за μ операций с преимуществом ε. ■

В алгоритме 2 описаны возможные действия криптоаналитика $\mathcal{A}$ по решению задачи $\mathrm{COA}_{\mathsf{DF96},d}(c_1,\ldots,c_l)$. Через $\mathrm{Res}(g(x),f(x))$ обозначен результант полиномов $f(x)$ и $g(x)$.

Предположим, что алгоритм 2 делает τ внешних итераций (строки 2–13), так что $|L|=l+\tau$. Тогда общее число операций с шифртекстами $\mu\leqslant\tau(l+\tau)(l+\tau-1)^2/2$.

Когда строится алгоритм факторизации, он должен иметь возможность генерировать шифртексты для данного n. Исходя из леммы 1, сгенерировав по равномерному распределению $\pi_1(x),\ldots,\pi_l(x)$ и $\rho_1(x),\ldots,\rho_l(x)$, мы можем рассматривать их как первые и вторые координаты шифртекстов DF96. При этом распределение ψ на множестве открытых текстов будет равномерным (сводимость от $\mathrm{COA}_{\mathsf{DF96},d}(c_1,\ldots,c_l)$ к задаче факторизации справедлива только для равномерного ψ).

Отметим, что обратная сводимость не имеет места. В самом деле, пусть атакующий знает разложение числа n, но не знает (r_p,r_q). Тогда, подставляя любые значения из

Алгоритм 2. $\mathcal{A}((\pi_1(x), \rho_1(x)), \ldots, (\pi_l(x), \rho_l(x)))$

1: $L := \pi_1(x), \ldots, \pi_l(x)$.
2: **Пока** `true`
3: $i \xleftarrow{\$} \{1, \ldots, |L|\}$; $j \xleftarrow{\$} \{1, \ldots, |L|\}$; $\circ \xleftarrow{\$} \{+, -, \cdot\}$;
4: $\gamma(x) := \pi_i(x) \circ \pi_j(x)$; $L := L \cup \{\gamma(x)\}$.
5: **Для всех** $i, j \in \{1, \ldots, |L|\}$:
6: **Для всех** $k, t \in \{1, \ldots, |L|\}$:
7: $\delta :=\leftarrow \text{НОД}(\text{Res}(\pi_i(x) - \pi_j(x), \pi_k(x) - \pi_t(x)), n)$.
8: **Если** $(1 < \delta < n) \wedge (\delta \bmod n = 0)$, **то**
9: $p = \delta$;
10: $r_p^{-1} = \text{НОД}((\pi_i(x) - \pi_j(x)) \bmod p, (\pi_k(x) - \pi_t(x)) \bmod p)$;
11: $q = n/p$;
12: $r_q^{-1} = \text{НОД}((\rho_i(x) - \rho_j(x)) \bmod q, (\rho_k(x) - \rho_t(x)) \bmod q)$.
13: **Вернуть** $p, (r_p^{-1}, r_q^{-1})$.

$\mathbb{Z}_p^*$ и $\mathbb{Z}_q^*$ вместо r_p и r_q соответственно, на шагах 1 и 2 алгоритма $\mathsf{Decrypt}_{\mathsf{DF96},d}$ атакующий может получать на выходе некоторый открытый текст, и у него нет критерия, чтобы отделить правильный вариант r_p и r_q от неправильного в случае равномерного распределения открытых текстов.

Если распределение открытых текстов отлично от равномерного и количество шифртекстов достаточно для надёжного различения этих распределений, то это может служить критерием правильности угадывания r_p и r_q. Но в этом случае средняя сложность атаки алгоритмом 2 снижается за счёт повышения вероятности появления шифртекстов c_i и c_j, таких, что $\mathsf{Decrypt}_{\mathsf{DF96},d}(c_i, (r_p, r_q)) = \mathsf{Decrypt}_{\mathsf{DF96},d}(c_j, (r_p, r_q))$. Таким образом, знание разложения n не исключает перебора попарных разностей и подсчёта их наибольшего общего делителя.

Заключение

Показана сводимость атаки на криптосистему DF96 только по шифртекстам к задаче разложения числа n на сомножители. Данный факт представляет интерес с точки зрения выяснения эквивалентности атак только по шифртекстам и с известными открытыми текстами на гомоморфные криптосистемы, множество открытых текстов которых — это $\mathbb{Z}_n$, а именно: эти атаки не эквивалентны для криптосистем такого типа.

В общем виде можно сказать, что данный факт исходит из того, что при фиксированном секретном ключе зашифрование действует сюръективно на множестве шифртекстов. Иными словами, для генерации корректных шифртекстов нет необходимости в знании секретного ключа — любому шифртексту при любом ключе соответствует некоторый открытый текст.

ЛИТЕРАТУРА

1. *Gentry C.* Fully Homomorphic encryption using ideal lattices // Proc. 41-th ACM Symp. STOC'09. Bethesda, USA, 2009. P. 169–178.
2. *Vaikuntanathan V.* Computing blindfolded: New developments in fully homomorphic encryption // Proc. 52nd Ann. Symp. FOCS. Palm Springs, CA, USA, 2011. P. 5–16.
3. *Трепачева А. В.* О соотношениях между атаками на симметричные шифры, гомоморфные над кольцом вычетов // Безопасность информационных технологий. 2017. Т. 24. № 2. С. 82–91.

4. *Domingo-Ferrer J.* A new privacy homomorphism and applications // Inform. Process. Lett. 1996. V. 60. No. 5. P. 277–282.
5. *Cheon J. H., Kim W.-H., and Nam H. S.* Known-plaintext cryptanalysis of the Domingo-Ferrer algebraic privacy homomorphism scheme // Inform. Process. Lett. 2006. V. 97. No. 3. P. 118–123.
6. *Трепачева А. В.* Улучшенная атака по известным открытым текстам на гомоморфную криптосистему Доминго-Феррера. // Труды ИСП РАН. 2014. Т. 26. № 5. С. 83–98.

УДК 512.548.7+004.056.55 DOI 10.17223/2226308X/16/26

ОБ ОДНОМ КВАЗИГРУППОВОМ АЛГОРИТМЕ ШИФРОВАНИЯ, СОХРАНЯЮЩЕГО ФОРМАТ

К. Д. Царегородцев

Рассматривается возможный подход к построению схем шифрования, сохраняющего формат, на основе квазигрупповых преобразований (левых и правых сдвигов на псевдослучайные элементы). Показано, что в случае функционального задания квазигруппы с помощью правильных семейств дискретных функций над прямым произведением групп обратное к левому (правому) сдвигу преобразование также задаётся правильным семейством.

Ключевые слова: *FPE, квазигруппа, правильное семейство.*

Шифрование, сохраняющее формат (FPE, Format Preserving Encryption [1], далее FPE-схема) — алгоритм, позволяющий зашифровывать сообщения из произвольного конечного множества M таким образом, что результат зашифрования также лежит в множестве M. Такой тип алгоритмов довольно востребован на практике, о чём свидетельствует большое количество работ, рассматривающих стойкость таких криптопримитивов [1–3]. При этом подобные алгоритмы часто подвержены специфическим атакам, связанным, в том числе, и с возможным относительно малым размером области определения [4, 5]. Известны «доказуемо стойкие» алгоритмы как для очень малых ($|M| \approx 2^{10}$), так и для очень больших областей определения (размер которых приближается к размеру области определения стандартных блочных шифров либо превышает их, wide-block encryption), в то время как для «средних» областей определения всё ещё не существует одного предпочтительного подхода. В этой работе мы рассмотрим возможный подход к построению FPE-схем, основанный на квазигрупповых операциях.

Определение 1. Квазигруппой $(Q, \circ)$ называется множество Q с заданной на нём бинарной операцией $\circ$ со следующим свойством: для любых a, b уравнения $a \circ x = b$, $x \circ a = b$ однозначно разрешимы (относительно x).

Другими словами, операции левого ($x \to L_a(x) = a \circ x$) и правого ($x \to R_a(x) = x \circ a$) сдвигов являются подстановками на множестве Q.

Основанные на квазигрупповых преобразованиях алгоритмы интенсивно изучались [6–9], некоторые из них предлагались в качестве кандидатов на международную стандартизацию (например, [10]). В работе [11] предложен следующий подход. Преобразуем элемент $m \in M$, где $(M, \circ)$ — некоторая квазигруппа, в элемент $c \in M$ следующим образом:

$$m \to c = L_{k_1,\ldots,k_\ell}(m) = k_1 \circ (k_2 \circ (\ldots (k_\ell \circ m) \ldots)), \qquad (1)$$

то есть элемент m последовательным применением левых сдвигов внутри квазигруппы M переводится в элемент c. В [11] предложены также варианты схемы, в которых

псевдослучайным образом в зависимости от параметров алгоритма производятся левые и правые сдвиги. Относительно приведённой FPE-схемы можно задать следующие вопросы:

— насколько L-преобразование «похоже» на случайную подстановку на M?
— как устроено обращение операции $m \to L_{k_1,\ldots,k_\ell}(m)$ (далее для краткости будем говорить «L-преобразование»)?

Частичный ответ на первый вопрос был дан ранее. L-преобразования рассматривались, в частности, в работах [7, 9, 12]. Так, в работе [12] среди прочего показано, что для любой квазигрупповой операции $\circ$ при случайном независимом выборе элементов $k_i \in M$ (при условии, что носитель распределения k_i достаточно большой) распределение элемента c экспоненциально быстро сходится к равновероятному распределению на множестве M. При этом результаты работы [12] неприменимы к ситуации, в которых противник может получать образы различных адаптивно выбираемых m. В [11] показано, что при порождении элементов k_i с помощью псевдослучайных функций (на основе мастер-ключа и настройки (tweak)) стойкость полученной FPE-схемы в стандартной модели $TPRP$ [13] может быть оценена через стойкость используемой псевдослучайной функции, а также через стойкость L-преобразования (1) в модели, где противнику даётся либо случайная подстановка $\pi \in S_M$, либо L-преобразование $L_{k_1,\ldots,k_\ell}$ для случайно выбранных $k_i \in M$, и его задачей является различение этих двух ситуаций. В той же работе указано, что стойкость полученной схемы сильно зависит от структуры квазигруппы и её свойств.

В настоящей работе рассмотрим ограниченный класс квазигрупп, порождённых «правильными семействами» дискретных функций [14].

Определение 2. Отображение $F\colon M^n \to M^n$ будем называть правильным семейством (размера n), если для любых двух неравных наборов $x, y \in M^n$ найдётся такой индекс i, что $x_i \neq y_i$, но $F_i(x_1, \ldots, x_n) = F_i(y_1, \ldots, y_n)$.

Пусть F, G — два правильных семейства размера n над прямым произведением H^n групп $(H, +)$. Тогда операция $\pi_F(x)$, определённая как

$$\pi_F = (x_1 + F_1(x_1, \ldots, x_n), \ldots, x_n + F_n(x_1, \ldots, x_n)),$$

является подстановкой на множестве H^n [14, теорема 8]. В таком случае можно рассмотреть следующую операцию умножения $\circ$ на $H^n \times H^n$:

$$(x, y) \to x \circ y = \pi_F(x) + \pi_G(y), \tag{2}$$

где $+$ понимается как покомпонентное сложение в группе H.

Определённое таким образом умножение в H^n может быть обращено. Рассмотрим подстановку π_F^{-1}. Можно показать, что существует правильное семейство $G = \widetilde{F}$, такое, что $\pi_F^{-1} = \pi_G$ (такое семейство можно назвать «дуальным», поскольку дважды применённая операция $F \to \widetilde{F}$ оставляет исходное семейство на месте).

Теорема 1. Если F — правильное семейство на группе H^n, то семейство $\widetilde{F}$, заданное как

$$\widetilde{F}(x) = (-x) + \pi_F^{-1}(x), \quad \pi_F(x) = x + F(x), \quad x \in H^n, \tag{3}$$

также является правильным на H^n.

Таким образом, если F и $\widetilde{F}$ — пара правильных семейств, связанных соотношением (3), и операция $\circ$ задается формулой (2), то операция $x \circ y$ обращается справа следующим образом:

$$x = \pi_{\widetilde{F}}\left((x \circ y) - \pi_G(y)\right).$$

Аналогичным образом операция $x \circ y$ может быть обращена слева с использованием «дуального» к G семейства $\widetilde{G}$.

Из теоремы 1 вытекает, что как L-преобразование, так и обратное к нему L^{-1} могут быть заданы с помощью правильных семейств дискретных функций. Это позволяет перейти от табличного задания квазигруппы к функциональному [14].

ЛИТЕРАТУРА

1. *Bellare M., Ristenpart T., Rogaway P., and Stegers T.* Format-preserving encryption // LNCS. 2009. V. 5867. P. 295–312.
2. *Lee J.-K., Koo B., Roh D., et al.* Format-preserving encryption algorithms using families of tweakable blockciphers // LNCS. 2014. V. 8949. P. 132–159.
3. *Dworkin M.* Recommendation for Block Cipher Modes of Operation: Methods for Format-Preserving Encryption. NIST Special Publication 800-38G. 2016.
4. *Hoang V. T., Tessaro S., and Trieu N.* The curse of small domains: New attacks on format-preserving encryption // LNCS. 2018. V. 10991. P. 221–251.
5. *Amon O., Dunkelman O., Keller N., et al.* Three third generation attacks on the format preserving encryption scheme FF3 // LNCS. 2021. V. 12697. P. 127–154.
6. *Марков В. Т., Михалёв А. В., Нечаев А. А.* Неассоциативные алгебраические структуры в криптографии и кодировании // Фундамент. и прикл. матем. 2016. Т. 21. № 4. С. 99–124.
7. *Markovski S. and Bakeva V.* Quasigroup string processing: Part 4 // Contributions. Sec. Natural Math. Biotechn. Sci. 2006. V. 27. No. 1–2. `http://csnmbs.manu.edu.mk/index.php/csnmbs/article/view/5`.
8. *Shcherbacov V.* Elements of Quasigroup Theory and Applications. N.Y.: Chapman and Hall/CRC, 2017. 598 p.
9. *Артамонов В. А.* Квазигруппы и их приложения // Чебышевский сборник. 2018. Т. 19. № 2. С. 111–122.
10. *Gligoroski D., Ødegård R. S., Mihova M., et al.* Cryptographic hash function Edon-R'. Proc. 1st Intern. Workshop on Security and Communication Networks. Trondheim, Norway, 2009. P. 1–9.
11. *Tsaregorodtsev K.* Format-preserving encryption: a survey // Математические вопросы криптографии. 2022. Т. 13. № 2. С. 133–153.
12. *Яшунский А. Д.* О скорости сходимости квазигрупповых сверток вероятностных распределений // Дискретная математика. 2022. Т. 34. № 3. С. 160–171.
13. *Liskov M., Rivest R., and Wagner D.* Tweakable block ciphers // J. Cryptology. 2011. V. 24. No. 3. P. 588–613.
14. *Носов В. А., Панкратьев А. Е.* О функциональном задании латинских квадратов // Интеллектуальные системы. Теория и приложения. 2008. Т. 12. № 1–4. С. 317–332.

УДК 519.7 DOI 10.17223/2226308X/16/27

ОБ ОДНОМ ПОДХОДЕ К ПОСТРОЕНИЮ КЛЮЧЕВОЙ ПСЕВДОСЛУЧАЙНОЙ ФУНКЦИИ НА ОСНОВЕ БЛОЧНОГО ШИФРА МАГМА

А. А. Щербаченко

Рассматривается способ преобразования блочного шифра Магма (алгоритм ГОСТ 34.12-2018 с длиной блока 64 бита) в ключевую псевдослучайную функцию MAGMA-PRF. Показано, что MAGMA-PRF является стойкой к некоторым конструктивным методам криптоанализа, которые применимы к базовому блочному шифру. Предложены способы использования MAGMA-PRF и в рамках доказуемого подхода к обоснованию стойкости показано, что в некоторых режимах работы (CTR, CTR-ACPKM, GCM) MAGMA-PRF имеет лучшие криптографические свойства, чем блочные шифры с такой же длиной блока.

Ключевые слова: *блочные шифры, режимы шифрования, алгоритм Магма, MAGMA-PRF, доказуемая стойкость.*

В работе [1] предложен способ построения ключевой псевдослучайной функции (PRF) на основе блочного шифра, а также исследованы свойства конструкции AES-PRF, в которой в качестве базового шифра использовался алгоритм AES. В настоящей работе исследуется возможность применения данного способа к отечественному блочному шифру Магма.

Блочный шифр Магма является международным стандартом и описан в ГОСТ 34.12-2018 [2] (алгоритм блочного шифрования с длиной блока $n = 64$ бита). Магма представляет собой двухъячеистую сбалансированную сеть Фейстеля, которая обрабатывает блоки следующим образом. Ключ K длиной 256 бит разбивается на подблоки $k_1, \ldots, k_8$ одинаковой длины (32 бита). Входной блок $x \in V^{64}$ ($V^n = \{0,1\}^n$) разбивается на две равные половины: $x = (x_l, x_r)$. Раундовое преобразование блока определяется как $F_{k_i}(x) = (x_r, x_l \oplus s(x_r \boxplus k_i) \lll_{11})$, где $\boxplus$ — операция сложения по модулю 2^{32}; s — замена полубайт блока по фиксированным таблицам замен; $\lll_{11}$ — циклический сдвиг на 11 в сторону старших бит (влево). Функция зашифрования определяется как $E_K(x) = S \circ F_{k_{32}} \circ \ldots \circ F_{k_2} \circ F_{k_1}(x)$, где S означает перестановку половин блока. Раундовые ключи $k_1, \ldots, k_{32}$ вычисляются по следующему правилу: $k_{i+8} = k_{i+16} = k_i$, $k_{i+24} = k_{9-i}$ при $i = 1, \ldots, 8$.

Описанный в [1] способ заключается в добавлении к выходу блочного шифра промежуточного состояния, полученного после r раундов шифрования. Обозначим через $E_K^{(r)}(x) = F_{k_r} \circ \ldots \circ F_{k_2} \circ F_{k_1}(x)$ функцию E_K, усеченную до r раундов ($1 \leqslant r \leqslant 31$). Рассмотрим преобразование

$$\text{MAGMA-PRF}_K^r(x) = E_K(x) \oplus E_K^{(r)}(x), \tag{1}$$

схематично представленное на рис. 1.

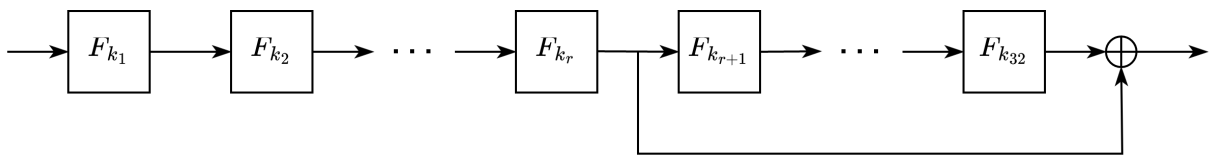

Рис. 1. Конструкция MAGMA-PRF^r

Хорошо известным фактом в теории доказуемой стойкости [3] является то, что при использовании PRF вместо подстановки (PRP) режимы шифрования, в которых примитив используется для выработки одноразовой гаммы, не подвержены эффекту «парадокса дней рождения», что позволяет шифровать большее количество материала на одном ключе. Данное утверждение верно в том случае, если используемая PRF является стойкой к построению эффективного различителя: не должно существовать атак, вероятность успеха которых существенным образом возрастает в зависимости от количества доступного противнику материала q (известных или адаптивно выбираемых пар блоков открытого текста/шифртекста) и при некотором значении q и вычислительных ресурсах t превышает вероятность успеха при тотальном опробовании ключей ($t/2^k$, где k — битовая длина ключа).

Противником $\mathcal{A}$ будем называть произвольный вероятностный алгоритм, решающий конкретную задачу по взлому исследуемой криптосистемы. Под оракулом $\mathcal{O}$ понимается вероятностный интерактивный алгоритм, моделирующий работу криптосистемы, с которым $\mathcal{A}$ взаимодействует (совершает запросы) по принципу «чёрного ящика»; это взаимодействие будем обозначать как $\mathcal{A}^{\mathcal{O}(\cdot)}$, где символ $\cdot$ означает интерфейс, к которому имеет доступ противник. Совокупность задачи, решаемой противником, и его возможностей по взаимодействию с криптосистемой будем называть моделью угроз. В задачах типа «различение» противник взаимодействует с одним из двух оракулов $\mathcal{O}_b$, $b \xleftarrow{\$} \{0,1\}$ (символом $\xleftarrow{\$}$ обозначаем случайный равновероятный выбор из множества), значение бита b противнику неизвестно; результат работы противника будем обозначать $\mathcal{A}^{\mathcal{O}_b(\cdot)} \Rightarrow \hat{b}$, где $\hat{b} \in \{0,1\}$ — предположение противника о бите b.

Определение 1. (PRF)-преобладанием противника $\mathcal{A}$, ограниченного t вычислительными ресурсами и q адаптивными запросами к соответствующему оракулу в виде n-битных блоков, в задаче различения ключевой функции $F : V^k \times V^n \to V^n$ и функции ρ, выбираемой случайным равновероятным образом из множества всех функций из V^n в V^n (модель угроз PRF), назовём величину

$$\mathrm{Adv}^{\mathrm{PRF}}_{\mathrm{F}}(\mathcal{A}) = \mathsf{P}\left[K \xleftarrow{\$} V^k; \mathcal{A}^{\mathrm{F}_K(\cdot)} \Rightarrow 1\right] - \mathsf{P}\left[\rho \xleftarrow{\$} \mathrm{Func}(V^n, V^n); \mathcal{A}^{\rho(\cdot)} \Rightarrow 1\right].$$

На конструкции типа (1) найдены универсальные атаки различения [1]. При $q = 2^{64}$ (противник обладает доступом к полной кодовой книге) преобладание противника в различении составит $1 - 2^{-64}$ за $t = 2^{64}$ операций, что практически равно 1. В случае $q < 2^{32}$ преобладание оценивается величиной $q^2/2^{128}$, в случае $2^{32} < q < 2^{64}$ — величиной $q/2^{96}$. Отметим, что даже при значениях q, близких к граничным, в том числе когда противник имеет доступ к *почти всей* кодовой книге ($q \leqslant 2^{64} - 1$), преобладание его по универсальному методу не превосходит 2^{-32}.

Рассмотрим некоторые конструктивные атаки на алгоритм Магма, которые предложены в литературе, а также их влияние на конструкцию MAGMA-PRF.

Метод Исобе [4] основан на использовании свойства шифра «точка отражения». Обозначим $E_K^{(i,j)}(x) = F_{k_j} \circ \ldots \circ F_{k_i}(x)$, $1 \leqslant i < j \leqslant 32$. Заметим, что в силу того, что в раундах 1–8, 9–16, 17–24 используются одинаковые последовательности раундовых ключей, для любого x справедливо $E_K^{(8)}(x) = E_K^{(9,16)}(x) = E_K^{(17,24)}(x)$, и функцию E_K можно переписать в виде $E_K(x) = D_K^{(8)} \circ S \circ E_K^{(8)} \circ E_K^{(8)} \circ E_K^{(8)}(x)$, где $D_K^{(8)}$ — функция, обратная к $E_K^{(8)}$. Тогда если найдётся такое x, называемое точкой отражения, что $E_K^{(24)}(x) = (y_l, y_l)$ (половины блоков совпадают), то на последних восьми раундах произойдёт частичное расшифрование результата 17–24 раундов, и результат всего

зашифрования будет соответствовать первым 16 раундам: $E_K(x) = E_K^{(16)}(x)$. Для случайно выбираемого x вероятность возникновения отражения составляет $P_{\text{ref}} = 2^{-32}$.

Конструкция MAGMA-PRF^{16} («сбалансированный» вариант) не является стойкой к различителю, основанному на методе Исобе. Если x является точкой отражения для E_K, то MAGMA-PRF$^{16}_K(x) = E_K(x) \oplus E_K^{(16)}(x) = 0$. Среди q доступных противнику пар возникнет порядка $q \cdot P_{\text{ref}}$ нулей, в то время как для случайной функции это число оценивается величиной $q \cdot 2^{-64}$.

Метод Динура — Дункельмана — Шамира [5] основан на использовании свойства шифра «фиксированная точка». Неподвижной точкой для $E_K^{(i,j)}$ назовём такой x, что $E_K^{(i,j)}(x) = x$. Тогда если x — неподвижная точка для $E_K^{(8)}$, то с учётом представления для E_K, указанного выше, имеем $E_K(x) = E_K^{(25,32)}(x)$. Для случайного x вероятность быть неподвижной точкой оценивается величиной $P_{\text{fix}} = 2^{-64}$.

Применим метод [5] к MAGMA-PRF^r при $r \in \{8, 24\}$. С вероятностью P_{fix} реализуется «неподвижная точка» для $E_K^{(8)}$, тогда к MAGMA-PRF применима та же атака, что и к блочному шифру. Трудоёмкость атаки на 8 раундов [5] оценивается величиной 2^{128}, при этом противник не знает, при какой паре реализовалась неподвижная точка, и опробует $q = 2^{64}$ пар. Вероятность успеха при доступном количестве материала $q < 2^{64}$ оценим величиной $\min(t/2^{192}, q/2^{64})$, что лучше универсального метода различения $(q/2^{96})$ при минимальных вычислительных ресурсах противника $t > 2^{96}/q$.

При других значениях r для MAGMA-PRF^r не было найдено возможностей применения рассмотренных методов для построения более эффективной атаки, чем тотальное опробование. Предполагается, что для того, чтобы отследить благоприятное для атаки событие типа «точка отражения» или «неподвижная точка», противник вынужден угадывать либо $E_K(x)$, либо промежуточное состояние $E_K^{(r)}(x)$, что повышает трудоёмкость методов до 2^{256}.

Представляется, что конструкция является стойкой к классическому дифференциальному и линейному методам криптоанализа, поскольку задействуются все 32 раунда и эффективность характеристики (линейной или дифференциальной) для MAGMA-PRF не превышает таковой для полнораундовой E_K (эта гипотеза требует дальнейших исследований). При этом остаётся открытым вопрос о стойкости конструкции к другим методам — например, в работе [6] показано, что конструкция AES-PRF при малых значениях r не является стойкой к методам, основанным на невозможных дифференциалах и невозможных линейных аналогах.

Поскольку при «крайних» значениях $r \in \{1, \ldots, 7\}$ или $r \in \{25, \ldots, 31\}$ в вычислении внутреннего состояния задействуются не все раундовые ключи (с прямой или обратной стороны), для противодействия возможным атакам, не учтённым в настоящей работе, представляется целесообразным выбирать значения r из промежутка $\{9, \ldots, 23\}$, за исключением 16, атака для которого описана ранее.

Таким образом, при $r \in \{9, \ldots, 15\} \cup \{17, \ldots, 23\}$ нам не удалось выявить атак, которые позволили бы отличить MAGMA-PRF от случайной функции с преобладанием выше, чем для универсальных методов. Дадим следующую эвристическую оценку преимущества противника, ограниченного t вычислительными операциями и q адаптивными запросами к оракулу (адаптивно выбираемыми парами «открытый текст/шифртекст»), в модели PRF для MAGMA-PRF^r при указанных значениях r и $q \leqslant 2^{64} - 1$:

$$\text{Adv}^{\text{PRF}}_{\text{MAGMA-PRF}}(t, q) \lessapprox \min\left(\frac{t}{2^{256}}, \frac{q}{2^{96}}\right). \qquad (2)$$

Покажем, что в случае, если эвристическая оценка (2) верна (т. е. не существует методов, позволяющих при тех же ресурсах достичь большего преобладания; знаком «$\lessapprox$» обозначаем знак «$\leqslant$» в предположении, что множество алгоритмов противника ограничено классом известных методов), то при использовании MAGMA-PRF режимы шифрования CTR, CTR-ACPKM, GCM являются стойкими даже при значениях q, близких к 2^{64}.

Определение 2. (IND-CPNA)-преобладанием противника $\mathcal{A}$ в задаче различения режима шифрования MODE, в котором используется ключевое преобразование F, и оракула \$ (модель угроз IND-CPNA), назовём величину

$$\mathrm{Adv}^{\text{IND-CPNA}}_{\mathrm{MODE}[F]}(\mathcal{A}) = \mathsf{P}\left[K \xleftarrow{\$} V^k; \mathcal{A}^{\mathrm{MODE}[F_K](\cdot,\cdot)} \Rightarrow 1\right] - \mathsf{P}\left[\mathcal{A}^{\$(\cdot,\cdot)} \Rightarrow 1\right].$$

Запрос противника $\mathcal{A}$ к оракулу состоит из сообщения M и уникальной (неповторяющейся) синхропосылки IV. Оракул $\mathrm{MODE}[\mathrm{F}_K]$ на запрос $\mathcal{A}$ возвращает результат зашифрования сообщения M на синхропосылке IV, оракул \$ — случайную равновероятную последовательность бит такой же длины. К ресурсам противника относятся: t — количество вычислительных операций, q' — максимальное число запросов к оракулу, l — максимальная длина сообщения в запросе (в n-битных блоках).

Далее будем обозначать через $\mathrm{Adv}^{\mathrm{TM}}_{\mathrm{MODE}[F]}(t, q', \dots) = \max\limits_{\mathcal{A}:\ t_{\mathcal{A}} \leqslant t, q'_{\mathcal{A}} \leqslant q', \dots} \mathrm{Adv}^{\mathrm{TM}}_{\mathrm{MODE}[F]}(\mathcal{A})$ наибольшее преобладание среди всех возможных противников, действующих в некоторой модели угроз TM и ограниченных ресурсами t, q' (и другими видами ресурсов, определяемыми в рамках модели угроз TM).

Режим CTR является одним из классических режимов шифрования и определён, в том числе, в ГОСТ 34.13-2018 [7] (режим гаммирования). При использовании в нём MAGMA-PRF верна следующая

Теорема 1. Для преобладания противника в модели угроз IND-CPNA для режима CTR[MAGMA-PRF] справедливо неравенство

$$\mathrm{Adv}^{\text{IND-CPNA}}_{\text{CTR[MAGMA-PRF]}}(t, q', l) \leqslant \mathrm{Adv}^{\mathrm{PRF}}_{\text{MAGMA-PRF}}(t', q = q' \cdot l),$$

где $t' = t + O(q)$.

Утверждение теоремы следует непосредственно из классического результата для режима CTR [3]. В силу особенностей режима CTR по ГОСТ 34.13-2018, в котором уникальная синхропосылка занимает половину блока счётчика, длина шифруемого сообщения l до смены синхропосылки не должна превышать 2^{32} блоков.

Режим CTR-ACPKM определён в Р 1323565.1.017-2018 [8] (режимы работы блочных шифров со сменой ключа). При использовании в нём MAGMA-PRF верна следующая

Теорема 2. Для преобладания противника в модели угроз IND-CPNA для режима CTR-ACPKM[MAGMA-PRF] справедливо неравенство

$$\mathrm{Adv}^{\text{IND-CPNA}}_{\text{CTR-ACPKM[MAGMA-PRF]}}(t, q', l) \leqslant m \cdot \mathrm{Adv}^{\mathrm{PRF}}_{\text{MAGMA-PRF}}(t', q = \sigma + 4),$$

где $\sigma = q' \cdot N/64$ — объём секции (в блоках); $m = \lceil l \cdot 64/N \rceil$ — количество секций в сообщении; N — длина одной секции в битах, параметр режима CTR-ACPKM; $t' = t + O(q)$.

Утверждение теоремы следует непосредственно из доказательства [9] для случая PRP, при этом при использовании PRF отсутствуют квадратичные слагаемые, отвечающие за переход от PRP к PRF в процессе сведéния.

Режим GCM является режимом аутентифицированного шифрования (AEAD), обеспечивающим одновременно конфиденциальность и целостность сообщений, и определён в NIST SP800-38D [10].

Для AEAD-режимов рассматриваются модели угроз Priv (IND-CPNA) и Auth. В модели Auth противник взаимодействует с парой оракулов: левый оракул Enc_K выполняет зашифрование и вычисление имитовставки, а правый оракул Dec_K проверяет имитовставку и выполняет расшифрование в случае её корректности либо возвращает ошибку. $\mathrm{Adv}^{\mathrm{Auth}}_{\mathrm{AEAD[F]}}(\mathcal{A})$ определяется как вероятность подделки имитовставки (навязывания сообщения) противником.

При использовании MAGMA-PRF в режиме GCM верна следующая

Теорема 3. Для режима GCM[MAGMA-PRF] справедливы неравенства

$$\mathrm{Adv}^{\mathrm{Priv}}_{\mathrm{GCM[MAGMA\text{-}PRF]}}(t, q', l) \leqslant \mathrm{Adv}^{\mathrm{PRF}}_{\mathrm{MAGMA\text{-}PRF}}(t', q = q' \cdot l),$$

$$\mathrm{Adv}^{\mathrm{Auth}}_{\mathrm{GCM[MAGMA\text{-}PRF]}}(t, q', \nu, l) \leqslant \mathrm{Adv}^{\mathrm{PRF}}_{\mathrm{MAGMA\text{-}PRF}}(t', q = q' + \nu + \sigma + 1) + \frac{\nu\,(l+1)}{2^{\tau}},$$

где ν — число запросов к правому оракулу (максимальное число попыток навязывания противником); σ — общая длина сообщений в блоках (с учётом ассоциированных данных); τ — длина имитовставки, параметр режима GCM; $t' = t + O(q)$.

Утверждение теоремы следует из результатов [1] для AES-PRF.

В перечисленных режимах работы MAGMA-PRF работает почти также быстро, как Магма. Для её вычисления требуется 64 бита дополнительной памяти, отведённых под хранение промежуточного состояния $E^{(r)}_K(x)$, и одна дополнительная операция XOR 64-битных слов $E^{(r)}_K(x)$ и $E_K(x)$.

Таким образом, показано, что использование конструкции MAGMA-PRF в режимах CTR, CTR-ACPKM и GCM позволяет обрабатывать количество материала, превышающее границу «парадокса дней рождения». Количество блоков сообщений, которые могут быть обработаны MAGMA-PRF на одном ключе, близко к 2^{64}. В то же время при использовании подстановки (в частности, базового блочного шифра Магма) в данных режимах количество материала, который допускается обрабатывать на одном ключе, много меньше и, как правило, оценивается величиной 2^{32}.

ЛИТЕРАТУРА

1. *Mennink B. and Neves S.* Optimal PRFs from blockcipher designs // IACR Trans. Symmetric Cryptology. 2017. No. 3. P. 228–252.
2. ГОСТ 34.12-2018. Информационная технология. Криптографическая защита информации. Блочные шифры. М.: Стандартинформ, 2018.
3. *Bellare M. and Rogaway P.* Introduction to Modern Cryptography. University of California at Davis, 2005. 283 p.
4. *Isobe T.* A single-key attack on the full GOST block cipher // J. Cryptol. 2013. V. 26. No. 1. P. 172–189.
5. *Dinur I., Dunkelman O., and Shamir A.* Improved attacks on full GOST // LNCS. 2012. V. 7549. P. 9–28.
6. *Derbez P., Iwata T., Sun L., et al.* Cryptanalysis of AES-PRF and its dual // IACR Trans. Symmetric Cryptology. 2018. No. 2. P. 161–191.

7. ГОСТ 34.13-2018. Информационная технология. Криптографическая защита информации. Режимы работы блочных шифров. М.: Стандартинформ, 2018.
8. Р 1323565.1.017-2018. Информационная технология. Криптографическая защита информации. Криптографические алгоритмы, сопутствующие применению алгоритмов блочного шифрования. М.: Стандартинформ, 2018.
9. *Ahmetzyanova L. R., Alekseev E. K., Sedov G. K., et al.* Practical significance of security bounds for standardized internally re-keyed block cipher modes // Математические вопросы криптографии. 2019. Т. 10. № 2. С. 31–46.
10. *Dworkin M.* NIST SP 800-38D. Recommendation for Block Cipher Modes of Operation: Galois/Counter Mode (GCM) and GMAC. Technical Report. Gaithersburg, MD, United States, 2007.

UDC 519.719.2 DOI 10.17223/2226308X/16/28

PUBLIC KEYS FOR E-COINS: PARTIALLY SOLVED PROBLEM USING SIGNATURE WITH RERANDOMIZABLE KEYS

A. A. Babueva, S. N. Kyazhin

We give an example of an existing cryptographic mechanism that can be considered as a partial solution to the problem "Public keys for e-coins" proposed at the International Olympiad in Cryptography NSUCRYPTO'2022. This mechanism is used with the class of signatures with rerandomizable keys and provides one of the two security properties required by the authors of the problem. The results of this paper contain a systematic description of security models that can be used to analyze signature with rerandomizable keys, which is of independent interest.

Keywords: *public key derivation, signature with rerandomizable keys, related key attack, BIP32, NSUCRYPTO.*

1. Introduction

The unsolved problem «Public keys for e-coins» [1] was given as one of the tasks at the International Olympiad in Cryptography NSUCRYPTO'2022. The problem is to propose a way to calculate the public keys $\mathsf{pk}_1, \ldots, \mathsf{pk}_n$ (corresponding to the private keys $\mathsf{sk}_1, \ldots, \mathsf{sk}_n$), which can be used to verify the transaction authenticity, based on a single master key pk_0 (i.e., *public key derivation* scheme): $\mathsf{pk}_i = f(\mathsf{pk}_{i-1}, T)$, $i = 1, \ldots, n$. The proposed method, according to the requirements proposed by the authors of the problem, should provide the following security properties:

1) knowing pk_0, f and T, it is impossible to find any private key sk_i, $i = 1, \ldots, n$;
2) it is impossible to recover sk_i, if the secret keys $\mathsf{sk}_1, \ldots, \mathsf{sk}_{i-1}, \mathsf{sk}_{i+1}, \ldots, \mathsf{sk}_n$ are known.

However, in general, such requirements are not enough for key derivation, since the security of the mechanisms used to authenticate transaction is often analyzed under the assumption of a random and independent key selection. The solution is to analyze the joint security of such mechanism and key derivation process. The influence of key derivation on the mechanism used for the transaction authenticity can be described with a well-known type of attacks: related key attacks [2].

The problem does not limit the mechanism that is used to authenticate the transaction, however, signature schemes are most often used for this purpose. In this paper, we describe the interface of a modified signature scheme (the so-called signature with rerandomizable keys), and also systematize security models for its analysis. All the models considered

describe a stronger property compared to the property 1. In addition, we give an example of the existing public key derivation mechanism for the ECDSA signature scheme [3] described in Bitcoin Information Proposal BIP32 [4], as well as the results of the ECDSA security analysis in one of the described models.

2. Signature with rerandomizable keys

Let's describe the interface of the modified signature scheme using the standard signature scheme interface. We adopt the definition of signature with rerandomizable keys from [5, 6].

Definition 1. A *signature* Sig is a tuple of the following algorithms:

— Sig.Gen is the probabilistic key generation algorithm, takes as input public parameters par, returns a pair $(\mathsf{sk}, \mathsf{pk})$ of secret and public keys;
— Sig.Sign is the probabilistic signing algorithm, takes as input a secret key sk and a message m, returns a signature σ;
— Sig.Verify is the deterministic verification algorithm, takes as input a public key pk, a signature σ, and a message m, returns 1 (accept) or 0 (reject).

Definition 2. A *signature with rerandomizable keys* RSig is a tuple of the following algorithms:

— RSig.Gen, RSig.Sign, RSig.Verify are the algorithms as defined above;
— RSig.RandSK is the probabilistic secret key rerandomization algorithm, takes as input a secret key sk and randomness $\rho \in \Omega$, returns a rerandomized secret key sk';
— RSig.RandPK is the probabilistic public key rerandomization algorithm, takes as input a public key pk and randomness $\rho \in \Omega$, returns a rerandomized public key pk'.

Remark 1. The public key derivation mechanism defines the mechanism for generating ρ values and is not part of the RSig.

Remark 2. In general, the Ω set from which the ρ values can be selected is limited. But we did not include this in the system interface due to a remark 1.

3. Security models for signature with rerandomizable keys

The standard security requirement for the signature scheme is the unforgeability property which is formalized by UF-CMA notion [7]. The adversary is allowed to obtain the signatures for adaptively chosen messages. It's task is to provide a forgery, i.e., (signature, message) pair which is correct and non-trivial.

The rerandomizable key usage expands the ways to define both the type of attack and the threat. In the current section we provide the survey of the known security models for such class of signature schemes.

The attack. The adversary is allowed to obtain the signatures computed not only with the original secret signing key sk, but also with the modified keys produced by RandSK algorithm.

We define two types of attack depending on which keys can be queried for the signatures:

— keys produced with an arbitrary randomness ρ from the predefined set Ω;
— keys produced with the randomness honestly chosen by the challenger during the game processing. The ability of the adversary to know these randomness values is captured by the access to the *Rand* oracle.

Further we differ these types of attack in the model name by RKA (Related Key Attack) and KRKA (Known Related Key Attack) respectively. Clearly, first type of attack is stronger, i.e., the security in $*$-RKA model implies the security in $*$-KRKA model.

The threat. One way to define the threat is to do it similarly to the standard UF-CMA notion, i.e., making the forgery (m^*, σ^*) for the original public key pk. The only ambiguity here is defining the triviality of the forgery.

We define two types of threat depending on the forgery for which message should be done for winning the game:

— message m^* should not be queried for signing;
— message m^* should not be queried for signing with the original signing key sk.

Further we differ these types of threat in the model name by wUF (weak UnForgeability) and UF (UnForgeability) respectively. Clearly, first type of threat is stronger (the model is weaker), i.e., the security in UF-$*$ model implies the security in wUF-$*$ model.

The formal definition of the corresponding security models is presented in Fig. 1. Here, the basic UF-CM-RKA model is defined by black color, it was introduced in [2]. The modification of this model by adding blue color strings defines the UF-CM-KRKA model [7], and by adding red color strings — wUF-CM-RKA model [8]. Finally, all strings together form the wUF-CM-KRKA model.

$\mathbf{Exp}^{\text{wUF-CM-KRKA}}_{\mathsf{RSig}}(\mathcal{A})$

1 : $(\mathsf{sk}, \mathsf{pk}) \xleftarrow{\$} \mathsf{RSig.Gen}(\,)$
2 : $\mathcal{L}, \mathcal{R} \leftarrow \varnothing$
3 : $(m^*, \sigma^*) \xleftarrow{\$} \mathcal{A}^{Sign, Rand}(\mathsf{pk})$
4 : **if** $m^* \in \mathcal{L}$: **return** 0
5 : $res \leftarrow \mathsf{RSig.Verify}(\mathsf{pk}, m^*, \sigma^*)$
6 : **return** res

Oracle $Sign(\rho, m)$

1 : **if** $(\rho \notin \Omega) \vee (\rho \notin \mathcal{R})$:
2 : **return** $\perp$
3 : $\mathsf{sk}' \leftarrow \mathsf{RSig.RandSK}(\mathsf{sk}, \rho)$
4 : $\sigma \leftarrow \mathsf{RSig.Sign}(\mathsf{sk}', m)$
5 : **if** $(\mathsf{sk}' = \mathsf{sk}) \vee (\mathsf{sk}' \neq \mathsf{sk})$:
6 : $\mathcal{L} \leftarrow \mathcal{L} \cup \{m\}$
7 : **return** σ

Oracle $Rand(\,)$

1 : $\rho \xleftarrow{\mathcal{U}} \Omega$
2 : $\mathcal{R} \leftarrow \mathcal{R} \cup \{\rho\}$
3 : **return** ρ

Fig. 1. UF-CM-RKA, UF-CM-KRKA, wUF-CM-RKA, wUF-CM-KRKA models definition

Another way to define the threat is to allow the adversary to make the forgery for any public key, not only the original key pk. In this case, the adversary returns the triple (ρ^*, m^*, σ^*), where ρ^* defines the public key pk^* for which the forgery is made. The randomness ρ^* should belong to the set Ω or should be obtained as a result of query to the *Rand* oracle (depending on RKA or KRKA attack type). Following the paper [7], we refer to the models with such definition of the threat by adding "s" (strong) before the attack type. Clearly, such models are stronger then the corresponding models without "s", i.e., the security in $*$-sRKA ($*$-sKRKA) model implies the security in $*$-RKA ($*$-KRKA) model.

Similarly to the previous models, there are two possible ways to determine the triviality of the forgery:

— pair (ρ^*, m^*) should be fresh, i.e., should not be queried to the *Sign* oracle (UF-* model);
— message m^* should be fresh, i.e., should not be queried to the *Sign* oracle (wUF-* model).

The formal definition of the corresponding security models is presented in Fig. 2. Here, the basic UF-CM-sRKA model is defined by black color, it was introduced in [2]. The modification of this model by adding blue color strings defines the UF-CM-sKRKA

model [7], and by adding red color strings — wUF-CM-sRKA model [6]. Finally, all strings together form the wUF-CM-sKRKA model [5].

$\mathbf{Exp}^{\text{wUF-CM-sKRKA}}_{\mathsf{RSig}}(\mathcal{A})$

1 : $(\mathsf{sk}, \mathsf{pk}) \xleftarrow{\$} \mathsf{RSig.Gen}(\ )$
2 : $\mathcal{L}, \mathcal{R} \leftarrow \varnothing$
3 : $(\rho^*, m^*, \sigma^*) \xleftarrow{\$} \mathcal{A}^{Sign, Rand}(\mathsf{pk})$
4 : **if** $(\rho^* \notin \Omega) \vee (\rho^* \notin \mathcal{R})$:
5 : **return** $\perp$
6 : **if** $((\rho^*, m^*) \in \mathcal{L}) \vee ((\cdot, m^*) \in \mathcal{L})$:
7 : **return** 0
8 : $\mathsf{pk}^* \leftarrow \mathsf{RSig.RandPK}(\mathsf{pk}, \rho^*)$
9 : $res \leftarrow \mathsf{RSig.Verify}(\mathsf{pk}^*, m^*, \sigma^*)$
10 : **return** res

Oracle $Sign(\rho, m)$

1 : **if** $(\rho \notin \Omega) \vee (\rho \notin \mathcal{R})$:
2 : **return** $\perp$
3 : $\mathsf{sk}' \leftarrow \mathsf{RSig.RandSK}(\mathsf{sk}, \rho)$
4 : $\sigma \leftarrow \mathsf{RSig.Sign}(\mathsf{sk}', m)$
5 : $\mathcal{L} \leftarrow \mathcal{L} \cup \{(\rho, m)\}$
6 : **return** σ

Oracle $Rand(\ )$

1 : $\rho \xleftarrow{\mathcal{U}} \Omega$
2 : $\mathcal{R} \leftarrow \mathcal{R} \cup \{\rho\}$
3 : **return** ρ

Fig. 2. UF-CM-sRKA, UF-CM-sKRKA, wUF-CM-sRKA, wUF-CM-sKRKA models definition

4. BIP32 scheme and its security

Let P be the generator of the elliptic curve point group, $(\mathsf{sk}, \mathsf{pk} = \mathsf{sk}P)$ — the ECDSA signature key pair, K — the HMAC [9] key.

The scheme described in BIP32 assumes the use of signature with rerandomizable keys based on ECDSA, where

$$\mathsf{RSig.RandSK}(\mathsf{sk}, \rho) = \mathsf{sk} + \rho,$$
$$\mathsf{RSig.RandPK}(\mathsf{pk}, \rho) = \mathsf{pk} + \rho P,$$

and the mechanism for generating ρ based on the algorithm HMAC.

Denote by $\mathsf{HMAC}_{l/2}(K, m)$ a function that returns the left $l/2$ bits of the l-bit result of the function $\mathsf{HMAC}(K, m)$. For simplicity, we will further omit the functions of converting a bit string into a group element and vice versa.

The mechanism of generating ρ_i, used to calculate the ith key pair $(\mathsf{sk}_i, \mathsf{pk}_i)$, $i = 1, \ldots, n$, from the key pair $(\mathsf{sk}_0, \mathsf{pk}_0 = \mathsf{sk}_0 P)$, is defined by the following function:

$$\rho_i = \mathsf{HMAC}_{l/2}(K, \mathsf{pk}_0 \| i).$$

Thus, in terms of the original problem, the key K plays the role of an parameter T, and the function f is represented as follows:

$$\mathsf{pk}_i = f(\mathsf{pk}_{i-1}, K) = \mathsf{pk}_{i-1} + \mathsf{HMAC}_{l/2}(K, \mathsf{pk}_0 \| i)P - \mathsf{HMAC}_{l/2}(K, \mathsf{pk}_0 \| i - 1)P.$$

Among the models described in the section 3, the UF-CM-sKRKA model seems relevant for analyzing this scheme, because:

— an actual threat is forgery with respect to at least one key pk_i, $i \in \{i, \ldots, n\}$ (strong threat is relevant);
— the adversary has the capability to get the ρ values, only calculated using the HMAC function (known related key attack is relevant).

The paper [7] shows that the ECDSA signature with rerandomizabe keys is secure in this model.

5. Conclusion

In this paper, we give an example of an existing cryptographic mechanism that can be considered as a partial solution to the problem "Public keys for e-coins" proposed at the International Olympiad in Cryptography NSUCRYPTO'2022. This mechanism is used with the class of signatures with rerandomizable keys and provides one of the two security properties required by the authors of the problem. The existence of mechanisms with the second property remains (hopefully temporarily) an unsolved problem.

The results of this paper contain a systematic description of security models that can be used to analyze signature with rerandomizable keys, which is of independent interest.

REFERENCES

1. Problem 10. "Public keys for e-coins". International Olympiad in Cryptography NSUCRYPTO'2022. https://nsucrypto.nsu.ru/archive/2022/round/2/section/0/task/10/.
2. *Bellare M., Cash D., and Miller R.* Cryptography secure against related-key attacks and tampering. LNCS, 2011, vol. 7073, pp. 486–503.
3. FIPS 186-5. Digital Signature Standard. https://csrc.nist.gov/publications/detail/fips/186/5/final.
4. BIP 32. Hierarchical Deterministic Wallets. https://github.com/bitcoin/bips/blob/master/bip-0032.mediawiki.
5. *Das P., Faust S., and Loss J.* A formal treatment of deterministic wallets. Proc. ACM SIGSAC Conf. CCS'19, N.Y., ACM, 2019, pp. 651–668.
6. *Fleischhacker N., Krupp J., Malavolta G., et al.* Efficient unlinkable sanitizable signatures from signatures with re-randomizable keys. LNCS, 2016, vol. 9614, pp. 310–330.
7. *Yuen Y. H. and Yiu S. M.* Strong known related-key attacks and the security of ECDSA. LNCS, 2019, vol. 11928, pp. 130–145.
8. *Morita H., Schuldt J. C. N., Matsuda T., et al.* On the security of the Schnorr signature scheme and DSA against related-key attacks. LNCS, 2016, vol. 9558, pp. 20–35.
9. *Bellare M., Canetti R., and Krawczyk H.* Keying hash functions for message authentication. LNCS, 1996, vol. 1109, pp. 1–15.

UDC 519.7 DOI 10.17223/2226308X/16/29

EFFICIENT MATRIX MULTIPLICATION FOR CRYPTOGRAPHY WITH A COMPANION MATRIX OVER $\mathbb{F}_2$[1]

S. Pal

A number of schemes in cryptography and other allied areas require operations on matrices that are computationally expensive. However, the computational load due to standard operations like multiplication can be drastically reduced by the choice of special matrices. One such special matrix is the companion matrix of a monic polynomial of degree n over a finite field. Due to its cyclic structure and sparseness property, such a matrix not only helps us to reduce the complexity of matrix multiplication but also can be applied for cryptographic purposes. In this paper, an algorithm is proposed for the multiplication of an arbitrary matrix with a companion matrix over a finite field of order p. In our algorithm, we not only reduce the complexity but also minimize the number of multiplication operations as much as possible. The complexity of multiplication of any $n \times n$ matrix with a companion

[1]The work was supported by the Mathematical Center in Akademgorodok under the agreement No. 075-15-2022-282 with the Ministry of Science and Higher Education of the Russian Federation.

matrix of a monic polynomial of degree n is $\mathcal{O}(n^2)$, whereas the complexity of standard matrix multiplication is $\mathcal{O}(n^3)$. Moreover, the number of multiplication operations is $n^2 - nt$, $0 \leqslant t < n$, and 0 for the fields $\mathbb{F}_p$ and $\mathbb{F}_2$ of order p and 2, respectively, which is far less than n^3 multiplications required for standard matrix multiplication.
Keywords: *companion matrix, matrix multiplication, cryptology.*

1. Matrix Multipication with the companion matrix of a monic polynomial

1.1. Motivation

Due to the rapid increase in 5G and 6G technologies, ARX (Addition, Rotation and XOR)-based schemes are more popular nowadays. Avoiding the multiplication operation provides the efficiency in the lightweight system. This is the reason why we focus on reducing multiplication operations as much as possible. Matrices are not generally used for cryptographic purposes due to their expensive operations like multiplication. Researcher [1] tried to find a suitable way to reduce the expensive operations by searching for special kinds of matrices. Our goal is making the matrices useful for cryptographic purposes. In the case of the multiplication of two matrices, all elements of both of these matrices are required. But the operations can be reduced by observing the elements and structures of the matrices. Moreover, multiplying a companion matrix of a monic polynomial over a finite field with another matrix does not require all elements of the companion matrix. It is explained in the contribution section that the new matrix will be obtained after multiplication by observing only the second matrix and the coefficients of the monic polynomial. For the sake of the simplicity of our work, we provide a modified definition of matrix multiplication, which is given below.

Definition 1. Matrix multiplication of the two $n \times n$ matrices involves the multiplication of i^{th} element of k^{th} row of the first matrix with i^{th} row of the second matrix, provides n new rows and the k^{th} row of the new matrix is obtained by adding them, where $1 \leqslant i, k \leqslant n$.

Definition 2. The companion matrix [2, 3] of a monic polynomial of degree n, i.e., $f(x) = a_0 + a_1x + \ldots + a_{n-1}x^{n-1} + x^n$ over a field of order p is the $n \times n$ matrix

$$C_f = \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{bmatrix}.$$

1.2. Contribution

Theorem 1. Multiplication of a companion matrix of a monic polynomial of degree n over finite field of order p by a second matrix gives a new matrix whose rows are as follows:

— The first row of the new matrix is equal to the r times of n^{th} row of the second matrix, where the constant coefficient of the monic polynomial is r, $0 < r < p$.
— The k^{th} row is equal to the $(k-1)^{th}$ row of the second matrix, if the k^{th} coefficient of the monic polynomial is zero, where $2 \leqslant k \leqslant n$.
— The k^{th} row of the new matrix is equal to the summation of the $(k-1)^{th}$ row and r times of n^{th} row of the second matrix, if the k^{th} coefficient of the monic polynomial is r, where $2 \leqslant k \leqslant n$ and $0 < r < p$.

Corollary 1. Multiplication of a companion matrix of a monic polynomial of degree n over a finite field of order 2 by a second matrix gives a new matrix whose rows are as follows:

— The first row is equal to the last, i.e., n^{th} row of the second matrix.
— The k^{th} row is equal to the $(k-1)^{th}$ row of the second matrix, if the k^{th} coefficient of the monic polynomial is zero, where $2 \leqslant k \leqslant n$.
— The k^{th} row is equal to the summation of the $(k-1)^{th}$ row and n^{th} row of the second matrix, if the k^{th} coefficient of the monic polynomial is one, where $2 \leqslant k \leqslant n$.

Now, we can explain our Algorithm 1 to calculate an $n \times n$ new matrix $D = C_f \times B$, where d_{ij} and b_{ij} are the coefficients of $n \times n$ matrices D and B over any field respectively, $1 \leqslant i, j \leqslant n$; C_f is a companion matrix of a monic polynomial $f(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n$ over field $\mathbb{F}_p$ of order p. For the sake of applicability, we take the constant term of the polynomial as non-zero for our work.

Algorithm 1.

Input: $f(x) = a_0 + a_1 x + \ldots + a_n x^n$ be a monic polynomial and B be any $n \times n$ matrix with coefficient b_{ij}, $0 \leqslant i, j \leqslant n-1$.
Output: New matrix D with coefficient d_{ij}, $0 \leqslant i, j \leqslant n-1$.

Set $a_i^{'} = p - a_i$, $0 \leqslant i \leqslant n-1$.
For $i = 0, \ldots, n-1$ **do**:
  For $j = 0, \ldots, n-1$ **do**:
    If $i = 0$, **then** $d_{ij} := (a_i^{'} \cdot b_{(n-1)j})$;
    else if $a_i = 0$, **then** $d_{ij} \leftarrow b_{(i-1)j}$;
      else $d_{ij} := (a_i^{'} \cdot b_{(n-1)j}) + b_{(i-1)j}$.

Example 1. Let C_f be the companion matrix of a monic polynomial $f(x) = x^3 + 4x^2 + 3$ of degree 3 over $\mathbb{F}_5$, i.e., $C_f = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$, and B be any 3×3 matrix over any field, i.e., $B = \begin{bmatrix} 2 & 1 & 4 \\ 1 & 3 & 5 \\ 2 & 0 & 3 \end{bmatrix}$, then $D = C_f B = \begin{bmatrix} 4 & 0 & 6 \\ 2 & 1 & 4 \\ 3 & 3 & 8 \end{bmatrix}$. Here, $D_1 = 2 \cdot B_3$, $D_2 = B_1$, $D_3 = B_2 + B_3$, where D_i and B_i, $1 \leqslant i \leqslant 3$, are the rows of matrices D and B, respectively.

2. Complexity calculation for matrix multiplications

In this section, we calculate the complexity of matrix multiplication followed by counting the number of operations required for multiplication of two $n \times n$ matrices using the standard matrix multiplication method and our method, which includes a special matrix, the companion matrices of a monic polynomial of degree n over a field of order 2 and p. Multiplication and division operations are computationally expensive compared to addition, subtraction, and shift operations. The table shows that only addition and multiplication operations are required for standard matrix multiplication, whereas addition, multiplication, and shift operations are required for matrix multiplication by our method over $\mathbb{F}_p$. Most importantly, only addition and shift operations are required for matrix multiplication by our method over $\mathbb{F}_2$.

Operations required for different type of matrix multiplications

Operations List			
Matrix multiplication	Number of additions	Number of multiplications	Number of shift operations
Standard	$n^3 - n^2$	n^3	0
Using companion matrix over $\mathbb{F}_2$	$n^2 - nt - n$	0	$nt + n$
Using companion matrix over $\mathbb{F}_p$	$n^2 - nt - n$	$n^2 - nt$	nt

Lemma 1. Standard matrix multiplication of two $n \times n$ matrices requires n^3 multiplications, $n^3 - n^2$ additions, $2n^3$ operations in total. The complexity is $\mathcal{O}(n^3)$.

Lemma 2. Matrix multiplication by a companion matrix of a monic polynomial of degree n over $\mathbb{F}_p$ requires nt shift operations, $n^2 - nt$ multiplications, and $n^2 - nt - n$ additions, $2n^2 - nt - n$ operations in total. Here, t is the number of rows of the companion matrix of the monic polynomial whose t^{th} coefficient is zero. The complexity is $\mathcal{O}(n^2)$.

Lemma 3. Matrix multiplication by a companion matrix of a monic polynomial of degree n over $\mathbb{F}_2$ requires $nt + n$ shift operations and $n^2 - nt - n$ additions, n^2 operations in total. Here, t is the number of rows of the companion matrix of the monic polynomial whose t^{th} coefficient is zero. The complexity is $\mathcal{O}(n^2)$.

REFERENCES

1. *Mahalanobis A.* Are matrices useful in public-key cryptography? Intern. Math. Forum, 2013, vol. 8, no. 39, pp. 1939–1953.
2. *Herstein I. N.* Topics in Algebra, 2nd ed. John Wiley & Sons, 2006.
3. *Ghorpade S. R. and Ram S.* Block companion Singer cycles, primitive recursive vector sequences and coprime polynomial pairs over finite fields. Finite Fields Their Appl., 2011, vol. 17, no. 5, pp. 461–472.

UDC 519.7 DOI 10.17223/2226308X/16/30

CRYPTANALYSIS OF LWE AND SIS-BASED CRYPTOSYSTEMS BY USING QUANTUM ANNEALING[1]

A. Qayyum, M. Haris

In the paper, we study lattice-based cryptographic problems, in particular Learning With Errors (LWE) and Short Integer Solution (SIS) lattice problems, which are considered to be known cryptographic primitives that are supposed to be secure against both classical and quantum attacks. We formulated the LWE and SIS problems as Mixed-Integer Programming (MIP) model and then converted them to Quadratic Unconstrained Binary Optimization (QUBO) problem, which can be solved by using a quantum annealer. Quantum annealing searches for the global minimum of an input objective function subjected to the given constraints to optimize the given model. We have estimated the q-bits required for the Quantum Processing Unit (QPU). Our results show that this approach can solve certain instances of the LWE and SIS problems efficiently.

Keywords: *post-quantum cryptography, lattice-based cryptography, learning with errorss, short integer solution, quadratic unconstraint binary optimization, quantum processing unit.*

[1]The work is supported by the Mathematical Center in Akademgorodok under the agreement No. 075-15-2022-282 with the Ministry of Science and Higher Education of the Russian Federation.

1. Introduction

Quantum computers hold a promise to solve many computational problems faster than classical computers. Because of this potential, researchers are interested in exploring quantum computing and their implications for different fields including cryptography. The advent of quantum computing has a serious threat to many cryptography foundations. Because of this concern, many researchers have started looking at Post Quantum Cryptography (PQC). In this scenario, the National Institute of Standards and Technology (NIST) motivated and invited researchers to submit the PQC algorithms. After analyzing these submissions, NIST announced the selected algorithms for PQC in 2022. The selected algorithm for public-key encryption and key establishment is CRYSTALS-Kyber [1]. The CRYSTALS-DILITHIUM [2], Falcon [3] and the SPHINCS+ [4] are selected for digital signatures. Most of the selected algorithms are based on the lattice. Therefore, we are exploring lattice-based cryptographic problems, specifically Learning with Errors (LWE) and Short Integer Solution (SIS) lattice problems.

Lattice-based cryptography is an asymmetric cryptography that uses mathematical structure called lattice to design cryptographic primitives. Its security is based on the hardness of certain lattice problems, such as the Shortest Vector Problem (SVP), Shortest Independent Vectors Problem (SIVP), Closest Vector Problem (CVP), Short Integer Solution (SIS), and Learning with Errors (LWE) [5]. To solve these lattice problems, there are several methods such as lattice basis reduction, sieving, and enumeration.

In the context of solving SIS and LWE problems, one method is known as lattice basis reduction, which includes algorithms such as LLL [6] and BKZ [7] to reduce the input lattice and extract short basis vectors that can be used to solve SIS and LWE. Another method is sieving, which includes the Gauss-Sieve algorithm [8] and Block-Kannan algorithm [9]. This technique is used to exploit the statistical properties and do the sampling from a high-dimensional lattice to find for the vectors close to the origin. The other technique is enumeration, which includes Voronoi algorithm [10], Babai's Nearest Plane [11]. This technique aims to find all lattice points within a certain radius and eliminate the irrelevant vectors.

In the thesis, we investigate the complexity of the implementation of LWE and SIS problems using quantum annealing. We propose Mixed-Integer Programming (MIP) model for SIS and LWE. MIP is a mathematical optimization technique which provides the solution for both discrete decision and continuous variables problems. We used the PuLP library of Python which is a widely-used tool for implementing MIP models. It offers efficient solvers that can effectively address optimization problems, such as our model for SIS and LWE. Furthermore, we convert them to their equivalent QUBO model and estimate the required q-bits.

2. Quantum annealing and adiabatic computing

Quantum annealing is a type of quantum computing that aims to solve optimization problems by minimizing the energy of a physical system. For optimization, quantum annealing searches for the global minimum of the input objective function. Quantum annealer put the states in superposition at the start. Then, these states alter by quantum physics, which is beyond our control. So we give the Quadratic Unconstrained Binary Optimization (QUBO) problem at the beginning, and the configuration at the end corresponds to the solution [12]. Quantum annealing is related to adiabatic quantum computing, which is a specific form of quantum annealing that works on the energy minimization process.

QUBO is a form of binary quadratic model which uses binary variables to represent the problem. It involves finding the values of binary variables that minimize a quadratic function. It is a unique class of equations that corresponds to the design of a quantum processing unit (QPU). The QPU is made up of q-bits and couplers which connect pairs of q-bits. We can consider the q-bits as variables in equation, and couplers as pairs of variables that are multiplied together making up a quadratic equation [13].

3. Learning with Errors (LWE)

A lattice is a geometric structure which contains a set of points in n-dimensional space. It is widely used in different areas of study including lattice-based cryptography. LWE problem is one of the lattice hard problems. It is considered NP-hard. It is used in many cryptographic schemes such as encryption, digital signatures, and key-exchange algorithms. Formally, lattice and LWE are defined as follows:

A lattice $\mathcal{L}$ in $\mathbb{R}^n$ is a discrete additive subgroup of $\mathbb{R}^n$. It is generated by a set of linearly independent vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$, called a basis for the lattice [14]:

$$\mathcal{L} = \left\{ \sum_{i=1}^{n} a_i \mathbf{b}_i : a_i \in \mathbb{Z} \right\}$$

In the LWE we have a list of "equations with errors" such that

$$\begin{aligned}
\sum_{j=1}^{n} a_{1,j} s_j &\approx_\chi b_1 \pmod q, \\
\sum_{j=1}^{n} a_{2,j} s_j &\approx_\chi b_2 \pmod q, \\
&\vdots \\
\sum_{j=1}^{n} a_{m,j} s_j &\approx_\chi b_m \pmod q,
\end{aligned}$$

where $i = 1, 2, \ldots, m$, $q = q(n) \leqslant \mathrm{poly}(n)$ is prime integer, $a_{1j}, a_{2j} \ldots a_{mj} \in \mathbb{Z}_q^n$ are chosen independently and uniformly, $s \in \mathbb{Z}_q^n$, and $b_i \in \mathbb{Z}_q$. The errors $e_i \in \mathbb{Z}_q$ in the equations are specified by a probability distribution $\chi : \mathbb{Z}_q \to \mathbb{R}_+$ on $\mathbb{Z}_q$. The $e_i \in \mathbb{Z}_p$ are chosen independently from χ [15]. The goal is to find the secret key s given a set of noisy linear equations.

LWE Problem Hardness

LWE is not proven NP-hard problem, but it is as hard as certain worst case lattice problems, such as CVP or SVP. It is proven that LWE is at least as hard as SIVP NP-hard problem [15]. We try to approximate the vector s_j by z_j by:

$$\begin{aligned}
b_i &= \sum_{j=1}^{n} a_{ij} s_j + e_i \mod q \text{ — actual value}, \\
\widetilde{b}_i &= \sum_{j=1}^{n} a_{ij} z_j + r_i \mod q \text{ — predicted value}.
\end{aligned}$$

Our goal is to minimize the difference between b_i and $\widetilde{b}_i$ such that the vector z_j becomes close to the vector s_j. We choose r_i parameter from the same distribution as e_i.

Mathematical Model of LWE Problem

Objective function is

$$\min \sum_{i=1}^{m} t_i.$$

The objective of our optimization model is to minimize the sum of t_i variables which we will define in (1) and (2), subjected to the following constraints:

— we ensure that for each j only one x_{jk} variable can take the value 1 and all other entries will be equal to 0:

$$\sum_{k=0}^{q-1} x_{jk} = 1;$$

— we define the value of z_j as a linear combination of the x_{jk} variables:

$$z_j = \sum_{k=0}^{q-1} (k \cdot x_{jk});$$

— the main constraint, which defines the LWE problem, is the following:

$$\sum_{j=1}^{n} a_{ij} z_j + r_i = D_i q + \widetilde{b}_i$$

Here, z_j are the unknowns that we want to find, all the operations are made modulo q, and we introduce additional variable D_i to linearize it;

— we define the value t_i, that is the the difference between actual b_i and predicted $\widetilde{b}_i$, which we want to minimize. The variable C_i is introduced to linearize the equation:

$$b_i - \widetilde{b}_i = C_i q + t_i; \tag{1}$$

— we set the upper bound for $\widetilde{b}_i$:

$$\widetilde{b}_i \leqslant q - 1;$$

— we set the upper bound for t_i:

$$t_i \leqslant q - 1. \tag{2}$$

The variables belong to the following sets:

$$x_{jk} \in \{0, 1\}, \quad C_i, D_i \in \mathbb{Z}, \quad z_j, \widetilde{b}_i \in \mathbb{Z}_q$$

The parameters belong to the following sets:

$$a_{ij}, b_i, r_i \in \mathbb{Z}_q, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n, \quad k = 0, 1, \ldots, q-1.$$

To perform the experiments for the above mathematical model, we fix the parameters $s_j \in \mathbb{Z}_q^n$, $a_{ij} \in \mathbb{Z}_q^n$ are chosen independently and uniformly, the parameter error $e_i \in \mathbb{Z}_q$ is chosen from the uniform distribution, and $b_i \in \mathbb{Z}_q$ is calculated by the following equation:

$$\sum_{j=1}^{n} a_{ij} s_j + e_i = F_i q + b_i$$

where $F_i \in \mathbb{Z}_q$. After fixing these parameters, we ran our MIP model and obtained the following results (the Table):

n, m, q	No. of runs	Mean $\sum t_i$	Max Objective Function	Mean $\sum s_j - z_j \mod q$
3, 4, 23	20	3.15	5	0.25
4, 5, 23	20	2.05	4	0.15
5, 6, 23	20	1.9	3	0.9
6, 7, 23	20	1.75	3	0.45

After verifying our linear mathematical model, we converted it to QUBO model by converting all the variables into binary variables and introducing the quadratic term in the objective function. We also estimated the number of q-bits of our QUBO model by calculating the q-bits for each binary and integer variables used in the objective function and all the subjected constraints.

Proposition 1. The minimal number of qubits required for the implementation of the considered QUBO model of LWE for QPU is at most

$$3nq + 17mq + m \times \left\lceil \log_2 \left(\left\lceil \frac{n(q-1)^2 + (q-1)}{q} \right\rceil \right) \right\rceil + \\ +4m \times \left\lceil \frac{n(q-1)^2 + (q-1)}{q} \right\rceil + 5m \lceil \log_2 q \rceil + mn \lceil \log_2 q \rceil .$$

The above estimation of q-bits is theoretical. In the future, we will convert our QUBO Model into Ising Model to observe the practical number of q-bits required for QPU with the help of D-wave Ocean Software.

4. Short Integer Solution (SIS)

SIS problem is one of the lattice hard problem. It is widely used in different cryptographic schemes such as encryption, digital signatures, key-exchange algorithms. The security of many lattice-based cryptographic schemes, such as the Ring Learning with Errors (RLWE) scheme, is based on the hardness of the SIS problem. Formally, we can define SIS as follows:

Given a matrix $A \in \mathbb{Z}^{n \times m}$ with $m > n$ and a positive integer β, find a non-zero vector $z \in \mathbb{Z}^n$ such that:

$$\|z\| \leqslant \beta,$$
$$Az = 0 \mod q,$$

where $\beta \geqslant \sqrt{m \log q}$, q is a prime integer and $\|.\|$ is Euclidean norm [5].

SIS Problem Hardness

SIS problem is not proven NP-hard problem but it is as hard as certain worst case lattice problems, such as CVP or SVP, which are known to be NP-hard problems [5]. We want to minimize $\|z\|$ under following constraints:

$$\|z\| \leqslant \beta, \quad Az = 0 \mod q.$$

But $\|z\|$ is non-linear function and we applied the techniques to linearize it.

Mathematical Model of SIS Problem

Objective function is

$$\min \sum_j (u_j + v_j).$$

The objective of our optimization model is to minimize the values of u_j and v_j variables, which we will define in (3), subjected to the following constraints:

— we ensure that for each j only one x_{jk} variable can take the value 1 and all other entries will be 0:

$$\sum_{k=0}^{q-1} \left(x_{jk}^{+} + x_{jk}^{-}\right) = 1, \quad j = 1, 2, \ldots, n;$$

— we define the value of z_j, which is a linear combination of the x_{jk} variables:

$$z_j = \sum_{k=0}^{q-1} k \left(x_{jk}^{+} - x_{jk}^{-}\right), \quad j = 1, 2, \ldots, n;$$

— we define the SIS problem, where z_j are the unknowns that we want to find:

$$\sum_j a_{ij} z_j = C_i \cdot q, \quad i = 1, 2, \ldots, m;$$

—

$$u_j = \sum_k k x_{j,k}^{+}, \quad v_j = \sum_k k x_{j,k}^{-}, \quad j = 1, 2, \ldots, n; \tag{3}$$

— we ensure that at least one element of the vector z is non-zero:

$$\sum_j (u_j + v_j) \geqslant 1;$$

— we ensure the upper bound of $u_j + v_j$:

$$u_j + v_j \leqslant q - 1, \quad u_j + v_j \leqslant \beta;$$

— $z_j, C_i \in \mathbb{Z}$; $x_{jk}^{+}, x_{jk}^{-} \in \{0, 1\}$, $u_j, v_j \in \mathbb{Z}_{\geqslant 0}$ — variables; $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, $k = 0, 1, \ldots, q-1$ — parameters.

We converted the above model to QUBO model by converting all the variables into binary variables and introducing the quadratic term in the objective function. We also estimated the number of q-bits of our QUBO Model by calculating the q-bits for each binary and integer variables used in the objective function and all the subjected constraints.

Proposition 2. The minimal number of qubits required for the implementation of the considered QUBO model of SIS for QPU is at most

$$16nq + 2m \times \left\lceil \frac{n(q-1)^2}{q} \right\rceil + mn \left\lceil \log_2 q \right\rceil.$$

5. Conclusion

We have investigated the complexity of implementation of LWE and SIS problems using quantum annealing. First, we introduced the Mixed-Integer Programming model for LWE and SIS. Next, we presented our experimental results by using Python library called PuLP. Finally, we formulated our mathematical model into QUBO and estimated the number of q-bits required for the Quantum Processing Unit to perform Quantum Annealing for our QUBO Models. After analyzing the LWE and SIS problem, we can conclude that the cryptanalysis on the certain instances of LWE problem is possible by using QUBO Model. In future, we will consider the r_i in terms of the probability distribution for the LWE problem. We will convert our QUBO Model into Ising Model to observe the practical number of q-bits required for QPU with the help of D-wave Ocean Software. We will also try to compare the performance of our quantum annealing-based algorithm with classical lattice algorithms.

Acknowledgment

We would like to express our sincere gratitude to our scientific supervisor A. V. Kutsenko for his guidance and support through out the research process. His valuables remarks play a vital role in shaping the direction of our thesis. We are thankful to him for believing on us and motivating us to achieve the results in our research.

REFERENCES

1. *Bos J., Ducas L., Kiltz E., et al.* CRYSTALS-Kyber: a CCA-secure module-lattice-based KEM. IEEE Europ. Symp. EuroS&P, London, UK, 2018, pp. 353–367.
2. *Ducas L., Kiltz E., Lepoint T., et al.* Crystals-dilithium: A lattice-based digital signature scheme. IACR Trans. Cryptographic Hardware Embedded Systems, 2018, no. 1, pp. 238–268.
3. *Fouque P. A., Hoffstein J., Kirchner P., et al.* Falcon: Fast-Fourier lattice-based compact signatures over NTRU. Submission to the NIST post-quantum cryptography standardization process, 2018. `https://www.di.ens.fr/~prest/Publications/falcon.pdf`.
4. *Bernstein D. J., Hulsing A., Kolbl S., et al.* The SPHINCS+ signature framework. Proc. 2019 ACM SIGSAC Conf. CCS'19, London, UK, 2019, pp. 2129–2146.
5. *Peikert C.* A decade of lattice cryptography. Found. Trends Theor. Comput. Sci., 2016, vol. 10, no. 4, pp. 283–424.
6. *Lentra A., Lenstra H., and Lovasz L.* Factoring polynomials with rational coefficients. Math. Ann., 1982, vol. 261, pp. 515–534.
7. *Chen Y. and Nguyen P. Q.* BKZ 2.0: Better lattice security estimates. LNCS, 2011, vol. 7073, pp. 1–20.
8. *Ishiguro T., Kiyomoto S., Miyake Y., and Takagi T.* Parallel gauss sieve algorithm: Solving the SVP challenge over a 128-dimensional ideal lattice. LNCS, 2014, vol. 8383, pp. 411–428.
9. *Hanrot G. and Stehle D.* Improved analysis of Kannan's shortest lattice vector algorithm. LNCS, 2007, vol. 4622, pp. 170–186.
10. *Doulgerakis E., Laarhoven T., and de Weger B.* Finding closest lattice vectors using approximate voronoi cells. LNCS, 2019, vol. 11505, pp. 3–22.
11. *Babai L.* On Lovász' lattice reduction and the nearest lattice point problem. Combinatorica, 1986, no. 6, pp. 1–13.
12. *Date P., Patton R., Schuman C., and Patok P.* Efficiently embedding QUBO problems on adiabatic quantum computers. Quantum Inform. Processing, 2019, no. 18, pp. 1–31.
13. *Glover F., Kochenberger G., and Hennig R.* Quantum bridge analytics I: a tutorial on formulating and using QUBO models. Ann. Operations Res., 2022, vol. 314, no. 1, pp. 141–183.
14. *Micciancio D. and Goldwasser S.* Complexity of Lattice Problems. N.Y., Springer, 2002.
15. *Regev O.* On lattices, learning with errors, random linear codes, and cryptography. J. ACM, 2009, vol. 56, no. 6, pp. 1–40.

Секция 4

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ И ПРОГРАММИРОВАНИЯ

УДК 519.682 DOI 10.17223/2226308X/16/31

АНАЛОГ ТЕОРЕМЫ КРОНЕКЕРА — КАПЕЛЛИ ДЛЯ СИСТЕМ НЕКОММУТАТИВНЫХ ЛИНЕЙНЫХ УРАВНЕНИЙ, ПОРОЖДАЮЩИХ ЛИНЕЙНЫЕ ЯЗЫКИ

О. И. Егорушкин, И. В. Колбасина, К. В. Сафонов

Продолжается исследование систем некоммутативных полиномиальных уравнений, которые интерпретируются как грамматики формальных языков. Такие системы решаются в виде формальных степенных рядов (ФСР), выражающих нетерминальные символы через терминальные символы алфавита и рассматриваемых как формальные языки. Всякому ФСР поставлен в соответствие его коммутативный образ, который получается в предположении, что все символы обозначают коммутативные переменные, действительные или комплексные. Рассматриваются уравнения, которые линейны по нетерминальным символам с полиномиальными коэффициентами от терминальных символов, а значит, эти системы порождают линейные формальные языки. Совместность системы некоммутативных полиномиальных уравнений не связана напрямую с совместностью её коммутативного образа, и потому в качестве аналога теоремы Кронекера — Капелли удаётся получить лишь достаточное условие несовместности некоммутативной системы.

Ключевые слова: *системы линейных уравнений, некоммутативные переменные, формальный степенной ряд, коммутативный образ.*

Теория формальных языков имеет фундаментальное значение для программирования и разработки информационных технологий. Она является основой разработки алгоритмов для поиска в Интернете, работы новостных агрегаторов, машинного перевода текстов, понимания генетического кода в биоинформатике, разработки и анализа языков программирования. Все эти приложения используют взаимосвязи языка (как множества возможных текстов) с грамматикой (сводом формальных правил, определяющих языковые конструкции и их равнозначимость). Более того, всюду нужны быстрые качественные алгоритмы формальных построений грамматики по языку и языка по грамматике и синтаксического анализа конструкций, невозможные без серьёзного теоретического обоснования.

Контекстно-свободные, в частности линейные, грамматики активно используются для решения задач, связанных с разработкой формальных языков и синтаксических анализаторов [1–3]. Одним из основных достоинств контекстно-свободных и линейных грамматик является возможность задания широкого класса языков при сохранении относительной компактности представления [4, 5].

Продолжая исследование, начатое в работах [1, 2], рассмотрим систему линейных уравнений с полиномиальными коэффициентами

$$a_{i1}(x)z_1b_{i1}(x) + \ldots + a_{in}(x)z_nb_{in}(x) = d_i(x), \quad i = 1, \ldots, k, \tag{1}$$

которая решается относительно некоммутативных символов $z = (z_1, \ldots, z_n)$ в виде ФСР, зависящих от некоммутативных символов $x = (x_1, \ldots, x_m)$, где a_{ij}, b_{ij}, d_i — многочлены. Такие системы имеют приложения в теории формальных языков, поскольку являются грамматиками, порождающими важный класс линейных языков [3, 4].

В рамках теории формальных языков и грамматик символы $x_1, \ldots, x_m$ называются терминальными и образуют словарь (алфавит) данного языка, тогда как символы $z_1, \ldots, z_n$ называются нетерминальными и необходимы для задания грамматических правил. Над всеми символами определены некоммутативное умножение (конкатенация) и коммутативное формальное сложение, а также коммутативная операция умножения на числовые коэффициенты, и потому можно рассматривать символьные многочлены и ФСР с действительными или комплексными коэффициентами. Наконец, мономы от терминальных символов интерпретируются как предложения языка, а каждый ФСР, который является решением системы (1), рассматривается как порождённый грамматикой формальный язык, т. е. формальная сумма всех «правильных» предложений этого языка [3, 4].

Исследовать символьную некоммутативную систему (1) достаточно трудно, поэтому рассмотрим её коммутативный образ [1, 2, 5].

Предположим, следуя [1], что все мономы от $x_1, \ldots, x_m$ занумерованы в лексикографическом порядке по возрастанию степеней в последовательность $u_0, u_1, \ldots$, играющую роль базиса, тогда каждый ряд s можно единственным образом записать в виде разложения по этому базису с числовыми коэффициентами $\langle s, u_i \rangle$ при мономах u_i:

$$s = \sum_{i=0}^{\infty} \langle s, u_i \rangle u_i. \tag{2}$$

Теперь поставим в соответствие ФСР (2) его коммутативный образ $\mathrm{ci}(s)$ — степенной ряд, который получается из s в предположении, что символы $x_1, \ldots, x_m$ (равно как и $z_1, \ldots, z_n$) обозначают коммутативные переменные, принимающие значения из поля комплексных чисел [5].

Очевидно, что коммутативным образом системы (1) является система коммутативных уравнений

$$\mathrm{ci}(a_{i1}(x))\,\mathrm{ci}(b_{i1}(x))z_1 + \ldots + \mathrm{ci}(a_{in}(x))\,\mathrm{ci}(b_{in}(x))z_n = \mathrm{ci}(d_i(x)), \quad i = 1, \ldots, k. \tag{3}$$

Назовём рангом матрицы, элементы которой являются многочленами от коммутативных переменных, число линейно независимых столбцов матрицы над полем комплексных чисел.

Для системы коммутативных уравнений (3) рассмотрим её матрицу

$$(\mathrm{ci}(a_{ij}(x))\,\mathrm{ci}(b_{ij}(x)),$$

а также расширенную матрицу $(\mathrm{ci}(a_{ij}(x))\,\mathrm{ci}(b_{ij}(x)) \mid d_i(x))$.

Некоммутативным аналогом теоремы Кронекера — Капелли является следующая

Теорема 1. Если для коммутативного образа (3) системы некоммутативных линейных уравнений (1) выполнено неравенство

$$\mathrm{rank}(\mathrm{ci}(a_{ij}(x))\,\mathrm{ci}(b_{ij}(x))) < \mathrm{rank}(\mathrm{ci}(a_{ij}(x))\,\mathrm{ci}(b_{ij}(x)) \mid \mathrm{ci}(d_i(x))),$$

то исходная система некоммутативных линейных уравнений (1) несовместна.

Замечание 1. В случае равенства рангов рассматриваемых матриц утверждение о совместности исходной некоммутативной системы неверно; нетрудно привести соответствующий пример.

Поскольку ФСР, которые являются компонентами решения системы (1), интерпретируются как формальные языки, теорема 1 позволяет установить случаи, когда линейная грамматика не порождает никакого линейного языка.

ЛИТЕРАТУРА

1. *Егорушкин О. И., Колбасина И. В., Сафонов К. В.* О совместности систем символьных полиномиальных уравнений и их приложении // Прикладная дискретная математика. Приложение. 2016. № 9. С. 119–121.
2. *Egorushkin O. I., Kolbasina I. V., and Safonov K. V.* On solvability of systems of symbolic polynomial equations // Журн. СФУ. Сер. Матем. и физ. 2016. Т. 9. Вып. 2. С. 166–172.
3. *Глушков В. М., Цейтлин Г. Е., Ющенко Е. Л.* Алгебра. Языки. Программирование. Киев: Наукова думка, 1973.
4. *Salomaa A. and Soitolla M.* Automata-Theoretic Aspects of Formal Power Series. N.Y.: Springer Verlag, 1978.
5. *Семёнов А. Л.* Алгоритмические проблемы для степенных рядов и контекстно-свободных грамматик // Докл. АН СССР. 1973. № 212. С. 50–52.

УДК 004.05 DOI 10.17223/2226308X/16/32

ОБ АЛГОРИТМАХ ПОИСКА КОМПЬЮТЕРНОЙ ИНФОРМАЦИИ

А. В. Жаркова, А. Г. Мусугалиева

Исследованы алгоритмы поиска подстроки в строке: наивный, Бойера — Мура, Кнута — Морриса — Пратта, Рабина — Карпа, а также применимые для них джокеры (символы подстановки, «совпадающие» с любым символом или их последовательностью). Реализована программа на языке C# для поиска файлов по различным параметрам, которая позволяет также сканировать заданную директорию с целью поиска вредоносных объектов. Проведены вычислительные эксперименты. Общее лучшее время поиска файлов (достаточно найти первое вхождение) оказалось с помощью алгоритма Бойера — Мура, худшее — с помощью алгоритма Рабина — Карпа. Для поиска файлов по небольшим заданным данным и параметрам можно использовать наивный поиск, для средних и больших данных и параметров при малых образцах лучше использовать алгоритм Кнута — Морриса — Пратта, при больших — Бойера — Мура.

Ключевые слова: *алгоритм Бойера — Мура, алгоритм Кнута — Морриса — Пратта, алгоритм Рабина — Карпа, кибербезопасность, поиск подстроки в строке, поиск файла, сканирование.*

В связи с большим ростом объёма данных существует много проблем с обработкой информации. Устройство может содержать несколько сотен тысяч файлов, среди которых часто требуется найти необходимый. Сейчас функции поиска файлов встроены во многие системы, но если поиск информации является ключевой задачей, например при поиске вредоносных объектов антивирусными программами, требуется знать принципы его организации.

Целью работы является изучение различных алгоритмов поиска файлов с заданными параметрами и создание программного продукта для поиска информации на устройстве.

Пусть задан образец $P[1 \dots m]$ и текст $T[1 \dots n]$, $m \leqslant n$. В задаче поиска подстрок требуется найти все вхождения образца в текст.

Рассмотрены следующие алгоритмы поиска подстроки в строке:

1) Наивный алгоритм: левый конец образца P ставится вровень с левым концом текста и все символы P сравниваются с соответствующими символами слева направо пока либо не встретится несовпадение, либо не исчерпается P, что будет означать, что найдено вхождение P в текст. Простейший алгоритм поиска подстрок оказывается неэффективным, поскольку информация о тексте, полученная для одного найденного вхождения, не используется при нахождении других вхождений. Сложность алгоритма $O(mn)$ [1, 2].
2) Алгоритм Бойера — Мура: прикладываем образец к началу текста и сравниваем с конца, при несовпадении очередного символа сдвигаем образец на максимальное значение из элемента массива сдвига (содержит величины, на которые может быть сдвинут образец при несовпадении очередного символа) и элемента массива прыжков (содержит величины, на которые можно сдвинуть образец, чтобы совместить ранее совпавшие символы с вновь совпадающими символами строки). Значение элемента массива сдвига вычисляется следующим образом: $s[P[i]] = m - i, i = 1, \dots, m$. Элементы массива прыжков изменяются в зависимости от повторений последних символов, алгоритм работает за время $O(m+n)$ [3].
3) Алгоритм Кнута — Морриса — Пратта (КМП): используется префиксная функция $\pi[q]$, значение которой равно длине наибольшего префикса образца P, который является суффиксом строки P_q, где P_q — k-символьный префикс $P[1 \dots k]$ образца $P[1 \dots m]$. В основной функции алгоритма сканируется текст слева направо и проверяется условие $P[q+1] = T[i]$, где $i = 1, \dots, n$; q — количество совпавших символов; вхождение образца найдено, если $q = m$. Время работы алгоритма равно $O(m+n)$ [2].
4) Алгоритм Рабина — Карпа: каждый символ представляет собой цифру в системе счисления с основанием d, пусть $d = 10$. Для заданного образца $P[1 \dots m]$ обозначим через p соответствующее ему десятичное значение. Аналогично для заданного текста $T[1 \dots n]$ обозначим через t_s десятичное значение подстроки $T[s+1 \dots s+m]$ длины m, $s = 0, 1, \dots, n-m$. Таким образом, s — допустимый сдвиг (найдено вхождение) тогда и только тогда, когда $t_s = p$. Время работы в наихудшем случае равно $O((n-m+1)m)$ [2].

Данные алгоритмы реализованы для поиска файлов, в связи с этим достаточно найти только первое вхождение образца в текст.

Для поиска также можно ввести специальный символ, называемый джокером, который «совпадает» с любым символом. Если число разрешённых джокеров не ограничено, то неизвестно, можно ли решить задачу за линейное время. Однако если число джокеров ограничено фиксированной постоянной (не зависящей от размера P), то образец можно найти за линейное время [1]. В разработанной программе при необходимости для замены любой строки символов, в том числе пустой, в начале или в конце образца используется джокер «*».

Согласно государственному стандарту РФ ГОСТ Р 51188–98 [4], при испытаниях программных средств на наличие компьютерных вирусов используют две основные

группы методов их обнаружения: программные и аппаратно-программные. К программным методам относятся:

— сканирование;
— обнаружение изменений;
— эвристический анализ;
— резидентные «сторожа»;
— вакцинирование программных средств.

Метод сканирования состоит в том, что специальная антивирусная программа, называемая сканером, последовательно просматривает проверяемые файлы в поиске так называемых «сигнатур» известных компьютерных вирусов. При этом под сигнатурой понимают уникальную последовательность байтов, принадлежащую конкретному известному компьютерному вирусу и не встречающуюся в других программах [4].

Разработана и реализована программа на языке C# для поиска файлов по различным параметрам (имя файла, его расширение, размер и содержимое) с помощью перечисленных алгоритмов поиска образца в тексте; программа позволяет также сканировать заданную директорию на поиск вредоносных объектов.

Проведены вычислительные эксперименты. В табл. 1 приведено среднее время работы поиска файлов по имени и расширению с помощью различных алгоритмов поиска подстроки в строке; l — длина образца. Поиск производился в директории размером 151 Мбайт, в которой находилось 412 файлов.

Таблица 1

Среднее время поиска файлов по имени и расширению

Алгоритм	Время работы, мс				
	По имени файла		По расширению файла	По имени и расширению файла	
	$1 \leqslant l \leqslant 10$	$11 \leqslant l \leqslant 50$		$3 \leqslant l \leqslant 20$	$21 \leqslant l \leqslant 50$
Наивный	18,5	3,61	33,58	11,47	3,64
Бойера — Мура	19,11	5,22	34,75	12,28	4,92
КМП	19,47	4,42	34,83	12,58	4,47
Рабина — Карпа	21,03	5,69	34,67	14,17	5,7

Можно заметить, что при поиске файлов в заданной директории по имени и расширению лучшее время во всех случаях показал наивный алгоритм. Это можно объяснить тем, что в остальных алгоритмах требуется предварительная обработка образца, что занимает время при выполнении, а нас интересует только наличие данного образца в тексте. В большинстве случаев хуже всех отработал поиск с помощью алгоритма Рабина — Карпа, при поиске по расширению файла — алгоритм Кнута — Морриса — Пратта.

В табл. 2 приведено среднее время работы поиска файлов по содержимому с помощью различных алгоритмов поиска подстроки в строке; l — длина образца. Поиск производился в директории размером 65,7 Мбайт, в которой находилось 120 .txt-файлов.

Можно заметить, что при поиске файлов по содержимому для коротких образцов быстрее всех отработал поиск с использованием алгоритма Кнута — Морриса — Пратта, для длинных — алгоритм Бойера — Мура, дольше всех в обоих случаях находились файлы с помощью алгоритма Рабина — Карпа.

Общее лучшее время поиска файлов оказалось с помощью алгоритма Бойера — Мура, худшее — с помощью алгоритма Рабина — Карпа. При этом из представленных случаев чаще всего лучше работал наивный поиск, хуже — также алгоритм Рабина — Карпа.

Таблица 2

Среднее время поиска файлов по содержимому

Алгоритм	Время работы, мс	
	$1 \leqslant l \leqslant 10$	$11 \leqslant l \leqslant 100$
Наивный	1471,39	1395,8
Бойера — Мура	464,81	453,5
Кнута — Морриса — Пратта	422,2	679,97
Рабина — Карпа	2463,11	62336,02

Таким образом, для поиска файлов по небольшим заданным данным и параметрам (например, поиск по имени, по расширению) можно вполне использовать наивный поиск, для средних и больших данных и параметров (например, поиск по содержимому) при малых образцах лучше использовать алгоритм Кнута — Морриса — Пратта, при больших — Бойера — Мура.

В программе предусмотрено также сканирование заданной директории на вредоносные объекты. Параметры для поиска вредоносных объектов хранятся в базе данных, которая представляет собой текстовый файл, основная база для которого взята из открытого источника антивируса ClamAV [5] и переделана под формат программы, её можно дополнять. Основное время работы этой части программы — это работа с антивирусной базой, которая содержит 4 млн строк, поэтому среднее время работы составляет от нескольких десятков до нескольких сотен секунд.

ЛИТЕРАТУРА

1. *Гасфилд Д.* Строки, деревья и последовательности в алгоритмах: Информатика и вычислительная биология. СПб.: Невский диалект, БХВ-Петербург, 2003.
2. *Кормен Т., Лейзерсон Ч., Ривест М.* Алгоритмы: построение и анализ. М.: Изд-во «Вильямс», 2005.
3. *Макконнелл Дж.* Основы современных алгоритмов. М.: Техносфера, 2004.
4. ГОСТ Р 51188-98. Защита информации. Испытания программных средств на наличие компьютерных вирусов. Типовое руководство. М.: ИПК Издательство стандартов, 2003.
5. `http://www.clamav.net/downloads` — ClamAV Download. 2017.

УДК 519.688 DOI 10.17223/2226308X/16/33

ОБ ОДНОМ ПРЕДСТАВЛЕНИИ ЭЛЕМЕНТОВ КОНЕЧНЫХ 2-ГРУПП В ВИДЕ БУЛЕВЫХ ВЕКТОРОВ

А. А. Кузнецов, А. С. Кузнецова

Предложен способ представления конечных 2-групп в виде булевых векторов. Пусть G — конечная (бернсайдова) 2-группа, порядок которой равен 2^k. Каждый элемент группы представим уникальным булевым вектором размерности k. Для вычисления произведения двух элементов используются аналоги полиномов Холла, только теперь в них вместо умножения и сложения над полем $\mathbb{Z}_2$ используются эквивалентные булевы (побитовые) операции «и» и «исключающее или». В задачах, требующих вычисления большого количества произведений элементов группы, описанный метод позволяет кардинально уменьшить время работы компьютерных программ.

Ключевые слова: *2-группа, булев вектор, побитовые операции, полиномы Холла.*

В последние десятилетия наблюдается рост исследований, связанных с разработкой новых криптосистем и протоколов обмена ключами, основанных на различных некоммутативных алгебраических системах. В частности, в работах [1–3] в качестве криптографических примитивов предложено использовать бернсайдовы группы периода $n = 3$. Для $n > 3$ вопрос пока не рассматривался. В связи с этим становится актуальной задача разработки ресурсоэффективных алгоритмов для работы с бернсайдовыми группами бóльших периодов.

Пусть $G = \langle a_1, a_2, \ldots, a_m \rangle$ — некоторая конечная (бернсайдова) 2-группа периода 2^n, где $a_1, \ldots, a_m$ — порождающие элементы группы и $|G| = 2^k$. Для G можно получить pc-представление элементов группы (Power Commutator presentation [4]). В этом случае

$$\forall g \in G \, (g = a_1^{x_1} \ldots a_k^{x_k}), \; x_i \in \mathbb{Z}_2.$$

Пусть $a_1^{x_1} \ldots a_k^{x_k}$ и $a_1^{y_1} \ldots a_k^{y_k}$ — два произвольных элемента группы G, записанные в pc-формате, и пусть их произведение равно $a_1^{z_1} \ldots a_k^{z_k}$.

Вычисление степеней z_i традиционно осуществляется на основе собирательного процесса Холла [4]. Однако существует более эффективный способ умножения элементов, основанный на полиномах Холла [4]:

$$z_i = x_i + y_i + p_i(x_1, \ldots, x_{i-1}, y_1, \ldots, y_{i-1}), \; x_i, y_i, z_i \in \mathbb{Z}_2.$$

Заметим, что операции умножения и сложения в поле $\mathbb{Z}_2$ тождественны булевым операциям «и» и «исключающее или» соответственно. Это позволяет естественным образом представлять элементы группы в виде булевых векторов размерности k.

В качестве примера рассмотрим максимальную двупорождённую конечную группу $G = \langle a_1, a_2 \rangle$ периода $2^2 = 4$, которую обычно обозначают $B(2,4)$ или $B_2(4)$. Порядок данной группы равен 2^{12}, и для каждого элемента из G существует уникальное pc-представление вида $a_1^{x_1} \ldots \; a_{12}^{x_{12}}$, где $x_i \in \mathbb{Z}_2$, $i = 1, 2, \ldots, 12$. Здесь a_1 и a_2 — порождающие элементы G, $a_3 \ldots, a_{12}$ вычисляются рекурсивно через a_1 и a_2.

Получим в системе компьютерной алгебры GAP pc-представление данной группы. Для краткости тривиальные коммутаторные соотношения не приводятся (например, $[a_4, a_1] = 1$ и др.):

$$a_1^2 = a_4, \; a_2^2 = a_5, \; a_3^2 = a_8 a_9 a_{10} a_{11} a_{12}, \; a_4^2 = 1, \; a_5^2 = 1, \; a_6^2 = a_{11}, \; a_7^2 = a_{11} a_{12},$$
$$a_i^2 = 1 \; (8 \leqslant i \leqslant 12), \; [a_3, a_1] = a_6, \; [a_3, a_2] = a_7, \; [a_4, a_2] = a_6 a_8 a_9 a_{10} a_{12}, \; [a_4, a_3] = a_8 a_{11},$$
$$[a_5, a_1] = a_7 a_8 a_9 a_{10}, \; [a_5, a_3] = a_{10} a_{11} a_{12}, \; [a_5, a_4] = a_9 a_{11}, [a_6, a_1] = a_8, \; [a_6, a_2] = a_9,$$
$$[a_6, a_3] = a_{11}, \; [a_6, a_4] = a_{11}, \; [a_6, a_5] = a_{11}, \; [a_7, a_1] = a_9 a_{12}, \; [a_7, a_2] = a_{10}, \; [a_7, a_3] = a_{11} a_{12},$$
$$[a_7, a_4] = a_{11} a_{12}, \; [a_7, a_5] = a_{11} a_{12}, \; [a_8, a_1] = a_{11}, \; [a_8, a_2] = a_{12}, \; [a_9, a_1] = a_{11} a_{12},$$
$$[a_9, a_2] = a_{11}, \; [a_{10}, a_1] = a_{12}, \; [a_{10}, a_2] = a_{11} a_{12}.$$

Вычислим полиномы Холла группы G для порождающих элементов a_1 и a_2 на основе алгоритма из [5]:

1) $a_1 \cdot a_1^{y_1} \ldots a_{12}^{y_{12}} = a_1^{z_1} \ldots a_{12}^{z_{12}}$,

$$\text{где} \;\; z_1 = y_1 + 1, \; z_2 = y_2, \; z_3 = y_3, \; z_4 = y_1 + y_4, \; z_5 = y_5, \; z_6 = y_6 + y_1 y_2, \; z_7 = y_7,$$
$$z_8 = y_8 + y_1 y_2 + y_1 y_3, \; z_9 = y_9 + y_1 y_2, \; z_{10} = y_{10} + y_1 y_2,$$
$$z_{11} = y_{11} + y_1 y_3 + y_1 y_2 y_3 + y_1 y_2 y_4 + y_1 y_2 y_5 + y_1 y_2 y_6, \; z_{12} = y_{12} + y_1 y_2;$$

2) $a_2 \cdot a_1^{y_1} \dots a_{12}^{y_{12}} = a_1^{z_1} \dots a_{12}^{z_{12}}$,

где $z_1 = y_1,\ z_2 = y_2 + 1,\ z_3 = y_1 + y_3,\ z_4 = y_4,\ z_5 = y_2 + y_5,\ z_6 = y_6,$
$z_7 = y_7 + y_1y_2,\ z_8 = y_8 + y_1y_3,\ z_9 = y_9 + y_1y_3 + y_2y_4,\ z_{10} = y_{10} + y_1y_2 + y_1y_3 + y_2y_3,$
$z_{11} = y_{11} + y_1y_2 + y_1y_3 + y_2y_3 + y_2y_4 + y_1y_2y_3 + y_1y_2y_4 + y_1y_2y_5 + y_1y_2y_7,$
$z_{12} = y_{12} + y_1y_2 + y_1y_3 + y_2y_3 + y_1y_2y_3 + y_1y_2y_4 + y_1y_2y_5 + y_1y_2y_7.$

Каждый элемент группы представляет собой битовую строку $(z_1, z_2, \dots, z_{12})$. Таким образом, для кодирования одного элемента в $B(2,4)$ требуется 12 бит. В общем случае, если порядок группы равен 2^k, то для хранения одного элемента потребуется k бит.

На компьютере операции над битами выполняются значительно быстрее, чем над целочисленными или строковыми типами данных. В задачах, требующих вычисления большого количества произведений элементов группы, описанный метод позволит кардинально уменьшить время работы компьютерных программ. Одной из таких проблем является задача поиска кратчайших маршрутов на графах Кэли, которые часто применяются при проектировании топологий для сетей межпроцессорного соединения в суперкомпьютерах, а также дата-центрах.

Кроме того, предложенное представление элементов группы в форме булевых векторов позволяет применять их даже на самых примитивных микроконтроллерах.

ЛИТЕРАТУРА

1. *Baumslag G., Fazio N., Nicolosi A. R., et al.* Generalized learning problems and applications to non-commutative cryptography // LNCS. 2011. V. 6980. P. 324–339.
2. *Fazio N., Iga K., Nicolosi A. R., et al.* Hardness of learning problems over Burnside groups of exponent 3 // Des. Codes Cryptogr. 2015. V. 75(1). P. 59–70.
3. *Kahrobaei D. and Noce M.* Algorithmic problems in Engel groups and cryptographic applications // Int. J. Group Theory. 2020. V. 9(4). P. 231–250.
4. *Sims C.* Computation with Finitely Presented Groups. Cambridge: Cambridge University Press, 1994. 628 p.
5. *Кузнецов А. А. Кузнецова А. С.* Быстрое умножение элементов в конечных двупорожденных группах периода пять // Прикладная дискретная математика. 2013. № 1(19). С. 110–116.

Секция 5

ПРИКЛАДНАЯ ТЕОРИЯ КОДИРОВАНИЯ И АВТОМАТОВ

УДК 003.26 DOI 10.17223/2226308X/16/34

ОБ ОДНОМ КЛАССЕ АЛГЕБРОГЕОМЕТРИЧЕСКИХ КОДОВ

М. М Глухов, К. Н Панков

В целях использования в кодовых системах открытого шифрования приводится семейство $[720, 3r-57, 720-3r]$-алгеброгеометрических кодов над полем GF(256), лежащих выше границы Варшамова — Гилберта вместе с двойственными к ним кодами при $81 \leqslant r \leqslant 197$.

Ключевые слова: *постквантовая криптография, алгеброгеометрические коды, система МакЭлиса, граница Варшамова — Гилберта.*

Первоначально идея использования линейных кодов для построения схем с открытым ключом была предложена в 1978 г. Р. МакЭлисом [1]. Открытым ключом в системе МакЭлиса служит замаскированная определённым образом порождающая матрица некоторого кода, а закрытым — маскирующие матрицы. Стойкость системы основана на сложности задачи декодирования «кода общего положения». В некотором смысле двойственная система открытого шифрования предложена в 1986 г. Г. Нидеррайтером [2]. В системе Нидеррайтера вместо порождающей матрицы кода используется проверочная, шифруемое сообщения играет роль вектора ошибок, а зашифрованный текст является синдромом вектора ошибок.

В упомянутых системах предлагалось использовать двоичные коды Гоппы, а попытки использовать некоторые классы кодов Рида — Соломона в системе Нидеррайтера и кодов Рида — Маллера в системе МакЭлиса [3] привели к снижению стойкости [4, 5].

Одним из участников конкурса NIST Post-Quantum Cryptography Standardization является проект Classic McEliece, в основе которого лежит схема Нидеррайтера на двоичных кодах Гоппы. Открытым ключом в такой схеме является основная часть систематической проверочной матрицы кода. При реализации с использованием данной схемы механизма инкапсуляции ключа авторами проекта предлагаются следующие величины параметров (n, t, m) (при этом $k = n - tm$, $d = 2t + 1$, а используемый $[n, k, d]_2$-код Гоппы задаётся унитарным неприводимым над $\mathrm{GF}(2^m)$ многочленом степени t): $(3488, 64, 12)$, $(4608, 96, 13)$, $(6688, 128, 13)$, $(6960, 119, 13)$, $(8192, 128, 13)$. Выбор классических кодов Гоппы обоснован, в частности, наличием эффективной процедуры декодирования.

Как видно из приведённых параметров, для построения подобных схем необходимы достаточно длинные коды. Такие коды можно найти в классах алгеброгеометрических кодов на кривых. Большое разнообразие кривых и возможность выбора точек дают дополнительную вариативность для схем типа Classic McEliece при использовании кодов на кривой.

В [6] приведены классы многочленов, на которых достигаются верхние оценки некоторых тригонометрических сумм над полями нечётной характеристики. В работе [7]

результат С. А. Степанова обобщён на случай произвольного конечного непростого поля. Эти результаты позволили построить большие классы длинных кодов на кривых, лежащих вблизи границы алгеброгеометрических кодов.

В частности, в [8] приведён класс алгеброгеометрических $[768, 3r-57, d]_{2^8}$-кодов при $114 < 3r < 768$ и $d = 768 - 3r$ на кривой

$$E : y^3 = f(x) = (x^{63} - 1)/(x^3 - 1).$$

Некоторые свойства этих кодов описаны в [9]. Параметры указанных кодов обусловлены тем, что данная кривая имеет в точности 768 различных $\mathrm{GF}(2^8)$-рациональных точек $P_1, \ldots, P_{768}$, а род кривой $g = 58$. Коды заданы двумя дивизорами: $D = P_1 + \ldots + P_{768}$ и $G = rP_\infty$, где P_∞ — бесконечно удалённая точка кривой. Кодовыми словами такого кода являются векторы значений функций $x^i y^j$ при $j = 0, 1, 2$, $i = 0, 1, \ldots, r - 20j$ в точках $P_1, \ldots, P_{768}$.

Если в качестве дивизора D взять сумму n произвольных различных точек кривой $D = P_{i_1} + \ldots + P_{i_n}$ при условии $2g-2 < 3r < n < 768$, то аналогичным образом получим $[n, 3r - 57, n - 3r]_{2^8}$-код.

Выберем достаточно большое число n, удовлетворяющее этим условиям и такое, что $\mathrm{C}_{768}^n > 2^{254}$. Несложно вычислить, что неравенство справедливо при $48 \leqslant n \leqslant 720$. Рассмотрим случай $n = 720$. Зафиксируем произвольные 720 точек на кривой E: $P_1, \ldots, P_{720}$ с координатами из $\mathrm{GF}(2^8)$. Векторы значений указанных функций в этих точках задают базис $[720, 3r - 57, 720 - 3r]_{2^8}$-кода при соответствующих значениях r. Двойственным к такому коду является $[720, 777 - 3r, 3r - 114]_{2^8}$-код.

Построенные коды не являются, очевидно, МДР-кодами. Но комбинаторными вычислениями можно показать, что при большинстве значений r они лежат выше границы из теоремы Варшамова — Гилберта, утверждающей, что существуют линейные $[n, k, d]_q$-коды, мощность которых удовлетворяет неравенству

$$q^k \geqslant q^n / \Big(\sum_{i=0}^{d-2} C_{n-1}^i (q-1)^i \Big),$$

а именно: справедлива следующая

Теорема 1. Пусть $E : y^3 = (x^{63} - 1)/(x^3 - 1)$ — кривая над полем $F = \mathrm{GF}(2^8)$, $P_1, \ldots, P_{720}$ — произвольные различные F-рациональные точки этой кривой, P_∞ — бесконечно удалённая точка кривой. Тогда алгеброгеометрические коды $C_r(D, G)$ на кривой E, определяемые дивизорами $D = P_1 + \ldots + P_{720}$ и $G = rP_\infty$ при всех натуральных r, $81 \leqslant r \leqslant 197$, являются $[720, 3r - 57, 720 - 3r]_{2^8}$-кодами, а их мощность, как и мощность двойственных к ним $[720, 777 - 3r, 3r - 114]_{2^8}$-кодов, удовлетворяет неравенству из теоремы Варшамова — Гилберта.

ЛИТЕРАТУРА

1. *McEllice R. J.* Public-key cryptosystem based on algebraic coding theory // DSN Progress Report. 1978. P. 42–44.
2. *Niederreiter H.* Knapsack-type cryptosystems and algebraic coding theory // Problems Control Inform. Theory. 1986. V. 15. No. 2. P. 159–166.
3. *Сидельников В. М.* Открытое шифрование на основе двоичных кодов Рида — Маллера // Дискретная математика. 1994. Т. 5. Вып. 2. С. 3–20.
4. *Сидельников В. М., Шестаков С. О.* О системе шифрования, построенной на основе обобщенных кодов Рида —- Соломона // Дискретная математика. 1992. Т. 4. Вып. 3. С. 57–63.

5. *Minder L. and Shokrollahi A.* Cryptanalysis of the Sidelnikov cryptosystem // LNCS. 2007. V. 4515. P. 347–360.
6. *Степанов С. А.* О нижних оценках сумм характеров над конечными полями // Дискретная математика. 1991. Т. 3. Вып. 2. С. 77–86.
7. *Глухов М. М.* Нижние оценки сумм характеров от многочленов над конечными полями // Дискретная математика. 1994. Т. 6. Вып. 3. С. 136–142.
8. *Ozbudak F. and Glukhov M.* Codes on superelliptic curves // Turkish J. Math. 1998. V. 2. No. 2. P. 223–234.
9. *Pankov K. N. and Glukhov M. M.* Estimation of the power of algebraic geometric codes designed to construct a post-quantum algorithm for ensuring information security of on-board systems // 2023 Systems of Signals Generating and Processing in the Field of On-board Communications. Moscow, 2023. P. 1–5.

УДК 621.391:519.725 DOI 10.17223/2226308X/16/35

СЕРИЯ КОРОТКИХ ТОЧНЫХ ФОРМУЛ ДЛЯ ПАРАМЕТРА БХАТТАЧАРЬИ КООРДИНАТНЫХ КАНАЛОВ[1]

С. Г. Колесников, В. М. Леонтьев

Пусть W — симметричный канал с двоичным входом и конечным выходным алфавитом. В 2007 г. Э. Ариканом обнаружено явление поляризации каналов, которое позволяет выделить из множества координатных каналов $W_N^{(i)}$, построенных по W, те, по которым предпочтительнее передавать информационные биты. Один из инструментов, позволяющих произвести разделение каналов на «плохие» и «хорошие», — это параметр Бхаттачарьи $Z\big(W_N^{(i)}\big)$. Однако его вычисление затруднено из-за большого числа требуемых операций сложения — порядка 2^{2N}, где N — длина кода. В работе И. Тала и А. Варди 2013 г. предложен метод оценки сверху и снизу вероятностей ошибок в каналах $W_N^{(i)}$, $1 \leqslant i \leqslant N$, имеющий сложность порядка $O(N\mu^2 \log \mu)$, где $\mu > \mu_0$, а число μ_0 не зависит от длины N. Однако число μ может быть достаточно большим и зависит, в частности, от требуемой точности. Ранее авторами в случае, когда W — двоичный симметричный канал без памяти, построены две серии точных формул для параметров Бхаттачарьи, требующих всё ещё экспоненциального, но много меньшего числа операций, чем в формулах из оригинальной статьи Э. Арикана. В настоящей работе для всякого $N = 2^n$ удалось построить серию из $n(n-1)/2$ точных формул, которые не содержат суммирования по переменным.

Ключевые слова: *полярный код, двоичный симметричный канал без памяти, параметр Бхаттачарьи.*

Пусть W — канал с двоичным входным алфавитом $X = \{0, 1\}$, конечным выходным алфавитом $Y = \{y_1, \ldots, y_s\}$ и переходными вероятностями $W(y_j \,|\, x_i)$, $1 \leqslant j \leqslant s$, $1 \leqslant i \leqslant 2$. Выражение

$$Z(W) = \sum_{y \in Y} \sqrt{W(y \mid 0)\, W(y \mid 1)}$$

называется параметром Бхаттачарьи канала W. Число $Z(W)$ даёт верхнюю границу вероятности принятия ошибочного решения по методу максимального правдоподобия. В теории полярных кодов [1, 2] предполагается, что i-й бит каждого сообщения длины

[1]Работа поддержана Красноярским математическим центром, финансируемым Минобрнауки РФ (Соглашение № 075-02-2023-936).

$N = 2^n$, где $n \in \mathbb{N}$, передаётся по координатному каналу $W_N^{(i)} : X \to Y^N \times X^{i-1}$, $1 \leqslant i \leqslant N$, с переходными вероятностями

$$W_N^{(i)}(y, u' \mid u_i) = \frac{1}{2^{N-1}} \sum_{u'' \in X^{N-i}} W^N(y \mid uG_N),$$

где $u = u'(u_i)u''$ — конкатенация векторов u', (u_i) и u''; G_N — поляризационная матрица с ядром Арикана; W^N — декартова степень канала W. Согласно определению,

$$Z\left(W_N^{(i)}\right) = \sum_{y \in Y} \sum_{u' \in X^{i-1}} \sqrt{W_N^{(i)}(y, u' \mid 0)\, W_N^{(i)}(y, u' \mid 1)}. \tag{1}$$

При построении полярного кода мы вычисляем параметры $Z\left(W_N^{(i)}\right)$ или находим их верхние и нижние оценки для того, чтобы принять решение о возможности использования канала $W_N^{(i)}$ для передачи по нему информационных бит. Однако здесь возникают трудности, поскольку правая часть (1) требует порядка 2^{2N} операций сложения.

В [3] для симметричного канала W предложены методы оценки снизу и сверху вероятности ошибки в каналах $W_N^{(i)}$ с вычислительной сложностью порядка $O(N\mu^2 \log \mu)$, где $\mu > \mu_0$, а константа μ_0 зависит от пропускной способности канала, скорости и вероятности блоковой ошибки кода, но не зависит от длины кода N. От величины μ зависит также точность вычислений. Но, например, при $p = 0{,}1$ имеем $Z\left(W_{2^{10}}^{(2^{10})}\right) \simeq 6{,}7 \cdot 10^{-228}$ и, чтобы достигнуть такой точности, потребуется достаточно большое μ.

С другой стороны, когда W — двоичный симметричный канал без памяти, в [4] установлено равенство

$$Z\left(W_N^{(i)}\right) = \sum_{y=(y_1,\ldots,y_{i-1})} 2\sqrt{Q_i(y,0)\, Q_i(y,1)},$$

где

$$Q_i(y, \varepsilon) = \sum_{(\varepsilon, u_{i+1}, \ldots, u_N)} p^{w(\tilde{y}G_N \oplus \tilde{u}G_N)} (1-p)^{N - w(\tilde{y}G_N \oplus \tilde{u}G_N)}, \quad \varepsilon = 0, 1.$$

Здесь $w(b)$ — вес Хемминга вектора b, а $\tilde{y}$ и $\tilde{u}$ — N-мерные векторы, получающиеся из векторов y и u дополнением нулями справа и слева соответственно. Векторы $\tilde{y}G_N$ и $\tilde{u}G_N$ порождают подпространство в пространстве V всех строк длины N над полем $\mathbb{Z}_2$. Удачные характеризации этих подпространств позволили авторам в [4] получить две серии точных формул для параметров $Z\left(W_N^{(i)}\right)$ при $i = N - 2^k + 1$, $0 \leqslant k \leqslant n$, и $i = N/2 - 2^k + 1$, $1 \leqslant k \leqslant n - 2$. Формулы первой серии требуют порядка $\binom{2^{n-k} + 2^k - 1}{2^k} 2^{2^k}$ операций сложения, а для второй — $\binom{2^{n-k-1} + 2^k - 1}{2^k} 2^{2^k}$.

Исследуя свойства поляризационной матрицы G_N, удалось описать подпространство U, порождённое последними $N - 2^m$ строками матрицы G_N, и два его подмножества: подпространство U_0, порождённое последними $N - 2^m - 1$ строками матрицы G_N, и подмножество U_1, состоящее из всех линейных комбинаций 2^m-й строки матрицы G_N с векторами из U_0. Оказалось, что

$$U = \left\{(u_1, \ldots, u_N) \in V : u_{(i-1)2^{n-m}+1} \oplus \ldots \oplus u_{i2^{n-m}} = 0,\ i = 1, \ldots, 2^m\right\},$$

$$U_c = \left\{u = (u_1, \ldots, u_N) \in U : \bigoplus_{i=1}^{2^m} \bigoplus_{j=1}^{2^{n-m-1}} u_{(i-1)2^m + j} = c\right\}, \qquad c = 0, 1.$$

Это позволило для любых целых неотрицательных чисел $t_1, \ldots, t_{2^m}$ определить в U_c количества векторов с t_i единицами в i-м блоке координат $i = 1, 2, \ldots, 2^m$. Оно выражается суммой

$$\frac{1}{2}\left[\prod_{k=1}^{2^m}\binom{2^{n-m}-1}{2t_k} + (-1)^{c+t_1+\cdots+t_{2^m}}\prod_{k=1}^{2^m}\binom{2^{n-m-1}-1}{t_k}\right].$$

Как следствие, удалось доказать следующую теорему:

Теорема 1. Пусть $m, n \in \mathbb{N}$, $m < n$ и $p \in [0, 1]$. Справедливо равенство

$$Z\left(W_{2^n}^{(2^m+1)}\right) =$$
$$= (1-p)^{-2^m}((A+B)^{2^m} - B^{2^m}) + \sqrt{\left(\frac{1+(1-2p)^{2^{n-m}}}{2}\right)^{2^{m+1}} - ((1-p)^2 - p^2)^{2^n}},$$

где

$$A = \frac{p(1-p)}{2}\left(1 + (1-2p)^{2^{n-m}-1}\right) + \frac{(1-p)^2}{2}\left(1 - (1-2p)^{2^{n-m}-1}\right),$$
$$B = \frac{1-p}{2}\left(1 + (1-2p)^{2^{n-m}-1}\right).$$

Ввиду теоремы 1 и равенства $Z\left(W_{2N}^{(2i)}\right) = \left[Z\left(W_N^{(i)}\right)\right]^2$ (см., например, [1]), можно говорить о серии из $n(n-1)/2$ точных формул. Численные эксперименты показывают, что значения

$$Z\left(W_{2^{n+j}}^{(2^j(2^m+1))}\right),$$

где $j, m, n \in \mathbb{N}$ и $m < n$, близки к единице.

ЛИТЕРАТУРА

1. *Arikan E.* Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels // IEEE Trans. Inform. Theory. 2009. No. 7. P. 3051–3073.
2. *Tal I. and Vardy A.* How to construct polar codes // IEEE Trans. Inform. Theory. 2013. No. 10. P. 6542–6582.
3. *Трифонов П. В.* Основы помехоустойчивого кодирования. СПб.: Университет ИТМО, 2022.
4. *Колесников С. Г., Леонтьев В. М.* Серии формул для параметров Бхаттачарьи в теории полярных кодов // Проблемы передачи информации. 2023. Т. 59. Вып. 1 (в печати).

УДК 519.17 DOI 10.17223/2226308X/16/36

ВЫЧИСЛЕНИЕ ПАР, ИСПРАВЛЯЮЩИХ ОШИБКИ, ДЛЯ АЛГЕБРОГЕОМЕТРИЧЕСКОГО КОДА[1]

Е. С. Малыгина, А. А. Кунинец

Для произвольного алгеброгеометрического кода и дуального к нему явно вычислены пары, исправляющие ошибки. Такая пара состоит из кодов, которые необходимы для эффективного алгоритма декодирования заданного кода. Вид пар

[1]Работа поддержана грантом РНФ, проект № 22-41-0441.

зависит от степеней дивизоров, с помощью которых строится как исходный код, так и один из кодов, входящих в пару. Кроме того, вычислены пары, исправляющие ошибки, для подполевых подкодов исходного алгеброгеометрического кода и дуального к нему.

Ключевые слова: *функциональное поле, алгеброгеометрический код, исправляющая ошибки пара, подполевой подкод.*

Введение

Исследование задачи декодирования кодов, построенных на алгебраических кривых, явилось очень востребованным за последние тридцать лет. Изначально Т. Хёхольдт и др. предложили синдромный алгоритм декодирования для кодов, ассоциированных с плоской кривой [1]. Затем А. Скоробогатов и С. Влэдуц обобщили этот алгоритм на произвольные кривые [2]. Далее Р. Пелликаан и Р. Кёттер независимо друг от друга предложили алгоритм декодирования, исключающий абстрактные понятия алгебраической геометрии и использующий пары, исправляющие ошибки [3, 4]. Парой, исправляющей ошибки, для кода $\mathcal{C}$ является пара кодов $\mathcal{A}$ и $\mathcal{B}$, удовлетворяющая некоторым ограничениям на размерность и минимальное расстояние и условию, что покомпонентное произведение кодовых слов $\mathcal{A}$ и $\mathcal{B}$ содержится в $\mathcal{C}^{\perp}$. Существование такой пары обеспечивает эффективный алгоритм декодирования для алгеброгеометрических кодов (АГ-кодов), который использует только методы линейной алгебры. Особый интерес представляет построение таких пар, поскольку сама пара является входным параметром для алгоритма декодирования.

Стоит также отметить, что пары, исправляющие ошибки, заслуживают внимания и с криптографической точки зрения, поскольку лежат в основе атаки на АГ-коды [5].

1. Предварительные сведения

1.1. Алгеброгеометрические коды

Введём ряд обозначений:

- $\mathbb{F}_q$ — конечное поле, состоящее из q элементов, где q — простое или степень простого числа;
- $F/\mathbb{F}_q$ — алгебраическое функциональное поле рода g;
- $P_1, P_2, \ldots, P_n$ — попарно различные точки поля $F/\mathbb{F}_q$ степени один;
- $D = P_1 + \ldots + P_n$ — дивизор функционального поля F;
- G — дивизор функционального поля F, такой, что $\operatorname{supp}(G) \cap \operatorname{supp}(D) = \varnothing$, где $\operatorname{supp}(*)$ — носитель соответствующего дивизора.

Определение 1. Алгеброгеометрическим кодом $\mathcal{C}_{\mathscr{L}}(D, G)$, ассоциированным с дивизорами D и G, называется подпространство в $\mathbb{F}_q^n$ вида

$$\mathcal{C}_{\mathscr{L}}(D, G) = \{(f(P_1), \ldots, f(P_n)) : f \in \mathscr{L}(G)\} \subseteq \mathbb{F}_q^n,$$

где $\mathscr{L}(G)$ — пространство Римана — Роха.

Более детально с основными понятиями теории функциональных полей можно ознакомиться в [6].

Отметим, что всякий код $\mathcal{C}_{\mathscr{L}}(D, G)$ можно задать параметрами $[n, k, d]$, где n — длина кода (число точек в записи дивизора D); $k = k(\mathcal{C})$ — размерность кода (размерность пространства Римана — Роха $\mathscr{L}(G)$ или $\dim G$); $d = d(\mathcal{C})$ — минимальное расстояние кода.

Согласно [6, Theorem 2.2.2], код $\mathcal{C}_{\mathscr{L}}(D,G)$ является $[n,k,d]$-кодом, причём

$$k = \deg G + 1 - g,\ d \geqslant n - \deg G,$$

если $2g-2 < \deg G < n$.

Дуальный код к коду $\mathcal{C}_{\mathscr{L}}(D,G)$ будем обозначать так:

$$\mathcal{C}_{\mathscr{L}}(D,G)^{\perp} = \{x \in \mathbb{F}_q^n : \forall c \in \mathcal{C}_{\mathscr{L}}(D,G)\,(\langle x,c\rangle = 0)\}.$$

Здесь $\langle x,c\rangle = \sum\limits_{i=1}^{n} x_1c_1 + \ldots, x_nc_n$. Тогда, согласно [6, Theorem 2.2.7], код $\mathcal{C}_{\mathscr{L}}(D,G)^{\perp}$ является $[n,k',d']$-кодом, причём

$$k' = n + g - 1 - \deg G,\ d' \geqslant \deg G - (2g-2),$$

если $2g-2 < \deg G < n$.

В общем случае рассматриваемые коды могу исправить до $\lfloor d(d')-1/2\rfloor$ ошибок, где d и d' — минимальные расстояния кодов $\mathcal{C}_{\mathscr{L}}(D,G)$ и $\mathcal{C}_{\mathscr{L}}(D,G)^{\perp}$ соответственно.

Будем называть код MDS-кодом, если его минимальное расстояние достигает границы Синглтона, т. е. $d(\mathcal{C}) = n + 1 - k(\mathcal{C})$.

1.2. Пары, исправляющие ошибки

Произведение Шура двух векторов $a,b \in \mathbb{F}_q^n$ определяется как произведение их соответствующих координат:

$$(a_1,\ldots,a_n) * (b_1,\ldots,b_n) = (a_1b_1,\ldots,a_nb_n),$$
$$(a_1,\ldots,a_n)^i = (a_1^i,\ldots,a_n^i).$$

Для кодов $\mathcal{A},\mathcal{B} \subset \mathbb{F}_q^n$ произведение Шура $\mathcal{A}*\mathcal{B}$ определяется следующим образом:

$$\mathcal{A}*\mathcal{B} = \mathrm{Span}_{\mathbb{F}_q}\{a*b : a \in \mathcal{A}, b \in \mathcal{B}\}.$$

Определение 2. Пусть $\mathcal{C} \subseteq \mathbb{F}_q^n$ — линейный код. Тогда пара линейных кодов $(\mathcal{A},\mathcal{B})$, где $\mathcal{A},\mathcal{B} \subset \mathbb{F}_q^n$, называется парой, исправляющей t ошибок, для кода $\mathcal{C}$, если выполняются следующие условия:

1) $\mathcal{A}*\mathcal{B} \subseteq \mathcal{C}^{\perp}$;
2) $\dim(\mathcal{A}) > t$;
3) $d(\mathcal{B}^{\perp}) > t$;
4) $d(\mathcal{A}) + d(\mathcal{C}) > n$.

В обозначениях определения считаем, что $d(\mathcal{C}) \geqslant 2t+1$.

В [7, 8] описаны условия существования пары $\mathcal{A}$ и $\mathcal{B}$, исправляющей t ошибок.

2. Основной результат

Несмотря на наличие ряда работ, посвящённых вопросу существования пар, исправляющих ошибки, для линейных кодов, ни в одной из них не представлено нахождение такой пары для произвольного АГ-кода. Исключением является теорема 14 [5], посвящённая криптоанализу криптосистемы Мак-Элиса, в которой рассмотрен общий вид пары, исправляющей ошибки, для дуального кода. В следующих теоремах мы не только описываем вид кодов в паре, исправляющей ошибки, для $\mathcal{C}_{\mathscr{L}}(D,G)$ и $\mathcal{C}_{\mathscr{L}}(D,G)^{\perp}$, но также задаём классификацию относительно рода функционального поля и степеней дивизоров, ассоциированных с кодами из пары. Отметим, что теорему 14 из [5] мы специализируем на случай принадлежности одного из кодов пары, исправляющей ошибки, к MDS-кодам.

Теорема 1. Пусть $\mathcal{C} = \mathcal{C}_{\mathscr{L}}(D, G)$ — АГ-код, ассоциированный с функциональным полем $F/\mathbb{F}_q$ рода g. Тогда парой, исправляющей $t = (n - \deg(G) - g - 1)/2$ ошибок, для кода $\mathcal{C}$ является:

— $\mathcal{A} = \mathcal{C}_{\mathscr{L}}(D, F)$ и $\mathcal{B} = \mathcal{C}_{\mathscr{L}}(D, G + F)^{\perp}$, причём
  1) $\mathcal{A}$ — MDS-код, если $\deg(F) = t + g$ и g — произвольный;
  2) $\mathcal{B}$ — MDS-код, если $\deg(F) = n - \deg(G) - t - 1$ и $g = 0$;
  3) $\mathcal{C}$ — MDS-кодом, если $\deg(G) = n - 2t - 1$, $\deg(F) = t$ и $g = 0$;

— $\mathcal{A} = \mathcal{C}_{\mathscr{L}}(D, F)^{\perp}$ и $\mathcal{B} = \mathcal{C}_{\mathscr{L}}(D, F - G)$, причём
  1) $\mathcal{A}$ — MDS-код, если $\deg(F) = n + g - t - 2$ и g — произвольный;
  2) $\mathcal{B}$ — MDS-код, если $\deg(G) < (n + 3)/3$, $\deg(F) = \deg(G) + t - 1$ и $g = 0$.

Теорема 2. Пусть $\mathcal{C} = \mathcal{C}_{\mathscr{L}}(D, G)$ — АГ-код, ассоциированный с функциональным полем $F/\mathbb{F}_q$ рода g, и $\mathcal{C}^{\perp} = \mathcal{C}_{\mathscr{L}}(D, G)^{\perp}$ — дуальный к $\mathcal{C}$. Тогда парой, исправляющей $t = (\deg(G) - 3g + 1)/2$ ошибок, для кода $\mathcal{C}^{\perp}$ является:

— $\mathcal{A} = \mathcal{C}_{\mathscr{L}}(D, F)$ и $\mathcal{B} = \mathcal{C}_{\mathscr{L}}(D, G - F)$, причём
  1) $\mathcal{A}$ — MDS-код, если $\deg(F) = t + g$ и g — произвольный;
  2) $\mathcal{B}$ — MDS-код, если $\deg(F) = \deg(G) - g - t + 1$ и $g = 0$.

Интересным объектом исследования относительно кодовых криптосистем являются подполевые подкоды, поскольку существует гипотеза, что именно такие коды являются стойкими к атаке на основе пар, исправляющих ошибки (по аналогии с классическими кодами Гоппы, являющимися подполевыми подкодами обобщённых кодов Рида — Соломона).

В действительности если $\mathcal{C}|\mathbb{F}_p$ — подполевой подкод кода $\mathcal{C} = \mathcal{C}_{\mathscr{L}}(D, G)$, определённого над $\mathbb{F}_q$, и $\mathbb{F}_p \subseteq \mathbb{F}_q$, то, согласно [7], пара, исправляющая t ошибок, для кода $\mathcal{C}$ является парой, исправляющей такое же количество ошибок, и для подполевого подкода $\mathcal{C}|\mathbb{F}_p$. При этом алгоритм декодирования будет работать над расширением $\mathbb{F}_q$ конечного поля $\mathbb{F}_p$ за время $\mathcal{O}((mn)^3)$, где $q = p^m$. Соответственно вопрос редукции сложности задачи декодирования подполевых подкодов сводится к нахождению пары, исправляющей ошибки, для подполевого подкода над $\mathbb{F}_p$.

Теорема 3. Пусть $\mathcal{C} = \mathcal{C}_{\mathscr{L}}(D, G)$ — АГ-код, ассоциированный с функциональным полем $F/\mathbb{F}_q$ рода g. Тогда парой, исправляющей $t = (n - \deg(G) - g - 1)/2$ ошибок, для кода $\mathcal{C}|\mathbb{F}_p$ является:

— при $\deg(G) = 1$ и $g = 0$:
$$\mathcal{A} = (\mathcal{C}_{\mathscr{L}}(D, F)^{\perp})|_{\mathbb{F}_p},$$
где $\deg(F) = t$;

— при $\deg(G) \leqslant n - g + 1$:
$$\mathcal{A} = (\mathcal{C}_{\mathscr{L}}(D, F)|_{\mathbb{F}_p})^{\perp},$$
где $t = n/2 - 1$ и $\deg(F) = n/2 + g - 1$;

— при $(n + 1)/3 - g \leqslant \deg(G) \leqslant n + 1 - g$:
$$\mathcal{A} = \mathcal{C}_{\mathscr{L}}(D, F)|_{\mathbb{F}_p},$$
где $\deg(F) = n + g - t - 1$;

— при $\deg(G) \leqslant n - g + 1$:
$$\mathcal{A} = ((\mathcal{C}_{\mathscr{L}}(D, F)^{\perp})|_{\mathbb{F}_p})^{\perp},$$
где $t = n/2$ и $\deg(F) = n/2 + g - 1$.

Во всех четырёх случаях $\mathcal{B} = (\mathcal{A} * \mathcal{C}|_{\mathbb{F}_p})^{\perp}$.

Теорема 4. Пусть $\mathcal{C} = \mathcal{C}_{\mathscr{L}}(D, G)$ — АГ-код, ассоциированный с функциональным полем $F/\mathbb{F}_q$ рода g, и $\mathcal{C}^{\perp}$ — дуальный к $\mathcal{C}$. Тогда парой, исправляющей $t = (\deg(G) - 3g + 1)/2$ ошибок, для кода $(\mathcal{C}^{\perp})|_{\mathbb{F}_p}$ является:

— при $2n/3 + 2g - 3 \leqslant \deg(G) \leqslant (2n+1)/3 + 3g - 2$ и $g = 0, 1$:

$$\mathcal{A} = (\mathcal{C}_{\mathscr{L}}(D, F)^{\perp})|_{\mathbb{F}_p},$$

где $\deg(F) = t + g$;

— при $3g - 3 \leqslant \deg(G) \leqslant (4n-1)/3 + 3g - 2$:

$$\mathcal{A} = (\mathcal{C}_{\mathscr{L}}(D, F)|_{\mathbb{F}_p})^{\perp},$$

где $t = n/2 - 1$ и $\deg(F) = n/2 + g - 1$.

В обоих случаях $\mathcal{B} = (\mathcal{A} * (\mathcal{C}^{\perp})|_{\mathbb{F}_p})^{\perp}$.

Заключение

Для обеспечения условия 2 в определении пары, исправляющей ошибки, мы ограничиваемся рассмотрением случаев, когда $\deg(F) = t + g$ и $\deg(F) = n + g - t - 2$, хотя данные значения являются нижней и верхней границами соответственно для $\deg(F)$ в зависимости от вида кода $\mathcal{A}$.

Весьма интересным представляется также вычисление пар, исправляющих ошибки, для трэйс-кодов (такие коды получены с помощью применения к кодовым словам кода $\mathcal{C}$, определённого над $\mathbb{F}_q$, функции следа $\mathrm{Tr} : \mathcal{C} \to \mathbb{F}_p$, где $\mathbb{F}_p \subseteq \mathbb{F}_q$), поскольку такие коды связаны соотношением $(\mathcal{C}|_{\mathbb{F}_p})^{\perp} = \mathrm{Tr}(\mathcal{C}^{\perp})$.

ЛИТЕРАТУРА

1. *Justesen J., Larsen K., Jensen H., et al.* Construction and decoding of a class of algebraic geometry codes // IEEE Trans. Inform. Theory. 1989. No. 35(4). P. 811–821.
2. *Skorobogatov A. N. and Vlădut S. G.* On the decoding of algebraic-geometric codes // IEEE Trans. Inform. Theory. 1990. No. 36(5). P. 1051–1060.
3. *Pellikaan R.* On decoding by error location and dependent sets of error positions // Discrete Math. 1992. No. 106–107. P. 369–381.
4. *Kötter R.* A unified description of an error locating procedure for linear codes // Proc. Algebraic Combinatorial Coding Theory III. Voneshta Voda, Bulgaria, 1992. P. 113–117.
5. *Couvreur A., Marquez-Corbella I., and Pellikaan R.* Cryptanalysis of McEliece cryptosystem based on algebraic geometry codes and their subcodes // IEEE Trans. Inform. Theory. 2017. No. 63. P. 5404–5418.
6. *Stichtenoth H.* Algebraic Function Fields and Codes. Berlin; Heidelberg: Springer, 1991.
7. *Pellikaan R.* On the existence of error-correcting pairs // Statistical Planning and Inference. 1996. No. 51. P. 229–242.
8. *Marquez-Corbella I. and Pellikaan R.* Error-correcting pairs: a new approach to code-based cryptography // 20th Conf. Appl. of Computer Algebra. 2014. `https://hal.science/hal-01088433`.

УДК 519.7 DOI 10.17223/2226308X/16/37

ПЕРИОДИЧЕСКИЕ СВОЙСТВА КОНЕЧНО-АВТОМАТНОГО ГЕНЕРАТОРА

П. К. Обухов, И. А. Панкратова

Изучаются периодические свойства двухкаскадного конечно-автоматного генератора. Получено значение максимального периода генератора и некоторые необходимые условия его достижения.

Ключевые слова: *конечно-автоматный генератор, подстановки, периодические последовательности.*

Рассматривается предложенный Г. П. Агибаловым двухкаскадный конечно-автоматный криптографический генератор $G = A_1 \cdot A_2$ [1], схема которого показана на рис. 1. Генератор представляет собой последовательное соединение автономного автомата $A_1 = (\mathbb{F}_2^n, \mathbb{F}_2, g_1, f_1)$ (с функцией переходов $g_1 : \mathbb{F}_2^n \to \mathbb{F}_2^n$ и функцией выходов $f_1 : \mathbb{F}_2^n \to \mathbb{F}_2$) и автомата $A_2 = (\mathbb{F}_2, \mathbb{F}_2^m, \mathbb{F}_2, g_2, f_2)$ (с функцией переходов $g_2 : \mathbb{F}_2 \times \mathbb{F}_2^m \to \mathbb{F}_2^m$ и функцией выходов $f_2 : \mathbb{F}_2 \times \mathbb{F}_2^m \to \mathbb{F}_2$), $n, m \geqslant 1$.

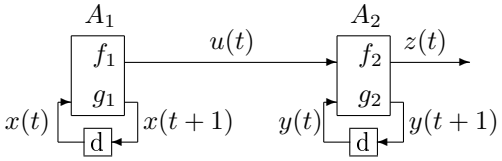

Рис. 1. Схема генератора G

Генератор функционирует в дискретном времени $t = 1, 2, \ldots$, в каждый момент t которого автомат A_1, находясь в состоянии $x(t) \in \mathbb{F}_2^n$, выдаёт выходной символ $u(t) = f_1(x(t))$ и переходит в следующее состояние $x(t+1) = g_1(x(t))$, а автомат A_2, находясь в состоянии $y(t) \in \mathbb{F}_2^m$, принимает от A_1 символ $u(t)$, выдаёт на выход генератора выходной символ $z(t) = f_2(u(t), y(t))$ и переходит в следующее состояние $y(t+1) = g_2(u(t), y(t))$. Последовательность $u(1) \ldots u(l)$, $l \in \mathbb{N}$, выходных символов автомата A_1 называется управляющей последовательностью автомата A_2, а последовательность $z(1) \ldots z(l)$ выходных символов автомата A_2 — выходной последовательностью генератора G. Ключом генератора может быть любое непустое подмножество множества $\{x(1), y(1), f_1, g_1, f_2, g_2\}$.

Периодом генератора назовём длину периода его выходной последовательности, получим его оценки и исследуем условия максимальности. Требование большого периода генератора необходимо для противостояния атаке на шифртекст, зашифрованный с повторным использованием ключевой последовательности [2, с. 139].

Утверждение 1. Если $(x(t), y(t)) = (x(l), y(l))$ для некоторых $t, l \in \mathbb{N}$, то период генератора делит $(l - t)$.

Доказательство. Пусть, для определённости, $l > t$.

По определению генератора, из условия $x(t) = x(l)$ получаем, во-первых, $x(t+1) = x(l+1)$, во-вторых, $u(t) = u(l)$. Из последнего равенства и условия $y(t) = y(l)$ следует, что $y(t+1) = y(l+1)$ и $z(t) = z(l)$. Рассуждая далее (имея в виду равенство $(x(t+1), y(t+1)) = (x(l+1), y(l+1))$), получаем $z(t+1) = z(l+1)$ и т. д. — отрезок выходной последовательности $z(t) \ldots z(l-1)$ будет повторяться через каждые $(l - t)$ шагов. ■

Следствие 1. Период генератора G не превосходит 2^{n+m}.

Доказательство. Следует из того, что количество различных пар $(x(t), y(t))$ равно 2^{n+m}. ■

Следующий пример показывает, что оценка следствия 1 достижима.

Пример 1. Пусть $n = m = 2$, функции переходов и выходов автоматов A_1, A_2 заданы табл. 1–3, начальные состояния $x(1) = y(1) = 00$.

Таблица 1
A_1

$x(t)$	00	01	10	11
$x(t+1)$	01	10	11	00
$u(t)$	1	0	1	1

Таблица 2
A_2 (g_2)

$u(t)$	$y(t)$			
	00	01	10	11
0	01	10	00	11
1	00	01	11	10

Таблица 3
A_2 (f_2)

$u(t)$	$y(t)$			
	00	01	10	11
0	0	0	0	0
1	1	1	1	0

Выходная последовательность генератора равна 1011101010010011... Через 16 тактов работы получим $x(17) = y(17) = 00$, и последовательность повторится.

Назовём траекторией генератора последовательность пар состояний $(x(t), y(t))$, $t = 1, 2, \ldots, 2^{n+m}$. Из утверждения 1 следует

Утверждение 2. Если период генератора G максимальный (равен 2^{n+m}), то его траектория содержит все возможные пары состояний из $\mathbb{F}_2^n \times \mathbb{F}_2^m$, а отображение

$$\begin{gathered}\psi_G : \mathbb{F}_2^n \times \mathbb{F}_2^m \to \mathbb{F}_2^n \times \mathbb{F}_2^m,\ \big(x(t), y(t)\big) \mapsto \big(x(t+1), y(t+1)\big),\\ \big(x(t+1), y(t+1)\big) = \big(g_1(x(t)), g_2(f_1(x(t)), y(t))\big),\end{gathered} \tag{1}$$

является полноцикловой подстановкой.

Заметим, что условие утверждения 2 является только необходимым, но недостаточным для максимальности периода генератора. Приведём ещё некоторые необходимые условия.

Утверждение 3. Если период генератора G равен 2^{n+m}, то:

1) функция g_1 является полноцикловой подстановкой;
2) изменение начального состояния $x(1)$ или $y(1)$ не влияет на период генератора;
3) функция f_1 — не константа;
4) хотя бы одна из подфункций $f_2(0, \cdot)$ и $f_2(1, \cdot)$ — не константа;
5) подфункции $g_2(0, \cdot)$ и $g_2(1, \cdot)$ функции переходов g_2 являются подстановками;
6) $y(2^n i + j) \neq y(2^n k + j)$ для всех $i, k \in \{0, \ldots, 2^m - 1\}$, $i \neq k$, $j = 1, \ldots, 2^n$.

Доказательство.
1) Предположим, что функция $g_1 : \mathbb{F}_2^n \to \mathbb{F}_2^n$ — не подстановка. Тогда

$$\exists a \in \mathbb{F}_2^n\, \forall b \in \mathbb{F}_2^n \big(g_1(b) \neq a\big).$$

Но тогда пара с первой компонентой a встретится в траектории (1) только один раз, если $x(1) = a$, и ни одного иначе. И то и другое противоречит тому, что траектория содержит все возможные пары из $\mathbb{F}_2^n \times \mathbb{F}_2^m$, в частности, $2^m \geqslant 2$ пар с первой компонентой a.

Предположим, что начальное состояние $x(1)$ принадлежит циклу длины меньше 2^n подстановки g_1. Тогда первые компоненты всех пар в траектории (1) принадлежат тому же циклу, что противоречит утверждению 2.

2) По утверждению 2, траектория генератора содержит все возможные пары состояний автоматов A_1 и A_2; кроме того, выходная последовательность $z(1)z(2)\dots z(2^{n+m})$ не содержит циклов. Изменение начального состояния $x(1)$ или $y(1)$ приведёт к циклическому сдвигу траектории и, как следствие, выходной последовательности, что никак не отразится на её периоде.

3) Предположим, что $f_1 = \operatorname{const} c$. Тогда выходная последовательность генератора вычисляется по формуле $z(t) = f_2(c, y(t))$, $t = 1, \dots$, и не может иметь период больше, чем 2^m — количество различных значений $y(t)$.

4) Предположим, что $f_2(0,\cdot) = \operatorname{const} c_1$ и $f_2(1,\cdot) = \operatorname{const} c_2$. Если $c_1 = c_2$, то $z(t) = c_1$ для всех $t = 1, 2, \dots$ и выходная последовательность имеет период 1. Если $c_1 \neq c_2$, то $z(t) = f_1(x(t))$ или $z(t) = \neg f_1(x(t))$ — и в том и в другом случае период выходной последовательности не больше 2^n.

5) Предположим, что подфункция $g_2(0, y) : \mathbb{F}_2^m \to \mathbb{F}_2^m$ не является подстановкой. Тогда

$$\exists a \in \mathbb{F}_2^m \, \forall b \in \mathbb{F}_2^m \, \big(g_2(0, b) \neq a\big).$$

Выберем значение $c \in \mathbb{F}_2^n$, что $f_1(c) = 0$ (ввиду п. 3 такое c обязательно найдётся); пусть $g_1(c) = d$. Но тогда пара (d, a) отсутствует в траектории (1), поскольку её предшественником может быть только пара (c, b) для некоторого $b \in \mathbb{F}_2^m$, а $\Psi_G\big((c, b)\big) = \big(g_1(c), g_2(0, b)\big) \neq (d, a)$ для любого $b \in \mathbb{F}_2^m$. Получили противоречие с утверждением 2. Для $g_2(1, y)$ доказательство аналогично.

6) Следует из утверждения 2 и того, что $x(2^n i + j) = x(2^n k + j)$ для всех $i, j, k \in \mathbb{N}$ (ввиду п. 1).

Утверждение 3 доказано. ■

Утверждение 4. Если в генераторе G функция g_1 и подфункции $g_2(0,\cdot)$, $g_2(1,\cdot)$ являются подстановками, то выходная последовательность генератора чисто периодическая.

Доказательство. Выходной символ генератора G вычисляется по формуле

$$z(t) = f_2(f_1(x(t)), y(t)),$$

т. е. зависит только от состояний автоматов A_1 и A_2. Поэтому наличие предпериода в выходной последовательности означает наличие предпериода в траектории, т. е. для некоторых $t, l \in \mathbb{N}$ выполнены условия $\Psi_G\big((x(t), y(t))\big) = (x(t+1), y(t+1)) = \Psi_G\big((x(l), y(l))\big)$ и $(x(t), y(t)) \neq (x(l), y(l))$. Тогда $g_1(x(t)) = x(t+1) = g_1(x(l))$, следовательно, $x(t) = x(l)$, а значит, $f_1(x(t)) = f_1(x(l)) = c$ для некоторого $c \in \{0, 1\}$.

Продолжаем далее: $g_2(c, y(t)) = y(t+1) = g_2(c, y(l))$, но при $y(t) \neq y(l)$ это невозможно, так как $g_2(c,\cdot)$ — подстановка. ■

ЛИТЕРАТУРА

1. *Боровкова И. В., Панкратова И. А., Семенова Е. В.* Криптоанализ двухкаскадного конечно-автоматного генератора с функциональным ключом // Прикладная дискретная математика. 2018. № 42. С. 48–56.
2. *Алферов А. П., Зубов А. Ю., Кузьмин А. С., Черемушкин А. В.* Основы криптографии. М.: Гелиос АРВ, 2002. 480 с.

СВЕДЕНИЯ ОБ АВТОРАХ

БАБУЕВА Александра Алексеевна — ведущий инженер-аналитик ООО «КРИПТО-ПРО», г. Москва. E-mail: **babueva@cryptopro.ru**

БАКСОВА Ирина Павловна — аспирантка кафедры дискретной математики ММФ МГУ, сотрудник лаборатории ТВП, г. Москва. E-mail: **ibaksova@bk.ru**

БАХАРЕВ Александр Олегович — магистрант ММФ НГУ, инженер Математического центра в Академгородке, г. Новосибирск. E-mail: **a.bakharev@g.nsu.ru**

БОБРОВСКИЙ Дмитрий Александрович — аспирант Финансового университета при Правительстве Российской Федерации, ведущий системный аналитик ООО «Код Безопасности», г. Москва. E-mail: **dabobrovskiy@gmail.com**

БОРЛАКОВ Радмир Русланович — студент ИКТИБ ЮФУ, г. Таганрог.

БУГРОВ Алексей Дмитриевич — кандидат физико-математических наук, г. Москва. E-mail: **Bugrovalexey1@yandex.ru**

БЫКОВ Денис Александрович — студент Новосибирского государственного университета, г. Новосибирск. E-mail: **den.bykov.2000i@gmail.com**

ГЛУХОВ Михаил Михайлович — кандидат физико-математических наук, Московский технический университет связи и информатики, г. Москва. E-mail: **glukhovmm@rambler.ru**

ДЕНИСОВ Олег Викторович — кандидат физико-математических наук, доцент, IT-специалист ООО «Инновационные телекоммуникационные технологии», г. Москва. E-mail: **denisovOleg@yandex.ru**

ЕГОРУШКИН Олег Игоревич — кандидат физико-математических наук, доцент Сибирского государственного университета науки и технологий имени академика М. Ф. Решетнёва, г. Красноярск. E-mail: **olegegoruschkin@yandex.ru**

ЖАРКОВА Анастасия Владимировна — кандидат физико-математических наук, доцент Саратовского национального исследовательского государственного университета имени Н. Г. Чернышевского, г. Саратов. E-mail: **ZharkovaAV3@gmail.com**

ЗАИКИН Олег Сергеевич — кандидат технических наук, ведущий сотрудник ИДСТУ СО РАН, г. Иркутск. E-mail: **oleg.zaikin@icc.ru**

ИЩУКОВА Евгения Александровна — кандидат технических наук, доцент, доцент ИКТИБ ЮФУ, г. Таганрог. E-mail: **uaishukova@sfedu.ru**

КАМЛОВСКИЙ Олег Витальевич — кандидат физико-математических наук, ведущий научный сотрудник ООО «Центр сертификационных исследований», г. Москва. E-mail: **ov-kam@yandex.ru**

КИРЮХИН Виталий Александрович — старший специалист ООО «СФБ Лаб», АО «ИнфоТеКС», г. Москва. E-mail: **vitaly.kiryukhin@sfblaboratory.ru**

КОЛБАСИНА Ирина Валерьевна — старший преподаватель Сибирского государственного университета науки и технологий имени академика М. Ф. Решетнёва, г. Красноярск. E-mail: **kabaskina@yandex.ru**

КОЛЕСНИКОВ Никита Сергеевич — младший научный сотрудник лаборатории «Математические методы защиты информации» Северо-Западного центра математических исследований имени С. Ковалевской, БФУ им. И. Канта, г. Калининград. E-mail: **nikolesnikov100@gmail.com**

КОЛЕСНИКОВ Сергей Геннадьевич — доктор физико-математических наук, доцент, профессор Сибирского государственного университета науки и технологий, профессор Сибирского федерального университета, г. Красноярск. E-mail: **sklsnkv@mail.ru**

КОЛОМЕЕЦ Николай Александрович — кандидат физико-математических наук, научный сотрудник Института математики им. С. Л. Соболева СО РАН, старший преподаватель кафедры теоретической кибернетики ММФ НГУ, г. Новосибирск. E-mail: **kolomeec@math.nsc.ru**

КОНДЫРЕВ Дмитрий Олегович — младший научный сотрудник Математического центра в Академгородке, г. Новосибирск. E-mail: **dkondyrev@gmail.com**

КОРЕНЕВА Алиса Михайловна — кандидат физико-математических наук, начальник отдела криптографического анализа ООО «Код Безопасности», доцент департамента ИБ Финансового университета при Правительстве РФ, г. Москва. E-mail: **A.Koreneva@securitycode.ru**

КУЗНЕЦОВ Александр Алексеевич — доктор физико-математических наук, профессор, директор Института космических исследований и высоких технологий Сибирского государственного университета науки и технологий имени академика М. Ф. Решетнева, г. Красноярск. E-mail: **alex_kuznetsov80@mail.ru**

КУЗНЕЦОВА Александра Сергеевна — кандидат физико-математических наук, доцент кафедры прикладной математики Сибирского государственного университета науки и технологий имени академика М. Ф. Решетнева, г. Красноярск. E-mail: **alexakuznetsova85@gmail.com**

КУНИНЕЦ Артем Андреевич — студент ОНК «Институт высоких технологий» БФУ им. И. Канта, г. Калининград. E-mail: **artkuninets@yandex.ru**

КУРОЧКИН Алексей Вячеславович — преподаватель МФТИ, ведущий системный аналитик ООО «Код Безопасности», г. Москва. E-mail: **kurochkin.av@phystech.edu**

КУЦЕНКО Александр Владимирович — кандидат физико-математических наук, научный сотрудник Института математики им. С. Л. Соболева СО РАН, ассистент механико-математического факультета Новосибирского государственного университета, г. Новосибирск. E-mail: **alexandrkutsenko@bk.ru**

КЯЖИН Сергей Николаевич — кандидат физико-математических наук, ведущий инженер-аналитик ООО «КРИПТО-ПРО», г. Москва. E-mail: **kyazhin@cryptopro.ru**

ЛЕОНТЬЕВ Владимир Маркович — аспирант Сибирского федерального университета, г. Красноярск. E-mail: **v.m.leontiev@outlook.com**

МАЛЫГИНА Екатерина Сергеевна — кандидат физико-математических наук, доцент ОНК «Институт высоких технологий», младший научный сотрудник лаборатории «Математические методы защиты и обработки информации» Северо-Западного центра математических исследований имени С. Ковалевской, БФУ им. И. Канта, г. Калининград. E-mail: **emalygina@kantiana.ru**

МАРО Екатерина Александровна — кандидат технических наук, доцент ИКТИБ ЮФУ, г. Таганрог. E-mail: **eamaro@sfedu.ru**

МЕДВЕДЕВ Анатолий Александрович — студент Национального исследовательского Томского государственного университета, г. Томск. E-mail: **anatoliy2499251007@gmail.com**

МОКРОУСОВ Антон Сергеевич — студент Новосибирского государственного университета, г. Новосибирск. E-mail: **settingx@mail.ru**

МУСУГАЛИЕВА Альбина Геннадьевна — выпускница кафедры теоретических основ компьютерной безопасности и криптографии Саратовского национального исследовательского государственного университета имени Н. Г. Чернышевского, г. Саратов. E-mail: **Albina_Musugalieva@mail.ru**

НОВОСЕЛОВ Семен Александрович — кандидат физико-математических наук, младший научный сотрудник лаборатории «Математические методы защиты информации» Северо-Западного центра математических исследований имени С. Ковалевской, БФУ им. И. Канта, г. Калининград. E-mail: **snovoselov@kantiana.ru**

ОБУХОВ Петр Кириллович — студент Национального исследовательского Томского государственного университета, г. Томск. E-mail: **petr5542@gmail.com**

ПАНАСЕНКО Сергей Петрович — кандидат технических наук, директор по научной работе компании «Актив», г. Москва. E-mail: **panasenko@guardant.ru**

ПАНКОВ Константин Николаевич — кандидат физико-математических наук, заведующий сектором постквантовой криптографии Квантового центра, ВРИО заведующего кафедрой Московского технического университета связи и информатики, эксперт ТК-159 и ТК-362, г. Москва. E-mail: **k.n.pankov@yandex.ru**

ПАНКРАТОВА Ирина Анатольевна — кандидат физико-математических наук, доцент, заведующая лабораторией компьютерной криптографии Национального исследовательского Томского государственного университета, г. Томск. E-mail: **pank@mail.tsu.ru**

ПАРФЕНОВ Денис Романович — студент второго курса магистратуры ФИТ НГУ, г. Новосибирск. E-mail: **d.parfenov@g.nsu.ru**

ПОГОРЕЛОВ Борис Александрович — доктор физико-математических наук, профессор, действительный член Академии криптографии Российской Федерации, г. Москва.

ПУДОВКИНА Марина Александровна — доктор физико-математических наук, профессор НИЯУ «МИФИ», г. Москва. E-mail: **maricap@rambler.ru**

РАЗЕНКОВ Семен Игоревич — студент Национального исследовательского Томского государственного университета, г. Томск. E-mail: **sirazenkov@stud.tsu.ru**

САФОНОВ Константин Владимирович — доктор физико-математических наук, профессор, заведующий кафедрой Сибирского государственного университета науки и технологий имени академика М. Ф. Решетнёва, г. Красноярск. E-mail: **safonovkv@rambler.ru**

СЕМЁНОВ Александр Анатольевич — кандидат технических наук, доцент, ведущий научный сотрудник Института динамики систем и теории управления им. В. М. Матросова СО РАН, г. Иркутск. E-mail: **biclop.rambler@yandex.ru**

СЕРГЕЕВ Андрей Михайлович — специалист ООО «СФБ Лаб», г. Москва.
E-mail: **Andrey.Segreev@sfblaboratory.ru**

СМИРНОВ Антон Михайлович — ассистент кафедры «Криптография и безопасность компьютерных систем» НИЯУ «МИФИ», г. Москва. E-mail: **smirnov.a98@yandex.ru**

ТАРАННИКОВ Юрий Валерьевич — кандидат физико-математических наук, доцент кафедры дискретной математики ММФ МГУ, г. Москва. E-mail: **yutarann@gmail.com**

ТОКАРЕВА Наталья Николаевна — кандидат физико-математических наук, доцент Новосибирского государственного университета, ведущий научный сотрудник Математического центра в Академгородке, г. Новосибирск; научный сотрудник лаборатории «Математические методы защиты информации» Северо-Западного центра математических исследований имени С. Ковалевской, БФУ им. И. Канта, г. Калининград. E-mail: **crypto1127@mail.ru**

ТРЕПАЧЕВА Алина Викторовна — программист, Южный федеральный университет, г. Ростов-на-Дону/Таганрог. E-mail: **alina1989malina@ya.ru**

ФИРСОВ Георгий Валентинович — старший системный программист ООО «Код Безопасности», Национальный исследовательский ядерный университет «МИФИ», г. Москва.
E-mail: **G.Firsov@securitycode.ru**

ХИЛЬЧУК Ирина Сергеевна — аспирант ММФ НГУ, младший научный сотрудник Математического центра в Академгородке, г. Новосибирск. E-mail: **irina.khilchuk@gmail.com**

ЦАРЕГОРОДЦЕВ Кирилл Денисович — аспирант Московского государственного университета им. М. В. Ломоносова, старший специалист-исследователь АО «НПК “Криптонит”», г. Москва.
E-mail: **kirill94_12@mail.ru**

ЧУХНО Андрей Борисович — преподаватель кафедры компьютерной безопасности МИЭМ НИУ ВШЭ, эксперт ООО «Код Безопасности», г. Москва. E-mail: **achuhno@hse.ru**

ШАПОРЕНКО Александр Сергеевич — аспирант ММФ НГУ, младший научный сотрудник Математического центра в Академгородке, г. Новосибирск. E-mail: **shaporenko.alexandr@gmail.com**

ЩЕРБАЧЕНКО Андрей Александрович — специалист ООО «СФБ-Лаб», г. Москва.
E-mail: **Andrey.Shcherbachenko@sfblaboratory.ru**

HARIS Muhammad — master's student of FIT NSU, Novosibirsk. E-mail: **m.kharis@g.nsu.ru**

PAL Santu — researcher at Novosibirsk State University and Sobolev Institute of Mathematics, Novosibirsk. E-mail: **santukgp@gmail.com**

QAYYUM Abdul — master's student of FIT NSU, Novosibirsk. E-mail: **a.kaiyum@g.nsu.ru**

АННОТАЦИИ ДОКЛАДОВ НА АНГЛИЙСКОМ ЯЗЫКЕ

SECTION 1

Baksova I. P., Tarannikov Yu. V. **BOUNDS ON THE NUMBER OF PARTITIONS OF THE VECTOR SPACE OVER A FINITE FIELD INTO AFFINE SUBSPACES OF THE SAME DIMENSION.** We give lower and upper bounds on the number of ordered $N_m^k(q)$ and unordered $\widetilde{N}_m^k(q)$ partitions of the space $\mathbb{F}_q^m$ into affine subspaces of the same dimension k. In particular, the asymptotics of the logarithm of the number of unordered partitions of the space $\mathbb{F}_3^m$ into one-dimensional affine subspaces is established:

$$\frac{m}{3}\cdot 3^m + c_1\cdot 3^m + o\left(3^m\right) \leqslant \log_3 \widetilde{N}_m^1(3) \leqslant \frac{m}{3}\cdot 3^m + c_2\cdot 3^m + o\left(3^m\right).$$

Also, we highlight the bounds

$$\log_q N_m^k(q) \gtrsim (m-k)q^{m-k}, \quad m-k\to\infty,$$
$$\log_3 N_m^k(3) \gtrsim 2\,(m-k)\,3^{m-k},$$
$$\log_q N_m^k(q) \gtrsim \left(m-\frac{q-1}{q}\,k\right)q^{m-k}, \quad k\to\infty, \quad m-k\to\infty$$
$$\log_q N_m^k(q) \leqslant (k+1)(m-k-\log_q e)q^{m-k} + O(q^{m-k}) + O(k(m-k)).$$

Keywords: *affine subspaces, partitions of a space, bounds, bent functions.*

Pogorelov B. A., Pudovkina M. A. **MULTIPERMUTATIONS AND PERFECT DIFFUSION OF PARTITIONS.** Multipermutations are introduced by C.-P. Schnorr and S. Vaudenay as formalization of perfect diffusion in block ciphers. In this paper, we consider an abelian group X and a set H of transformations on X^2 introduced by S. Vaudenay. Any bijective transformation from H is a multipermutation. Multipermutations from H are defined by orthomorphisms on X. The set H is nonempty iff there exists an orthomorphism on X. We consider a set W of distinct cosets of W_0 in X. We describe multipermutations from H such that they perfectly diffuse one of partitions W^2 or $X\times W$. As an example, we prove that 8-bit and 16-bit transformations of CS-cipher perfectly diffuse such partitions.
Keywords: *multipermutation, orthomorphism, Quasi-Hadamard transformation, perfect diffusion of partitions, CS-cipher.*

SECTION 2

Bugrov A. D. **PROPERTIES OF CLASSES OF BOOLEAN FUNCTIONS CONSTRUCTED FROM SEVERAL LINEAR RECURRENCES OVER THE RING OF INTEGERS MODULO 2^n.** A class of Boolean functions constructed from high-coordinate sequences of linear recurrences over the ring $\mathbb{Z}_{2^n}$ is defined. Various coordinate sets are used to isolate the coordinate sequences. It is shown that this class consists of functions that are significantly removed from the class of all affine functions.

Keywords: *linear recurrent sequences, coordinate sequences, Boolean functions, nonlinearity of Boolean functions.*

Bykov D. A. **ON TIGHTNESS OF THE LOWER BOUND FOR THE NUMBER OF BENT FUNCTIONS AT THE MINIMUM DISTANCE FROM A BENT FUNCTION FROM THE MAIORANA —MCFARLAND CLASS.** The lower bound $2^{2n+1} - 2^n$ for the number of bent functions at the minimum distance from a bent function from the Maiorana — McFarland class $\mathcal{M}_{2n}$ in $2n$ variables is investigated. A criterion for the reachability of this lower bound for functions in algebraic representation is presented. It is constructively proven that it is accurate for $n = p^k$, where $p \neq 2, 3$ is prime and k is natural. It is shown that a necessary condition for the reachability of the bound is the construction of a function from $\mathcal{M}_{2n}$ using an APN permutation whose set of values on any affine subspace of dimension 3 is not an affine subspace.
Keywords: *bent function, Boolean function, minimum distance, Maiorana — McFarland class, lower bound.*

Kamlovskii O. V., Pankov K. N. **SOME CLASSES OF RESILIENT FUNCTIONS OVER GALOIS RINGS AND THEIR LINEAR CHARACTERISTICS.** Let $R = \mathrm{GR}(q^l, p^l) = \{r_1, \ldots, r_{q^l}\}$ be a Galois ring. Let $A_n(R)$ be a set of all affine functions $g(\mathbf{x}) = a_0 + a_1x_1 + \ldots + a_nx_n = a_0 + \langle \mathbf{a}, \mathbf{x} \rangle$, where $\mathbf{x} = (x_1, \ldots, x_n)$, $a_0 \in R$, $\mathbf{a} = (a_1, \ldots, a_n) \in R^n$. We present some classes of resilient function $f : R^n \to R$ with the specified value of linear characteristic $C(f)$, where $C(f) = \max\limits_{a \in R \setminus \{0\}} \max\limits_{g \in A_n(R)} \left| \sum\limits_{x_1, \ldots, x_n \in R} \chi(af(\mathbf{x}) - g(\mathbf{x})) \right|$ and χ is the canonical additive character of the ring R. In the paper, we describe the function f using a branching construction of the given functions $f_1, \ldots, f_{r_{q^l}}$ in $n-1$ variables. We prove that in the case when the functions $f_1, \ldots, f_{r_{q^l}}$ are k-resilient, the resulting function f is also k-resilient. Moreover, $C(f) \leqslant C(f_{r_1}) + \ldots + C(f_{r_{q^l}})$. We also describe the function $f(\mathbf{x}, \mathbf{y}) = \langle \varphi(\mathbf{x}), \mathbf{y} \rangle + h(\mathbf{x})$, where $n = 2k$, $\varphi : R^k \to R^k$, $h : R^k \to R$, $\mathbf{x}$, $\mathbf{y} \in R^k$. It is known that in the case $\varphi(R^k) \subset (R^*)^k$ (R^* is the group of units in the ring R) the function f is $(k-1)$-resilient. We prove that in the case $|\varphi^{-1}(\mathbf{c})| \leqslant t$ for all $\mathbf{c} \in R^k$ the enequality $C(f) \leqslant q^{k(2l-1)}$ is true.
Keywords: *discrete functions, resilient functions, Galois rings, linear characteristic of functions.*

Kolomeec N. A. **ON PRESERVING THE STRUCTURE OF A SUBSPACE BY A VECTORIAL BOOLEAN FUNCTION.** We consider the following property of a function $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$: F preserves the structure of an affine subspace $U \subseteq \mathbb{F}_2^n$ if $F(U) = \{F(x) : x \in U\}$ is an affine subspace of $\mathbb{F}_2^m$. The connection between this property and the existence of component functions of F whose restriction to the subspace is constant is established. Estimations for the nonlinearity and the order of differential uniformity of such F are provided. We also prove that the set of dimensions of affine subspaces whose structure is preserved by the multiplicative inversion function is the smallest among all one-to-one monomial functions.
Keywords: *affine subspaces, invariant subspaces, nonlinearity, differential uniformity, APN functions, monomial functions*

Kutsenko A. V. **GRAM MATRICES OF BENT FUNCTIONS AND PROPERTIES OF SUBFUNCTIONS OF QUADRATIC SELF-DUAL BENT FUNCTIONS.** A Boolean function in even number of variables n is called a bent function if it has flat Walsh — Hadamard spectrum consisting of numbers $\pm 2^{n/2}$. A bent function is

called self-dual if it coincides with its dual bent function. Previously the author obtained a sufficient condition for subfunctions in $n-2$ variables of a self-dual bent function in n variables, obtained by fixing the first two variables, to be bent. In this paper, we prove that for quadratic self-dual bent functions this condition is not necessary for $n \geqslant 6$. The concept of the Gram matrices of Boolean functions is introduced, the general form of the Gram matrix of a bent function and its dual function are obtained. It is proved that if the Gram matrix of a bent function in n variables is non-invertible, then its subfunctions in $n-2$ variables, obtained by fixing the first two variables, are bent functions. It is also proved that the subfunctions of its dual bent function are also bent functions.

Keywords: *self-dual bent function, subfunction, Gram matrix, quadratic function, 4-decompositions.*

Pankratova I. A., Medvedev A. A. **CONSTRUCTION OF A SUBSTITUTION ON $\mathbb{F}_2^n$ BASED ON A SINGLE BOOLEAN FUNCTION.** The following construction of a vector Boolean function is considered: $F(x) = \big(f(x), f(\pi(x)), f(\pi^2(x)), \ldots, f(\pi^{n-1}(x))\big)$, where $\pi \in \mathbb{S}_n$, f is a n-dimensional Boolean function. Some necessary conditions for F to be a bijection are proved, namely: f must be balanced, $f(0^n) \neq f(1^n)$, π must be full cycle substitution, $f \neq$ const on any cycle of substitution π', where $\pi'(a_1 \ldots a_n) = (a_{\pi(1)} \ldots a_{\pi(n)})$ for all $a_1 \ldots a_n \in \mathbb{F}_2^n$.

Keywords: *bijection, vector Boolean function.*

SECTION 3

Denisov O. V. **DISTINGUISHING ATTACK ON FOUR ROUNDS OF THE LUBY — RACKOFF CIPHER BY DIFFERENTIALS OF TWO-BLOCK TEXTS.** We show that the Luby — Rackoff cipher (Feistel network with block length $n = 2m$, random independent round functions $f^1, \ldots, f^R : \mathbb{Z}_2^m \to \mathbb{Z}_2^m$) is a Markov cipher. Matrices $\mathbb{P}^R$ of R-round transition probabilities of differentials have been found for $R = 1, 2, 4$ and arbitrary m. Let $(p_2(\Delta y), y \in \mathbb{X}')$ be the conditional probability distribution of the 4-round output differential under the fixed input differential $\Delta x = (\Delta x_0, \Delta x_1)$, $p_1(\Delta y) = (M^2-1)^{-1}$ — the uniform distribution on the $\mathbb{X}' = \mathbb{Z}_2^{2m} \setminus \{\mathbf{0}\}$, $M = 2^m$. We have obtained expressions for Kullback divergences between the distributions and prove that: 1) if $(\Delta x_0 = 0 \wedge \Delta x_1 \neq 0) \vee (\Delta x_0 \neq 0 \wedge \Delta x_1 \neq 0)$, then $K(2:1) \sim (2\ln 2 - 1)M^{-2}$, $K(1:2) \sim (1 - \ln 2)M^{-2}$ as $M \to \infty$; 2) if $\Delta x_0 \neq 0$, $\Delta x_1 = 0$, then $K(2:1) \sim$ $\sim (2\ln 2 - 1)M^{-1}$, $K(1:2) \sim (1 - \ln 2)M^{-1}$. Therefore, in the second case distinguishing attack (based on the statistics of the logarithm of the likelihood ratio in the model of independent two-block texts) will be more powerful. In particular, for error probabilities $\alpha = \beta = 0{,}1$ and large M, the average values of text amounts are approximately equal $T_1(M) = \dfrac{0{,}8\ln 9}{1 - \ln 2} M = 5{,}72M$, $T_2(M) = \dfrac{0{,}8\ln 9}{2\ln 2 - 1} M = 4{,}55M$. In statistical experiments for $12 \leqslant n \leqslant 44$ empirical probabilities of errors are close to the specified α, β and amounts of texts are close to the calculated values $T_1(M), T_2(M)$.

Keywords: *Markov cipher, Feistel network, Luby — Rackoff cipher, Kullback divergence, sequential distinguishing attack.*

Zaikin O. S. **INVERTING 29-STEP MD5 COMPRESSION FUNCTION VIA SAT.** The cryptographic hash function MD5 was proposed in 1992. Its key component is a 64-step compression function. The compression function is still preimage resistant, that is why its step-reduced versions are usually investigated in this context. In 2007, the 26-step version of the MD5 compression function was inverted via SAT. In 2012, 27- and 28-step

versions were inverted via SAT as well. In the paper, an approach to forming 32 intermediate inversion problems between two subsequent steps of the MD5 compression function is proposed. SAT encodings of such problems were constructed between 28 and 29 steps. Several simplest problems were leveraged for tuning a modern SAT solver. As a result, the 29-step version of the MD5 compression function was inverted for the first time.
Keywords: *cryptographic hash function, MD5, algebraic cryptanalysis, logical cryptanalysis, SAT.*

Ishchukova E. A., Borlakov R. R. **COMPARATIVE ANALYSIS OF THE GRAPHIC INFORMATION TRANSFORMATION QUALITY USING BLOCK CIPHERS.** The paper presents the results of practical experiments on encryption algorithms (DES, AES, Magma) in the Electronic Code Book (ECB) mode. The influence of encryption on the quality of graphic information conversion depending on its properties is shown. During the experiment, the hypothesis was tested. According to our hypothesis, the quality of encryption depends not only on the encryption algorithm and its mode, but also on the properties of the information being converted itself. It was experimentally demonstrated that the quality of information conversion is influenced by such parameters as: the number of colors in the color palette, the number of large and small objects in the picture, the number of pixels, the presence or absence of a background, and others.
Keywords: *encryption, block cipher, encryption mode, electronic code book, graphic information.*

Kolomeec N. A. **ON THE NUMBER OF IMPOSSIBLE DIFFERENTIALS OF SOME ARX TRANSFORMATION.** The additive differential probabilities of the function $(x \oplus y) \lll r$ are considered, where $x, y \in \mathbb{Z}_2^n$ and $1 \leqslant r < n$. They are interesting in the context of differential cryptanalysis of ciphers whose schemes consist of additions modulo 2^n, bitwise XORs ($\oplus$) and bit rotations ($\lll r$). We calculate the number of all impossible differentials, i.e. differentials with probability 0, for all possible r and n. The limit of the ratio of this number to the number of all differentials as r and $n - r$ tend to ∞ equals $38/245$. We also compare the given numbers and the number of impossible differentials for the function $x \oplus y$.
Keywords: *ARX, differential probabilities, XOR, modular addition, bit rotations, impossible differentials.*

Kondyrev D. O. **EFFICIENCY ANALYSIS OF CRYPTOGRAPHIC ALGORITHMS FOR APPLICATIONS IN ZK-SNARK.** The paper presents a comparative efficiency analysis of various cryptographic algorithms in terms of applications in zk-SNARK based systems. To conduct experiments with measuring parameters, an infrastructure based on ZoKrates has been developed. A series of measurements with different input data was carried out for each algorithm. The number of constraints in the R1CS representation of the algorithm, the length of the proof key and the verification key, the running time of the setup phase of the protocol, and the proof generation time have been measured. As a result, we have obtained experimental data that allow us to determine the boundaries of the practical applicability of algorithms in distributed ledgers.
Keywords: *distributed ledgers, zero-knowledge proof, zk-SNARK, R1CS, algorithm efficiency.*

Koreneva A. M., Firsov G. V. **ON ONE BLOCK CIPHER MODE OF OPERATION FOR PROTECTION OF BLOCK-ORIENTED SYSTEM STORAGE DEVICES.** At the end of 2022, standardization recommendations were adopted in the

Russian Federation that define the mode of operation of a block cipher to protect block-oriented storage devices. This mode is called Disk Encryption with Counter. It has several operational characteristics, that complicate its use for system partition encryption. Therefore, the synthesis of alternative modes for full disk encryption is in demand. In the most of exisiting software for system partition encryption XTS mode is used, but it has several properties, that lead to degradation of its cryptographic qualities. This paper proposes a provably secure modification of XTS mode — XEH (Xor-Encrypt-Hash) mode. Furthermore, XEH's performance characteristics have been investigated.
Keywords: *full disk encryption, block cipher mode of operation, symmetric cryptography, cryptographic protection of information, block-oriented storage devices.*

Kurochkin A. V., Chukhno A. B., Bobrovskiy D. A. **CONSTRUCTION OF THE DIFFERENCE RELATION FOR THE KB-256 ALGORITHM.** The difference relation for the KB-256 algorithm has been constructed. The probability of executing the difference relation for 15 out of 16 rounds is not less than 2^{-134}.
Keywords: *differential, linear methods of cryptographic analysis, KB-256.*

Malygina E. S., Kutsenko A. V., Novoselov S. A., Kolesnikov N. S., Bakharev A. O., Khilchuk I. S., Shaporenko A. S., Tokareva N. N. **MAIN APPROACHES IN POST-QUANTUM CRYPTOGRAPHY: DESCRIPTION, A COMPARATIVE STUDY.** Post-quantum cryptography is an area of theoretical and applied research with the goal to develop cryptographic systems that are secure against both quantum and classical computers. Now, among the most promising directions one can mention the lattice-based cryptography, code-based cryptography and isogenies. This paper is a review, it includes a summary of two papers previously prepared by the authors and devoted to the description of the main approaches to the construction of post-quantum cryptosystems. Hard problems from these areas are considered, known results on resilience and performance of the corresponding cryptosystems are analyzed.
Keywords: *post-quantum cryptography, lattice-based cryptography, error-correcting codes, isogenies, quantum computer.*

Maro E. A., Zaikin O. S. **ALGEBRAIC CRYPTANALYSIS OF 9 ROUNDS OF LIGHTWEIGHT BLOCK CIPHER SIMON32/64.** A lightweight block cipher Simon32/64 from the Simon family is considered. Its full version consists of 32 rounds. Cryptanalysis of the first 8 rounds has already been repeatedly performed via SAT, i.e., by reducing to the Boolean satisfiability problem and applying SAT solvers. However, for 9 rounds this is still a challenging problem for the SAT approach. In the paper, a SAT encoding for cryptanalysis of the first 9 rounds of Simon32/64 is constructed. Three classes of cryptanalysis problems are formed depending on how the plain text is chosen. Provided that 16 out of 64 bits of each secret key are known, all the problems were solved via a parallel SAT solver.
Keywords: *lightweight block cipher, Simon family of ciphers, algebraic cryptanalysis, SAT-solver.*

Mokrousov A. S., Kolomeec N. A. **ON ADDITIVE DIFFERENTIALS THAT GO THROUGH ARX TRANSFROMATION WITH HIGH PROBABILITY.** In the paper, we consider additive differential probabilities of the function $(x \oplus y) \lll r$, where $x, y \in \mathbb{Z}_2^n$ and $1 \leqslant r < n$. They are interesting in the context of differential cryptanalysis of ciphers that use addition modulo 2^n, bitwise XOR ($\oplus$) and bit rotations ($\lll r$) as basic operations. All differentials up to argument symmetries whose probability exceeds 1/4 are

obtained. The possible values of their probabilities are $1/3+4^{2-i}/6$ for $i \in \{1, \ldots, n\}$, which coincide with the differentials probabilities of the function $x \oplus y$. We describe differentials with each of these probabilities and calculate the number of them. It is proven that the number of all considered differentials is equal to $48n - 68$ for $n \geqslant 2$.
Keywords: *ARX, differential probabilities, XOR, modular addition, bit rotations.*

Panasenko S. P. **SYMMETRIC LIGHTWEIGHT CRYPTOGRAPHY: PRINCIPLES, APPROACHES, AND TRADEOFFS.** This review is devoted to the development and standardization of lightweight symmetric cryptographic algorithms. The paper briefly describes the reasons for extracting lightweight cryptoalgorithms into a separate class of algorithms designated for use in devices with limited resources. The main directions of standardization of lightweight cryptoalgorithms are listed, both at the level of world and national standards of a number of countries. The main methods of developing lightweight cryptoalgorithms are described, both on the basis of existing general-purpose algorithms and original ones. In the final section, the main trends of the expected further development in the direction of lightweight symmetric cryptography are formulated, including possible deviation from universal lightweight cryptographic standards towards specialized ones, increasing requirements for lightweight cryptographic algorithms in terms of their cryptographic strength, and the formation of new applications of such algorithms, which may affect the methods of their development and the requirements imposed on them.
Keywords: *lightweight cryptography, symmetric encryption, hash functions.*

Parfenov D. R., Bakharev A. O. **ADDITIONAL OPTIMIZATION OF THE GUARANTEED NUMBER OF ACTIVATIONS IN XS-CIRCUITS COMPUTATION ALGORITHM.** We propose an additional optimization to the Guaranteed Number of Activations (GNA) computation algorithm. The main idea of this optimization is to replace linear dependence checks based on the matrix rank computations with suffix checks of paths corresponding to partitions in the search tree. Proposed algorithm has been implemented and is two times faster than the previous solution. Using an optimized version of the GNA computation algorithm, we carried out several computational experiments. As a result, we refuted several hypothesis and proposed a scalable XS-circuit construction with an optimal GNA value.
Keywords: *guaranteed number of activations,* XS*-circuit, differential cryptanalysis.*

Pudovkina M. A., Smirnov A. M. **THE BOOMERANG ATTACK ON THE 4-ROUND LILLIPUT-TBC-II-256 CIPHER.** Lilliput-AE is a tweakable block cipher submitted as a candidate to the NIST lightweight cryptography standardization process. It is an OCB based authenticated encryption scheme using the block cipher Lilliput with a tweakey schedule (LILLIPUT-TBC). It has 128-bit blocks and supports key sizes of 128, 192, and 256 bits. Lilliput-AE has two particular authenticated encryption modes: Lilliput-I and Lilliput-II based respectively on a nonce-respecting mode and a nonce-misuse resistant mode. In this paper, we present an attack on the 4-round LILLIPUT-TBC-II-256 cipher with 256-bit security level using boomerang technique based on Yoyo tricks, which were firstly presented at ASIACRYPT 2017 to attack the AES block cipher. The attack requires 2^{180} encryptions. The data complexity is 2^{24} texts and the memory complexity is $2^{24,3}$ bit. The main result is obtained due to the simultaneous use of boomerang technique and the property of the diffusion transformation.
Keywords: *lightweight cipher, yoyo tricks, authenticated encryption, linear transformation, S-box, OFB mode, boomerang technique.*

Razenkov S. I. **FPGA IMPLEMENTATION OF AN SD-CARD ENCRYPTOR BASED ON MAGMA CIPHER IN COUNTER MODE.** A hardware implementation of an SDSC card encryptor based on Magma cipher in counter mode is described. Implementations using two different FPGA architectures and synthesis software have similar hardware resource utilization. Keystream generation does not depend on the data which allows to increase the clock frequency of the generator. It was shown that this approach significantly reduces device operation time.
Keywords: *FPGA, Magma cipher, counter mode.*

Semenov A. A. **THE USE OF BACKDOORS TO ESTIMATE THE HARDNESS OF PROPOSITIONAL PROOFS AND CRYPTOGRAPHIC ATTACKS.** In the paper, we consider the problem of constructing tree-like unsatisfiability proof certificates under the assumption that this proof is generated by a SAT solver based on the CDCL algorithm. Such tree-like representations are convenient when it is necessary to evaluate how hard it is to prove the unsatisfiability of a specific formula, or to estimate the runtime of some cryptographic attack mounted using the SAT solver. We propose tree-like descriptions of CDCL scenarios in application to both unsatisfiable formulas arising in, e.g. symbolic verification, and to satisfiable formulas encoding the problems of inversion of discrete functions (including cryptographic ones). We prove a number of properties of the introduced tree-like structures. In particular, we formulate the basic property of the class of cryptographic attacks based on inverse backdoor sets in the language of the proposed structures.
Keywords: *Boolean satisfiability (SAT), propositional proof system, backdoor, CDCL algorithm.*

Sergeev A. M., Kiryukhin V. A. **KEY-RECOVERY SECURITY OF KEYED HASH FUNCTIONS BASED ON GOST 34.11-2018 ("STREEBOG").** Keyless hash function GOST 34.11-2018 ("Streebog") is used in many keyed cryptoalgorithms, including HMAC-Streebog and Streebog-K. Using the provable security approach, we obtain the upper bounds on the probability of recovering the secret key for the two algorithms mentioned. We also propose a sandwich-like method of converting "Streebog" to the keyed cryptoalgorithm (conventionally called Streebog-S) without changing the hash function itself. Streebog-S is a secure pseudorandom function and a secure message authentication code. Unlike HMAC-Streebog and Streebog-K, the only key-recovery method for Streebog-S is straightforward guessing. This statement holds under the assumption that the similar is true for the underlying iteratively applied compression function.
Keywords: *Streebog, HMAC, provable security.*

Trepacheva A. V. **ON THE SECURITY OF DOMINGO-FERRER'S HOMOMORPHIC CRYPTOSYSTEM AGAINST CIPHERTEXT-ONLY ATTACK.** The paper proposes an analysis of the security of the Domingo-Ferrer's homomorphic encryption scheme against the ciphertext-only attack. This cryptosystem provides a good counterexample to the equivalence hypothesis of ciphertext-only attack and known plaintext attack on encryption schemes, that are homomorphic over the residue ring modulo a hardly-factorizable number.
Keywords: *homomorphic encryption, cryptanalysys, ciphertext-only attack, Domingo-Ferrer's encryption scheme, factorization problem.*

Tsaregorodtsev K. D. **ON THE ONE QUASIGROUP BASED FORMAT PRESERVING ENCRYPTION ALGORITHM.** One of the possible approaches to the

construction of "medium-sized" format preserving encryption (FPE) schemes is analyzed, which can be described as follows. Let us assume that there is a quasigroup $(M, \circ)$, where M is a "medium-sized" set (i.e., $|M| = 2^{15}$ and above), and we want to construct a tweakable encryption scheme $E_k^\tau \colon M \to M$. Then with the help of k and τ one can generate (using some pseudorandom function) a series of pseudorandom elements $k_i \in M$. To encrypt $m \in M$, one then applies a series of left shifts, i.e., $c \leftarrow k_1 \circ (\ldots (k_\ell \circ m) \ldots) \in M$. The security of this method depends on the security of a pseudorandom function and the security of distinguishing a series of left shifts from the random permutation on M. We show that if one uses functional representation of a quasigroup operation using the proper families of discrete functions over the product of Abelian groups H^n, then left (right) shift, as well as its inverse, can be specified using proper families representation of an operation. A family of functions $F \colon M^n \to M^n$ is called proper iff for any $x, y \in M^n$ there exists i such that $x_i \neq y_i$, but $F_i(x_1, \ldots, x_n) = F_i(y_1, \ldots, y_n)$. If $M = H^n$, where $(H, +)$ is a group, then one can define the following map: $\pi_F = (x_1 + F_1(x_1, \ldots, x_n), \ldots, x_n + F_n(x_1, \ldots, x_n))$, which is a permutation in case of a proper family F. Then we can define a quasigroup operation $x \circ y = \pi_F(x) + \pi_G(y)$, where F and G are two proper families. The following theorem is proven: if F is a proper family over H^n, then the family $\widetilde{F}(x) = (-x) + \pi_F^{-1}(x)$, where $\pi_F(x) = x + F(x)$, $x \in H^n$, is also proper. This theorem allows us to invert the $\circ$ operation using the functional representation: $x = \pi_{\widetilde{F}}((x \circ y) - \pi_G(y))$.
Keywords: *FPE, quasigroup, proper family.*

Shcherbachenko A. A. **AN APPROACH TO CONSTRUCT A KEYED PRF FROM THE "MAGMA" BLOCK CIPHER.** On the basis of recently proposed results for AES, we present new construction, MAGMA-PRF, based on Russian standardized block cipher "MAGMA". We show that MAGMA-PRF is secure against known attacks, which are applicable to plain "MAGMA". We also show that MAGMA-PRF is secure in CTR, CTR-ACPKM, and GCM modes of operations, which, instantiated with PRF instead of PRP, are proven to have better cryptographic properties.
Keywords: *block cipher, encryption modes, MAGMA, MAGMA-PRF, provable security.*

Babueva A. A., Kyazhin S. N. **PUBLIC KEYS FOR E-COINS: PARTIALLY SOLVED PROBLEM USING SIGNATURE WITH RERANDOMIZABLE KEYS.** We give an example of an existing cryptographic mechanism that can be considered as a partial solution to the problem "Public keys for e-coins" proposed at the International Olympiad in Cryptography NSUCRYPTO'2022. This mechanism is used with the class of signatures with rerandomizable keys and provides one of the two security properties required by the authors of the problem. The results of this paper contain a systematic description of security models that can be used to analyze signature with rerandomizable keys, which is of independent interest.
Keywords: *public key derivation, signature with rerandomizable keys, related key attack, BIP32, NSUCRYPTO.*

Pal S. **EFFICIENT MATRIX MULTIPLICATION FOR CRYPTOGRAPHY WITH A COMPANION MATRIX OVER $\mathbb{F}_2$.** A number of schemes in cryptography and other allied areas require operations on matrices that are computationally expensive. However, the computational load due to standard operations like multiplication can be drastically reduced by the choice of special matrices. One such special matrix is the companion matrix of a monic polynomial of degree n over a finite field. Due to its cyclic structure and sparseness property, such a matrix not only helps us to reduce the complexity of matrix

multiplication but also can be applied for cryptographic purposes. In this paper, an algorithm is proposed for the multiplication of an arbitrary matrix with a companion matrix over a finite field of order p. In our algorithm, we not only reduce the complexity but also minimize the number of multiplication operations as much as possible. The complexity of multiplication of any $n \times n$ matrix with a companion matrix of a monic polynomial of degree n is $\mathcal{O}(n^2)$, whereas the complexity of standard matrix multiplication is $\mathcal{O}(n^3)$. Moreover, the number of multiplication operations is $n^2 - nt$, $0 \leqslant t < n$, and 0 for the fields $\mathbb{F}_p$ and $\mathbb{F}_2$ of order p and 2, respectively, which is far less than n^3 multiplications required for standard matrix multiplication.
Keywords: *companion matrix, matrix multiplication, cryptology.*

Qayyum A., Haris M. **CRYPTANALYSIS OF LWE AND SIS-BASED CRYPTOSYSTEMS BY USING QUANTUM ANNEALING.** In the paper, we study lattice-based cryptographic problems, in particular Learning With Errors (LWE) and Short Integer Solution (SIS) lattice problems, which are considered to be known cryptographic primitives that are supposed to be secure against both classical and quantum attacks. We formulated the LWE and SIS problems as Mixed-Integer Programming (MIP) model and then converted them to Quadratic Unconstrained Binary Optimization (QUBO) problem, which can be solved by using a quantum annealer. Quantum annealing searches for the global minimum of an input objective function subjected to the given constraints to optimize the given model. We have estimated the q-bits required for the Quantum Processing Unit (QPU). Our results show that this approach can solve certain instances of the LWE and SIS problems efficiently.
Keywords: *post-quantum cryptography, lattice-based cryptography, learning with errorss, short integer solution, quadratic unconstraint binary optimization, quantum processing unit.*

SECTION 4

Egorushkin O. I., Kolbasina I. V., Safonov K. V. **AN ANALOGUE OF THE KRONECKER — CAPPELLI THEOREM FOR SYSTEMS OF NONCOMMUTATIVE LINEAR EQUATIONS GENERATING LINEAR LANGUAGES.** The paper continues the study of systems of noncommutative polynomial equations, which are interpreted as grammars of formal languages. Such systems are solved in the form of formal power series (FPS), which express non-terminal symbols in terms of the terminal symbols of the alphabet and are considered as formal languages. Each FPS is associated with its commutative image, which is obtained under the assumption that all symbols denote commutative variables, real or complex. In this paper, we consider equations that are linear in nonterminal symbols with polynomial coefficients in terminal symbols, which means that these systems generate linear formal languages. As is known, the compatibility of a system of noncommutative polynomial equations is not directly related to the compatibility of its commutative image, and therefore, as an analogue of the Kronecker — Cappelli theorem, it is only possible to obtain a sufficient condition for the inconsistency of a noncommutative system.
Keywords: *systems of linear equations, noncommutative variables, formal power series, commutative image.*

Zharkova A. V., Musugalieva A. G. **ABOUT ALGORITHMS FOR SEARCHING COMPUTER INFORMATION.** Due to the large growth in the volume of data, there are many problems with information processing. If the search for information is a key

task of the system, for example, the search for malware by antivirus programs, you need to know the principles of its organization. In this paper, we study the algorithms for searching a substring in a string: the naive search algorithm, the Knuth — Morris — Pratt algorithm, the Boyer — Moore algorithm, the Rabin — Karp algorithm, as well as wildcards applicable to them (substitution characters, "matching" with any character or their sequence). As a result, a C# program has been developed and implemented to search for files by various parameters (file name, extension, size and content) using the above algorithms and methods. The program allows you to scan a given directory to search for malware. Computational experiments were carried out, including changing the maximum number of characters of the sample and text, the corresponding conclusions were drawn. The overall best file search time (it is enough to find the first occurrence) turned out to be using the Boyer — Moore algorithm, the worst — using the Rabin — Karp algorithm. To search for files for small given data and parameters, you can use naive search, for medium and large data and parameters for small samples it is better to use the Knuth — Morris — Pratt algorithm, for large ones — Boyer — Moore algorithm.

Keywords: *Boyer — Moore algorithm, cybersecurity, file search, Knuth — Morris — Pratt algorithm, Rabin — Karp algorithm, scanning, substring search in a string.*

Kuznetsov A. A., Kuznetsova A. S. **ON ONE REPRESENTATION OF ELEMENTS OF FINITE 2-GROUPS IN THE FORM OF BOOLEAN VECTORS.** In this paper, we propose a way to represent elements of finite 2-groups as Boolean vectors. Let G be some finite (Burnside) 2-group and its order is 2^k. In this case, each element of the group will be represented by a unique Boolean (bit) vector of dimension k. To calculate the product of two elements, we use analogues of Hall polynomials but now instead of multiplication and addition over the field $\mathbb{Z}_2$ we use the equivalent Boolean (bitwise) operations "and", as well as "exclusive or". Note that operations on bits are much faster on a computer than on integer or string data types. For problems requiring the calculation of a large number of products of group elements the method will dramatically reduce the running time of computer programs.

Keywords: *2-group, Boolean vector, Hall polynomials.*

SECTION 5

Glukhov M. M., Pankov K. N. **ON A CLASS OF ALGEBRAIC GEOMETRIC CODES.** In the paper, we present a family of algebraic geometric codes over GF(256) that lie above the Gilbert — Varshamov bound together with dual codes. The family of these codes can be used to construct a post-quantum algorithm of the classic McEllice type. The following theorem holds for them: Let $E : y^3 = (x^{63} - 1)/(x^3 - 1)$ be a curve over the field $F = \mathrm{GF}(256)$, $P_1, \ldots, P_{720}$ are arbitrary distinct F-rational points of this curve, P_∞ is the point at infinity. Then the algebraic geometric codes $C_r(D, G)$ on the curve defined by the divisors $D = P_1 + \ldots + P_{720}$ and $G = rP_\infty$ for all integers r, $81 \leqslant r \leqslant 197$, are $[720, 3r - 57, 720 - 3r]_{2^8}$-codes, and their cardinality, as well as the cardinality of their dual $[720, 777 - 3r, 3r - 114]_{2^8}$-codes, satisfies the Gilbert — Varshamov bound.

Keywords: *post-quantum cryptography, error-correcting codes, algebraic geometric codes, Gilbert — Varshamov bound.*

Kolesnikov S. G., Leontiev V. M. **A SERIES OF SHORT EXACT FORMULAS FOR THE BHATTACHARYA PARAMETER OF COORDINATE CHANNELS.** Let W be a symmetric channel with binary input alphabet and a finite output alphabet. In 2007,

E. Arikan discovered the phenomenon of channel polarization, which makes it possible to single out those synthetic channels $W_N^{(i)}$, constructed by W, through which it is preferable to transmit information bits. One of the tools for splitting channels into "bad" and "good" is the Bhattacharya parameter $Z(W_N^{(i)})$. However, the calculation of $Z(W_N^{(i)})$ is difficult, since it requires about 2^{2N} addition operations, where N is the code length. In 2013, I. Tal and A. Vardy proposed a method for estimating from above and below the error probabilities in the channels $W_N^{(i)}$, $1 \leqslant i \leqslant N$, which has a complexity $O(N\mu^2 \log \mu)$, where $\mu > \mu_0$ and the number μ_0 does not depend on N. However, the number μ can be quite great and depends, in particular, on the required precision. Previously, in the case where W is a memoryless binary symmetric channel, the authors constructed two series of exact formulas for the Bhattacharya parameters, which still require an exponential but much less number of operations than in the formulas from Arikan's original paper. In the present paper, for every $N = 2^n$, we construct a series of $n(n-1)/2$ exact formulas that do not contain summation over variables.

Keywords: *polar code, binary memoryless symmetric channel, Bhattacharya parameter.*

Malygina E. S., Kuninets A. A. **CALCULATION OF ERROR-CORRECTING PAIRS FOR AN ALGEBRAIC-GEOMETRIC CODE.** Error-correcting pairs are calculated explicitly for an arbitrary algebraic-geometric code and its dual code. Such a pair consists of codes that are necessary for an effective decoding algorithm for a given code. The type of pairs depends on the degrees of divisors with which both the original code and one of the codes from error-correcting pair are constructed. So for the algebraic-geometric code $\mathcal{C}_{\mathscr{L}}(D, G)$ of the length n associated with a functional field $F/\mathbb{F}_q$ of genus g the error-correcting pair with number of errors $t = (n - \deg(G) - g - 1)/2$ is $(\mathcal{C}_{\mathscr{L}}(D, F), \mathcal{C}_{\mathscr{L}}(D, G+F)^{\perp})$ or $(\mathcal{C}_{\mathscr{L}}(D, F)^{\perp}, \mathcal{C}_{\mathscr{L}}(D, F-G))$. For the dual code $\mathcal{C}_{\mathscr{L}}(D, G)^{\perp}$ the error-correcting pair with number of errors $t = (\deg(G) - 3g + 1)/2$ is $\mathcal{C}_{\mathscr{L}}(D, F), \mathcal{C}_{\mathscr{L}}(D, G - F))$. Considering each component of pair as MDS-code we obtain additional conditions on degrees of divisors G and F. In addition, error-correcting pairs are calculated for subfield subcodes $\mathcal{C}_{\mathscr{L}}(D, G)|_{\mathbb{F}_p}$ and $\mathcal{C}_{\mathscr{L}}(D, G)^{\perp}|_{\mathbb{F}_p}$ where $\mathbb{F}_p$ is a subfield of $\mathbb{F}_q$. The form of a first component in the pair depends on degrees of divisors G and F and in some cases on genus g.

Keywords: *function field, algebraic-geometric code, error-correcting pair, subfield subcodes.*

Obukhov P. K., Pankratova I. A. **PERIODIC PROPERTIES OF A FINITE AUTOMATON GENERATOR.** The periodic properties of a two-stage finite automaton generator $G = A_1 \cdot A_2$ are studied, where $A_1 = (\mathbb{F}_2^n, \mathbb{F}_2, g_1, f_1)$ (it is autonomous), $g_1 : \mathbb{F}_2^n \to \mathbb{F}_2^n$, $f_1 : \mathbb{F}_2^n \to \mathbb{F}_2$, $A_2 = (\mathbb{F}_2, \mathbb{F}_2^m, \mathbb{F}_2, g_2, f_2)$, $g_2 : \mathbb{F}_2 \times \mathbb{F}_2^m \to \mathbb{F}_2^m$, $f_2 : \mathbb{F}_2 \times \mathbb{F}_2^m \to \mathbb{F}_2$, $n, m \geqslant 1$. It is obtained that the maximum value of the generator period is 2^{n+m}. Some necessary conditions for its achievement are formulated, namely: 1) the function g_1 is a full cycle substitution; 2) changing the initial state $x(1)$ or $y(1)$ does not affect the period of the generator; 3) function f_1 is not a constant; 4) at least one of the subfunctions $f_2(0, \cdot)$ and $f_2(1, \cdot)$ is not a constant; 5) the subfunctions $g_2(0, \cdot)$ and $g_2(1, \cdot)$ of the transition function g_2 are substitutions; 6) $y(2^n i + j) \neq y(2^n k + j)$ for all $i, k \in \{0, \ldots, 2^m - 1\}$, $i \neq k$, $j = 1, \ldots, 2^n$.

Keywords: *finite automaton generator, substitutions, periodic sequences.*