МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ПРИКЛАДНАЯ ДИСКРЕТНАЯ МАТЕМАТИКА

Научный журнал

2024 № 66

Зарегистрирован в Федеральной службе по надзору
в сфере связи и массовых коммуникаций

Свидетельство о регистрации ПИ № ФС 77-33762 от 16 октября 2008 г.

Подписной индекс в объединённом каталоге «Пресса России» 38696

УЧРЕДИТЕЛЬ
Томский государственный университет

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

В журнале публикуются результаты фундаментальных и прикладных научных исследований отечественных и зарубежных ученых, включая студентов и аспирантов, в области дискретной математики и её приложений в криптографии, компьютерной безопасности, кибернетике, информатике, программировании, теории надёжности, интеллектуальных системах.

Периодичность выхода журнала: 4 номера в год.

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

Подписано к печати 04.12.2024. Формат $60 \times 84\frac{1}{8}$. Усл. п. л. 14,3. Тираж 300 экз.
Заказ № 6142. Цена свободная. Дата выхода в свет 12.12.2024.

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

СОДЕРЖАНИЕ

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПРИКЛАДНОЙ ДИСКРЕТНОЙ МАТЕМАТИКИ

МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

ПРИКЛАДНАЯ ТЕОРИЯ АВТОМАТОВ

ПРИКЛАДНАЯ ТЕОРИЯ ГРАФОВ

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ И ПРОГРАММИРОВАНИЯ

CONTENTS

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПРИКЛАДНОЙ ДИСКРЕТНОЙ МАТЕМАТИКИ

УДК 512+512.5+512.54+512.54.03 DOI 10.17223/20710410/66/1

ОБ УРАВНЕНИЯХ В СВОБОДНЫХ ГРУППАХ С КОММУТАНТНЫМИ ОГРАНИЧЕНИЯМИ НА РЕШЕНИЯ

А. И. Зеткина

Ярославский государственный университет, г. Ярославль, Россия

E-mail: a.zetkina1@uniyar.ac.ru

Описан полиномиальный алгоритм, позволяющий по произвольному разрешенному относительно неизвестных уравнению вида $w(x_1, \ldots, x_n) = [a, b]$, где $w(x_1, \ldots, x_n)$ — групповое слово в алфавите неизвестных, а $[a, b]$ — коммутатор свободных образующих a и b свободной группы F_2, определить, существует ли решение этого уравнения, удовлетворяющее условию $x_1, \ldots, x_n \in F_2^{(1)}$, где $F_2^{(1)}$ — коммутант группы F_2. Установлено существование полиномиального алгоритма, позволяющего по произвольному разрешенному относительно неизвестных уравнению вида $w(x_1, \ldots, x_n) = g(a, b)$, где $g(a, b)$ — элемент длины меньше 4 свободной группы F_2, определить, существует ли решение этого уравнения, удовлетворяющее условию $x_1, \ldots, x_t \in F_2^{(1)}$, где t — произвольное фиксированное число между 1 и n. Доказана алгоритмическая разрешимость аналогичной проблемы для уравнений $w(x_1, a, b) = 1$ с одной переменной x_1.

Ключевые слова: *свободная группа, уравнение в свободной группе.*

ON EQUATIONS IN FREE GROUPS WITH COMMUTANT RESTRICTIONS ON SOLUTIONS

A. I. Zetkina

Yaroslavl State University, Yaroslavl, Russia

A polynomial algorithm has been constructed that allows, given an arbitrary equation of the form $w(x_1, \ldots, x_n) = [a, b]$, resolved with respect to unknowns, where $w(x_1, \ldots, x_n)$ is a group word in the alphabet of unknowns and $[a, b]$ is the commutator of free generators a and b of the free group F_2, to determine whether there is a solution to this equation that satisfies the condition $x_1 \ldots, x_n \in F_2^{(1)}$, where $F_2^{(1)}$ is the commutator of group F_2. The existence of a polynomial algorithm has been established that allows, given an arbitrary equation of the form $w(x_1, \ldots, x_n) = g(a, b)$, where $g(a, b)$ is an element of length less than 4 of the free group F_2, to determine whether a solution to this equation exists, that satisfies the condition $x_1, \ldots, x_t \in F_2^{(1)}$, where t is an arbitrary fixed number between 1 and n. The algorithmic solvability of a similar problem has been proven for the equations $w(x_1, a, b) = 1$ with one variable x_1.

Keywords: *free group, equation in a free group.*

Введение

Через F_m будем обозначать свободную группу ранга m со свободными образующими $a_1, \ldots, a_m$. При $m = 2$ вместо a_1 и a_2 будем писать a и b соответственно.

Определим некоторые понятия, относящиеся к системам уравнений в свободных группах.

Определение 1. Системой уравнений с неизвестными $x_1, \ldots, x_n$ в свободной группе F_m называется выражение вида

$$\mathop{\&}_{i=1}^{k}\Big(w_i(x_1, \ldots, x_n, a_1, \ldots, a_m) = u_i(x_1, \ldots, x_n, a_1, \ldots, a_m)\Big), \tag{1}$$

где $w_i(x_1, \ldots, x_n, a_1, \ldots, a_m)$ и $u_i(x_1, \ldots, x_n, a_1, \ldots, a_m)$ — слова в алфавите

$$\{x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}, a_1, a_1^{-1}, \ldots, a_m, a_m^{-1}\}.$$

Определение 2. Набор $\langle g_1, \ldots, g_n\rangle$ элементов группы F_m называется решением системы (1), если при любом $i = 1, \ldots, k$ в группе F_m выполняется равенство

$$w_i(g_1, \ldots, g_n, a_1, \ldots, a_m) = u_i(g_1, \ldots, g_n, a_1, \ldots, a_m).$$

Определение 3. Две системы уравнений с одними и теми же неизвестными называются эквивалентными, если множества их решений совпадают.

Используя уравнение

$$[x, a_1] = ([x, a_2]\, y^2)^2,$$

имеющее в свободной группе F_m при любом $m \geqslant 2$ лишь тривиальное решение $x{=}y{=}1$, любую систему уравнений (1) можно заменить одним равносильным уравнением. Построение по (1) равносильного уравнения ведётся индукцией по k. При $k = 2$ система уравнений

$$\mathop{\&}_{i=1}^{2}\Big(w_i(x_1, \ldots, x_n, a_1, \ldots, a_m) = u_i(x_1, \ldots, x_n, a_1, \ldots, a_m)\Big)$$

равносильна одному уравнению

$$\begin{gathered}[w_1(x_1, \ldots, x_n, a_1, \ldots, a_m)u_1^{-1}(x_1, \ldots, x_n, a_1, \ldots, a_m), a_1] = \\ = ([w_1(x_1, \ldots, x_n, a_1, \ldots, a_m)u_1^{-1}(x_1, \ldots, x_n, a_1, \ldots, a_m), a_2] \\ (w_2(x_1, \ldots, x_n, a_1, \ldots, a_m)u_2^{-1}(x_1, \ldots, x_n, a_1, \ldots, a_m))^2)^2.\end{gathered}$$

Для уравнений в свободных группах традиционно рассматриваются две основные задачи: проблема существования решения и проблема описания множества всех решений.

Исследование разрешимости уравнений в свободных группах было начато американскими математиками в конце 50-х годов в связи с проблемой разрешимости элементарных теорий свободных групп, поставленной А. Тарским [1], остававшейся открытой почти полвека и решённой в начале 2000-х годов О. Харлампович и А. Мясниковым [2].

Сначала исследовались лишь отдельные уравнения, а в 1960 г. Р. Линдон [3] нашёл для произвольного уравнения с одним неизвестным описание множества всех его решений с помощью параметрических слов, т. е. выражений, полученных из образующих рассматриваемой свободной группы с помощью операций группового умножения и возведения в степень с переменным целочисленным показателем. Позже А. А. Лоренц [4] и К. И. Аппель [5] уточнили это описание, доказав, что общее решение любого уравнения с одним неизвестным в свободной группе представимо конечным числом формул вида

AB^tC, где A, B, C — конкретные слова, а t — параметр, принимающий произвольные целочисленные значения. Дальнейшее продвижение в этом вопросе достигнуто в 1970 г. Ю. И. Хмелевским [6].

В 1982 г. Г. С. Маканин [7] получил полное решение проблемы распознавания разрешимости уравнений в свободной группе. Он доказал, что если уравнение с длиной записи d имеет решение в свободной группе, то длина каждой компоненты минимального (по максимальной длине компоненты) решения не превосходит числа $\Phi(d)$, где $\Phi(x)$ — некоторая рекурсивная функция. Это даёт переборный алгоритм для распознавания разрешимости произвольного уравнения в свободной группе.

Вскоре после опубликования работы [7] удалось на том же пути доказать разрешимость экзистенциональной (универсальной) и позитивной теорий любой свободной группы [8]. При доказательстве разрешимости позитивной теории свободной группы Г. С. Маканин использовал результат Ю. И. Мерзлякова [9] об устранимости кванторов общности в позитивных формулах, относящихся к свободным группам.

А. А. Разборов [10] дал описание множества решений произвольной совместной системы уравнений в свободной группе.

1. Уравнения с ограничениями на решения

После построения Г. С. Маканиным [7] алгоритма, позволяющего по произвольной системе уравнений в свободной группе F_m определить, имеет ли она решение, особый интерес стал представлять вопрос о существовании аналогичных алгоритмов для уравнений в свободных группах с различными «не слишком сложными» ограничениями на решения.

Вопрос о разрешимости позитивной теории свободной группы был сведён Ю. И. Мерзляковым [9] к следующей проблеме: существует ли алгоритм, позволяющий для произвольного уравнения

$$w(x_1, \ldots, x_n, a_1, \ldots, a_m) = 1$$

в свободной группе счётного ранга, где n и m — произвольные натуральные числа, определить, имеет ли оно такое решение $g_1, \ldots, g_n$, что

$$g_1 \in F_{m_1},\ g_2 \in F_{m_2},\ \ldots,\ g_n \in F_{m_n},$$

где $m_1 \leqslant m_2 \leqslant \ldots \leqslant m_n$; F_{m_i} — свободная группа с образующими $a_1, \ldots, a_{m_i}$.

Г. С. Маканин в [8] построил искомый алгоритм и тем самым доказал разрешимость позитивной теории свободной группы.

Известно, что вопрос о точности матричного представления Гасснер [11, 12] группы крашеных кос эквивалентен вопросу об отсутствии нетривиального решения в свободной группе F_m уравнения

$$x_1 a_1 x_1^{-1} \cdot x_2 a_2 x_2^{-1} \ldots x_m a_m x_m^{-1} = a_1 \cdot a_2 \ldots a_m,$$

удовлетворяющего условию $x_1 \in F_m^{(2)}$, $\ldots$, $x_n \in F_m^{(2)}$, где $F_m^{(2)}$ — второй коммутант свободной группы F_m. Напомним, что для произвольной группы G через $G^{(2)}$ обозначается её второй коммутант, т. е. $G^{(2)} = [G^{(1)}, G^{(1)}]$, где $G^{(1)} = [G, G]$ — коммутант группы G.

Обобщая эти ситуации, Г. С. Маканин поставил в «Коуровской тетради» [13] следующую проблему для уравнений в свободных группах:

9.25. Указать алгоритм, который по уравнению

$$w(x_1, \ldots, x_n, a_1, \ldots, a_m) = 1$$

в свободной группе F_m и списку конечно порождённых подгрупп $H_1, \ldots, H_n$ группы F_m позволял бы узнать, существует ли решение этого уравнения с условием $x_1 \in H_1, \ldots, x_n \in H_n$.

Первые положительные результаты в решении этой проблемы были получены А. Ш. Малхасяном [14].

В. Диекерт [15] показал, что проблема определения по произвольному уравнению

$$w(x_1, \ldots, x_n, a_1, \ldots, a_m) = 1$$

в свободной группе F_n и списку *регулярных подмножеств* (языков) $H_1, \ldots, H_n$ группы F_m, существует ли решение этого уравнения с условием $x_1 \in H_1, \ldots, x_n \in H_n$, разрешима и принадлежит классу PSPACE. Так как конечно порождённые подгруппы являются регулярными подмножествами, тем самым решается и проблема Г. С. Маканина.

Представляет интерес дальнейшее исследование различных обобщений проблемы Г. С. Маканина для свободных групп, получающихся путем ослабления ограничений, налагаемых на подгруппы $H_1, \ldots, H_n$.

Одна из причин, по которым в формулировке задачи 9.25 речь идёт именно о конечно порождённых подгруппах, заключается в том, что для конечно порождённых подгрупп свободной группы разрешима проблема вхождения.

В то же время проблема вхождения разрешима и для многих бесконечно порождённых подгрупп свободной группы, причём, например, для первого $F_m^{(1)}$ и второго $F_m^{(2)}$ коммутантов свободной группы F_m проблема вхождения решается значительно проще, чем для некоторых конечно порождённых подгрупп. Поэтому представляется достаточно естественным следующее обобщение задачи 9.25:

9.25а. Существует ли алгоритм, который по уравнению

$$w(x_1, \ldots, x_n, a_1, \ldots, a_m) = 1$$

в свободной группе F_m и списку подгрупп $H_1, \ldots, H_n$ с разрешимыми проблемами вхождения позволял бы узнать, существует ли решение этого уравнения с условием $x_1 \in H_1, \ldots, x_n \in H_n$?

2. Уравнения, разрешенные относительно неизвестных

В ряде работ, например в [3, 6, 16–19], рассматриваются уравнения вида

$$w(x_1, \ldots, x_n) = g,$$

где $w(x_1, \ldots, x_n)$ — групповое слово в алфавите неизвестных, а g — элемент свободной группы F_m. Такие уравнения получили название *уравнений, разрешенных относительно неизвестных, уравнений с правой частью или однокоэффициентных уравнений*. Проблема разрешимости таких уравнений получила название *проблемы подстановки* или *проблемы эндоморфной сводимости* [17, 18].

В [19] получен следующий результат:

Теорема 1 [19]. В свободной группе F_2 со свободными образующими a и b можно построить такое разрешенное относительно неизвестных уравнение

$$w(x, x_1, \ldots, x_n) = [a, b]$$

с неизвестными $x_1, x_2, \ldots, x_n$ и параметром x, что невозможно создать алгоритм, позволяющий для произвольного натурального числа k определить, существует ли решение уравнения

$$w(a^k, x_1, \ldots, x_n) = [a, b],$$

удовлетворяющее условию $x_1, \ldots, x_t \in F_2^{(1)}$, где t — некоторое фиксированное число между 1 и n; $[a,b] = a^{-1}b^{-1}ab$ — коммутатор элементов a и b.

Естественно возникает вопрос о том, каким может быть t.

С. И. Адян предложил исследовать прежде всего предельные случаи: $t = 1$ и $t = n$.

Теорема 2. Существует полиномиальный алгоритм, позволяющий для произвольного разрешенного относительно неизвестных уравнения в свободной группе F_2 со свободными образующими a и b

$$w(x_1, \ldots, x_n) = [a,b] \tag{2}$$

с неизвестными $x_1, x_2, \ldots, x_n$ определить, существует ли решение этого уравнения, удовлетворяющее условию $x_1, \ldots, x_n \in F_2^{(1)}$.

Доказательство. Известно [20], что коммутант $F_2^{(1)} = [F_2, F_2]$ свободной группы F_2 свободно порождается нетривиальными коммутаторами

$$c_{i,j} = [a^i, b^j],$$

где i и j — произвольные отличные от нуля целые числа. Поэтому вопрос о разрешимости в свободной группе F_2 уравнения (2) с указанными ограничениями на решения сводится к вопросу о разрешимости в свободной бесконечно порождённой группе $F_2^{(1)}$ уравнения

$$w(x_1, \ldots, x_n) = c_{1,1}. \tag{3}$$

Если $x_1^0, \ldots, x_n^0$ — решение уравнения (3), то, заменив в выражении элементов $x_1^0, \ldots, x_n^0$ через свободные образующие $c_{i,j}$ все $c_{i,j}$, кроме $c_{1,1}$, на единицу, получим новое решение

$$c_{1,1}^{\sigma_{c_{1,1}}(x_1^0)}, \ldots, c_{1,1}^{\sigma_{c_{1,1}}(x_n^0)}$$

этого уравнения, где через $\sigma_{c_{1,1}}(x_i^0)$ обозначается сумма показателей степени свободной образующей $c_{1,1}$ в выражении элемента x_i^0 через свободные образующие $c_{i,j}$. Значит, целые числа $\sigma_{c_{1,1}}(x_1^0)$, $\ldots$, $\sigma_{c_{1,1}}(x_n^0)$ являются решением уравнения

$$\sigma_{(x_1)}(w)y_1 + \ldots + \sigma_{(x_n)}(w)y_n = 1. \tag{4}$$

Верно и обратное.

Таким образом, вопрос о разрешимости в свободной группе F_2 уравнения (2) с указанными ограничениями на решения равносилен вопросу о разрешимости в целых числах линейного уравнения (4). Последний вопрос решается полиномиальным алгоритмом. ■

Слово $[a,b]$, стоящее в правой части уравнения из теоремы 1, имеет длину 4. Как показывает следующая теорема, это наименьшая возможная длина.

Теорема 3. Существует полиномиальный алгоритм, позволяющий по произвольному разрешенному относительно неизвестных уравнению вида

$$w(x_1, \ldots, x_n) = g(a,b),$$

где $w(x_1, \ldots, x_n)$ — групповое слово в алфавите неизвестных $\{x_1, x_2, \ldots, x_n\}$, $g(a,b)$ — элемент длины меньше 4 свободной группы F_2 со свободными образующими a и b, определить, существует ли решение этого уравнения, удовлетворяющее условию

$$x_1, \ldots, x_t \in F_2^{(1)}, \tag{5}$$

где t — произвольное фиксированное число между 1 и n.

Доказательство. Если g — групповое слово длины меньше 4 в алфавите $\{a, b\}$ свободных образующих группы F_2, то нетрудно убедиться, что g — степень A^k некоторого примитивного элемента A группы F_2.

Покажем, что уравнение $w(x_1, \ldots, x_n) = g$, т. е. уравнение

$$w(x_1, \ldots, x_n) = A^k, \tag{6}$$

имеет решение в группе F_2, удовлетворяющее условию (5), тогда и только тогда, когда в циклической группе F_1 с образующим элементом a разрешимо уравнение

$$w(1, \ldots, 1, x_{t+1}, \ldots, x_n) = a^k. \tag{7}$$

Вопрос о разрешимости уравнения (7) сводится к вопросу о разрешимости в целых числах линейного уравнения с целыми коэффициентами, который полиномиально разрешим.

Предположим, что A и B — система свободных образующих группы F_2 (A — примитивный элемент этой группы), а φ и ψ — такие автоморфизмы этой группы, что

$$\varphi(A) = a, \quad \varphi(B) = b; \quad \psi(a) = A, \quad \psi(b) = B.$$

Пусть $g_1, \ldots, g_n$ — решение в группе F_2 уравнения (6), удовлетворяющее условию $g_1, \ldots, g_t \in F_2^{(1)}$. Применив к равенству $w(g_1, \ldots, g_n) = A^k$ автоморфизм φ, получим

$$w(\varphi(g_1), \ldots, \varphi(g_n)) = a^k.$$

К последнему равенству применим гомоморфизм φ_1 группы F_2 на группу F_1, заданный равенствами $\varphi_1(a) = a$, $\varphi_1(b) = 1$, и получим

$$w(\varphi_1(\varphi(g_1)), \ldots, \varphi_1(\varphi(g_n))) = a^k.$$

Если $g \in F_2^{(1)}$, то $\varphi_1(\varphi(g)) = 1$, поэтому $\varphi_1(\varphi(g_1)) = 1$, $\ldots$, $\varphi_1(\varphi(g_t)) = 1$. Значит, $\varphi_1(\varphi(g_{t+1})), \ldots, \varphi_1(\varphi(g_n))$ — решение уравнения (7) в группе F_1.

Обратно, если $h_{t+1}, \ldots, h_n$ — решение уравнения (7) в группе F_1, то, применив к равенству

$$w(1, \ldots, 1, h_{t+1}, \ldots, h_n) = a^k$$

в группе F_2 автоморфизм ψ этой группы, получим

$$w(1, \ldots, 1, \psi(h_{t+1}), \ldots, \psi(h_n)) = A^k,$$

значит, $g_1 = 1$, $\ldots$, $g_t = 1$, $g_{t+1} = \psi(h_{t+1})$, $\ldots$, $g_n = \psi(h_n)$ — решение в группе F_2 уравнения (6), удовлетворяющее условию $g_1, \ldots, g_t \in F_2^{(1)}$. ■

Рассмотрим аналогичный вопрос для уравнений с одним неизвестным. Как и ранее, через $F_n^{(1)}$ обозначаем коммутант свободной группы F_n.

Теорема 4. Существует полиномиальный алгоритм, позволяющий по любому уравнению с одним неизвестным

$$w(x_1, a_1, \ldots, a_n) = 1 \tag{8}$$

в свободной группе F_n определить, имеет ли оно такое решение x_1, что $x_1 \in F_n^{(1)}$.

Доказательство. А. А. Лоренц [4] и К. И. Аппель [5] доказали, что множество решений уравнения с одним неизвестным задаётся конечным множеством параметрических слов, т. е. слов вида $A\,B^{\lambda}\,C$, где λ — целочисленный параметр. Д. Бормотов, Р. Гилман и А. Мясников [21] разработали полиномиальный алгоритм построения по уравнению с одним неизвестным соответствующего множества параметрических слов.

Тем самым вопрос о существовании решения уравнения (8) в свободной группе F_n с условием $x_1 \in F_n^{(1)}$ полиномиально сводится к определению, существует ли такое целое число λ, что $A\,B^{\lambda}\,C \in F_n^{(1)}$, или $B^{\lambda} = A^{-1}C^{-1}$ в $F_n/F_n^{(1)}$, т. е. задача о существовании у уравнения (8) решения с условием $x_1 \in F_n^{(1)}$ сводится к *проблеме степеней* для группы $F_n/F_n^{(1)}$: существует ли такое целое число λ, что $B^{\lambda} = A^{-1}\,C^{-1}$ в $F_n/F_n^{(1)}$. Для завершения доказательства достаточно заметить, что проблема степеней для групп $F_n/F_n^{(1)}$ полиномиально разрешима. ■

Заключение

Полученные в работе результаты об *алгоритмической* ***разрешимости*** проблемы совместности для некоторых типов уравнений в свободных группах с ограничениями на решения вместе с результатами работы [19] об *алгоритмической* ***неразрешимости*** проблемы совместности для аналогичных уравнений с ограничениями на решения близкого типа могут рассматриваться как некоторый вклад в реализацию сформулированной в 60-е годы XX в. выдающимся отечественным математиком Сергеем Ивановичем Адяном «Программы уточнения границы между алгоритмически разрешимыми и алгоритмически неразрешимыми проблемами».

Выражаю благодарность рецензентам, сделавшим ряд полезных замечаний по оформлению работы.

ЛИТЕРАТУРА

1. *Tarski A., Mostowski A., and Robinson R. M.* Undecidable Theories. Amsterdam: North-Holland Publ. Company, 1953. XI+98 p.
2. *Kharlampovich O. and Myasnikov A.* Elementary theory of free non-abelian groups // J. Algebra. 2006. V. 302. P. 451–552.
3. *Lyndon R. C.* Equations in free groups // Trans. Amer. Math. Soc. 1960. V. 96. P. 445–457.
4. *Лоренц А. А.* О представлении множеств решений систем уравнений с одним неизвестным в свободных группах // Докл. АН СССР. 1968. Т. 178. № 2. С. 290–292.
5. *Appel K. I.* One-variable equations in free groups // Proc. Amer. Math. Soc. 1968. V. 19. P. 912–918.
6. *Хмелевский Ю. И.* Системы уравнений в свободной группе. I, II // Изв. АН СССР. Сер. матем. 1971. Т. 35. № 6. С. 1237–1268; 1972. Т. 36. № 1. С. 110–179.
7. *Маканин Г. С.* Уравнения в свободной группе // Изв. АН СССР. Сер. матем. 1982. Т. 46. № 6. С. 1199–1273.
8. *Маканин Г. С.* Разрешимость универсальной и позитивной теорий свободной группы // Изв. АН СССР. Сер. матем. 1984. Т. 48. № 4. С. 735–749.
9. *Мерзляков Ю. И.* Позитивные формулы на свободных группах // Алгебра и логика. 1966. Т. 5. Вып. 4. С. 25–42.
10. *Разборов А. А.* О системах уравнений в свободной группе // Изв. АН СССР. Сер. матем. 1984. Т. 48. № 4. С. 779–832.
11. *Gassner B. J.* On braid groups // Abh. Math. Sem. Univ. Hamburg. 1961. V. 25. P. 10–22.

12. *Birman J. S.* Braids, Links and Mapping Class Groups. AM-82. V. 82. Princeton: Princeton University Press, 1974.
13. Коуровская тетрадь: нерешенные вопросы теории групп. 11 изд., доп. Новосибирск: ИМ СО РАН, 1990.
14. *Малхасян А. Ш.* О разрешимости в подгруппах уравнений в свободной группе // Прикладная математика. 1986. Вып. 2. С. 42–47.
15. *Diekert V.* Makanin's Algorithm for Solving Word Equations with Regular Constraints. University of Stuttgart, Faculty of Computer Science. Technical Report No. 1998/02. 43 p.
16. *Мальцев А. И.* Об уравнении $zxyx^{-1}y^{-1}z^{-1} = aba^{-1}b^{-1}$ в свободной группе // Алгебра и логика. 1962. Т. 1. № 5. С. 45–50.
17. *Schupp P. E.* On the substitution problem for free groups // Proc. Amer. Math. Soc. 1969. V. 23. P. 421–423.
18. *Edmunds C. C.* On the endomorphisms problem for free group // Com. Algebra. 1975. No. 3. P. 7–20.
19. *Дурнев В. Г.* К проблеме разрешимости уравнений с одним коэффициентом // Матем. заметки. 1996. Т. 59. № 6. С. 832–845.
20. *Magnus W., Karrass A., and Solitar D.* Combinatorial Group Theory. N.Y.: Interscience Publ., 1966.
21. *Bormotov D., Gilman R., and Myasnikov A.* Solving one-variable equation in free groups // J. Group Theory. 2009. V. 12. No. 2. P. 317–330.

REFERENCES

1. *Tarski A., Mostowski A., and Robinson R. M.* Undecidable Theories. Amsterdam, North-Holland Publ. Company, 1953, XI+98 p.
2. *Kharlampovich O. and Myasnikov A.* Elementary theory of free non-abelian groups. J. Algebra, 2006, vol. 302, pp. 451–552.
3. *Lyndon R. C.* Equations in free groups. Trans. Amer. Math. Soc., 1960, vol. 96, pp. 445–457.
4. *Lorents A. A.* O predstavlenii mnozhestv resheniy sistem uravneniy s odnim neizvestnym v svobodnykh gruppakh [On the representation of solution sets of systems of equations with one unknown in free groups]. Dokl. AN SSSR, 1968, vol. 178, no. 2, pp. 290–292. (in Russian)
5. *Appel K. I.* One-variable equations in free groups. Proc. Amer. Math. Soc., 1968, vol. 19, pp. 912–918.
6. *Khmelevskiy Yu. I.* Systems of equations in a free group. I, II. Math. USSR-Izv., 1971, vol. 5, no. 6, pp. 1245–1276; 1972, vol. 6, no. 1, pp. 109–180.
7. *Makanin G. S.* Equations in a free group. Math. USSR-Izv., 1983, vol. 21, no. 3, pp. 483–546.
8. *Makanin G. S.* Decidability of the universal and positive theories of a free group. Math. USSR-Izv., 1985, vol. 25, no. 1, pp. 75–88.
9. *Merzlyakov Yu. I.* Pozitivnye formuly na svobodnykh gruppakh [Positive formulas on free groups]. Algebra i Logika, 1966, vol. 5, iss. 4, pp. 25–42. (in Russian)
10. *Razborov A. A.* On systems of equations in a free group., Math. USSR-Izv., 1985, vol. 25, no. 1, pp. 115–162.
11. *Gassner B. J.* On braid groups. Abh. Math. Sem. Univ. Hamburg, 1961, vol. 25, pp. 10–22.
12. *Birman J. S.* Braids, Links and Mapping Class Groups. AM-82, vol. 82, Princeton, Princeton University Press, 1974.
13. Kourovskaya tetrad': nereshennye voprosy teorii grupp [Kurovskaya Notebook: Unsolved Questions in Group Theory]. 11th ed. Novosibirsk: IM SB RAS, 1990. (in Russian)

14. *Malkhasyan A. Sh.* O razreshimosti v podgruppakh uravneniy v svobodnoy gruppe [On the solvability in subgroups of equations in a free group]. Prikladnaya Matematika, 1986, iss. 2, pp. 42–47. (in Russian)
15. *Diekert V.* Makanin's Algorithm for Solving Word Equations with Regular Constraints. University of Stuttgart, Faculty of Computer Science, Technical Report No. 1998/02, 43 p.
16. *Maltsev A. I.* Ob uravnenii $zxyx^{-1}y^{-1}z^{-1} = aba^{-1}b^{-1}$ v svobodnoy gruppe [About the equation $zxyx^{-1}y^{-1}z^{-1} = aba^{-1}b^{-1}$ in a free group]. Algebra i Logika, 1962, vol. 1, no. 5, pp. 45–50. (in Russian)
17. *Schupp P. E.* On the substitution problem for free groups. Proc. Amer. Math. Soc., 1969, vol. 23, pp. 421–423.
18. *Edmunds C. C.* On the endomorphisms problem for free group. Com. Algebra, 1975, no. 3, pp. 7–20.
19. *Durnev V. G.* On the solvability problem for equations with a single coefficient. Math. Notes, 1996, vol. 59, no. 6, pp. 601–610.
20. *Magnus W., Karrass A., and Solitar D.* Combinatorial Group Theory. N.Y., Interscience Publ., 1966.
21. *Bormotov D., Gilman R., and Myasnikov A.* Solving one-variable equation in free groups. J. Group Theory, 2009, vol. 12, no. 2, pp. 317–330.

УДК 510.67 DOI 10.17223/20710410/66/2

АКСИОМАТИЗИРУЕМОСТЬ И РАЗРЕШИМОСТЬ УНИВЕРСАЛЬНЫХ ТЕОРИЙ НАСЛЕДСТВЕННЫХ КЛАССОВ МОДЕЛЕЙ КОНЕЧНЫХ И БЕСКОНЕЧНЫХ ЯЗЫКОВ[1]

А. В. Ильев

Институт математики им. С. Л. Соболева СО РАН, г. Омск, Россия

E-mail: artyom_iljev@mail.ru

Изучаются наследственные классы алгебраических систем языка $\mathrm{L} = \mathrm{L}_{\mathrm{fin}} \cup \mathrm{L}_{\infty}$, где $\mathrm{L}_{\mathrm{fin}} = \langle R_1, R_2, \ldots, R_m, = \rangle$ и $\mathrm{L}_{\infty} = \langle R_{m+1}, R_{m+2}, \ldots \rangle$, причём в L_{∞} число предикатов каждой местности конечно, все предикаты упорядочены по возрастанию своих местностей и обладают свойством неповторения элементов. Класс L-систем называется наследственным, если он замкнут относительно подсистем. Доказано, что класс L-систем является наследственным тогда и только тогда, когда он может быть определён в терминах запрещённых подсистем. Класс L-систем называется универсально аксиоматизируемым, если существует такое множество универсальных предложений Z языка L, что этот класс состоит из всех систем, удовлетворяющих множеству Z. Рассмотрены вопросы универсальной аксиоматизируемости наследственных классов L-систем. Показано, что наследственный класс L-систем универсально аксиоматизируем, если и только если он может быть определён в терминах конечных запрещённых подсистем. Доказана разрешимость универсальной теории произвольного аксиоматизируемого наследственного класса L-систем, множество минимальных запрещённых подсистем которого рекурсивно.

Ключевые слова: *алгебраическая система, наследственный класс, универсальная теория, универсальная аксиоматизируемость, разрешимость.*

AXIOMATIZABILITY AND DECIDABILITY OF UNIVERSAL THEORIES OF HEREDITARY CLASSES OF MODELS OF FINITE AND INFINITE LANGUAGES

A. V. Ilev

Sobolev Institute of Mathematics SB RAS, Omsk, Russia

In the paper, hereditary classes of L-structures are studied with language of the form $\mathrm{L} = \mathrm{L}_{\mathrm{fin}} \cup \mathrm{L}_{\infty}$, where $\mathrm{L}_{\mathrm{fin}} = \langle R_1, R_2, \ldots, R_m, = \rangle$ and $\mathrm{L}_{\infty} = \langle R_{m+1}, R_{m+2}, \ldots \rangle$, and also in L_{∞} the number of predicates of each arity is finite, all predicates are ordered in ascending of their arities and satisfy the non-element repetition property. A class of L-structures is called hereditary if it is closed under substructures. It is proved that the class of L-structures is hereditary if and only if it can be defined in terms of forbidden substructures. A class of L-structures is called universally axiomatizable if there is a set Z of universal L-sentences such that the class consists of all structures satisfying Z. The problems of the universal axiomatizability of hereditary classes of L-structures are considered in the paper. It is shown that hereditary class

[1]Работа выполнена в рамках госзадания ИМ СО РАН, проект FWNF-2022-0003.

of L-structures is universally axiomatizable if and only if it can be defined in terms of finite forbidden substructures. It is proved that the universal theory of any axiomatizable hereditary class of L-structures with a recursive set of minimal forbidden substructures is decidable.

Keywords: *structure, hereditary class, universal theory, universal axiomatizability, decidability.*

Введение

Целью работы является обобщение ранее полученных результатов для графов, гиперграфов и матроидов, а также построение алгоритмов для их конструктивного доказательства. Представленные механизмы могут быть полезны в дальнейшем при исследовании как сугубо теоретических задач, например решения систем уравнений над соответствующими алгебраическими системами [1, 2], так и практических задач в тех вопросах, когда вместо перебора конкретных объектов уместно рассмотреть лишь отдельные их ключевые свойства [3, 4].

В настоящее время в теории графов активно используются алгебраические и логические методы, в том числе методы теории моделей. Сформировалось целое направление исследований, которое получило название алгебраической теории графов, и можно также говорить о формировании особого раздела теории графов — логической теории графов [5]. Напомним, что обыкновенный граф рассматривается как алгебраическая система, язык которой состоит из предиката равенства и бинарного предиката смежности вершин, удовлетворяющего аксиомам иррефлексивности и симметричности. Известно, что теория графов неразрешима, так же как и теория конечных графов [6].

Традиционный интерес вызывают вопросы аксиоматизируемости и универсальной аксиоматизируемости различных классов графов [7–9]. Так, в [10] обсуждаются вопросы аксиоматизируемости наследственных классов графов, определённых в терминах запрещённых порождённых подграфов; в [11] — вопросы аксиоматизируемости наследственных классов графов, определённых в терминах любых запрещённых подграфов. В связи с этим естественным образом возникает задача о нахождении критерия аксиоматизируемости наследственных классов произвольных бесконечных алгебраических систем и их определения в терминах запрещённых подсистем, по аналогии с графами. В частности, такая задача актуальна для более сложных объектов, например гиперграфов, аксиоматизируемость хорновых классов которых исследована в [12], и класса матроидов фиксированного ранга, аксиоматизаруемость которого показана в [13]. Ряд общих вопросов аксиоматизируемости универсальных классов рассмотрен в [14], однако предложенные там подходы не содержат конкретной алгоритмической реализации.

Особое место в теории моделей занимает изучение универсальных теорий. С помощью известной процедуры скулемизации можно перейти от любой теории к универсальной теории в расширенном языке [15]. Кроме того, некоторые общие проблемы разрешимости удаётся интерпретировать как проблемы разрешимости универсальных теорий. Повышенный интерес к универсальным теориям вызывает их применение в логическом программировании и теории баз данных [16]. Разрешимость универсальной теории графов и универсальной теории произвольного аксиоматизируемого наследственного класса графов, множество минимальных запрещённых подграфов которого рекурсивно, доказана в [17].

В данной работе методами теории моделей изучаются наследственные классы алгебраических систем языка $\mathrm{L} = \mathrm{L}_{\mathrm{fin}} \cup \mathrm{L}_{\infty}$, где $\mathrm{L}_{\mathrm{fin}} = \langle R_1, R_2, \ldots, R_m, = \rangle$ и

$\mathrm{L}_\infty = \langle R_{m+1}, R_{m+2}, \ldots \rangle$, причём в L_∞ число предикатов каждой местности конечно, все предикаты упорядочены по возрастанию своей местности и обладают свойством неповторения элементов. В п. 1 приведены основные сведения из теории моделей. В п. 2 рассмотрены вопросы аксиоматизируемости наследственных классов L-систем и показано, что всякий наследственный класс L-систем универсально аксиоматизируем тогда и только тогда, когда он может быть определён в терминах конечных запрещённых подсистем. В п. 3 содержится основной результат работы — доказана разрешимость универсальной теории произвольного аксиоматизируемого наследственного класса L-систем, множество минимальных запрещённых подсистем которого рекурсивно.

1. Предварительные сведения

Напомним основные определения теории моделей.

Языком, или *сигнатурой* $\mathrm{L} = \mathrm{R} \cup \mathrm{F} \cup \mathrm{C}$, называется совокупность следующих множеств:

1) множества *предикатных символов* R;
2) множества *функциональных символов* F;
3) множества *константных символов* C,

причём с каждым предикатным символом $R \in \mathrm{R}$ и с каждым функциональным символом $F \in \mathrm{F}$ однозначно связывается натуральное число n_R или n_F — *арность*, или *местность*.

Алгебраическая система языка L, или L-*система*, — это последовательность

$$\mathcal{A} = \langle A;\ R^{\mathcal{A}}, F^{\mathcal{A}}, c^{\mathcal{A}} \rangle,$$

в которой A — непустое множество, называемое *основным множеством*, или *носителем* системы $\mathcal{A}$; каждому предикатному символу $R \in \mathrm{R}$ соответствует n_R-местное отношение $R^{\mathcal{A}} \subseteq A^{n_R}$; каждому функциональному символу $F \in \mathrm{F}$ соответствует n_F-местная функция $F^{\mathcal{A}} : A^{n_F} \to A$; каждому константному символу $c \in \mathrm{C}$ соответствует некоторый элемент $c^{\mathcal{A}} \in A$. В дальнейшем при описании L-систем используем краткую запись $\mathcal{A} = \langle A, \mathrm{L} \rangle$. Алгебраическая система $\mathcal{A}$ называется *моделью*, если в ней отсутствуют функции.

Алгебраические системы $\mathcal{A} = \langle A, \mathrm{L} \rangle$ и $\mathcal{B} = \langle B, \mathrm{L} \rangle$ языка L называются *изоморфными*, если существует изоморфизм $f : A \to B$, сохраняющий их предикаты и функции.

L-система $\mathcal{A} = \langle A, \mathrm{L} \rangle$ называется *подсистемой* L-системы $\mathcal{B} = \langle B, \mathrm{L} \rangle$ (обозначается $\mathcal{A} \subseteq \mathcal{B}$), если:

1) $A \subseteq B$;
2) функции и предикаты в $\mathcal{A}$ являются ограничениями на A соответствующих функций и предикатов в $\mathcal{B}$;
3) множество A замкнуто относительно функций.

Если $\mathcal{A} \subseteq \mathcal{B}$ и $A \subset B$, то $\mathcal{A}$ называется *собственной подсистемой* $\mathcal{B}$.

Формулой языка L называется формула исчисления предикатов первого порядка с равенством, внелогические константы которой содержатся в L. Формулу без свободных переменных называют *предложением*. Истинность предложения φ в алгебраической системе $\mathcal{A}$ обозначается через $\mathcal{A} \models \varphi$. Предложение φ называется *универсальным предложением*, или $\forall$-*предложением*, если $\varphi = \forall x_1 \ldots \forall x_n\ \psi$, где ψ — бесквантворная формула, не содержащая других переменных, кроме $x_1, \ldots, x_n$. Предложение φ называется *экзистенциальным предложением*, или $\exists$-*предложением*, если $\varphi = \exists x_1 \ldots \exists x_n\ \psi$, где ψ — бесквантворная формула, не содержащая других переменных, кроме $x_1, \ldots, x_n$.

Под *классом* алгебраических систем в дальнейшем будем понимать *абстрактный класс*, т. е. такое семейство L-систем, которое вместе с любой алгебраической системой содержит все изоморфные ей L-системы. Класс алгебраических систем называется *наследственным*, если он замкнут относительно подсистем.

Класс $\boldsymbol{K}$ алгебраических систем называется *аксиоматизируемым*, если существует такое множество предложений Z языка L, что для любой системы $\mathcal{A}$

$$\mathcal{A} \in \boldsymbol{K} \Leftrightarrow \mathcal{A} \models \varphi \text{ для всех } \varphi \in Z.$$

Множество предложений Z называется *множеством аксиом* для класса $\boldsymbol{K}$. Если для $\boldsymbol{K}$ существует конечное множество аксиом, то класс $\boldsymbol{K}$ называется *конечно аксиоматизируемым*. Если для $\boldsymbol{K}$ существует множество аксиом, состоящее только из $\forall$-предложений, то класс $\boldsymbol{K}$ называется *универсально аксиоматизируемым*, или $\forall$-*аксиоматизируемым*. Если для класса $\boldsymbol{K}$ существует *рекурсивное множество аксиом* Z, т. е. Z — система аксиом класса $\boldsymbol{K}$, и существует алгоритм, который по любому предложению языка L позволяет узнать, принадлежит оно множеству Z или нет, то класс $\boldsymbol{K}$ называется *рекурсивно аксиоматизируемым*.

Предложения φ_1 и φ_2 языка L будем называть *эквивалентными* на классе $\boldsymbol{K}$ алгебраических систем языка L, если для любой системы $\mathcal{A}$ класса $\boldsymbol{K}$

$$\mathcal{A} \models \varphi_1 \Leftrightarrow \mathcal{A} \models \varphi_2.$$

Пусть $S(\mathrm{L})$ — множество всех предложений языка L; $\boldsymbol{K}$ — некоторый класс L-систем. *Элементарной теорией* (или просто *теорией класса*) $\boldsymbol{K}$ называется множество $\mathrm{Th}(\boldsymbol{K})$ всех предложений из $S(\mathrm{L})$, истинных во всех системах из $\boldsymbol{K}$. Если существует алгоритм, который позволяет ответить на вопрос, принадлежит или нет произвольное предложение из $S(\mathrm{L})$ теории $\mathrm{Th}(\boldsymbol{K})$, то эта теория называется *разрешимой*. Множество всех $\forall$-предложений теории $\mathrm{Th}(\boldsymbol{K})$ называется *универсальной теорией*, или $\forall$-*теорией класса* $\boldsymbol{K}$. Множество всех $\exists$-предложений теории $\mathrm{Th}(\boldsymbol{K})$ называется *экзистенциальной теорией*, или $\exists$-*теорией класса* $\boldsymbol{K}$.

Пусть $\mathcal{H}$ — произвольное множество L-систем. Тогда класс $\mathrm{Forb}(\mathcal{H})$, который состоит из всех L-систем, не содержащих подсистем из $\mathcal{H}$ и им изоморфных, может быть определён заданием L-систем $\mathcal{A} \in \mathcal{H}$ в качестве *запрещённых подсистем*. Будем говорить, что класс L-систем $\boldsymbol{K}$ *определим в терминах запрещённых подсистем*, если $\boldsymbol{K} = \mathrm{Forb}(\mathcal{H})$ для некоторого множества $\mathcal{H}$. Чтобы определить *класс запрещённых подсистем* $\boldsymbol{H}$ для класса $\boldsymbol{K}$, необходимо для множества $\mathcal{H}$ взять его замыкание относительно изоморфизма.

Множество L-систем $\mathcal{H}$ называется *множеством минимальных запрещённых подсистем* для класса $\boldsymbol{K}$, если $\boldsymbol{K} = \mathrm{Forb}(\mathcal{H})$ и при этом для любой L-системы $\mathcal{A} \in \mathcal{H}$ всякая её собственная подсистема $\mathcal{A}_1 \notin \mathcal{H}$, как и все L-системы, изоморфные $\mathcal{A}_1$.

Утверждение 1. Пусть $\boldsymbol{K} = \mathrm{Forb}(\boldsymbol{H})$. Класс $\boldsymbol{H}$ является классом минимальных запрещённых подсистем для класса $\boldsymbol{K}$ тогда и только тогда, когда $\boldsymbol{H}$ является замыканием относительно изоморфизма минимального по включению множества $\mathcal{H}$ запрещённых подсистем для класса $\boldsymbol{K}$, т. е. $\boldsymbol{K} = \mathrm{Forb}(\mathcal{H})$ и $\boldsymbol{K} \neq \mathrm{Forb}(\mathcal{H}_1)$ для всех $\mathcal{H}_1 \subset \mathcal{H}$.

Доказательство.

Необходимость. Предположим противное: для любого минимального по включению множества $\mathcal{H}$ запрещённых подсистем для класса $\boldsymbol{K}$, замыканием которого относительно изоморфизма является класс $\boldsymbol{H}$, существует множество $\mathcal{H}_1 \subset \mathcal{H}$, такое, что $\boldsymbol{K} = \mathrm{Forb}(\mathcal{H}_1)$. Тогда существует L-система $\mathcal{A} \in \boldsymbol{H}$, такая, что $\mathcal{A} \in \mathcal{H} \setminus \mathcal{H}_1$. При этом

$\mathcal{A} \notin \mathrm{Forb}(\mathcal{H}) = \mathrm{Forb}(\mathcal{H}_1)$, т. е. существует L-система $\mathcal{A}_1 \in \mathcal{H}_1$, такая, что $\mathcal{A}_1$ является собственной подсистемой $\mathcal{A}$. Но поскольку $\mathcal{A}_1 \in \mathcal{H}$, а значит, и $\mathcal{A}_1 \in \boldsymbol{H}$, получаем противоречие с тем, что $\boldsymbol{H}$ — класс минимальных запрещённых подсистем для класса $\boldsymbol{K}$.

Достаточность. Предположим противное: существует L-система $\mathcal{A} \in \boldsymbol{H}$ и её собственная подсистема $\mathcal{A}_1 \in \boldsymbol{H}$. Без ограничения общности рассмотрим такие из них, которые обе принадлежат множеству $\mathcal{H}$. Рассмотрим множество $\mathcal{H}_1 = \mathcal{H} \setminus \{\mathcal{A}\}$. Очевидно, что $\boldsymbol{K} = \mathrm{Forb}(\mathcal{H}_1)$, поскольку все L-системы, не содержащие в качестве подсистемы $\mathcal{A}$, не должны содержать в качестве подсистемы и $\mathcal{A}_1 \in \mathcal{H}_1$. Получили противоречие с условием минимальности по включению множества $\mathcal{H}$. ■

Множество запрещённых подсистем языка L называется *рекурсивным*, если существует геделевская нумерация g этих подсистем, такая, что множество их номеров является рекурсивным, т. е. существует алгоритм, позволяющий узнать, принадлежит ли произвольное натуральное число множеству номеров.

Утверждение 2 (критерий $\forall$-аксиоматизируемости) [15]. Пусть $\boldsymbol{K}$ — аксиоматизируемый класс алгебраических систем языка L. Класс $\boldsymbol{K}$ является универсально аксиоматизируемым тогда и только тогда, когда он замкнут относительно подсистем.

В настоящей работе будем рассматривать только два типа алгебраических систем. Во-первых, модели $\mathcal{A} = \langle A, \mathrm{L_{fin}}\rangle$ конечных языков с равенством, в которых $\mathrm{L_{fin}} = \langle R_1, R_2, \ldots, R_m, =\rangle$. Во-вторых, модели $\mathcal{A} = \langle A, \mathrm{L}\rangle$ бесконечных языков с равенством вида $\mathrm{L} = \mathrm{L_{fin}} \cup \mathrm{L}_\infty$, где $\mathrm{L}_\infty = \langle R_{m+1}, R_{m+2}, \ldots\rangle$, причём в L_∞ число предикатов каждой местности конечно, все предикаты упорядочены по возрастанию своей местности и обладают свойством неповторения элементов, т. е. для всех $R_k \in \mathrm{L}_\infty$ выполнены следующие условия:

$$\forall x_1 \ldots \forall x_l \; [R_k(x_1, \ldots, x_l) \to \bigwedge_{i \neq j} (x_i \neq x_j)].$$

Обозначим через n_k местность предиката R_k.

Поскольку $\mathrm{L_{fin}} \subset \mathrm{L}$ для всех $m \in \mathbb{N}$, в дальнейшем конечный случай не выделяется в формулировках утверждений и их доказательствах, верных для любого рассматриваемого языка L.

2. Аксиоматизируемые наследственные классы

Рассмотрим несколько утверждений, необходимых для указания связи между наследственными классами моделей языка L и их запрещёнными подсистемами.

Лемма 1. Абстрактный класс L-систем $\boldsymbol{K}$ является наследственным тогда и только тогда, когда он может быть определён в терминах запрещённых подсистем.

Доказательство.

Необходимость. Пусть $\boldsymbol{K}$ — наследственный класс L-систем, т. е. для любых L-систем $\mathcal{A}_1$ и $\mathcal{A}_2$ если $\mathcal{A}_1 \in \boldsymbol{K}$ и $\mathcal{A}_2$ является произвольной подсистемой $\mathcal{A}_1$, то $\mathcal{A}_2 \in \boldsymbol{K}$. Рассмотрим класс $\boldsymbol{H}$ — дополнение к $\boldsymbol{K}$ в классе всех L-систем. Поскольку $\mathcal{A}_2 \notin \boldsymbol{H}$, то $\mathcal{A}_1 \in \mathrm{Forb}(\boldsymbol{H})$ и поэтому $\boldsymbol{K} \subseteq \mathrm{Forb}(\boldsymbol{H})$.

Теперь рассмотрим произвольную L-систему $\mathcal{A}_3 \in \mathrm{Forb}(\boldsymbol{H})$, т. е. такую, что всякая её подсистема не содержится в $\boldsymbol{H}$, в том числе и сама $\mathcal{A}_3$. Но тогда $\mathcal{A}_3 \in \boldsymbol{K}$ и, следовательно, $\mathrm{Forb}(\boldsymbol{H}) \subseteq \boldsymbol{K}$.

Таким образом, $\boldsymbol{K} = \mathrm{Forb}(\boldsymbol{H})$, т. е. класс $\mathbf{K}$ может быть определён в терминах запрещённых подсистем языка L.

Достаточность. Пусть $\boldsymbol{K} = \mathrm{Forb}(\boldsymbol{H})$ — класс, определимый в терминах запрещённых подсистем языка L. Предположим, что существует L-система $\mathcal{A}_1 \in \boldsymbol{K}$ и её подсистема $\mathcal{A}_2$, такая, что $\mathcal{A}_2 \notin \boldsymbol{K}$. Тогда L-система $\mathcal{A}_2$ содержит подсистему $\mathcal{A}_3$, такую, что $\mathcal{A}_3 \in \boldsymbol{H}$. Но поскольку $\mathcal{A}_3$ является также подсистемой $\mathcal{A}_1$, то $\mathcal{A}_1 \notin \mathrm{Forb}(\boldsymbol{H})$ — противоречие.

Таким образом, для любой L-системы $\mathcal{A}_1 \in \boldsymbol{K}$ всякая её подсистема содержится в $\boldsymbol{K}$. Следовательно, $\boldsymbol{K}$ — наследственный класс L-систем. ■

Лемма 2. Пусть $\boldsymbol{K} = \mathrm{Forb}(\boldsymbol{H})$, причём все L-системы класса $\boldsymbol{H}$ конечны, $\mathcal{A}$ — бесконечная L-система, каждая конечная подсистема которой принадлежит классу $\boldsymbol{K}$. Тогда $\mathcal{A}$ также принадлежит классу $\boldsymbol{K}$.

Доказательство. Предположим противное: $\mathcal{A} \notin \boldsymbol{K}$. Тогда существует её подсистема $\mathcal{A}_1$, такая, что $\mathcal{A}_1 \in \boldsymbol{H}$. Но по условию леммы $\mathcal{A}_1$ должна быть конечна — противоречие с тем, что все конечные подсистемы $\mathcal{A}$ принадлежат классу $\boldsymbol{K}$ и, следовательно, не принадлежат классу $\boldsymbol{H}$.

Таким образом, $\mathcal{A} \in \boldsymbol{K}$. ■

Теорема 1. Наследственный класс L-систем (универсально) аксиоматизируем тогда и только тогда, когда он может быть определён в терминах конечных запрещённых подсистем, причём обязательно существует минимальное по включению множество таких подсистем.

Доказательство.

Необходимость. По определению наследственный класс L-систем замкнут относительно подсистем, следовательно, в силу критерия универсальной аксиоматизируемости (утверждения 2) любой аксиоматизируемый наследственный класс L-систем является $\forall$-аксиоматизируемым, поэтому любая его аксиома может считаться $\forall$-предложением.

Тогда множество запрещённых подсистем наследственного класса $\boldsymbol{K}$, которое существует по лемме 1, можно задать следующим образом.

Для каждой аксиомы φ определим конечное множество $\mathcal{H}_\varphi$ запрещённых подсистем с числом элементов от 1 до p, где p — количество переменных в данной аксиоме, при помощи алгоритма 1. Затем объединим множества $\mathcal{H}_\varphi$ для всех аксиом $\{\varphi\}$ и для каждого набора изоморфных L-систем из этого объединения исключим все, кроме одной. В результате получим множество $\mathcal{H}$ конечных запрещённых подсистем языка L для данного наследственного класса $\boldsymbol{K}$.

Для доказательства существования минимального по включению множества запрещённых подсистем воспользуемся утверждением 1 и алгоритмом 2, который из множества конечных запрещённых подсистем $\mathcal{H}$ выделяет множество $\mathcal{H}_{\min}^{\leqslant k}$ минимальных запрещённых подсистем с числом элементов не большим k.

Поскольку множество $\mathcal{H}$ не содержит изоморфных L-систем и с учётом ограничений, наложенных на предикаты языка L, число k-элементных запрещённых подсистем всегда конечно, алгоритм 2 корректно работает для любого $k \in \mathbb{N}$. Если объединить последовательно найденные алгоритмом 2 множества $\mathcal{H}_{\min}^{\leqslant k}$ для всех k, для которых выполнено вложение $\mathcal{H}_{\min}^{\leqslant k} \subseteq \mathcal{H}_{\min}^{\leqslant k+1}$, то получится минимальное по включению множество $\mathcal{H}_{\min}$ запрещённых подсистем для класса $\boldsymbol{K}$.

Достаточность. Рассмотрим произвольный наследственный класс L-систем $\boldsymbol{K} = \mathrm{Forb}(\mathcal{H})$, где $\mathcal{H}$ — минимальное по включению множество конечных запрещённых подсистем. Любой конечной L-системе $\mathcal{A}$ можно поставить в соответствие условие существования подсистемы, изоморфной ей, при помощи алгоритма 3.

Алгоритм 1. Построение множества запрещённых подсистем $\mathcal{H}_\varphi$ для аксиомы φ

Вход: $\forall$-предложение φ.
Выход: Конечное множество $\mathcal{H}_\varphi$.

1: Отрицание аксиомы $\neg\varphi$ эквивалентно предложению $\exists x_1 \ldots \exists x_p\ \psi$, где ψ — бескванторная формула, находящаяся в предварённой дизъюнктивной форме (ПДФ), т. е. $\psi = \bigvee\limits_r \psi_r$, где ψ_r — конъюнкты.

2: Каждый конъюнкт ψ_r, не содержащий множителей $(x_i = x_j)$ и $(x_i \neq x_j)$, дополнить условием $(x_i = x_j) \vee (x_i \neq x_j)$. Полученное предложение, эквивалентное $\neg\varphi$, привести к ПДФ и обозначить $\neg\varphi_1$.

3: **Для всех** $R_k \in \mathrm{L}_{\mathrm{fin}}$ и конечного числа предикатов $R_k \in \mathrm{L}_\infty$, местность которых не превосходит p, каждый конъюнкт предложения $\neg\varphi_1$, не содержащий множителей $R_k(t_1, \ldots, t_l)$ и $\neg R_k(t_1, \ldots, t_l)$, дополнить условием $R_k(t_1, \ldots, t_l) \vee \neg R_k(t_1, \ldots, t_l)$ для любых $\{t_1, \ldots, t_l\} \subseteq \{x_1, \ldots, x_p\}$, где $l = n_k$. Затем перейти к предложению $\neg\varphi_2$, эквивалентному $\neg\varphi_1$, находящемуся в ПДФ.

4: Каждый конъюнкт предложения $\neg\varphi_2$ либо с точностью до изоморфизма задаёт подсистему языка L с фиксированным числом элементов от 1 до p и всеми возможными фиксированными наборами элементов, удовлетворяющих или не удовлетворяющих предикатам R_k языка L, либо противоречит наложенным на L-системы ограничениям и удаляется. В итоге получается предложение $\neg\varphi_3$ в ПДФ, по конъюнктам которого строится множество запрещённых подсистем $\mathcal{H}_\varphi$.

Алгоритм 2. Построение множества $\mathcal{H}_{\min}^{\leqslant k}$ минимальных запрещённых подсистем

Вход: Множество конечных запрещённых подсистем $\mathcal{H}$.
Выход: Множество $\mathcal{H}_{\min}^{\leqslant k}$.

1: **Для всех** $i = 1, \ldots, k-1$:
Просмотреть все $(i+1)$-элементные запрещённые подсистемы из $\mathcal{H}$, составляющие множество $\mathcal{H}^{i+1}$. Удалить из него все L-системы, содержащие подсистемы из $\mathcal{H}^1, \ldots, \mathcal{H}^i$.

2: $\mathcal{H}_{\min}^{\leqslant k} := \bigcup\limits_{i=1}^{k} \mathcal{H}^i$.

Алгоритм 3. Построение предложения φ, означающего условие существования подсистемы $\mathcal{A}$

Вход: Конечная L-система $\mathcal{A}$.
Выход: $\exists$-предложение φ.

1: $\varphi := \exists x_1 \ldots \exists x_p\ \psi$, где p — число элементов L-системы $\mathcal{A}$; ψ — пустой конъюнкт.

2: В конъюнкт ψ добавить условия попарного различия переменных $x_1, \ldots, x_p$.

3: **Для всех** предикатов $R_k \in \mathrm{L}_{\mathrm{fin}}$ и конечного числа предикатов $R_k \in \mathrm{L}_\infty$, местность которых не превосходит p, а также всех возможных множеств $\{t_1, \ldots, t_l\} \subseteq \{x_1, \ldots, x_p\}$, где $l = n_k$, в конъюнкт ψ добавить множители $R_k(t_1, \ldots, t_l)$ или $\neg R_k(t_1, \ldots, t_l)$ в зависимости от того, как соответствующие этим переменным элементы представлены в системе $\mathcal{A}$.

Тогда аксиоматика класса $\boldsymbol{K}$ должна состоять из множества аксиом, каждая из которых соответствует одной из запрещённых подсистем из множества $\mathcal{H}$, т. е. аксиомами являются отрицания соответствующих предложений $\varphi = \exists x_1 \ldots \exists x_p\ \psi$. При этом в силу леммы 2 таких аксиом достаточно, чтобы выделить не только конечные L-системы, принадлежащие классу $\boldsymbol{K}$, но и бесконечные. Таким образом, любая L-система, удовлетворяющая множеству аксиом $\{\neg\varphi\}_{\mathcal{H}}$, содержится в наследственном классе $\boldsymbol{K}$, причём все аксиомы являются $\forall$-предложениями, т. е. класс $\boldsymbol{K}$ универсально аксиоматизируем. ■

Для наглядности разберём в качестве примеров два наследственных класса гиперграфов с рёбрами конечной мощности.

Гиперграф с рёбрами конечной мощности — это алгебраическая система $H = \langle V, \mathrm{L}_H\rangle$, носитель которой V — непустое множество вершин, а язык $\mathrm{L}_H = \langle E_1, E_2, \ldots, =\rangle$ состоит из счётного множества предикатов, местность каждого из которых совпадает с его порядковым номером, и предиката равенства; каждый предикат $E_n(x_1, \ldots, x_n)$ означает, что элементы $x_1, \ldots, x_n$ лежат в ребре гиперграфа мощности n, т. е. предикаты $E_n(x_1, \ldots, x_n)$ удовлетворяют условиям *неупорядоченности* и *неповторения элементов* для всех $n \in \mathbb{N}$:

(H1) $\forall x_1 \ldots \forall x_n\ [E_n(x_1, \ldots, x_n) \Rightarrow \bigwedge\limits_{\pi} E_n(\pi(x_1), \ldots, \pi(x_n))]$, где π — любая перестановка $x_1, \ldots, x_n$;

(H2) $\forall x_1 \ldots \forall x_n\ [E_n(x_1, \ldots, x_n) \Rightarrow \bigwedge\limits_{p \neq q} (x_p \neq x_q)]$.

Подгиперграф — это подсистема, полученная из исходного гиперграфа удалением вершин вместе со всеми инцидентными рёбрами.

Пример 1. Гиперграф называется *линейным*, если его рёбра пересекаются максимум по одной вершине, т. е. *линейный гиперграф* — это алгебраическая система $H = \langle V, \mathrm{L}_H\rangle$, которая является гиперграфом и для всех $k, m \in \mathbb{N}$ удовлетворяет условию

(H3′) $\forall x_1 \ldots \forall x_k\ \forall y_1 \ldots \forall y_m\ [E_k(x_1, \ldots, x_k) \wedge E_m(y_1, \ldots, y_m) \wedge (x_p = y_g) \Rightarrow \bigwedge\limits_{\substack{i \neq p \\ j \neq q}} (x_i \neq y_j)]$.

Запрещёнными подгиперграфами для данного класса являются все гиперграфы, у которых хотя бы одна пара рёбер содержит в пересечении не менее двух вершин. Множество *минимальных запрещённых подгиперграфов для класса линейных гиперграфов* состоит из гиперграфов на l вершинах ($l \geqslant 3$), обладающих следующими свойствами:

— существует пара рёбер, которая содержит в пересечении не менее двух вершин;
— при удалении любой вершины в полученном гиперграфе не будет ни одной пары рёбер, которая бы содержала в пересечении не менее двух вершин.

Перечислим с точностью до изоморфизма все минимальные запрещённые подгиперграфы для класса линейных гиперграфов с числом вершин, не превосходящим 3:

— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_3\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1\}, \{v_2\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1\}, \{v_3\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1\}, \{v_2\}, \{v_3\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_1\}\}$;

— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_2\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_1\}, \{v_2\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_2\}, \{v_3\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_1\}, \{v_2\}, \{v_3\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_1\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_1\}, \{v_2\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_1\}, \{v_2\}, \{v_3\}\}$.

Пример 2. Гиперграф называется *антицепью*, если никакое из его рёбер не является подмножеством другого ребра, т. е. *антицепь* — это алгебраическая система $H = \langle V, \mathrm{L}_H \rangle$, которая является гиперграфом и для всех $k, m \in \mathbb{N}$ удовлетворяет условию

(НЗ″) $\forall x_1 \ldots \forall x_k \; \forall y_1 \ldots \forall y_m \; [E_k(x_1, \ldots, x_k) \Rightarrow \neg E_{k+m}(x_1, \ldots, x_k, y_1, \ldots, y_m)]$.

Запрещёнными подгиперграфами для данного класса будут все гиперграфы, у которых хотя бы одно ребро содержится в другом ребре. Множество *минимальных запрещённых подгиперграфов для класса антицепей* состоит из гиперграфов на l вершинах ($l \geqslant 2$), обладающих следующими свойствами:

— существует ребро, содержащее все вершины гиперграфа;
— существует хотя бы одно ребро мощности меньшей l;
— никакие два ребра мощности меньшей l не содержатся друг в друге.

Перечислим с точностью до изоморфизма все минимальные запрещённые подгиперграфы для класса антицепей с числом вершин, не превосходящим 3:

— $V = \{v_1, v_2\}$, $E = \{\{v_1, v_2\}, \{v_1\}\}$;
— $V = \{v_1, v_2\}$, $E = \{\{v_1, v_2\}, \{v_1\}, \{v_2\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1\}, \{v_2\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1\}, \{v_2\}, \{v_3\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_3\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}\}$;
— $V = \{v_1, v_2, v_3\}$, $E = \{\{v_1, v_2, v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}\}$.

Замечание 1. Классы линейных гиперграфов и антицепей имеют рекурсивные множества минимальных запрещённых подгиперграфов.

Действительно, всем гиперграфам можно поставить в соответствие экзистенциальные предложения, означающие существование изоморфных им подгиперграфов, причём все эти предложения обладают уникальными номерами (см., например, нумерацию формул в [6, с. 42]) и таким особым видом, что их можно алгоритмически выделить среди всех предложений языка L_H. По определению множества минимальных запрещённых подгиперграфов для класса линейных гиперграфов члены этого множества можно алгоритмически выделить среди всех гиперграфов, проверив выполнение соответствующих условий. Таким образом, для любого натурального числа можно алгоритмически установить, является ли оно номером какого-либо подгиперграфа из множества минимальных запрещённых подгиперграфов для класса линейных гиперграфов. Для класса антицепей рассуждения полностью аналогичны.

3. Разрешимость универсальных теорий наследственных классов

Установление разрешимости теории какого-либо класса $\boldsymbol{K}$ алгебраических систем позволяет сделать вывод о принципиальной возможности получения исчерпывающего перечня свойств, присущих всем системам этого класса. Поскольку разрешимые теории в чистом виде встречаются довольно редко, то доказательство разрешимости универсальной теории и построение соответствующего алгоритма является актуальной задачей.

Теорема 2. Универсальная теория произвольного аксиоматизируемого наследственного класса L-систем, множество минимальных запрещённых подсистем которого рекурсивно, разрешима.

Доказательство. В силу теоремы 1 мы рассматриваем случай, когда $\mathrm{Th}_\forall(\boldsymbol{K})$ — универсальная теория произвольного наследственного класса L-систем $\boldsymbol{K}$, определённого в терминах конечных запрещённых подсистем. Рассмотрим алгоритм 4 проверки предложения на принадлежность $\mathrm{Th}_\forall(\boldsymbol{K})$, на вход которому подаётся произвольное универсальное предложение φ. Его отрицание $\neg\varphi$ преобразуется в предложение, эквивалентное на классе всех L-систем и находящееся в предварённой дизъюнктивной форме. Алгоритм пытается построить L-систему класса $\boldsymbol{K}$, на которой предложение $\neg\varphi$ истинно. Если это удаётся, то предложение φ не принадлежит универсальной теории $\mathrm{Th}_\forall(\boldsymbol{K})$ и алгоритм выдаёт ответ «НЕТ». В противном случае φ принадлежит этой теории и алгоритм выдаёт ответ «ДА».

Алгоритм 4. Проверка универсального предложения φ на принадлежность $\mathrm{Th}_\forall(\boldsymbol{K})$

Вход: Предложение φ.
Выход: Ответ «ДА» или «НЕТ».

1: Для предложения φ построить его отрицание $\neg\varphi = \exists x_1 \ldots \exists x_p\ \psi$, где ψ — бескванторная формула, и преобразовать в эквивалентное предложение $\neg\varphi_1$, находящееся в ПДФ: $\neg\varphi_1 = \exists x_1 \ldots \exists x_p \bigvee\limits_r \psi_r$, где ψ_r — конъюнкты.

2: Просмотреть все конъюнкты предложения $\neg\varphi_1$. Если в конъюнкте ψ_r содержатся переменные x_i и x_j, но нет ни множителя $(x_i = x_j)$, ни множителя $(x_i \neq x_j)$, то заменить конъюнкт ψ_r на дизъюнкцию $[\psi_r \wedge (x_i = x_j)] \vee [\psi_r \wedge (x_i \neq x_j)]$. Эта процедура продолжается, пока возможно. Получим эквивалентное предложение $\neg\varphi_2$, в каждом конъюнкте которого все его переменные будут связаны между собой равенствами и неравенствами.

3: Просмотреть все конъюнкты предложения $\neg\varphi_2$:

   1) в каждом конъюнкте ψ_r, содержащем множитель $(x_i = x_j)$, заменить все вхождения переменной x_j на x_i в остальных множителях конъюнкта ψ_r и удалить исходное равенство и повторяющиеся множители;
   2) если в конъюнкте ψ_r содержатся множители вида $(t = t)$, где $t \in \{x_1, \ldots, x_p\}$, то удалить их как избыточные;
   3) если в конъюнкте ψ_r содержатся множители вида $(t \neq t)$, то удалить конъюнкт из предложения как тождественно ложный.

   Данная процедура продолжается до тех пор, пока из всех конъюнктов не исключатся все равенства и неравенства вида $(t \neq t)$, где $t \in \{x_1, \ldots, x_p\}$. В итоге получим эквивалентное предложение $\neg\varphi_3$ в ПДФ, в котором каждый конъюнкт содержит условия попарного различия всех входящих в него переменных.

4: Просмотреть все конъюнкты предложения $\neg\varphi_3$ и в них все предикаты $R_k \in \mathrm{L_{fin}}$:

1) если в конъюнкте ψ_r содержатся переменные $t_1, \ldots, t_l$, где $t_i \in \{x_1, \ldots, x_p\}$, но нет ни множителя $R_k(t_1, \ldots, t_l)$, ни множителя $\neg R_k(t_1, \ldots, t_l)$, то заменить конъюнкт ψ_r на дизъюнкцию $[\psi_r \wedge R_k(t_1, \ldots, t_l)] \vee [\psi_r \wedge \neg R_k(t_1, \ldots, t_l)]$. Эта процедура продолжается, пока возможно, для всех наборов входящих в конъюнкт переменных $\{t_1, \ldots, t_l\} \subseteq \{x_1, \ldots, x_p\}$, где $l = n_k$ и переменные в наборе могут повторяться;
2) если в конъюнкте ψ_r одновременно содержатся множители $R_k(t_1, \ldots, t_l)$ и $\neg R_k(t_1, \ldots, t_l)$, то удалить его из предложения как тождественно ложный.

В итоге получим эквивалентное предложение $\neg\varphi_4$ в ПДФ, в котором каждый конъюнкт содержит условия удовлетворения или неудовлетворения всех наборов входящих в него переменных всем предикатам $R_k \in \mathrm{L}_{\mathrm{fin}}$.

5: Просмотреть все конъюнкты предложения $\neg\varphi_4$ и в них все предикаты $R_k \in \mathrm{L}_\infty$:

1) если в конъюнкте ψ_r содержатся переменные $t_1, \ldots, t_l$, где $t_i \in \{x_1, \ldots, x_p\}$, но нет ни множителя $R_k(t_1, \ldots, t_l)$, ни множителя $\neg R_k(t_1, \ldots, t_l)$ при $n_k \leqslant p$, то заменить конъюнкт ψ_r на дизъюнкцию $[\psi_r \wedge R_k(t_1, \ldots, t_l)] \vee [\psi_r \wedge \neg R_k(t_1, \ldots, t_l)]$. Эта процедура продолжается, пока возможно, для всех наборов входящих в конъюнкт переменных $\{t_1, \ldots, t_l\} \subseteq \{x_1, \ldots, x_p\}$, где $l = n_k$ и рассматриваются только наборы неповторяющихся переменных;
2) если в конъюнкте ψ_r содержатся множители вида $\neg R_k(t_1, \ldots, t, \ldots, t, \ldots, t_l)$, то удалить их как избыточные;
3) если в конъюнкте ψ_r содержится множитель вида $R_k(t_1, \ldots, t, \ldots, t, \ldots, t_l)$ или одновременно содержатся множители $R_k(t_1, \ldots, t_l)$ и $\neg R_k(t_1, \ldots, t_l)$, то удалить конъюнкт из предложения как ложный на любой L-системе.

В итоге получим эквивалентное предложение $\neg\varphi_5$ в ПДФ, в котором каждый конъюнкт содержит условия удовлетворения или неудовлетворения всех возможных наборов входящих в него переменных всем необходимым предикатам $R_k \in \mathrm{L}_\infty$.

Если на каком-то из шагов 3–5 будут удалены все конъюнкты предложения, то предложение φ истинно для всех L-систем и, следовательно, принадлежит $\mathrm{Th}_\forall(\boldsymbol{K})$. Алгоритм завершает работу и выдаёт ответ «ДА».

Иначе получается предложение $\neg\varphi_5$ в ПДФ, эквивалентное предложению $\neg\varphi$, причём каждый конъюнкт предложения $\neg\varphi_5$ однозначно задаёт условие существования некой подсистемы языка L. Переход на шаг 6.

6: Просмотреть все конъюнкты предложения $\neg\varphi_5$. Для текущего конъюнкта ψ_r построить L-систему, которая задаётся его условием, и проверить её на принадлежность классу $\boldsymbol{K}$.

Если L-система принадлежит классу $\boldsymbol{K}$, то алгоритм завершает работу и выдаёт ответ «НЕТ».

Если L-система не принадлежит классу $\boldsymbol{K}$, то перейти к следующему конъюнкту.

Если все конъюнкты предложения $\neg\varphi_5$ просмотрены и ни для одного из них не удалось построить модели из класса $\boldsymbol{K}$, то алгоритм завершает работу и выдаёт ответ «ДА».

Обработка текущего конъюнкта ψ_r:

1) построить q-элементную L-систему $A_r = \langle A, \mathrm{L}\rangle$, элементы которой взаимно однозначно соответствуют переменным конъюнкта;
2) поскольку множество минимальных запрещённых подсистем класса $\boldsymbol{K}$ состоит из конечных L-систем и является рекурсивным, существует процеду-

ра, позволяющая узнать, принадлежит ли произвольная конечная L-система этому множеству. С помощью этой процедуры определить множество всех аксиом $\{\theta\}$, отрицания которых соответствуют минимальным запрещённым подсистемам класса $\boldsymbol{K}$, имеющим не более q элементов;

3) чтобы убедиться в принадлежности L-системы A_r классу $\boldsymbol{K}$, нужно проверить, не имеет ли она запрещённых подсистем, соответствующих множеству предложений $\{\theta\}$, т. е. нет ли такого экзистенциального предложения $\neg\theta$, которое истинно на L-системе A_r. Для этого для каждой аксиомы θ, содержащей n переменных ($n \leqslant q$), рассматриваются все возможные соответствия между переменными $\{x_1, x_2, \ldots, x_n\}$ аксиомы θ и элементами $\{1, 2, \ldots, q\}$ L-системы A_r (см. таблицу для $n = 3$, $q = 4$). Для каждого такого соответствия проверяется истинность $\neg\theta$ на L-системе A_r.

№	x_1	x_2	x_3
1	1	2	3
2	1	2	4
3	1	3	2
4	1	3	4
5	1	4	2
6	1	4	3
...	...	...	...
23	4	3	1
24	4	3	2

Если хотя бы одно из предложений $\{\neg\theta\}$ окажется истинным на L-системе A_r, то она не принадлежит классу $\boldsymbol{K}$. Перейти к следующему конъюнкту предложения $\neg\varphi_5$.

Если все предложения $\{\neg\theta\}$ окажутся ложными на L-системе A_r, то она не содержит запрещённых подсистем для класса $\boldsymbol{K}$ и принадлежит $\boldsymbol{K}$. Таким образом, для предложения $\neg\varphi$ построена модель из класса $\boldsymbol{K}$.

Предполагается, что на каждом шаге алгоритм удаляет из конъюнктов текущего предложения все повторяющиеся множители при их возникновении. Эти изменения никак не влияют на эквивалентность предложения на L-системах, далее не будем акцентировать на них внимание в силу их естественности.

В конечном итоге будет получен ответ на вопрос о принадлежности универсального предложения φ теории $\mathrm{Th}_\forall(\boldsymbol{K})$. ■

Следующие примеры демонстрируют работу алгоритма 4 из теоремы 2.

Пример 3. Алгоритмом рассматривается универсальная теория линейных гиперграфов. Предложение φ имеет вид

$$\forall x_1 \forall x_2 \forall x_3 \ \big[(x_1 = x_2) \vee (x_1 = x_3) \vee \neg E_3(x_1, x_2, x_3) \vee \neg E_3(x_1, x_3, x_2) \vee$$
$$\vee \neg E_3(x_2, x_1, x_3) \vee \neg E_3(x_2, x_3, x_1) \vee \neg E_3(x_3, x_1, x_2) \vee \neg E_3(x_3, x_2, x_1) \vee$$
$$\vee E_2(x_2, x_3) \vee E_2(x_3, x_2) \vee \neg E_2(x_1, x_2) \vee \neg E_2(x_2, x_1) \vee \bigvee_{i=1,2,3} \neg E_1(x_i)\big].$$

На шаге 1 формулируется его отрицание $\neg\varphi$, которое совпадает с $\neg\varphi_1$:

$$\exists x_1\exists x_2\exists x_3\ \big[(x_1\neq x_2)\wedge(x_1\neq x_3)\wedge E_3(x_1,x_2,x_3)\wedge E_3(x_1,x_3,x_2)\wedge$$
$$\wedge E_3(x_2,x_1,x_3)\wedge E_3(x_2,x_3,x_1)\wedge E_3(x_3,x_1,x_2)\wedge E_3(x_3,x_2,x_1)\wedge$$
$$\wedge E_2(x_1,x_2)\wedge E_2(x_2,x_1)\wedge\neg E_2(x_2,x_3)\wedge\neg E_2(x_3,x_2)\wedge\bigwedge_{i=1,2,3}E_1(x_i)\big].$$

На шаге 2 единственный конъюнкт $\psi_{1,1}$ предложения $\neg\varphi_1$ заменяется на дизъюнкцию

$$\big(\psi_{1,1}\wedge(x_2=x_3)\big)\vee\big(\psi_{1,1}\wedge(x_2\neq x_3)\big).$$

На шаге 3 после удаления равенств и повторяющихся множителей получается предложение $\neg\varphi_3$:

$$\exists x_1\exists x_2\exists x_3\ \big[\big((x_1\neq x_2)\wedge E_3(x_1,x_2,x_2)\wedge E_3(x_2,x_1,x_2)\wedge$$
$$\wedge E_3(x_2,x_2,x_1)\wedge E_2(x_1,x_2)\wedge E_2(x_2,x_1)\wedge\neg E_2(x_2,x_2)\wedge\bigwedge_{i=1,2}E_1(x_i)\big)\vee$$
$$\vee\big((x_1\neq x_2)\wedge(x_1\neq x_3)\wedge(x_2\neq x_3)\wedge E_3(x_1,x_2,x_3)\wedge E_3(x_1,x_3,x_2)\wedge$$
$$\wedge E_3(x_2,x_1,x_3)\wedge E_3(x_2,x_3,x_1)\wedge E_3(x_3,x_1,x_2)\wedge E_3(x_3,x_2,x_1)\wedge$$
$$\wedge E_2(x_1,x_2)\wedge E_2(x_2,x_1)\wedge\neg E_2(x_2,x_3)\wedge\neg E_2(x_3,x_2)\wedge\bigwedge_{i=1,2,3}E_1(x_i)\big)\big].$$

На шаге 4 предложение $\neg\varphi_4$ совпадает с $\neg\varphi_3$, поскольку $\mathrm{L}_{\mathrm{fin}}=\varnothing$.

На шаге 5 конъюнкт $\psi_{4,1}$ предложения $\neg\varphi_4$ удаляется как ложный, поскольку он содержит множители $E_3(x_1,x_2,x_2)$, $E_3(x_2,x_1,x_2)$ и $E_3(x_2,x_2,x_1)$, а конъюнкт $\psi_{4,2}$ заменяется на дизъюнкцию

$$\big(\psi_{4,2}\wedge E_2(x_1,x_3)\wedge E_2(x_3,x_1)\big)\vee\big(\psi_{4,2}\wedge E_2(x_1,x_3)\wedge\neg E_2(x_3,x_1)\big)\vee$$
$$\vee\big(\psi_{4,2}\wedge\neg E_2(x_1,x_3)\wedge E_2(x_3,x_1)\big)\vee\big(\psi_{4,2}\wedge\neg E_2(x_1,x_3)\wedge\neg E_2(x_3,x_1)\big).$$

Таким образом, получается предложение $\neg\varphi_5$, состоящее из четырёх конъюнктов:

$$\exists x_1\exists x_2\exists x_3\ \big[\psi_{5,1}\vee\psi_{5,2}\vee\psi_{5,3}\vee\psi_{5,4}\big].$$

На шаге 6 алгоритм строит L-системы для каждого из конъюнктов предложения $\neg\varphi_5$. Затем доказывается, что системы для конъюнктов $\psi_{5,1}$ и $\psi_{5,4}$ не являются линейными гиперграфами, а системы для конъюнктов $\psi_{5,2}$ и $\psi_{5,3}$ не являются гиперграфами вообще.

Алгоритм завершает работу и выдаёт ответ «ДА», т. е. исходное предложение φ принадлежит универсальной теории линейных гиперграфов.

Пример 4. Алгоритмом рассматривается универсальная теория антицепей. Предложение φ имеет вид

$$\forall x_1\forall x_2\forall x_3\ \big[(x_1=x_2)\vee(x_1=x_3)\vee(x_2=x_3)\vee E_3(x_1,x_2,x_3)\vee E_3(x_1,x_3,x_2)\vee$$
$$\vee E_3(x_2,x_1,x_3)\vee E_3(x_2,x_3,x_1)\vee E_3(x_3,x_1,x_2)\vee E_3(x_3,x_2,x_1)\vee E_2(x_1,x_3)\vee$$
$$\vee\neg E_2(x_1,x_2)\vee\neg E_2(x_2,x_1)\vee\neg E_2(x_2,x_3)\vee\neg E_2(x_3,x_2)\vee E_1(x_1)\vee E_1(x_2)\big].$$

На шаге 1 формулируется его отрицание $\neg\varphi$, которое совпадает с $\neg\varphi_1$:

$$\exists x_1\exists x_2\exists x_3\ \big[(x_1\neq x_2)\wedge(x_1\neq x_3)\wedge(x_2\neq x_3)\wedge\neg E_3(x_1,x_2,x_3)\wedge\neg E_3(x_1,x_3,x_2)\wedge$$
$$\wedge\neg E_3(x_2,x_1,x_3)\wedge\neg E_3(x_2,x_3,x_1)\wedge\neg E_3(x_3,x_1,x_2)\wedge\neg E_3(x_3,x_2,x_1)\wedge E_2(x_1,x_2)\wedge$$
$$\wedge E_2(x_2,x_1)\wedge E_2(x_2,x_3)\wedge E_2(x_3,x_2)\wedge\neg E_2(x_1,x_3)\wedge\neg E_1(x_1)\wedge\neg E_1(x_2)\big].$$

На шаге 2 предложение $\neg\varphi_2$ совпадает с $\neg\varphi_1$; на шаге 3 предложение $\neg\varphi_3$ совпадает с $\neg\varphi_2$; на шаге 4 предложение $\neg\varphi_4$ совпадает с $\neg\varphi_3$.

На шаге 5 единственный конъюнкт $\psi_{4,1}$ предложения $\neg\varphi_4$ заменяется на дизъюнкцию

$$\big(\psi_{4,1} \wedge E_2(x_3, x_1) \wedge E_1(x_3)\big) \vee \big(\psi_{4,1} \wedge E_2(x_3, x_1) \wedge \neg E_1(x_3)\big) \vee$$
$$\vee \big(\psi_{4,1} \wedge \neg E_2(x_3, x_1) \wedge E_1(x_3)\big) \vee \big(\psi_{4,1} \wedge \neg E_2(x_3, x_1) \wedge \neg E_1(x_3)\big).$$

Таким образом, получается предложение $\neg\varphi_5$, состоящее из четырёх конъюнктов:

$$\exists x_1 \exists x_2 \exists x_3 \ \big[\psi_{5,1} \vee \psi_{5,2} \vee \psi_{5,3} \vee \psi_{5,4}\big].$$

На шаге 6 алгоритм строит L-системы для каждого из конъюнктов предложения $\neg\varphi_5$. Затем доказывается, что системы для конъюнктов $\psi_{5,1}$ и $\psi_{5,2}$ не являются гиперграфами, система для конъюнкта $\psi_{5,3}$ является гиперграфом, но не является антицепью, однако система для конъюнкта $\psi_{5,4}$ является антицепью.

Алгоритм завершает работу и выдаёт ответ «НЕТ», т. е. исходное предложение φ не принадлежит универсальной теории антицепей.

Заключение

В работе рассмотрены вопросы аксиоматизируемости и разрешимости универсальных теорий наследственных классов алгебраических систем языка $\mathrm{L} = \mathrm{L}_{\mathrm{fin}} \cup \mathrm{L}_{\infty}$. Предложенные идеи доказательства соответствующих результатов применимы при решении систем уравнений на гиперграфах, матроидах и других объектах, которые можно определить с использованием языка логики первого порядка этого вида. Остаются открытыми вопросы применения таких языковых конструкций для построения более эффективных алгоритмов решения практических задач, чем известные.

ЛИТЕРАТУРА

1. *Ильев А. В., Ильев В. П.* Алгоритмы решения систем уравнений над различными классами конечных графов // Прикладная дискретная математика. 2021. № 53. С. 89–102.
2. *Никитин А. Ю., Рыбалов А. Н.* О сложности проблемы разрешимости систем уравнений над конечными частичными порядками // Прикладная дискретная математика. 2018. № 39. С. 94–98.
3. *Балджанова Р. В., Ильев А. В., Ильев В. П.* О сложности кластеризации графа в задаче с ограничениями на размеры кластеров // Прикладная дискретная математика. 2023. № 60. С. 76–84.
4. *Il'ev A. V. and Il'ev V. P.* Bounds for the clustering complexity in a graph clustering problem with clusters of bounded size // J. Math. Sci. 2023. V. 275. P. 78–84.
5. *Зыков А. А.* Основы теории графов. М.: Вузовская книга, 2004. 664 с.
6. *Ершов Ю. Л., Лавров И. А., Тайманов А. Д., Тайцлин М. А.* Элементарные теории // Успехи матем. наук. 1965. Т. 20. № 4. С. 37–108.
7. *Bozapalidis A. and Kalampakas A.* An axiomatization of graphs // Acta Inform. 2004. V. 41. P. 19–61.
8. *Taylor W.* Atomic compactness and graph theory // Fundamenta Mathematicae. 1969. V. LXV. P. 139–145.
9. *Yamamoto M., Nishizaki S., Hagiya M., and Toda Y.* Formalization of planar graphs // LNCS. 1995. V. 971. P. 369–384.
10. *Caicedo X.* Finitely axiomatizable quasivarieties of graphs // Algebra Univers. 1995. V. 34. No. 2. P. 314–321.

11. *Ильев А. В.* Об аксиоматизируемости наследственных классов графов и матроидов // Сиб. электрон. матем. изв. 2016. Т. 13. С. 137–147.
12. *Ham L. and Jackson M.* Axiomatisability and hardness for universal Horn classes of hypergraphs // Algebra Univers. 2018. V. 79. Art. 30.
13. *Il'ev A. V. and Il'ev V. P.* On axiomatizability and decidability of universal theories of hereditary classes of matroids // J. Physics: Conf. Ser. 2019. V. 1210. Art. 012056.
14. *Stronkowski M. M.* Axiomatizations of universal classes through infinitary logic // Algebra Univers. 2018. V. 79. Art. 26.
15. *Ершов Ю. Л., Палютин Е. А.* Математическая логика. М.: Наука, 1987. 336 с.
16. *Горбунов В. А.* Алгебраическая теория квазимногообразий. Новосибирск: Научная книга, 1999. 368+xii с.
17. *Ильев А. В.* Разрешимость универсальных теорий и аксиоматизируемость наследственных классов графов // Труды института математики и механики УрО РАН. 2016. Т. 22. № 1. С. 100–111.

REFERENCES

1. *Il'ev A. V. and Il'ev V. P.* Algoritmy resheniya sistem uravneniy nad razlichnymi klassami konechnykh grafov [Algorithms for solving systems of equations over various classes of finite graphs]. Prikladnaya Diskretnaya Matematika, 2021, no. 53, pp. 89–102. (in Russian)
2. *Nikitin A. Yu. and Rybalov A. N.* O slozhnosti problemy razreshimosti sistem uravneniy nad konechnymi chastichnymi poryadkami [On complexity of the satisfiability problem of systems over finite posets]. Prikladnaya Diskretnaya Matematika, 2018, no. 39, pp. 94–98. (in Russian)
3. *Baldzhanova R. V., Il'ev A. V., and Il'ev V. P.* O slozhnosti klasterizatsii grafa v zadache s ogranicheniyami na razmery klasterov [On the complexity of graph clustering in the problem with bounded cluster sizes]. Prikladnaya Diskretnaya Matematika, 2023, no. 60, pp. 76–84. (in Russian)
4. *Il'ev A. V. and Il'ev V. P.* Bounds for the clustering complexity in a graph clustering problem with clusters of bounded size. J. Math. Sci., 2023, vol. 275, pp. 78–84.
5. *Zykov A. A.* Osnovy teorii grafov [Fundamentals of Graph Theory]. Moscow, Vuzovskaia kniga, 2004. 664 p. (in Russian)
6. *Ershov Yu. L., Lavrov I. A., Taimanov A. D., and Taitslin M. A.* Elementary theories. Russian Math. Surveys, 1965, vol. 20, iss. 4, pp. 35–105.
7. *Bozapalidis A. and Kalampakas A.* An axiomatization of graphs. Acta inform., 2004, vol. 41, pp. 19–61.
8. *Taylor W.* Atomic compactness and graph theory. Fundamenta Mathematicae, 1969, vol. LXV, pp. 139–145.
9. *Yamamoto M., Nishizaki S., Hagiya M., and Toda Y.* Formalization of planar graphs. LNCS, 1995, vol. 971, pp. 369–384.
10. *Caicedo X.* Finitely axiomatizable quasivarieties of graphs. Algebra Univers., 1995, vol. 34, no. 2, pp. 314–321.
11. *Il'ev A. V.* Ob aksiomatiziruemosti nasledstvennykh klassov grafov i matroidov [On axiomatizability of hereditary classes of graphs and matroids]. Siberian Electronic Math. Reports, 2016, vol. 13, pp. 137–147. (in Russian)
12. *Ham L. and Jackson M.* Axiomatisability and hardness for universal Horn classes of hypergraphs. Algebra Univers., 2018, vol. 79, art. 30.
13. *Il'ev A. V. and Il'ev V. P.* On axiomatizability and decidability of universal theories of hereditary classes of matroids. J. Physics: Conf. Ser., 2019, vol. 1210, art. 012056.

14. *Stronkowski M. M.* Axiomatizations of universal classes through infinitary logic. Algebra Univers., 2018, vol. 79, art. 26.
15. *Ershov Yu. L. and Palyutin E. A.* Matematicheskaia logika [Mathematical Logic]. Moscow, Nauka, 1987. 336 p. (in Russian)
16. *Gorbunov V. A.* Algebraic Theory of Quasivarieties. New York, Plenum, 1998. 298+xii p.
17. *Il'ev A. V.* Razreshimost' universal'nykh teoriy i aksiomatiziruemost' nasledstvennykh klassov grafov [Decidability of universal theories and axiomatizability of hereditary classes of graphs]. Trudy Instituta Matematiki i Mekhaniki UrO RAN, 2016, vol. 22, no. 1, pp. 100–111. (in Russian)

УДК 511.33 DOI 10.17223/20710410/66/3

ПЕРИОДИЧЕСКИЕ МУЛЬТИПЛИКАТИВНЫЕ АРИФМЕТИЧЕСКИЕ ФУНКЦИИ

Е. В. Кайгородов

Горно-Алтайский государственный университет, г. Горно-Алтайск, Россия

E-mail: gazetaintegral@gmail.com

Вводятся понятия периодической мультипликативной функции, основного модуля такой функции, простейшей периодической мультипликативной функции. Изучаются основные свойства периодических мультипликативных функций, а также даётся их полное описание через характеры Дирихле. В частности, доказывается, что всякая отличная от единичной периодическая мультипликативная функция единственным образом представляется в виде произведения простейших периодических мультипликативных функций, причём основные модули таких функций представляют собой степени простых чисел, произведение которых есть каноническое разложение основного модуля исходной функции. На основании этого представления исследование периодических мультипликативных функций сводится к исследованию простейших периодических мультипликативных функций. Полученные результаты подводят к полному описанию периодических мультипликативных функций.

Ключевые слова: *арифметическая функция, периодическая мультипликативная функция, характер Дирихле, L-функция Дирихле.*

PERIODIC MULTIPLICATIVE ARITHMETIC FUNCTIONS

E. V. Kaigorodov

Gorno-Altaisk State University, Gorno-Altaisk, Russia

The notions of a periodic multiplicative function, the main modulus of such function, and the simplest periodic multiplicative function have been introduced. The basic properties of periodic multiplicative functions are studied, and a complete description of such functions through Dirichlet characters is given. In particular, it has been proven that any periodic multiplicative function other than unitary can be uniquely represented as a product of the simplest periodic multiplicative functions, and the principal modules of such functions are powers of prime numbers, the product of which is the canonical decomposition of the principal module of the original function. Based on this representation, the study of periodic multiplicative functions is reduced to the study of the simplest periodic multiplicative functions. The obtained results lead to a complete description of periodic multiplicative functions.

Keywords: *arithmetic function, periodic multiplicative function, Dirichlet character, Dirichlet L-function.*

Введение

Определение 1. Периодической мультипликативной функцией будем называть числовую функцию $f(n)$ натурального аргумента со следующими свойствами:

1) $f(n)$ отлична от тождественного нуля;
2) $f(n)$ мультипликативна, то есть $f(mn) = f(m)f(n)$ для любых взаимно простых натуральных чисел m и n;
3) $f(n)$ периодична, то есть найдётся такое натуральное число m, называемое периодом, что $f(n+m) = f(n)$ для любого n.

Пусть периодические мультипликативные функции $f(n)$ и $g(n)$ имеют период m. Нетрудно видеть, что в силу периодичности обеих функций они тождественно равны тогда и только тогда, когда их значения совпадают на некоторой полной системе вычетов по модулю m. Это значит, что любая периодическая мультипликативная функция однозначно определяется заданием своих значений на полной системе вычетов по модулю m.

Наименьший среди всех периодов обозначим буквой k и назовём основным модулем периодической мультипликативной функции $f(n)$. Чтобы указать, что число k является основным модулем функции $f(n)$, будем эту функцию записывать в виде $f(n,k)$. Эти и все другие обозначения, а также терминология в данной работе стандартны и заимствованы из [1].

Систематически изучать периодические мультипликативные функции начал, по-видимому, Г. Канольд, установивший в [2, 3] их простейшие свойства. Периодические арифметические функции привлекали внимание Т. М. Апостола — им посвящён § 27.10 фундаментального справочника коллектива авторов [4]. А. Конси и Т. Макгенри в [5] описали интересные примеры использования некоторых мультипликативных арифметических функций вместе с шифром Хилла в криптографии с открытым ключом.

В настоящее время изучение арифметических функций, их важнейших классов и свойств оправдано потребностями практики. На наш взгляд, именно актуальные проблемы криптографии побуждают сейчас специалистов заниматься вопросами описания конкретных классов функций и конструированием новых арифметических функций с определёнными свойствами, с использованием которых впоследствии, возможно, будут разработаны криптосистемы для постквантовой криптографии. Так, в декабре 2022 г. группа китайских ученых предложила способ взлома 2048-битного ключа криптосистемы RSA и допустила возможность реализации этого способа в будущем на квантовом компьютере с 372 физическими кубитами и глубиной квантовой схемы более 1000 [6]. Это обстоятельство даёт мощный толчок к ускорению теоретико-числовых исследований в области постквантовой криптографии. Крайне необходимо, чтобы такая система была построена и получила уверенное развитие к моменту квантового взлома.

Известно, что ключ дешифрования криптосистемы RSA определяется по функции Эйлера $\varphi(n)$, которая представляет собой классический пример мультипликативной арифметической функции в теории чисел. Свойства функции Эйлера позволили ей сыграть важную роль в построении названной криптосистемы. Эти факты наводят на мысль о вероятном создании в обозримом будущем криптосистем, в которых найдут применение новые арифметические функции с наперёд заданными свойствами, «заточенными» под специфику разрабатываемых квантовых алгоритмов шифрования. Задача поиска и изучения таких функций в некоторой мере смежна с проблемой описания конкретных классов арифметических (в частности, мультипликативных) функ-

ций. Понимание строения периодических мультипликативных арифметических функций может помочь в дальнейшем получить новые арифметические функции, пригодные для использования в постквантовых криптосистемах.

Следуя в основных чертах методу первой главы книги Н. Г. Чудакова [1], можно полностью описать строение периодических мультипликативных функций.

1. Основные свойства периодических мультипликативных функций

Теорема 1 [1, теорема 2]. Произведение конечного числа периодических мультипликативных функций есть также периодическая мультипликативная функция, основной модуль которой равен делителю наименьшего общего кратного основных модулей сомножителей, этот модуль равен произведению основных модулей сомножителей, если последние попарно взаимно просты.

Следующее простое утверждение по существу совпадает с теоремой 1 работы [3].

Теорема 2. Пусть имеем периодическую мультипликативную функцию $f(n,k)$. Существует характер Дирихле $\chi(n,k')$, где $k' \mid k$, такой, что $f(n,k) = \chi(n,k')$ для взаимно простых n и k.

Доказательство. Рассмотрим функцию $g(n) = f(n,k)\chi_0(n,k'')$, где k'' — произведение всех простых чисел, входящих в каноническое разложение k; $\chi_0(n,k'')$ — главный характер. Из теоремы 1 следует, что $g(n)$ — периодическая мультипликативная функция, причём её основной модуль k' делит наименьшее общее кратное чисел k и k'', равное k. Но если $p \nmid k$, то, как легко показать индукцией, $f(p^\alpha) = (f(p))^\alpha$: для $\alpha = 1$ это очевидно, а если это верно для данного α, то $(f(p))^{\alpha+1} = (f(p))^\alpha f(p) = f(p^\alpha)f(p+k) = f(p^{\alpha+1} + p^\alpha k) = f(p^{\alpha+1})$. Если же $p \mid k$, то $g(p^\alpha) = 0$. Значит, функция $g(n)$ вполне мультипликативна и является характером Дирихле. ■

Лемма 1. Если $(a,m) = 1$, n — натуральное число, то найдётся такое натуральное число x, что $(a + mx, n) = 1$.

Доказательство. Следует из теоремы Дирихле о простых числах в арифметической прогрессии. ■

Лемма 2. Пусть $A_i(x) = a_i + m_i x$, $d_i = (a_i, m_i)$; $(d_i, d_j) = 1$ при $i \neq j$, $i,j = 1,2,\ldots,\nu$. Тогда существует ν натуральных чисел $x_1, x_2, \ldots, x_\nu$, таких, что числа $A_1(x_1), A_2(x_2), \ldots, A_\nu(x_\nu)$ попарно взаимно просты.

Доказательство. Положим $a_i = a'_i d_i$, $m_i = m'_i d_i$, $A_i(x) = A'_i(x)d_i$. Тогда $A_i(x) = d_i(a'_i + m'_i x)$, причём $(a'_i, m'_i) = 1$. По лемме 1 найдётся такое натуральное число x_1, что $A'_1(x_1)$ взаимно просто с $d_1 d_2 \ldots d_\nu$. Далее, найдётся такое натуральное число x_2, что $A'_2(x_2)$ взаимно просто с $d_1 d_2 \ldots d_\nu A'_1(x_1)$, и т. д. Наконец, найдётся такое натуральное число x_ν, что $A'_\nu(x_\nu)$ взаимно просто с $d_1 d_2 \ldots d_\nu A'_1(x_1) A'_2(x_2) \ldots A'_{\nu-1}(x_{\nu-1})$. Очевидно, числа $x_1, x_2, \ldots, x_\nu$ — искомые. ■

Теорема 3. Пусть k — основной модуль периодической мультипликативной функции $f(n,k)$ и $k = k_1 k_2 \ldots k_\nu$, где все k_i попарно взаимно просты. Тогда существует единственная система периодических мультипликативных функций $f_1(n), f_2(n), \ldots, f_\nu(n)$, основные модули которых соответственно равны $k_1, k_2, \ldots, k_\nu$ и таковы, что $f(n) = f_1(n)f_2(n)\ldots f_\nu(n)$. При этом области значений функций $f_1(n)$, $f_2(n)$, …, $f_\nu(n)$ суть части области значений функции $f(n)$.

Доказательство. Пусть дано произвольное целое число $i \leqslant \nu$. Для каждого n определим (как в [1, теорема 3]) число n_i условиями

$$n_i \equiv n \pmod{k_i}, \quad n_i \equiv 1 \pmod{k_j} \text{ для всех } i \neq j. \tag{1}$$

Полагаем теперь $f_i(n) = f(n_i)$. Все n_i для данного n образуют один класс вычетов по модулю k, поэтому функция $f_i(n)$ определена однозначно. Ясно, что $f_i(1) = 1$, откуда следует, что функция $f_i(n)$ отлична от тождественного нуля.

Для произвольных взаимно простых чисел m и n имеем по определению m_i и n_i:

$$\begin{aligned} m_i &\equiv m \pmod{k_i}, \qquad m_i \equiv 1 \pmod{k_j}, \\ n_i &\equiv n \pmod{k_i}, \qquad n_i \equiv 1 \pmod{k_j}, \quad i \neq j. \end{aligned} \tag{2}$$

Перемножая сравнения, получаем $m_i n_i \equiv mn \pmod{k_i}$, $m_i n_i \equiv 1 \pmod{k_j}$, $i \neq j$. Отсюда имеем $f_i(mn) = f(m_i n_i) = f(m_i) f(n_i) = f_i(m) f_i(n)$, если числа m_i и n_i взаимно просты. Однако их можно такими выбрать. Действительно, если $(m_i, n_i) = d > 1$, то $(d, k) = 1$, так как в противном случае существовало бы такое простое число p, что $p \mid d$, $p \mid k$ и $p \mid k_i$ — потому что в силу (2) $p \not| k_j$ при $i \neq j$ и, наконец, $p \mid m_i$, $p \mid n_i$, откуда $p \mid m$ и $p \mid n$.

Из взаимной простоты чисел d и k по лемме 2 следует существование таких натуральных чисел x и y, что $(m_i + kx, n_i + ky) = 1$, поскольку $((m_i, k), (n_i, k)) = (d, k) = 1$.

Мультипликативность функции $f_i(n)$ доказана. Периодичность доказывается так же, как в [1, теорема 3].

Полагая $i = 1, 2, \ldots, \nu$, получим ν функций $f_1(n), f_2(n), \ldots, f_\nu(n)$. Покажем, что $f(n) = f_1(n) \cdot f_2(n) \cdot \ldots \cdot f_\nu(n)$. Это можно сделать по аналогии с [1, теорема 3], если доказать, что для данного n числа $n_1, n_2, \ldots, n_\nu$ можно выбрать попарно взаимно простыми. Положим для этого $K_i = k/k_i$. Рассмотрим арифметические прогрессии $n_i + kx$, где $i = 1, 2, \ldots, \nu$, и обозначим (n_i, k) через d_i. Имеем $(n_i, k_i) = d_i$, так как $(n_i, K_i) = 1$ в силу (1). Но $(k_i, k_j) = 1$ при $i \neq j$, а $d_i \mid k_i$, поэтому $(d_i, d_j) = 1$. Остаётся применить лемму 2.

Конец доказательства переносится сюда из [1] без изменений. ■

Теорема 3 показывает, что всякая отличная от единичной периодическая мультипликативная функция единственным образом представляется произведением вида

$$f(n, p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_\nu^{\alpha_\nu}) = f(n, p_1^{\alpha_1}) f(n, p_2^{\alpha_2}) \ldots f(n, p_\nu^{\alpha_\nu}).$$

2. Строение периодических мультипликативных функций

Изучим строение простейших периодических мультипликативных функций, то есть функций вида $f(n) = f(n, p^\alpha)$. Будем считать, что $\alpha \geqslant 1$ (иначе получается единичная функция). Примем обозначения: $\chi(n, p^\beta)$, $\beta \leqslant \alpha$, — характер Дирихле, соответствующий периодической мультипликативной функции $f(n, p^\alpha)$ в силу теоремы 2; $a_\nu = f(p^\nu)$, $\nu = 1, 2, \ldots$

Теорема 4. При $\nu \geqslant \alpha$ справедливо равенство $a_\nu = a_\alpha$. Если при этом характер $\chi(n)$ неглавный, то $a_\alpha = 0$.

Доказательство. Так как p^α — период функции $f(n)$, то при $\nu \geqslant \alpha$ получаем $a_\nu = f(p^\nu) = f(p^\alpha) = a_\alpha$. Пусть теперь характер $\chi(n, p^\beta)$ неглавный, тогда существует такое c, что $\chi(c) \neq 0$ и $\chi(c) \neq 1$. Для этого c будем иметь $\chi(c) a_\alpha = f(cp^\alpha) = f(p^\alpha) = a_\alpha$, откуда $a_\alpha = 0$. ■

Теорема 5. Пусть

$$f(n) = \begin{cases} \chi(n, p^\gamma), & \text{если } p \nmid n, \\ a_\nu, & \text{если } n = p^\nu, \end{cases}$$

причём $a_\beta \neq a_{\beta+1} = a_{\beta+2} = \ldots$ и $a_\beta = 0$, если характер $\chi(n)$ неглавный. Тогда $f(n)$ есть периодическая мультипликативная функция основного модуля p^α, где $\alpha = \beta + \gamma$.

Доказательство. Мультипликативность $f(n)$ очевидна. Непосредственной проверкой легко установить, что $p^{\beta+\gamma}$ есть период $f(n)$. Допустим теперь, что $p^{\beta+\gamma-1}$ — тоже период $f(n)$. Тогда если $\gamma = 1$, то $a_{\beta+1} = f(p^{\beta+1}) = f(p^\beta) = a_\beta$, что противоречит определению функции $f(n)$. Если $\gamma > 1$, то характер $\chi(n)$ неглавный, то есть $a_\beta = 0$. Возьмём число n, для которого $\chi(n) \neq \chi(n + p^{\gamma-1})$, очевидно, $p \nmid n$. Мы допустили, что $p^{\beta+\gamma-1}$ — период $f(n)$, поэтому

$$a_\beta \chi(n + p^{\gamma-1}) = f(p^\beta(n + p^{\gamma-1})) = f(np^\beta) = a_\beta \chi(n),$$

откуда $a_\beta = 0$, поскольку $\chi(n) \neq \chi(n + p^{\gamma-1})$. Снова пришли к противоречию. ■

Теоремы 1–5 полностью описывают строение периодических мультипликативных функций через характеры Дирихле.

Теперь можно установить связь между функциями

$$L(s, f) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

и классическими L-функциями Дирихле.

Пусть k — основной модуль периодической мультипликативной функции $f(n)$, а $\chi(n)$ — характер Дирихле, соответствующий функции $f(n)$ в силу теоремы 2. При $\operatorname{Re} s > 1$ имеем

$$L(s, f) = \prod_p \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \ldots\right) =$$

$$= \prod_{p \nmid k} \left(1 + \frac{\chi(p)}{p^s} + \frac{\chi(p^2)}{p^{2s}} + \ldots\right) \prod_{p \mid k} \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \ldots\right) =$$

$$= L(s, \chi) \prod_{p \mid k} \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \ldots\right).$$

Пусть $p^\alpha \mid k$. По теореме 3 запишем $f(n, k) = f_1(n, p^\alpha) f_2(n, k/p^\alpha)$. Значит,

$$f(p^\nu) = f_1(p^\nu) f_2(p^\nu) = a_\nu \chi_2(p^\nu) = a_\nu (\chi_2(p))^\nu = a_\alpha (\chi_2(p))^\nu$$

при $\nu \geqslant \alpha$. Поэтому

$$\sum_{\nu=0}^{\infty} \frac{(f(p))^\nu}{p^{\nu s}} = \sum_{\nu=0}^{\alpha-1} a_\nu \left(\frac{\chi_2(p)}{p^s}\right)^\nu + a_\alpha \sum_{\nu=\alpha}^{\infty} \left(\frac{\chi_2(p)}{p^s}\right)^\nu$$

является рациональной функцией (и даже многочленом, если окажется $a_\alpha = 0$) от p^{-s}. Таким образом, $L(s, f) = L(s, \chi) F(s)$, где χ — соответствующий f характер; $F(s)$ — произведение рациональных функций (и даже многочленов, если характер χ первообразный) от p^{-s} по всем p, делящим основной модуль периодической мультипликативной функции $f(n)$. Легко видеть, что в случае непервообразного характера $\chi(n)$ полюсы функции $F(s)$ гасятся тривиальными нулями $L(s, \chi)$, лежащими на прямой $\operatorname{Re} s = 0$.

Заключение

В работе полностью описано строение периодических мультипликативных арифметических функций через характеры Дирихле. Для решения этой задачи использованы методы теории характеров, предложенные советским математиком Н. Г. Чудаковым. Применение этих методов к изучению периодических мультипликативных арифметических функций дало вполне удовлетворительный результат, что говорит об их универсальности и потенциальной возможности распространения на смежные проблемы аналитической и мультипликативной теории чисел.

ЛИТЕРАТУРА

1. *Чудаков Н. Г.* Введение в теорию L-функций Дирихле. М.: ОГИЗ, 1947. 202 с.
2. *Kanold H. J.* Über periodische multiplikative zahlentheoretische Funktionen // Math. Ann. 1961. V. 144. P. 135–141. (in German)
3. *Kanold H. J.* Über periodische zahlentheoretische Funktionen // Math. Ann. 1962. V. 147. P. 269–274. (in German)
4. *Frank W. O., Daniel W. L., Ronald F. B., and Charles W. C.* NIST Handbook of Mathematical Functions. 1st. ed. Cambridge: Cambridge University Press, 2010. 966 p.
5. *Conci A. and MacHenry T.* Cryptography and multiplicative arithmetic functions // 2015 IEEE Intern. Conf. on Industrial Technology (ICIT). Seville, Spain, 2015. P. 1515–1519.
6. *Yan B., Tan Z., Wei S., et al.* Factoring Integers with Sublinear Resources on a Superconducting Quantum Processor. `https://arxiv.org/abs/2212.12372`. 2022.

REFERENCES

1. *Chudakov N. G.* Vvedenie v teoriyu L-funktsiy Dirikhle [Introduction to the theory of Dirichlet L-functions]. Moscow, OGIZ Publ., 1947. 202 p. (in Russian)
2. *Kanold H. J.* Über periodische multiplikative zahlentheoretische Funktionen. Math. Ann., 1961, vol. 144, pp. 135–141.
3. *Kanold H. J.* Über periodische zahlentheoretische Funktionen. Math. Ann., 1962, vol. 147, pp. 269–274.
4. *Frank W. O., Daniel W. L., Ronald F. B., and Charles W. C.* NIST Handbook of Mathematical Functions, 1st. ed. Cambridge, Cambridge University Press, 2010. 966 p.
5. *Conci A. and MacHenry T.* Cryptography and multiplicative arithmetic functions. 2015 IEEE Intern. Conf. on Industrial Technology (ICIT), Seville, Spain, 2015, pp. 1515–1519.
6. *Yan B., Tan Z., Wei S., et al.* Factoring Integers with Sublinear Resources on a Superconducting Quantum Processor. `https://arxiv.org/abs/2212.12372`, 2022.

УДК 512.532.3 DOI 10.17223/20710410/66/4

МУЛЬТИПЛИКАТИВНЫЕ ПОЛУГРУППЫ ВЫЧЕТОВ С ПЛАНАРНЫМИ ГРАФАМИ КЭЛИ

О. В. Князев, Д. В. Соломатин

Омский государственный педагогический университет, г. Омск, Россия

E-mail: solomatin_dv@omgpu.ru

Изучаются мультипликативные полугруппы вычетов, допускающие планарные графы Кэли. Доказано, что мультипликативная полугруппа кольца вычетов $\mathbb{Z}_n$ допускает планарный граф Кэли тогда и только тогда, когда $n = 4, 6, 8$ или n — простое число. Приведены примеры минимальных систем образующих мультипликативных полугрупп вычетов по некоторым модулям и их графы Кэли, иллюстрирующие полученные результаты.

Ключевые слова: *вычет, мультипликативная полугруппа вычетов, порождающее множество полугруппы, граф Кэли полугруппы, планарный граф.*

MULTIPLICATIVE RESIDUE SEMIGROUPS WITH PLANAR CAYLEY GRAPHS

O. V. Knyazev, D. V. Solomatin

Omsk State Pedagogical University, Omsk, Russia

We study multiplicative residue semigroups that admit planar Cayley graphs. It is proved that the multiplicative semigroup of the residue ring $\mathbb{Z}_n$ admits a planar Cayley graph if and only if $n = 4, 6, 8$ or n is a prime number. Examples of minimal systems of generators of multiplicative residue semigroups with respect to some modules and their Cayley graphs are given, illustrating the obtained results.

Keywords: *residue, multiplicative semigroup of residues, generating semigroup set, Cayley graph of semigroup, planar graph.*

Введение

Задача проверки изоморфизма графов является одной из важнейших задач в теории графов, среди имеющих прикладное и теоретическое значение. Она заключается в определении, являются ли два графа «одним и тем же» графом с точностью до нумерации вершин. В случае планарных графов задача проверки их изоморфизма разрешима за полиномиальное время [1]. Это означает, что существуют эффективные алгоритмы для решения этой задачи для планарных графов. Однако в общем случае изоморфность двух планарных графов Кэли $\mathrm{Cay}(S_1, X_1)$ и $\mathrm{Cay}(S_2, X_2)$ конечных полугрупп S_1 и S_2 с множествами образующих X_1 и X_2 не означает изоморфности полугрупп S_1 и S_2, а степень полинома в сложности алгоритма проверки изоморфизма двух полугрупп, допускающих планарный граф Кэли, зависит ещё и от числа различных минимальных множеств образующих этих полугрупп. Дело в том, что упомянутые алгоритмы используют специфичные структурные характеристики графов, поэтому область их

применения ограничена. Тем не менее отдельный интерес с прикладной точки зрения вызывает проблема изоморфизма конечных полугрупп. Эта проблема является важной в области теории полугрупп и теории вычислений, она изучалась многими исследователями.

Одним из важных классов графов, определяемых полугруппами, являются графы Кэли, поскольку они имеют многочисленные приложения [2, 3]. Один из возможных подходов к решению проблемы изоморфизма конечных полугрупп заключается в проверке изоморфности графов Кэли пар рассматриваемых полугрупп. Если при этом графы планарные, то существует полиномиальный алгоритм проверки их изоморфности. Поэтому возникает естественный вопрос: для каких конечных полугрупп граф Кэли планарен? Кроме того, решения этого вопроса могут пригодиться в таких прикладных областях, как криптография, проектирование топологии сетей и конечных автоматов. Известно, что мультипликативные полугруппы вычетов могут использоваться в криптографических алгоритмах. Структура этих полугрупп в сочетании со свойствами планарных графов Кэли может помочь в разработке безопасных криптографических систем.

Кольцо вычетов, группа вычетов, мультипликативная полугруппа вычетов относятся к ярким представителям модулярной математики, практическая значимость которой не вызывает сомнений. Изучение с разных сторон этих представителей позволит шире использовать прикладные возможности этой теории.

Планарные графы Кэли полезны при проектировании эффективных и надёжных топологий коммуникационных сетей. Свойства мультипликативных полугрупп вычетов могут быть использованы для оптимизации маршрутизации и проверки связности в этих сетях. Графы Кэли полугрупп тесно связаны также с конечными автоматами. Эта связь может использоваться для моделирования и анализа вычислительных процессов, особенно при проектировании эффективных алгоритмов и автоматов.

Изучение свойств планарных графов Кэли полугрупп может привести к новым открытиям в теории графов, таким, как понимание хроматического числа, связности и других инвариантов графов. Фундаментальные математические исследования взаимосвязей между мультипликативными полугруппами вычетов и планарными графами Кэли могут привести к новым открытиям в алгебре и комбинаторике, потенциально решая открытые проблемы или приводя к новым гипотезам. Не случайно обобщённые графы Кэли полугрупп всё чаще формируют тематику научных и исследовательских статей [4]. Упомянем также тот факт, что предмет геометрической теории групп и полугрупп составляет изучение конечно-порождённых групп и полугрупп путём отыскания связей между их алгебраическими свойствами и топологическими, геометрическими свойствами пространств, на которых такие системы действуют, либо самих систем, рассматриваемых как геометрические объекты, что обычно делается рассмотрением графа Кэли и соответствующей словарной метрики.

В работе [5] описаны кольца вычетов с циклической группой обратимых элементов, у которых мультипликативная полугруппа допускает планарный граф Кэли. С течением времени в материалах работы научного Алгебраического семинара ОмГПУ [6] накоплено большое количество фактов и примеров, связанных с кольцами вычетов, графами Кэли их мультипликативных полугрупп, планарностью этих графов. И теперь мы можем дать описание всех колец вычетов, у которых граф Кэли мультипликативной полугруппы допускает плоскую укладку.

1. Предварительные сведения

Графом Кэли полугруппы S относительно множества образующих её элементов X называется ориентированный мультиграф $\mathrm{Cay}(S, X)$, множество вершин которого равно S, а рёбра начинаются в вершине $a \in S$, заканчиваются в вершине $b \in S$ и помечены элементом $x \in X$ тогда и только тогда, когда $ax = b$. *Основой* ориентированного мультиграфа с помеченными рёбрами называется обыкновенный граф, полученный из исходного путём удаления всех меток и петель и заменой всех дуг, соединяющих одни и те же вершины в некотором направлении, одним ребром, соединяющим те же вершины. Говорим, что *полугруппа допускает планарный граф* Кэли, если основа её графа Кэли относительно некоторого множества образующих элементов является планарным графом, то есть может быть отображена на плоскость так, что вершинам графа соответствуют некоторые точки плоскости, а рёбрам графа соответствуют непрерывные плоские линии без самопересечений, не имеющие общих точек, кроме, возможно, общих вершин. Пара графов *гомеоморфны* друг другу, если они получены из одного и того же графа путём подразбиения его рёбер. Согласно критерию Понтрягина — Куратовского, *граф планарен* тогда и только тогда, когда он не содержит подграфов, гомеоморфных полному графу пятого порядка K_5 или полному двудольному графу $K_{3,3}$, содержащему по три вершины в каждой из долей. Другими словами, планарный граф характеризуется запрещёнными минорами K_5 и $K_{3,3}$. *Внешнепланарный граф* — такой граф, который допускает плоскую укладку, в которой все вершины принадлежат внешней грани. Внешнепланарные графы можно охарактеризовать двумя запрещёнными минорами K_4 и $K_{2,3}$.

Описание допускающих плоские графы Кэли конечных коммутативных групп (как и некоммутативных) известно давно. Пусть C_n — мультипликативная циклическая группа порядка n. Заметим, что $C_m \times C_n \cong C_{mn}$ при взаимно простых m и n. Из результатов работы [7, § 3, следствие 3] следует

Теорема 1. Конечная коммутативная группа G планарна тогда и только тогда, когда $G = C_2 \times C_{2n}$, или $G = C_2 \times C_2 \times C_2$, или $G = C_n$, где $n \in \mathbb{N}$.

Заметим, что справедливость этой теоремы следует и из более общего результата [8, § 5], который гласит: конечная группа G планарна тогда и только тогда, когда $G = G_1 \times G_2$, где $G_1 = C_1$ или C_2, а $G_2 = C_n$, D_n, S_4, A_4 или A_5. Здесь $D_n = \langle x, y \mid x^n = y^2 = (xy)^2 = 1 \rangle$ — диэдральная группа, элементами D_n являются автоморфизмы графа, состоящего только из цикла с n вершинами, таким образом, $|D_n| = 2n$; A_n и S_n — знакопеременная и симметрическая группы подстановок степени n.

Напомним, что *вычетом* целого числа a по модулю n называется остаток от деления a на n и обозначается через $a \bmod n$. Мультипликативная полугруппа $\mathbb{Z}_n$ вычетов по модулю n — это множество $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ всех вычетов по модулю n с операцией умножения: $a \cdot b = ab \bmod n$. В частности, $a \cdot b = ab$, если $ab < n$. Обозначим через $\mathbb{Z}_n^{-1}$ группу всех обратимых в Z_n элементов; она состоит из всех элементов полугруппы $\mathbb{Z}_n$, взаимно простых с числом n, и только из таких. Поэтому её порядок $|Z_n^{-1}|$ равен $\varphi(n)$, где φ — функция Эйлера.

Подполугруппу S_1 полугруппы S называют *выпуклой* (*фильтром*), если из того, что $x, y \in S$ и $x \cdot y \in S_1$, следует $x, y \in S_1$, то есть произведение элементов из S может принадлежать S_1 только тогда, когда эти элементы уже лежат внутри S_1.

Очевидна следующая лемма, так как только произведение взаимно простых с n чисел даёт взаимно простое с n число.

Лемма 1. Группа обратимых элементов $\mathbb{Z}_n^{-1}$ является выпуклой подполугруппой в полугруппе $\mathbb{Z}_n$.

Из определения выпуклой подполугруппы следует

Лемма 2. Пусть S_1 является выпуклой подполугруппой в полугруппе S. Тогда в любой системе образующих полугруппы S обязана присутствовать в качестве подсистемы система образующих полугруппы S_1.

Следствие 1. Всякая система образующих мультипликативной полугруппы вычетов $\mathbb{Z}_n$ содержит систему образующих подгруппы $\mathbb{Z}_n^{-1}$.

Нам потребуется ещё одно вспомогательное утверждение:

Лемма 3. Необходимым условием планарности графа Кэли коммутативной конечной полугруппы является не только планарность, но и внешнепланарность графа Кэли всякой её собственной выпуклой подполугруппы.

Доказательство. Заметим, что граф Кэли коммутативной полугруппы является связным. Пусть S_1 — собственная выпуклая подполугруппа конечной полугруппы S. Нетрудно понять, что $S \setminus S_1$ — непустой идеал полугруппы S, содержащий хотя бы один образующий полугруппы S. Рассмотрим граф Кэли полугруппы S. В качестве подграфа он содержит граф Кэли её собственной выпуклой подполугруппы S_1. Если предположить, что граф Кэли подполугруппы S_1 при любой раскладке не становится внешнепланарным, то часть графа Кэли полугруппы S, соответствующая идеалу $S\setminus S_1$, чтобы не нарушить планарность, должна расположиться на одной грани графа подполугруппы S_1. Но это не спасает планарность графа Кэли полугруппы S, так как всякая вершина графа Кэли подполугруппы S_1 смежна хотя бы с одной вершиной из $S \setminus S_1$. Значит, граф Кэли её собственной выпуклой подполугруппы S_1 должен быть внешнепланарным. ■

Из следствия 1 и лемм 1 и 3 получаем

Следствие 2. Необходимым условием планарности графа Кэли мультипликативной полугруппы вычетов $\mathbb{Z}_n$ является внешнепланарность графа Кэли её подгруппы $\mathbb{Z}_n^{-1}$.

Результатом этого следствия и теоремы 1 является

Лемма 4. Если граф Кэли мультипликативной полугруппы вычетов $\mathbb{Z}_n$ является планарным, то $\mathbb{Z}_n^{-1} = C_2 \times C_{2k}$, или $\mathbb{Z}_n^{-1} = C_2 \times C_2 \times C_2$, или $\mathbb{Z}_n^{-1} = C_m$, где n — произвольное натуральное число; m и $2k$ — порядки соответствующих циклических подгрупп.

Теорема 2 [9, теорема 1]. Конечный моноид S, являющийся произведением неодноэлементных циклических моноидов, допускает внешнепланарный граф Кэли тогда и только тогда, когда выполняется хотя бы одно из следующих условий:

1) $S = \langle a \mid a^3 = a\rangle \times \left\langle b \mid b^{h+t} = b^h\right\rangle^1$, где для натуральных h, t имеют место неравенства $h \leqslant 2$ и $h + t \leqslant 4$;
2) $S = \langle a \mid a^{1+m} = a\rangle^{+1} \times \left\langle b \mid b^{h+t} = b^h\right\rangle^i$, где $i \in \{1, +1\}$ и для натуральных m, h, t выполняется одно из следующих ограничений:
   2.1. $m = 1$, $t \leqslant 2$;
   2.2. $i = 1$, $m \leqslant 2$, $h = 1$, $t = 2$;
   2.3. $i = 1$, $m = 2$, $h = 1$, $t \leqslant 2$;
3) $S = \left\langle a_0 \mid a_0^{r+m} = a_0^r\right\rangle^1 \times \prod\limits_{i=1}^{n} \langle a_i \mid a_i^2 = a_i\rangle^{+1}$, где для натуральных r, n, m выполнено $n = m = 1$, или $n - 1 = r = m = 1$, или $n = m - 1 = 1$.

Нетрудные рассуждения, опирающиеся на эту теорему и на [10, теорема 3.1] для $C_2 \times C_2$, дают следствие 3. Более того, если $\mathbb{Z}_n^{-1} \cong C_2 \times C_{2m}$, то, в соответствии с леммой 3, граф Кэли полугруппы $\mathbb{Z}_n^{-1}$ должен быть внешнепланарным. По [10, теорема 3.1] приходим к выводу, что такая полугруппа допускает внешнепланарный граф Кэли только при $m = 1$.

Следствие 3. Среди групп $C_2 \times C_{2n}$, $C_2 \times C_2 \times C_2$ и C_n внешнепланарный граф Кэли имеют только группы C_n или $C_2 \times C_2$, где n — произвольное натуральное число.

Нам также понадобится

Теорема 3 [5, теорема 3]. Граф Кэли мультипликативной полугруппы вычетов $\mathbb{Z}_n$ с циклической группой $\mathbb{Z}_n^{-1}$ планарен тогда и только тогда, когда $n = 4$, или $n = 6$, или n — простое число.

2. Основной результат

Теорема 4. Мультипликативная полугруппа вычетов $\mathbb{Z}_n$ допускает планарный граф Кэли тогда и только тогда, когда $n \in \{4, 6, 8\} \cup P$, где P — множество всех простых чисел.

Доказательство. Пусть $\mathbb{Z}_n$ — мультипликативная полугруппа классов вычетов по модулю n с планарным графом Кэли. Тогда по следствию 2, лемме 4 и следствию 3 подгруппа $\mathbb{Z}_n^{-1}$ равна C_n или $C_2 \times C_2$, где n — произвольное натуральное число.

Пусть сначала $\mathbb{Z}_n^{-1} = C_n$. Тогда попадаем в условие теоремы 3, и в этом случае $n = 4$, или $n = 6$, или n — простое число.

Пусть теперь $\mathbb{Z}_n^{-1} = C_2 \times C_2$. Отсюда $|\mathbb{Z}_n^{-1}| = \varphi(n) = 4$. Решив уравнение $\varphi(n) = 4$ относительно n, получаем, что $n \in \{5, 8, 10, 12\}$. Проверка убеждает, что $\mathbb{Z}_5^{-1} = C_4$. Группа $\mathbb{Z}_5^{-1}$ — циклическая группа и подходит по теореме 3. Группа $\mathbb{Z}_8^{-1}$ изоморфна группе $C_2 \times C_2$. Планарность её графа Кэли можно проверить, посмотрев на рис. 1. Группа $\mathbb{Z}_{10}^{-1}$ изоморфна C_4. Отсеиваем её по теореме 3. Группа $\mathbb{Z}_{12}^{-1} \cong C_2 \times C_2$. Убедимся в непланарности графа Кэли соответствующей ей полугруппы Z_{12} ниже.

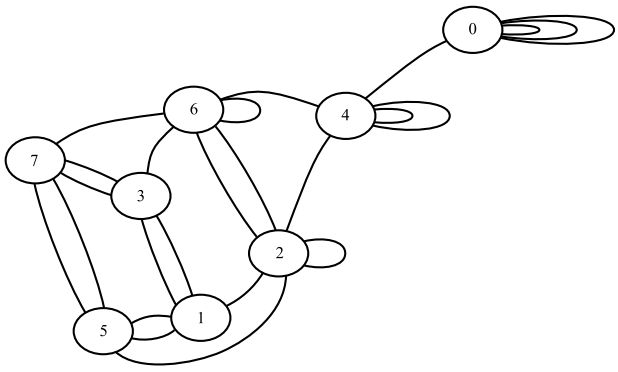

Рис. 1. Cay($\mathbb{Z}_8, \{2, 3, 5\}$)

Рассмотрим все возможные варианты общего случая, когда $\mathbb{Z}_n^{-1} \cong C_2 \times C_2$. В этом случае в $\mathbb{Z}_n^{-1}$ имеется цикл из не менее чем четырёх элементов. Наличие подграфа $K_{3,3}$ в основе графа Кэли мультипликативной полугруппы классов вычетов по модулю n, приводящее к потере планарности, обеспечивается тремя пунктами. Во-первых, нужно, чтобы в подгруппе обратимых был цикл из не менее чем четырёх элементов, что обеспечивается в данном случае теоремой Машке. Во-вторых, нужны два разных образующих из идеала $\mathbb{Z}_n \setminus \mathbb{Z}_n^{-1}$. Эти образующие, будучи возведёнными в некоторые

степени, породят идемпотенты, а общее число идемпотентов в полугруппе $\mathbb{Z}_n$ по составному модулю $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$, где p_i — попарно различные простые числа, равно 2^k. Заметим, что при составном модуле $k > 1$, следовательно, ниже групповой части в решётке полугруппы $\mathbb{Z}_n$ имеются ещё хотя бы два идемпотента, $e_1 = p_1^{\beta_{11}} p_2^{\beta_{12}} \cdots p_k^{\beta_{1k}} z_1$ и $e_2 = p_1^{\beta_{21}} p_2^{\beta_{22}} \cdots p_k^{\beta_{2k}} z_2$, где $z_i \in \mathbb{Z}_n^{-1}$, $\beta_{ij} \in \{0, \alpha_j\}$ при $1 \leqslant i \leqslant 2$ и $1 \leqslant j \leqslant k$, принадлежащих классам $\bar{e_1}$ и $\bar{e_2}$ соответственно. И в-третьих, между какими-либо из вершин тех разных двух классов нужен маршрут. Последнее выполняется, так как, в частности, элемент 0 обязательно должен порождаться произведением образующих либо сам находится среди образующих, но и тогда умножение на 0 сформирует ребро, связывающее 0 с любой другой вершиной. Таким образом обосновывается наличие подграфа $K_{3,3}$ на вершинах $\{1, ab, d\}$ в одной доле и вершинах $\{a, b, c\}$ в другой доле, где $\{1, a, b, ab\} \subseteq \mathbb{Z}_n^{-1}$, $c \in \bar{e_1}$, $d \in \bar{e_2}$, следовательно, граф Кэли рассматриваемой полугруппы не является планарным.

Группа D_2 изоморфна четверной группе Клейна, которая встречается во многих разделах математики, например, ей изоморфна приведённая система вычетов по модулю 8, состоящая из классов 1, 3, 5, 7, и приведённая система вычетов по модулю 12, состоящая из классов 1, 5, 7, 11. Графы Кэли планарной полугруппы $\mathbb{Z}_8$ и непланарной полугруппы $\mathbb{Z}_{12}$ представлены на рис. 1 и 2 соответственно.

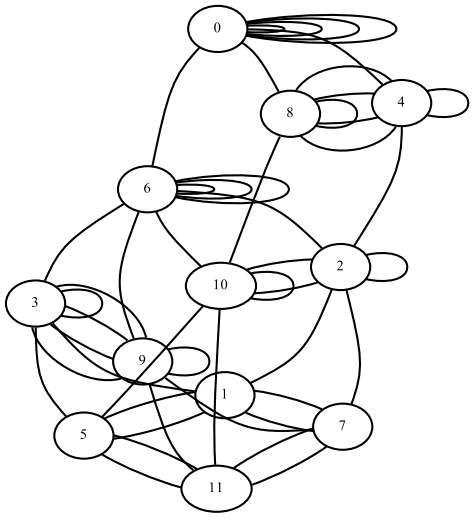

Рис. 2. Cay$(\mathbb{Z}_{12}, \{2, 3, 5, 7\})$

В доказательстве этой теоремы можно обойтись без теоремы 2 и следствия 3. Действительно, в общем случае непланарность графа Кэли полугруппы $\mathbb{Z}_n$, имеющей D_2 в качестве подгруппы, обеспечивается, если основа этого графа содержит подграф, гомеоморфный полному двудольному графу $K_{3,3}$, содержащему по три вершины в каждой из долей. В самом деле, если $D_2 = \langle a, b \rangle$, что выполнено при взаимно простых с модулем образующих, $(a, n) = 1 = (b, n)$, то основа графа Кэли Cay$(\mathbb{Z}_n, X)$, где $\{a, b\} \subseteq X$, содержит простой цикл на вершинах $\{a, b, 1, ab\}$, причём $a^2 = b^2 = 1$. Например, в $\mathbb{Z}_8$ можно видеть $a = 3$, $b = 5$, $ab = 7$, а в $\mathbb{Z}_{12}$ этот цикл содержит вершины $a = 5$, $b = 7$, $ab = 11$. Далее, каждая вершина этого цикла соединена с вершинами из смежных классов $cD_2 = \{cv : v \in D_2\}$ и $dD_2 = \{dv : v \in D_2\}$, где $(c, n) \neq 1 \neq (d, n)$. При $c = d = 2$ в $\mathbb{Z}_8$ наблюдается совпадение классов $cD_2 = dD_2 = \{2, 6\}$, а при $c = 2 \neq 3 = d$ в $\mathbb{Z}_{12}$ есть $cD_2 = \{2, 10\} \neq \{3, 9\} = dD_2$. Наконец, каждый элемент из cD_2 соединён с 0 маршрутом, имеющим общую вершину с маршрутом от элемен-

тов из dD_2 до 0, поэтому между вершинами подграфов cD_2 и dD_2 существует путь. В примере с $\mathbb{Z}_{12}$ таким путём может быть выбрана простая цепь на множестве вершин $\{2, 6, 3\}$.

Таким образом, подграф, гомеоморфный графу $K_{3,3}$, формируется в основе графа Кэли полугруппы $\mathbb{Z}_n$, содержащей подгруппу D_2, на вершинах $\{d, ab, 1\}$ в одной своей доле и вершинах $\{c, a, b\}$ в другой доле, где $c \neq d$ — необратимые образующие полугруппы $\mathbb{Z}_n$, а обратимыми её образующими являются $a \neq b$. В случае, когда $c = d$, из комбинаторных рассуждений можно получить, что в полугруппе $\mathbb{Z}_n$, имеющей D_2 в качестве подгруппы, содержится восемь элементов: $\{1 = a^2 = b^2, a, b, ab, 1c = bc, ac = abc, c^2, c^3\}$, ведь если $1c \neq bc$, то возможно $1c = ac$, тогда получим ту же полугруппу с точностью до изоморфизма, меняющего местами a и b, а если $1c \neq bc$ и $1c \neq ac$, то в графе Кэли возникает конфигурация $K_{3,3}$, описанная выше в предположении, что отделившийся элемент $d \neq c$ — это $d = bc$ или $d = ac$. Аналогична ситуация с $\mathbb{Z}_n^{-1} = \mathbb{Z}_2 \times D_2$. Для обнаружения $K_{3,3}$ в этом случае берутся a и b образующими из подгруппы $\{e\} \times D_2$, где e — нейтральный элемент группы $\mathbb{Z}_2$.

Учитывая теорему 3, приходим к выводу, что единственной имеющей нециклическую группу обратимых элементов $\mathbb{Z}_n^{-1}$ мультипликативной полугруппой вычетов $\mathbb{Z}_n$, допускающей планарный граф Кэли, является полугруппа $\mathbb{Z}_8$. ■

Заметим, что в силу следствия 3 при доказательстве теоремы не было необходимости рассматривать случай полугруппы $C_2 \times C_2 \times C_2$, который приводит нас к $n = 24$, так как в этом случае в полугруппе $\mathbb{Z}_n$ подгруппа $\mathbb{Z}_n^{-1}$ должна иметь мощность 8. Получается, что $\varphi(n) = 8$, тогда $n = \varphi^{-1}(8) \in \{15, 16, 20, 24, 30\}$. Принимая во внимание нижнюю оценку функции Эйлера $\varphi(n) \geqslant \sqrt{n}$, перечислены все возможные значения. При этом для модуля $n \in \{15, 16, 20, 30\}$ группа обратимых элементов является группой $C_2 \times C_4$, а для $n = 24$ имеем $\mathbb{Z}_n^{-1} \cong C_2 \times C_2 \times C_2$ и непланарный граф Кэли относительно любой системы образующих, в частности, представленный на рис. 3 относительно одной из таких систем.

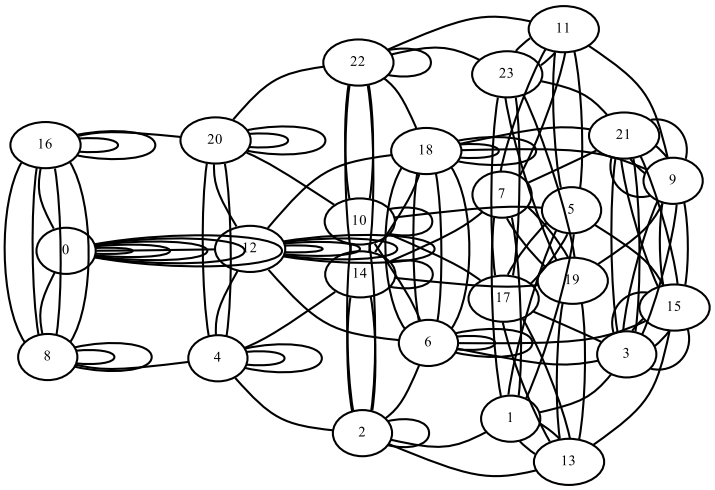

Рис. 3. Cay($\mathbb{Z}_{24}, \{2, 3, 5, 7, 13\}$)

Заключение

Представленный на рис. 1 планарный граф Кэли полугруппы $\mathbb{Z}_8$ — не единственный возможный планарный граф этой полугруппы. Плоская укладка существует относительно таких минимальных множеств образующих, как $\langle 2, 3, 5\rangle = \{2, 4, 0\}\cdot\{3, 1\}\cdot\{5, 1\}$,

$\langle 3,5,6\rangle = \{3,1\}\cdot\{5,1\}\cdot\{6,4,0\}$, $\langle 2,5,7\rangle = \{2,4,0\}\cdot\{5,1\}\cdot\{7,1\}$, $\langle 5,6,7\rangle = \{5,1\}\cdot\{6,4,0\}\cdot$ $\cdot\,\{7,1\}$. При этом для оставшихся из всех возможных минимальных множеств образующих ($\langle 2,3,7\rangle = \{2,4,0\}\cdot\{3,1\}\cdot\{7,1\}$, $\langle 3,6,7\rangle = \{3,1\}\cdot\{6,4,0\}\cdot\{7,1\}$) граф Кэли полугруппы $\mathbb{Z}_8$ не является планарным. Ознакомиться со всеми графами Кэли мультипликативных полугрупп классов вычетов по модулю до 24 включительно относительно всех возможных систем их образующих можно в материалах работы научного Алгебраического семинара ОмГПУ на сайте [6].

В данной работе описаны мультипликативные полугруппы вычетов с планарными графами Кэли. Полугруппы изоморфны, если их графы Кэли изоморфны, а для планарных графов проблема изоморфизма решается за полиномиальное время. То есть из представленных результатов следует существование полиномиальных алгоритмов для проблемы изоморфизма рассматриваемых полугрупп. Для описанных полугрупп оказалось, что проблема изоморфизма решается за время $O(1)$, достаточно сравнить модули, по которым формируются классы вычетов, являющиеся элементами анализируемых полугрупп. Но это лишь первый шаг, в других классах конечных полугрупп проблема не столь тривиальна.

ЛИТЕРАТУРА

1. *Hopcroft J. E. and Wong J. K.* Linear time algorithm for isomorphism of planar graphs (Preliminary Report) // Proc. STOC'74. Seattle, Washington, USA, 1974. P. 172–184.
2. *Kelarev A. V.* Labelled Cayley graphs and minimal automata // Australas. J. Comb. 2004. V. 30. P. 95–102.
3. *Kelarev A., Ryan J., and Yearwood J.* Cayley graphs as classifiers for data mining: the influence of asymmetries // Discrete Math. 2009. V. 309(17). P. 5360–5369.
4. *Zhu Y.* Generalized Cayley graphs of semigroups I // Semigroup Forum. 2012. V. 84. P. 131–143.
5. *Мартынов Л. М., Соломатин Д. В.* Полугруппы вычетов с циклическими группами обратимых элементов, допускающие планарные графы Кэли // Вестник Омского университета. 2012. № 2. С. 57–62.
6. `https://school.omgpu.ru/course/view.php?id=2219` — Научный Алгебраический семинар ОмГПУ на базе математического кружка «Мир Математики». 2024.
7. *Беленкова Ж. Т., Романьков В. А.* Плоские графы Кэли конечных групп. Препринт. Омск: ОмГУ, 1997. 8 с.
8. *Maschke H.* The representation of finite groups, especially of the rotation groups of the regular bodies of three- and four-dimensional space, by Cayley's color diagrams // Amer. J. Math. 1896. V. 18. No. 2. P. 156–194.
9. *Соломатин Д. В.* Прямые произведения циклических моноидов, допускающие внешнепланарные графы Кэли и их обобщения // Вестник Тверского государственного университета. Сер. Прикладная математика. 2023. № 4. С. 43–56.
10. *Соломатин Д. В.* Строение полугрупп, допускающих внешнепланарные графы Кэли // Сиб. электрон. матем. изв. 2011. Т. 8. С. 191–212.

REFERENCES

1. *Hopcroft J. E. and Wong J. K.* Linear time algorithm for isomorphism of planar graphs (Preliminary Report). Proc. STOC'74, Seattle, Washington, USA, 1974, pp. 172–184.
2. *Kelarev A. V.* Labelled Cayley graphs and minimal automata. Australas. J. Comb., 2004, vol. 30, pp. 95–102.

3. *Kelarev A., Ryan J., and Yearwood J.* Cayley graphs as classifiers for data mining: the influence of asymmetries. Discrete Math, 2009, vol. 309(17), pp. 5360–5369.
4. *Zhu Y.* Generalized Cayley graphs of semigroups I. Semigroup Forum, 2012, vol. 84, pp. 131–143.
5. *Martynov L. M. and Solomatin D. V.* Polugruppy vychetov c tsiklicheskimi gruppami obratimykh elementov, dopuskayushchie planarnye grafy Keli [Semigroups of residues with cyclic groups of invertible elements admitting planar Cayley graphs]. Herald of Omsk University, 2012, vol. 2, pp. 57–62. (in Russian)
6. `https://school.omgpu.ru/course/view.php?id=2219` — Scientific Algebraic Seminar of Omsk State Pedagogical University, 2024.
7. *Belenkova Zh. T. and Roman'kov V. A.* Ploskie grafy Keli konechnykh grupp [Plane Cayley Graphs of Finite Groups]. Preprint, Omsk, OmSU, 1997. 8 p. (in Russian)
8. *Maschke H.* The representation of finite groups, especially of the rotation groups of the regular bodies of three- and four-dimensional space, by Cayley's color diagrams. Amer. J. Math., 1896, vol. 18, no. 2, pp. 156–194.
9. *Solomatin D. V.* Pryamye proizvedeniya ciklicheskih monoidov, dopuskayushchie vneshneplanarnye grafy Keli i ih obobshcheniya [Direct products of cyclic monoids admitting outerplanar Cayley graphs and their generalizations]. Herald of Tver State University, Ser. Appl. Math., 2023, no. 4, pp. 43–56. (in Russian)
10. *Solomatin D. V.* Stroenie polugrupp, dopuskayushchikh vneshneplanarnye grafy Keli [Semigroups with outerplanar Cayley graphs]. Sib. Elektron. Mat. Izv., 2011, vol. 8, pp. 191–212. (in Russian)

УДК 519.719.2 DOI 10.17223/20710410/66/5

ОРТОМОРФИЗМЫ ГРУПП С МИНИМАЛЬНО ВОЗМОЖНЫМИ ПОПАРНЫМИ РАССТОЯНИЯМИ

С. В. Спиридонов

Лаборатория ТВП, г. Москва, Россия

E-mail: SpiridonovSV00@yandex.ru

Изучаются ортоморфизмы групп, находящиеся на минимально возможном расстоянии друг от друга по метрике Кэли. Описан класс преобразований, переводящих произвольный заданный ортоморфизм в множество всех ортоморфизмов, находящихся от исходного на минимально возможном расстоянии Кэли, равном двум. С помощью спектрально-разностного метода построения подстановок над обобщённой группой кватернионов Q_{4n}, $4n = 2^t$ $(t = 4, \ldots, 8)$, найдены ортоморфизмы с близкими к оптимальным значениями разностных характеристик.

Ключевые слова: *ортоморфизм, латинский квадрат, ортогональные латинские квадраты, метрика Кэли, s-бокс, нелинейное преобразование, подстановка, обобщённая группа кватернионов.*

ORTHOMORPHISMS OF GROUPS WITH MINIMAL POSSIBLE PAIRWISE DISTANCES

S. V. Spiridonov

TVP Laboratory, Moscow, Russia

Orthomorphisms of groups, which are at the minimum possible distance from each other according to the Cayley metric are studied. A class of transformations is described that map an arbitrary given orthomorphism into the set of all orthomorphisms that are at the minimum possible Cayley distance of two from the original. Using the spectral-difference method for constructing substitutions over the generalized quaternion group Q_{4n}, where $4n = 2^t$ $(t = 4, \ldots, 8)$, orthomorphisms with values of difference characteristics close to optimal have been found.

Keywords: *orthomorphism, Latin square, orthogonal Latin squares, Cayley metric, s-box, nonlinear transformation, substitution, generalized quaternion group.*

Введение

Понятие ортоморфизма впервые введено в работах [1, 2]. В терминах вполне перестановочных многочленов поля $\mathbb{F}_q$ ортоморфизмы детально изучались в [3, 4], некоторые подходы к построению ортоморфизмов приведены в [5, 6]. Ортоморфизмы активно изучались в 50–60-е годы XX в. в связи с некоторыми техническими приложениями как частный случай «проблемы параллельных перепаек» [7]. Ортоморфизмы находят широкое применение во многих криптографических конструкциях [8, 9]. Их изучение тесно связано с задачами построения кодов аутентификации [10, 11], систем ортогональных латинских квадратов [12–14] и квазигрупп [15, 16]. Необходимость развития методов построения ортоморфизмов над произвольными группами возникает в связи

с появлением ряда работ о неабелевых группах наложения ключа и их использовании при синтезе шифрсистем [17–19].

В данной работе результаты [6] обобщаются на произвольную группу G. Предложены алгоритмы, позволяющие для заданного ортоморфизма получить все ортоморфизмы, находящиеся на минимально возможном расстоянии от него. Описан класс преобразований, переводящих произвольный заданный ортоморфизм в множество всех ортоморфизмов, находящихся от исходного на минимально возможном расстоянии Кэли, равном двум. Полученные результаты иллюстрируются примерами над обобщённой группой кватернионов Q_{4n}.

Работа имеет следующую структуру. Пункт 1 содержит основные определения и обозначения, необходимые для дальнейшего изложения результатов. В п. 2 приведены утверждения, касающиеся свойств ортоморфизмов, находящихся на минимально возможном расстоянии Кэли друг от друга, и алгоритмы построения всех таких ортоморфизмов. Пункт 3 содержит результаты по построению ортоморфизмов с близкими к оптимальным значениями разностных характеристик.

1. Основные определения и обозначения

В работе используются следующие обозначения:

$(G, *)$	—	конечная группа с бинарной операцией $*$ и нейтральным элементом e;
$G^\times$	—	множество элементов группы G без нейтрального элемента e;
$S(G)$	—	симметрическая группа на множестве G;
A_m^k	—	число различных размещений из m элементов по k;
C_m^k	—	число различных сочетаний из m элементов по k;
Q_{4n}	—	обобщённая группа кватернионов порядка $4n$.

Определение 1. Подстановка $g \in S(G)$ называется *ортоморфизмом* группы G, если отображение $g' : G \to G$, определяемое условием $g'(x) = x^{-1} * g(x)$, где x^{-1} — элемент, обратный для $x \in G$ относительно операции $*$, является подстановкой из $S(G)$.

Множество всех ортоморфизмов группы G обозначим через $\mathrm{Orth}(G)$. На элементах группы G вводится произвольное отношение порядка:

$$G = \{e = z_0, z_1, \ldots, z_{n-1}\}, \quad |G| = n, \quad z_i < z_{i+1}, \quad i = 0, 1, \ldots, n-2.$$

Определение 2. *Расстоянием Кэли* между подстановками $g, h \in S(G)$ называется число

$$\tau(g, h) = \sum_{i=1}^{m} k_i(l_i - 1) = n - \sum_{i=1}^{m} k_i,$$

где k_i — число циклов длины l_i в разложении подстановки $h^{-1}g$ в произведение независимых циклов, т. е. цикловая структура подстановки $h^{-1}g$ имеет следующий вид:

$$[h^{-1}g] = \left[l_1^{k_1}, l_2^{k_2}, \ldots, l_m^{k_m}\right].$$

Нетрудно видеть, что $\tau(g, h)$ — это минимальное число транспозиций, переводящих подстановку g в h.

Определение 3. *Расстоянием Хемминга* между подстановками $g, h \in S(G)$ называется число

$$\chi(g, h) = |\{x \in G : g(x) \neq h(x)\}|.$$

Без существенных изменений для случая произвольной группы G справедливо утверждение из работы [6].

Утверждение 1. Если $g, h \in \mathrm{Orth}(G)$, $g \neq h$, то $\tau(g,h) \geqslant 2$.

Доказательство. Заметим, что для любых подстановок $g, h \in S(G)$, $g \neq h$, справедливо $\tau(g,h) \geqslant 1$. Пусть $\tau(g,h) = 1$. Тогда существуют такие $x_1, x_2 \in G$, что $x_1 \neq x_2$ и $h = (x_1, x_2)g$. Ортоморфизм h может быть записан в виде таблицы:

$$h = \begin{pmatrix} \dots & x_1 & \dots & x_2 & \dots \\ \dots & g(x_2) & \dots & g(x_1) & \dots \end{pmatrix}.$$

Так как $g \in \mathrm{Orth}(G)$, для построенной по определению 1 подстановки g' справедливо равенство $g'(x_1) = x_1^{-1} * g(x_1)$. Так как $h \in \mathrm{Orth}(G)$, то имеет место $g'(x_1) = x_1^{-1} * g(x_2)$ либо $g'(x_1) = x_2^{-1} * g(x_1)$. Следовательно,

$$\begin{cases} g'(x_1) = x_1^{-1} * g(x_1), \\ g'(x_1) = x_1^{-1} * g(x_2) \end{cases} \quad \text{либо} \quad \begin{cases} g'(x_1) = x_1^{-1} * g(x_1), \\ g'(x_1) = x_2^{-1} * g(x_1). \end{cases}$$

В первом случае имеем $g(x_1) = g(x_2)$ — противоречие, так как g — подстановка. Во втором случае $x_1 = x_2$ — противоречие условию $x_1 \neq x_2$. ■

Будем говорить, что ортоморфизмы $g, h \in \mathrm{Orth}(G)$, $g \neq h$, находятся на минимально возможном расстоянии друг от друга, если $\tau(g,h) = 2$.

Нетрудно видеть, что ортоморфизмы $g, h \in \mathrm{Orth}(G)$, находящиеся на расстоянии Кэли $\tau(g,h) = 2$, имеют расстояние Хемминга $\chi(g,h) = 3$ или $\chi(g,h) = 4$.

Пусть $I_i(g) = \{h \in \mathrm{Orth}(G) : \tau(g,h) = 2,\ \chi(g,h) = i+2\}$, $i = 1, 2$. Через $I(g)$ будем обозначать множество ортоморфизмов, находящихся на минимально возможном расстоянии Кэли от g, то есть

$$I(g) = \{h \in \mathrm{Orth}(G) : \tau(g,h) = 2\} = I_1(g) \cup I_2(g).$$

Изучение стойкости блочных шифрсистем с неабелевой группой наложения ключа относительно разностного метода криптоанализа может потребовать рассмотрения двух различных разностных характеристик перемешивающих отображений.

Определение 4. *Разностными характеристиками* $p_g^{(1)}$ и $p_g^{(2)}$ подстановки $g \in S(G)$ называются величины:

$$p_g^{(i)} = \max_{\alpha,\beta \in G^\times} {p_{\alpha,\beta}^g}^{(i)},\ i = 1, 2,$$

где

$${p_{\alpha,\beta}^g}^{(1)} = |G|^{-1} \cdot |\{x \in G : g(x)^{-1} * g(x * \alpha) = \beta\}|,$$
$${p_{\alpha,\beta}^g}^{(2)} = |G|^{-1} \cdot |\{x \in G : g(\alpha * x) * g(x)^{-1} = \beta\}|.$$

Замечание 1. Нетрудно видеть, что справедливы неравенства

$$p_g^{(1)} \leqslant 1,\ p_g^{(2)} \leqslant 1.$$

Верхняя оценка достижима, например, для тождественной подстановки.

Определение 5. *Обобщённой группой кватернионов* Q_{4n} с бинарной операцией $*$ и нейтральным элементом e называется неабелева конечная группа порядка $4n$, порождённая двумя элементами x и y:

$$\langle x, y \mid x^{2n} = y^4 = 1,\ x^n = y^2,\ y^{-1} * x * y = x^{-1} \rangle.$$

Из определения следует, что элементы Q_{4n} можно записать в виде

$$x^k y^j,\ 0 \leqslant k \leqslant 2n-1,\ j \in \{0,1\}.$$

Элементам Q_{4n} сопоставим элементы из $\mathbb{Z}_{4n}$ с помощью функции $\varphi : Q_{4n} \to \mathbb{Z}_{4n}$:

$$\varphi(x^k y^j) = 2n\, j + k.$$

Всюду далее элементы $z \in Q_{4n}$ обобщённой группы кватернионов будем записывать в виде их образа $\varphi(z) \in \mathbb{Z}_{4n}$.

2. Построение ортоморфизмов, находящихся на минимально возможном расстоянии от данного ортоморфизма

Приведём алгоритмы построения множеств $I_1(g)$ и $I_2(g)$ для произвольного ортоморфизма g, имеющие меньшую трудоёмкость по сравнению с наивным алгоритмом, основанным на переборе всех A_n^4 размещений.

2.1. Ортоморфизмы на расстоянии Хемминга, равном 3

Алгоритм 1 строит множества $I_1(g)$ для произвольного ортоморфизма g. Полагаем, что G — конечная группа порядка n, $n \geqslant 4$.

Алгоритм 1.

Вход: Ортоморфизм $g \in \mathrm{Orth}(G)$.
Выход: Список $I_1(g)$.

1: $i := 0$, $I_1(g) := \varnothing$.
2: **Если** $i = \mathrm{A}_{n-1}^2$, **то**
   закончить работу, на выход подать элементы списка I_1.
3: **Если** $i < \mathrm{A}_{n-1}^2$, **то**
4: выбрать новую упорядоченную пару $(x_1, x_2) \in G^2$ со свойством $\max\{x_1, x_2\} \neq z_{n-1}$;
5: $i := i + 1$;
6: перейти на шаг 7.
7: $x_3 := g(x_2) * g(x_1)^{-1} * x_1$.
8: **Если** $x_3 > \max\{x_1, x_2\}$, **то**
   перейти на шаг 9, **иначе** перейти на шаг 2.
9: $i := i + 1$.
10: **Если** $g(x_1)^{-1} * x_2 * x_1^{-1} * g(x_3) = e$ и $g(x_3)^{-1} * x_1 * x_2^{-1} * g(x_2) = e$, **то**
11: $h := (x_3, x_2)(x_2, x_1)g$; добавить h в список $I_1(g)$.
12: Перейти на шаг 2.

Пример 1. Рассмотрим ортоморфизм $g \in Q_{16}$, заданный табл. 1.

Таблица 1

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$g(x)$	0	2	1	5	8	a	9	d	3	c	6	f	e	7	b	4
$g'(x)$	0	1	7	2	c	d	b	e	9	5	8	4	6	a	3	f

Результатом применения алгоритма 1 к ортоморфизму $g \in \mathrm{Orth}(Q_{16})$ является множество $I_1(g)$, состоящее из четырёх ортоморфизмов (табл. 2).

Таблица 2

Ортоморфизм																Транспозиции
1	0	2	5	8	a	9	d	3	c	6	f	e	7	b	4	(0,1)(1,2)
4	2	1	5	8	a	9	d	3	c	6	0	e	7	b	f	(0,b)(b,f)
0	2	1	5	9	8	a	d	3	c	6	f	e	7	b	4	(4,5)(5,6)
0	2	1	5	8	e	9	d	a	c	6	f	3	7	b	4	(5,8)(8,c)

Следующее утверждение описывает свойства ортоморфизмов множества $I_1(g)$ и показывает корректность работы алгоритма 1.

Утверждение 2. Пусть $g \in \mathrm{Orth}(G)$, $h \in S(G)$ и существуют такие попарно различные элементы $x_1, x_2, x_3 \in G$, что $h = (x_3, x_2)(x_2, x_1)g$. Тогда следующие условия эквивалентны:

1) $h \in \mathrm{Orth}(G)$;
2) имеют место равенства

$$\begin{cases} g(x_1)^{-1} * x_2 * g'(x_3) = e, \\ g(x_2)^{-1} * x_3 * g'(x_1) = e, \\ g(x_3)^{-1} * x_1 * g'(x_2) = e. \end{cases} \tag{1}$$

Доказательство. Пусть h — ортоморфизм. Покажем, что выполнены равенства (1). Заметим, что для элементов $h'(x_1)$, $h'(x_2)$, $h'(x_3)$ справедливы следующие соотношения:

$$\begin{array}{lll} h'(x_1) \neq g'(x_1), & h'(x_2) \neq g'(x_2), & h'(x_3) \neq g'(x_3), \\ h'(x_1) \neq g'(x_3), & h'(x_2) \neq g'(x_1), & h'(x_3) \neq g'(x_2). \end{array}$$

Следовательно, $h'(x_1) = g'(x_2)$, $h'(x_2) = g'(x_3)$, $h'(x_3) = g'(x_1)$ и

$$\begin{array}{lll} h'(x_1) = x_1^{-1} * g(x_3), & h'(x_2) = x_2^{-1} * g(x_1), & h'(x_3) = x_3^{-1} * g(x_2), \\ h'(x_1) = x_2^{-1} * g(x_2), & h'(x_2) = x_3^{-1} * g(x_3), & h'(x_3) = x_1^{-1} * g(x_1). \end{array}$$

Справедливы равенства

$$\begin{aligned} e = h'(x_1)^{-1} * h'(x_1) = g(x_3)^{-1} * x_1 * g'(x_2), \\ e = h'(x_2)^{-1} * h'(x_2) = g(x_1)^{-1} * x_2 * g'(x_3), \\ e = h'(x_3)^{-1} * h'(x_3) = g(x_2)^{-1} * x_3 * g'(x_1). \end{aligned}$$

Обратно, пусть выполнены равенства (1). Тогда

$$\begin{aligned} h'(x_1) = x_1^{-1} * g(x_3) = g'(x_2), \\ h'(x_2) = x_2^{-1} * g(x_1) = g'(x_3), \\ h'(x_3) = x_3^{-1} * g(x_2) = g'(x_1). \end{aligned}$$

Следовательно, h — ортоморфизм. ■

Обозначим через t_1 трудоёмкость алгоритма 1.

Утверждение 3. При $n \to \infty$ для величины t_1 справедлива оценка $t_1 = O(n^2)$.

Доказательство. Трудоёмкость алгоритма 1 оценивается произведением числа A^2_{n-1} повторений шага 2 и фиксированного числа элементарных операций на шагах 7 и 9 алгоритма. ■

Замечание 2. Трудоёмкость алгоритма 1, по крайней мере, в n раз меньше трудоёмкости алгоритма, основанного на переборе всех A^3_n размещений элементов $x_1, x_2, x_3 \in G$, вычислении подстановки h и проверке свойства $h \in \mathrm{Orth}(G)$.

Из утверждения 3 следуют оценки числа ортоморфизмов в списке $I_1(g)$ на выходе алгоритма 1.

Следствие 1. Для любого $g \in \mathrm{Orth}(G)$ справедливы неравенства

$$0 \leqslant |I_1(g)| \leqslant \mathrm{A}^2_{n-1}.$$

Табл. 3 иллюстрирует достижимость нижних и верхних оценок числа $|I_1(g)|$ для обобщённой группы кватернионов Q_{4n}, $n = 2, 4, 8, 16, 32, 64$.

Таблица 3

Группа G	Нижняя оценка $\lvert I_1(g)\rvert$	Наименьшее найденное значение $\lvert I_1(g)\rvert$	Наибольшее найденное значение $\lvert I_1(g)\rvert$	Верхняя оценка $\lvert I_1(g)\rvert$
Q_8	0	0	0	42
Q_{16}	0	0	9	210
Q_{32}	0	0	7	930
Q_{64}	0	0	15	3906
Q_{128}	0	0	29	16002
Q_{256}	0	0	43	64770

2.2. Ортоморфизмы на расстоянии Хемминга, равном 4

Алгоритм 2 строит множество $I_2(g)$ для произвольного ортоморфизма g. Как и ранее, полагаем, что G — конечная группа порядка n, $n \geqslant 4$.

Пример 2. Рассмотрим ортоморфизм $g \in Q_8$, заданный табл. 4.

Таблица 4

x	0	1	2	3	4	5	6	7
$g(x)$	0	2	4	6	1	3	7	5
$g'(x)$	0	1	6	7	5	4	3	2

Результатом применения алгоритма 2 к ортоморфизму $g \in \mathrm{Orth}(Q_8)$ является множество $I_2(g)$, состоящее из восьми ортоморфизмов (табл. 5).

Таблица 5

Ортоморфизм								Транспозиции
2	0	4	6	1	3	5	7	$(0,1)(6,7)$
4	2	0	6	1	5	7	3	$(0,2)(5,7)$
6	4	2	0	1	3	7	5	$(0,3)(1,2)$
3	2	4	7	1	0	6	5	$(0,5)(3,6)$
7	2	4	3	1	6	0	5	$(0,6)(3,5)$
0	6	4	2	7	3	1	5	$(1,3)(4,6)$
0	2	6	4	3	1	7	5	$(2,3)(4,5)$
0	2	4	6	5	7	3	1	$(4,7)(5,6)$

Нетрудно видеть, что полученные ортоморфизмы задают латинские квадраты, ортогональные к таблице Кэли группы Q_8, например:

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & 3 & 0 & 5 & 6 & 7 & 4 \\ 2 & 3 & 0 & 1 & 6 & 7 & 4 & 5 \\ 3 & 0 & 1 & 2 & 7 & 4 & 5 & 6 \\ 4 & 7 & 6 & 5 & 2 & 1 & 0 & 3 \\ 5 & 4 & 7 & 6 & 3 & 2 & 1 & 0 \\ 6 & 5 & 4 & 7 & 0 & 3 & 2 & 1 \\ 7 & 6 & 5 & 4 & 1 & 0 & 3 & 2 \end{pmatrix}, \quad \begin{pmatrix} 4 & 2 & 0 & 6 & 1 & 5 & 7 & 3 \\ 7 & 3 & 1 & 5 & 2 & 4 & 6 & 0 \\ 6 & 0 & 2 & 4 & 3 & 7 & 5 & 1 \\ 5 & 1 & 3 & 7 & 0 & 6 & 4 & 2 \\ 2 & 6 & 4 & 0 & 5 & 3 & 1 & 7 \\ 1 & 7 & 5 & 3 & 6 & 2 & 0 & 4 \\ 0 & 4 & 6 & 2 & 7 & 1 & 3 & 5 \\ 3 & 5 & 7 & 1 & 4 & 0 & 2 & 6 \end{pmatrix}.$$

Алгоритм 2.

Вход: Ортоморфизм $g \in \mathrm{Orth}(G)$.
Выход: Список $I_2(g)$.

1: $i := 0$, $I_2(g) := \varnothing$.
2: **Если** $i = \mathrm{C}_n^3 - 1$, **то**
закончить работу, на выход подать элементы списка I_2.
3: **Если** $i < \mathrm{C}_n^3 - 1$, **то**
4: выбрать новое сочетание $(x_1, x_3, x_4) \in G^3$, удовлетворяющее условиям $x_1 < x_3 < x_4$, $x_1 \neq z_{n-3}$;
5: $i := i + 1$;
6: перейти на шаг 7.
7: $x_2 := g(x_1) * g(x_4)^{-1} * x_4$; $\widetilde{x}_2 := g(x_1) * g(x_3)^{-1} * x_3$.
8: **Если** $x_1 < x_2$ или $x_1 < \widetilde{x}_2$, **то**
перейти на шаг 9, **иначе** перейти на шаг 2.
9: Проверить справедливость систем равенств (2) и (3):

$$\begin{cases} \left[\begin{cases} g(x_4)^{-1} * x_3 * x_1^{-1} * g(x_1) = e, \\ g(x_3)^{-1} * x_4 * x_2^{-1} * g(x_2) = e, \end{cases} \right. \\ \left. \left[\begin{cases} g(x_4)^{-1} * x_3 * x_2^{-1} * g(x_2) = e, \\ g(x_3)^{-1} * x_4 * x_1^{-1} * g(x_1) = e, \end{cases} \right. \\ g(x_2)^{-1} * x_1 * x_3^{-1} * g(x_3) = e; \end{cases} \tag{2}$$

$$\begin{cases} \left[\begin{cases} g(x_4)^{-1} * x_3 * x_1^{-1} * g(x_1) = e, \\ g(x_3)^{-1} * x_4 * \widetilde{x}_2^{-1} * g(\widetilde{x}_2) = e, \end{cases} \right. \\ \left. \left[\begin{cases} g(x_4)^{-1} * x_3 * \widetilde{x}_2^{-1} * g(\widetilde{x}_2) = e, \\ g(x_3)^{-1} * x_4 * x_1^{-1} * g(x_1) = e, \end{cases} \right. \\ g(x_1)^{-1} * \widetilde{x}_2 * x_3^{-1} * g(x_3) = e. \end{cases} \tag{3}$$

10: **Если** выполнена система (2), **то**
11: для элемента x_2 вычислить $h = (x_4, x_3)(x_2, x_1)g$, добавить h в список $I_2(g)$.
12: **Если** выполнена система (3), **то**
13: для элемента $\widetilde{x}_2$ вычислить $\widetilde{h} = (x_4, x_3)(\widetilde{x}_2, x_1)g$, добавить $\widetilde{h}$ в список $I_2(g)$.
14: Перейти на шаг 2.

Следующее утверждение описывает свойства ортоморфизмов множества $I_2(g)$ и показывает корректность работы алгоритма 2.

Утверждение 4. Пусть $g \in \mathrm{Orth}(G), h \in S(G)$, существуют такие попарно различные элементы $x_1, x_2, x_3, x_4 \in G$, что $h = (x_4, x_3)(x_2, x_1)g$. Тогда следующие условия эквивалентны:

1) $h \in \mathrm{Orth}(G)$;
2) справедлива одна из систем равенств (4) или (5):

$$\begin{cases} \left[\begin{cases} g(x_4)^{-1} * x_3 * g'(x_1) = e, \\ g(x_3)^{-1} * x_4 * g'(x_2) = e, \end{cases} \\ \left[\begin{cases} g(x_4)^{-1} * x_3 * g'(x_2) = e, \\ g(x_3)^{-1} * x_4 * g'(x_1) = e, \end{cases} \\ g(x_1)^{-1} * x_2 * g'(x_4) = e, \\ g(x_2)^{-1} * x_1 * g'(x_3) = e; \end{cases} \tag{4}$$

$$\begin{cases} \left[\begin{cases} g(x_4)^{-1} * x_3 * g'(x_1) = e, \\ g(x_3)^{-1} * x_4 * g'(x_2) = e, \end{cases} \\ \left[\begin{cases} g(x_4)^{-1} * x_3 * g'(x_2) = e, \\ g(x_3)^{-1} * x_4 * g'(x_1) = e, \end{cases} \\ g(x_1)^{-1} * x_2 * g'(x_3) = e, \\ g(x_1)^{-1} * x_2 * g'(x_3) = e. \end{cases} \tag{5}$$

Доказательство. Пусть h — ортоморфизм. Покажем, что выполнено условие 2. Элементы $h'(x_1)$, $h'(x_2)$, $h'(x_3)$, $h'(x_4)$ удовлетворяют следующим условиям:

$$\begin{array}{llll} h'(x_1) \neq g'(x_1), & h'(x_2) \neq g'(x_2), & h'(x_3) \neq g'(x_3), & h'(x_4) \neq g'(x_3), \\ h'(x_1) \neq g'(x_2), & h'(x_2) \neq g'(x_3), & h'(x_3) \neq g'(x_4), & h'(x_4) \neq g'(x_4). \end{array}$$

Следовательно,

$$\big(h'(x_1), h'(x_2), h'(x_3), h'(x_4)\big) \in \{\big(g'(x_3), g'(x_4), g'(x_1), g'(x_2)\big), \big(g'(x_3), g'(x_4), g'(x_2), g'(x_1)\big), \\ \big(g'(x_4), g'(x_3), g'(x_1), g'(x_2)\big), \big(g'(x_4), g'(x_3), g'(x_2), g'(x_1)\big)\}.$$

Если $\big(h'(x_1), h'(x_2), h'(x_3), h'(x_4)\big) = \big(g'(x_3), g'(x_4), g'(x_1), g'(x_2)\big)$, то

$$\begin{array}{ll} \begin{cases} h'(x_1) = x_1^{-1} * g(x_2), \\ h'(x_1) = x_3^{-1} * g(x_3); \end{cases} & \begin{cases} h'(x_2) = x_2^{-1} * g(x_1), \\ h'(x_2) = x_4^{-1} * g(x_4); \end{cases} \\ \begin{cases} h'(x_3) = x_3^{-1} * g(x_4), \\ h'(x_3) = x_1^{-1} * g(x_1); \end{cases} & \begin{cases} h'(x_4) = x_4^{-1} * g(x_3), \\ h'(x_3) = x_2^{-1} * g(x_2), \end{cases} \end{array}$$

поэтому

$$\begin{aligned} e &= h'(x_1)^{-1} * h'(x_1) = g(x_2)^{-1} * x_1 * x_3^{-1} * g(x_3) = g(x_2)^{-1} * x_1 * g'(x_3), \\ e &= h'(x_2)^{-1} * h'(x_2) = g(x_1)^{-1} * x_2 * x_4^{-1} * g(x_4) = g(x_1)^{-1} * x_2 * g'(x_4), \\ e &= h'(x_3)^{-1} * h'(x_3) = g(x_4)^{-1} * x_3 * x_1^{-1} * g(x_1) = g(x_4)^{-1} * x_3 * g'(x_1), \\ e &= h'(x_4)^{-1} * h'(x_4) = g(x_3)^{-1} * x_4 * x_2^{-1} * g(x_2) = g(x_3)^{-1} * x_4 * g'(x_2). \end{aligned}$$

Аналогично рассматриваются оставшиеся три случая.

Обратно, пусть выполнены следующие равенства из п. 2:

$$g(x_2)^{-1} * x_1 * g'(x_3) = e,$$
$$g(x_1)^{-1} * x_2 * g'(x_4) = e,$$
$$g(x_4)^{-1} * x_3 * g'(x_1) = e,$$
$$g(x_3)^{-1} * x_4 * g'(x_2) = e.$$

Покажем, что $h \in \mathrm{Orth}(G)$.

Так как $h = (x_4, x_3)(x_2, x_1)g$, то элементы $h'(x_i)$, $i = 1, 2, 3, 4$, имеют вид

$$h'(x_1) = x_1^{-1} * g(x_2), \qquad h'(x_2) = x_2^{-1} * g(x_1),$$
$$h'(x_3) = x_3^{-1} * g(x_4), \qquad h'(x_4) = x_4^{-1} * g(x_3).$$

Из равенства $g(x_2)^{-1} * x_1 * g'(x_3) = e$ следует, что $g'(x_3) = x_1^{-1} * g(x_2) = h'(x_1)$.
Из равенства $g(x_1)^{-1} * x_2 * g'(x_4) = e$ следует, что $g'(x_4) = x_2^{-1} * g(x_1) = h'(x_2)$.
Из равенства $g(x_4)^{-1} * x_3 * g'(x_1) = e$ следует, что $g'(x_1) = x_3^{-1} * g(x_4) = h'(x_1)$.
Из равенства $g(x_3)^{-1} * x_4 * g'(x_2) = e$ следует, что $g'(x_2) = x_4^{-1} * g(x_3) = h'(x_4)$.
Следовательно, h — ортоморфизм.

Случаи выполнения других наборов равенств п. 2 утверждения 4 проверяются аналогично. ■

Обозначим через t_2 трудоёмкость алгоритма 2.

Утверждение 5. При $n \to \infty$ для величины t_2 справедлива оценка $t_2 = O(n^3)$.

Доказательство. Трудоёмкость алгоритма 2 оценивается произведением числа $(\mathrm{C}_n^3 - 1)$ повторений шага 2 и фиксированного числа элементарных операций на шагах 7 и 9 алгоритма. ■

Замечание 3. Трудоёмкость алгоритма 2, по крайней мере, в n раз меньше трудоёмкости алгоритма, основанного на переборе всех A_n^4 размещений элементов $x_1, x_2, x_3, x_4 \in G$, вычислении подстановки h и проверке свойства $h \in \mathrm{Orth}(G)$.

Из утверждения 5 следуют оценки числа ортоморфизмов в списке $I_2(g)$ на выходе алгоритма 2.

Следствие 2. Для любого $g \in \mathrm{Orth}(G)$ справедливы неравенства

$$0 \leqslant |I_2(g)| \leqslant \mathrm{C}_n^3 - 1.$$

В табл. 6 приведены данные о достижимости нижних и верхних оценок числа $|I_2(g)|$ для обобщённой группы кватернионов Q_{4n}, $n = 2, 4, 8, 16, 32, 64$.

Таблица 6

Группа G	Нижняя оценка $\lvert I_2(g)\rvert$	Наименьшее найденное значение $\lvert I_2(g)\rvert$	Наибольшее найденное значение $\lvert I_2(g)\rvert$	Верхняя оценка $\lvert I_2(g)\rvert$
Q_8	0	8	8	55
Q_{16}	0	2	31	559
Q_{32}	0	10	46	4959
Q_{64}	0	32	83	41663
Q_{128}	0	42	103	341375
Q_{256}	0	78	170	2763519

3. Экспериментальные результаты

Ортоморфизмы, описанные в [9, 18], обладают близкими к 1 значениями разностных характеристик. Примеры таких ортоморфизмов содержатся в табл. 7–11.

Алгоритмы 1 и 2 настоящей работы были использованы в составе спектрально-разностного метода [20, 21] для построения ортоморфизмов $g \in Q_{2^n}$ обобщённой группы кватернионов с близкими к оптимальным значениями разностных характеристик $p_g^{(1)}$ и $p_g^{(2)}$. В табл. 12–16 приведены примеры ортоморфизмов обобщённой группы кватернионов $g \in \mathrm{Orth}(Q_{4n})$, где $4n = 2^t$ $(t = 4, 5, 6, 7, 8)$, с близкими к оптимальным значениями разностных характеристик. Такие подстановки являются перспективными для использования в качестве нелинейных перемешивающих преобразований в блочных шифрсистемах с операцией наложения ключа из группы Q_{4n}, где $4n = 2^t$, $t \geqslant 3$.

Таблица 7

$g \in \mathrm{Orth}(Q_{16})$																$p_g^{(1)}$	$p_g^{(2)}$
0	2	4	6	8	a	c	e	1	3	5	7	d	f	9	b	12/16	12/16
0	2	4	6	c	e	8	a	9	b	d	f	1	3	5	7	12/16	12/16

Таблица 8

$g \in \mathrm{Orth}(Q_{32})$																$p_g^{(1)}$	$p_g^{(2)}$
0	2	4	6	8	a	c	e	10	12	14	16	18	1a	1c	1e	28/32	24/32
1	3	5	7	9	b	d	f	19	1b	1d	1f	11	13	15	17		
0	2	4	6	8	a	c	e	18	1a	1c	1e	10	12	14	16	28/32	24/32
11	13	15	17	19	1b	1d	1f	1	3	5	7	9	b	d	f		

Таблица 9

$g \in \mathrm{Orth}(Q_{64})$																$p_g^{(1)}$	$p_g^{(2)}$
0	2	4	6	8	a	c	e	10	12	14	16	18	1a	1c	1e	60/64	48/64
20	22	24	26	28	2a	2c	2e	30	32	34	36	38	3a	3c	3e		
1	3	5	7	9	b	d	f	11	13	15	17	19	1b	1d	1f		
31	33	35	37	39	3b	3d	3f	21	23	25	27	29	2b	2d	2f		
0	2	4	6	8	a	c	e	10	12	14	16	18	1a	1c	1e	60/64	48/64
30	32	34	36	38	3a	3c	3e	20	22	24	26	28	2a	2c	2e		
21	23	25	27	29	2b	2d	2f	31	33	35	37	39	3b	3d	3f		
1	3	5	7	9	b	d	f	11	13	15	17	19	1b	1d	1f		

Таблица 10

$g \in \mathrm{Orth}(Q_{128})$																$p_g^{(1)}$	$p_g^{(2)}$
0	2	4	6	8	a	c	e	10	12	14	16	18	1a	1c	1e		
20	22	24	26	28	2a	2c	2e	30	32	34	36	38	3a	3c	3e		
40	42	44	46	48	4a	4c	4e	50	52	54	56	58	5a	5c	5e		
60	62	64	66	68	6a	6c	6e	70	72	74	76	78	7a	7c	7e	$\frac{124}{128}$	$\frac{96}{128}$
1	3	5	7	9	b	d	f	11	13	15	17	19	1b	1d	1f		
21	23	25	27	29	2b	2d	2f	31	33	35	37	39	3b	3d	3f		
61	63	65	67	69	6b	6d	6f	71	73	75	77	79	7b	7d	7f		
41	43	45	47	49	4b	4d	4f	51	53	55	57	59	5b	5d	5f		
0	2	4	6	8	a	c	e	10	12	14	16	18	1a	1c	1e		
20	22	24	26	28	2a	2c	2e	30	32	34	36	38	3a	3c	3e		
60	62	64	66	68	6a	6c	6e	70	72	74	76	78	7a	7c	7e		
40	42	44	46	48	4a	4c	4e	50	52	54	56	58	5a	5c	5e	$\frac{124}{128}$	$\frac{96}{128}$
41	43	45	47	49	4b	4d	4f	51	53	55	57	59	5b	5d	5f		
61	63	65	67	69	6b	6d	6f	71	73	75	77	79	7b	7d	7f		
1	3	5	7	9	b	d	f	11	13	15	17	19	1b	1d	1f		
21	23	25	27	29	2b	2d	2f	31	33	35	37	39	3b	3d	3f		

Таблица 11

$g \in \mathrm{Orth}(Q_{256})$																$p_g^{(1)}$	$p_g^{(2)}$
0	2	4	6	8	a	c	e	10	12	14	16	18	1a	1c	1e		
20	22	24	26	28	2a	2c	2e	30	32	34	36	38	3a	3c	3e		
40	42	44	46	48	4a	4c	4e	50	52	54	56	58	5a	5c	5e		
60	62	64	66	68	6a	6c	6e	70	72	74	76	78	7a	7c	7e		
80	82	84	86	88	8a	8c	8e	90	92	94	96	98	9a	9c	9e		
a0	a2	a4	a6	a8	aa	ac	ae	b0	b2	b4	b6	b8	ba	bc	be		
c0	c2	c4	c6	c8	ca	cc	ce	d0	d2	d4	d6	d8	da	dc	de		
e0	e2	e4	e6	e8	ea	ec	ee	f0	f2	f4	f6	f8	fa	fc	fe	$\frac{252}{256}$	$\frac{192}{256}$
1	3	5	7	9	b	d	f	11	13	15	17	19	1b	1d	1f		
21	23	25	27	29	2b	2d	2f	31	33	35	37	39	3b	3d	3f		
41	43	45	47	49	4b	4d	4f	51	53	55	57	59	5b	5d	5f		
61	63	65	67	69	6b	6d	6f	71	73	75	77	79	7b	7d	7f		
c1	c3	c5	c7	c9	cb	cd	cf	d1	d3	d5	d7	d9	db	dd	df		
e1	e3	e5	e7	e9	eb	ed	ef	f1	f3	f5	f7	f9	fb	fd	ff		
81	83	85	87	89	8b	8d	8f	91	93	95	97	99	9b	9d	9f		
a1	a3	a5	a7	a9	ab	ad	af	b1	b3	b5	b7	b9	bb	bd	bf		
0	2	4	6	8	a	c	e	10	12	14	16	18	1a	1c	1e		
20	22	24	26	28	2a	2c	2e	30	32	34	36	38	3a	3c	3e		
40	42	44	46	48	4a	4c	4e	50	52	54	56	58	5a	5c	5e		
60	62	64	66	68	6a	6c	6e	70	72	74	76	78	7a	7c	7e		
c0	c2	c4	c6	c8	ca	cc	ce	d0	d2	d4	d6	d8	da	dc	de		
e0	e2	e4	e6	e8	ea	ec	ee	f0	f2	f4	f6	f8	fa	fc	fe		
80	82	84	86	88	8a	8c	8e	90	92	94	96	98	9a	9c	9e		
a0	a2	a4	a6	a8	aa	ac	ae	b0	b2	b4	b6	b8	ba	bc	be	$\frac{252}{256}$	$\frac{192}{256}$
81	83	85	87	89	8b	8d	8f	91	93	95	97	99	9b	9d	9f		
a1	a3	a5	a7	a9	ab	ad	af	b1	b3	b5	b7	b9	bb	bd	bf		
c1	c3	c5	c7	c9	cb	cd	cf	d1	d3	d5	d7	d9	db	dd	df		
e1	e3	e5	e7	e9	eb	ed	ef	f1	f3	f5	f7	f9	fb	fd	ff		
1	3	5	7	9	b	d	f	11	13	15	17	19	1b	1d	1f		
21	23	25	27	29	2b	2d	2f	31	33	35	37	39	3b	3d	3f		
41	43	45	47	49	4b	4d	4f	51	53	55	57	59	5b	5d	5f		
61	63	65	67	69	6b	6d	6f	71	73	75	77	79	7b	7d	7f		

Таблица 12

$g \in \mathrm{Orth}(Q_{16})$																$p_g^{(1)}$	$p_g^{(2)}$
c	f	3	6	e	a	0	4	1	9	5	7	d	2	8	b	3/16	3/16
c	0	8	6	4	e	b	5	1	3	7	a	f	9	2	d	3/16	3/16

Таблица 13

$g \in \mathrm{Orth}(Q_{32})$																$p_g^{(1)}$	$p_g^{(2)}$
10	b	a	5	1c	c	8	16	18	14	1a	4	15	13	d	1	4/32	4/32
3	2	0	e	9	12	1e	1f	17	f	1d	1b	11	7	19	6		
f	2	18	1c	1f	13	4	17	10	1	0	6	c	a	1b	16	4/32	4/32
1e	15	5	7	1d	b	d	3	14	e	11	9	12	1a	19	8		

Таблица 14

$g \in \mathrm{Orth}(Q_{64})$																$p_g^{(1)}$	$p_g^{(2)}$
20	2c	28	24	e	2	3a	5	35	38	21	16	17	31	a	14	4/64	5/64
11	23	2e	26	1c	32	6	2a	18	4	1	c	0	29	7	30		
34	3	1d	2f	2b	37	d	1f	22	12	15	3c	19	1b	3e	f		
39	8	33	3d	1a	3b	b	3f	9	13	25	27	36	1e	10	2d		
32	3f	28	24	e	2	23	5	2c	38	21	27	17	12	1c	1a	4/64	5/64
11	31	35	26	a	20	6	2a	18	4	1	c	0	29	7	30		
34	3	2e	2f	2b	3d	d	1f	22	3a	15	14	8	1b	3e	f		
39	b	33	37	3c	3b	19	1d	9	10	25	16	36	1e	13	2d		

Таблица 15

$g \in \mathrm{Orth}(Q_{128})$																$p_g^{(1)}$	$p_g^{(2)}$
7a	46	33	5e	56	42	c	20	10	71	44	52	15	79	1c	55		
4c	a	24	36	63	6a	41	6	30	39	7d	48	5f	73	3c	3e		
7c	2d	32	21	f	4a	54	2e	12	65	75	1f	18	40	5c	3		
60	49	25	5d	68	2f	6e	34	6b	47	2c	22	37	35	14	1a	$\frac{5}{128}$	$\frac{6}{128}$
31	7	4	4e	0	7b	d	8	11	74	66	4b	23	29	77	7e		
62	4d	6c	16	27	b	57	26	3f	13	19	72	61	3b	3d	1		
2a	2	5b	67	59	6f	3a	70	28	64	6d	51	5	2b	17	7f		
9	45	1d	43	1e	58	1b	4f	53	76	78	5a	69	50	e	38		
a	46	32	5e	56	72	9	20	6d	7a	6e	73	15	79	1c	5a		
68	39	4f	36	35	6a	3d	6	30	31	3c	2d	45	5d	4c	3e		
71	7d	65	62	f	48	6c	2e	12	54	74	7e	18	41	c	76		
60	49	25	66	44	2f	34	e	6b	22	2c	16	37	2	14	7	$\frac{5}{128}$	$\frac{6}{128}$
52	1f	2a	4e	0	7b	8	26	47	33	75	4b	23	29	51	11		
21	4d	43	d	27	b	70	1a	3f	13	19	4a	40	3b	1e	1		
4	5b	2b	67	59	6f	3a	5f	28	7c	42	77	5	61	17	57		
5c	64	1d	24	50	58	1b	7f	53	3	78	55	69	63	10	38		

Таблица 16

$g \in \mathrm{Orth}(Q_{256})$																$p_g^{(1)}$	$p_g^{(2)}$
3c	7a	49	fe	6d	45	e6	d9	b5	12	14	96	82	a2	d4	80		
41	2	8	e1	38	ca	c0	f8	67	d2	94	36	c	60	0	ef		
e8	8d	8c	25	b8	4a	cc	70	78	39	46	dc	d1	6a	5a	be		
64	f4	f1	90	f0	1	bd	43	1f	a9	13	a8	50	d8	4c	6e		
de	e4	c4	bf	24	2a	d0	93	cd	b1	e3	6	f3	52	61	e0		
6f	d5	99	1d	28	1a	10	a1	30	35	ab	8b	a5	9e	4d	ba		
54	22	68	44	c8	da	af	1e	9	56	47	d6	86	f6	16	2e		
5e	a	7	76	c9	c2	48	4e	84	32	b4	f2	2f	ac	fc	31	$\frac{5}{256}$	$\frac{7}{256}$
9c	eb	3a	26	23	21	5b	6b	a0	72	9f	4b	f5	e2	e	8e		
53	cb	15	c6	f	d7	2d	a4	91	fb	75	58	f9	7b	33	3f		
f7	7e	3	2b	d3	b	bc	b0	51	db	ad	cf	74	34	a6	8f		
8a	63	b3	85	e9	88	c1	9d	3d	73	b6	37	79	3b	7d	a7		
d	fd	fa	c7	65	1c	ec	19	55	b7	5	df	ce	a3	c3	7c		
9a	aa	66	71	b2	5f	b9	3e	42	e7	5c	27	89	dd	6c	77		
81	83	87	97	17	2c	ae	4f	29	ee	95	57	5d	11	98	40		
18	1b	4	7f	69	62	e5	ea	92	ff	c5	20	ed	bb	9b	59		
3c	7a	49	fe	6d	aa	e6	d9	b5	5b	14	96	82	a2	5f	80		
41	2	8	e1	38	ca	c0	f8	67	d2	94	36	c	60	0	ef		
e8	a1	8c	25	b8	4a	cc	70	78	39	6b	dc	92	6a	5a	be		
64	f4	f1	f	f0	1	bd	43	1f	a9	13	a8	50	d8	4c	6e		
de	e4	c4	59	24	2a	2c	93	cd	b1	e3	6	f3	52	61	e0		
6f	d5	99	1d	28	1a	10	8d	30	b6	ab	8b	a5	9e	4d	ba		
54	7d	68	44	c8	da	af	d3	9	56	47	d6	86	f6	16	2e		
5e	a	eb	76	c9	c2	7f	4e	84	32	b4	f2	2f	ac	fc	31	$\frac{5}{256}$	$\frac{7}{256}$
9c	77	3a	26	23	21	45	4b	a0	72	9f	d4	f5	e2	e	8e		
53	cb	15	c6	90	d7	2d	a4	91	fb	75	58	f9	7b	33	3f		
f7	7e	3	2b	65	b	bc	b0	51	db	ad	cf	74	34	a6	8f		
8a	63	57	85	e9	88	c1	9d	46	73	35	37	79	3b	22	a7		
d	fd	fa	c7	b3	1c	ec	19	55	b7	5	df	ce	a3	c3	7c		
9a	7	66	71	b2	3d	b9	3e	42	e7	5c	27	89	dd	4f	1e		
81	83	87	97	e5	d0	ae	6c	29	ee	95	12	5d	11	98	40		
18	1b	4	48	69	62	17	ea	d1	ff	c5	20	ed	bb	9b	bf		

Заключение

В работе изложены алгоритмы построения для произвольного ортоморфизма произвольной группы множеств $I_1(g)$ и $I_2(g)$, содержащих ортоморфизмы, находящиеся на минимально возможном расстоянии от исходного, приведены обоснование и трудоёмкость данных алгоритмов. Кроме этого, с помощью спектрально-разностного метода построены ортоморфизмы обобщённой группы кватернионов порядка 16, 32, 64, 128, 256 с близкими к оптимальным значениями разностных характеристик $p_g^{(1)}$ и $p_g^{(2)}$.

Автор выражает благодарность научному руководителю А. В. Менячихину за постановку задачи и Д. А. Бурову за ценные замечания.

ЛИТЕРАТУРА

1. *Johnson D. M., Dulmage A. L., and Mendelsohn N. S.* Orthomorphisms of groups and orthogonal Latin squares. I // Canad. J. Math. 1961. V. 13. P. 356–372.
2. *Mann H. B.* On orthogonal Latin squares // Bull. Amer. Math. Soc. 1944. V. 50. P. 249–257.
3. *Niederreiter H. and Robinson K.* Bol loops of order pq // Math. Proc. Cambr. Phil. Soc. 1981. V. 89. P. 241–256.
4. *Niederreiter H. and Robinson K.* Complete mappings of finite fields // J. Austral. Math. Soc. Ser. A. 1982. V. 33. No. 2. P. 197–212.

5. *Менячихин А. В.* Метод ограниченного дефицита и задача построения ортоморфизмов и почти ортоморфизмов абелевых групп // Дискретная математика. 2019. Т. 31. № 3. С. 58–77.
6. *Менячихин А. В.* Ортоморфизмы абелевых групп с минимально возможными попарными расстояниями // Дискретная математика. 2018. Т. 30. № 4. С. 55–65.
7. *Сачков В. Н.* Цепи Маркова итерационных систем преобразований // Тр. по дискр. матем. 2002. Т. 6. С. 165–183.
8. *Evans A. B.* Applications of complete mappings and orthomorphisms of finite groups // Quasigroups Relat. Syst. 2015. V. 23. P. 5–30.
9. *Evans A. B.* Orthomorphism Graphs of Groups. Lecture Notes in Math. Berlin: Springer, 1992. V. 1535.
10. *Зубов А. Ю.* Математика кодов аутентификации. М.: Гелиос АРВ, 2007. 480 с.
11. *Черемушкин А. В.* Криптографические протоколы. Основные свойства и уязвимости. М.: Изд. центр «Академия», 2009. 272 с.
12. *Тришин А. Е.* Способ построения ортогональных латинских квадратов на основе подстановочных двучленов конечных полей // Обозр. прикл. и промышл. матем. 2008. Т. 15. № 4. С. 764–765.
13. *Тужилин М. Э.* Латинские квадраты и их применение в криптографии // Прикладная дискретная математика. 2012. № 3(17). С. 47–52.
14. *Denes J. and Keedwell A. D.* Latin Squares and their Applications. Budapest: Academiai Kiado, 2015. 545 p.
15. *Глухов М. М.* О методах построения систем ортогональных квазигрупп с использованием групп // Математические вопросы криптографии. 2011. Т. 2. № 4. С. 5–24.
16. *Глухов М. М.* О применениях квазигрупп в криптографии // Прикладная дискретная математика. 2008. № 2(2). С. 28–32.
17. *Погорелов Б. А., Пудовкина М. А.* Вариации ортоморфизмов и псевдоадамаровых преобразований на неабелевой группе // Прикладная дискретная математика. Приложение. 2019. № 12. С. 24–27.
18. *Погорелов Б. А., Пудовкина М. А.* Классы кусочно-квазиаффинных преобразований на обобщенной 2-группе кватернионов // Дискретная математика. 2022. Т. 34. № 1. С. 103–125.
19. *Погорелов Б. А., Пудовкина М. А.* Классы кусочно-квазиаффинных подстановок на диэдральной, полудиэдральной и модулярной максимально-циклической 2-группах // Дискретная математика. 2022. Т. 34. № 2. С. 50–66.
20. *Menyachikhin A. V.* Spectral-linear and sectral-differntial methods for generating S-boxes having almost optimal cryptographic parameters // Матем. вопр. криптогр. 2017. Т. 8. № 2. С. 97–116.
21. *Menyachikhin A. V.* The change in linear and differential characteristics of substitution after the multiplication by transposition // Матем. вопр. криптогр. 2020. Т. 11. № 2. С. 111–123.

REFERENCES

1. *Johnson D. M., Dulmage A. L., and Mendelsohn N. S.* Orthomorphisms of groups and orthogonal Latin squares. I. Canad. J. Math., 1961, vol. 13, pp. 356–372.
2. *Mann H. B.* On orthogonal Latin squares. Bull. Amer. Math. Soc., 1944, vol. 50, pp. 249–257.
3. *Niederreiter H. and Robinson K.* Bol loops of order *pq*. Math. Proc. Cambr. Phil. Soc., 1981, vol. 89, pp. 241–256.
4. *Niederreiter H. and Robinson K.* Complete mappings of finite fields // J. Austral. Math. Soc., Ser. A, 1982, vol. 33, no. 2, pp. 197–212.

5. *Menyachikhin A. V.* The limited deficit method and the problem of constructing orthomorphisms and almost orthomorphisms of Abelian groups. Discrete Math. Appl., 2021, vol. 31, no. 5, pp. 327–343.
6. *Menyachikhin A. V.* Orthomorphisms of Abelian groups with minimum possible pairwise distances. Discrete Math. Appl., 2020, vol. 30, no. 3, pp. 177–186.
7. *Sachkov V. N.* Tsepi Markova iteratsionnykh sistem preobrazovaniy [Markov chains of iterative transformation systems]. Tr. po Diskr. Matem., 2002, vol. 6, pp. 165–183. (in Russian)
8. *Evans A. B.* Applications of complete mappings and orthomorphisms of finite groups. Quasigroups and Relat. Syst., 2015, vol. 23, pp. 5–30.
9. *Evans A. B.* Orthomorphism Graphs of Groups. Lecture Notes in Math., Berlin, Springer, 1992, vol. 1535.
10. *Zubov A. Yu.* Matematika kodov autentifikatsii [Mathematics of Authentication Codes]. Moscow, Gelios ARV Publ., 2007. 480 p.
11. *Cheremushkin A. V.* Kriptograficheskiye protokoly. Osnovnyye svoystva i uyazvimosti [Cryptographic Protocols. Basic Properties and Vulnerabilities]. Moscow, Akademiya Publ., 2009. 272 p.
12. *Trishin A. E.* Sposob postroyeniya ortogonal'nykh latinskikh kvadratov na osnove podstanovochnykh dvuchlenov konechnykh poley [A method for constructing orthogonal Latin squares based on wildcard binomials of finite fields]. Obozr. Prikl. i Promyshl. Matem., 2008, vol. 15, no. 4, pp. 764–765. (in Russian)
13. *Tuzhilin M. E.* Latinskiye kvadraty i ikh primeneniye v kriptografii [Latin squares and their applications in cryptography]. Prikladnaya Diskretnaya Matematika, 2012, no. 3(17), pp. 47–52. (in Russian)
14. *Denes J. and Keedwell A. D.* Latin Squares and their Applications. Budapest, Academiai Kiado, 2015. 545 p.
15. *Glukhov M. M.* O metodakh postroyeniya sistem ortogonal'nykh kvazigrupp s ispol'zovaniyem grupp [On a method of construction of orthogonal quasigroup systems by means of groups.] Mat. Vopr. Kriptogr., 2011, vol. 2, no. 4, pp. 5–24. (in Russian)
16. *Glukhov M. M.* O primeneniyakh kvazigrupp v kriptografii [Some applications of quasigroups in cryptography]. Prikladnaya Diskretnaya Matematika, 2008, no. 2(2), pp. 28–32. (in Russian)
17. *Pogorelov B. A. and Pudovkina M. A.* Variatsii ortomorfizmov i psevdoadamarovykh preobrazovaniy na neabelevoy gruppe [Variations of orthomorphisms and pseudo-Hadamard transformations on nonabelian groups]. Prikladnaya Diskretnaya Matematika. Prilozhenie, 2019, no. 12, pp. 24–27. (in Russian)
18. *Pogorelov B. A. and Pudovkina M. A.* Classes of piecewise-quasiaffine transformations on the generalized 2-group of quaternions. Discrete Math. Appl., 2023, vol. 33, no. 5, pp. 299–316.
19. *Pogorelov B. A. and Pudovkina M. A.* Classes of piecewise quasiaffine transformations on dihedral, quasidihedral and modular maximal-cyclic 2-groups. Discrete Math. Appl., 2024, vol. 34, no. 1, pp. 15–27.
20. *Menyachikhin A. V.* Spectral-linear and spectral-differntial methods for generating S-boxes having almost optimal cryptographic parameters. Mat. Vopr. Kriptogr., 2017, vol. 8, no. 2, pp. 97–116.
21. *Menyachikhin A. V.* The change in linear and differential characteristics of substitution after the multiplication by transposition. Mat. Vopr. Kriptogr., 2020, vol. 11, no. 2, pp. 111–123.

МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

УДК 004.056 DOI 10.17223/20710410/66/6

АТАКИ НА ПРОТОКОЛЫ АУТЕНТИФИЦИРОВАННОЙ ВЫРАБОТКИ ОБЩЕГО КЛЮЧА ПРИ НАВЯЗЫВАНИИ БУДУЩИХ ОТКРЫТЫХ ЭФЕМЕРНЫХ КЛЮЧЕЙ

Е. К. Алексеев, С. Н. Кяжин, С. В. Смышляев

ООО «КРИПТО-ПРО», г. Москва, Россия

E-mail: alekseev@cryptopro.ru, kyazhin@cryptopro.ru, svs@cryptopro.ru

Исследуется построение атак на протоколы аутентифицированной выработки общего ключа при наличии у нарушителя возможности навязывать участнику использование эфемерных открытых значений. Обосновывается актуальность рассмотрения указанной возможности. Описываются атаки на протоколы SIGMA, SIGMA-R, STS-MAC, «Эхинацея-3» и постквантовый протокол BKM-KK. Приводятся рассуждения о конструктивных особенностях протоколов, позволяющих защититься от атак такого типа.

Ключевые слова: *криптография, криптографический протокол, аутентифицированная выработка общего ключа, атака, навязывание открытых эфемерных ключей.*

FORCING FUTURE PUBLIC EPHEMERAL KEYS TO ATTACK AUTHENTICATED KEY ESTABLISHMENT PROTOCOLS

E. K. Alekseev, S. N. Kyazhin, S. V. Smyshlyaev

CryptoPro LLC, Moscow, Russia

This paper studies the security of the authenticated key establishment protocols against the adversary who has the capability to force the participants to use of ephemeral public values. The paper substantiates the relevance of considering this capability, describes, in particular, attacks on the SIGMA, SIGMA-R, STS-MAC, Echinacea-3 protocols and the post-quantum BKM-KK protocol, and discusses the design features of protocols that allow to protect against attacks of this type.

Keywords: *cryptography, cryptographic protocol, authenticated key establishment, attack, forcing public ephemeral keys.*

Введение

Одним из первых шагов при проведении криптографического анализа или синтеза любой системы является определение потенциальных возможностей нарушителя по взаимодействию с системой на качественном уровне. Важности этого этапа посвящён ряд как иностранных, так и отечественных работ [1–4].

В работе [5] рассматриваются вопросы определения возможностей нарушителя для протоколов аутентифицированной выработки общего ключа (Authenticated Key Establishment, AKE) и впервые, насколько известно авторам, упоминается возможность нарушителя навязывать открытые эфемерные значения, которые стороны будут использовать при будущем взаимодействии. Подчеркнём, что данная возможность подразумевает подмену значения открытого эфемерного ключа не в процессе передачи по каналу связи, а до начала взаимодействия, непосредственно на стороне участника, который использует его при выполнении протокола, считая корректным элементом эфемерной ключевой пары. При этом в указанных работах данная возможность нарушителя не исследуется ни в части её актуальности на практике, ни в части уязвимости к ней известных AKE-протоколов, ни в части подходов к обеспечению защиты от атак, основанных на её применении.

В настоящей работе приведены следующие результаты первых шагов по исследованию такой возможности нарушителя. В п. 1 анализируется актуальность такой возможности на практике. Показывается, что построение протоколов, стойких по отношению к нарушителям с такими возможностями, позволит улучшить эффективность протоколов и снизить требования к защищённому хранению данных сторонами взаимодействия. Указанные улучшения могут стать особенно актуальными для случая применения постквантовых криптографических алгоритмов, открытые ключи которых зачастую в десятки раз больше ключей классических механизмов. В п. 2 приводятся атаки на такие известные протоколы, как SIGMA, SIGMA-R, STS-MAC, «Эхинацея-3» и постквантовый протокол BKM-KK. В п. 3 обсуждаются вопросы построения AKE-протоколов, защищённых от атак со стороны нарушителей, которые обладают возможностью навязывать сторонам открытые эфемерные значения.

1. Актуальность рассмотрения возможности навязывания открытых эфемерных значений

Источником возможностей нарушителя, в первую очередь, является порядок применения криптографического механизма в более высокоуровневой системе или условия его эксплуатации на практике. Если не учесть какую-либо существенную возможность, можно получить расхождение прогноза о стойкости, сформированного в результате проведения криптографического анализа, с действительностью, как, например, получилось когда-то с подпротоколом Record протокола TLS 1.0 [6, 7]. Процесс выявления потенциальных возможностей нарушителя, по всей видимости, невозможно формализовать, и его успешность зависит лишь от опыта работающих аналитиков.

При этом в процессе синтеза криптосистемы целесообразно наделять потенциального нарушителя максимально широким спектром возможностей. Действительно, чем больше возможностей у нарушителя, относительно которого удалось обосновать стойкость новой криптосистемы, тем меньше требований будет предъявлено к порядку её эксплуатации. Например, если криптосистема стойка относительно нарушителя, который может узнавать какие-то промежуточные значения, возникающие при реализации протокола, то участники могут не беспокоиться о безопасном удалении этих значений. Таким образом, некоторые возможности, которые в настоящее время кажутся нереальными на практике, при их рассмотрении могут открывать существенные перспективы в части создания эффективных криптосистем. Возможность нарушителя, которой посвящена настоящая работа, относится именно к таким.

Любой современный AKE-протокол предполагает использование участниками эфемерных, то есть одноразовых, значений. Это могут быть и случайно сгенерированные

битовые строки (например, значения client_random и server_random в протоколах TLS Handshake разных версий), и более сложные объекты, как, например, эфемерные ключевые пары, используемые для выработки общего секрета по схеме Диффи — Хеллмана. В большинстве реализаций такие параметры порождаются с помощью различного типа датчиков случайных чисел непосредственно в процессе взаимодействия по протоколу. Однако стойкие современные AKE-протоколы зачастую требуют выполнения большого объёма вычислений, что при их реализации на низкоресурсных устройствах приводит к поиску путей для оптимизации. Одним из таких путей, который явно нарушает изначальную конструкцию протокола, является использование эфемерных ключевых пар в нескольких сеансах. Про такой сценарий написан отдельный раздел 2.12 документа [8], определяющего протокол IKEv2, он обсуждается также в [9]. Такая оптимизация предполагает, что эфемерный ключ должен где-то храниться длительное (по сравнению с выполнением одного сеанса протокола) время. При этом храниться ключевая пара должна в защищённой памяти, размер которой также может быть ограничен на низкоресурсных устройствах. Таким образом, следующий шаг оптимизации состоит в том, чтобы хранить защищённо лишь закрытую часть эфемерного ключа. При этом у нарушителя может быть возможность не только узнать соответствующую открытую его часть (такая возможность уже рассматривалась в работе [10]), но и подменить её. Отметим, что работа носит теоретический характер, в ней не рассматриваются практические сценарии навязывания (подмены) сеансовой ключевой информации. Другим подходом к оптимизации вычислений, который нарушает спецификацию протокола в меньшей степени, является предварительное вычисление какого-то количества эфемерных ключей для будущих сеансов. В таком сценарии у низкоресурсных устройств тем более может возникнуть соблазн хранить хотя бы открытые части в памяти, доступной для нарушителя.

С другой стороны, если какой-нибудь протокол является стойким (в требуемом от него смысле) относительно нарушителя, который может навязывать сторонам открытые эфемерные значения, то при его реализации использование предвычисленных значений, хранящихся в памяти, доступной для нарушителя, является вполне безопасным путём оптимизации. С переходом на постквантовые криптографические алгоритмы такая возможность может стать востребованной не только низкоресурсными устройствами, но и более привычными платформами. Это объясняется тем, что размеры открытых ключей некоторых алгоритмов достигают сотен тысяч байтов при том, что открытые ключи классических алгоритмов, основанных на эллиптических кривых, занимают несколько десятков байтов. Ниже мы отдельно рассуждаем о стойкости протоколов, основанных на постквантовых алгоритмах, и приводим пример атаки на один из таких протоколов.

2. Атаки на известные AKE-протоколы

Опишем атаки на известные AKE-протоколы, в ходе которых нарушитель навязывает участникам взаимодействия открытые эфемерные значения. Рассматриваются протоколы из трёх классов, которые определяются типом ключа, используемым для аутентификации сторон:

1) ключ подписи;
2) скаляр (используемый при вычислениях, задаваемых непосредственно протоколом);
3) ключ механизма инкапсуляции ключа (Key Encapsulation Mechanism, KEM).

В табл. 1 перечислены описанные атаки: в первом столбце указан класс атакуемого протокола, во втором — протокол, в третьем (объединённом) — используемые возможности нарушителя («*» означает навязывание участникам взаимодействия открытых эфемерных значений без знания соответствующих им закрытых значений), в четвёртом — реализуемая угроза:

— AUTH — ложная аутентификация от имени одного участника;
— MITM — ложная аутентификация от имени двух участников;
— KCI — ложная аутентификация после вскрытия долговременного ключа проверяющего;
— SEC — нарушение секретности ключей;
— PFS — нарушение «свойства PFS», т. е. секретности ключей, выработанных до вскрытия долговременного ключа.

В последних двух столбцах приведены ссылки на описание атак.

Таблица 1

Атаки на AKE-протоколы, описанные в настоящей работе

Тип долговр. ключей	**Протокол**	**Возможности нарушителя**					**Угроза**	**Опис. атаки, разд.**	**Схема атаки, рис.**
		Навяз. откр. эфемерного ключа *инициатору*	Навяз. откр. эфемерного ключа *ответчику*	Изменение сообщений в канале	Компрометация долговременного ключа	Компрометация сеансового ключа			
Подпись	SIG-DH+	✓					AUTH	2.1	2
			✓						
	SIGMA		✓	✓					4
		✓	✓	✓			MITM		5
	STS-MAC		✓	✓			AUTH		
		✓	✓	✓			MITM		
	«Эхинацея-3»		✓	✓			AUTH		
		✓	✓	✓			MITM		
	SIGMA-R	✓		✓			AUTH		7
Скаляр	TS3	✓					AUTH	2.2	8
			✓						
	CF	✓*							10
			✓*						
	SK6	✓*			✓		KCI		12
		✓*	✓*		✓		PFS		
		✓*	✓*			✓	SEC		
KEM	BKM-KK		✓				AUTH	2.3	14

Порядок наименования и описания (в том числе обозначения, использование классов базовых криптографических механизмов без уточнения конкретных их представителей и т. д.) протоколов в целом соответствует работе [11], но для краткости протоколы описаны сразу в варианте с предвычислением эфемерных значений. Операции чтения из защищённой памяти и памяти, доступной для нарушителя, обозначены как $e \leftarrow sMEM$ и $E \leftarrow MEM$ соответственно. Важно отметить: мы считаем, что сторона во время работы по протоколу не проверяет соответствие закрытого и открытого

эфемерных ключей, так как это свело бы на нет почти всё преимущество от предварительных вычислений. Навязывание нарушителем открытых эфемерных значений указано в рамках, как действие, выполняемое соответствующей стороной. В рамках также указаны значения, изменяемые нарушителем в пересылках.

2.1. Протоколы, использующие схему подписи

Опишем атаки с навязыванием открытых эфемерных значений на протоколы SIG-DH+ [12], SIGMA [13] и SIGMA-R [13]. В результате всех указанных атак реализуется угроза ложной аутентификации.

Протокол SIG-DH+ (рис. 1).

$A : sk_A^s, pk_A^s$		$B : sk_B^s, pk_B^s$
$w_A \leftarrow sMEM$		
$W_A \leftarrow MEM$	$\xrightarrow{A, W_A}$	$w_B \leftarrow sMEM$
		$W_B \leftarrow MEM$
Verify σ_B	$\xleftarrow{B, W_B, \sigma_B}$	$\sigma_B \leftarrow \mathsf{Sig}_{sk_B^s}(B, W_B, W_A, A)$
$\sigma_A \leftarrow \mathsf{Sig}_{sk_A^s}(A, W_A, W_B, B)$	$\xrightarrow{\sigma_A}$	Verify σ_A
$Q \leftarrow w_A \cdot W_B$		$Q \leftarrow w_B \cdot W_A$
$K \leftarrow \mathsf{KDF}(Q, W_A, W_B, A, B)$		$K \leftarrow \mathsf{KDF}(Q, W_A, W_B, A, B)$
return K		**return** K

Рис. 1. Схема протокола SIG-DH+ с предвычислениями эфемерных ключей

Атака на протокол SIG-DH+ требует от нарушителя лишь навязать стороне A значение W'_A с известным нарушителю дискретным логарифмом w'_A ($W'_A = w'_A \cdot P$) вместо эфемерного ключа W_A. Передаваемые по каналу связи значения менять не требуется. При этом нарушитель может вычислить ключ K, равный тому, который вычислит сторона B, так как $Q = w_B \cdot W'_A = w'_A \cdot W_B$, а w'_A нарушителю известно. Таким образом, нарушитель вырабатывает общий ключ со стороной B, но B думает, что выработал его со стороной A. Схема атаки представлена на рис. 2.

Существенным отличием в описанном сценарии атаки от обычной замены W_A на W'_A при передаче по каналу связи является то, что при подмене в канале связи сторона A будет вычислять значение подписи σ_A от истинного значения W_A, поэтому на стороне B проверка этой подписи завершится ошибкой.

Легко видеть, что аналогичная атака работает при навязывании открытых эфемерных значений стороне B.

Уязвимыми к аналогичному сценарию атаки оказываются и другие протоколы, в которых по каналу не пересылаются значения, полученные с помощью выработанного по схеме Диффи — Хеллмана секрета (для протокола SIG-DH+ это $Q = w_A \cdot W_b = = w_b \cdot W_A$). В качестве примера можно привести протокол TS3-1 [14, 15] (упрощённое описание можно найти в [11]).

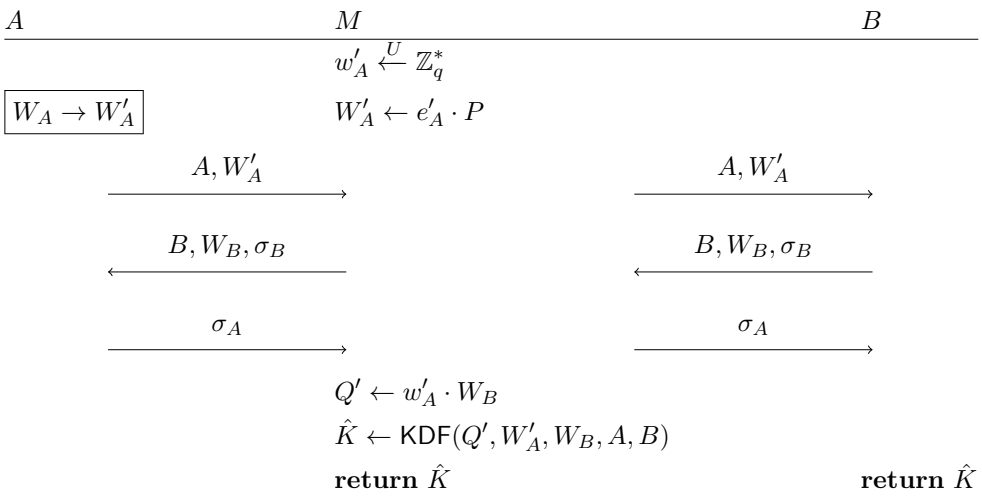

Рис. 2. Схема атаки на протокол SIG-DH+ с навязыванием эфемерных значений стороне A

Протоколы SIGMA. Перейдём к описанию атак, требующих от нарушителя не только навязывания открытых эфемерных значений, но и изменения сообщений, передаваемых по каналу связи. Напомним, что протокол SIGMA лег в основу протоколов TLS 1.3 [16], IKEv2 [8] и стандартизированного в России протокола «Эхинацея-3». Описание протокола SIGMA представлено на рис. 3.

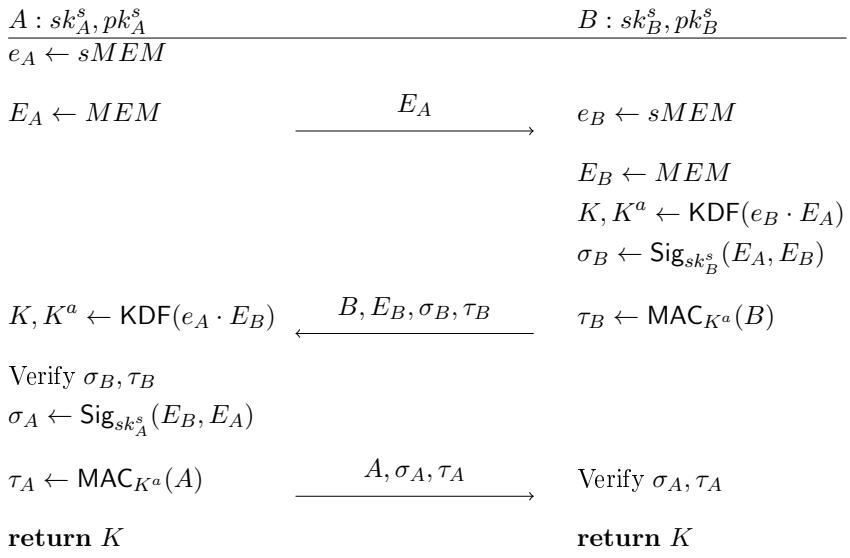

Рис. 3. Схема протокола SIGMA с предвычислениями

Пересылки, осуществляемые при реализации атаки на протокол SIGMA в варианте с навязыванием эфемерных открытых ключей стороне B, представлены на рис. 4. В результате нарушитель вырабатывает общий ключ со стороной A, а сторона A думает, что выработала его со стороной B.

Аналогичная атака при навязывании открытых эфемерных ключей стороне A не сработает, так как сторона B сформирует значение τ_B на ключе $\mathsf{KDF}(e_B \cdot E'_A)$, а сторона A будет проверять его на ключе $\mathsf{KDF}(e_A \cdot E_B)$. При этом ни значение e_A, ни

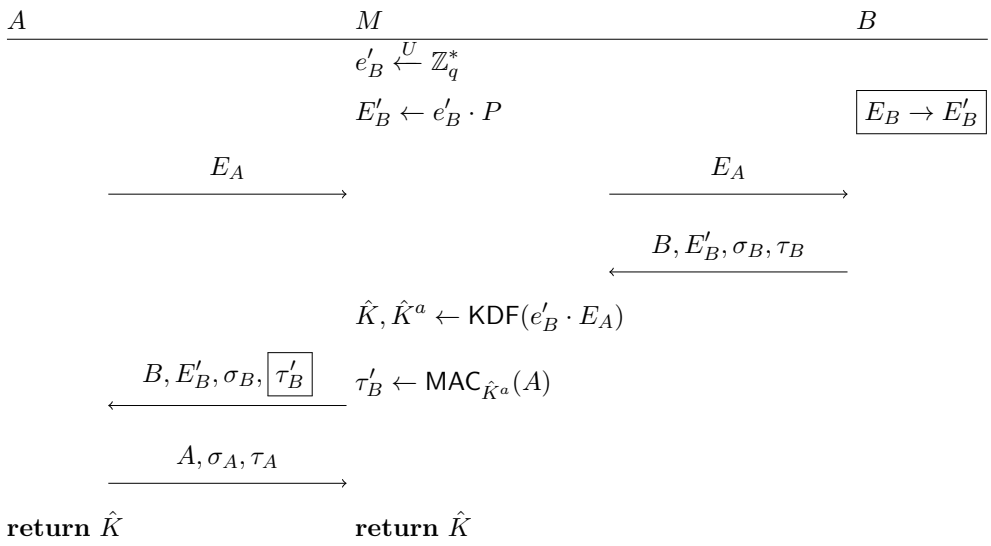

Рис. 4. Схема атаки на протокол SIGMA с навязыванием эфемерных значений стороне B

значение e_B нарушителю не известно, поэтому подменить в канале значение τ_B на то, которое успешно проверится на стороне A, он не может.

Рассмотрим атаку на протокол SIGMA при наличии у нарушителя возможности навязывать открытые эфемерные значения обеим сторонам. В этом случае получается реализовать угрозу ложной аутентификации ещё и от лица A. Схематичное описание атаки представлено на рис. 5.

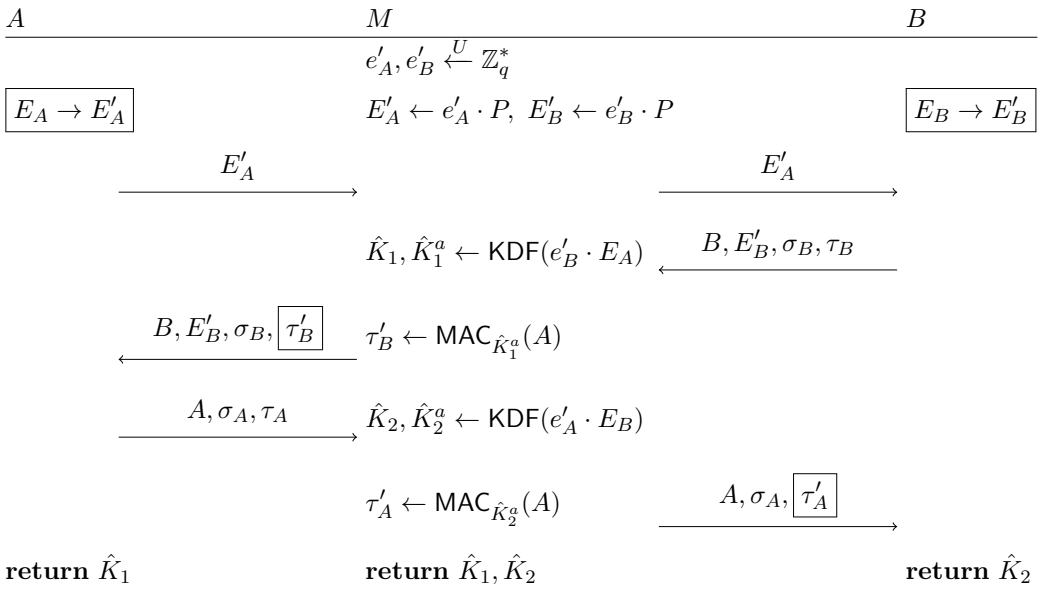

Рис. 5. Схема атаки на протокол SIGMA с навязыванием эфемерных значений двум сторонам

Протоколы STS-MAC [17] и «Эхинацея-3» [18] незначительно отличаются от протокола SIGMA (упрощённое описание можно найти в [11]), вследствие чего обе атаки на протокол SIGMA работают и для них.

Протокол SIGMA-R. Наиболее сложной с точки зрения требуемых от нарушителя действий является атака на протокол SIGMA-R, описание которого представлено на рис. 6. Сценарий атаки на протокол SIGMA-R приведён на рис. 7.

$A: sk_A^s, pk_A^s$		$B: sk_B^s, pk_B^s$
$e_A \leftarrow sMEM$		
$n_A, E_A \leftarrow MEM$	$\xrightarrow{E_A, n_A}$	$e_B \leftarrow sMEM$
$K, K_A^a, K_B^a, K_A^{ae}, K_B^{ae} \leftarrow \mathsf{KDF}(e_A \cdot E_B)$	$\xleftarrow{E_B, n_B}$	$n_B, E_B \leftarrow MEM$
$\tau_A \leftarrow \mathsf{MAC}_{K_A^a}(A)$ $\sigma_A \leftarrow \mathsf{Sig}_{sk_A^s}(n_B, E_A)$		
$C_A \leftarrow \mathcal{AE}.\mathsf{Enc}_{K_A^{ae}}(\varepsilon, (A, \sigma_A, \tau_A))$	$\xrightarrow{C_A}$	$K, K_A^a, K_B^a, K_A^{ae}, K_B^{ae} \leftarrow \mathsf{KDF}(e_B \cdot E_A)$
		$A, \sigma_A, \tau_A \leftarrow \mathcal{AE}.\mathsf{Dec}_{K_A^{ae}}(\varepsilon, C_A)$ Verify σ_A, τ_A $\tau_B \leftarrow \mathsf{MAC}_{K_B^a}(B)$ $\sigma_B \leftarrow \mathsf{Sig}_{sk_B^s}(n_A, E_B)$
$B, \sigma_B, \tau_B \leftarrow \mathcal{AE}.\mathsf{Dec}_{K_B^{ae}}(\varepsilon, C_B)$	$\xleftarrow{C_B}$	$C_B \leftarrow \mathcal{AE}.\mathsf{Enc}_{K_B^{ae}}(\varepsilon, (B, \sigma_B, \tau_B))$
Verify σ_B, τ_B **return** K		**return** K

Рис. 6. Схема протокола SIGMA-R с предвычислениями

Сценарий атаки на протокол SIGMA-R усложняется по сравнению с атакой на протокол SIGMA из-за того, что в протоколе SIGMA-R третья и четвёртая пересылки зашифровываются на ключах, выработанных из ключа Диффи — Хеллмана. При навязывании эфемерных ключей стороне A нарушителю, чтобы выдать себя за A перед B, нужно корректно сформировать вторую пересылку C_A, которую направляет сторона A и в которой содержатся значения подписи и имитовставки. Имитовставку нарушитель может посчитать сам, но подпись стороны A он должен получить из оригинального сообщения C_A, сформированного стороной A. Чтобы знать ключ, на котором будет формироваться C_A, нарушитель меняет в канале значение E_B. За счёт этого он может получить «честное» значение σ_A, а дальше уже сформировать C'_A на том ключе, на котором его будет расшифровывать сторона B.

В схематичном описании атаки при использовании функции KDF указаны только те из вырабатываемых ключей, которые нарушитель использует в дальнейшем для осуществления атаки.

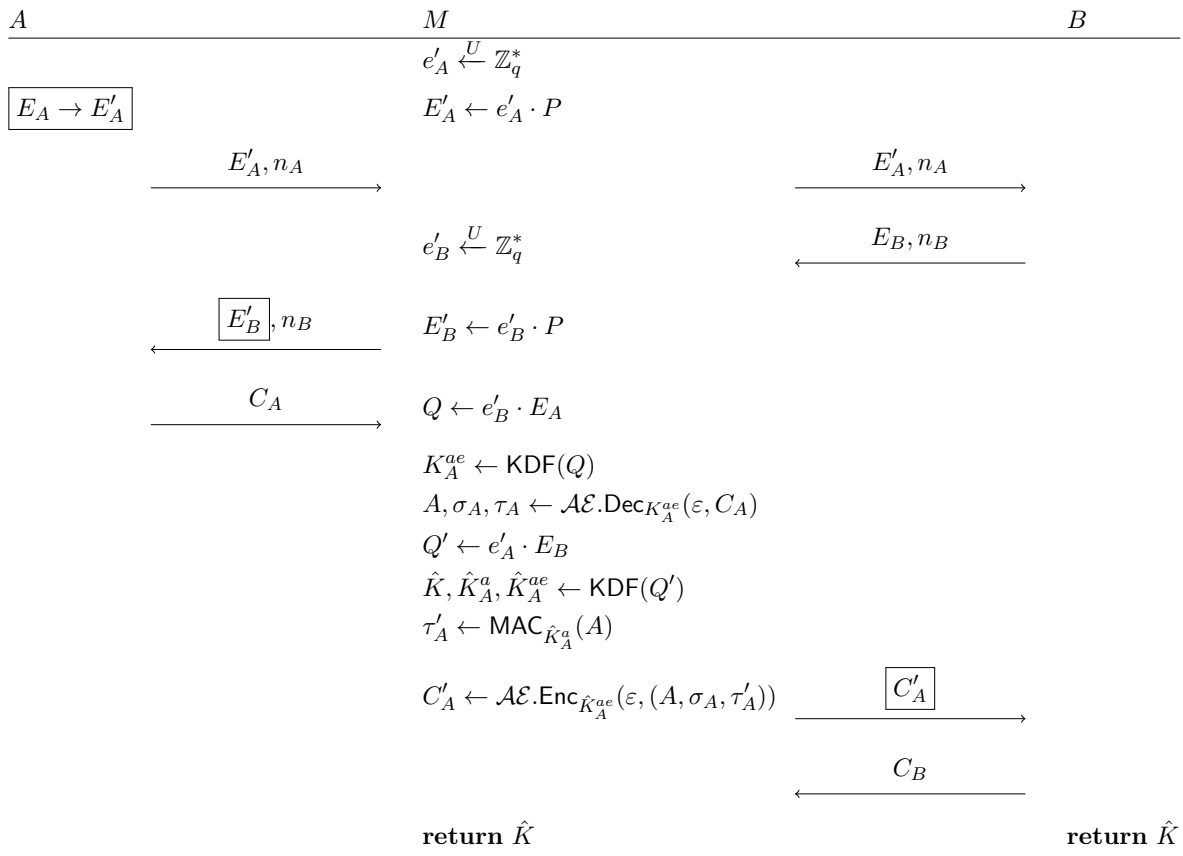

Рис. 7. Схема атаки на протокол SIGMA-R с навязыванием эфемерных значений стороне B

2.2. Протоколы, использующие долговременные скалярные величины

Опишем атаки с навязыванием открытых эфемерных значений на протоколы TS3 [14, 15], CF [19] и SK6 [20]. Для протоколов TS3 и CF атаки приводят к реализации угрозы ложной аутентификации, а для протокола SK6 — к реализации угрозы KCI и нарушению свойства PFS.

Протокол TS3 (рис. 8). Сценарий атаки такой же, как для протокола SIG-DH+, его применение становится возможным из-за того, что по каналу не пересылаются сообщения, зависящие от выработанного ключа, а аутентифицирующее сторону значение зависит лишь от идентификаторов сторон и их эфемерных ключей.

Протокол CF. Схожий сценарий работает для протокола CF, его описание представлено на рис. 9, схема атаки — на рис. 10.

Особенность сценария атаки против этого протокола в том, что нарушитель навязывает стороне A эфемерный ключ $E'_A = -X_A + P$, дискретный логарифм которого ему неизвестен. Однако за счёт того, что единственным не передаваемым по каналу связи аргументом функции выработки итогового сеансового ключа является точка $W = (e_B + x_B)(E_A + X_A)$, нарушитель может её вычислить при таком значении E'_A.

$A : x_A, X_A$		$B : x_B, X_B$
$e_A \leftarrow sMEM$		
$E_A \leftarrow MEM$		
$K^a \leftarrow \mathsf{KDF}(x_A \cdot X_B)$		
$\tau_A \leftarrow \mathsf{MAC}_{K^a}(A, B, E_A)$	$\xrightarrow{A, E_A, \tau_A}$	$K^a \leftarrow \mathsf{KDF}(x_B \cdot X_A)$
		Verify τ_A
		$e_B \leftarrow sMEM$
		$E_B \leftarrow MEM$
Verify τ_B	$\xleftarrow{B, E_B, \tau_B}$	$\tau_B \leftarrow \mathsf{MAC}_{K^a}(B, A, E_B)$
$K \leftarrow e_A \cdot E_B$		$K \leftarrow e_B \cdot E_A$
return K		**return** K

A	M	B
	$e_A \xleftarrow{U} \mathbb{Z}_q^*$	
$\boxed{E_A \to E'_A}$	$E'_A = e'_A \cdot P$	
$\xrightarrow{A, E'_A, \tau_A}$	$\xrightarrow{A, E'_A, \tau_A}$	
	$\hat{K} \leftarrow e'_A \cdot E_B$ $\xleftarrow{B, E_B, \tau_B}$	
	return $\hat{K}$	**return** $\hat{K}$

Рис. 8. Схема протокола TS3 с предвычислениями и схема атаки на него

$A : (x_A, X_A), (sk_A^s, pk_A^s)$		$B : (x_B, X_B), (sk_B^s, pk_B^s)$
$e_A \leftarrow sMEM$		
$E_A \leftarrow MEM$		
$\sigma_A \leftarrow \mathsf{Sig}_{sk_A^s}(E_A)$	$\xrightarrow{A, E_A, \sigma_A}$	Verify σ_A
		$e_B \leftarrow sMEM$
		$E_B \leftarrow MEM$
Verify σ_B	$\xleftarrow{B, E_B, \sigma_B}$	$\sigma_B \leftarrow \mathsf{Sig}_{sk_B^s}(E_B)$
$W \leftarrow (e_A + x_A) \cdot (E_B + X_B)$		$W \leftarrow (e_B + x_B) \cdot (E_A + X_A)$
$K \leftarrow \mathsf{KDF}(A, B, W, E_A)$		$K \leftarrow \mathsf{KDF}(A, B, W, E_A)$
return K		**return** K

Рис. 9. Схема протокола CF с предвычислениями

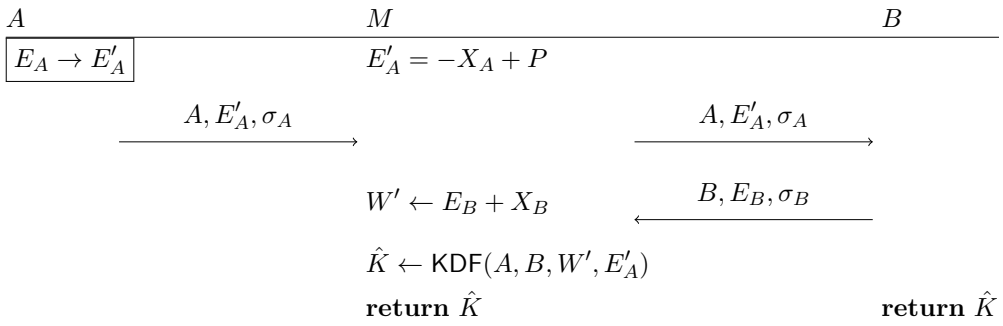

Рис. 10. Схема атаки на протокол CF с навязыванием эфемерных значений стороне A

Протокол SK6. В атаке на протокол SK6, представленный на рис. 11, нарушитель реализует угрозу KCI.

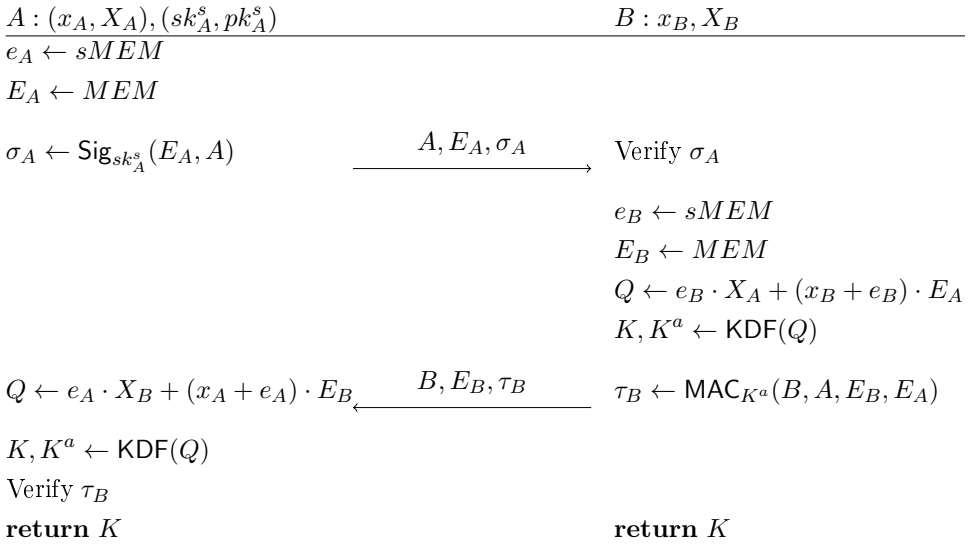

Рис. 11. Схема протокола SK6 с предвычислениями

Заметим, что протокол SK6 уязвим по отношению к угрозе KCI, если вскрыть скалярную часть долговременного ключа инициатора, то есть стороны A (без использования возможности навязывания эфемерных ключей). Поэтому рассмотрение аналогичных атак относительно стороны A с использованием соответствующей возможности нарушителя не представляет интереса. Атака с навязыванием открытых эфемерных ключей стороне A по реализации угрозы KCI относительно стороны B представлена на рис. 12. Такая атака становится возможной за счёт того, что общий ключ вычисляется по формуле $K = \mathsf{KDF}(e_B \cdot X_A + (x_B + e_B)E_A)$. Заметим, что при $E_A = -X_A$ аргумент функции KDF равен $-x_B \cdot X_A$ и не зависит от эфемерных ключей сторон.

Упомянутое свойство того, как вычисляется аргумент функции KDF, позволяет также построить атаку на протокол SK6, в результате которой нарушается свойство PFS. Для этого нарушителю понадобится навязывать открытые эфемерные значения обеим сторонам: стороне A навязывается ключ $E'_A = -X_A$, а стороне B — ключ $E'_B = -X_B$. В сеансе, где A и B используют оба этих навязанных ключа, будет выработан ключ $K = \mathsf{KDF}(-x_A \cdot X_B) = \mathsf{KDF}(-x_B \cdot X_A)$. Если после такого сеанса нарушитель

вскроет долговременный скалярный ключ какой-либо из сторон A или B, то он сможет вычислить выработанный ключ K.

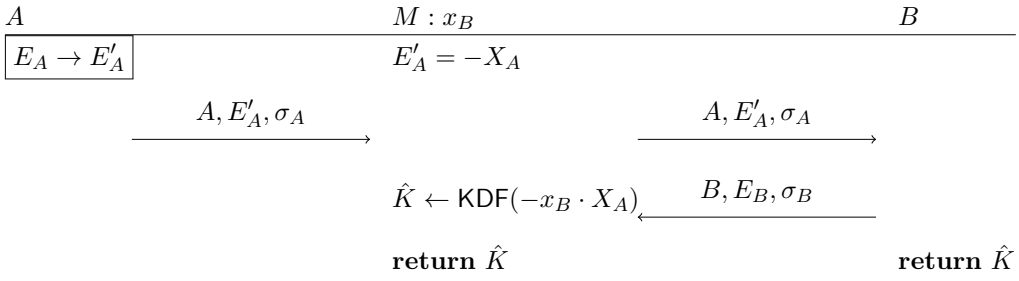

Рис. 12. Схема атаки на протокол SK6 с навязыванием эфемерных значений стороне A

Отметим, что указанное свойство позволяет нарушителю, навязывая эфемерные ключи обеим сторонам, заставить стороны A и B в любом количестве сеансов выработать одинаковый ключ, что также является уязвимостью протокола (приводит к нарушению секретности ключа при наличии у нарушителя возможности вскрывать ключи некоторых сеансов).

2.3. Протоколы, использующие KEM

Опишем атаку с навязыванием открытых эфемерных значений на протокол BKM-KK [21]. Под этим сокращённым названием мы имеем в виду протокол, который в оригинальной работе называется «Optimised KEM-based UM protocol». Схема протокола представлена на рис. 13. В данном протоколе и долговременные, и эфемерные ключи — это ключи механизма инкапсуляции ключа.

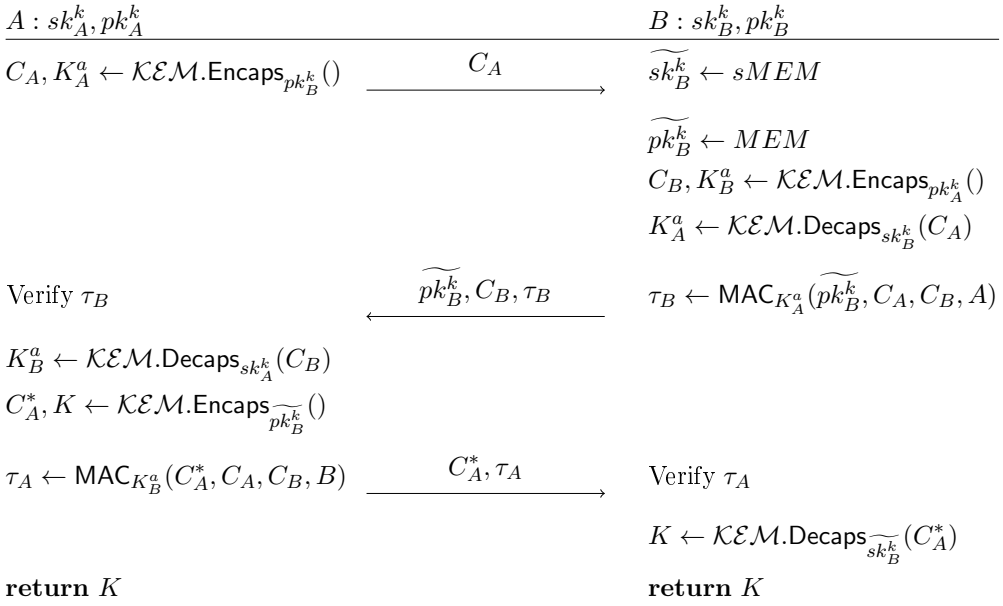

Рис. 13. Схема протокола BKM-KK с предвычислениями

Действия нарушителя при реализации атаки на протокол BKM-KK схематично приведены на рис. 14. В результате навязывания эфемерных ключей схемы KEM стороне B нарушителю удаётся выработать общий ключ со стороной A, причём A думает,

что выработала ключ со стороной B. В работе [21] приведено ещё пять протоколов, идейно схожих с протоколом BKM-KK, но отличающихся тем, какие механизмы используются для аутентификации сторон. Описанная атака применима ко всем этим протоколам.

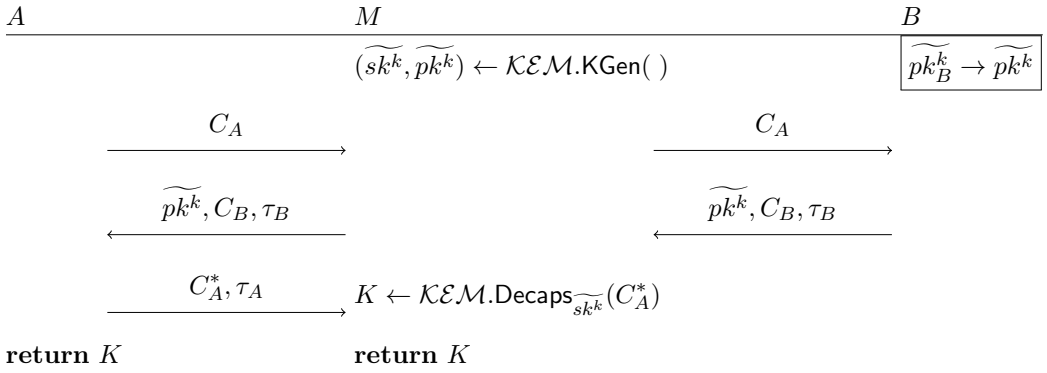

Рис. 14. Схема атаки на протокол BKM-KK с навязыванием эфемерных значений стороне A

3. Подходы к обеспечению стойкости

Приведём примеры подходов, с помощью которых можно защититься от атак, описанных в предыдущем пункте. В качестве примеров протоколов, строение которых не позволяет реализовать ни один из описанных сценариев атак, используются протоколы SIGMA-opt1 [13], «Лимонник-3» [18] и KEMTLS [22].

При построении атак на протоколы, использующие схему подписи, одной из особенностей уязвимых протоколов явилось то, что стороны вычисляли подпись от значений, не зависящих от закрытого эфемерного ключа:

— в протоколе SIG-DH+ подпись вычисляется от идентификаторов сторон и открытых эфемерных ключей;
— в протоколе SIGMA подпись вычисляется от открытых эфемерных ключей сторон;
— в протоколе SIGMA-R подпись вычисляется от своего открытого эфемерного ключа и случайности, присланной другой стороной.

При этом рассматриваемая возможность нарушителя не позволяет считать, что сторона использует соответствующий закрытый эфемерный ключ, если она прислала корректную подпись под соответствующим открытым ключом. Один из подходов к проверке соответствия закрытого и открытого ключей состоит в том, чтобы подписывать значение, которое можно вычислить только с использованием корректного закрытого ключа. Этим значением может быть имитовставка, вычисленная с помощью выработанного общего ключа (или производного от него) от данных, переданных в канале связи. Примером протокола, в котором делается именно так, является протокол SIGMA-opt1 [13] (его описание приведено на рис. 15). Причиной модификации протокола SIGMA в оригинальной работе называется экономия размера передаваемых по каналу связи данных. Однако, как показано выше, такой подход имеет преимущества и с криптографической точки зрения.

$A: sk_A^s, pk_A^s$		$B: sk_B^s, pk_B^s$
$e_A \leftarrow sMEM$		
$E_A \leftarrow MEM$	$\xrightarrow{E_A}$	$e_B \leftarrow sMEM$
		$E_B \leftarrow MEM$
		$K, K^a \leftarrow \mathsf{KDF}(e_B \cdot E_A)$
$K, K^a \leftarrow \mathsf{KDF}(e_A \cdot E_B)$	$\xleftarrow{B, E_B, \sigma_B}$	$\sigma_B \leftarrow \mathsf{Sig}_{sk_B^s}(E_A, E_B, \mathsf{MAC}_{K^a}(B))$
Verify σ_B		
$\sigma_A \leftarrow \mathsf{Sig}_{sk_A^s}(E_B, E_A, \mathsf{MAC}_{K^a}(A))$	$\xrightarrow{A, \sigma_A}$	Verify σ_A
return K		**return** K

Рис. 15. Схема протокола SIGMA-opt1 с предвычислениями

Одной из причин уязвимости протоколов, использующих в качестве долговременных скалярные ключи, является то, что общий ключ вычисляется по формуле, которая позволяет манипулировать значением ключа, изменяя открытые эфемерные ключи. В протоколе CF ключ удалось сделать вычисляемым по открытым параметрам, а в протоколе SK6 — сделать его не зависящим от эфемерных ключей. Одним из подходов к тому, чтобы нарушить это свойство общего ключа, является использование при его вычислении компонент, зависящих от долговременных и эфемерных ключей как отдельных аргументов функции KDF. Недостатком такого метода является увеличение в некоторых случаях количества реализуемых операций вычисления кратной точки. Примером протокола, для которого не удаётся применить описанные сценарии атак, является другой стандартизированный в России AKE-протокол «Лимонник-3», схема которого представлена на рис. 16. Ключ в этом протоколе вычисляется по формуле (для стороны A)

$$K = \mathsf{KDF}(e_A \cdot X_B, x_A \cdot E_B, E_A, E_B).$$

Легко видеть, что навязывание, скажем, E'_B изменит лишь значение второго аргумента, но это не приносит пользы при построении атаки, так как ключи, на которых вычисляется и проверяется имитовставка, будут различны, а нарушитель это значение перевычислить не сможет, так как не знает ни одного из закрытых значений для вычисление первого аргумента функции KDF. Сценарий по реализации угрозы KCI также не реализуем, так как нарушитель, атакующий сторону A, не сможет вычислить компоненту $e_A \cdot X_B$. Что касается трудоёмкости вычислений, заметим: в отличие от протокола CF ($K = \mathsf{KDF}(A, B, (e_A + x_A)(E_B + X_B), E_A)$), в протоколе «Лимонник-3» для получения ключа требуется вычислить две кратные точки, а не одну.

В качестве протокола, использующего схемы KEM, для которого не удаётся построить атаку с навязыванием открытых эфемерных значений, приведём протокол KEMTLS [22], описанный на рис. 17. Уязвимостью протокола BKM-KK, позволившей провести на него атаку, стало то, что ключи, которые получены в результате выполнения алгоритма декапсуляции, используются сторонами исключительно для подтверждения владения закрытым ключом путём вычисления имитовставок от передаваемых в канале данных. При этом эти ключи никак не влияют на результат работы протокола, то есть на итоговый сеансовый ключ K. В протоколе KEMTLS ключи K_A и K_B, полученные в результате декапсуляции на долговременных закрытых ключах

$A : x_A, X_A$		$B : x_B, X_B$
$e_A \leftarrow sMEM$		
$E_A \leftarrow MEM$	$\xrightarrow{A, E_A}$	$e_B \leftarrow sMEM$
		$E_B \leftarrow MEM$
		$Q \leftarrow e_B \cdot X_A$
		$R \leftarrow x_B \cdot E_A$
		$K, K^a \leftarrow \mathsf{KDF}(Q, R, A, B)$
$Q \leftarrow x_A \cdot E_B$	$\xleftarrow{B, E_B, \tau_B}$	$\tau_B \leftarrow \mathsf{MAC}_{K^a}(\mathsf{c}_1, E_B, E_A, B, A)$
$R \leftarrow e_A \cdot X_B$		
$K, K^a \leftarrow \mathsf{KDF}(Q, R, A, B)$		
Verify τ_B		
$\tau_A \leftarrow \mathsf{MAC}_{K^a}(\mathsf{c}_2, E_A, E_B, A, B)$	$\xrightarrow{\tau_A}$	Verify τ_A
return K		**return** K

Рис. 16. Схема протокола «Лимонник-3» с предвычислениями

сторон, используются не только для аутентификации сторон, но и замешиваются в вырабатываемый ключ K: $K = \mathsf{KDF}(\widetilde{K_B}, K_A, K_B)$. Это не позволяет нарушителю узнать сеансовый ключ и подтвердить его значение, используя закрытый ключ схемы KEM, соответствующий навязанному одной из сторон открытому эфемерному ключу.

$A : sk_A^k, pk_A^k$		$B : sk_B^k, pk_B^k$
$\widetilde{sk_A^k} \leftarrow sMEM$		
$\widetilde{pk_A^k} \leftarrow MEM$		
$r_A \xleftarrow{U} \{0,1\}^{256}$	$\xrightarrow{\widetilde{pk_A^k}, r_A}$	$\widetilde{C_B}, \widetilde{K_B} \leftarrow \mathcal{KEM}.\mathsf{Encaps}_{\widetilde{pk_A^k}}(\,)$
$\widetilde{K_B} \leftarrow \mathcal{KEM}.\mathsf{Decaps}_{\widetilde{sk_A^k}}(\widetilde{C_B})$	$\xleftarrow{B, \widetilde{C_B}, r_B}$	$r_B \xleftarrow{U} \{0,1\}^{256}$
$C_A, K_A \leftarrow \mathcal{KEM}.\mathsf{Encaps}_{pk_B^k}(\,)$	$\xrightarrow{A, C_A}$	$K_A \leftarrow \mathcal{KEM}.\mathsf{Decaps}_{sk_B^k}(C_A)$
$K_B \leftarrow \mathcal{KEM}.\mathsf{Decaps}_{sk_A^k}(C_B)$	$\xleftarrow{C_B}$	$C_B, K_B \leftarrow \mathcal{KEM}.\mathsf{Encaps}_{pk_A^k}(\,)$
$K, K^a \leftarrow \mathsf{KDF}(\widetilde{K_B}, K_A, K_B)$		$K, K^a \leftarrow \mathsf{KDF}(\widetilde{K_B}, K_A, K_B)$
$src \leftarrow \widetilde{pk_A^k}, r_A, B, \widetilde{C_B}, r_B, A, C_A, C_B$		$src \leftarrow \widetilde{pk_A^k}, r_A, B, \widetilde{C_B}, r_B, A, C_A, C_B$
$\tau_A \leftarrow \mathsf{MAC}_{K^a}(src)$	$\xrightarrow{\tau_A}$	Verify τ_A
Verify τ_B	$\xleftarrow{\tau_B}$	$\tau_A \leftarrow \mathsf{MAC}_{K^a}(src, \tau_A)$
return K		**return** K

Рис. 17. Схема криптографического ядра протокола KEMTLS с предвычислениями (вариант двусторонней аутентификации без механизмов обеспечения анонимности)

Заключение

В работе приведены атаки на ряд AKE-протоколов (в частности, на протоколы SIGMA, SIGMA-R, STS-MAC, «Эхинацея-3» и постквантовый протокол BKM-KK), осуществление которых становится возможным при наличии у нарушителя способности навязывать участнику использование будущих открытых эфемерных значений. Приведены идеи противодействия таким атакам. Применение описанных идей позволяет защититься от приведённых в настоящей работе сценариев атак, но не гарантирует того, что нет какого-то иного успешного, но пока незамеченного сценария атаки.

Одним из способов достичь большей уверенности в стойкости AKE-протоколов, да и вообще произвольных криптосистем, является метод теоретико-сложностных сведений задачи взлома протокола к решению каких-либо математических задач, сложность которых проверена годами исследований. Однако для этого, прежде всего, необходимо разработать формальную модель безопасности таких систем. Таким образом, открытыми задачами по теме настоящей работы является разработка формальной модели безопасности, учитывающей возможность нарушителя навязывать участникам открытые эфемерные значения, а также построение сведения задачи взлома протоколов из п. 3 (или любых других протоколов) к известным трудным математическим задачам. Целесообразно также сформулировать набор синтезных принципов, следование которым позволит обеспечить защиту от атак из рассматриваемого класса.

ЛИТЕРАТУРА

1. *Алексеев Е. К.* Что плохого можно сделать, неправильно используя криптоалгоритмы? Симпозиум CTCrypt 2019. `https://ctcrypt.ru/files/files/2019/materials/29_Alekseyev.pdf`. 2019.
2. *Алексеев Е. К., Ахметзянова Л. Р., Божко А. А., Грибоедова Е. С.* Теоретическая криптография в реальных условиях. Блог компании КриптоПро. `https://cryptopro.ru/blog/2019/11/19/teoreticheskaya-cryptografiya-v-realnykh-usloviyakh`. 2020.
3. *Царегородцев К. Д., Грибоедова Е. С.* Еще раз о важности построения модели противника на примере протокола аутентификации 5G-AKA // Конференция РусКрипто'2022. `https://ruscrypto.ru/resource/archive/rc2022/files/02_tsaregorodsev_griboedova.pdf`. 2022.
4. *Degabriele J. P., Paterson K. G., and Watson G. J.* Provable security in the real world // IEEE Security & Privacy. 2011. V. 9. No. 3. P. 33–41.
5. *Алексеев Е. К., Ахметзянова Л. Р., Божко А. А. и др.* О возможностях нарушителя при атаках на некоторый класс протоколов аутентифицированной выработки общего ключа. Конференция РусКрипто'2022. `https://ruscrypto.ru/resource/archive/rc2022/files/02_alekseyev_akhmetzyanova_kutsenok_kyazhin.pdf`. 2022.
6. *Krawczyk H.* The order of encryption and authentication for protecting communications (or: How secure is SSL?) // LNCS. 2001. V. 2139. P. 310–331.
7. *Canvel B., Hiltgen A., Vaudenay S., and Vuagnoux M.* Password interception in a SSL/TLS channel // LNCS. 2003. V. 2729. P. 583–599.
8. *Kaufman C., Hoffman P., Nir Y., et al.* Internet Key Exchange Protocol Version 2 (IKEv2). RFC 7296. 2014.
9. *Sheffer Y. and Fluhrer S.* Additional Diffie — Hellman Tests for the Internet Key Exchange Protocol Version 2 (IKEv2). RFC 6989. 2013.
10. *Seye P. B. and Sarr A. P.* Enhanced modelling of authenticated key exchange security // LNCS. 2017. V. 10547. P. 36–52.

11. *Alekseev E. K., Babueva A. A., and Zazykina O. A.* AKE Zoo: 100 Two-Party Protocols (to be continued). Cryptology ePrint Archive. 2023. Paper 2023/1044.
12. *Huang H. and Cao Z.* Authenticated Key Exchange Protocols with Enhanced Freshness Properties. Cryptology ePrint Archive. 2009. Paper 2009/505.
13. *Krawczyk H.* SIGMA: The 'SIGn-and-MAc' approach to authenticated Diffie — Hellman and its use in the IKE protocols // LNCS. 2003. V. 2729. P. 400–425.
14. *Jeong I. R., Katz J., and Lee D. H.* One-round protocols for two-party authenticated key exchange // LNCS. 2004. V. 3089. P. 220–232.
15. *Jeong I. R., Katz J., and Lee D. H.* One-Round Protocols for Two-Party Authenticated Key Exchange. `https://www.cs.umd.edu/~jkatz/papers/1round_AKE.pdf`. 2008.
16. *Rescorla E.* The Transport Layer Security (TLS) Protocol Version 1.3. RFC 8446. 2018.
17. *Diffie W., Van Oorschot P. C., and Wiener M. J.* Authentication and authenticated key exchanges // Des. Codes Cryptogr. 1992. V. 2. P. 107–125.
18. Информационная технология. Криптографическая защита информации. Схемы выработки общего ключа с аутентификацией на основе открытого ключа. Р 1323565.1.004-2017. М.: Стандартинформ, 2017.
19. *Cremers C. and Feltz M.* One-round Strongly Secure Key Exchange with Perfect Forward Secrecy and Deniability. Cryptology ePrint Archive. 2011. Paper 2011/300.
20. *Song B. and Kim K.* Two-pass authenticated key agreement protocol with key confirmation // LNCS. 2000. V. 1977. P. 237–249.
21. *Boyd C., Kock B., and Millerjord L.* Modular Design of KEM-Based Authenticated Key Exchange. Cryptology ePrint Archive. 2023. Paper 2023/167.
22. *Schwabe P., Stebila D., and Wiggers T.* Post-quantum TLS without handshake signatures // Proc. 2020 ACM SIGSAC Conf. CCS'20. USA, 2020. P. 1461–1480.

REFERENCES

1. *Alekseev E. K.* Chto plokhogo mozhno sdelat', nepravil'no ispol'zuya kriptoalgoritmy? [What bad things can be done by using cryptoalgorithms incorrectly?] CTCrypt 2019 Symp. `https://ctcrypt.ru/files/files/2019/materials/29_Alekseyev.pdf`. 2019. (in Russian)
2. *Alekseev E. K., Akhmetzyanova L. R., Bozhko A. A., and Griboedova E. S.* Teoreticheskaya kriptografiya v real'nykh usloviyakh [Theoretical cryptography in the real world]. CryptoPro Blog. `https://cryptopro.ru/blog/2019/11/19/teoreticheskaya-cryptografiya-v-realnykh-usloviyakh`. 2020. (in Russian)
3. *Tsaregorodtsev K. D. and Griboedova E. S.* Yeshche raz o vazhnosti postroyeniya modeli protivnika na primere protokola autentifikatsii 5G-AKA [On the importance of making an adversary model, once again, for the 5G-AKA authentication protocol example]. RusCrypto'2022 Conf. `https://ruscrypto.ru/resource/archive/rc2022/files/02_tsaregorodsev_griboedova.pdf`. 2022. (in Russian)
4. *Degabriele J. P., Paterson K. G., and Watson G. J.* Provable security in the real world. IEEE Security & Privacy, 2011, vol. 9, no. 3, pp. 33–41.
5. *Alekseev E. K., Akhmetzyanova L. R., Bozhko A. A., et al.* O vozmozhnostyakh narushitelya pri atakakh na nekotoryy klass protokolov autentifitsirovannoy vyrabotki obshchego klyucha [On the adversary capabilities needed to attack a certain class of authenticated key establishment protocols]. RusCrypto'2022 Conf. `https://ruscrypto.ru/resource/archive/rc2022/files/02_alekseyev_akhmetzyanova_kutsenok_kyazhin.pdf`. 2022. (in Russian)
6. *Krawczyk H.* The order of encryption and authentication for protecting communications (or: How secure is SSL?). LNCS, 2001, vol. 2139, pp. 310–331.

7. *Canvel B., Hiltgen A., Vaudenay S., and Vuagnoux M.* Password interception in a SSL/TLS channel. LNCS, 2003, vol. 2729, pp. 583–599.
8. *Kaufman C., Hoffman P., Nir Y., et al.* Internet Key Exchange Protocol Version 2 (IKEv2). RFC 7296, 2014.
9. *Sheffer Y. and Fluhrer S.* Additional Diffie — Hellman Tests for the Internet Key Exchange Protocol Version 2 (IKEv2). RFC 6989, 2013.
10. *Seye P. B. and Sarr A. P.* Enhanced modelling of authenticated key exchange security. LNCS, 2017, vol. 10547, pp. 36–52.
11. *Alekseev E. K., Babueva A. A., and Zazykina O. A.* AKE Zoo: 100 Two-Party Protocols (to be continued). Cryptology ePrint Archive, 2023, Paper 2023/1044.
12. *Huang H. and Cao Z.* Authenticated Key Exchange Protocols with Enhanced Freshness Properties. Cryptology ePrint Archive, 2009, Paper 2009/505.
13. *Krawczyk H.* SIGMA: The 'SIGn-and-MAc' approach to authenticated Diffie — Hellman and its use in the IKE protocols. LNCS, 2003, vol. 2729, pp. 400–425.
14. *Jeong I. R., Katz J., and Lee D. H.* One-round protocols for two-party authenticated key exchange. LNCS, 2004, vol. 3089, pp. 220–232.
15. *Jeong I. R., Katz J., and Lee D. H.* One-Round Protocols for Two-Party Authenticated Key Exchange. `https://www.cs.umd.edu/~jkatz/papers/1round_AKE.pdf`. 2008.
16. *Rescorla E.* The Transport Layer Security (TLS) Protocol Version 1.3. RFC 8446, 2018.
17. *Diffie W., Van Oorschot P. C., and Wiener M. J.* Authentication and authenticated key exchanges. Des. Codes Cryptogr., 1992, vol. 2, pp. 107–125.
18. Informatsionnaya tekhnologiya. Kriptograficheskaya zashchita informatsii. Skhemy vyrabotki obshchego klyucha s autentifikatsiyey na osnove otkrytogo klyucha [Information Technology. Information Cryptographic Protection. Public Key Based on the Authenticated Key Agreement Schemes]. R 1323565.1.004-2017. Moscow, Standartinform Publ., 2017. (in Russian)
19. *Cremers C. and Feltz M.* One-round Strongly Secure Key Exchange with Perfect Forward Secrecy and Deniability. Cryptology ePrint Archive, 2011, Paper 2011/300.
20. *Song B. and Kim K.* Two-pass authenticated key agreement protocol with key confirmation. LNCS, 2000, vol. 1977, pp. 237–249.
21. *Boyd C., Kock B., and Millerjord L.* Modular Design of KEM-Based Authenticated Key Exchange. Cryptology ePrint Archive, 2023, Paper 2023/167.
22. *Schwabe P., Stebila D., and Wiggers T.* Post-quantum TLS without handshake signatures. Proc. 2020 ACM SIGSAC Conf. CCS'20, USA, 2020, pp. 1461–1480.

ПРИКЛАДНАЯ ТЕОРИЯ АВТОМАТОВ

УДК 519.7 DOI 10.17223/20710410/66/7

О СВОЙСТВАХ КОНЕЧНО-АВТОМАТНОГО ГЕНЕРАТОРА[1]

А. О. Бахарев*, Р. О. Запанов*, С. Е. Зинченко*, И. А. Панкратова**, Е. С. Прудников**

*Новосибирский государственный университет, г. Новосибирск, Россия
**Томский государственный университет, г. Томск, Россия

E-mail: a.bakharev@g.nsu.ru, rinchinzapanov@yandex.ru, s.zinchenko@alumni.nsu.ru, pank@mail.tsu.ru, egorprudnikov71@gmail.com

Рассматриваются периодические свойства двухкаскадного конечно-автоматного криптографического генератора. Сформулированы некоторые необходимые условия того, что выходная последовательность генератора имеет период максимально возможной длины. Получены также достаточные условия, на основании которых предложен способ построения такого генератора. Доказано, что для любой двоичной последовательности, период которой равен степени двойки, существует генератор, выдающий её.

Ключевые слова: *конечный автомат, криптографический генератор, криптоавтомат, период последовательности.*

ON THE PROPERTIES OF A FINITE-STATE GENERATOR

A. O. Bakharev*, R. O. Zapanov*, S. E. Zinchenko*, I. A. Pankratova**, E. S. Prudnikov**

*Novosibirsk State University, Novosibirsk, Russia
**Tomsk State University, Tomsk, Russia

The periodic properties of a two-stage finite-state 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), $A_2 = (\mathbb{F}_2, \mathbb{F}_2^m, \mathbb{F}_2, g_2, f_2)$, $n, m \geqslant 1$. Some necessary conditions for such a generator with the maximum period of 2^{n+m} have been formulated, namely: 1) the output sequence of A_1 is purely periodic and the period length is 2^n; 2) the substitution G_u transforming any initial state $y(1)$ of the automaton A_2 into the state $y(2^n+1)$ is a full-cycle substitution; 3) the function f_1 has an odd weight; 4) the substitutions $g(0,\cdot)$ and $g(1,\cdot)$ have different parities. Some sufficient conditions have been also formulated, for example, in addition to conditions 1–4, the function $g_2(u, y)$ must be injective in u and the weight of the function f_2 must be odd. Two methods for constructing a generator having maximum period have been proposed. It has been proved that, for any binary sequence whose period is a power of two, there exists a generator that produces it.

Keywords: *finite state machine, cryptographic generator, cryptoautomaton, sequence period.*

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

Введение

В работе [1] Г.П. Агибаловым введено понятие криптоавтомата как класса автоматных сетей с ключом, который может включать в себя начальные состояния компонент сети и их функции переходов и выходов. Этому определению криптоавтомата соответствуют различные криптографические примитивы: генераторы ключевого потока MUGI [2] и KNOT [3] — в поточных шифрах, симметричный конечно-автоматный шифр Закревского [4], конечно-автоматные криптосистемы с открытым ключом для шифрования и цифровой подписи семейства FAPKC [5] и другие.

Рассматриваемый в данной работе генератор является, с одной стороны, частным случаем последовательной композиции $A_1 \cdot A_2$ в модели [1] (A_1 — автомат Мура); с другой — обобщением конечно-автоматного генератора (δ, τ)-шагов [6] (функция переходов автомата A_2 произвольна). В [7] изучены некоторые задачи криптоанализа генератора, в [8] — его периодические свойства. В продолжение этих исследований в работе получены условия максимальности периода выходной последовательности генератора и предложены способы построения генераторов с таким свойством.

1. Базовые определения и обозначения

Пусть $\mathbb{F}_2 = \{0, 1\}$; *весом* $\mathrm{wt}(f)$ булевой функции $f : \mathbb{F}_2^n \to \mathbb{F}_2$, $n \in \mathbb{N}$, будем называть

$$\mathrm{wt}(f) = |\,\{x \in \mathbb{F}_2^n : f(x) = 1\}\,|.$$

Последовательность $\{u_i : 1, 2, \dots\}$, $u_i \in \mathbb{F}_2$, называется *периодической*, если для некоторых $i_0, t \in \mathbb{N}$ выполнено $u_i = u_{i+t}$ для всех $i \geqslant i_0$. Минимальное t с таким свойством называется *периодом* последовательности. Периодическая последовательность называется *чисто периодической*, если $i_0 = 1$.

Пусть $\sigma \in \mathbb{S}_n$ — подстановка степени n, $\sigma = \tau_1 \circ \dots \circ \tau_k$ — произвольное разложение σ в произведение транспозиций. Число $\mathsf{sgn} = (-1)^k$ называется *знаком* σ, полностью определяется подстановкой σ и не зависит от способа разложения в произведение транспозиций.

Лемма 1 [9]. Пусть $\sigma \in \mathbb{S}_n$ и c — число попарно независимых циклов в σ. Тогда $\mathsf{sgn}(\sigma) = (-1)^{n-c}$.

2. Конечно-автоматный генератор

Схема двухкаскадного конечно-автоматного криптографического генератора $G = A_1 \cdot A_2$ представлена на рис. 1: это 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 $f_2 : \mathbb{F}_2 \times \mathbb{F}_2^m \to \mathbb{F}_2$), $n, m \geqslant 1$.

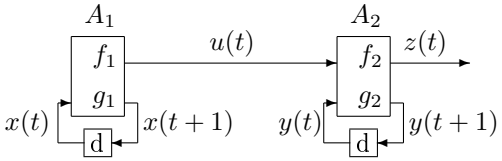

Рис. 1. Схема генератора G

В каждый момент времени $t = 1, 2, \dots$ 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 символ $z(t) = f_2(u(t), y(t))$ и переходит в следующее состояние $y(t+1) = g_2(u(t), y(t))$. Ключом генератора может быть любое непустое подмножество множества $\{x(1), y(1), f_1, g_1, f_2, g_2\}$.

Периодом генератора назовём период его выходной последовательности $z(1)z(2)\ldots$ В [8] доказано, что период генератора не превосходит 2^{n+m}. Сформулируем условия достижения верхней оценки.

3. Необходимые условия максимальности периода генератора

Обозначим через $g_2^\delta = g_2(\delta, \cdot)$, $\delta \in \{0, 1\}$, подфункции функции g_2.

Утверждение 1 [8]. Если период генератора G равен 2^{n+m}, то:

1) функция g_1 является полноцикловой подстановкой;
2) подфункции g_2^0 и g_2^1 являются подстановками;
3) $y(2^n i + j) \neq y(2^n k + j)$ для всех $i, k \in \{0, \ldots, 2^m - 1\}$, $i \neq k$, $j = 1, \ldots, 2^n$;
4) выходная последовательность $z(1)z(2)\ldots$ генератора чисто периодическая.

Дополним список необходимых условий:

Утверждение 2. Если период генератора G равен 2^{n+m}, то последовательность $u(1)u(2)\ldots$ чисто периодическая и её период равен 2^n.

Доказательство. Из [10, утверждение 6.1, п. 2], п. 1 утверждения 1 и формулы $u(t) = f_1(g_1(t))$ следует, что последовательность $\{u(t) : t = 1, 2, \ldots\}$ чисто периодическая и её минимальный период s делит 2^n. Предположим, что $s < 2^n$, тогда $s \,|\, 2^{n-1}$ и $u(j) = u(2^{n-1} i + j)$ для всех i, j.

По п. 3 утверждения 1 для любого j, $1 \leqslant j \leqslant 2^n$, значения $y(2^n i + j)$, $i = 0, \ldots, 2^m - 1$, попарно различны, а всего таких значений 2^m. Значит, для $j = 2^{n-1} + 1$ найдётся i, такое, что $y(1) = y(2^n i + 2^{n-1} + 1) = y(2^{n-1} k + 1)$, где $k = 2i + 1$. Из того, что $u(1) = u(2^{n-1} k + 1)$, и описания функционирования генератора заключаем, что $z(1) = z(2^{n-1} k + 1)$ и $y(2) = y(2^{n-1} k + 2)$. Продолжая по индукции, получим: $z(j) = z(2^{n-1} k + j)$ для всех j, т. е. период генератора делит $2^{n-1} k = 2^n i + 2^{n-1} \leqslant$ $\leqslant 2^n (2^m - 1) + 2^{n-1} = 2^{n+m} - 2^{n-1}$, что противоречит условию. ■

Для $u = u(1)u(2)\ldots u(2^n)$ обозначим через $G_u : \mathbb{F}_2^m \to \mathbb{F}_2^m$ композицию подстановок

$$G_u = g_2^{u(1)} \circ \ldots \circ g_2^{u(2^n)},$$

другими словами, $G_u(y(1)) = y(2^n + 1)$ для всех $y(1) \in \mathbb{F}_2^m$, т. е. G_u — это подстановка на множестве состояний автомата A_2, переводящая любое его начальное состояние в то, в которое автомат перейдёт после одного цикла последовательности на входе.

В следующем утверждении приведено, в том числе (в п. 2), решение задачи 9 из второго раунда десятой Международной олимпиады по криптографии Non-Stop University CRYPTO [11, 12].

Утверждение 3. Если период генератора G равен 2^{n+m}, то:

1) подстановка G_u полноцикловая;
2) вес функции f_1 нечётный;
3) подстановки g_2^0 и g_2^1 имеют разную чётность.

Доказательство.

1) Следует из п. 3 утверждения 1.

2) Пусть $\mathrm{wt}(f_1) = k$. Из утверждения 2 следует, что в отрезке $u = u(1)u(2)\ldots u(2^n)$ содержатся значения функции f_1 на всех наборах значений её аргументов, т. е. k единиц и $2^n - k$ нулей. Следовательно,

$$\mathsf{sgn}(G_u) = \mathsf{sgn}(g_2^1)^k \cdot \mathsf{sgn}(g_2^0)^{2^n-k}. \tag{1}$$

С другой стороны, по лемме 1 и ввиду п. 1

$$\mathsf{sgn}(G_u) = (-1)^{2^m-1} = -1. \tag{2}$$

Это возможно только при нечётном k.

3) Если $\mathsf{sgn}(g_2^0) = \mathsf{sgn}(g_2^1)$, то $\mathsf{sgn}(G_u) = \mathsf{sgn}(g_2^0)^{2^n} = 1$ — противоречие с (2). ■

Условия утверждений 1–3 не являются достаточными для максимальности периода генератора, в частности, потому, что не накладывают никаких ограничений на функцию выходов автомата A_2. В следующем пункте эти условия дополняются ещё двумя, что даёт достаточные (но теперь не необходимые) условия максимальности периода.

4. Достаточные условия максимальности периода генератора

Утверждение 4. Пусть для генератора G выполнены следующие условия:

1) последовательность $u(1)u(2)\ldots$ чисто периодическая с периодом 2^n;
2) подфункции g_2^0 и g_2^1 — подстановки, их композиция $G_u = g_2^{u(1)} \circ \ldots \circ g_2^{u(2^n)}$ — полноцикловая подстановка;
3) функция $g_2(u, y) : \mathbb{F}_2 \times \mathbb{F}_2^m \to \mathbb{F}_2^m$ инъективна по переменной u.

Тогда период последовательности состояний $(y(t) : t \in \mathbb{N})$ равен 2^{n+m}, а отображение

$$\psi_G : \mathbb{F}_2^n \times \mathbb{F}_2^m \to \mathbb{F}_2^n \times \mathbb{F}_2^m,\ \big(x(t), y(t)\big) \mapsto \big(x(t+1), y(t+1)\big),$$
$$\big(x(t+1), y(t+1)\big) = \big(g_1(x(t)), g_2(f_1(x(t)), y(t))\big)$$

является полноцикловой подстановкой.

Доказательство. Обозначим через π период последовательности $(y(t) : t \in \mathbb{N})$. Из условий 1 и 3 ввиду [10, следствие 1 теоремы 7.5] получаем, что $2^n \,|\, \pi$. Условие 2 означает, что значения $y(1), y(2^n + 1), \ldots, y(2^n(2^m - 1) + 1)$ попарно различны. Следовательно, $\pi \geqslant 2^{n+m}$. Неравенство $\pi \leqslant 2^{n+m}$ очевидно в силу формулы $y(t+1) = g_2(f_1(x(t)), y(t))$ и того, что количество разных пар $(x(t), y(t))$ равно 2^{n+m}.

Поскольку $u(t) = f_1(x(t))$, из условия 1 следует, что период последовательности $(x(t) : t \in \mathbb{N})$ равен 2^n. Тогда по [10, утверждение 6.2, п. 1] получаем, что период последовательности $\big((x(t), y(t)) : t \in \mathbb{N}\big)$ равен $\mathrm{НОК}(2^n, 2^{n+m}) = 2^{n+m}$, т. е. отображение ψ_G — полноцикловая подстановка. ■

Утверждение 5. Если генератор G удовлетворяет условиям утверждения 4 и вес функции f_2 нечётный, то период генератора равен 2^{n+m}.

Доказательство. Обозначим: $k = \mathrm{wt}(f_1)$; N — количество единиц в отрезке $z = z(1)z(2)\ldots z(2^{n+m})$; π — период этой последовательности; N_0 и N_1 — веса подфункций $f_2(0, y)$ и $f_2(1, y)$ соответственно. Из условия и формул (1) и (2) следует, что k нечётно; из того, что нечётен вес $\mathrm{wt}(f_2) = N_0 + N_1$, заключаем, что числа N_0 и N_1 имеют разную чётность.

Период последовательности $\big((x(t), y(t)) : t \in \mathbb{N}\big)$ равен 2^{n+m} (следует из утверждения 4). Тогда ввиду $z(t) = f_2(f_1(x(t)), y(t))$ выполняется

$$N = (2^n - k)N_0 + kN_1,$$

т. е. N нечётно. По [10, утверждение 6.1, п. 2] имеем $\pi \,|\, 2^{n+m}$; при $\pi < 2^{n+m}$ период повторяется в отрезке z чётное число раз, следовательно, и число N должно быть чётным. Значит, $\pi = 2^{n+m}$. ■

Условия утверждений 4 и 5 являются достаточными, но не необходимыми для максимальности периода генератора; так, в примере 1 функция g_2 не инъективна по u и вес функции f_2 чётный, однако период генератора равен 2^{n+m}.

Пример 1. Пусть $n = 1$, $m = 2$, $g_1(x) = x \oplus 1$, $f_1(x) = x$, $x(1) = 0$, $y(1) = 00$, функции g_2 и f_2 заданы табл. 1 и 2.

Таблица 1
Функция g_2

$u(t)$	$y(t)$			
	00	01	10	11
0	01	10	11	00
1	01	10	00	11

Таблица 2
Функция f_2

$u(t)$	$y(t)$			
	00	01	10	11
0	0	1	0	0
1	0	0	1	0

Тогда $(u(t) : t = 1, 2) = (0, 1)$, $(y(t) : t = 1, \ldots, 8) = (00, 01, 10, 11, 11, 00, 01, 10)$, $G_u = (00, 10, 11, 01)$ — полноцикловая; $(z(t) : t \in \mathbb{N}) = 00000011\ldots$ — периодическая с периодом $8 = 2^{n+m}$.

5. Построение генераторов максимального периода

Утверждение 6. Если $\mathrm{wt}(f_1) = k$ — нечётное число, то для любых $n, m \geqslant 1$ существуют такие функции g_1, g_2, f_2, что период генератора G равен 2^{n+m}.

Доказательство. Сопоставим векторам $x \in \mathbb{F}_2^n$ и $y \in \mathbb{F}_2^m$ числа из $\mathbb{Z}_{2^n}$ и $\mathbb{Z}_{2^m}$, двоичными представлениями которых они являются, и определим функции

$$g_1(x) = (x + 1) \bmod 2^n, \quad g_2(u, y) = (y + u) \bmod 2^m, \quad f_2(u, y) = \mathbb{I}[y = 0],$$

где $\mathbb{I}$ — функция-индикатор. Отметим, что g_1 и g_2^1 являютя полноцикловыми подстановками, g_2^0 — тождественное отображение. Пусть, без ограничения общности, $y(1) = 0$ и, следовательно, $z(1) = 1$; период генератора G (последовательности $(z(t) : t \in \mathbb{N})$) обозначим через π.

По построению $z(t) = z(t + 2^{n+m})$, $t \in \mathbb{N}$. Тогда $\pi = 2^p$ для $p \leqslant n + m$; $z(1 + 2^p) = 1$, а значит,

$$y(1 + 2^p l) = 0, \quad l \in \mathbb{N}. \tag{3}$$

Поскольку за 2^n тактов работы генератора последовательность u на входе автомата A_2 пробегает значения функции f_1 на всех наборах, имеем $y(1 + 2^n) = k \bmod 2^m$ и $y(1 + 2^n) \neq 0$ ввиду нечётности k.

Если $p < n$, то $y(1 + 2^n) = y(1 + 2^p 2^{n-p}) = 0$ по (3) — противоречие.

Если $p > n$, то $y(1 + 2^p) = y(1 + 2^n 2^{p-n}) = 2^{p-n}k = 0 \bmod 2^m$, что невозможно при нечётном k и $p < n + m$. Следовательно, $\pi = 2^{n+m}$. ■

Замечание 1. Для функции f_2, построенной в доказательстве утверждения 6, не выполнены условия утверждения 5, так как $\mathrm{wt}(f_2) = 2$ — чётное число.

Утверждения 1 (п. 1) и 3 (п. 2) задают необходимые требования к автомату A_1 для построения генератора максимального периода. Для функции переходов g_2 автомата A_2 должны выполняться условия 2 и 3 утверждения 4, при этом условие 2 зависит от выходной последовательности автомата A_1. Опишем способ построения такой функции g_2, что условия утверждения 4 выполнены для любого автомата A_1, удовлетворяющего необходимым условиям.

Утверждение 7. Пусть функция $g_2(u, y) : \mathbb{F}_2 \times \mathbb{F}_2^m \to \mathbb{F}_2^m$ такова, что g_2^0 — полноцикловая подстановка, g_2^1 — любая её чётная степень (или наоборот). Тогда:

1) функция $g_2(u, y)$ инъективна по u;
2) композиция $G_u = g_2^{u(1)} \circ \ldots \circ g_2^{u(2^n)}$ — полноцикловая подстановка для любой последовательности $u = u(1) \ldots u(2^n)$ нечётного веса.

Доказательство. Как и при доказательстве утверждения 6, будем отождествлять векторы $y \in \mathbb{F}_2^m$ и числа из $\mathbb{Z}_{2^m}$. Пусть, без ограничения общности, $g_2^0(y) = (y+1) \bmod 2^m$ и $g_2^1 = (g_2^0)^l$, l — чётное. Тогда $g_2^1(y) = (y+l) \bmod 2^m$.

1) При чётном l для любого $y \in \mathbb{F}_2^m$ имеем

$$g_2(0, y) = (y+1) \bmod 2^m \neq (y+l) \bmod 2^m = g_2(1, y).$$

Следовательно, функция $g_2(u, y)$ инъективна по u.

2) Пусть последовательность u содержит k единиц. Тогда $G_u = (g_2^0)^s$, где $s = 2^n - k + lk$ — нечётное в силу чётности l и нечётности k. Рассмотрим r-ю степень этой подстановки: $G_u^r = (g_2^0)^{sr}$; и уравнение $G_u^r(y) = y$:

$$G_u^r(y) = y + sr = y \pmod{2^m},$$

которое верно только при $r = 0 \pmod{2^m}$. Следовательно, G_u — полноцикловая. ■

Из утверждений 4, 5 и 7 получаем метод построения генератора G, имеющего максимальный период (алгоритм 1).

Алгоритм 1. Построение генератора G максимального периода

Вход: $n, m \in \mathbb{N}$.
Выход: функции переходов и выходов автоматов A_1 и A_2.
Выбрать:
1: $g_1 : \mathbb{F}_2^n \to \mathbb{F}_2^n$ — произвольная полноцикловая подстановка;
2: $f_1 : \mathbb{F}_2^n \to \mathbb{F}_2$ — произвольная функция нечётного веса;
3: $g_2^0 : \mathbb{F}_2^m \to \mathbb{F}_2^m$ — произвольная полноцикловая подстановка;
g_2^1 — любая её чётная степень (или наоборот);
4: $f_2 : \mathbb{F}_2^{m+1} \to \mathbb{F}_2$ — произвольная функция нечётного веса.

Алгоритм 1, очевидно, не обладает свойством полноты (с его помощью нельзя построить все возможные генераторы максимального периода), поскольку условия, на которые он опирается, не являются необходимыми. Однако, если не фиксировать значения n и m, то для любой двоичной последовательности, период которой равен степени двойки, можно предложить генератор, выдающий её.

Утверждение 8. Для любых $l \geqslant 2$ и $z = z(1) \ldots z(2^l) \in \mathbb{F}_2^l$ существует двухкаскадный конечно-автоматный генератор, выходная последовательность которого равна $z \cdot z \cdot \ldots$ (здесь «$\cdot$» — операция конкатенации).

Доказательство. Опишем способ построения генератора G.

Положим $n = 1$, $m = l - 1$; как и прежде, будем отождествлять векторы из $\mathbb{F}_2^m$ и числа из $\mathbb{Z}_{2^m}$. Зададим функции автомата A_1 как $g_1(x) = x \oplus 1$, $f_1(x) = x$, начальное состояние $x(0) = 0$; тогда на вход автомата A_2 поступит последовательность $u = 010101 \ldots$

Определим функцию $g_2(u, y) = (y + u) \bmod 2^m$ и начальное состояние $y(0) = 0$. Тогда пара «вход/состояние» автомата A_2 за первые 2^l шагов будет пробегать значения

$$(0,0),\ (1,0),\ (0,1)\ (1,1),\ \ldots,\ (0, 2^m - 1),\ (1, 2^m - 1).$$

Для получения последовательности z на выходе генератора осталось положить $f_2(0, y) = z(2y + 1)$ и $f_2(1, y) = z(2y + 2)$, $y = 0, 1, \ldots, 2^m - 1$. ■

Замечание 2. Если последовательность z в утверждении 8 апериодична, то способом, описанным в его доказательстве, получим генератор максимального периода.

Заключение

Рассмотрены условия, при которых двухкаскадный конечно-автоматный генератор имеет максимальный период; дополнен список необходимых условий и получены некоторые достаточные, на основании которых предложен простой метод построения такого генератора. Однако найденные необходимые условия не являются достаточными и наоборот, т. е. критерий максимальности периода генератора не сформулирован.

Некоторые из результатов работы докладывались на конференции SIBECRYPT'24, их краткое изложение можно найти в [13].

ЛИТЕРАТУРА

1. *Агибалов Г. П.* Криптоавтоматы с функциональными ключами // Прикладная дискретная математика. 2017. № 36. С. 59–72.
2. *Watanabe D., Furuya S., Yoshida H., et al.* A new keystream generator MUGI // LNCS. 2002. V. 2365. P. 179–194.
3. *Joux A. and Muller F.* Loosening the KNOT // LNCS. 2003. V. 2887. P. 87–99.
4. *Закревский А. Д.* Метод автоматической шифрации сообщений // Прикладная дискретная математика. 2009. № 2(4). С. 127–137.
5. *Tao R.* Finite Automata and Application to Cryptography. TSINGHUA University Press, 2009. 406 p.
6. *Агибалов Г. П., Панкратова И. А.* О двухкаскадных конечно-автоматных криптографических генераторах и методах их криптоанализа // Прикладная дискретная математика. 2017. № 35. С. 38–47.
7. *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.
8. *Обухов П. К., Панкратова И. А.* Периодические свойства конечно-автоматного генератора // Прикладная дискретная математика. Приложение. 2023. № 16. С. 141–143.
9. *Кострикин А. И.* Введение в алгебру. Ч. 1: Основы алгебры: учебник для вузов. М.: Физматлит, 2000. 272 с.
10. *Фомичев В. М.* Методы дискретной математики в криптологии. М.: Диалог-МИФИ, 2010. 424 с.
11. *Idrisova V. A., Tokareva N. N., Gorodilova A. A., et al.* Mathematical problems and solutions of the Ninth International Olympiad in Cryptography NSUCRYPTO // Прикладная дискретная математика. 2023. № 62. С. 29–54.
12. `https://nsucrypto.nsu.ru/`
13. *Прудников Е. С.* Конечно-автоматные генераторы максимального периода // Прикладная дискретная математика. Приложение. 2024. № 17. С. 152–154.

REFERENCES

1. *Agibalov G. P.* Kriptoavtomaty s funktsional'nymi klyuchami [Cryptautomata with functional keys]. Prikladnaya Diskretnaya Matematika, 2017, no. 36, pp. 59–72. (in Russian)
2. *Watanabe D., Furuya S., Yoshida H., et al.* A new keystream generator MUGI. LNCS, 2002, vol. 2365, pp. 179–194.

3. *Joux A. and Muller F.* Loosening the KNOT. LNCS, 2003, vol. 2887, pp. 87–99.
4. *Zakrevskiy A. D.* Metod avtomaticheskoy shifratsii soobshcheniy [The method for messages automatic encryption]. Prikladnaya Diskretnaya Matematika, 2009, no. 2(4), pp. 127–137. (in Russian)
5. *Tao R.* Finite Automata and Application to Cryptography. TSINGHUA University Press, 2009. 406 p.
6. *Agibalov G. P. and Pankratova I. A.* O dvukhkaskadnykh konechno-avtomatnykh kriptograficheskikh generatorakh i metodakh ikh kriptoanaliza [About 2-cascade finite automata cryptographic generators and their cryptanalysis]. Prikladnaya Diskretnaya Matematika, 2017, no. 35, pp. 38–47. (in Russian)
7. *Borovkova I. V., Pankratova I. A., and Semenova E. V.* Kriptoanaliz dvukhkaskadnogo konechno-avtomatnogo generatora s funktsional'nym klyuchom [Cryptanalysis of 2-cascade finite automata generator with functional key]. Prikladnaya Diskretnaya Matematika, 2018, no. 42, pp. 48–56. (in Russian)
8. *Obukhov P. K. and Pankratova I. A.* Periodicheskie svoystva konechno-avtomatnogo generatora [Periodic properties of a finite automaton generator]. Prikladnaya Diskretnaya Matematika. Prilozhenie, 2023, no. 16, pp. 141–143. (in Russian)
9. *Kostrikin A. I.* Vvedenie v algebru. Ch. 1: Osnovy algebry [Introduction to Algebra. P. 1: Fundamentals of Algebra. Textbook for Universities.] Moscow, Fizmatlit Publ., 2000. 272 p. (in Russian)
10. *Fomichev V. M.* Metody diskretnoy matematiki v kriptologii [Methods of Discrete Mathematics in Cryptology]. Moscow, Dialog-MIFI Publ., 2010. 424 p. (in Russian)
11. *Idrisova V. A., Tokareva N. N., Gorodilova A. A., et al.* Mathematical problems and solutions of the Ninth International Olympiad in Cryptography NSUCRYPTO. Prikladnaya Diskretnaya Matematika, 2023, no. 62, pp. 29–54.
12. `https://nsucrypto.nsu.ru/`
13. *Prudnikov E. S.* Konechno-avtomatnye generatory maksimal'nogo perioda [Finite-state generators with maximal period]. Prikladnaya Diskretnaya Matematika. Prilozhenie, 2024, no. 17, pp. 152–154. (in Russian)

ПРИКЛАДНАЯ ТЕОРИЯ ГРАФОВ

УДК 519.172.3 DOI 10.17223/20710410/66/8

О СУЩЕСТВОВАНИИ СИЛЬНО РЕГУЛЯРНЫХ ОРГРАФОВ С НАБОРОМ ПАРАМЕТРОВ (22, 9, 6, 3, 4)

В. А. Бызов, И. А. Пушкарев

Вятский государственный университет, г. Киров, Россия

E-mail: vbyzov@yandex.ru

Показано существование семейства сильно регулярных орграфов с набором параметров (22, 9, 6, 3, 4). Матрицы смежности найденных орграфов состоят из циркулянтных блоков 3×3. Группой автоморфизмов всех найденных орграфов является группа $\mathbb{Z}_3$. Структура полученных орграфов описана при помощи понятий скелета и оснастки.

Ключевые слова: *сильно регулярный орграф, циркулянтная матрица, компактификация матриц, группа автоморфизмов, изоморфные орграфы.*

ON THE EXISTENCE OF DIRECTED STRONGLY REGULAR GRAPHS WITH PARAMETERS (22, 9, 6, 3, 4)

V. A. Byzov, I. A. Pushkarev

Vyatka State University, Kirov, Russia

The paper shows the existence of the family of directed strongly regular graphs with parameters (22, 9, 6, 3, 4). The adjacency matrices of the found digraphs consist of 3×3 circulant blocks. The automorphism group of all the digraphs found is the group $\mathbb{Z}_3$. The structure of the resulting digraphs has been described using the concepts of skeleton and rigging.

Keywords: *directed strongly regular graph, circulant matrix, compactification of matrices, automorphism group, isomorphic digraphs.*

Введение

Дадим основные понятия и обозначения, необходимые для дальнейшего изложения.

В работе рассматриваются исключительно ориентированные графы (далее — орграфы) без петель и кратных дуг одного направления. Вершины таких орграфов будем нумеровать натуральными числами, начиная с единицы. Матрицы смежности будем формировать одним из стандартных для орграфов способов, а именно: писать единицу в i-й строке и j-м столбце, если в орграфе существует дуга, идущая из вершины i в вершину j. Все остальные элементы матрицы равны нулю.

Для единичной матрицы порядка n будем использовать обозначение I_n, для квадратной матрицы порядка n, целиком состоящей из единиц, — обозначение J_n. Запись $J_{k,l}$ обозначает матрицу $k \times l$, состоящую из единиц.

Понятие сильно регулярного орграфа введено А. М. Дювалем в работе [1] как ориентированное обобщение концепции сильно регулярных неориентированных графов [2]. Приведём два эквивалентных определения такого орграфа.

Определение 1. Сильно регулярным орграфом с набором параметров (v, k, t, λ, μ) называется орграф на v вершинах, удовлетворяющий следующему набору условий:

1) полустепени исхода и полустепени захода всех вершин равны k;
2) для любой вершины x имеется ровно t путей вида $x \to z \to x$;
3) для любой дуги $x \to y$ есть ровно λ путей вида $x \to z \to y$;
4) если в орграфе нет дуги $x \to y$, то имеется ровно μ путей вида $x \to z \to y$.

Определение 2. Сильно регулярным орграфом с набором параметров (v, k, t, λ, μ) называется орграф, матрица смежности A которого удовлетворяет следующим соотношениям:

$$A^2 = tI_v + \lambda A + \mu(J_v - I_v - A); \tag{1}$$

$$AJ_v = J_v A = kJ_v. \tag{2}$$

Орграфы описанного типа будем обозначать $\mathrm{dsrg}(v, k, t, \lambda, \mu)$. В работе [1] описан ряд необходимых условий, которым должны удовлетворять параметры сильно регулярного орграфа. Но, как и в случае сильно регулярных графов, есть большое число наборов параметров, для которых неизвестно, существует ли сильно регулярный орграф с данными параметрами. А. Э. Брауэр на сайте [3] систематизирует информацию о том, для каких наборов параметров орграфы существуют (и приводит эти орграфы), а для каких вопрос их существования пока открыт. До текущего момента наименьшим таким случаем (в смысле количества вершин) был вопрос существования орграфа $\mathrm{dsrg}(22, 9, 6, 3, 4)$. В этой работе построено семейство орграфов с таким набором параметров.

Замечание 1. Отметим некоторую несогласованность в порядке перечисления параметров сильно регулярного орграфа разными авторами. В основополагающей работе [1] используется следующий порядок параметров: (v, k, μ, λ, t). В других работах (например, [4]) параметры перечисляются в порядке (v, k, t, λ, μ). Второй способ нам кажется более удачным, потому что он лучше согласуется с общепринятым обозначением сильно регулярных графов $\mathrm{srg}(v, k, \lambda, \mu)$, поэтому в работе используются именно такие обозначения.

Циркулянтная матрица (или циркулянт) — это квадратная матрица вида

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \\ a_n & a_1 & \dots & a_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & \dots & a_1 \end{pmatrix}, \tag{3}$$

то есть это матрица, в которой каждая строка, начиная со второй, получается циклическим сдвигом предыдущей строки на одну позицию вправо.

Множество циркулянтных матриц порядка n с элементами из $\mathbb{Z}$ (со стандартными операциями умножения и сложения) образует кольцо, которое изоморфно факторкольцу $\mathbb{Z}[x]/(x^n - 1)$ [5]. При этом изоморфизме матрице (3) соответствует многочлен $a_1 + a_2 x + \ldots + a_n x^{n-1}$.

Назовём *компактификацией* блочной матрицы M, состоящей из циркулянтов, замену всех циркулянтов на соответствующие им при описанном изоморфизме многочлены. Полученную матрицу будем обозначать $M(x)$. Компактификация матриц согласована с операциями сложения и умножения матриц, т. е. если квадратные блочные матрицы M_1, M_2, M_3 и M_4 одного порядка, состоящие из циркулянтов $m \times m$, таковы, что $M_3 = M_1 + M_2$ и $M_4 = M_1 \cdot M_2$, то $M_3(x) \equiv M_1(x) + M_2(x) \pmod{x^m - 1}$ и $M_4(x) \equiv M_1(x)\, M_2(x) \pmod{x^m - 1}$.

В качестве примера рассмотрим матрицу смежности S графа Шрикханде [2, 6], который является сильно регулярным с набором параметров $(16, 6, 2, 2)$. При подходящей нумерации вершин она может быть представлена в следующем виде:

$$S = \left(\begin{array}{cccc|cccc|cccc|cccc}
0&0&0&0&0&0&1&1&0&0&1&1&0&1&0&1\\
0&0&0&0&1&0&0&1&1&0&0&1&1&0&1&0\\
0&0&0&0&1&1&0&0&1&1&0&0&0&1&0&1\\
0&0&0&0&0&1&1&0&0&1&1&0&1&0&1&0\\
\hline
0&1&1&0&0&0&0&0&0&1&0&1&0&0&1&1\\
0&0&1&1&0&0&0&0&1&0&1&0&1&0&0&1\\
1&0&0&1&0&0&0&0&0&1&0&1&1&1&0&0\\
1&1&0&0&0&0&0&0&1&0&1&0&0&1&1&0\\
\hline
0&1&1&0&0&1&0&1&0&0&0&0&1&1&0&0\\
0&0&1&1&1&0&1&0&0&0&0&0&0&1&1&0\\
1&0&0&1&0&1&0&1&0&0&0&0&0&0&1&1\\
1&1&0&0&1&0&1&0&0&0&0&0&1&0&0&1\\
\hline
0&1&0&1&0&1&1&0&1&0&0&1&0&0&0&0\\
1&0&1&0&0&0&1&1&1&1&0&0&0&0&0&0\\
0&1&0&1&1&0&0&1&0&1&1&0&0&0&0&0\\
1&0&1&0&1&1&0&0&0&0&1&1&0&0&0&0
\end{array}\right).$$

Матрица S может быть компактифицирована в матрицу

$$S(x) = \begin{pmatrix}
0 & x^2 + x^3 & x^2 + x^3 & x + x^3 \\
x + x^2 & 0 & x + x^3 & x^2 + x^3 \\
x + x^2 & x + x^3 & 0 & 1 + x \\
x + x^3 & x + x^2 & 1 + x^3 & 0
\end{pmatrix}.$$

При этом исходная матрица удовлетворяет уравнению $S^2 = 4I_{16} + 2J_{16}$, следовательно, матрица $S(x)$ удовлетворяет сравнению

$$S^2(x) \equiv 4I_4 + 2(1 + x + x^2 + x^3)J_4 \pmod{x^4 - 1}.$$

1. Поиск орграфов dsrg$(22, 9, 6, 3, 4)$

В соответствии с определением 2 надо найти бинарную квадратную матрицу порядка 22, удовлетворяющую условиям (1) и (2). Полный перебор по всем таким матрицам занял бы необозримо долгое время, поэтому будем осуществлять поиск только среди матриц, имеющих специальный вид. Идея описываемого подхода позаимствована из работы О. Гриценко [7].

Будем искать матрицу смежности A орграфа dsrg$(22, 9, 6, 3, 4)$ в виде (4). Первый столбец и первая строка выбраны таким образом, чтобы автоматически выполнялись первые два условия из определения 1, в них на месте многоточий в приведённой записи стоят нули. Матрицы K_{ij} $(1 \leqslant i,\, j \leqslant 7)$ являются циркулянтами третьего порядка:

$$A = \left(\begin{array}{c|ccc|ccc|ccc|c|ccc} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \dots & 0 & 0 & 0 \\ \hline 1 & & & & & & & & & & & & & \\ 1 & & K_{11} & & & K_{12} & & & K_{13} & & \dots & & K_{17} & \\ 1 & & & & & & & & & & & & & \\ \hline 1 & & & & & & & & & & & & & \\ 1 & & K_{21} & & & K_{22} & & & K_{23} & & \dots & & K_{27} & \\ 1 & & & & & & & & & & & & & \\ \hline 0 & & & & & & & & & & & & & \\ 0 & & K_{31} & & & K_{32} & & & K_{33} & & \dots & & K_{37} & \\ 0 & & & & & & & & & & & & & \\ \hline 1 & & & & & & & & & & & & & \\ 1 & & K_{41} & & & K_{42} & & & K_{43} & & \dots & & K_{47} & \\ 1 & & & & & & & & & & & & & \\ \hline \vdots & & \vdots & & & \vdots & & & \vdots & & \ddots & & \vdots & \\ \hline 0 & & & & & & & & & & & & & \\ 0 & & K_{71} & & & K_{72} & & & K_{73} & & \dots & & K_{77} & \\ 0 & & & & & & & & & & & & & \end{array}\right) \tag{4}$$

Для искомой матрицы A условия (1) и (2) примут следующий вид:

$$A^2 + A = 2I_{22} + 4J_{22}; \tag{5}$$

$$AJ_{22} = J_{22}A = 9J_{22}. \tag{6}$$

Запишем матрицу A в виде

$$A = \begin{pmatrix} 0 & B \\ D & C \end{pmatrix},$$

где C — подматрица матрицы A, полученная вычёркиванием первой строки и первого столбца; $B = (11\dots10\dots0)$ — первая строка матрицы A без элемента A_{11}; $D = (11\dots10001110\dots0)^{\mathrm{T}}$ — первый столбец матрицы A без элемента A_{11}. При таком представлении равенство (5) равносильно системе следующих четырёх уравнений:

$$\begin{cases} BD = 6, \\ BC + B = 4J_{1,21}, \\ CD + D = 4J_{21,1}, \\ DB + C^2 + C = 2I_{21} + 4J_{21}. \end{cases} \tag{7}$$

Первое равенство в системе (7) выполняется автоматически, это следует из вида матриц B и D.

Из второго равенства в (7) следует, что

$$\sum_{i=1}^{9} C_{ij} = \begin{cases} 3, & \text{если } j \leqslant 9, \\ 4, & \text{если } j \geqslant 10. \end{cases} \tag{8}$$

Поскольку искомая матрица удовлетворяет соотношению $AJ_{22} = 9J_{22}$, из (8) следует

$$\sum_{i=10}^{21} C_{ij} = 5 \quad \text{при} \quad 1 \leqslant j \leqslant 21. \tag{9}$$

Из третьего равенства в (7) получаем

$$\sum_{j\in\{1,\dots,12\}\setminus\{7,8,9\}} C_{ij} = \begin{cases} 3, & \text{если } i \in \{1,\dots,12\}\setminus\{7,8,9\}, \\ 4 & \text{иначе.} \end{cases} \tag{10}$$

Из равенств (6) и (10) следует, что

$$\sum_{j\notin\{1,\dots,12\}\setminus\{7,8,9\}} C_{ij} = 5 \quad \text{при} \quad 1 \leqslant i \leqslant 21. \tag{11}$$

Рассмотрим матрицу H_7, имеющую следующий вид:

$$H_7 = \begin{pmatrix} 3 & 3 & 3 & 4 & 4 & 4 & 4 \\ 3 & 3 & 3 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 & 4 & 4 & 4 \\ 3 & 3 & 3 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 & 4 & 4 & 4 \end{pmatrix}.$$

Введём обозначение: $H_{21} = 4J_{21} - DB$. Прямым вычислением можно убедиться, что матрица H_{21} получается заменой в матрице H_7 всех чисел 3 на блоки $3J_3$ и всех чисел 4 на блоки $4J_3$.

Последнее уравнение системы (7) можно переписать в виде

$$C^2 + C = 2I_{21} + H_{21}. \tag{12}$$

Применим описанный во введении метод. Так как матрица C состоит из 7×7 циркулянтных блоков третьего порядка, то она может быть компактифицирована в матрицу $C(x)$ с элементами из $\mathbb{Z}[x]/(x^3 - 1)$. Уравнение (12) преобразуется в следующее:

$$C^2(x) + C(x) \equiv 2I_7 + (1 + x + x^2)H_7 \pmod{x^3 - 1}. \tag{13}$$

Так как $x = 1$ является корнем многочлена $x^3 - 1$, то при подстановке $x = 1$ в (13) получаем следующее верное равенство:

$$C^2(1) + C(1) = 2I_7 + 3H_7. \tag{14}$$

Так как все коэффициенты многочлена $C(x)$ являются нулями или единицами, то элементы матрицы $C(1)$ являются целыми числами от 0 до 3. Диагональные элементы матрицы $C(1)$ не больше двух, потому что в матрице смежности искомого орграфа на диагонали стоят нули.

Авторами написана программа, которая осуществляет поиск всех матриц седьмого порядка с элементами от 0 до 3 (диагональными — от 0 до 2), удовлетворяющих условиям (14) и (8)–(11), переписанным для компактифицированной матрицы. Для реализации поиска использована библиотека программирования в ограничениях Artelys Kalis [8].

В результате поиска получено 10338 матриц, подходящих на роль матрицы $C(1)$. Заметим, что при наличии матрицы $C(1)$ область поиска матрицы смежности A искомого орграфа $\mathrm{dsrg}(22, 9, 6, 3, 4)$ значительно сужается: появляются ограничения на суммы элементов в строках циркулянтных блоков матрицы A. Исходя из этого, список

матриц $C(1)$ был сокращён до 144 матриц, для которых может быть найдена матрица A, удовлетворяющая условиям (5) и (6).

Для каждой из этих 144 матриц $C(1)$ программно были найдены все подходящие матрицы смежности A. Таких матриц получилось 384. В следующем пункте проанализированы полученные орграфы.

Запуск программы осуществлялся на компьютере с процессором Intel Core i5-7400 (3,00 ГГц), объём оперативной памяти равен 32 ГБ. Распределение этапов поиска по времени выглядит следующим образом:

1) поиск 10338 матриц, подходящих на роль матрицы $C(1)$, — 125 часов;
2) поиск матриц A для найденных матриц $C(1)$ — 40 часов.

Замечание 2. Стоит отметить, что при помощи разработанной программы были найдены не все возможные матрицы смежности орграфа $\mathrm{dsrg}(22, 9, 6, 3, 4)$, у которых основная часть состоит из циркулянтов 3×3, а только такие, у которых первая строка и первый столбец имеют заданный вид (см. (4)).

2. Анализ построенных примеров на изоморфность

Приведём в качестве примера одну из матриц смежности, найденных при помощи программы:

$$A = \begin{pmatrix}
0&1&1&1&1&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0\\
1&0&1&1&0&0&0&0&0&1&1&0&0&0&0&0&0&1&0&1&1&1\\
1&1&0&1&0&0&0&1&0&0&0&1&0&0&0&0&0&0&1&1&1&1\\
1&1&1&0&0&0&0&0&1&0&0&0&1&0&0&0&1&0&0&1&1&1\\
1&0&0&1&0&0&0&1&1&0&0&1&1&0&0&1&1&0&0&1&0&0\\
1&1&0&0&0&0&0&0&1&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&0&1&0&0&0&0&1&0&1&1&1&0&0&1&0&0&0&1&0&0&1\\
0&0&0&0&1&1&1&0&0&0&0&1&0&1&1&1&1&1&0&0&0&0\\
0&0&0&0&1&1&1&0&0&0&0&0&1&1&1&1&0&1&1&0&0&0\\
0&0&0&0&1&1&1&0&0&0&1&0&0&1&1&1&1&0&1&0&0&0\\
1&1&0&0&0&0&0&0&0&1&0&1&1&0&0&0&0&1&0&1&1&1\\
1&0&1&0&0&0&0&1&0&0&1&0&1&0&0&0&0&0&1&1&1&1\\
1&0&0&1&0&0&0&0&1&0&1&1&0&0&0&0&1&0&0&1&1&1\\
0&0&1&1&0&1&0&1&0&0&0&1&0&0&0&0&1&0&1&1&0&1\\
0&1&0&1&0&0&1&0&1&0&0&0&1&0&0&0&1&1&0&1&1&0\\
0&1&1&0&1&0&0&0&0&1&1&0&0&0&0&0&0&1&1&0&1&1\\
0&0&0&1&1&1&1&1&0&1&0&0&0&1&1&1&0&0&0&0&0&0\\
0&1&0&0&1&1&1&1&1&0&0&0&0&1&1&1&0&0&0&0&0&0\\
0&0&1&0&1&1&1&0&1&1&0&0&0&1&1&1&0&0&0&0&0&0\\
0&1&0&0&1&0&0&0&0&1&1&1&0&1&1&0&0&1&1&0&0&0\\
0&0&1&0&0&1&0&1&0&0&0&1&1&0&1&1&1&0&1&0&0&0\\
0&0&0&1&0&0&1&0&1&0&1&0&1&1&0&1&1&1&0&0&0&0
\end{pmatrix}. \tag{15}$$

Разработанная программа позволяет построить целое семейство орграфов требуемого вида. При этом среди построенных орграфов будет довольно много пар изоморфных, и задача выяснения количества построенных орграфов с точностью до изоморфизма не вполне тривиальна. Перебирать все возможные биекции множеств вершин в поисках изоморфизма непрактично и не поучительно, некоторые особенности структуры построенных орграфов позволяют сделать это более эффективно.

Определение 3.

1) По построению группа автоморфизмов каждого найденного орграфа содержит (при надлежащей нумерации вершин) перестановку

$$(1)(2,3,4)(5,6,7)(8,9,10)(11,12,13)(14,15,16)(17,18,19)(20,21,22).$$

Фактически, во всех построенных примерах группа автоморфизмов орграфа совпадает с группой $\mathbb{Z}_3$, порождённой этой перестановкой.
2) Соответственно множество вершин орграфа разбивается на восемь подмножеств: одно — одноэлементное (эту вершину будем называть *особой*) и семь трёхэлементных (будем называть их *этажами*).
3) Множество дуг, ведущих с одного этажа на другой, инвариантно относительно группы перестановок.
4) Одним из вариантов множеств дуг, допускаемых предыдущим пунктом, является *полный переход*, когда из каждой вершины одного этажа идёт ровно по одной дуге в каждую вершину другого этажа.
5) Полным переходом назовём также ситуацию, когда вместо одного из этажей фигурирует одноэлементное множество из особой вершины (в этом случае дуг не девять, а три).
6) Дополнительно отметим, что «внутри» этажа также могут присутствовать три или шесть дуг орграфа (цикл или два встречных цикла), но множество дуг внутри этажа может быть и пустым.
7) *Скелетом* орграфа назовём орграф с восемью вершинами, соответствующими множествам вершин, семь из которых могут быть раскрашены в четыре цвета, соответствующие наличию или отсутствию на этаже внутренних дуг. Дугами этого орграфа являются (в случае наличия таковых) полные переходы.
8) Орграф с теми же вершинами и (неформально говоря) всеми остальными дугами рассматриваемого орграфа (существует несколько вариантов *неполных* переходов с одного этажа на другой, которые будем обозначать как разные цвета дуг) назовём *оснасткой*.

Найденные при помощи программы орграфы разбиваются на 16 групп, каждая группа состоит из 24 изоморфных орграфов. На рис. 1–8 приведены скелеты и оснастки, соответствующие восьми группам (на оснастках особая вершина не изображена, поскольку она является изолированной). Остальные восемь орграфов получаются путём смены направлений всех дуг.

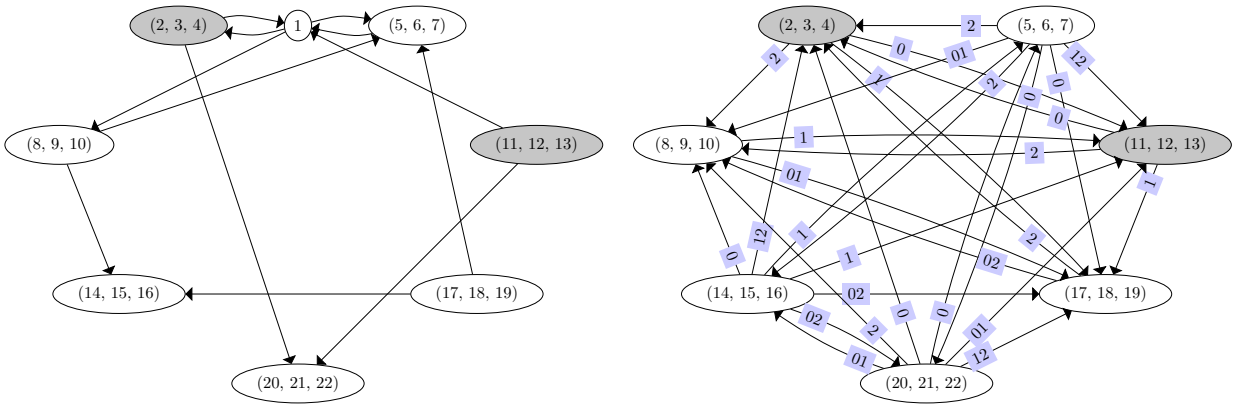

Рис. 1. Скелет и оснастка первого орграфа

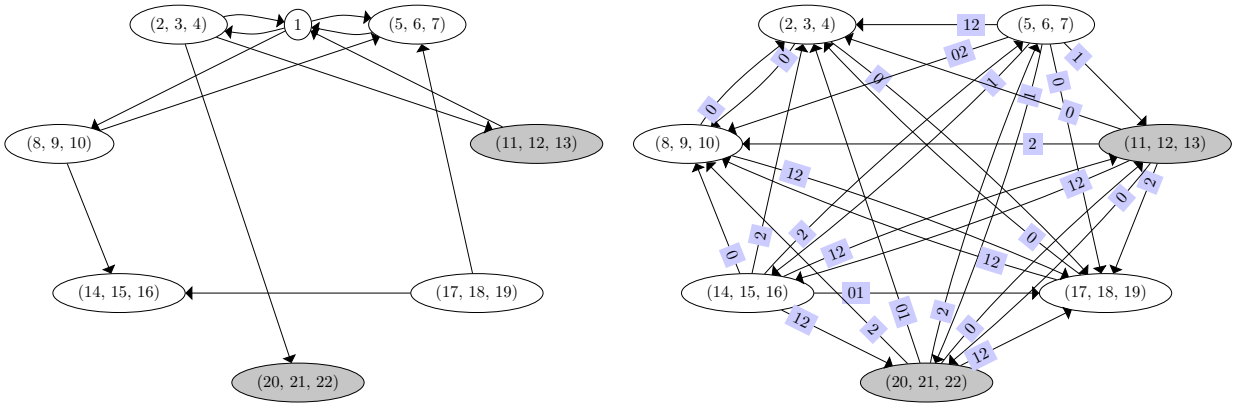

Рис. 2. Скелет и оснастка второго орграфа

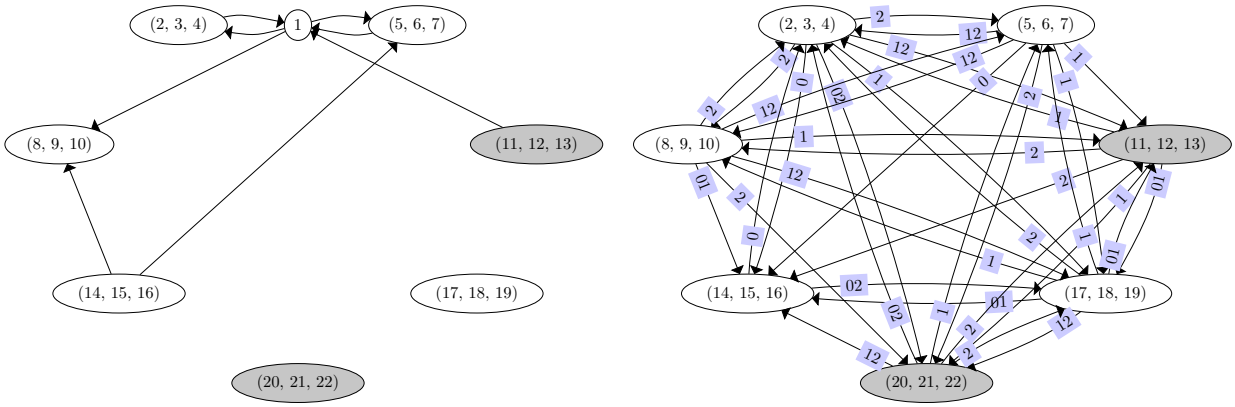

Рис. 3. Скелет и оснастка третьего орграфа

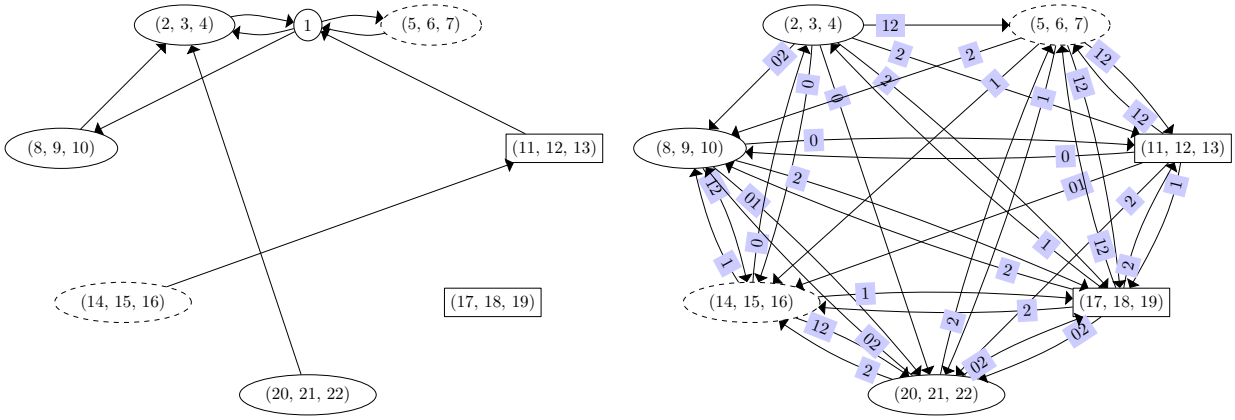

Рис. 4. Скелет и оснастка четвёртого орграфа

На рисунках используются следующие обозначения:

1) Если внутри этажа ни одна вершина не соединена с другой, то этаж изображается в виде незакрашенного эллипса со сплошной границей (например, $(5, 6, 7)$ на рис. 1).
2) Если внутри этажа каждые две вершины соединены парой дуг, то этаж изображается в виде закрашенного эллипса (например, $(2, 3, 4)$ на рис. 1).

3) Если на этаже находится один цикл $(3k+2) \to (3k+3) \to (3k+4) \to (3k+2)$, $k \in \{0, 2, \ldots, 6\}$, то этаж изображается в виде эллипса с пунктирной границей (например, $(5, 6, 7)$ на рис. 4).
4) Если на этаже находится один цикл $(3k+4) \to (3k+3) \to (3k+2) \to (3k+4)$, $k \in \{0, 2, \ldots, 6\}$, то этаж изображается в виде прямоугольника (например, $(11, 12, 13)$ на рис. 4).
5) Если подпись на дуге $(a_0, a_1, a_2) \to (b_0, b_1, b_2)$ в оснастке содержит цифру $s \in \in \{0, 1, 2\}$, то в исходном орграфе от каждой вершины a_i идёт дуга к вершине $b_{(i+s) \bmod 3}$ при $i \in \{0, 1, 2\}$. Например, в оснастке на рис. 1 от этажа $(14, 15, 16)$ идёт дуга к этажу $(2, 3, 4)$ с подписью 12, что означает наличие в исходном орграфе дуг $14 \to 3$, $15 \to 4$, $16 \to 2$ и $14 \to 4$, $15 \to 2$, $16 \to 3$.

Скелет и оснастка графа, матрица смежности которого приведена в (15), изображены на рис. 1.

Отметим, что скелет орграфа не определяет однозначно его оснастку: пятый и шестой орграфы (рис. 5 и 6) имеют одинаковые скелеты, но разные оснастки; одинаковые скелеты также у седьмого и восьмого орграфов (рис. 7 и 8).

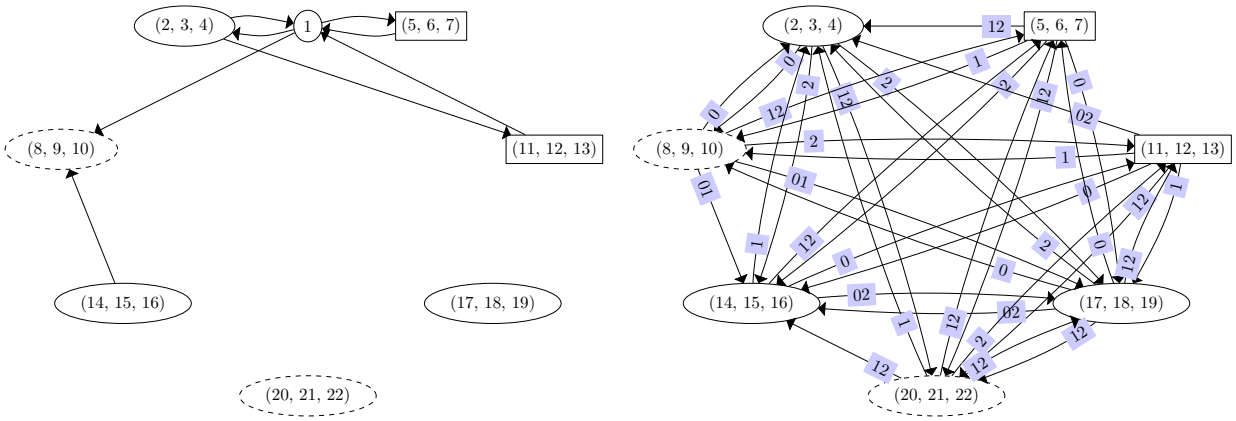

Рис. 5. Скелет и оснастка пятого орграфа

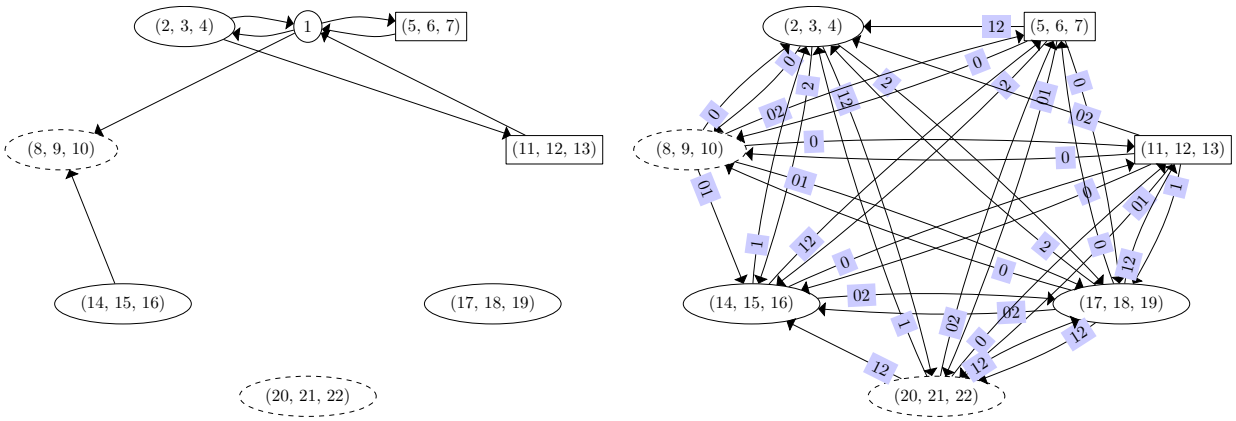

Рис. 6. Скелет и оснастка шестого орграфа

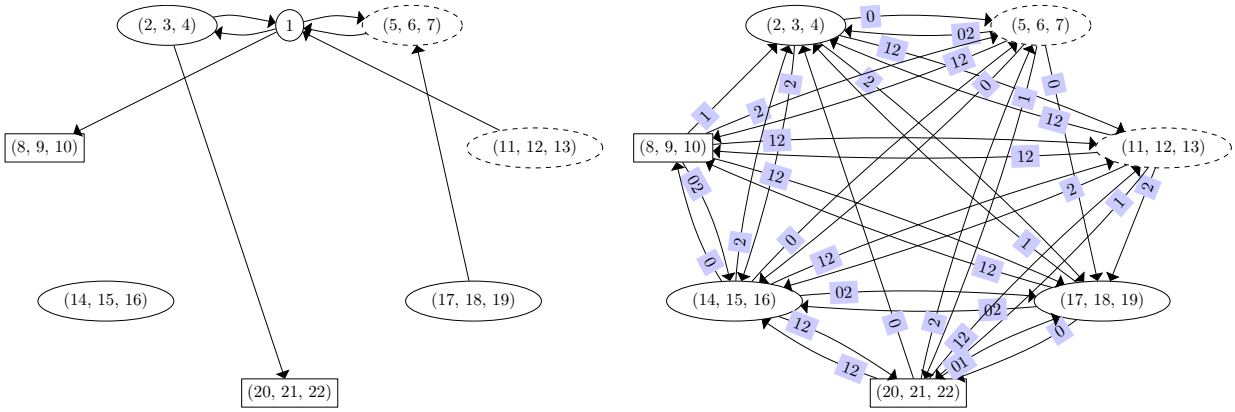

Рис. 7. Скелет и оснастка седьмого орграфа

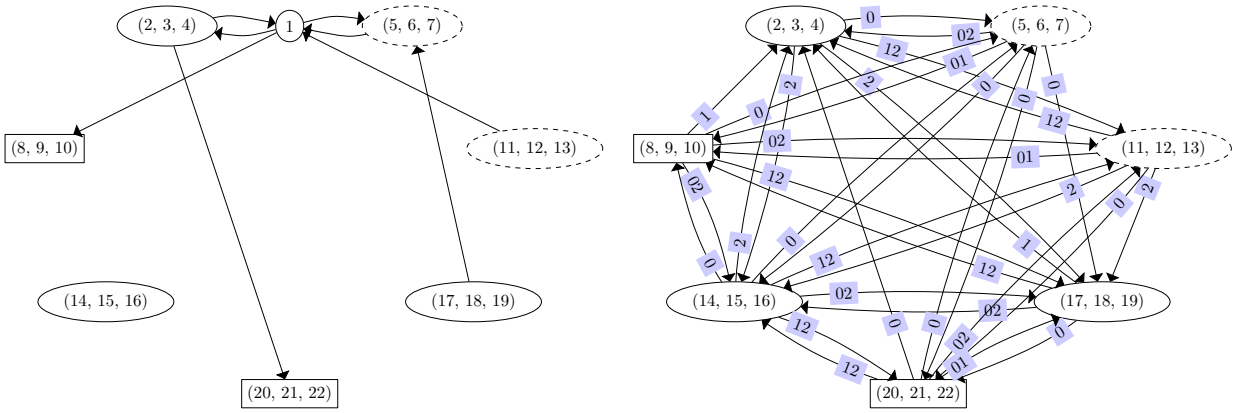

Рис. 8. Скелет и оснастка восьмого орграфа

Замечание 3. Исследование найденных графов на изоморфность было проведено и программно при помощи функции is_isomorphic из библиотеки NetworkX [9]. Но авторам кажется, что теоретический анализ структуры графов более информативен, чем простая констатация их количества при помощи программы. Главная цель теоретического анализа — построение системы инвариантов, аналогичные которой могут встретиться и в других задачах.

Замечание 4. В [1] доказано: если орграф G является сильно регулярным с набором параметров (v, k, t, λ, μ), то орграф G', являющийся дополнением G, также сильно регулярный с набором параметров $(v, v-k-1, v-2k+t-1, v-2k+\mu-2, v-2k+\lambda)$ (матрица смежности графа G' равна $A' = J_v - I_v - A$, где A — матрица смежности графа G). Из этого следует, что в работе автоматически построено семейство сильно регулярных орграфов с набором параметров $(22, 12, 9, 6, 7)$.

Заключение

В работе построено семейство сильно регулярных орграфов с набором параметров (22, 9, 6, 3, 4), вопрос о существовании которых ранее считался открытым. Построение матрицы основано на идее О. Гриценко [7]. Кроме того, предложена некоторая естественная структура, позволяющая решить вопрос об их изоморфности простым, интуитивно наглядным способом (не требующим переборной программы). Структурная декомпозиция оказалась нетривиальной в том смысле, что первый структурный элемент (скелет) не всегда определяет второй (оснастку) однозначно, и полезной в том

смысле, что (теоретически) позволяет вручную проверить соответствие орграфа требуемым ограничениям.

Очевидно, что описанная конструкция (разбиение части матрицы без первой строки и первого столбца на циркулянтные блоки) применима только для наборов параметров (v, k, t, λ, μ), для которых числа $v-1$, k, t делятся на размер блока. Эта ситуация нередко встречается среди нерешённых на данный момент задач, поэтому метод, использованный авторами, потенциально применим и к ним, однако о других продвижениях такого типа авторам неизвестно.

ЛИТЕРАТУРА

1. *Duval A. M.* A directed graph version of strongly regular graphs // J. Combinat. Theory. Ser. A. 1988. V. 47. No. 1. P. 71–100.
2. *Brouwer A. E. and Maldeghem H. V.* Strongly Regular Graphs. Cambridge: Cambridge University Press, 2022. 425 p.
3. `https://homepages.cwi.nl/~aeb/math/dsrg/dsrg.html` — Parameters of directed strongly regular graphs. 2024.
4. *Brouwer A. E., Crnkovic D., and Svob A.* A Construction of Directed Strongly Regular Graphs with Parameters (63,11,8,1,2). `http://arxiv.org/abs/2404.03100v2`. 2024.
5. *Kra I. and Simanca S. R.* On circulant matrices // Notices Amer. Math. Soc. 2012. V. 59. No. 3. P. 368–377.
6. *Shrikhande S. S.* The uniqueness of the L_2 association scheme // Ann. Math. Statistics. 1959. V. 30. No. 3. P. 781–798.
7. *Gritsenko O.* On Strongly Regular Graph with Parameters (65; 32; 15; 16). `https://arxiv.org/abs/2102.05432`. 2021.
8. `https://www.artelys.com/solvers/kalis/` — Artelys Kalis. 2024.
9. *Hagberg A. A., Schult D. A., and Swart P. J.* Exploring network structure, dynamics, and function using NetworkX // Proc. SciPy 2008. Pasadena, California, August 19–24, 2008. P. 11–15.

REFERENCES

1. *Duval A. M.* A directed graph version of strongly regular graphs. J. Combinat. Theory, Ser. A, 1988, vol. 47, no. 1, pp. 71–100.
2. *Brouwer A. E. and Maldeghem H. V.* Strongly Regular Graphs. Cambridge, Cambridge University Press, 2022. 425 p.
3. `https://homepages.cwi.nl/~aeb/math/dsrg/dsrg.html` — Parameters of directed strongly regular graphs, 2024.
4. *Brouwer A. E., Crnkovic D., and Svob A.* A Construction of Directed Strongly Regular Graphs with Parameters (63,11,8,1,2). `http://arxiv.org/abs/2404.03100v2`, 2024.
5. *Kra I. and Simanca S. R.* On circulant matrices. Notices Amer. Math. Soc., 2012, vol. 59, no. 3, pp. 368–377.
6. *Shrikhande S. S.* The uniqueness of the L_2 association scheme. Ann. Math. Statistics, 1959, vol. 30, no. 3, pp. 781–798.
7. *Gritsenko O.* On Strongly Regular Graph with Parameters (65; 32; 15; 16). `https://arxiv.org/abs/2102.05432`, 2021.
8. `https://www.artelys.com/solvers/kalis/` — Artelys Kalis, 2024.
9. *Hagberg A. A., Schult D. A., and Swart P. J.* Exploring network structure, dynamics, and function using NetworkX. Proc. SciPy 2008, Pasadena, California, August 19–24, 2008, pp. 11–15.

УДК 519.87:519.6:519.178 DOI 10.17223/20710410/66/9

ОТКРЫТИЕ БЕСКОНЕЧНЫХ СЕМЕЙСТВ ОПТИМАЛЬНЫХ ДВУХКОНТУРНЫХ КОЛЬЦЕВЫХ СЕТЕЙ С ЗАДАННЫМ ШАБЛОНОМ ОБРАЗУЮЩИХ[1]

Э. А. Монахова, О. Г. Монахов

Институт вычислительной математики и математической геофизики СО РАН, г. Новосибирск, Россия

E-mail: emilia@rav.sscc.ru

Оптимальные кольцевые циркулянтные сети степени четыре рассматриваются как модели надёжных сетей связи с минимальными задержками для сетей на кристалле и мультипроцессорных кластерных систем. Проведён поиск аналитически задаваемых бесконечных семейств оптимальных графов на основе анализа базы данных оптимальных описаний двухконтурных кольцевых циркулянтных сетей. Путём интеграции визуализации данных и аналитических описаний оптимальных графов построены и теоретически обоснованы новые бесконечные семейства оптимальных сетей с линейной образующей вида $s = 4d+\alpha$, где d — диаметр графа. Предложенный подход получения семейств оптимальных сетей является новым и представляет интерес для дальнейшего изучения свойств оптимальных двухконтурных кольцевых сетей.

Ключевые слова: *датасет оптимальных сетей, неориентированные двухконтурные кольцевые сети, циркулянтные сети, минимальный диаметр.*

DISCOVERY OF INFINITE FAMILIES OF OPTIMAL DOUBLE-LOOP NETWORKS WITH A GIVEN TEMPLATE OF GENERATORS

E. A. Monakhova, O. G. Monakhov

Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia

Optimal ring circulant networks of degree four are considered as models of reliable communication networks with minimal delays for networks on a chip and multiprocessor cluster systems. Based on the analysis of a data set of optimal descriptions of double-loop networks, a search has been carried out for analytically determined infinite families of optimal graphs. By integrating data visualization and analytical descriptions of optimal graphs, new infinite families of optimal networks with a linear generator of the form $s = 4d + \alpha$, where d is the diameter of the graph, have been constructed and theoretically justified. The proposed approach to obtaining families of optimal networks is new and is of interest for further studies of the properties of optimal double-loop networks.

Keywords: *dataset of optimal networks, undirected double-loop networks, circulant networks, minimum diameter.*

[1]Работа выполнена при финансовой поддержке бюджетного проекта ИВМиМГ СО РАН (код проекта FWNM-2022-0005).

Введение

Двухконтурные кольцевые сети (циркулянтные сети степени четыре) находят широкое применение при проектировании телекоммуникационных сетей, построении суперкомпьютеров, в прикладных задачах криптографии, а также исследуются как сети связи в сетях на кристалле в качестве замены традиционно используемых в них двумерных решёток и торов, имеющих существенно большие задержки при одинаковом числе узлов [1–8]. Удобство таких сетей обусловлено свойствами симметричности, высокой связности и масштабируемости, что позволяет применять их в центрах коллективного пользования, беспроводных сенсорных и нейронных сетях [9–12].

Двухконтурная кольцевая циркулянтная сеть (undirected double-loop network) представляет собой неориентированный граф $C(N;1,s)$, $1 < s < N/2$, с множеством вершин $V = \{0,1,\ldots,N-1\}$ и рёбер $E = \{(i,j) : i-j \equiv \pm 1 \pmod N,\ i-j \equiv \pm s \pmod N\}$, где $1, s$ — образующие; N — порядок графа. Пример двухконтурной кольцевой сети с числом узлов $N = 18$ представлен на рис. 1. Задержки при передаче информации в сети и организации коллективных обменов и настройки в системе оцениваются диаметром графа связей (длиной максимального кратчайшего пути между любыми двумя узлами) [3, 4, 13]. Наибольший интерес представляют такие сети, диаметр которых совпадает с теоретической нижней границей диаметра либо отличен от неё на единицу. Данные сети носят названия оптимальных и субоптимальных соответственно. Не меньший интерес для исследователей представляет определение аналитических формул для описания оптимальных сетей (семейств оптимальных сетей). Точная нижняя граница диаметра циркулянтов степени четыре получена в [14, 15]: $D(N) = \lceil(-1+\sqrt{2N-1})/2\rceil$. Известна верхняя граница максимально возможного числа вершин в циркулянтных графах степени четыре с диаметром $d \geqslant 1$: $N_d = 2d^2+2d+1$.

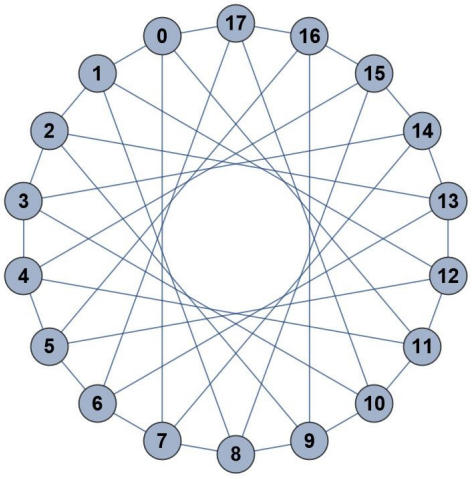

Рис. 1. Двухконтурная кольцевая сеть $C(18;1,7)$

Далее *оптимальным* называется граф $C(N;1,s)$ диаметра $d(C(N;1,s)) = D(N)$, *субоптимальным* — граф диаметра $d(C(N;1,s)) = D(N)+1$. В [16] выдвинута следующая гипотеза: по крайней мере, субоптимальные графы вида $C(N;1,s)$ существуют для любых N. Справедливость гипотезы проверена для всех $N \leqslant 8\cdot 10^6$, а также показано, что при $N \leqslant 8\cdot 10^6$ число значений N, для которых нет оптимальных графов, не превышает 6 %.

Оптимальное семейство циркулянтных сетей степени четыре и любого порядка $N > 4$ найдено в [14] и переоткрыто в [15, 17]: $\{C(N;d,d+1), d \geqslant 1\}$, где образую-

щая d — ближайшее целое к $(-1+\sqrt{2N-1})/2$. В [14] доказано, что все графы семейства имеют одновременно минимальный диаметр и среднее расстояние между вершинами. Для данного семейства сетей известны аналитические алгоритмы парной маршрутизации с константной оценкой сложности [3, 17, 18]. Отметим, что для всех $N_{d-1} < N \leqslant 2d^2+1$ оптимальным является также описание $(N; d-1, d)$ [3].

Из описаний циркулянтных графов вида $C(N; d, d+1)$ или $C(N; d-1, d)$ можно получить изоморфные описания путём умножения их образующих на элементы $t \leqslant \lfloor N/2 \rfloor$ приведённой системы вычетов по модулю N. Но такой метод не может быть использован для получения оптимальных описаний графов $C(N; 1, s)$ при любых N, поскольку для некоторых N они либо не существуют, либо существуют, но не изоморфны описаниям указанных видов.

1. Проблема поиска семейств оптимальных двухконтурных кольцевых сетей

В работе [19] дан подробный анализ существующих в литературе подходов к построению бесконечных семейств оптимальных (или субоптимальных) графов $C(N; 1, s)$, а также вопросов, связанных с применимостью их как моделей сетей связи многопроцессорных систем. Первый результат аналитического решения проблемы получен в [20]:

Теорема 1 [20]. Циркулянты $C(2d^2+2d+1; 1, 2d+1)$ оптимальны при любых целых $d \geqslant 1$.

Это семейство переоткрыто в [21, 22] и активно исследовалось для различных применений (см. обзор в [3]). В [23] данное семейство рассмотрено в качестве топологии сетей на кристалле. В большинстве работ, посвящённых поиску бесконечных семейств оптимальных графов $C(N; 1, s)$, используются теоретические верхние оценки диаметра или образующих [16, 24–29]. В [30] реализованы генетические алгоритмы поиска семейств и построены семейства оптимальных графов с линейными образующими $s = 4d \pm \alpha$ и $s = 6d \pm \alpha$. В [16, 24, 27, 29, 31] найдены или исследуются семейства графов с линейными образующими $s = 2d \pm \alpha$, где d — диаметр; в [32–34] — семейства графов с квадратичными образующими от диаметра. Для ряда семейств графов с образующей $s = 4d \pm \alpha$ найдены эффективные алгоритмы маршрутизации [27, 33], а также совершенные доминирующие множества вершин (perfect dominating set) [31].

В [35] впервые построен датасет оптимальных графов $C(N; 1, s)$ до 50 тысяч вершин (`https://github.com/mila0411/Double-loop-networks/tree/main/Dataset`). Точки (N, s, d) датасета соответствуют параметрам описаний оптимальных графов: N — порядок графа; s — образующая; d — диаметр. Для каждого значения N показаны все образующие $s \leqslant N/2$, которые определяют граф минимально возможного диаметра при данном N.

Первоначальный анализ датасета с целью открытия аналитически описываемых семейств оптимальных графов проведён в [35] с помощью подхода, основанного на шаблонах с недоопределёнными коэффициентами и использующего для поиска перспективных шаблонов алгоритмы метаэвристического поиска. В [19] рассмотрен другой метод автоматизации поиска семейств оптимальных графов в датасете, основанный на последовательном делении параметров оптимальных графов для построения коэффициентов полиномов их порядков и образующих, и разработаны алгоритмы поиска семейств оптимальных графов, отличающихся видом образующих с линейной или квадратичной функцией от диаметра. Основной идеей алгоритмов является определение параметров кривой, которая описывает семейство оптимальных графов $C(N; 1, s)$, с последующей проверкой принадлежности других точек датасета этой кривой. Прин-

ципы, положенные в основу этих алгоритмов, были успешно применены для автоматизированного поиска семейств в датасете другого класса графов — оптимальных хордальных кольцевых сетей [36].

В настоящей работе предложен новый подход к решению проблемы, который использует объединение аналитических результатов, полученных с помощью предложенного ранее алгоритма поиска, с графической визуализацией данных из датасета. Далее этот подход мы реализуем на множестве оптимальных графов, имеющих линейные образующие определённого вида.

2. Поиск семейств оптимальных графов с заданным шаблоном образующих

На рис. 2 в координатах (N, s) представлен фрагмент графической визуализации данных из датасета с оптимальными образующими $s \leqslant 500$. На полученном графике наблюдается интересная картина. Оптимальные образующие линейного вида от диаметра разбились на отдельные ярко выраженные полосы (сегменты), а именно: на нижней полосе точек отображены графы с линейными образующими вида $s = 2d \pm \alpha$, на следующем сегменте — графы с образующими $s = 4d \pm \alpha$ и далее выделяются отдельно сегменты для графов с $s = 6d \pm \alpha$ и $s = 8d \pm \alpha$. Таким образом, обнаружено свойство устойчивого существования образующих вида $s = kd \pm \alpha$ оптимальных графов с чётными коэффициентами $k \geqslant 2$. С другой стороны, отсутствуют устойчивые семейства оптимальных описаний с нечётными значениями k. Дополнительной проверкой существования в датасете оптимальных образующих вида $s = 3d \pm \alpha$, $s = 5d \pm \alpha$ и $s = 7d \pm \alpha$ мы получили, что при малых диаметрах существуют такие образующие, но при росте N они быстро заканчиваются и не входят в устойчивый режим. Поэтому для поиска семейств оптимальных графов с линейными образующими рассматриваются образующие $s = kd + \alpha$, где k принимает только чётные значения, а α — любые целые значения с $|\alpha| < d$.

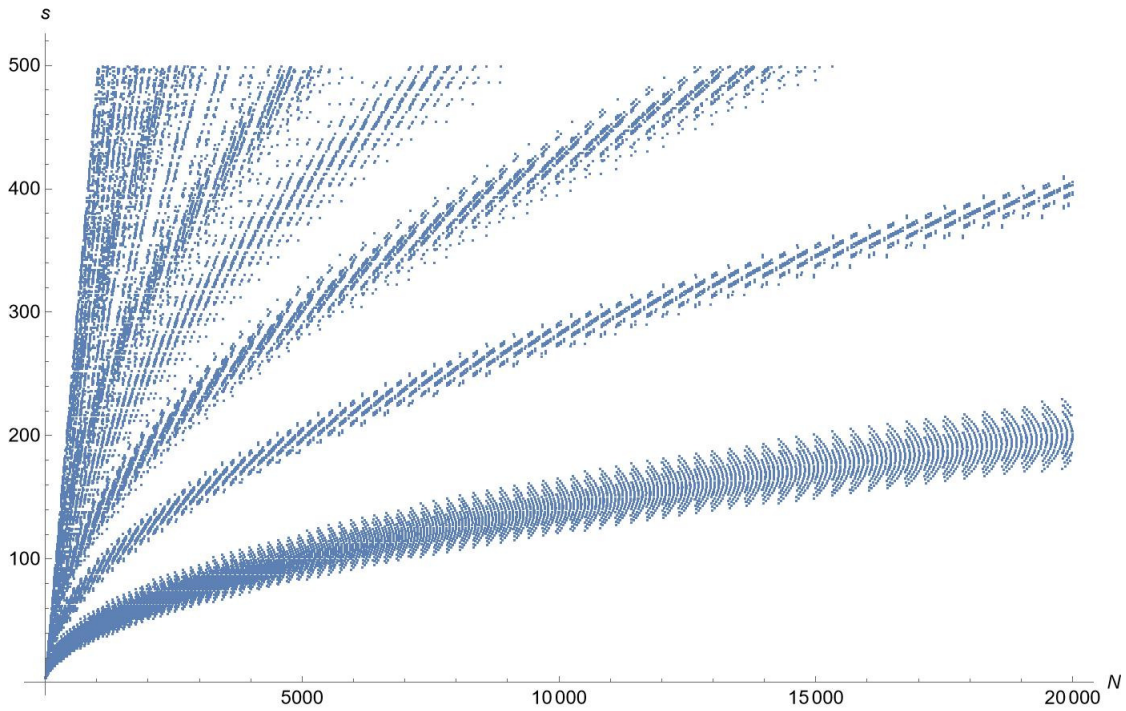

Рис. 2. Точки датасета для графов $C(N; 1, s)$, $s \leqslant 500$

Множество оптимальных двухконтурных кольцевых циркулянтов с образующими $s = 2d + \alpha$ аналитически описано Д. Тзвиели [16, теоремы 4.2, 4.6]. В настоящей работе будем исследовать в датасете существование множества оптимальных графов с образующими $s = 4d + \alpha$.

В [19] приведён алгоритм поиска семейств двухконтурных кольцевых графов с линейными образующими, который был применён к поиску семейств оптимальных графов с линейными образующими $s = 4d+\alpha$, $|\alpha| < d$. Для решения задачи отдельно рассмотрены поиск графов с чётными и нечётными диаметрами, при этом использовано значение периода повторяемости членов семейства $p = 2$. После этого для дальнейшего исследования отобраны аналитические описания тех семейств, которые соответствуют устойчивому режиму повторяемости на большом диапазоне диаметров. Ниже представлены списки полученных аналитических описаний 72 семейств оптимальных графов, отдельно для чётных и нечётных диаметров. Элементы списков упорядочены в порядке возрастания параметра α.

Список семейств оптимальных графов $C(N; 1, s)$ с чётным диаметром d:

$$(s, N) \in \big\{(4d-16, 2d^2-3d/2-25), (4d-14, 2d^2-3d/2-20), (4d-14, 2d^2-d/2-22),$$
$$(4d-12, 2d^2-d/2-17), (4d-8, 2d^2-3d/2-5), (4d-8, 2d^2-3d/2-4),$$
$$(4d-7, 2d^2-d-4), (4d-6, 2d^2-3d/2-3), (4d-6, 2d^2-3d/2-2),$$
$$(4d-6, 2d^2-d/2-3), (4d-5, 2d^2-3), (4d-5, 2d^2-d-2), (4d-4, 2d^2-d/2-2),$$
$$(4d-4, 2d^2+d/2-2), (4d-3, 2d^2-1), (4d-2, 2d^2+d/2-1), (4d-2, 2d^2+3d/2-1),$$
$$(4d, 2d^2-3d/2), (4d, 2d^2-3d/2+1), (4d, 2d^2+3d/2), (4d+1, 2d^2-d-1),$$
$$(4d+1, 2d^2-d), (4d+2, 2d^2-3d/2-2), (4d+2, 2d^2-3d/2-1), (4d+2, 2d^2-d/2-1),$$
$$(4d+2, 2d^2-d/2), (4d+3, 2d^2-1), (4d+3, 2d^2-d-2), (4d+3, 2d^2-d-1),$$
$$(4d+4, 2d^2-d/2-3), (4d+4, 2d^2-d/2-2), (4d+4, 2d^2+d/2-1), (4d+5, 2d^2-3),$$
$$(4d+6, 2d^2+d/2-4), (4d+8, 2d^2-3d/2-10), (4d+10, 2d^2-3d/2-17)\big\}.$$

Список семейств оптимальных графов $C(N; 1, s)$ с нечётным диаметром d:

$$(s, N) \in \big\{(4d-20, 2d^2-3d/2-83/2), (4d-18, 2d^2-3d/2-69/2),$$
$$(4d-12, 2d^2-3d/2-25/2), (4d-11, 2d^2-d-12), (4d-10, 2d^2-3d/2-19/2),$$
$$(4d-10, 2d^2-d/2-21/2), (4d-9, 2d^2-d-8), (4d-8, 2d^2-d/2-15/2),$$
$$(4d-8, 2d^2+d/2-17/2), (4d-6, 2d^2+d/2-11/2), (4d-4, 2d^2-3d/2-1/2),$$
$$(4d-4, 2d^2-3d/2+1/2), (4d-3, 2d^2-d), (4d-2, 2d^2-3d/2-1/2),$$
$$(4d-2, 2d^2-3d/2+1/2), (4d-2, 2d^2-d/2-1/2), (4d-2, 2d^2-d/2+1/2),$$
$$(4d-1, 2d^2), (4d-1, 2d^2-d), (4d, 2d^2-d/2-1/2), (4d, 2d^2-d/2+1/2),$$
$$(4d, 2d^2+d/2+1/2), (4d+1, 2d^2), (4d+1, 2d^2+d), (4d+2, 2d^2+d/2-1/2),$$
$$(4d+2, 2d^2+3d/2+1/2), (4d+3, 2d^2+d), (4d+4, 2d^2-3d/2-7/2),$$
$$(4d+4, 2d^2-3d/2-5/2), (4d+4, 2d^2+3d/2-1/2), (4d+5, 2d^2-d-4),$$
$$(4d+6, 2d^2-3d/2-15/2), (4d+6, 2d^2-3d/2-13/2), (4d+6, 2d^2-d/2-9/2),$$
$$(4d+7, 2d^2-d-8), (4d+8, 2d^2-d/2-19/2)\big\}.$$

По сравнению с [30], где для поиска семейств с образующей $s = 4d + \alpha$ использовались генетические алгоритмы, дополнительно найдено 16 новых аналитически описываемых семейств указанного вида.

Для проверки существования найденных семейств при диаметрах графов, выходящих за границы датасета при $d > 158$, на основе программы [37] разработан новый алгоритм проверки оптимальных описаний из списков семейств для $N > 50000$. Алгоритм реализован на языке Си и использует специальную программу определения диаметров циркулянтных графов по их описанию и сравнения полученных диаметров с нижней границей $D(N)$. Семейства проверены выборочно для значений диаметров $d = 200, 201, 300, 301, 400, 401, 500, 501$ (при больших d временны́е затраты на проверку становятся значительными), их оптимальность подтвердилась для этих диаметров.

3. Общие формулы описания семейств оптимальных графов с линейной образующей $s = 4d + \alpha$

На рис. 3 представлена графическая визуализация точек датасета (N, s), соответствующих описаниям оптимальных графов с линейными образующими вида $s = 4d+\alpha$ при диаметрах $142 \leqslant d \leqslant 158$ и $N \leqslant 50000$, где $d = 158$ — верхняя граница диаметра графов из датасета. На графике видно, что с ростом N (и d) стабильно повторяются одинаковые графические конфигурации оптимальных описаний как для чётных, так и для нечётных диаметров. На рис. 4 показаны отдельно точки датасета для чётных и нечётных диаметров.

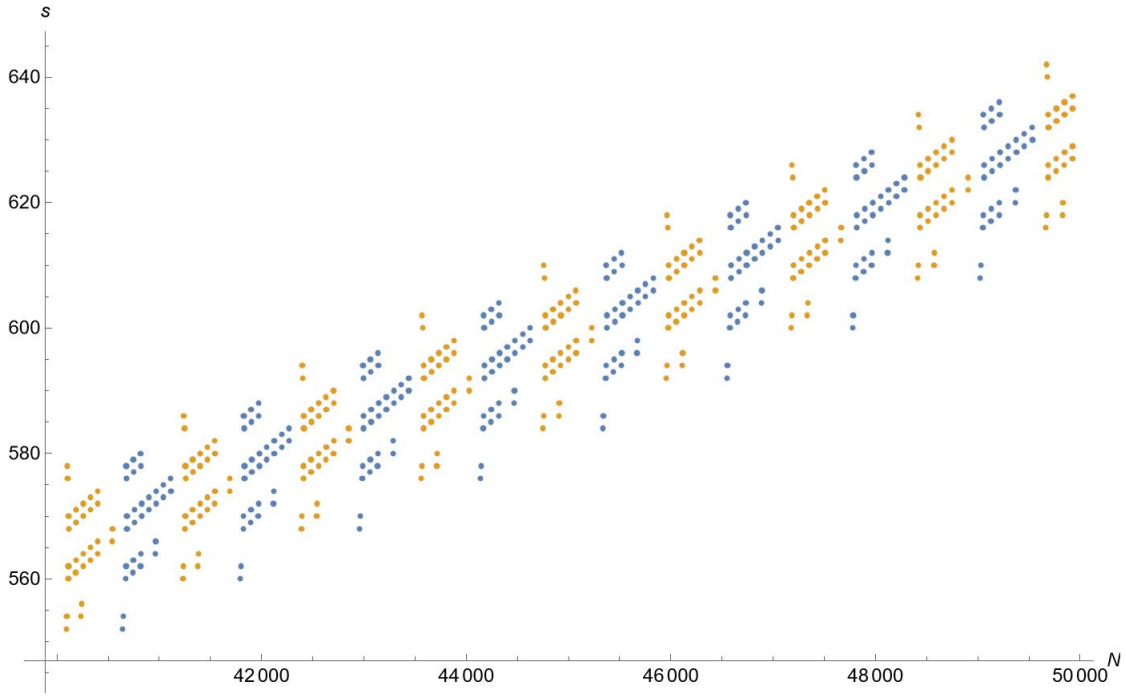

Рис. 3. Визуализация точек датасета в устойчивом режиме

После этого была решена задача отображения описаний оптимальных семейств из полученных списков в точки на графике. Как видно из списков, некоторым точкам с одинаковой образующей соответствуют по 2, 3 или 4 значения порядков графов. Далее мы выделили в качестве базовых те семейства, которые можно описать общими шаблонами (формулами с параметром), учитывая закономерность их появления вдоль линий (обозначены цифрами от 1 до 8) на рис. 4. В табл. 1 и 2 приведены найденные шаблоны для базовых семейств с чётным и нечётным диаметрами соответственно. Здесь n — номер линии, которой соответствуют формулы для s и N; i — параметр, определяющий на n-й линии местоположение базового семейства.

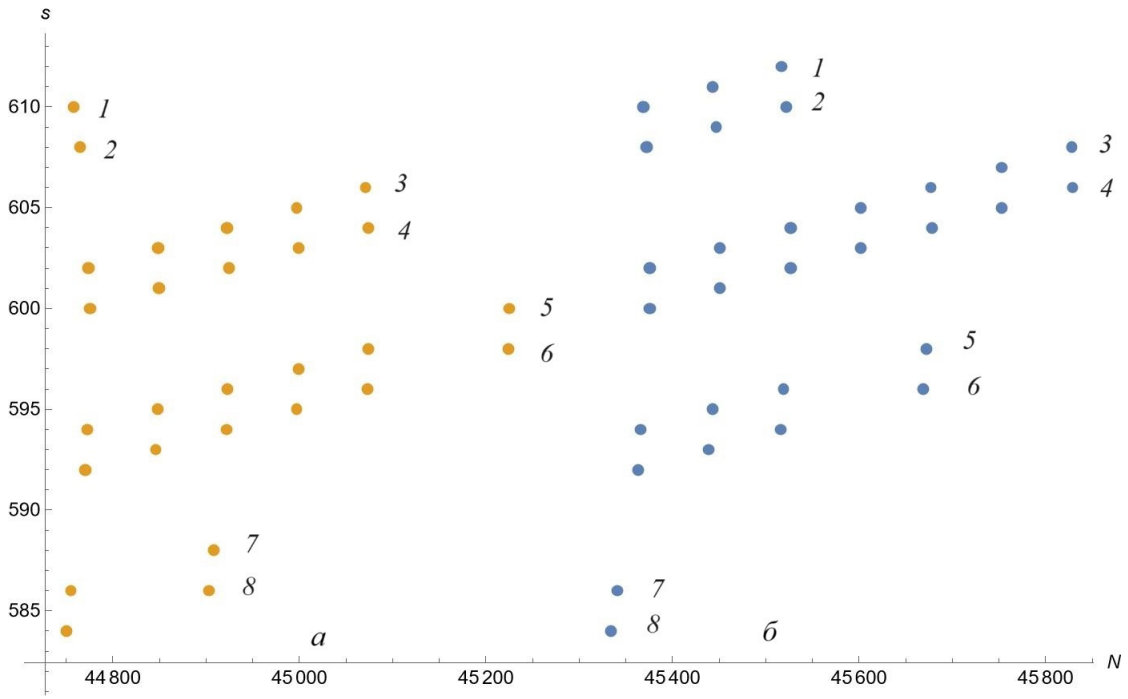

Рис. 4. Визуализация точек датасета для $d = 150$ и 151: *а* — чётные, *б* — нечётные диаметры

Таблица 1

Шаблоны для базовых семейств оптимальных графов $C(N;1,s)$ с чётным d

n	s	N	i
1	$4d+10+i$	$(d/2-2)s+3d/2+3+\lceil i/2\rceil$	0
2	$4d+8+i$	$(d/2-2)s+5d/2+6+\lfloor i/2\rfloor$	0
3	$4d+2+i$	$(d/2-1)s+3d/2+\lceil i/2\rceil$	$0,1,2,3,4$
4	$4d+i$	$(d/2-1)s+5d/2+1+\lfloor i/2\rfloor$	$0,1,2,3,4$
5	$4d-6+i$	$d/2s+3d/2-3+\lceil i/2\rceil$	$0,1,2,3,4,6$
6	$4d-8+i$	$d/2s+5d/2-4+\lfloor i/2\rfloor$	$0,1,2,3,4,6$
7	$4d-14+i$	$(d/2+1)s+3d/2-6+\lceil i/2\rceil$	$0,2$
8	$4d-16+i$	$(d/2+1)s+5d/2-9+\lfloor i/2\rfloor$	$0,2$

Таблица 2

Шаблоны для базовых семейств оптимальных графов $C(N;1,s)$ с нечётным d

n	s	N	i
1	$4d+6+i$	$(d-3)s/2+3d/2+3/2+\lceil i/2\rceil$	$0,1,2$
2	$4d+4+i$	$(d-3)s/2+5d/2+7/2+\lfloor i/2\rfloor$	$0,1,2$
3	$4d-2+i$	$(d-1)s/2+3d/2-3/2+\lceil i/2\rceil$	$0,\ldots,6$
4	$4d-4+i$	$(d-1)s/2+5d/2-3/2+\lfloor i/2\rfloor$	$0,\ldots,6$
5	$4d-10+i$	$(d+1)s/2+3d/2-9/2+\lceil i/2\rceil$	$0,1,2,4$
6	$4d-12+i$	$(d+1)s/2+5d/2-13/2+\lfloor i/2\rfloor$	$0,1,2,4$
7	$4d-18+i$	$(d+3)s/2+3d/2-15/2+\lceil i/2\rceil$	0
8	$4d-20+i$	$(d+3)s/2+5d/2-23/2+\lfloor i/2\rfloor$	0

Определение общих шаблонов для базовых семейств оптимальных графов из табл. 1 и 2 позволит найти доказательства их существования при любых диаметрах. Дополнительные семейства из списков, не вошедшие в число базовых, будут исследованы

в дальнейшем. В п. 6 приведены условия их оптимальности, которые установлены эмпирически и подтверждены экспериментально.

Необходимым условием оптимальности аналитически описываемого семейства циркулянтных графов диаметра d является условие $N_{d-1} < N \leqslant N_d$, или

$$2d^2 - 2d + 2 \leqslant N \leqslant 2d^2 + 2d + 1. \tag{1}$$

Второе дополнительное условие $s < N/2$ следует из известного свойства циркулянтов [2, 3]: циркулянты $C(N; 1, s)$ и $C(N; 1, N - s)$ изоморфны. Таким образом, все функции $N(d)$ порядков графов оптимальных семейств необходимо должны удовлетворять порогу существования (1) и можно таким способом определить для каждого аналитически описываемого семейства минимальный диаметр d_m, начиная с которого семейство может быть оптимальным. Далее исследуем, какие из построенных семейств графов являются бесконечными, то есть существуют при любых диаметрах, больших d_m.

4. Новые бесконечные оптимальные семейства графов чётного диаметра с образующей $s = 4d + \alpha$

Сначала рассмотрим семейства графов с чётным диаметром.

Для следующих 12 семейств оптимальных графов с образующей $s = 4d + \alpha$ существование при любых диаметрах $d \geqslant d_m$ следует в силу изоморфизма их описаний описаниям вида $(N; d, d + 1)$ или $(N; d - 1, d)$.

Теорема 2. Пусть d — чётное число. Тогда существуют такие числа d_m, что при любом $d \geqslant d_m$ семейства циркулянтов

$$C(2d^2 - 3d/2 - 4; 1, 4d - 8), \quad C(2d^2 - 3d/2 - 3; 1, 4d - 6), \quad C(2d^2 - d/2 - 3; 1, 4d - 6),$$
$$C(2d^2 - d/2 - 2; 1, 4d - 4), \quad C(2d^2 + d/2 - 2; 1, 4d - 4), \quad C(2d^2 + d/2 - 1; 1, 4d - 2),$$
$$C(2d^2 + 3d/2 - 1; 1, 4d - 2), \quad C(2d^2 + 3d/2; 1, 4d)$$

оптимальны.

Доказательство. Для всех указанных графов $(N, d+1) = 1$ и $s(d+1) \bmod N = d$. Таким образом, графы $C(N; 1, s)$ изоморфны оптимальным графам $C(N; d, d+1)$. Числа d_m определяются из условия (1). Например, определим d_m для графов семейства $C(2d^2 - 3d/2 - 4; 1, 4d - 8)$: имеем $-3d/2 - 4 \geqslant -2d + 2$, отсюда $d \geqslant 12$. Аналогично определяются значения d_m для остальных семейств. ■

Теорема 3. Пусть d — чётное число. Тогда существуют такие числа d_m, что при любом $d \geqslant d_m$ отпимальны следующие семейства циркулянтов:

$$C(2d^2 - 3d/2; 1, 4d), \qquad C(2d^2 - 3d/2 - 1; 1, 4d + 2),$$
$$C(2d^2 - d/2 - 1; 1, 4d + 2), \qquad C(2d^2 - d/2 - 2; 1, 4d + 4).$$

Доказательство. Для всех указанных графов $(N, d-1) = 1$ и $s(d-1) \bmod N = d$. Таким образом, графы $C(N; 1, s)$ изоморфны оптимальным графам $C(N; d-1, d)$. Значения d_m для всех графов семейств определяются из условия (1). ■

Оптимальность семейства $C(2d^2 - d - 1; 1, 4d + 3)$ при любых чётных $d \geqslant 6$ доказана в [16, пример 5.2].

Следует отметить, что графы всех указанных семейств имеют одновременно минимальный диаметр, совпадающий с $D(N)$, и минимальное среднее расстояние между

вершинами. Докажем оптимальность семейств из табл. 1 с параметрами $n = 1, 3, 5, 7$ и $i = 0$.

Теорема 4. При любом чётном $d \geqslant d_m$ семейства циркулянтов

$$C(2d^2 - 3d/2 - 17; 1, 4d + 10), \qquad C(2d^2 - 3d/2 - 2; 1, 4d + 2),$$
$$C(2d^2 - 3d/2 - 3; 1, 4d - 6), \qquad C(2d^2 - 3d/2 - 20; 1, 4d - 14)$$

оптимальны, где $d_m = 38, 8, 10, 44$ соответственно.

Доказательство. Данные семейства соответствуют в табл. 1 значениям $n = 1, 3, 5, 7$ и $i = 0$. Значения d_m для них получены в силу (1).

Обозначим $\lfloor N/s \rfloor = b$, $N \bmod s = r$. Параметры графов семейств приведены в табл. 3.

Таблица 3

n	b	r	s
1	$d/2 - 2$	$3d/2 + 3$	$4d + 10$
3	$d/2 - 1$	$3d/2$	$4d + 2$
5	$d/2$	$3d/2 - 3$	$4d - 6$
7	$d/2 + 1$	$3d/2 - 6$	$4d - 14$

Все графы данных семейств соответствуют случаю $r < s/2$.

Пусть $D(v)$ означает длину кратчайшего пути (расстояние) из 0 в вершину v, $v = 1, 2, \ldots, N - 1$. Разместим вершины графа на линии, вершины пометим от 0 до N (вершина N по модулю N соответствуют вершине 0). Чтобы доказать, что графы данных семейств имеют диаметр, равный d, будем рассматривать шаги по образующей s на всём расстоянии от 0 до N, а затем их продолжение из N в $2N$, и образующей $-s$ из N в 0, а затем их продолжение из 0 в $-N$.

Интервал вершин $ks \leqslant v \leqslant (k + 1)s$, $k = 0, 1, \ldots, \lfloor b/2 \rfloor$, шагами образующих s и $-s$ разбивается на четыре участка (рис. 5). Вверху обозначены результаты шагов образующих на k-м интервале, внизу — длины участков. Для доказательства теоремы достаточно показать, что расстояния от 0 до всех вершин в интервале $ks \leqslant v \leqslant (k+1)s$ не превышают значения d. Для этого используем следующее свойство циркулянтов $C(N; 1, s)$: если в нём u и v — номера вершин, $u < v$, то $\max\limits_{u \leqslant x \leqslant v} D(x) = \lfloor (v - u + D(u) + D(v))/2 \rfloor$.

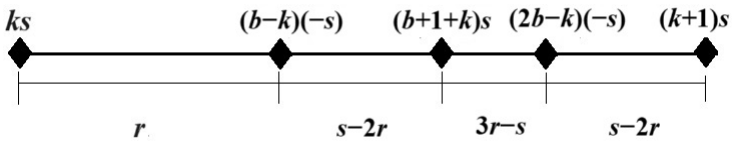

Рис. 5. Интервал вершин $\{ks, \ldots, (k + 1)s\}$

Имеем $D(ks) = k$, $D((b-k)(-s)) = b-k$, $D((b+1+k)s) = b+1+k$, $D((2b-k)(-s)) = 2b-k$, $D((k+1)s) = k+1$. Зная длины всех участков на рис. 5 и применяя указанное свойство, получим максимальные значения функции $D(v)$ для вершин всех четырёх участков:

$$\begin{gathered}\max_{ks\leqslant x\leqslant ks+r} D(x)[1] = \lfloor (b+r)/2 \rfloor,\\ \max_{ks+r\leqslant x\leqslant (k+1)s-r} D(x)[2] = b - r + \lceil s/2 \rceil,\\ \max_{(k+1)s-r\leqslant x\leqslant ks+2r} D(x)[3] = \lfloor (3(b+r) - s + 1)/2 \rfloor,\\ \max_{ks+2r\leqslant x\leqslant (k+1)s} D(x)[4] = b - r + \lceil s/2 \rceil.\end{gathered} \tag{2}$$

Подставив в (2) значения b, r и s из табл. 3, получим максимум расстояний от нуля до вершин всех четырёх участков (табл. 4):

Таблица 4

n	$\max D(x)[1]$	$\max D(x)[2]$	$\max D(x)[3]$	$\max D(x)[4]$
1	d	d	$d-3$	d
3	$d-1$	d	$d-2$	d
5	$d-2$	d	$d-1$	d
7	$d-3$	d	d	d

Теорема 4 доказана. ■

Докажем оптимальность семейств из табл. 1 с параметрами $n = 2, 4, 6, 8$ и $i = 0$.

Теорема 5. При любом чётном $d \geqslant d_m$ семейства циркулянтов

$$\begin{aligned} &C(2d^2 - 3d/2 - 10; 1, 4d + 8), && C(2d^2 - 3d/2 + 1; 1, 4d),\\ &C(2d^2 - 3d/2 - 4; 1, 4d - 8), && C(2d^2 - 3d/2 - 25; 1, 4d - 16)\end{aligned}$$

оптимальны, где $d_m = 24, 4, 12, 54$ соответственно.

Доказательство. Данные семейства соответствуют в табл. 1 значениям $n = 2, 4, 6, 8$ и $i = 0$. Значения d_m для них получены в силу (1). Метод доказательства и основные обозначения такие же, как в теореме 4. При этом табл. 3 заменяется на табл. 5, рис. 5 — на рис. 6, (2) заменяется на (3). В отличие от теоремы 4, графы данных семейств соответствуют случаю $r > s/2$.

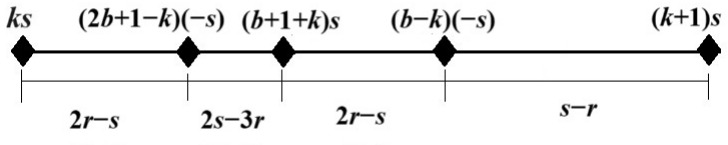

Рис. 6. Интервал вершин $\{ks, \ldots, (k+1)s\}$

Таблица 5

n	b	r	s
2	$d/2 - 2$	$5d/2 + 6$	$4d + 8$
4	$d/2 - 1$	$5d/2 + 1$	$4d$
6	$d/2$	$5d/2 - 4$	$4d - 8$
8	$d/2 + 1$	$5d/2 - 9$	$4d - 16$

Максимальные значения функции $D(v)$:

$$\begin{gathered}\max_{ks\leqslant x\leqslant(k-1)s+2r} D(x)[1] = b + r - \lfloor s/2 \rfloor,\\ \max_{(k-1)s+2r\leqslant x\leqslant(k+1)s-r} D(x)[2] = \lfloor 3(b-r)/2 + s + 1 \rfloor,\\ \max_{(k+1)s-r\leqslant x\leqslant ks+r} D(x)[3] = b + r - \lfloor s/2 \rfloor,\\ \max_{ks+r\leqslant x\leqslant(k+1)s} D(x)[4] = \lfloor (b - r + s + 1)/2 \rfloor.\end{gathered} \tag{3}$$

Подставив значения b, r и s из табл. 5 в (3), получим максимум расстояний от нуля до вершин всех четырёх участков (табл. 6).

Таблица 6

n	$\max D(x)[1]$	$\max D(x)[2]$	$\max D(x)[3]$	$\max D(x)[4]$
2	d	$d-3$	d	d
4	d	$d-2$	d	$d-1$
6	d	$d-1$	d	$d-2$
8	d	d	d	$d-3$

Теорема 5 доказана. ■

5. Новые бесконечные оптимальные семейства графов нечётного диаметра с образующей $s = 4d + \alpha$

Перейдём к доказательствам бесконечности существования семейств оптимальных графов нечётного диаметра.

Применив при любых нечётных $d \geqslant 5$ теорему 5.1 из [16], как следствие получим доказательство оптимальности шести семейств из нашего списка:

$$\begin{array}{lll} C(2d^2-d;1,4d-1), & C(2d^2;1,4d+1), & C(2d^2+d;1,4d+3),\\ C(2d^2-d;1,4d-3), & C(2d^2;1,4d-1), & C(2d^2+d;1,4d+1). \end{array}$$

Эти семейства соответствуют следующим параметрам в табл. 2: $n = 3$, $i = 1$; $n = 3$, $i = 3$; $n = 3$, $i = 5$; $n = 4$, $i = 1$; $n = 4$, $i = 3$; $n = 4$, $i = 5$. Далее доказывается оптимальность при всех нечётных $d \geqslant d_m$ других 12 семейств графов из списка.

Теорема 6. Пусть d — нечётное число. Тогда существуют числа d_m, такие, что при любом $d \geqslant d_m$ следующие семейства циркулянтов оптимальны:

$$\begin{array}{lll} C(2d^2-3d/2+1/2;1,4d-4), & C(2d^2-3d/2-1/2;1,4d-2), & C(2d^2-d/2+1/2;1,4d-2),\\ C(2d^2-d/2-1/2;1,4d), & C(2d^2+d/2+1/2;1,4d), & C(2d^2+d/2-1/2;1,4d+2),\\ C(2d^2+3d/2+1/2;1,4d+2), & C(2d^2+3d/2-1/2;1,4d+4). & \end{array}$$

Доказательство. Для всех указанных графов $(N,d) = 1$ и $sd \bmod N = d+1$. Таким образом, графы $C(N;1,s)$ изоморфны оптимальным графам $C(N;d,d+1)$. Числа d_m для всех семейств определяются из условия (1). ■

Теорема 7. Пусть d — нечётное число. Тогда существуют числа d_m, такие, что при любом $d \geqslant d_m$ следующие семейства циркулянтов оптимальны:

$$\begin{array}{ll} C(2d^2-3d/2-1/2;1,4d-4), & C(2d^2-3d/2+1/2;1,4d-2),\\ C(2d^2-d/2-1/2;1,4d-2), & C(2d^2-d/2+1/2;1,4d). \end{array}$$

Доказательство. Для всех указанных графов $(N, d) = 1$ и $sd \bmod N = d - 1$. Таким образом, графы $C(N; 1, s)$ изоморфны оптимальным графам $C(N; d-1, d)$. Для оптимальных графов диаметра d должно выполняться условие (1). Определим d_m для графов семейства $C(2d^2-3d/2+1/2; 1, 4d-2)$. Имеем $-3d/2+1/2 \geqslant -2d+2$, отсюда с учётом условия $s < N/2$ получаем $d \geqslant 5$. Аналогично определяются значения d_m для остальных семейств. ■

Графы всех указанных семейств имеют одновременно минимальный диаметр, совпадающий с $D(N)$, и минимальное среднее расстояние между вершинами.

Теорема 8. При любом нечётном $d \geqslant d_m$ семейства циркулянтов

$$C(2d^2 - 3d/2 - 15/2; 1, 4d + 6), \qquad C(2d^2 - 3d/2 - 1/2; 1, 4d - 2),$$
$$C(2d^2 - 3d/2 - 19/2; 1, 4d - 10), \qquad C(2d^2 - 3d/2 - 69/2; 1, 4d - 18)$$

оптимальны, где $d_m = 19, 5, 23, 73$ соответственно.

Доказательство. Данные семейства соответствуют значениям $n = 1, 3, 5, 7$ и $i = 0$ в табл. 2. Значения d_m для них получены в силу (1). Доказательство данной теоремы идентично доказательству теоремы 4, включая обозначения, рис. 5, функции (2) и табл. 4, с той разницей, что табл. 3 заменяется на табл. 7. ■

Таблица 7
Параметры семейств из теоремы 8

n	b	r	s
1	$(d-3)/2$	$(3d+3)/2$	$4d+6$
3	$(d-1)/2$	$(3d-3)/2$	$4d-2$
5	$(d+1)/2$	$(3d-9)/2$	$4d-10$
7	$(d+3)/2$	$(3d-15)/2$	$4d-18$

Теорема 9. При любом нечётном $d \geqslant d_m$ семейства циркулянтов

$$C(2d^2 - 3d/2 - 5/2; 1, 4d + 4), \qquad C(2d^2 - 3d/2 + 1/2; 1, 4d - 4),$$
$$C(2d^2 - 3d/2 - 25/2; 1, 4d - 12), \qquad C(2d^2 - 3d/2 - 83/2; 1, 4d - 20)$$

оптимальны, где $d_m = 9, 5, 29, 87$ соответственно.

Доказательство. Данные семейства соответствуют значениям $n = 2, 4, 6, 8$ и $i = 0$ в табл. 2. Значения d_m для них получены в силу (1). Отметим, что для $n = 4$ необходимо учесть дополнительное условие $s < N/2$. Основные обозначения и метод доказательства такие же, как в теореме 5. При этом рис. 6, функции (3), табл. 6 такие же, как в теореме 5, а табл. 5 заменяется на табл. 8. ■

Таблица 8
Параметры семейств из теоремы 9

n	b	r	s
2	$d/2-3/2$	$5d/2+7/2$	$4d+4$
4	$d/2-1/2$	$5d/2-3/2$	$4d-4$
6	$d/2+1/2$	$5d/2-13/2$	$4d-12$
8	$d/2+3/2$	$5d/2-23/2$	$4d-20$

Оптимальность графов оставшихся семейств при любых диаметрах, превышающих нижний порог, может быть доказана путём построения соответствующих схем шагов образующих в прямом и обратном направлениях от нулевой вершины.

Важно отметить следующее свойство оптимальных графов $C(N;1,s)$ с образующей $s=4d+\alpha$. При поиске кратчайших путей в графе достаточно рассмотреть шаги по образующим s $(-s)$ на расстоянии от 0 до $2N$ (от N до $-N$), что составляет два оборота по N. Это свойство позволяет сократить время поиска кратчайших путей при применении таких графов в качестве подсистемы связей в сетях на кристалле, если реализуются алгоритмы маршрутизации с многократными оборотами по N, как, например, в [38, 39].

6. Условие оптимальности семейств графов с образующей $s=4d+\alpha$

Определим условие оптимальности, которое должно выполняться на всём множестве базовых семейств оптимальных графов с образующей $s=4d+\alpha$ для любых диаметров и порядков этих графов. Для этого рассмотрим функции (2) и (3). Участки вершин, на которых достигается значение диаметра d, — это второй и четвёртый в (2) и первый и третий в (3).

Получены следующие условия оптимальности, которые должны выполняться для всех рассмотренных базовых семейств:

$$\lfloor N/s\rfloor - N \bmod s + \lceil s/2\rceil = d \quad \text{при} \quad N \bmod s < s/2; \tag{4}$$

$$\lfloor N/s\rfloor + N \bmod s - \lfloor s/2\rfloor = d \quad \text{при} \quad N \bmod s > s/2. \tag{5}$$

Условие (4) соответствует точкам базовых семейств на линиях с $n=1,3,5,7$, а условие (5) — точкам базовых семейств на линиях с $n=2,4,6,8$. Эти условия определяют соотношения между N, s и d для оптимальных графов с $d=D(N)$.

Проверив выполнение условий оптимальности (4) и (5) на всём датасете, мы получили график рис. 7. Множество графов оптимальных семейств с образующей $s=4d+\alpha$ соответствует точкам второй линии снизу на рис. 7.

Условия оптимальности (6) и (7) для дополнительных (не базовых) семейств оптимальных графов найдены экспериментально:

$$\lfloor N/s\rfloor - N \bmod s + \lceil s/2\rceil = d-1 \quad \text{при} \quad N \bmod s < s/2; \tag{6}$$

$$\lfloor N/s\rfloor + N \bmod s - \lfloor s/2\rfloor = d-1 \quad \text{при} \quad N \bmod s > s/2. \tag{7}$$

Выполнимость их также проверена на всём датасете, получен график рис. 8, аналогичный графику на рис. 7, с ним не пересекающийся и дополняющий его.

Из полученных графиков следует, что найденным условиям оптимальности удовлетворяет не только всё множество семейств оптимальных графов с образующей $s=4d+\alpha$, но и представительная часть графов датасета с линейными образующими других видов, что даёт возможность поиска новых описаний аналитических семейств для других классов образующих.

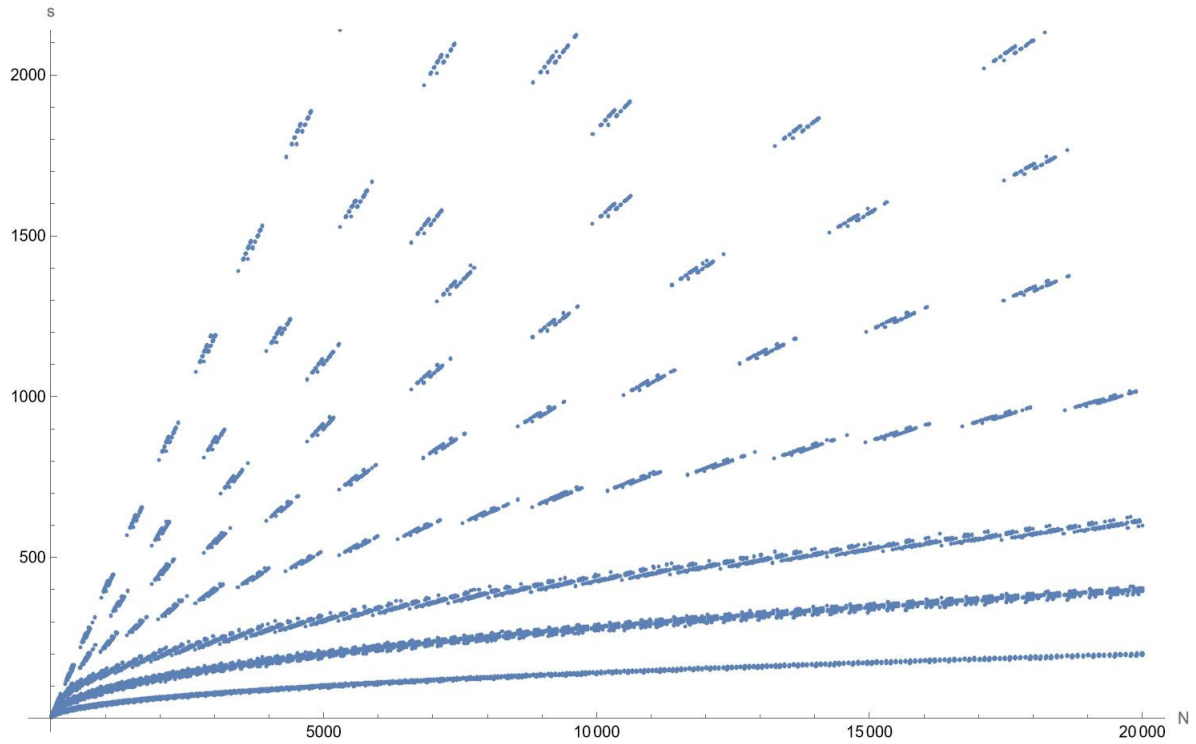

Рис. 7. Выполнение условий (4) и (5) на точках датасета, фрагмент с $N \leqslant 20000$

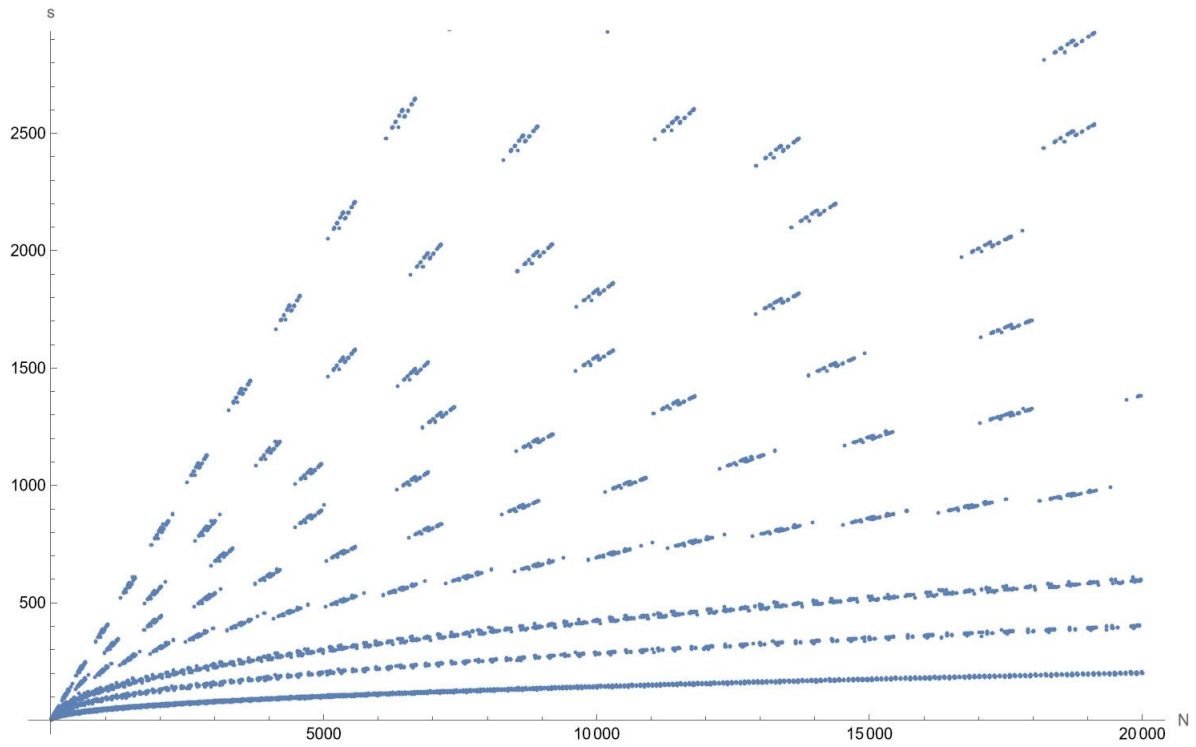

Рис. 8. Выполнение условий (6) и (7) на точках датасета, фрагмент с $N \leqslant 20000$

Заключение

Предложен новый метод открытия аналитических зависимостей параметров описаний семейств оптимальных двухконтурных кольцевых графов, представляющих практический интерес при моделировании систем связи для сетей на кристалле и кластеров

мультипроцессорных систем. Новый подход объединяет визуализацию точек датасета оптимальных графов с их алгоритмически найденными аналитическими описаниями и построение общих схем оптимальных аналитических описаний устойчивых рядов (серий) оптимальных графов. Построены и теоретически обоснованы новые бесконечные семейства оптимальных сетей с линейной образующей вида $s = 4d + \alpha$, где d — диаметр графа. Большая часть найденных семейств сетей имеет наилучшие возможные значения диаметра и среднего расстояния между узлами, что обеспечивает минимизацию задержек в сети и, в конечном итоге, повышение производительности системы [4]. Важным преимуществом применения семейств оптимальных графов с образующей $s = 4d + \alpha$ в качестве топологии в сетях на кристалле является, как показано для некоторых подсемейств в [27, 33], возможность разработки для них эффективных алгоритмов маршрутизации.

Для будущих исследований остаётся открытым вопрос — какие ещё принципы построения семейств оптимальных (субоптимальных) двухконтурных кольцевых графов могут быть реализованы при анализе полученного датасета. Планируется также рассмотреть построение бесконечных семейств оптимальных графов с линейными образующими вида $s = 6d + \alpha$, $8d + \alpha$ и др., которые также, как показывает визуализация на рис. 2, образуют устойчивые, повторяющиеся с ростом диаметра, конфигурации оптимальных описаний.

ЛИТЕРАТУРА

1. *Bermond J.-C., Comellas F., and Hsu D. F.* Distributed loop computer networks: a survey // J. Parallel Distrib. Comput. 1995. No. 24 (1). P. 2–10.
2. *Hwang F. K.* A survey on multi-loop networks // Theoret. Comput. Sci. 2003. V. 299. P. 107–121.
3. *Монахова Э. А.* Структурные и коммуникативные свойства циркулянтных сетей // Прикладная дискретная математика. 2011. № 3. С. 92–115.
4. *Huang X., Ramos A. F., and Deng Y.* Optimal circulant graphs as low-latency network topologies // J. Supercomput. 2022. V. 78. P. 13491–13510.
5. *Monakhova E. A., Monakhov O. G., and Romanov A. Yu.* Routing algorithms in optimal degree four circulant networks based on relative addressing: Comparative analysis for networks-on-chip // IEEE Trans. Netw. Sci. Eng. 2023. V. 10. No. 1. P. 413–425.
6. *Perez-Roses H., Bras-Amoros M., and Seradilla-Merinero J. M.* Greedy routing in circulant networks // Graphs Combinatorics. 2022. V. 38. Iss. 3. `https://doi.org/10.1007/s00373-022-02489-9`.
7. *Pai K.-J., Yang J.-S., Chen G.-Y., and Chang J.-M.* Configuring protection routing via completely independent spanning trees in dense Gaussian on-chip networks // IEEE Trans. Netw. Sci. Eng. 2022. V. 9. No. 2. P. 932–946.
8. *Chen B.-X., Meng J.-X., and Xiao W.-J.* A constant time optimal routing algorithm for undirected double-loop networks // Proc. 1$^{\text{th}}$ Int. Conf. Mobile Ad-hoc and Sensor Networks MSN 2005, Wuhan, China, December 2005. P. 309–316.
9. *Hoffmann R., Deserable D., and Seredynski F.* Cellular automata rules solving the wireless sensor network coverage problem // Natural Computing. 2022. V. 21. P. 417–447.
10. *Chen Y. B., Li Y., and Zheng X.* Research on undirected double-loop data center networks // Proc. Int. Conf. Advanced Cloud and Big Data. Huangshan, China, 2014. P. 180–183.
11. *Erickson A., Stewart I. A., Navaridas J., and Kiasari A. E.* The stellar transformation: From interconnection networks to datacenter networks // Computer Networks. 2017. V. 113. P. 29–45.

12. *Fei J. and Lu C.* Adaptive sliding mode control of dynamic systems using double loop recurrent neural network structure // IEEE Trans. Neural Netw. Learn. Syst. 2018. V. 29. P. 1275–1286.
13. *Monakhov O. G., Monakhova E. A., Romanov A. Y., et al.* Adaptive dynamic shortest path search algorithm in networks-on-chip based on circulant topologies // IEEE Access. 2021. V. 9. P. 160836–160846.
14. *Монахова Э. А.* Об аналитическом описании оптимальных двумерных диофантовых структур однородных вычислительных систем // Вычислительные системы. 1981. № 90. С. 81–91.
15. *Boesch F. and Wang J.-F.* Reliable circulant networks with minimum transmission delay // IEEE Trans. Circuits Syst. 1985. V. 32. No. 12. P. 1286–1291.
16. *Tzvieli D.* Minimal diameter double-loop networks. 1. Large infinite optimal families // Networks. 1991. V. 21. P. 387–415.
17. *Liu H., Li X., and Wang S.* Construction of dual optimal bidirectional double-loop networks for optimal routing // Mathematics. 2022. V. 10. P. 1–17.
18. *Monakhova E. A., Romanov A. Y., and Lezhnev E. V.* Shortest path search algorithm in optimal two-dimensional circulant networks: Implementation for networks-on-chip // IEEE Access. 2020. V. 8. P. 215010–215019.
19. *Монахова Э. А., Монахов О. Г.* Анализ базы данных оптимальных двухконтурных кольцевых сетей // Прикладная дискретная математика. 2024. № 64. С. 56–71.
20. *Монахова Э. А.* Синтез оптимальных диофантовых структур // Вычислительные системы. 1979. № 80. С. 18–35.
21. *Bermond J. C., Illiades G., and Peyrat C.* An optimization problem in distributed loop computer networks // Ann. New York Acad. Sci. 1989. V. 555. C. 45–55.
22. *Yebra J. L. A., Fiol M. A., Morillo P., and Alegre I.* The diameter of undirected graphs associated to plane tessellations // Ars Combinatoria. 1985. No. 20B. P. 159–172.
23. *Sukhov A. M., Romanov A. Y., and Amerikanov A. A.* The problem of a symmetric graph with a maximum number of vertices and minimum diameter // Lobachevskii J. Math. 2023. V. 44. P. 5453–5459.
24. *Du D.-Z., Hsu D. F., Li Q., and Xu J.* A combinatorial problem related to distributed loop networks // Networks. 1990. V. 20. P. 173–180.
25. *Li Y., Chen Y., Tai W., and Wang R.* The minimum distance diagram and diameter of undirected double-loop networks // Proc. 3rd Inter. Conf. ICMEMTC. Taiyuan, China, 2016. P. 1682–1687.
26. *Loudiki L., Kchikech M., and Essaky E. H.* A New Approach for Computing the Distance and the Diameter in Circulant Graphs. `https://arxiv.org/abs/2210.11116`. 2022.
27. *Jha P. K.* Dense bipartite circulants and their routing via rectangular twisted torus // Discr. Appl. Math. 2014. V. 166. P. 141–158.
28. *Jha P. K. and Smith J. D. H.* Cycle Kronecker products that are representable as optimal circulants // Discr. Appl. Math. 2015. V. 181. P. 130–138.
29. *Liu H., Yang Y., and Hu M.* Tight optimal infinite families of undirected double-loop networks // Systems Engineering Theory and Practice. 2002. V. 1. P. 75–79.
30. *Монахова Э. А., Монахов О. Г.* Эволюционный синтез семейств оптимальных двумерных циркулянтных сетей // Вестник СибГУТИ. 2014. № 2. С. 72–81.
31. *Jha K. P.* Tight-optimal circulants vis-a-vis twisted tori // Discr. Appl. Math. 2014. V. 175. P. 24–34.
32. *Chen B.-X., Meng J.-X., and Xiao W.-J.* Some new optimal and suboptimal infinite families of undirected double-loop networks // DMTCS. 2006. V. 8. P. 299–312.

33. *Jha P. K.* Dimension-order routing algorithms for a family of minimal-diameter circulants // J. Inter. Networks. 2013. V. 14. No. 1. Article 1350002.
34. *Bermond J.-C. and Tzvieli D.* Minimal diameter double-loop networks: dense optimal families // Networks. 1991. V. 21. P. 1–9.
35. *Монахова Э. А., Монахов О. Г.* Генерация и анализ датасета оптимальных двухконтурных циркулянтных сетей // Программная инженерия. 2024. Т. 15. № 8. С. 402–410.
36. *Монахова Э. А., Монахов О. Г.* Метод автоматического поиска семейств оптимальных хордальных кольцевых сетей // Дискретн. анализ и исслед. опер. 2024. Т. 31. № 1. С. 85–108.
37. *Монахов О. Г., Монахова Э. А.* Программа вычисления характеристик регулярных графов с параметрическим описанием. Свидетельство об официальной регистрации программ на ЭВМ № 2013619128. М.: Федеральная служба по интеллектуальной собственности, патентам и товарным знакам, 2013.
38. *Romanov A. Y.* Development of routing algorithms in networks-on-chip based on ring circulant topologies // Heliyon. 2019. V. 5. Iss. 4. `https://www.sciencedirect.com/science/article/pii/S2405844018355208`.
39. *Liu H. and Yang Y.* On the fault-tolerant routing in distributed loop networks // J. Electronics (China). 2000. V. 17. P. 84–89.

REFERENCES

1. *Bermond J.-C., Comellas F., and Hsu D. F.* Distributed loop computer networks: a survey. J. Parallel Distrib. Comput., 1995, no. 24 (1), pp. 2–10.
2. *Hwang F. K.* A survey on multi-loop networks. Theoret. Comput. Sci., 2003, vol. 299, pp. 107–121.
3. *Monakhova E. A.* Strukturnye i kommunikativnye svoystva tsirkulyantnykh setey [Structural and communicative properties of circulant networks]. Prikladnaya Diskretnaya Matematika, 2011, no. 3, pp. 92–115. (in Russian)
4. *Huang X., Ramos A. F., and Deng Y.* Optimal circulant graphs as low-latency network topologies. J. Supercomput., 2022, vol. 78, pp. 13491–13510.
5. *Monakhova E. A., Monakhov O. G., and Romanov A. Yu.* Routing algorithms in optimal degree four circulant networks based on relative addressing: Comparative analysis for networks-on-chip. IEEE Trans. Netw. Sci. Eng., 2023, vol. 10, no. 1, pp. 413–425.
6. *Perez-Roses H., Bras-Amoros M., and Seradilla-Merinero J. M.* Greedy routing in circulant networks. Graphs Combinatorics, 2022, vol. 38, iss. 3, `https://doi.org/10.1007/s00373-022-02489-9`.
7. *Pai K.-J., Yang J.-S., Chen G.-Y., and Chang J.-M.* Configuring protection routing via completely independent spanning trees in dense Gaussian on-chip networks. IEEE Trans. Netw. Sci. Eng., 2022, vol. 9, no. 2, pp. 932–946.
8. *Chen B.-X., Meng J.-X., and Xiao W.-J.* A constant time optimal routing algorithm for undirected double-loop networks. Proc. 1th Int. Conf. Mobile Ad-hoc and Sensor Networks MSN 2005, Wuhan, China, December 2005, pp. 309–316.
9. *Hoffmann R., Deserable D., and Seredynski F.* Cellular automata rules solving the wireless sensor network coverage problem. Natural Computing, 2022, vol. 21, pp. 417–447.
10. *Chen Y. B., Li Y., and Zheng X.* Research on undirected double-loop data center networks. Proc. Int. Conf. Advanced Cloud and Big Data, Huangshan, China, 2014, pp. 180–183.
11. *Erickson A., Stewart I. A., Navaridas J., and Kiasari A. E.* The stellar transformation: From interconnection networks to datacenter networks. Computer Networks, 2017, vol. 113, pp. 29–45.

12. *Fei J. and Lu C.* Adaptive sliding mode control of dynamic systems using double loop recurrent neural network structure. IEEE Trans. Neural Netw. Learn. Syst., 2018, vol. 29, pp. 1275–1286.
13. *Monakhov O. G., Monakhova E. A., Romanov A. Y., et al.* Adaptive dynamic shortest path search algorithm in networks-on-chip based on circulant topologies. IEEE Access, 2021, vol. 9, pp. 160836–160846.
14. *Monakhova E. A.* Ob analiticheskom opisanii optimal'nykh dvumernykh diofantovykh struktur odnorodnykh vychislitel'nykh sistem [On analytical representation of optimal two-dimensional Diophantine structures of homogeneous computer systems]. Vychislitel'nye Sistemy, 1981, no. 90, pp. 81–91. (in Russian)
15. *Boesch F. and Wang J.-F.* Reliable circulant networks with minimum transmission delay. IEEE Trans. Circuits Syst., 1985, vol. 32, no. 12, pp. 1286–1291.
16. *Tzvieli D.* Minimal diameter double-loop networks. 1. Large infinite optimal families. Networks, 1991, vol. 21, pp. 387–415.
17. *Liu H., Li X., and Wang S.* Construction of dual optimal bidirectional double-loop networks for optimal routing. Mathematics, 2022, vol. 10, pp. 1–17.
18. *Monakhova E. A., Romanov A. Y., and Lezhnev E. V.* Shortest path search algorithm in optimal two-dimensional circulant networks: Implementation for networks-on-chip. IEEE Access, 2020, vol. 8, pp. 215010–215019.
19. *Monakhova E. A. and Monakhov O. G.* Analiz bazy dannykh optimal'nykh dvukhkonturnykh kol'tsevykh setey [Database analysis of optimal double-loop networks]. Prikladnaya Diskretnaya Matematika, 2024, no. 64, pp. 56–71. (in Russian)
20. *Monakhova E. A.* Sintez optimal'nykh diofantovykh struktur [Synthesis of optimal Diophantine structures]. Vychislitel'nye Sistemy, 1979, no. 80, pp. 18–35. (in Russian)
21. *Bermond J. C., Illiades G., and Peyrat C.* An optimization problem in distributed loop computer networks. Ann. New York Acad. Sci., 1989, vol. 555, pp. 45–55.
22. *Yebra J. L. A., Fiol M. A., Morillo P., and Alegre I.* The diameter of undirected graphs associated to plane tessellations. Ars Combinatoria, 1985, no. 20B, pp. 159–172.
23. *Sukhov A. M., Romanov A. Y., and Amerikanov A. A.* The problem of a symmetric graph with a maximum number of vertices and minimum diameter. Lobachevskii J. Math., 2023, vol. 44,.pp. 5453–5459.
24. *Du D.-Z., Hsu D. F., Li Q., and Xu J.* A combinatorial problem related to distributed loop networks. Networks, 1990, vol. 20, pp. 173–180.
25. *Li Y., Chen Y., Tai W., and Wang R.* The minimum distance diagram and diameter of undirected double-loop networks. Proc. 3rd Inter. Conf. ICMEMTC, Taiyuan, China, 2016, pp. 1682–1687.
26. *Loudiki L., Kchikech M., and Essaky E. H.* A New Approach for Computing the Distance and the Diameter in Circulant Graphs. `https://arxiv.org/abs/2210.11116`. 2022.
27. *Jha P. K.* Dense bipartite circulants and their routing via rectangular twisted torus. Discr. Appl. Math., 2014, vol. 166, pp. 141–158.
28. *Jha P. K. and Smith J. D. H.* Cycle Kronecker products that are representable as optimal circulants. Discr. Appl. Math., 2015, vol. 181, pp. 130–138.
29. *Liu H., Yang Y., and Hu M.* Tight optimal infinite families of undirected double-loop networks. Systems Engineering Theory and Practice, 2002, vol. 1, pp. 75–79.
30. *Monakhova E. A. and Monakhov O. G.* Evolyutsionnyy sintez semeystv optimal'nykh dvumernykh tsirkulyantnykh setey [Evolutionary synthesis of families of optimal two-dimensional circulant networks.] Vestnik SibGUTI, 2014, no. 2, pp. 72–81. (in Russian)

31. *Jha K. P.* Tight-optimal circulants vis-a-vis twisted tori. Discrete Appl. Math., 2014, vol. 175, pp. 24–34.
32. *Chen B.-X., Meng J.-X., and Xiao W.-J.* Some new optimal and suboptimal infinite families of undirected double-loop networks. DMTCS, 2006, vol. 8, pp. 299–312.
33. *Jha P. K.* Dimension-order routing algorithms for a family of minimal-diameter circulants. J. Inter. Networks, 2013, vol. 14, no. 1, article. 1350002.
34. *Bermond J.-C. and Tzvieli D.* Minimal diameter double-loop networks: dense optimal families. Networks, 1991, vol. 21, pp. 1–9.
35. *Monakhova E. A. and Monakhov O. G.* Generatsiya i analiz dataseta optimal'nykh dvukhkonturnykh tsirkulyantnykh setey [Generation and analysis of a dataset of optimal double-loop circulant networks]. Programmnaya Inzheneriya, 2024, vol. 15, no. 8, pp. 402–410. (in Russian)
36. *Monakhova E. A. and Monakhov O. G.* A method for automatic search for families of optimal chordal ring networks J. Appl. Industr. Math., 2024, vol. 18, no. 1, pp. 122–136.
37. *Monakhov O. G., Monakhova E. A.* Programma vychisleniya kharakteristik regulyarnykh grafov s parametricheskim opisaniem. Svidetel'stvo ob ofitsial'noy registratsii programm na EVM № 2013619128 [Program for calculating characteristics of regular graphs with parametric description. Certificate of official registration of computer programs No. 2013619128]. Moscow, Federal Service for Intellectual Property, Patents and Trademarks, 2013. (in Russian)
38. *Romanov A. Y.* Development of routing algorithms in networks-on-chip based on ring circulant topologies. Heliyon, 2019, vol. 5, iss. 4, `https://www.sciencedirect.com/science/article/pii/S2405844018355208`.
39. *Liu H. and Yang Y.* On the fault-tolerant routing in distributed loop networks. J. Electronics (China), 2000, vol. 17, pp. 84–89.

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ И ПРОГРАММИРОВАНИЯ

УДК 510.52 DOI 10.17223/20710410/66/10

О ГЕНЕРИЧЕСКОЙ СЛОЖНОСТИ ПРОБЛЕМЫ ДИСКРЕТНОГО ЛОГАРИФМА В ПОСЛЕДОВАТЕЛЬНОСТЯХ ЛЮКА[1]

А. Н. Рыбалов

Институт математики им. С. Л. Соболева СО РАН, г. Омск, Россия

E-mail: alexander.rybalov@gmail.com

Изучается генерическая сложность проблемы дискретного логарифма в последовательностях Люка. Эта проблема была использована в 1990-е годы новозеландским криптографом П. Смитом для создания аналога классического протокола Диффи — Хеллмана, в котором возведение в степень по целому модулю заменяется на операцию сложения элементов последовательности Люка. Доказывается, что при условии трудноразрешимости проблемы дискретного логарифма в последовательностях Люка в худшем случае и P = BPP существует подпроблема этой проблемы, для которой нет полиномиального генерического алгоритма. Таким образом, обосновано применение данной алгоритмической проблемы в криптографии с открытым ключом, где важна генерическая трудноразрешимость, то есть трудноразрешимость для почти всех входов. Для доказательства используется метод генерической амплификации, который позволяет строить генерически трудные проблемы из проблем, трудных в худшем случае. Основным этапом этого метода является объединение эквивалентных входов в достаточно большие множества. Эквивалентность входов означает, что рассматриваемая проблема на них решается одинаково.

Ключевые слова: *генерическая сложность, последовательности Люка, дискретный логарифм.*

ON THE GENERIC COMPLEXITY OF THE DISCRETE LOGARITHM PROBLEM IN LUCAS SEQUENCES

A. N. Rybalov

Sobolev Institute of Mathematics, Omsk, Russia

In this paper, we study the generic complexity of the discrete logarithm problem over Lucas sequences. This problem was exploited in the 1990s by the New Zealand cryptographer P. Smith to create an analogue of the classic Diffie — Hellman protocol, in which exponentiation modulo an integer is replaced by the operation of adding the elements of the Lucas sequence. It is proved that, given the worst-case intractability of the discrete logarithm problem in Lucas sequences and P = BPP, there exists a subproblem of this problem for which there is no polynomial generic algorithm.

[1]Работа поддержана грантом Российского научного фонда № 22-11-20019.

Thus, this result justifies the application of this algorithmic problem to public-key cryptography, where generic hardness is very important, i.e., hardness for almost all inputs. To prove the theorem, we use the method of generic amplification, which allows us to construct generically hard problems from problems that are hard in the classical sense. The main component of this method is the cloning technique, which combines the input data of a problem into sufficiently large sets of equivalent input data. Equivalence is understood in the sense that the problem is solved similarly for them.

Keywords: *generic complexity, Lucas sequences, discrete logarithm.*

Введение

Последовательности Люка — это линейные рекуррентные последовательности целых чисел второго порядка, являющиеся обобщением классической последовательности чисел Фибоначчи. В 1990-е годы новозеландский криптограф П. Смит предложил использовать последовательности Люка для создания криптосистемы с открытым ключом [1], являющейся аналогом знаменитой системы RSA, а также для создания аналога протокола Диффи — Хеллмана [2], основанного на проблеме дискретного логарифма в этих последовательностях. Основной чертой этих алгоритмов является использование операции сложения элементов последовательностей Люка вместо возведения в степень по целому модулю, как в классических алгоритмах RSA и Диффи — Хеллмана. Впоследствии проводился криптоанализ предложенных систем [3], однако они до сих пор считаются стойкими.

В современной криптографии важно, чтобы алгоритмическая проблема, лежащая в основе стойкости криптосистемы с открытым ключом, являясь (гипотетически) трудной в классическом смысле, оставалась трудной и в генерическом смысле [4], т. е. для почти всех входов. Это объясняется тем, что при случайной генерации ключей в криптографическом алгоритме происходит генерация входа алгоритмической проблемы, лежащей в основе алгоритма. Если эта проблема будет легкоразрешимой почти всегда, то для почти всех таких входов её можно будет быстро решить и ключи почти всегда будут нестойкими. Поэтому проблема должна быть трудной для почти всех входов. Например, таким поведением обладают классические алгоритмические проблемы криптографии: проблема распознавания квадратичных вычетов [5], проблема дискретного логарифма [6], проблема извлечения корня в группах вычетов [7] (обращения функции RSA).

В данной работе изучается генерическая сложность проблемы дискретного логарифма в последовательностях Люка. Доказывается, что при условии её трудноразрешимости в худшем случае и P = BPP можно выделить подпроблему этой проблемы, для которой не существует полиномиального генерического алгоритма. Класс BPP состоит из проблем, разрешимых за полиномиальное время на вероятностных машинах Тьюринга. Считается, что класс BPP совпадает с классом P, то есть любой полиномиальный вероятностный алгоритм можно эффективно дерандомизировать, построив полиномиальный алгоритм, не использующий генератор случайных чисел и решающий ту же самую проблему. Хотя равенство P = BPP до сих пор не доказано, имеются веские основания в его пользу [8].

1. Предварительные сведения

Пусть I — некоторое множество входов. Для подмножества $S \subseteq I$ определим *последовательность относительных плотностей*

$$\rho_n(S) = \frac{|S_n|}{|I_n|}, \quad n = 1, 2, 3, \ldots,$$

где I_n — множество входов размера n; $S_n = S \cap I_n$. *Асимптотической плотностью* множества S назовём верхний предел

$$\rho(S) = \overline{\lim_{n\to\infty}} \rho_n(S).$$

Множество S называется *генерическим*, если $\rho(S)$=1, и *пренебрежимым*, если $\rho(S)$=0. Очевидно, что S генерическое тогда и только тогда, когда его дополнение $I \setminus S$ пренебрежимо.

Алгоритм $\mathcal{A}$ с множеством входов I и множеством выходов $J \cup \{?\}$ $(? \notin J)$ называется *генерическим*, если

1) $\mathcal{A}$ останавливается на всех входах из I;
2) множество $\{x \in I : \mathcal{A}(x) \neq ?\}$ является генерическим.

Генерический алгоритм $\mathcal{A}$ вычисляет функцию $f : I \to I$, если для всех $x \in I$

$$\Big(\mathcal{A}(x) \neq ?\Big) \Rightarrow \Big(\mathcal{A}(x) = f(x)\Big).$$

Напомним некоторые понятия классической теории сложности вычислений. *Время работы* $t_M(x)$ машины Тьюринга M на входе $x \in I$ — это число шагов машины от начала работы до остановки. Машина Тьюринга M *полиномиальна*, если существует полином $p(n)$, такой, что для любого $x \in I$ имеет место $t_M(x) < p(\mathrm{size}(x))$. Класс P состоит из подмножеств I, распознаваемых полиномиальными машинами Тьюринга.

Вероятностная машина Тьюринга — это машина Тьюринга, в программе которой допускаются пары недетерминированных правил, которые одновременно применимы в данной ситуации. В процессе работы такой машины с вероятностью 1/2 выбирается первое правило и с вероятностью 1/2 — второе. Время работы $t_M(x, \tau)$ вероятностной машины Тьюринга на входе x зависит от вычислительного пути (последовательности выполненных команд) τ. Вероятностная машина Тьюринга M называется *полиномиальной*, если существует полином $p(n)$, такой, что для любого x и для любого вычислительного пути τ машины M на x имеет место $t_M(x, \tau) < p(\mathrm{size}(x))$.

Обозначим через $\Pr[M(x) = y]$ вероятность того, что машина M на входе x выдаёт ответ y. Вероятностная машина M *вычисляет* функцию $f : I \to I$, если для любого $x \in I$ выполняется

$$\big(f(x) = y\big) \Rightarrow \Pr[M(x) = y] > 2/3.$$

Проблема распознавания множества $S \subseteq I$ принадлежит классу BPP, если существует вероятностная полиномиальная машина Тьюринга M, вычисляющая характеристическую функцию множества S:

$$\chi_S(x) = \begin{cases} 1, & \text{если } x \in S, \\ 0, & \text{если } x \notin S. \end{cases}$$

Вероятностные машины Тьюринга формализуют понятие алгоритма, использующего генератор случайных чисел. Класс BPP — это класс проблем, эффективно решаемых такими вероятностными алгоритмами.

2. Последовательности Люка и проблема дискретного логарифма

Последовательность Люка — это рекуррентная последовательность пар целых чисел $\big(U_n(P,Q), V_n(P,Q)\big)$, определяемая следующим образом:

$$\begin{gathered} U_0(P,Q) = 0, \quad U_1(P,Q) = 1, \\ V_0(P,Q) = 2, \quad V_1(P,Q) = P, \\ U_{n+2}(P,Q) = P \cdot U_{n+1}(P,Q) - Q \cdot U_n(P,Q), \quad n \geqslant 0, \\ V_{n+2}(P,Q) = P \cdot V_{n+1}(P,Q) - Q \cdot V_n(P,Q), \quad n \geqslant 0. \end{gathered}$$

Числа Фибоначчи являются частным случаем последовательности Люка $\{U_n(1,-1)\}$. Последовательности Люка обладают массой интересных свойств [3]. Перечислим те из них, которые нам понадобятся в дальнейшем. Пусть p — простое число, тогда для любых n, P, Q имеет место

$$V_n(P \bmod p, Q \bmod p) = V_n(P,Q) \bmod p.$$

Для любых n, k, P, Q выполняется

$$V_{nk}(P,Q) = V_n(V_k(P,Q), Q^k). \tag{1}$$

По простому модулю p последовательность $\{V_n(P,Q) : n = 1, \ldots\}$ периодична с некоторым периодом, являющимся делителем числа $p^2 - 1$. Другими словами, имеет место

$$V_n(P,Q) \bmod p = V_{n \bmod (p^2-1)}(P,Q) \bmod p.$$

Опишем реализацию протокола Диффи — Хеллмана (LUCDIF) с помощью последовательностей Люка [2]:

1) Сначала Алиса выбирает простое число p, число $g < p$ и секретное число $a < p$.
2) Затем Алиса вычисляет $V_a(g,1) \bmod p$.
3) После этого Алиса отправляет Бобу сообщение $(V_a(g,1) \bmod p, p, g)$.
4) Боб выбирает своё секретное число $b < p$. Используя его, он получает общий секретный ключ $V_b(V_a(g,1),1) \bmod p$.
5) Затем Боб отправляет Алисе сообщение $V_b(g,1) \bmod p$.
6) Алиса, в свою очередь, вычисляет общий секретный ключ $V_a(V_b(g,1),1) \bmod p$.

Из свойства (1) следует, что общий ключ у них будет один и тот же:

$$V_b(V_a(g,1),1) = V_{ab}(g,1) = V_a(V_b(g,1),1).$$

Криптостойкость этого протокола базируется на сложности проблемы дискретного логарифма для последовательности Люка $\{V_n(g,1) : n \in \mathbb{N}\}$. Опишем эту проблему.

Пусть $L(g,p)$ — множество всех чисел Люка $V_n(g,1)$ по модулю p. *Проблема дискретного логарифма для последовательности Люка* состоит в вычислении функции $\mathrm{dll} : I \to \mathbb{N}$, где I — множество троек (a,g,p), таких, что p — фиксированное простое число, g — фиксированное натуральное число $g < p$, a — произвольное число из $L(g,p)$. Функция dll определяется следующим образом:

$$\mathrm{dll}(a,g,p) = x \;\Leftrightarrow\; a = V_x(g,p).$$

Под размером входа понимается число разрядов в двоичной записи числа p. В настоящее время неизвестно [3] полиномиальных алгоритмов (даже вероятностных) для вычисления функции dll.

Для изучения генерической сложности этой проблемы необходимо провести стратификацию на множестве входов. Рассмотрим любую бесконечную последовательность простых чисел

$$\pi = \{p_1, p_2, \ldots, p_n, \ldots\},$$

удовлетворяющую условию $2^n \leqslant p_n < 2^{n+1}$ для любого n. Будем называть такую последовательность *экспоненциальной*. Определим функцию dll_π как ограничение функции dll на множество троек (a, g, p), таких, что $p \in \pi$. Для этой функции множество всех входов размера n состоит из троек (a, g, p) с фиксированными g, p и произвольным $a \in L(g, p)$. Очевидно, что проблема вычисления dll_π является подпроблемой вычисления dll. Следующая лемма показывает, что некоторые такие подпроблемы так же вычислительно трудны, как и оригинальная проблема.

Лемма 1. Если не существует полиномиального вероятностного алгоритма для вычисления функции dll, то найдётся такая экспоненциальная последовательность простых чисел π, что и для вычисления функции dll_π нет полиномиального вероятностного алгоритма.

Доказательство. Пусть $P_1, P_2, \ldots$ — все полиномиальные вероятностные алгоритмы. Из предположения о несуществовании полиномиального вероятностного алгоритма для вычисления dll следует, что для любого алгоритма P_n найдётся бесконечно много троек (a, g, p), для которых он не может вычислить dll. Тогда можно выбрать последовательность $\pi' = \{p_1, p_2, \ldots\}$ так, чтобы алгоритм P_n не вычислял dll для (a_n, g_n, p_n) и для любого n выполнялось бы $p_{n+1} > 2p_n$, и расширить π' до экспоненциальной последовательности π, добавив, где нужно, новые члены. Заметим, что для вычисления функции dll_π не существует полиномиального алгоритма. ■

3. Основной результат

Теорема 1. Пусть π — любая экспоненциальная последовательность простых чисел. Если существует полиномиальный генерический алгоритм, вычисляющий функцию dll_π, то существует полиномиальный вероятностный алгоритм, вычисляющий dll_π для всех входов.

Доказательство. Пусть полиномиальный генерический алгоритм $\mathcal{A}$ вычисляет функцию dll_π. Построим вероятностный полиномиальный алгоритм $\mathcal{B}$, вычисляющий dll_π на всём множестве входов. Алгоритм $\mathcal{B}$ на входе (a, g, p) размера n работает следующим образом:

1) Повторяет n раз:
2) генерирует случайно и равномерно $b \in \{1, \ldots, p^2 - 2\}$;
3) с помощью алгоритма Евклида вычисляет $(b, p^2 - 1)$;
4) если $(b, p^2 - 1) = 1$, переходит на шаг 7;
5) возвращается на шаг 2;
6) если после n попыток не получено b, взаимно простое с $p^2 - 1$, выдаёт 0;
7) вычисляет $a' = V_b(a, 1) \bmod p$;
8) запускает алгоритм $\mathcal{A}$ на (a', g, p);
9) если $\mathcal{A}(a', g, p) = y \in \mathbb{N}$, то по свойству (1)

$$a' = V_b(a, 1) \bmod p = V_b(V_x(g, 1), 1) \bmod p = V_{bx}(g, 1) \bmod p.$$

Значит, $y = bx \pmod{(p^2 - 1)}$, откуда $x = yb^{-1} \bmod (p^2 - 1)$ находится эффективно с помощью расширенного алгоритма Евклида;

10) если $\mathcal{A}(a', g, p) = ?$, то выдаёт 0.

Алгоритм $\mathcal{B}$ может выдать неправильный ответ на шаге 6 или шаге 10. Докажем, что вероятность этого меньше $1/2$.

Вероятность выдать ответ на шаге 6 равна

$$\Big(1-\frac{\varphi(p^2-1)}{p^2-3}\Big)^n < \Big(1-\frac{C}{\log n}\Big)^n < e^{-Cn/\log n} < 1/4$$

для достаточно больших n. Здесь $\varphi(x)$ — функция Эйлера (количество натуральных чисел, меньших x и взаимно простых с x); использована следующая оценка для функции Эйлера [9]:

$$\varphi(x) > \frac{Cx}{\log\log x}$$

для некоторой константы $C > 0$.

Оценим вероятность выдать ответ на шаге 10. Значение $a' = V_b(a,1) \bmod p = V_{bx \bmod (p^2-1)}(g,1) \bmod p$ пробегает все элементы $L(g,p)$, так как $bx \bmod (p^2-1)$ при $b \in \{1,\dots,p^2-2\}$, таких, что $(b, p^2-1) = 1$, принимает все значения из множества $\{1,\dots,p-1\}$. Поэтому множество

$$\big\{(a',g,p) : b \in \{1,\dots,p^2-2\}\big\}$$

совпадает с множеством всех входов размера n. Но алгоритм $\mathcal{A}$ генерический, поэтому доля тех входов (a',g,p), на которых он выдаёт неопределённый ответ, стремится к нулю с ростом n и с некоторого момента становится меньше $1/4$. ■

Непосредственным следствием теоремы 1 является следующая

Теорема 2. Если для вычисления функции dll не существует полиномиального вероятностного алгоритма, то существует экспоненциальная последовательность простых чисел π, такая, что для вычисления функции dll_π не существует генерического полиномиального алгоритма.

Автор выражает благодарность рецензенту за полезные замечания и предложения по улучшению текста статьи.

ЛИТЕРАТУРА

1. *Smith P. J. and Lennon M. J. J.* LUC: a new public key system // Proc. IFIP/Sec'93. Toronto, Canada, 1993. P. 103–117.
2. *Smith P. and Skinner C.* A public-key cryptosystem and a digital signature system based on the Lucas function analogue to discrete logarithms // LNCS. 1995. V. 917. P. 355–364.
3. *Bleichenbacher D., Bosma W., and Lenstra A.* Some remarks on Lucas-based cryptosystems // LNCS. 1995. V. 963. P. 386–396.
4. *Kapovich I., Miasnikov A., Schupp P., and Shpilrain V.* Generic-case complexity, decision problems in group theory and random walks // J. Algebra. 2003. V. 264. No. 2. P. 665–694.
5. *Рыбалов А. Н.* О генерической сложности проблемы распознавания квадратичных вычетов // Прикладная дискретная математика. 2015. № 2 (28). С. 54–58.
6. *Рыбалов А. Н.* О генерической сложности проблемы дискретного логарифма // Прикладная дискретная математика. 2016. № 3 (33). С. 93–97.
7. *Рыбалов А. Н.* О генерической сложности проблемы извлечения корня в группах вычетов // Прикладная дискретная математика. 2017. № 38. С. 95–100.
8. *Impagliazzo R. and Wigderson A.* P = BPP unless E has subexponential circuits: Derandomizing the XOR Lemma. Proc. 29th STOC. El Paso, ACM, 1997. P. 220–229.
9. *Rosser J. B. and Schoenfeld L.* Approximate formulas for some functions of prime numbers // Illinois J. Math. 1962. V. 6. No. 1. P. 64–94.

REFERENCES

1. *Smith P. J. and Lennon M. J. J.* LUC: a new public key system. Proc. IFIP/Sec'93, Toronto, Canada, 1993, pp 103–117.
2. *Smith P. and Skinner C.* A public-key cryptosystem and a digital signature system based on the Lucas function analogue to discrete logarithms. LNCS, 1995, vol. 917, pp. 355–364.
3. *Bleichenbacher D., Bosma W., and Lenstra A.* Some remarks on Lucas-based cryptosystems. LNCS, 1995, vol. 963, pp. 386–396.
4. *Kapovich I., Miasnikov A., Schupp P., and Shpilrain V.* Generic-case complexity, decision problems in group theory and random walks. J. Algebra, 2003, vol. 264, no. 2, pp. 665–694.
5. *Rybalov A. N.* O genericheskoy slozhnosti problemy raspoznavaniya kvadratichnykh vychetov [On generic complexity of the quadratic residuosity problem]. Prikladnaya Diskretnaya Matematika, 2015, no. 2 (28), pp. 54–58. (in Russian)
6. *Rybalov A. N.* O genericheskoy slozhnosti problemy diskretnogo logarifma [On generic complexity of the discrete logarithm problem]. Prikladnaya Diskretnaya Matematika, 2016, no. 3 (33), pp. 93–97. (in Russian)
7. *Rybalov A. N.* O genericheskoy slozhnosti problemy izvlecheniya kornya v gruppakh vychetov [On generic complexity of the problem of finding roots in groups of residues]. Prikladnaya Diskretnaya Matematika, 2017, no. 38, pp. 95–100. (in Russian)
8. *Impagliazzo R. and Wigderson A.* P = BPP unless E has subexponential circuits: Derandomizing the XOR Lemma. Proc. 29th STOC, El Paso, ACM, 1997, pp. 220–229.
9. *Rosser J.B. and Schoenfeld L.* Approximate formulas for some functions of prime numbers. Illinois J. Math., 1962, vol. 6, no. 1, pp. 64–94.

СВЕДЕНИЯ ОБ АВТОРАХ

АЛЕКСЕЕВ Евгений Константинович — кандидат физико-математических наук, начальник отдела криптографических исследований ООО «КРИПТО-ПРО», г. Москва. E-mail: **alekseev@cryptopro.ru**

БАХАРЕВ Александр Олегович — аспирант Новосибирского государственного университета, г. Новосибирск. E-mail: **a.bakharev@g.nsu.ru**

БЫЗОВ Виктор Александрович — кандидат физико-математических наук, доцент кафедры прикладной математики и информатики Вятского государственного университета, г. Киров. E-mail: **vbyzov@yandex.ru**

ЗАПАНОВ Ринчин Олегович — студент Новосибирского государственного университета, г. Новосибирск. E-mail: **rinchinzapanov@yandex.ru**

ЗЕТКИНА Алёна Игоревна — начальник отдела организации НИРС Управления научных исследований и инноваций Ярославского государственного университета им. П. Г. Демидова, г. Ярославль. E-mail: **a.zetkina1@uniyar.ac.ru**

ЗИНЧЕНКО Сергей Евгеньевич — студент Новосибирского государственного университета, г. Новосибирск. E-mail: **s.zinchenko@alumni.nsu.ru**

ИЛЬЕВ Артём Викторович — кандидат физико-математических наук, научный сотрудник Института математики им. С. Л. Соболева СО РАН, г. Омск.
E-mail: **artyom_iljev@mail.ru**

КАЙГОРОДОВ Евгений Владимирович — кандидат физико-математических наук, доцент кафедры математики, физики и информатики Горно-Алтайского государственного университета, г. Горно-Алтайск. E-mail: **gazetaintegral@gmail.com**

КНЯЗЕВ Олег Викторович — кандидат физико-математических наук, доцент, г. Омск. E-mail: **knyazev50@rambler.ru**

КЯЖИН Сергей Николаевич — кандидат физико-математических наук, заместитель начальника отдела криптографических исследований ООО «КРИПТО-ПРО», г. Москва. E-mail: **kyazhin@cryptopro.ru**

МОНАХОВ Олег Геннадьевич — кандидат технических наук, ведущий научный сотрудник Института вычислительной математики и математической геофизики СО РАН, г. Новосибирск.. E-mail: **monakhov@rav.sscc.ru**

МОНАХОВА Эмилия Анатольевна — кандидат технических наук, доцент, ведущий научный сотрудник Института вычислительной математики и математической геофизики СО РАН, г. Новосибирск.. E-mail: **emilia@rav.sscc.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**

ПРУДНИКОВ Егор Сергеевич — студент Национального исследовательского Томского государственного университета, г. Томск.
E-mail: **egorprudnikov71@gmail.com**

ПУШКАРЕВ Игорь Александрович — кандидат физико-математических наук, доцент, доцент кафедры прикладной математики и информатики Вятского государственного университета, г. Киров. E-mail: **god_sha@mail.ru**

РЫБАЛОВ Александр Николаевич — кандидат физико-математических наук, старший научный сотрудник лаборатории комбинаторных и вычислительных методов алгебры и логики Института математики им. С. Л. Соболева СО РАН, г. Омск.
E-mail: **alexander.rybalov@gmail.com**

СМЫШЛЯЕВ Станислав Витальевич — доктор физико-математических наук, генеральный директор ООО «КРИПТО-ПРО», г. Москва. E-mail: **svs@cryptopro.ru**

СОЛОМАТИН Денис Владимирович — кандидат физико-математических наук, доцент, доцент кафедры математики и методики обучения математике Омского государственного педагогического университета, г. Омск. E-mail: **denis_2001j@bk.ru**

СПИРИДОНОВ Сергей Викторович — студент, Лаборатория ТВП, г. Москва. E-mail: **SpiridonovSV00@yandex.ru**