

ПРИКЛАДНАЯ ДИСКРЕТНАЯ МАТЕМАТИКА

Научный журнал

2018

№ 42

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

РЕДАКЦИОННАЯ КОЛЛЕГИЯ ЖУРНАЛА
«ПРИКЛАДНАЯ ДИСКРЕТНАЯ МАТЕМАТИКА»

Агибалов Г. П., д-р техн. наук, проф. (главный редактор); Девянин П. Н., д-р техн. наук, чл.-корр. Академии криптографии РФ (зам. гл. редактора); Черемушкин А. В., д-р физ.-мат. наук, чл.-корр. Академии криптографии РФ (зам. гл. редактора); Панкратова И. А., канд. физ.-мат. наук, доц. (отв. секретарь); Агиевич С. В., канд. физ.-мат. наук; Алексеев В. Б., д-р физ.-мат. наук, проф.; Быкова В. В., д-р физ.-мат. наук, проф.; Глухов М. М., д-р физ.-мат. наук, академик Академии криптографии РФ; Евдокимов А. А., канд. физ.-мат. наук, проф.; Колесникова С. И., д-р техн. наук; Крылов П. А., д-р физ.-мат. наук, проф.; Логачев О. А., канд. физ.-мат. наук, доц.; Мясников А. Г., д-р физ.-мат. наук, проф.; Романьков В. А., д-р физ.-мат. наук, проф.; Салий В. Н., канд. физ.-мат. наук, проф.; Сафонов К. В., д-р физ.-мат. наук, проф.; Фомичев В. М., д-р физ.-мат. наук, проф.; Харин Ю. С., д-р физ.-мат. наук, чл.-корр. НАН Беларуси; Чеботарев А. Н., д-р техн. наук, проф.; Шоломов Л. А., д-р физ.-мат. наук, проф.

Адрес редакции и издателя: 634050, г. Томск, пр. Ленина, 36
E-mail: vestnik_pdm@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*
Верстка *И. А. Панкратовой*

Подписано к печати 17.12.2018. Формат 60 × 84 $\frac{1}{8}$. Усл. п. л. 15,5. Тираж 300 экз.
Заказ № 3583. Цена свободная. Дата выхода в свет 21.12.2018.

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

СОДЕРЖАНИЕ

| | |
|--|-----|
| ПАМЯТИ МИХАИЛА МИХАЙЛОВИЧА ГЛУХОВА | 5 |
| ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПРИКЛАДНОЙ ДИСКРЕТНОЙ МАТЕМАТИКИ | |
| Миронкин В. О. Об оценках распределения длины отрезка аperiodичности в графе k -кратной итерации равновероятного случайного отображения | 6 |
| Чередник И. В. Один подход к построению кратно транзитивного множества блочных преобразований | 18 |
| МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ | |
| Боровкова И. В., Панкратова И. А., Семенова Е. В. Криптоанализ двух- каскадного конечно-автоматного генератора с функциональным ключом | 48 |
| Agibalov G. P. ElGamal cryptosystems on Boolean functions | 57 |
| ПРИКЛАДНАЯ ТЕОРИЯ ГРАФОВ | |
| Ильев А. В., Ильев В. П. Об одной задаче кластеризации графа с частичным обучением | 66 |
| Ключарёв П. Г. Детерминированные методы построения графов Рамануджана, предназначенных для применения в криптографических алгоритмах, основан- ных на обобщённых клеточных автоматах | 76 |
| ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ В ДИСКРЕТНОЙ МАТЕМАТИКЕ | |
| Кузнецов А. А., Кузнецова А. С. Ресурсно-эффективный алгоритм для ис- следования роста в конечных двупорождённых группах периода 5 | 94 |
| ДИСКРЕТНЫЕ МОДЕЛИ РЕАЛЬНЫХ ПРОЦЕССОВ | |
| Газдюк Е. П., Жихаревич В. В., Никитина О. М., Остапов С. Э. Модели- рование движения одноклеточного микроорганизма «Amoeba Proteus» мето- дом подвижных клеточных автоматов | 104 |
| МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ | |
| Нефёдов В. Н., Смерчинская С. О., Яшина Н. П. Построение агрегирован- ного отношения, минимально удалённого от экспертных предпочтений | 120 |
| СВЕДЕНИЯ ОБ АВТОРАХ | 133 |

CONTENTS

| | |
|---|-----|
| IN MEMORY OF MIKHAIL MIKHAILOVICH GLUKHOV | 5 |
| THEORETICAL BACKGROUNDS OF APPLIED DISCRETE MATHEMATICS | |
| Mironkin V. O. On estimations of distribution of the length of aperiodicity segment in the graph of k -fold iteration of uniform random mapping | 6 |
| Cherednik I. V. One approach to constructing a multiply transitive class of block transformations | 18 |
| MATHEMATICAL METHODS OF CRYPTOGRAPHY | |
| Borovkova I. V., Pankratova I. A., Semenova E. V. Cryptanalysis of 2-cascade finite automata generator with functional key | 48 |
| Agibalov G. P. ElGamal cryptosystems on Boolean functions | 57 |
| APPLIED GRAPH THEORY | |
| Ilev A. V., Il'ev V. P. On a semi-superwized graph clustering problem | 66 |
| Klyucharev P. G. Deterministic methods of Ramanujan graph construction for use in cryptographic algorithms based on generalized cellular automata | 76 |
| COMPUTATIONAL METHODS IN DISCRETE MATHEMATICS | |
| Kuznetsov A. A., Kuznetsova A. S. A resource-efficient algorithm for study the growth in finite two-generator groups of exponent 5 | 94 |
| DISCRETE MODELS FOR REAL PROCESSES | |
| Hazdiuk Ye. P., Zhikharevich V. V., Nikitina O. M., Ostapov S. E. The unicellular microorganisms "Amoeba Proteus" locomotion simulation with the use of movable cellular automata method | 104 |
| MATHEMATICAL BACKGROUNDS OF INTELLIGENT SYSTEMS | |
| Nefedov V. N., Smerchinskaya S. O., Yashina N. P. Constructing an aggre- gated relation with a minimum distance from the expert preferences | 120 |
| BRIEF INFORMATION ABOUT THE AUTHORS | 133 |

ПАМЯТИ МИХАИЛА МИХАЙЛОВИЧА ГЛУХОВА

9 декабря с. г. скончался выдающийся математик и криптограф Советского Союза и России доктор физико-математических наук, профессор, действительный член Академии криптографии Российской Федерации Глухов Михаил Михайлович. Мировая наука, её актуальнейшие области знаний — конечные алгебраические структуры и их приложения в криптографии — понесли величайшую и невозполнимую утрату. Ушёл из жизни Учёный, чьи теоремы о шифрах, не распространяющих искажений шифротекстов, и чьи книги по алгебре и дискретной математике в криптографии стали для нас и наших учеников неисчерпаемым источником новых знаний и образцом настоящей математической культуры. Редакция журнала «Прикладная дискретная математика», членом которой он был, Оргкомитет Всероссийской научной конференции Sibecrypt по компьютерной безопасности и криптографии, в состав которого он входил, лаборатория Компьютерной Криптографии Томского государственного университета, которая благотворно вдохновляется его научной биографией и трудами, горько скорбят в связи с этой утратой. Мы приносим глубочайшие соболезнования всем нашим коллегам, знавшим Михаила Михайловича и сотрудничавшим с ним в науке и образовании, и вместе с ними разделяем огромную горе его родных и близких. Да будет вечной и светлой память о Вас, дорогой Михаил Михайлович!

10.12.2018

Г. П. Агибалов, И. А. Панкратова



ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПРИКЛАДНОЙ ДИСКРЕТНОЙ МАТЕМАТИКИ

УДК 519.212.2+519.719.2

ОБ ОЦЕНКАХ РАСПРЕДЕЛЕНИЯ ДЛИНЫ ОТРЕЗКА АПЕРИОДИЧНОСТИ В ГРАФЕ k -КРАТНОЙ ИТЕРАЦИИ РАВНОВЕРОЯТНОГО СЛУЧАЙНОГО ОТОБРАЖЕНИЯ

В. О. Миронкин

*Национальный исследовательский университет «Высшая школа экономики»,
г. Москва, Россия*

Работа посвящена исследованию случайной величины $\tau_{fk}(x_0)$, равной длине отрезка аперIODИЧНОСТИ произвольной вершины $x_0 \in S = \{1, \dots, n\}$, $n \in \mathbb{N}$, в графе k -кратной итерации равновероятного случайного отображения $f : S \rightarrow S$. Отрезком аперIODИЧНОСТИ, начинающимся в вершине $x_0 \in S$, называется отрезок выходящей из x_0 траектории от x_0 до её первого самопересечения. Исследовано поведение локальной вероятности $\mathbb{P}\{\tau_{fk}(x_0) = z\}$ как функционала от $z \in S$ при фиксированных значениях параметров $k, n \in \mathbb{N}$. Получена двусторонняя оценка $\mathbb{P}\{\tau_{fk}(x_0) = z\}$ для произвольных $k \in \mathbb{N}$, $x_0, z \in S$, таких, что $kz < n$. Для случаев простого k и $k^2z \leq n$ получены эффективно вычисляемые для используемых на практике значений n (2^{256} и более) двусторонние оценки $\mathbb{P}\{\tau_{fk}(x_0) = z\}$, выраженные в элементарных функциях. Для произвольных $k \in \mathbb{N}$, $x_0, z \in S$ выписаны двусторонние оценки для функции распределения $F_{\tau_{fk}(x_0)}(z)$ в случаях $kz < n/2$ и $kz \leq \sqrt{n}$.

Ключевые слова: *равновероятное случайное отображение, итерация случайного отображения, граф отображения, отрезок аперIODИЧНОСТИ, локальная вероятность, распределение.*

DOI 10.17223/20710410/42/1

ON ESTIMATIONS OF DISTRIBUTION OF THE LENGTH OF APERIODICITY SEGMENT IN THE GRAPH OF k -FOLD ITERATION OF UNIFORM RANDOM MAPPING

V. O. Mironkin

National Research University Higher School of Economics, Moscow, Russia

E-mail: mironkin.v@mail.ru

Given $k, n \in \mathbb{N}$, $x_0 \in S = \{1, \dots, n\}$, and $f : S \rightarrow S$, define $x_{i+1} = f^k(x_i)$ for every $i \in \{0, 1, \dots\}$ and $\tau_{fk}(x_0)$ as the least integer i such that $f^k(x_i) = x_j$ for some j , $j < i$. For the local probability $\mathbb{P}\{\tau_{fk}(x_0) = z\}$ and for the distribution function $F_{\tau_{fk}(x_0)}(z)$, the following estimates are obtained. If $kz < n$, then

$$\begin{aligned} \mathbb{P} \left\{ \tau_{fk}(x_0) = z \right\} &> \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} = z}} e^{-(1+\frac{m}{n})\frac{m^2}{2n}} + \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} < z}} \frac{1}{r+k} e^{-(1+\frac{r}{n})\frac{r^2}{2n}} \left(1 - \left(1 - \frac{r+k}{n} \right)^k \right), \\ \mathbb{P} \left\{ \tau_{fk}(x_0) = z \right\} &< \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} = z}} e^{-\frac{(m-1)^2}{2n}} + \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} < z}} \frac{1}{r} e^{-\frac{(r-1)^2}{2n}} \left(1 - \left(1 - \frac{r}{n} \right)^k \right), \end{aligned}$$

where $r = m + \left(z - \frac{m}{(m,k)} - 1 \right) k$. If $k^2 z \leq n$, then

$$\begin{aligned} \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} = z}} e^{-(1+\frac{m}{n})\frac{m^2}{2n}} + \left(1 - \frac{k^2 z}{2n} \right) \frac{k}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} < z}} e^{-(1+\frac{r}{n})\frac{r^2}{2n}} < \\ < \mathbb{P} \left\{ \tau_{fk}(x_0) = z \right\} < \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} = z}} e^{-\frac{(m-1)^2}{2n}} + \frac{k}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} < z}} e^{-\frac{(r-1)^2}{2n}}, \end{aligned}$$

which, for a prime k , is expressed in elementary functions and efficiently computable for used in practice values of n (2^{256} and more). Also, if $kz \leq \sqrt{n}$, then

$$\sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} \leq z}} \frac{r}{n} \left(1 - \frac{r(m+r)}{2n} \right) e^{-(1+\frac{m}{n})\frac{m^2}{2n}} < F_{\tau_{fk}(x_0)}(z) < \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} \leq z}} \frac{r+1}{n} e^{-\frac{(m-1)^2}{2n}},$$

where $r = m + \left(z - \frac{m}{(m,k)} \right) k$. In some cases, the obtained results allow to estimate the allowable period of usage of the encryption keys generated by iterative algorithms and to build criteria for quality assessment of random sequences.

Keywords: equiprobable random mapping, iteration of random mapping, graph of a mapping, aperiodicity segment, local probability, distribution.

Введение

Изучение моделей отображений, построенных на основе случайных равновероятных отображений, представляет собой важную задачу современной теории вероятностей и криптографии. Методика решения большого класса задач, связанных со случайными отображениями, достаточно полно проработана [1, 2], что позволяет строить новые математические модели [3, 4], адекватно описывающие особенности функционирования современных криптографических примитивов. Подобные модели находят применение, например, в рамках исследований свойств и характеристик итерационных алгоритмов преобразования данных, алгоритмов хэширования и выработки псевдослучайных последовательностей.

Данная работа посвящена изучению класса отображений, состоящего из итераций равновероятного случайного отображения, и продолжает цикл исследований [5–10].

Рассмотрим конечное множество $S = \{1, \dots, n\}$, $n > 1$, и пространство равновероятных случайных отображений $(\Omega, \mathcal{F}, \mathbb{P})$, в котором пространством элементарных исходов Ω является множество всех n^n отображений S в себя, алгеброй событий \mathcal{F} — множество всех подмножеств Ω , а вероятностная мера \mathbb{P} задана следующим образом:

$$\mathbb{P}(f) = \frac{1}{n^n}, \quad f \in \Omega. \quad (1)$$

Определение 1. Пусть $f : S \rightarrow S$. Графом отображения f называется ориентированный граф $G_f = (S, E_f)$ с множеством вершин S и множеством ориентированных рёбер $E_f = \{(x, f(x)) : x \in S\} \subset S^2$.

Для произвольного $k \in \mathbb{N}$ через f^k обозначим k -кратную итерацию $\underbrace{f(\dots(f(x)\dots))}_k$ отображения f и введём множества отображений

$$\Omega_k = \{f^k : f \in \Omega\}.$$

Будем считать, что f^0 — тождественное отображение $S \rightarrow S$.

Замечание 1. Распределение f^k не является равновероятным ни на Ω , ни на Ω_k .

Вопросы, связанные с описанием момента первого возвращения в пройденную траекторию, начатую в произвольной вершине $x_0 \in S$ графа G_{f^k} , при действии случайного отображения f^k , $k \geq 1$,

$$x_{i+1} = f^k(x_i), i = 0, 1, 2, \dots,$$

представляют как теоретический, так и практический интерес для ряда приложений криптографии [11–13]. Так, например, при исследовании криптографических свойств итерационных алгоритмов выработки производных ключей одной из наиболее существенных характеристик является число тактов работы алгоритма до появления ранее использованного ключа [14]. В частности, указанная характеристика позволяет определить допустимый срок эксплуатации долговременных ключей шифрования в зависимости от используемого алгоритма выработки производных ключей.

Определение 2. Отрезком аперидичности, начинающимся в вершине $x_0 \in S$ графа G_f , называется отрезок выходящей из x_0 траектории от x_0 до её первого самопересечения.

Через $\tau_f(x_0)$ обозначим случайную величину, равную длине отрезка аперидичности в графе G_f , начинающегося в вершине $x_0 \in S$:

$$\tau_f(x_0) = \min_{t \in \mathbb{N}} \{t : f^t(x_0) \in \{x_0, f(x_0), \dots, f^{t-1}(x_0)\}\}.$$

При этом справедливо соотношение

$$\tau_f(x_0) = \alpha_f(x_0) + \beta_f(x_0),$$

где $\alpha_f(x_0)$ — высота вершины x_0 (расстояние от x_0 до ближайшей циклической вершины); $\beta_f(x_0)$ — длина цикла компоненты G_f , содержащей вершину x_0 .

Замечание 2. Случайные величины $\tau_f(x_0)$, $\alpha_f(x_0)$, $\beta_f(x_0)$ зависят от параметра n . Однако с целью упрощения восприятия материала данная зависимость отражаться не будет.

1. Оценки локальной вероятности

Для любых $i_0, i_1 \in \mathbb{Z}$, $i_0 > i_1$, положим $\prod_{j=i_0}^{i_1} (\dots) \equiv 1$. В работе [5] получено точное выражение для локальной вероятности $\mathbb{P}\{\tau_{f^k}(x_0) = z\}$:

$$\mathbb{P}\{\tau_{f^k}(x_0) = z\} = \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} = z}} \prod_{i=1}^{m-1} \left(1 - \frac{i}{n}\right) + \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} < z}} \sum_{v=1}^k \prod_{i=1}^{m+(z-\frac{m}{(m,k)})k-v} \left(1 - \frac{i}{n}\right), \quad (2)$$

которое может быть использовано при описании вероятностных свойств ряда криптографических алгоритмов. Так, например, для алгоритмов выработки ключевых последовательностей величина $P\{\tau_{fk}(x_0) = 1\}$ соответствует событию, заключающемуся в выборе самого «слабого» долговременного ключа x_0 , порождающего вырожденную последовательность производных ключей вида x_0, x_0, \dots , а величина $P\{\tau_{fk}(x_0) = n\}$ — событию, заключающемуся в выработке максимального числа неповторяющихся ключей, позволяющих обеспечить шифрование максимального объёма данных.

В таблице для случая $n = 256$ представлены приближённые значения вероятности $P\{\tau_{fk}(x_0) = z\}$ при некоторых z, k и произвольном выборе $x_0 \in S$.

| $z \setminus k$ | 1 | 2 | 3 | 4 |
|-----------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1 | $3,9 \cdot 10^{-3}$ | $7,8 \cdot 10^{-3}$ | $7,8 \cdot 10^{-3}$ | $1,2 \cdot 10^{-2}$ |
| 2 | $7,8 \cdot 10^{-3}$ | $1,9 \cdot 10^{-2}$ | $3,0 \cdot 10^{-2}$ | $4,9 \cdot 10^{-2}$ |
| 2^2 | $1,5 \cdot 10^{-2}$ | $4,0 \cdot 10^{-2}$ | $5,6 \cdot 10^{-2}$ | $8,7 \cdot 10^{-2}$ |
| 2^4 | $3,9 \cdot 10^{-2}$ | $4,0 \cdot 10^{-2}$ | $2,7 \cdot 10^{-2}$ | $2,3 \cdot 10^{-2}$ |
| 2^6 | $4,5 \cdot 10^{-5}$ | $2,7 \cdot 10^{-6}$ | $2,3 \cdot 10^{-6}$ | $2,3 \cdot 10^{-6}$ |
| 2^8 | $2,7 \cdot 10^{-110}$ | $1,0 \cdot 10^{-112}$ | $1,0 \cdot 10^{-112}$ | $1,0 \cdot 10^{-112}$ |

Анализ выражения (2) показал, что при фиксированных n и k максимум вероятности $P\{\tau_{fk}(x_0) = z\}$ достигается при z , принадлежащем некоторой окрестности $\sqrt{\frac{\pi n}{2k}}$. При этом в случае простого k величина $P\{\tau_{fk}(x_0) = z\}$ монотонно возрастает до своего единственного максимума и затем монотонно убывает, а в случае составного k ведёт себя немонотонно (рис. 1).

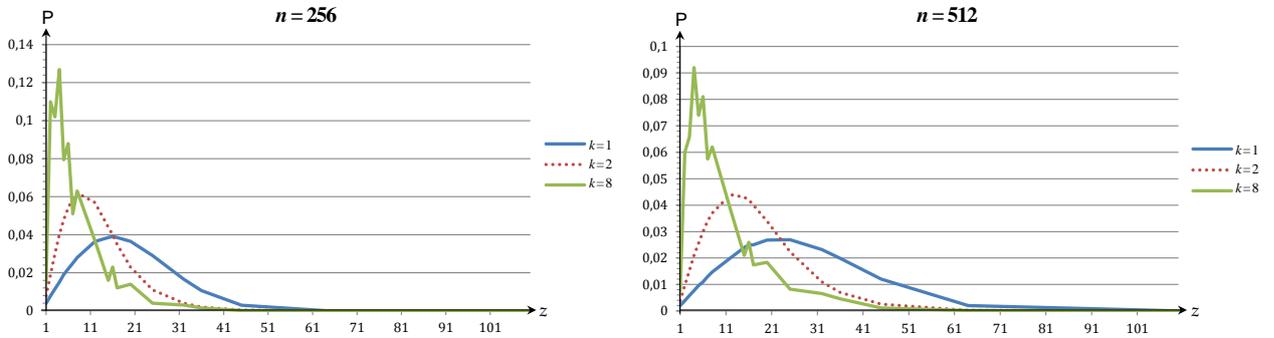


Рис. 1. Зависимость $P\{\tau_{fk}(x_0) = z\}$ от величины z

Следует заметить, что выражение (2) имеет достаточно сложный аналитический вид, что может создавать трудности при вычислении $P\{\tau_{fk}(x_0) = z\}$ для больших значений n . В связи с этим становится актуальной задача получения оценочных выражений величины $P\{\tau_{fk}(x_0) = z\}$, эффективно вычисляемых без привлечения высокопроизводительных ЭВМ.

Утверждение 1. Пусть случайное отображение $f : S \rightarrow S$ имеет распределение (1) на Ω . Тогда при любых таких $k \in \mathbb{N}$, $x_0, z \in S$, что $kz < n$, справедливы следующие неравенства:

$$P\{\tau_{fk}(x_0) = z\} > \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} = z}} e^{-(1+\frac{m}{n})\frac{m^2}{2n}} + \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} < z}} \frac{1}{r+k} e^{-(1+\frac{r}{n})\frac{r^2}{2n}} \left(1 - \left(1 - \frac{r+k}{n}\right)^k\right); \quad (3)$$

$$\mathbb{P} \left\{ \tau_{fk}(x_0) = z \right\} < \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} = z}} e^{-\frac{(m-1)^2}{2n}} + \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} < z}} \frac{1}{r} e^{-\frac{(r-1)^2}{2n}} \left(1 - \left(1 - \frac{r}{n} \right)^k \right), \quad (4)$$

где $r = m + \left(z - \frac{m}{(m,k)} - 1 \right) k$.

Доказательство. Представим выражение (2) в эквивалентном виде:

$$\mathbb{P} \left\{ \tau_{fk}(x_0) = z \right\} = \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} = z}} \prod_{i=1}^{m-1} \left(1 - \frac{i}{n} \right) + \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} < z}} \sum_{v=1}^k \prod_{i=1}^{m + \left(z - \frac{m}{(m,k)} - 1 \right) k + v - 1} \left(1 - \frac{i}{n} \right). \quad (5)$$

Рассмотрим второе слагаемое в (5) и преобразуем сумму вида $\sum_{v=1}^k \prod_{i=1}^{r+v-1} \left(1 - \frac{i}{n} \right)$, где $r \geq 1$, таким образом:

$$\sum_{v=1}^k \prod_{i=1}^{r+v-1} \left(1 - \frac{i}{n} \right) = \prod_{i=1}^{r-1} \left(1 - \frac{i}{n} \right) \sum_{w=0}^{k-1} \prod_{j=0}^w \left(1 - \frac{r+j}{n} \right).$$

При этом для $0 < \frac{r}{n} < \frac{r+k}{n} < 1$ справедливы следующие оценки сверху и снизу:

$$\begin{aligned} \sum_{w=0}^{k-1} \prod_{j=0}^w \left(1 - \frac{r+j}{n} \right) &< \sum_{w=0}^{k-1} \left(1 - \frac{r}{n} \right)^w = \frac{n}{r} \left(1 - \left(1 - \frac{r}{n} \right)^k \right), \\ \sum_{w=0}^{k-1} \prod_{j=0}^w \left(1 - \frac{r+j}{n} \right) &> \sum_{w=0}^{k-1} \left(1 - \frac{r+k}{n} \right)^w = \frac{n}{r+k} \left(1 - \left(1 - \frac{r+k}{n} \right)^k \right). \end{aligned} \quad (6)$$

Тогда с учётом (6) и двустороннего неравенства [15]

$$e^{-(1+\frac{r}{n})\frac{r^2}{2n}} \leq \prod_{i=1}^{r-1} \left(1 - \frac{i}{n} \right) \leq e^{-\frac{(r-1)^2}{2n}},$$

справедливого для $1 \leq r \leq n$, получаем выражение

$$\frac{n}{r+k} e^{-(1+\frac{r}{n})\frac{r^2}{2n}} \left(1 - \left(1 - \frac{r+k}{n} \right)^k \right) < \sum_{v=1}^k \prod_{i=1}^{r+v-1} \left(1 - \frac{i}{n} \right) < \frac{n}{r} e^{-\frac{(r-1)^2}{2n}} \left(1 - \left(1 - \frac{r}{n} \right)^k \right). \quad (7)$$

Положим $r = m + \left(z - \frac{m}{(m,k)} - 1 \right) k$. Из ограничений на суммирование во втором слагаемом в (5) $\left(\frac{m}{(m,k)} < z \text{ и } m \geq 1 \right)$ следует, что $r \geq 1$. Далее из условия $kz < n$ получаем цепочку неравенств:

$$r+k = m + \left(z - \frac{m}{(m,k)} \right) k = kz + m \left(1 - \frac{k}{(m,k)} \right) \leq kz < n.$$

Таким образом, $0 < \frac{r}{n} < \frac{r+k}{n} < 1$ и поэтому неравенство (7) выполняется для искомой вероятности. ■

При $k = 1$ результат утверждения 1 позволяет получить известное неравенство [4]:

$$\frac{z}{n} e^{-(1+\frac{z}{n})\frac{z^2}{2n}} < \mathbf{P} \{ \tau_f(x_0) = z \} < \frac{z}{n} e^{-\frac{(z-1)^2}{2n}}.$$

Следует также заметить, что в случаях, когда число итераций k отображения f невелико или когда множество S имеет большую мощность, оценочные выражения для величины $\mathbf{P} \{ \tau_f(x_0) = z \}$ имеют более простой вид.

Следствие 1. Пусть случайное отображение $f : S \rightarrow S$ имеет равномерное распределение (1) на Ω . Тогда при любых таких $k \in \mathbb{N}$ и $x_0, z \in S$, что $k^2 z \leq n$, справедливо двойное неравенство

$$\begin{aligned} & \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} = z}} e^{-(1+\frac{m}{n})\frac{m^2}{2n}} + \left(1 - \frac{k^2 z}{2n}\right) \frac{k}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} < z}} e^{-(1+\frac{r}{n})\frac{r^2}{2n}} < \\ & < \mathbf{P} \{ \tau_{f^k}(x_0) = z \} < \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} = z}} e^{-\frac{(m-1)^2}{2n}} + \frac{k}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} < z}} e^{-\frac{(r-1)^2}{2n}}, \end{aligned} \quad (8)$$

где $r = m + \left(z - \frac{m}{(m,k)} - 1\right)k$.

Доказательство. Для оценки второй группы слагаемых в (3) и (4) рассмотрим двустороннее неравенство [16]

$$1 - kx \leq (1 - x)^k \leq 1 - kx + C_k^2 x^2, \quad (9)$$

справедливое при $-1 < x < 1$, где $k \in \mathbb{N}$. При этом левое неравенство в (9) является содержательным при $1 - kx \geq 0$, т. е. при $x \leq 1/k$. Покажем, что $x = r/n$ удовлетворяет неравенству $x \leq 1/k$. Учитывая, что $1 - \frac{k}{(m,k)} \leq 0$, имеем цепочку неравенств

$$\frac{r}{n} = \frac{m + \left(z - \frac{m}{(m,k)} - 1\right)k}{n} \leq \frac{kz + m \left(1 - \frac{k}{(m,k)}\right) - k}{n} \leq \frac{kz - k}{n} < \frac{kz}{n} \leq \frac{1}{k}.$$

Тогда искомая оценка сверху величины $\mathbf{P} \{ \tau_{f^k}(x_0) = z \}$ следует из соотношений (3), (4) и (9) при $x = r/n$.

Содержательность правого неравенства в (9) имеет место при $1 - kx + C_k^2 x^2 \leq 1$, т. е. в случае $0 \leq x \leq \frac{2}{k-1}$, что справедливо для $x = \frac{r+k}{n}$:

$$0 < \frac{r+k}{n} = \frac{m + \left(z - \frac{m}{(m,k)}\right)k}{n} = \frac{kz + m \left(1 - \frac{k}{(m,k)}\right)}{n} \leq \frac{kz}{n} \leq \frac{2}{k-1}.$$

В итоге оценка снизу для $\mathbf{P} \{ \tau_{f^k}(x_0) = z \}$ следует из соотношений (3), (4) и (9) при $x = \frac{r+k}{n}$. ■

Замечание 3. Точность неравенства (9) возрастает с уменьшением значения kx (увеличением n , соответственно). Поэтому для значений параметра n , используемых в ряде практических приложений и имеющих порядок 2^{256} и выше, выражение (8) позволяет получать достаточно точные приближения.

Отметим, что неравенства (3), (4) и (8) существенно зависят от канонического разложения k . Для значений k с малым числом собственных делителей соответствующие нерегулярные суммы заметно упрощаются.

Следствие 2. Пусть случайное отображение $f : S \rightarrow S$ имеет равномерное распределение (1) на Ω и k — простое. Тогда при любых таких $x_0, z \in S$, что $k^2 z \leq n$, справедливы следующие оценки сверху:

— при $z = 1$

$$\mathbb{P} \{ \tau_{fk}(x_0) = 1 \} < \frac{1}{n} \left(1 + e^{-\frac{(k-1)^2}{2n}} \right);$$

— при $z > 1, k \nmid z$

$$\mathbb{P} \{ \tau_{fk}(x_0) = z \} < \frac{kz - z + 2}{n} e^{-\frac{(z-2)^2}{2n}} + \frac{kz - k + 1}{n} e^{-\frac{k^2(z-2)^2}{2n}};$$

— при $z > 1, k|z$

$$\mathbb{P} \{ \tau_{fk}(x_0) = z \} < \frac{kz - z + 1}{n} e^{-\frac{(z-2)^2}{2n}} + \frac{kz - k + 1}{n} e^{-\frac{k^2(z-2)^2}{2n}}.$$

Доказательство. Пусть $z = 1$. При этом $\frac{m}{(m, k)} \geq 1$. Тогда равенство $\frac{m}{(m, k)} = 1$ выполняется только при $m \in \{1, k\}$, поэтому с учётом (8) оценка сверху принимает следующий вид:

$$\mathbb{P} \{ \tau_{fk}(x_0) = 1 \} < \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m, k)} = 1}} e^{-\frac{(m-1)^2}{2n}} = \frac{1}{n} \left(1 + e^{-\frac{(k-1)^2}{2n}} \right).$$

Пусть теперь $z > 1$, тогда

$$\frac{m}{(m, k)} = \begin{cases} m, & k \nmid m, \\ \frac{m}{k}, & k|m, \end{cases} \Rightarrow r = \begin{cases} m(1-k) + k(z-1), & k \nmid m, \\ k(z-1), & k|m. \end{cases}$$

Рассмотрим случай $k \nmid z$. Группируя слагаемые, стоящие в правой части неравенства (8) и соответствующие значениям m , таким, что $k \nmid m$ и $k|m$, получаем цепочку неравенств:

$$\begin{aligned} \mathbb{P} \{ \tau_{fk}(x_0) = z \} &< \frac{1}{n} \left(e^{-\frac{(z-1)^2}{2n}} + e^{-\frac{(kz-1)^2}{2n}} \right) + \frac{k}{n} \left(z - \frac{z-1}{k} \right) e^{-\frac{(z-2)^2}{2n}} + \\ &+ \frac{k(z-1)}{n} e^{-\frac{(k(z-1)-1)^2}{2n}} < \frac{kz - z + 2}{n} e^{-\frac{(z-2)^2}{2n}} + \frac{kz - k + 1}{n} e^{-\frac{k^2(z-2)^2}{2n}}. \end{aligned}$$

Для случая $k|z$, рассуждая аналогично, получаем

$$\begin{aligned} \mathbb{P} \{ \tau_{fk}(x_0) = z \} &< \frac{1}{n} e^{-\frac{(kz-1)^2}{2n}} + \frac{k}{n} \left(z - \frac{z-1}{k} \right) e^{-\frac{(z-2)^2}{2n}} + \\ &+ \frac{k(z-1)}{n} e^{-\frac{(k(z-1)-1)^2}{2n}} < \frac{kz - z + 1}{n} e^{-\frac{(z-2)^2}{2n}} + \frac{kz - k + 1}{n} e^{-\frac{k^2(z-2)^2}{2n}}. \end{aligned}$$

Следствие доказано. ■

Следствие 3. Пусть случайное отображение $f : S \rightarrow S$ имеет равномерное распределение (1) на Ω и k — простое. Тогда при любых таких $x_0, z \in S$, что $k^2 z < n$, справедливы следующие оценки снизу:

— при $z = 1$

$$\mathbb{P} \{ \tau_{fk}(x_0) = 1 \} > \frac{1}{n} \left(1 + e^{-\left(1 + \frac{k}{n}\right) \frac{k^2}{2n}} \right);$$

— при $z > 1, k \nmid z$

$$\mathbb{P} \{ \tau_{fk}(x_0) = z \} > \frac{1}{n} e^{-\left(1 + \frac{z}{n}\right) \frac{z^2}{2n}} + \frac{(2k-1)(z-1)(2n-k^2z) + 2n}{2n^2} e^{-\left(1 + \frac{kz}{n}\right) \frac{k^2 z^2}{2n}};$$

— при $z > 1, k|z$

$$\mathbb{P} \{ \tau_{fk}(x_0) = z \} > \frac{(2k-1)(z-1)(2n-k^2z) + 2n}{2n^2} e^{-\left(1 + \frac{kz}{n}\right) \frac{k^2 z^2}{2n}}.$$

Доказательство. Для $z = 1$ неравенство очевидно. Для $z > 1$, повторяя рассуждения следствия 2, при $k \nmid z$ получаем

$$\begin{aligned} \mathbb{P} \{ \tau_{fk}(x_0) = z \} &> \frac{1}{n} \left(e^{-\left(1 + \frac{z}{n}\right) \frac{z^2}{2n}} + e^{-\left(1 + \frac{kz}{n}\right) \frac{k^2 z^2}{2n}} \right) + \\ &+ \left(1 - \frac{k^2 z}{2n} \right) \frac{k}{n} \left((z-1) \left(1 - \frac{1}{k} \right) e^{-\frac{(k(z-2)+1)^2}{2n}} + (z-1) e^{-\frac{k^2(z-1)^2}{2n}} \right) > \\ &> \frac{1}{n} e^{-\left(1 + \frac{z}{n}\right) \frac{z^2}{2n}} + \frac{(2k-1)(z-1)(2n-k^2z) + 2n}{2n^2} e^{-\left(1 + \frac{kz}{n}\right) \frac{k^2 z^2}{2n}}. \end{aligned}$$

Для случая $k|z$ имеем

$$\begin{aligned} \mathbb{P} \{ \tau_{fk}(x_0) = z \} &> \frac{1}{n} e^{-\left(1 + \frac{kz}{n}\right) \frac{k^2 z^2}{2n}} + \\ &+ \left(1 - \frac{k^2 z}{2n} \right) \frac{k}{n} \left((z-1) \left(1 - \frac{1}{k} \right) e^{-\frac{(k(z-2)+1)^2}{2n}} + (z-1) e^{-\frac{k^2(z-1)^2}{2n}} \right) > \\ &> \frac{(2k-1)(z-1)(2n-k^2z) + 2n}{2n^2} e^{-\left(1 + \frac{kz}{n}\right) \frac{k^2 z^2}{2n}}. \end{aligned}$$

Следствие доказано. ■

Замечание 4. Оценки, полученные в следствиях 2 и 3, могут быть эффективно вычислены для используемых на практике значений n .

Через $\nu_{fk}^{(z)}$, $k \in \mathbb{N}$, обозначим случайную величину, равную числу вершин в случайном графе G_{fk} , длина отрезка аperiodичности которых равна $z \in S$. Так как $\mathbb{E} \nu_{fk}^{(z)} = n \mathbb{P} \{ \tau_{fk}(x_0) = z \}$, то с соответствующей поправкой на множитель n для величины $\mathbb{E} \nu_{fk}^{(z)}$ справедливы все представленные выше результаты.

2. Оценки функции распределения

Из работы [5] известно точное выражение для функции распределения $F_{\tau_{fk}(x_0)}(z)$, $z \in S$, случайной величины $\tau_{fk}(x_0)$:

$$F_{\tau_{fk}(x_0)}(z) = \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} \leq z}} \sum_{t=0}^{\left(z - \frac{m}{(m,k)}\right)k} \prod_{i=1}^{m+t-1} \left(1 - \frac{i}{n} \right), \quad (10)$$

имеющее также ряд практических приложений. В частности, значения $F_{\tau_{fk}(x_0)}(z)$ могут быть использованы при оценке среднего допустимого объёма ключевого множества, вырабатываемого на основе фиксированного долговременного ключа с помощью некоторого итерационного алгоритма.

По причине, связанной со сложностью вычисления (10), получим оценочные выражения для величины $F_{\tau_{fk}(x_0)}(z)$.

Утверждение 2. Пусть случайное отображение $f : S \rightarrow S$ имеет распределение (1) на Ω . Тогда при любых таких $k \in \mathbb{N}$, $x_0, z \in S$, что $kz < n/2$, справедливо следующее двустороннее неравенство:

$$F_{\tau_{fk}(x_0)}(z) > \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} \leq z}} e^{-(1+\frac{m}{n})\frac{m^2}{2n}} \left(\frac{1}{n} + \frac{1}{m+r} \left(1 - \left(1 - \frac{m+r}{n} \right)^r \right) \right); \quad (11)$$

$$F_{\tau_{fk}(x_0)}(z) < \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} \leq z}} e^{-\frac{(m-1)^2}{2n}} \left(\frac{1}{n} + \frac{1}{m} \left(1 - \left(1 - \frac{m}{n} \right)^r \right) \right). \quad (12)$$

Здесь $r = m + \left(z - \frac{m}{(m,k)} \right) k$.

Доказательство. Выделим в (10) слагаемые, соответствующие значению $t = 0$:

$$F_{\tau_{fk}(x_0)}(z) = \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} \leq z}} \prod_{i=1}^{m-1} \left(1 - \frac{i}{n} \right) + \frac{1}{n} \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} \leq z}} \sum_{t=1}^{(z-\frac{m}{(m,k)})k} \prod_{i=1}^{m+t-1} \left(1 - \frac{i}{n} \right).$$

Рассмотрим сумму вида $\sum_{t=1}^r \prod_{i=1}^{m+t-1} \left(1 - \frac{i}{n} \right)$, где $m \geq 1$. При этом выполняется неравенство $0 < \frac{m}{n} < \frac{m+r}{n} < 1$. Действительно,

$$0 < \frac{m}{n} < \frac{m+r}{n} = \frac{2m + \left(z - \frac{m}{(m,k)} \right) k}{n} = \frac{kz + m \left(2 - \frac{k}{(m,k)} \right)}{n} \leq \frac{kz + m}{n} \leq \frac{2kz}{n} < 1.$$

Тогда, рассуждая, как в утверждении 1, получаем искомое выражение. ■

Следствие 4. Пусть случайное отображение $f : S \rightarrow S$ имеет распределение (1) на Ω . Тогда при любых таких $k \in \mathbb{N}$, $x_0, z \in S$, что $kz \leq \sqrt{n}$, справедливо

$$\sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} \leq z}} \frac{r}{n} \left(1 - \frac{r(m+r)}{2n} \right) e^{-(1+\frac{m}{n})\frac{m^2}{2n}} < F_{\tau_{fk}(x_0)}(z) < \sum_{\substack{m \geq 1, \\ \frac{m}{(m,k)} \leq z}} \frac{r+1}{n} e^{-\frac{(m-1)^2}{2n}},$$

где $r = m + \left(z - \frac{m}{(m,k)} \right) k$.

Доказательство. Оценим снизу выражение $\frac{1}{n} + \frac{1}{m+r} \left(1 - \left(1 - \frac{m+r}{n} \right)^r \right)$, стоящее под знаком суммы в (11). Для $k = r$ правое неравенство в (9) содержательно при условии $x \leq \frac{2}{r-1}$, которое выполняется для $x = \frac{m+r}{n}$. Действительно,

$$\frac{m+r}{n} = \frac{2m + \left(z - \frac{m}{(m,k)} \right) k}{n} = \frac{kz + m \left(2 - \frac{k}{(m,k)} \right)}{n} \leq \frac{kz + m}{n} \leq \frac{2kz}{n} \leq \frac{2}{r-1},$$

где последнее неравенство справедливо в условиях следствия $kz(r-1) < (kz)^2 \leq n$. Тогда имеет место цепочка соотношений

$$\begin{aligned} \frac{1}{n} + \frac{1}{m+r} \left(1 - \left(1 - \frac{m+r}{n} \right)^r \right) &> \frac{1}{m+r} \left(1 - \left(1 - \frac{r(m+r)}{n} + \frac{r(r-1)(m+r)^2}{2n^2} \right) \right) = \\ &= \frac{r}{n} - \frac{r(r-1)(m+r)}{2n^2} > \frac{r}{n} - \frac{r^2(m+r)}{2n^2}, \end{aligned}$$

с учётом которой из (11) следует искомая нижняя оценка.

Левое неравенство в (9) для $k=r$ содержательно при $x \leq \frac{1}{r}$, что выполняется для $x = \frac{m}{n}$:

$$\frac{m}{n} \leq \frac{1}{r} \leq \frac{2}{r-1},$$

поскольку $mr \leq kZR < (kz)^2 \leq n$. Поэтому, оценивая сверху выражение $\frac{1}{n} + \frac{1}{m} \left(1 - \left(1 - \frac{m}{n} \right)^r \right)$ с помощью (9), получаем искомую оценку. ■

Заключение

В работе [5] для локальной вероятности и распределения длины отрезка аперIODичности получены точные выражения, которые находят применение в рамках исследований и развития модели k -кратной итерации равновероятного случайного отображения [17, 18]. Однако на практике размер задач (величина n), как правило, достигает значений порядка 2^{256} и выше, вследствие чего вычисление по полученным в [5] формулам на персональных ЭВМ (без привлечения дополнительного оборудования) становится затруднительным.

Результаты настоящей работы позволяют эффективно оценивать значения соответствующих характеристик для практических задач современной криптографии. В частности, они могут быть использованы при построении статистических критериев выявления неравновероятности в последовательностях, формируемых на основе итерационных алгоритмов.

ЛИТЕРАТУРА

1. Колчин В. Ф. Случайные отображения. М.: Наука, 1984.
2. Harris B. Probability distributions related to random mapping // Ann. Math. Statist. 1960. V. 31. No. 4. P. 1045–1062.
3. Dalal A. and Schmutz E. Compositions of random functions on a finite set // Electr. J. Comb. 2002. V. 9. No. R26.
4. Flajolet P. and Odlyzko A. Random mapping statistics // LNCS. 1989. V. 434. P. 329–354.
5. Зубков А. М., Миронкин В. О. Распределение длины отрезка аперIODичности в графе k -кратной итерации случайного равновероятного отображения // Математические вопросы криптографии. 2017. Т. 8. № 4. С. 63–74.
6. Миронкин В. О., Михайлов В. Г. О множестве образов k -кратной итерации равновероятного случайного отображения // Математические вопросы криптографии. 2018. Т. 9. № 3. С. 99–108.
7. Миронкин В. О. Исследование свойств и характеристик степени случайного отображения // Обзорение прикладной и промышленной математики. 2014. Т. 21. № 1. С. 70–73.
8. Миронкин В. О. Вероятностные характеристики слоев в графе случайного отображения // Обзорение прикладной и промышленной математики. 2015. Т. 22. № 1. С. 80–82.

9. *Миронкин В. О.* Совместная вероятность длин отрезков аperiodичности двух вершин в графе степени случайного отображения // Обозрение прикладной и промышленной математики. 2015. Т. 22. № 4. С. 482–484.
10. *Миронкин В. О.* Об особенностях строения графа степени случайного отображения // Обозрение прикладной и промышленной математики. 2016. Т. 23. № 1. С. 57–62.
11. *Oechslin P.* Making a faster cryptanalytic time-memory trade-off // LNCS. 2003. V. 2729. P. 617–630.
12. *Pilshchikov D. V.* Estimation of the characteristics of time-memory-data tradeoff methods via generating functions of the number of particles and the total number of particles in the Galton-Watson process // Математические вопросы криптографии. 2014. Т. 5. № 2. С. 103–108.
13. *Pilshchikov D. V.* On the limiting mean values in probabilistic models of time-memory-data tradeoff methods // Математические вопросы криптографии. 2015. Т. 6. № 2. С. 59–65.
14. *Миронкин В. О.* О некоторых вероятностных характеристиках алгоритма выработки ключа «CRYPTOPRO KEY MESHING» // Проблемы информационной безопасности. Компьютерные системы. 2015. № 4. С. 140–146.
15. *Токарева Н. Н.* Симметричная криптография. Краткий курс: учебное пособие. Новосибирск: Новосиб. гос. ун-т, 2012.
16. *Сачков В. Н.* Вероятностные методы в комбинаторном анализе. М.: Наука, 1978.
17. *Зубков А. М., Серов А. А.* Совокупность образов подмножества конечного множества при итерациях случайных отображений // Дискретная математика. 2014. Т. 26. № 4. С. 43–50.
18. *Пильщикова Д. В.* Асимптотическое поведение мощности полного прообраза случайного множества при итерациях отображений конечного множества // Математические вопросы криптографии. 2017. Т. 8. № 1. С. 95–106.

REFERENCES

1. *Kolchin V. F.* Sluchaynie otobrazeniya [Random Mappings]. Moscow, Nauka Publ., 1984. (in Russian)
2. *Harris B.* Probability distributions related to random mapping. Ann. Math. Statist., 1960, vol. 31, no. 4, pp. 1045–1062.
3. *Dalal A. and Schmutz E.* Compositions of random functions on a finite set. Electr. J. Comb., 2002, vol. 9, no. R26.
4. *Flajolet P. and Odlyzko A.* Random mapping statistics. LNCS, 1989, vol. 434, pp. 329–354.
5. *Zubkov A. M. and Mironkin V. O.* Raspredelenie dlini otrezka aperiodichnosti v grafe k -kratnoy iteracii sluchaynogo ravnoveryatnogo otobrazeniya [Distribution of the length of aperiodicity segment in the graph of k -fold iteration of uniform random mapping]. Mat. Vopr. Kriptogr., 2017, vol. 8, no. 4, pp. 63–74. (in Russian)
6. *Mironkin V. O. and Mikhailov V. G.* O mnozestve obrazov k -kratnoy iteracii ravnoveryatnogo sluchaynogo otobrazeniya [On the sets of images of k -fold iteration of uniform random mapping]. Mat. Vopr. Kriptogr., 2018, vol. 9, no. 3, pp. 99–108. (in Russian)
7. *Mironkin V. O.* Issledovanie svoystv i harakteristik stepeni sluchaynogo otobrazeniya [Investigation of properties and characteristics of iteration of random mapping]. Obozrenie Prikladnoj i Promyshlennoj Matematiki, 2014, vol. 21, no. 1, pp. 70–73. (in Russian)
8. *Mironkin V. O.* Veroyatnostnie harakteristiki slojov v grafe sluchaynogo otobrazeniya [Probabilistic characteristics of layers in a random mapping draph]. Obozrenie Prikladnoj i Promyshlennoj Matematiki, 2015, vol. 22, no. 1, pp. 80–82. (in Russian)
9. *Mironkin V. O.* Sovmestnaya veroyatnost dlin otrezkov aperiodichnosti dvuh vershin v grafe stepeni sluchaynogo otobrazeniya [The joint probability of lengths of aperiodicity segments

- of two vertices in the graph of iteration of random mapping]. *Obozrenie Prikladnoj i Promyshlennoj Matematiki*, 2015, vol. 22, no. 4, pp. 482–484. (in Russian)
10. *Mironkin V. O.* Ob osobennostiah stroeniya grafa stepeni sluchaynogo otobrazeniya [On singularities of the structure of the graph of iteration of random mapping]. *Obozrenie Prikladnoj i Promyshlennoj Matematiki*, 2016, vol. 23, no. 1, pp. 57–62. (in Russian)
 11. *Oechskin P.* Making a faster cryptanalytic time-memory trade-off. *LNCS*, 2003, vol. 2729, pp. 617–630.
 12. *Pilshchikov D. V.* Estimation of the characteristics of time-memory-data tradeoff methods via generating functions of the number of particles and the total number of particles in the Galton-Watson process. *Mat. Vopr. Kriptogr.*, 2014, vol. 5, no. 2, pp. 103–108.
 13. *Pilshchikov D. V.* On the limiting mean values in probabilistic models of time-memory-data tradeoff methods. *Mat. Vopr. Kriptogr.*, 2015, vol. 6, no. 2, pp. 59–65.
 14. *Mironkin V. O.* O nekotoryh veroyatnostnih harakteristikah algoritma virabotki klucha “CRYPTOPRO KEY MESHING” [On some probabilistic characteristics of key derivation function “CRYPTOPRO KEY MESHING”]. *Problemy Informacionnoj Bezopasnosti. Komp’yuternye Sistemy*, 2015, no. 4, pp. 140–146. (in Russian)
 15. *Tokareva N. N.* Simmetrichnaya kriptografiya. Kratkiy kurs: uchebnoe posobie [Symmetric Cryptography. A Short Course: a Tutorial]. Novosibirsk, NSU Publ., 2012. (in Russian)
 16. *Sachkov V. N.* Veroyatnostnie metodi kombinatornogo analiza [Probabilistic Methods in Combinatorial Analysis]. Moscow, Nauka Publ., 1978. (in Russian)
 17. *Zubkov A. M. and Serov A. A.* Sovokupnost obrazov podmnozestva konechnogo mnozestva pri iteracijah sluchaynih otobrazeni [Images of subset of finite set under iterations of random mappings]. *Discr. Math.*, 2014, vol. 26, no. 4, pp. 43–50. (in Russian)
 18. *Pilshchikov D. V.* Asimptoticheskoe povedenie mochnosti polnogo proobraza sluchaynogo mnozestva pri iteracijah otobrazeniy konechnogo mnozestva [Asymptotic behaviour of the complete preimage cardinality for the image of a random set under iterations of mappings of a finite set]. *Mat. Vopr. Kriptogr.*, 2017, vol. 8, no. 1, pp. 95–106. (in Russian)

УДК 519.714.5

ОДИН ПОДХОД К ПОСТРОЕНИЮ КРАТНО ТРАНЗИТИВНОГО МНОЖЕСТВА БЛОЧНЫХ ПРЕОБРАЗОВАНИЙ

И. В. Чередник

Российский технологический университет (МИРЭА), г. Москва, Россия

Продолжается исследование множества преобразований $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$, реализуемых сетью Σ с одной бинарной квазигрупповой операцией F . В случае произвольного $k \geq 2$ определяются условия k -транзитивности этого множества и предлагается эффективный способ проверки этих условий. Приводится алгоритм построения таких сетей Σ , у которых множество преобразований $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ является k -транзитивным.

Ключевые слова: *сети, квазигруппы, блочные преобразования, k -транзитивное множество блочных преобразований.*

DOI 10.17223/20710410/42/2

ONE APPROACH TO CONSTRUCTING A MULTIPLY TRANSITIVE CLASS OF BLOCK TRANSFORMATIONS

I. V. Cherednik

Russian Technological University (MIREA), Moscow, Russia

E-mail: p.n.v.k.s@mail.ru

In this work, we continue to study the cryptographic properties of block transformations of a new type, which can be used to construct hash functions and block ciphers. Let Ω be an arbitrary finite set, $\mathcal{Q}(\Omega)$ be the collection of all binary quasigroups defined on the set Ω , and $\Sigma^F : \Omega^n \rightarrow \Omega^n$ be a mapping that is implemented by a network Σ of width n with one binary operation $F \in \mathcal{Q}(\Omega)$. The network Σ is called *bijective* if the mapping Σ^F is bijective for each $F \in \mathcal{Q}(\Omega)$ and all finite sets Ω . The networks Σ_1, Σ_2 are called *equivalent* if the map Σ_1^F of Σ_1 coincides with the map Σ_2^F of Σ_2 for each $F \in \mathcal{Q}(\Omega)$ and for all finite sets Ω . It is not difficult to define the elementary networks by analogy with the elementary matrices and prove that every bijective network Σ is equivalent to a unique product of elementary networks. This product is called the *canonical representation* of Σ and its length is denoted by $\|\Sigma\|$. A bijective network Σ is called *k -transitive for Ω* if the family $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ is k -transitive. We prove that the bijective network Σ is k -transitive for all sufficiently large finite sets iff Σ is k -transitive for some finite set Ω such that $|\Omega| \geq k\|\Sigma\| + kn$. In addition, we propose an effective method for verifying the network's k -transitivity for all sufficiently large finite sets, namely, the bijective network Σ is k -transitive for Ω such that $|\Omega| \geq k\|\Sigma\| + kn$ whenever it is k -transitive for some $(k + 1)$ -element subset of Ω . Also, we describe an algorithm for constructing k -transitive networks. For a given bijective network Σ of a width n , the algorithm adds $6n - 7$ elementary networks to the canonical representation of Σ without changing the existing contents. As a result of these modifications, we obtain a bijective network $\widehat{\Sigma}$ that is k -transitive for every sufficiently large finite set Ω , namely for $|\Omega| \geq k\|\widehat{\Sigma}\| + kn$.

Keywords: *network, quasigroup, block transformation, k-transitive class of block transformations.*

Введение

В работе продолжается исследование множества преобразований $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$, реализуемых сетью Σ с одной бинарной квазигрупповой операцией F , начатое в [1]. Напомним основные определения и необходимые результаты из [1].

Произвольная бинарная операция $F: \Omega \times \Omega \rightarrow \Omega$ называется квазигруппой на множестве Ω , если уравнения вида

$$F(x, b) = c, \quad F(a, y) = c$$

однозначно разрешимы при любых $a, b, c \in \Omega$ [2]. Множество всех квазигрупп, заданных на множестве Ω , будем обозначать $\mathcal{Q}(\Omega)$.

Пусть $\{x_1, \dots, x_n\}$ — множество переменных и $*$ — символ бинарной операции. Множество всех формул в алфавите $\{x_1, \dots, x_n, *\}$ будем обозначать \mathcal{W} . При сопоставлении символу $*$ конкретной бинарной квазигруппы $F \in \mathcal{Q}(\Omega)$ формула $w(x_1, \dots, x_n)$ реализует отображение $w^F: \Omega^n \rightarrow \Omega$, а набор формул $(w_1, \dots, w_m) \in \mathcal{W}^m$ — отображение $(w_1^F, \dots, w_m^F): \Omega^n \rightarrow \Omega^m$.

Определение 1. Пусть $(v_1, \dots, v_k) \in \mathcal{W}^k$ и в наборе $(w_1, \dots, w_m) \in \mathcal{W}^m$ каждая из формул $w_j, j \in \{1, \dots, m\}$, либо имеет вид $v_{i_1} * v_{i_2}, i_1 \neq i_2, i_1, i_2 \in \{1, \dots, k\}$, либо является некоторой формулой $v_i, i \in \{1, \dots, k\}$. Тогда будем говорить, что набор формул (w_1, \dots, w_m) является *результатом преобразования* набора формул (v_1, \dots, v_k) .

Один из способов построения произвольного набора формул (w_1, \dots, w_m) заключается в последовательном преобразовании набора переменных (x_1, \dots, x_n) . Для исследования свойств отображений одного класса, соответствующего определённому набору формул, введём дополнительное представление процесса преобразований набора формул, которое отличается большей наглядностью.

Определение 2. Пусть $t, n_0, n_1, \dots, n_t \in \mathbb{N}$ и

$$X_0 = \{x_1^{(0)}, x_2^{(0)}, \dots, x_{n_0}^{(0)}\}, X_1 = \{x_1^{(1)}, x_2^{(1)}, \dots, x_{n_1}^{(1)}\}, \dots, X_t = \{x_1^{(t)}, x_2^{(t)}, \dots, x_{n_t}^{(t)}\}$$

— семейство попарно непересекающихся конечных непустых множеств. Тогда *квазигрупповой сетью* (далее — просто *сетью*) *длины* t будем называть простой ориентированный граф Σ с множеством вершин $X_0 \cup X_1 \cup \dots \cup X_t$, содержащий только рёбра вида $(x_i^{(s-1)}, x_j^{(s)})$, $s \in \{1, \dots, t\}$, с тем ограничением, что степень захода каждой вершины $x_j^{(s)}$, $s \in \{1, \dots, t\}$, равна 1 или 2. При этом если степень захода вершины $x_j^{(s)}$ равна 2, то рёбра $(x_{i_1}^{(s-1)}, x_j^{(s)})$ и $(x_{i_2}^{(s-1)}, x_j^{(s)})$ имеют различные метки из множества $\{l, r\}$. Число $\max\{n_0, \dots, n_t\}$ будем называть *шириной* сети Σ . Множества X_0 и X_t называются множествами начальных и конечных вершин соответственно. Подграф Σ_s сети Σ , основанный на множестве вершин $X_{s-1} \cup X_s$, будем называть s -м *слоем* сети Σ . Сеть Σ называется *однослойной*, если она имеет длину 1.

Определение 3. Пусть Σ и Σ' — сети с множествами вершин $X = X_0 \cup X_1 \cup \dots \cup X_s$ и $X' = X'_0 \cup X'_1 \cup \dots \cup X'_t$ соответственно и $X \cap X' = X_s = X'_0$. Тогда естественным образом можно определить сеть длины $s + t$ с множеством вершин $X_0 \cup X_1 \cup \dots \cup X_s \cup X'_1 \cup \dots \cup X'_t$, которую будем называть *произведением* сетей Σ и Σ' и обозначать $\Sigma \cdot \Sigma'$.

Непосредственно из определений 2 и 3 следует, что произвольная сеть Σ длины t является произведением однослойных сетей: $\Sigma = \Sigma_1 \cdot \dots \cdot \Sigma_t$.

Определение 4. Пусть (v_1, \dots, v_n) — произвольный набор формул и Σ — однослойная сеть с множеством вершин $\{x_1^0, \dots, x_n^0\} \cup \{x_1^1, \dots, x_m^1\}$. Тогда определим набор формул (w_1, \dots, w_m) по следующим правилам:

- если вершине x_j^1 инцидентно единственное ребро (x_i^0, x_j^1) , то полагаем $w_j = v_i$;
- если вершине x_j^1 инцидентны рёбра $(x_{i_1}^0, x_j^1)$ и $(x_{i_2}^0, x_j^1)$ с метками l и r соответственно, то полагаем $w_j = v_{i_1} * v_{i_2}$.

При этом будем говорить, что однослойная сеть Σ описывает преобразование набора формул (v_1, \dots, v_n) в набор формул (w_1, \dots, w_m) . Произвольная сеть Σ является произведением однослойных сетей, являющихся её слоями, и потому естественным образом описывает последовательность преобразований набора формул.

Пусть $F \in \mathcal{Q}(\Omega)$ — произвольная квазигруппа и сеть Σ описывает последовательность преобразований набора переменных (x_1, \dots, x_n) в набор формул (w_1, \dots, w_m) . Тогда отображение $(w_1^F, \dots, w_m^F): \Omega^n \rightarrow \Omega^m$ будем обозначать Σ^F .

Нетрудно понять, что если $\Sigma = \Sigma_1 \cdot \Sigma_2$, то при выборе любой квазигруппы F справедливо соответствующее равенство отображений $\Sigma^F = \Sigma_1^F \cdot \Sigma_2^F$.

Определение 5. Будем говорить, что сети Σ и Σ' эквивалентны для множества Ω , если при выборе любой квазигруппы $F \in \mathcal{Q}(\Omega)$ отображения Σ^F и Σ'^F совпадают. Будем говорить, что сети Σ и Σ' эквивалентны, если они эквивалентны для любого множества.

Замечание 1. Если сети Σ и Σ' описывают преобразование набора переменных (x_1, \dots, x_n) в наборы формул (w_1, \dots, w_m) и (w'_1, \dots, w'_m) соответственно, то совпадение указанных наборов формул является достаточным условием для эквивалентности сетей Σ и Σ' . Более того, верно и обратное утверждение (теорема 7 в [1]).

Определение 6. Сеть Σ будем называть *биективной для множества Ω* , если при выборе любой квазигруппы $F \in \mathcal{Q}(\Omega)$ отображение Σ^F является биективным. Сеть Σ будем называть *биективной*, если она биективна для любого множества.

Очевидно, что для биективности сети Σ с множеством вершин $X_0 \cup X_1 \cup \dots \cup X_t$ необходимо, чтобы множества начальных и конечных вершин были равноможны, то есть выполнялось равенство $|X_0| = |X_t|$.

Определение 7. Сеть Σ с множеством вершин $X_0 \cup X_1 \cup \dots \cup X_t$ будем называть *сетью постоянной ширины*, если $|X_0| = |X_1| = \dots = |X_t|$.

В данной работе рассматриваются только сети постоянной ширины, поэтому будем использовать термин «*сеть*», подразумевая при этом *сеть постоянной ширины*.

Определение 8. Пусть Σ — однослойная сеть с множеством вершин $X_0 \cup X_1$. Вершину $x_i^{(0)} \in X_0$ сети Σ будем называть *неподвижной*, если Σ содержит ребро $(x_i^{(0)}, x_i^{(1)})$. Сеть Σ будем называть *элементарной*, если все вершины из множества X_0 неподвижны и ровно одна вершина из множества X_1 имеет степень захода 2.

Элементарную сеть с множеством вершин $X_0 \cup X_1$, которая содержит рёбра $(x_i^{(0)}, x_i^{(1)})$ и $(x_j^{(0)}, x_i^{(1)})$, будем обозначать $\Sigma_i^{\{i,j\}}$. В случае, когда ребро $(x_i^{(0)}, x_i^{(1)})$ имеет метку l , обозначение можно уточнить как $\Sigma_i^{(i,j,l)}$, а если оно имеет метку r — как $\Sigma_i^{(j,i,r)}$.

Произвольная элементарная сеть всегда является биективной. Ещё одним важным примером биективных сетей являются сети с множеством вершин $X_0 \cup X_1 \cup \dots \cup X_t$, у которых степень захода каждой вершины $x_j^{(s)}$, $s \in \{1, \dots, t\}$, равна 1. Такие сети

будем называть *перестановочными*. Произвольная перестановочная сеть определяет отображение $\Omega^n \rightarrow \Omega^n$, не зависящее от выбора квазигруппы $F \in \mathcal{Q}(\Omega)$ и действующее на множестве Ω^n как перестановка координат вектора. Отсюда следует, что любая перестановочная сеть эквивалентна однослойной перестановочной сети.

Элементарные и перестановочные сети являются примерами простейших биективных сетей, однако этих примитивов достаточно для реализации произвольной биективной сети.

Теорема 1 (следствие 7 в [1]). Сеть Σ является биективной в том и только в том случае, когда она эквивалентна произведению

$$\Sigma_{R_1} \cdot \dots \cdot \Sigma_{R_t} \cdot \Pi_R \text{ (или } \Pi_L \cdot \Sigma_{L_1} \cdot \dots \cdot \Sigma_{L_t}),$$

где $\Sigma_{R_1}, \dots, \Sigma_{R_t}$ ($\Sigma_{L_1}, \dots, \Sigma_{L_t}$) — элементарные сети, а Π_R (Π_L) — однослойная перестановочная сеть. При этом произведение определено однозначно с точностью до возможной перестановки элементарных сетей, а количество элементарных сетей в произведении равно количеству вершин сети Σ со степенью захода 2.

Количество вершин сети Σ со степенью захода 2 будем называть *весом сети* Σ или её *сложностью* и обозначать $\|\Sigma\|$.

Учитывая теорему 1, не ограничивая общности, можно считать, что произвольная биективная сеть Σ представляет собой произведение $\Sigma_1 \cdot \dots \cdot \Sigma_t \cdot \Pi$ с множеством вершин $X_0 \cup X_1 \cup \dots \cup X_t \cup X_{t+1}$, где $\Sigma_1, \dots, \Sigma_t$ — элементарные сети; Π — однослойная перестановочная сеть. Также, не ограничивая общности, можно считать, что $\Omega \subset \mathbb{N}$.

Определение 9. Если для элементов $y_1, y_2, y_3 \in \mathbb{N}$ и частично определённого отображения $F: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ выполняется соотношение $F(y_1, y_2) = y_3$, то будем говорить, что элементы y_1 и y_2 *содержатся в области определения отображения F* , а элемент y_3 — *в области значений отображения F* .

Определение 10. Частично определённое отображение $F: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$, удовлетворяющее условию

$$(F(y_1, y_2) = F(y'_1, y'_2)) \implies ((y_1, y_2) = (y'_1, y'_2) \text{ или } (y_1 \neq y'_1, y_2 \neq y'_2))$$

при всех допустимых $y_1, y_2, y'_1, y'_2 \in \mathbb{N}$, будем называть *частично определённым отображением без противоречий* (или *частично определённым непротиворечивым отображением*).

Определение 11. Разметкой сети $\Sigma = \Sigma_1 \cdot \dots \cdot \Sigma_t \cdot \Pi$ будем называть произвольное отображение $\mu: X_0 \cup X_1 \cup \dots \cup X_t \cup X_{t+1} \rightarrow \mathbb{N}$. Пусть $F: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ — частично определённое отображение. Тогда разметку μ сети Σ , которая удовлетворяет следующим условиям:

- для всех $s \in \{1, \dots, t\}$ и $i \in \{1, \dots, n\}$:
 - если $\deg^- x_i^{(s)} = 1$, то $\mu(x_i^{(s)}) = \mu(x_i^{(s-1)})$;
 - если $\deg^- x_i^{(s)} = 2$ и рёбра $(x_i^{(s-1)}, x_i^{(s)})$, $(x_j^{(s-1)}, x_i^{(s)})$ имеют метки l и r соответственно, то $\mu(x_i^{(s)}) = F(\mu(x_i^{(s-1)}), \mu(x_j^{(s-1)}))$;
 - если $\deg^- x_i^{(s)} = 2$ и рёбра $(x_i^{(s-1)}, x_i^{(s)})$, $(x_j^{(s-1)}, x_i^{(s)})$ имеют метки r и l соответственно, то $\mu(x_i^{(s)}) = F(\mu(x_j^{(s-1)}), \mu(x_i^{(s-1)}))$;
- если перестановочная сеть Π содержит рёбра $(x_{i_k}^{(t)}, x_k^{(t+1)})$, $k \in \{1, \dots, n\}$, то выполняются равенства $\mu(x_k^{(t+1)}) = \mu(x_{i_k}^{(t)})$, $k \in \{1, \dots, n\}$,

будем называть *правильной относительно F* . При этом само отображение F будем называть *правилом разметки μ* .

Определение 12. Пусть μ — разметка сети Σ с правилом F и при этом никакое сужение частичного отображения F не является правилом разметки μ . Тогда будем говорить, что F является *минимальным правилом* разметки μ . Нетрудно понять, что минимальное правило разметки μ определено однозначно. Правильную разметку μ будем называть *непротиворечивой*, если её минимальное правило является непротиворечивым отображением.

Определение 13. Если для разметки μ сети Σ выполняется система равенств $\{\mu(x_{i_j}^{(s_j)}) = v_j : j \in J\}$, то будем говорить, что μ — *разметка сети Σ с условиями* $\{\mu(x_{i_j}^{(s_j)}) = v_j : j \in J\}$. Система равенств $\mu(x_1^{(0)}) = v_1, \dots, \mu(x_n^{(0)}) = v_n$ называется *начальным условием* разметки μ , при этом говорят, что μ — *разметка с начальным условием* (v_1, \dots, v_n) .

Каждая правильная разметка сети Σ однозначно определяется своим начальным условием и правилом. В тех случаях, когда при некоторой разметке вершин $x_1^{(0)}, \dots, x_n^{(0)}$ сети Σ для полного задания правильной разметки не хватает области определения частично определённого отображения F , можно непротиворечивым образом расширить область определения F и тем самым определить разметку с правилом F . Поясним это подробнее.

Пусть задана начальная разметка $\mu(x_1^{(0)}) = v_1, \dots, \mu(x_n^{(0)}) = v_n$ сети Σ и $\mathbb{N} \setminus \{v_1, \dots, v_n\} \supset \{y_1, y_2, \dots\}$ — счётное множество меток, которые не содержатся ни в области определения, ни в области значений правила F (при этом возможно, что метки v_1, \dots, v_n также не содержатся ни в области определения, ни в области значений правила F). Тогда продолжим разметку μ сети Σ по следующему правилу:

- для всех $s \in \{1, \dots, t\}$ при $\Sigma_s = \Sigma_i^{\{i,j\}}$ положим $\mu(x_l^{(s)}) = \mu(x_l^{(s-1)})$, если $l \neq i$, а для разметки вершины $x_i^{(s)}$ возможны следующие варианты:
 - если $\Sigma_s = \Sigma_i^{\{i,j\}}$ и значение $F(\mu(x_i^{(s-1)}), \mu(x_j^{(s-1)}))$ определено, то пометим вершину $x_i^{(s)}$ меткой $F(\mu(x_i^{(s-1)}), \mu(x_j^{(s-1)}))$, иначе пометим вершину $x_i^{(s)}$ ранее не использованной меткой y_s и определим $F(\mu(x_i^{(s-1)}), \mu(x_j^{(s-1)})) = y_s$;
 - если $\Sigma_s = \Sigma_i^{\{j,i\}}$ и значение $F(\mu(x_j^{(s-1)}), \mu(x_i^{(s-1)}))$ определено, то пометим вершину $x_i^{(s)}$ меткой $F(\mu(x_j^{(s-1)}), \mu(x_i^{(s-1)}))$, иначе пометим вершину $x_i^{(s)}$ ранее не использованной меткой y_s и определим $F(\mu(x_j^{(s-1)}), \mu(x_i^{(s-1)})) = y_s$;
- если перестановочная сеть Π содержит рёбра $(x_{i_k}^{(t)}, x_k^{(t+1)})$, $k \in \{1, \dots, n\}$, то положим $\mu(x_k^{(t+1)}) = \mu(x_{i_k}^{(t)})$, $k \in \{1, \dots, n\}$.

При проведении разметки μ сети Σ описанным способом частично определённое (непротиворечивое) отображение F корректным образом продолжается до частично определённого (непротиворечивого) отображения, которое будем обозначать $F_{\Sigma, \mu}$, при этом построенная разметка μ является правильной относительно $F_{\Sigma, \mu}$.

Определение 14. Описанную процедуру продолжения разметки μ и расширения области определения F будем называть *свободным продолжением начальной разметки $\mu(x_1^{(0)}) = v_1, \dots, \mu(x_n^{(0)}) = v_n$ и её правила F относительно сети Σ* .

Пусть η и μ — разметки сети Σ и для отображения $\sigma_\mu: \mathbb{N} \rightarrow \mathbb{N}$ справедливо соотношение $\sigma_\mu \circ \eta = \mu$, то есть при всех $s \in \{0, \dots, t+1\}$ и $i \in \{1, \dots, n\}$ выполняется равенство $\sigma_\mu(\eta(x_i^{(s)})) = \mu(x_i^{(s)})$. Тогда будем обозначать это условие как $\sigma_\mu: \eta \rightarrow \mu$.

Определение 15. Правильную разметку η сети Σ с начальным условием (v_1, \dots, v_n) будем называть *свободной*, если для любой правильной разметки μ сети Σ с начальным условием (v_1, \dots, v_n) существует отображение σ_μ , удовлетворяющее условию $\sigma_\mu: \eta \rightarrow \mu$.

Непосредственно из определения свободной разметки следует, что при условии существования свободной разметки сети Σ с начальным условием (v_1, \dots, v_n) определена однозначно с точностью до обратимого переобозначения меток.

Теорема 2 (теорема 3 в [1]). Пусть разметка η получена в результате свободно-го продолжения начальной разметки $\eta(x_1^{(0)}) = v_1, \dots, \eta(x_n^{(0)}) = v_n$ и пустого правила G относительно сети Σ . Тогда η — свободная разметка сети Σ с начальным условием (v_1, \dots, v_n) , а отображение $G_{\Sigma, \eta}$ — её минимальное правило.

Определение 16. Биективную сеть Σ будем называть *транзитивной для множества Ω* , если множество отображений $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ является транзитивным.

Нетрудно понять, что сама природа множества Ω в данном определении не играет никакой роли, поэтому корректно говорить, что биективная сеть Σ является транзитивной для множеств мощности q . По-прежнему будем считать, что $\Omega \subset \mathbb{N}$, а для множества $\{1, \dots, q\}$ будем использовать обозначение Ω_q .

Определение 17. Разметку μ сети Σ с условиями

$$\mu(x_1^{(0)}) = v_1, \dots, \mu(x_n^{(0)}) = v_n, \mu(x_1^{(t)}) = w_1, \dots, \mu(x_n^{(t)}) = w_n$$

будем называть *разметкой сети Σ с ограничениями* $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$. При этом будем говорить, что сеть Σ *допускает* разметку μ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$.

Теорема 3 (теорема 11 в [1]). Сеть Σ допускает правильные непротиворечивые разметки при всех возможных ограничениях $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ из \mathbb{N} в том и только в том случае, когда сеть Σ допускает правильные непротиворечивые разметки при всех возможных ограничениях $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$ из Ω_2 .

Следствие 1 (следствие 12 в [1]). Пусть Σ — биективная сеть ширины n и Ω — множество мощности не менее чем $\|\Sigma\| + n$. Тогда следующие утверждения эквивалентны:

- 1) сеть Σ является транзитивной для множества Ω ;
- 2) сеть Σ допускает правильную непротиворечивую разметку элементами множества Ω при любых ограничениях $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ из множества Ω ;
- 3) сеть Σ допускает правильную непротиворечивую разметку элементами множества Ω при любых ограничениях $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$ из множества $\Omega_2 \subset \Omega$;
- 4) множество преобразований $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ действует транзитивным образом на подмножестве $\Omega_2^n \subset \Omega^n$.

1. 2-Разметка биективных сетей

Определение 18. Биективную сеть Σ будем называть *2-транзитивной для множества Ω* , если множество отображений $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ является 2-транзитивным.

Нетрудно понять, что сама природа конечного множества Ω в данном определении не играет никакой роли, поэтому корректно говорить, что биективная сеть Σ является 2-транзитивной для множеств мощности q . Не ограничивая общности, будем считать, что $\Omega \subset \mathbb{N}$. Очевидно

Утверждение 1. Пусть Σ — произвольная 2-транзитивная для множества Ω сеть, а Π_1, Π_2 — произвольные перестановочные сети, для которых корректно определить произведения $\Pi_1 \cdot \Sigma$ и $\Sigma \cdot \Pi_2$. Тогда сети $\Pi_1 \cdot \Sigma$ и $\Sigma \cdot \Pi_2$ также 2-транзитивны для множества Ω .

Не ограничивая общности, будем считать, что произвольная биективная сеть Σ представляет собой произведение элементарных сетей $\Sigma_1 \cdot \dots \cdot \Sigma_t$ с множеством вершин $X_0 \cup X_1 \cup \dots \cup X_t$.

Введенный в [1] аппарат разметки сетей на самом деле позволяет проверять не только транзитивность сети, но и более сложное свойство k -транзитивности при $k \geq 2$. Так, например, из 2-транзитивности биективной сети Σ для множества Ω следует, что для любых $(v_{11}, \dots, v_{1n}) \neq (v_{21}, \dots, v_{2n}), (w_{11}, \dots, w_{1n}) \neq (w_{21}, \dots, w_{2n}) \in \Omega^n$ сеть Σ допускает пару правильных разметок с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ и общим непротиворечивым правилом.

Определение 19. Произвольную пару $\mu = (\mu_1, \mu_2)$ разметок сети Σ будем называть 2-разметкой сети Σ . При этом метки разметок μ_1 и μ_2 будем называть метками 2-разметки μ . Пусть $F: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ — частично определённое отображение. Тогда 2-разметку $\mu = (\mu_1, \mu_2)$ сети Σ будем называть *правильной относительно F* , если каждая из разметок μ_1 и μ_2 является правильной относительно F , а отображение F будем называть *правилом 2-разметки μ* .

Определение 20. Пусть μ — 2-разметка сети Σ с правилом F и никакое сужение частичного отображения F не является правилом 2-разметки μ . Тогда будем говорить, что F является *минимальным правилом 2-разметки μ* . Правильную 2-разметку μ будем называть *непротиворечивой*, если её минимальное правило является непротиворечивым отображением.

Определение 21. Если для 2-разметки μ сети Σ выполняются системы равенств $\{\mu_1(x_{i_j}^{(s_j)}) = v_j : j \in J_1\}$ и $\{\mu_2(x_{i_j}^{(s_j)}) = v_j : j \in J_2\}$, то будем говорить, что μ является 2-разметкой сети Σ с условиями $\{\mu_1(x_{i_j}^{(s_j)}) = v_j : j \in J_1\}$ и $\{\mu_2(x_{i_j}^{(s_j)}) = v_j : j \in J_2\}$. Системы равенств $\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}$ и $\mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}$ называются *начальным условием 2-разметки μ* , при этом говорят, что μ — 2-разметка с начальным условием (v_{11}, \dots, v_{1n}) и (v_{21}, \dots, v_{2n}) .

В дальнейшем систему равенств $\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}$ и $\mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}$ также будем называть *начальной разметкой сети Σ* (несмотря на то, что словосочетание «начальная разметка» не определено), полагая при этом, что задана 2-разметка $\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}$ и $\mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}$ начальных вершин сети Σ .

Каждая правильная 2-разметка сети Σ однозначно определяется своим начальным условием. В тех случаях, когда при некоторой 2-разметке вершин $x_1^{(0)}, \dots, x_n^{(0)}$ сети Σ

для полного задания правильной 2-разметки не хватает области определения имеющегося правила, можно предложить два способа продолжения начальной разметки и непротиворечивого расширения области определения правила. Пусть задана начальная 2-разметка $\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}$ и $\mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}$ сети Σ , а также $\mathbb{N} \setminus \{v_{11}, v_{21}, \dots, v_{1n}, v_{2n}\} \supset \{y_{11}, y_{21}, y_{12}, y_{22}, \dots\}$ — счётное множество меток, которые не содержатся ни в области определения, ни в области значений правила F (при этом возможно, что метки $v_{11}, v_{21}, \dots, v_{1n}, v_{2n}$ также не содержатся ни в области определения, ни в области значений правила F).

Последовательное свободное продолжение разметки. С использованием множества меток $\{y_{11}, y_{12}, \dots\}$ проведём свободное продолжение начальной разметки $\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}$ и её правила F относительно сети Σ . Затем, используя множество меток $\{y_{21}, y_{22}, \dots\}$, проведём свободное продолжение начальной разметки $\mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}$ и её правила F_{Σ, μ_1} относительно сети Σ . В результате получим 2-разметку $\mu = (\mu_1, \mu_2)$ сети Σ с начальным условием (v_{11}, \dots, v_{1n}) и (v_{21}, \dots, v_{2n}) .

Нетрудно понять, что при проведении 2-разметки μ сети Σ описанным способом частично определённое (непротиворечивое) отображение F корректным образом продолжается до частично определённого (непротиворечивого) отображения $(F_{\Sigma, \mu})_{\Sigma, \mu_2}$, которое будем обозначать $F'_{\Sigma, \mu}$, а построенная 2-разметка μ является правильной относительно $F'_{\Sigma, \mu}$.

Описанную процедуру продолжения 2-разметки μ и расширения области определения F будем называть *последовательным свободным продолжением начальной 2-разметки* $\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}$ и $\mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}$ и её правила F относительно сети Σ . При этом будем говорить, что 2-разметка μ получена в результате последовательного свободного продолжения начальной 2-разметки $\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}$ и $\mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}$ и её правила F относительно сети Σ .

Параллельное свободное продолжение разметки. Продолжим начальную 2-разметку μ сети Σ , пользуясь следующими правилами при всех $s \in \{1, \dots, t\}$:

- если $\Sigma_s = \Sigma_i^{\{i_1, i_2\}}$, то при $l \neq i$ положим $\mu_1(x_l^{(s)}) = \mu_1(x_l^{(s-1)})$ и $\mu_2(x_l^{(s)}) = \mu_2(x_l^{(s-1)})$;
 - если $\Sigma_s = \Sigma_i^{\{i_1, i_2\}}$ и значение $F(\mu_1(x_{i_1}^{(s-1)}), \mu_1(x_{i_2}^{(s-1)}))$ определено, то положим $\mu_1(x_i^{(s)}) = F(\mu_1(x_{i_1}^{(s-1)}), \mu_1(x_{i_2}^{(s-1)}))$, в противном случае положим $\mu_1(x_i^{(s)}) = y_{1s}$ и доопределим $F(\mu_1(x_{i_1}^{(s-1)}), \mu_1(x_{i_2}^{(s-1)})) = y_{1s}$;
- после этого, если значение $F(\mu_2(x_{i_1}^{(s-1)}), \mu_2(x_{i_2}^{(s-1)}))$ определено, то положим $\mu_2(x_i^{(s)}) = F(\mu_2(x_{i_1}^{(s-1)}), \mu_2(x_{i_2}^{(s-1)}))$, в противном случае положим $\mu_2(x_i^{(s)}) = y_{2s}$ и доопределим $F(\mu_2(x_{i_1}^{(s-1)}), \mu_2(x_{i_2}^{(s-1)})) = y_{2s}$.

Нетрудно понять, что при проведении 2-разметки μ сети Σ описанным способом частично определённое (непротиворечивое) отображение F корректным образом продолжается до частично определённого (непротиворечивого) отображения, которое будем обозначать $F''_{\Sigma, \mu}$, а построенная 2-разметка μ является правильной относительно $F''_{\Sigma, \mu}$.

Описанную процедуру продолжения 2-разметки μ и расширения области определения F будем называть *параллельным свободным продолжением начальной 2-разметки* $\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}$ и $\mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}$ и её правила F относительно сети Σ . При этом будем говорить, что 2-разметка μ получена в резуль-

тате параллельного свободного продолжения начальной 2-разметки $\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}$ и $\mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}$ и её правила F относительно сети Σ .

Теорема 4. Пусть 2-разметка μ' получена в результате последовательного свободного продолжения начальной 2-разметки $\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}$ и $\mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}$ и правила F относительно сети Σ , а 2-разметка μ'' — в результате параллельного свободного продолжения начальной 2-разметки $\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}$ и $\mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}$ и правила F относительно сети Σ . Тогда 2-разметки μ' и μ'' отличаются только обратимой заменой меток.

Доказательство. Пусть $y_{2j_1}, \dots, y_{2j_r}$ — все метки вида y_{2*} , которые содержатся в разметке μ'' , и $\mu''(x_{i_1}^{(s_1)}) = y_{2j_1}, \dots, \mu''(x_{i_r}^{(s_r)}) = y_{2j_r}$ — первые появления указанных меток в разметке μ'' . Тогда нетрудно понять, что в 2-разметке μ'' отсутствуют метки $y_{1s_1}, \dots, y_{1s_r}$ и обратимая замена меток $y_{2j_1} \rightarrow y_{1s_1}, \dots, y_{2j_r} \rightarrow y_{1s_r}$ переводит 2-разметку μ'' в 2-разметку μ' . ■

Замечание 2. Вообще говоря, свободное продолжение 2-разметки можно было определить не только последовательным и параллельным способами, но и любым «промежуточным способом», использующим произвольную последовательность продолжения обеих компонент 2-разметки; в результате получилась бы 2-разметка, отличающаяся от последовательного (параллельного) свободного продолжения только обратимым переобозначением меток. Другими словами, главное в свободном продолжении 2-разметки — это не порядок продолжения, а его «свобода» в каждый момент.

Пусть $\eta = (\eta_1, \eta_2)$ и $\mu = (\mu_1, \mu_2)$ — 2-разметки сети Σ и для отображения σ выполняются соотношения $\sigma \circ \eta_1 = \mu_1$ и $\sigma \circ \eta_2 = \mu_2$. Будем обозначать это условие $\sigma: \eta \rightarrow \mu$.

Определение 22. Правильную 2-разметку η сети Σ с начальным условием (v_{11}, \dots, v_{1n}) и (v_{21}, \dots, v_{2n}) будем называть *свободной*, если для любой правильной 2-разметки μ сети Σ с начальным условием (v_{11}, \dots, v_{1n}) и (v_{21}, \dots, v_{2n}) существует отображение σ_μ , удовлетворяющее условию $\sigma_\mu: \eta \rightarrow \mu$.

Непосредственно из определения свободной 2-разметки следует, что при условии существования свободная 2-разметка сети Σ с начальным условием (v_{11}, \dots, v_{1n}) и (v_{21}, \dots, v_{2n}) определена однозначно с точностью до обратимого переобозначения остальных меток.

Теорема 5. Пусть 2-разметка η получена в результате параллельного свободного продолжения начальной 2-разметки $\eta_1(x_1^{(0)}) = v_{11}, \dots, \eta_1(x_n^{(0)}) = v_{1n}$ и $\eta_2(x_1^{(0)}) = v_{21}, \dots, \eta_2(x_n^{(0)}) = v_{2n}$ и пустого правила G относительно сети Σ . Тогда η — свободная 2-разметка сети Σ , а отображение $G_{\Sigma, \eta}$ — её минимальное правило.

Доказательство. Из определения процедуры свободного продолжения начальной 2-разметки $\eta_1(x_1^{(0)}) = v_{11}, \dots, \eta_1(x_n^{(0)}) = v_{1n}$ и $\eta_2(x_1^{(0)}) = v_{21}, \dots, \eta_2(x_n^{(0)}) = v_{2n}$ и её пустого правила G относительно сети Σ следует, что 2-разметка η является непротиворечивой, а её минимальное правило $G_{\Sigma, \eta}$ удовлетворяет условию

$$(G_{\Sigma, \eta}(z_1, z_2) = G_{\Sigma, \eta}(z_3, z_4)) \implies ((z_1, z_2) = (z_3, z_4)) \quad (1)$$

при всех допустимых $z_1, z_2, z_3, z_4 \in \mathbb{N}$.

Пусть μ — произвольная правильная 2-разметка сети Σ с тем же начальным условием, что и 2-разметка η . Тогда для доказательства существования отображения $\sigma_\mu: \eta \rightarrow \mu$ достаточно показать, что при совпадении меток $\eta_{i_1}(x_i^{(s)}) = \eta_{i_2}(x_j^{(r)})$ также выполняется равенство соответствующих меток $\mu_{i_1}(x_i^{(s)}) = \mu_{i_2}(x_j^{(r)})$.

Не ограничивая общности, будем считать, что $\Sigma = \Sigma_1 \cdot \dots \cdot \Sigma_t$. Докажем утверждение индукцией по длине произведения $\Sigma_1 \cdot \dots \cdot \Sigma_t$. База индукции при $t = 1$ очевидна.

Пусть $\Sigma = \Sigma_1 \cdot \dots \cdot \Sigma_{t-1} \cdot \Sigma_t$ — сеть длины $t > 1$ и $\Sigma_t = \Sigma_k^{\{k,l\}}$. Рассмотрим все возможные варианты для пары вершин $x_i^{(s)}$ и $x_j^{(r)}$:

- 1) Если $r, s < t$, то истинность утверждения следует из предположения индукции.
- 2) Если $r < t, s = t$ и $k \neq i$, то выполняется равенство $\eta_{i_1}(x_i^{(s-1)}) = \eta_{i_2}(x_j^{(r)})$ и остаётся воспользоваться предположением индукции.
- 3) Если $r = t, s < t$ и $k \neq j$, то выполняется равенство $\eta_{i_1}(x_i^{(s)}) = \eta_{i_2}(x_j^{(r-1)})$ и остаётся воспользоваться предположением индукции.
- 4) Если $r = s = t$ и $k \notin \{i, j\}$, то выполняется равенство $\eta_{i_1}(x_i^{(s-1)}) = \eta_{i_2}(x_j^{(r-1)})$ и остаётся воспользоваться предположением индукции.
- 5) В случае, когда $s = t$ и $\Sigma_t = \Sigma_i^{\{i,l\}}$, не ограничивая общности, будем считать, что $\Sigma_t = \Sigma_i^{(i,l)}$. Из определения процедуры свободного продолжения разметки следует, что $\eta_{i_1}(x_i^{(s)}) \notin \{v_{11}, v_{21}, \dots, v_{1n}, v_{2n}\}$, а минимальное правило 2-разметки η удовлетворяет условию (1). Значит, равенство меток $\eta_{i_1}(x_i^{(s)}) = \eta_{i_2}(x_j^{(r)})$ влечёт за собой совпадение упорядоченных наборов меток $(\eta_{i_1}(x_i^{(s-1)}), \eta_{i_1}(x_l^{(s-1)}))$ и, не ограничивая общности, $(\eta_{i_2}(x_j^{(r')}), \eta_{i_2}(x_{l'}^{(r')}))$, где $r' \in \{0, \dots, r-1\}$ — наибольшее со свойством $\eta_{i_2}(x_j^{(r')}) \neq \eta_{i_2}(x_j^{(r)})$. По предположению индукции упорядоченные наборы меток $(\mu_{i_1}(x_i^{(s-1)}), \mu_{i_1}(x_l^{(s-1)}))$ и $(\mu_{i_2}(x_j^{(r')}), \mu_{i_2}(x_{l'}^{(r')}))$ также совпадают и, следовательно, выполняются равенства $\mu_{i_1}(x_i^{(s)}) = \mu_{i_2}(x_j^{(r'+1)}) = \mu_{i_2}(x_j^{(r)})$.
- 6) Случай, когда $r = t$ и $\Sigma_t = \Sigma_j^{\{j,l\}}$, доказывается аналогично 5.

Теорема доказана. ■

Замечание 3. Поскольку свободная 2-разметка сети Σ с начальным условием (v_{11}, \dots, v_{1n}) и (v_{21}, \dots, v_{2n}) определена однозначно с точностью до переобозначений, то, не ограничивая общности, можно считать, что произвольная свободная 2-разметка η сети Σ может быть получена при помощи параллельного (последовательного) свободного продолжения начальной 2-разметки $\eta_1(x_1^{(0)}) = v_{11}, \dots, \eta_1(x_n^{(0)}) = v_{1n}$ и $\eta_2(x_1^{(0)}) = v_{21}, \dots, \eta_2(x_n^{(0)}) = v_{2n}$ и её пустого правила G относительно сети Σ .

Теорема 6. Пусть η — свободная 2-разметка сети Σ , μ — правильная 2-разметка сети Σ и возможно определить отображение σ_μ по правилу $\sigma_\mu(\eta_1(x_i^{(0)})) = \mu_1(x_i^{(0)})$, $\sigma_\mu(\eta_2(x_i^{(0)})) = \mu_2(x_i^{(0)})$, $i \in \{1, \dots, n\}$. Тогда отображение σ_μ допускает продолжение, удовлетворяющее условию $\sigma_\mu: \eta \rightarrow \mu$.

Доказательство. Достаточно показать, что при совпадении меток $\eta_{i_1}(x_i^{(s)}) = \eta_{i_2}(x_j^{(r)})$ выполняется равенство соответствующих меток $\mu_{i_1}(x_i^{(s)}) = \mu_{i_2}(x_j^{(r)})$. Это устанавливается индукцией по длине сети Σ аналогично доказательству теоремы 5. ■

Следствие 2. В условиях теоремы 6 если G и F — минимальные правила 2-разметок η и μ соответственно, то при всех допустимых $z_i, z_j \in \mathbb{N}$ выполняется равенство $\sigma_\mu(G(z_i, z_j)) = F(\sigma_\mu(z_i), \sigma_\mu(z_j))$.

В заключение сформулируем и докажем одно простое утверждение, которое необходимо для лучшего понимания свободной 2-разметки.

Утверждение 2. Пусть $\eta = (\eta_1, \eta_2)$ — свободная 2-разметка сети Σ . Тогда каждая из разметок η_1 и η_2 является свободной разметкой сети Σ .

Доказательство. Пусть $\mu = (\mu_1, \mu_2)$ — 2-разметка сети Σ , полученная в результате последовательного свободного продолжения начальной разметки

$$\begin{aligned}\mu_1(x_1^{(0)}) = \eta_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = \eta_1(x_n^{(0)}) = v_{1n}, \\ \mu_2(x_1^{(0)}) = \eta_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = \eta_2(x_n^{(0)}) = v_{2n}\end{aligned}$$

и пустого правила. Тогда, согласно определению процедуры последовательного свободного продолжения 2-разметки, разметка μ_1 является свободной разметкой сети Σ с начальным условием (v_{11}, \dots, v_{1n}) .

Согласно определению свободной 2-разметки, существует отображение σ_μ , удовлетворяющее условию $\sigma_\mu: \eta \rightarrow \mu$, в частности, отображение σ_μ удовлетворяет условию $\sigma_\mu: \eta_1 \rightarrow \mu_1$. Поскольку μ_1 — свободная разметка сети Σ с начальным условием (v_{11}, \dots, v_{1n}) , нетрудно понять, что η_1 также является свободной разметкой сети Σ . Аналогичным образом доказывается, что разметка η_2 является свободной разметкой сети Σ с начальным условием (v_{21}, \dots, v_{2n}) . ■

2. 2-Транзитивность сетей

Определение 23. 2-разметку $\mu = (\mu_1, \mu_2)$ сети Σ будем называть *2-разметкой сети Σ с ограничениями* $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$, если μ_1 и μ_2 — разметки сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ соответственно. При этом будем говорить, что сеть Σ *допускает* 2-разметку μ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$.

2-Разметку $\mu = (\mu_1, \mu_2)$ сети Σ с ограничениями будем называть *нетривиальной*, если μ_1 и μ_2 — различные разметки сети Σ , и *тривиальной* в противном случае. Для правильной непротиворечивой 2-разметки μ сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ её тривиальность равносильна двум равенствам $(v_{11}, \dots, v_{1n}) = (v_{21}, \dots, v_{2n})$ и $(w_{11}, \dots, w_{1n}) = (w_{21}, \dots, w_{2n})$. Поэтому далее будем подразумевать, что $(v_{11}, \dots, v_{1n}) \neq (v_{21}, \dots, v_{2n})$ и $(w_{11}, \dots, w_{1n}) \neq (w_{21}, \dots, w_{2n})$, когда будем говорить о *нетривиальной* правильной непротиворечивой 2-разметке μ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$.

Если биективная сеть Σ является 2-транзитивной для некоторого множества Ω , то для любых $(v_{11}, \dots, v_{1n}) \neq (v_{21}, \dots, v_{2n}), (w_{11}, \dots, w_{1n}) \neq (w_{21}, \dots, w_{2n}) \in \Omega^n$ существует такая квазигруппа $F \in \mathcal{Q}(\Omega)$, для которой выполняются равенства

$$\Sigma^F(v_{11}, \dots, v_{1n}) = (w_{11}, \dots, w_{1n}) \text{ и } \Sigma^F(v_{21}, \dots, v_{2n}) = (w_{21}, \dots, w_{2n}).$$

В таком случае квазигруппа F определяет нетривиальную правильную и непротиворечивую 2-разметку сети Σ элементами множества Ω с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$. Другими словами, существование нетривиальной правильной непротиворечивой 2-разметки сети Σ элементами множества Ω при произвольных ограниче-

ниях $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ является необходимым условием того, чтобы сеть Σ была 2-транзитивной для множества Ω .

Теорема 7. Пусть Σ — биективная сеть ширины n и Ω — множество мощности строго больше чем $2\|\Sigma\|$. Тогда следующие утверждения эквивалентны:

- 1) сеть Σ является 2-транзитивной для множества Ω ;
- 2) сеть Σ допускает нетривиальную правильную непротиворечивую 2-разметку элементами множества Ω при любых ограничениях $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$.

Доказательство.

1 \Rightarrow 2. Очевидно.

2 \Rightarrow 1. Каждой правильной непротиворечивой 2-разметке сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ соответствует непротиворечивое правило, определенное не более чем на $2\|\Sigma\| \leq |\Omega| - 1$ наборах. Согласно гипотезе Эванса, данное правило продолжается до квазигруппы на множестве Ω . ■

Следствие 3. Для биективной сети Σ следующие утверждения эквивалентны:

- 1) сеть Σ является 2-транзитивной для некоторого множества, мощность которого строго больше чем $2\|\Sigma\| + 2n$;
- 2) сеть Σ является 2-транзитивной для произвольного множества, мощность которого строго больше чем $2\|\Sigma\| + 2n$.

Устранение противоречий в разметке. Пусть η — произвольная 2-разметка сети $\Sigma = \Sigma_1 \cdot \dots \cdot \Sigma_t$. Свяжем с 2-разметкой η отношение $G \subset \mathbb{N}^3$, определённое следующим образом: G содержит тройку (y_l, y_r, y_q) в том и только в том случае, когда существует $s \in \{1, \dots, t\}$, для которого $\Sigma_s = \Sigma_m^{(i,j)}$ и выполняется хотя бы одно из равенств

$$(\eta_1(x_i^{(s-1)}), \eta_1(x_j^{(s-1)}), \eta_1(x_m^{(s)})) = (y_l, y_r, y_q) \text{ или } (\eta_2(x_i^{(s-1)}), \eta_2(x_j^{(s-1)}), \eta_2(x_m^{(s)})) = (y_l, y_r, y_q).$$

Если в отношении G содержатся две тройки, отличающиеся только в одной координате, например (y_l, y_r, y_q) и (y_l, y_r, y_p) , то, заменив в 2-разметке η все метки y_p на y_q , получим 2-разметку $\eta^{(1)}$, в которой используется на одну метку меньше, чем в 2-разметке η . Если в отношении $G^{(1)}$, соответствующем 2-разметке $\eta^{(1)}$, присутствуют две тройки, отличающиеся только в одной координате, то повторим описанные действия, и так далее.

Таким способом построим последовательность 2-разметок $\eta = \eta^{(0)}, \eta^{(1)}, \dots$ сети Σ , в которой каждая следующая 2-разметка использует на одну метку меньше, чем предыдущая. По этой причине последовательность 2-разметок оборвётся на некотором конечном шаге, например с номером k , в том смысле, что в отношении $G^{(k)}$, соответствующем 2-разметке $\eta^{(k)}$, не найдётся двух троек, отличающихся только в одной координате. Построенная таким образом 2-разметка $\eta^{(k)}$ будет правильной и непротиворечивой разметкой сети Σ , хотя, возможно, тривиальной.

Описанную процедуру будем называть *устранением противоречий в 2-разметке η* . При этом будем говорить, что 2-разметка $\eta^{(d)}$, $d \in \{0, 1, \dots, k\}$, получена из 2-разметки η *устранением противоречий*.

Лемма 1. Пусть η — произвольная 2-разметка сети Σ , μ — правильная непротиворечивая 2-разметка сети Σ и существует отображение $\sigma_\mu: \eta \rightarrow \mu$. Тогда для любой 2-разметки $\tilde{\eta}$, полученной из 2-разметки η устранением противоречий, также выполняется условие $\sigma_\mu: \tilde{\eta} \rightarrow \mu$.

Доказательство. Пусть $\eta = \eta^{(0)}, \eta^{(1)}, \dots, \eta^{(k)} = \tilde{\eta}$ — последовательность 2-разметок сети Σ , полученная в результате последовательного устранения противоречий в 2-разметке η . Для доказательства утверждения методом математической индукции достаточно показать, что для 2-разметки $\eta^{(1)}$ также выполняется условие $\sigma_\mu: \eta^{(1)} \rightarrow \mu$.

При построении 2-разметки $\eta^{(1)}$ в отношении G , соответствующем 2-разметке η , выбираются две тройки, отличающиеся только в одной координате. Рассмотрим все возможные случаи:

- 1) Если выбранная пара имеет вид (y_l, y_r, y_q) и (y_l, y_r, y_p) , то $\sigma_\mu(y_q) = \sigma_\mu(y_p)$, поскольку 2-разметка μ является правильной. Тогда при замене в 2-разметке η всех меток y_p на y_q получается 2-разметка $\eta^{(1)}$, для которой, очевидно, выполняется условие $\sigma_\mu: \eta^{(1)} \rightarrow \mu$.
- 2) Если выбранная пара имеет вид (y_l, y_r, y_q) и (y_l, y_p, y_q) , то $\sigma_\mu(y_r) = \sigma_\mu(y_p)$, поскольку 2-разметка μ является непротиворечивой. Тогда при замене в 2-разметке η всех меток y_p на y_r получается 2-разметка $\eta^{(1)}$, для которой, очевидно, выполняется условие $\sigma_\mu: \eta^{(1)} \rightarrow \mu$.
- 3) Если выбранная пара имеет вид (y_l, y_r, y_q) и (y_p, y_r, y_q) , то $\sigma_\mu(y_l) = \sigma_\mu(y_p)$, поскольку 2-разметка μ является непротиворечивой. Тогда при замене в 2-разметке η всех меток y_p на y_l получается 2-разметка $\eta^{(1)}$, для которой, очевидно, выполняется условие $\sigma_\mu: \eta^{(1)} \rightarrow \mu$.

Лемма доказана. ■

Следствие 4. Пусть η — произвольная 2-разметка сети Σ . Тогда правильная непротиворечивая 2-разметка $\tilde{\eta}$, полученная из 2-разметки η устранением противоречий, определена однозначно с точностью до обратимого переобозначения меток.

Доказательство. Пусть $\tilde{\eta}$ и $\hat{\eta}$ — две правильные непротиворечивые 2-разметки, полученные из 2-разметки η устранением противоречий. Тогда существуют отображения $\sigma_{\tilde{\eta}}$ и $\sigma_{\hat{\eta}}$, удовлетворяющие условиям $\sigma_{\tilde{\eta}}: \eta \rightarrow \tilde{\eta}$ и $\sigma_{\hat{\eta}}: \eta \rightarrow \hat{\eta}$. Согласно лемме 1, отображения $\sigma_{\tilde{\eta}}$ и $\sigma_{\hat{\eta}}$ также удовлетворяют условиям $\sigma_{\tilde{\eta}}: \hat{\eta} \rightarrow \tilde{\eta}$ и $\sigma_{\hat{\eta}}: \tilde{\eta} \rightarrow \hat{\eta}$, что и доказывает утверждение следствия. ■

Определение 24. Правильную непротиворечивую 2-разметку η сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ будем называть *свободной 2-разметкой* сети Σ с ограничениями, если для любой правильной непротиворечивой 2-разметки μ сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ существует отображение $\sigma_\mu: \eta \rightarrow \mu$.

Из определения свободной разметки следует, что, при условии существования, свободная 2-разметка сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ определена однозначно с точностью до обратимого переобозначения остальных меток.

Теорема 8. Если сеть Σ допускает нетривиальную правильную непротиворечивую 2-разметку с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$, то существует свободная 2-разметка сети Σ с указанными ограничениями.

Доказательство. Для удобства будем считать, что $\mathbb{N} \setminus \{v_{11}, v_{21}, \dots, v_{1n}, v_{2n}, w_{11}, w_{21}, \dots, w_{1n}, w_{2n}\} = \{y_{11}, y_{21}, y_{12}, y_{22}, \dots\}$. Пусть свободная 2-разметка $\eta' = (\eta'_1, \eta'_2)$ сети Σ получена в результате свободного продолжения начальной 2-разметки $\eta'_1(x_1^{(0)}) = v_{11}, \dots, \eta'_1(x_n^{(0)}) = v_{1n}$ и $\eta'_2(x_1^{(0)}) = v_{21}, \dots, \eta'_2(x_n^{(0)}) = v_{2n}$ с использованием меток $y_{11}, y_{21}, y_{12}, y_{22}, \dots$. Тогда для любой правильной непротиворечивой 2-разметки μ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ существует отображение σ_μ , удовлетворяющее условию $\sigma_\mu: \eta' \rightarrow \mu$. Продолжим указанное отображение σ_μ по правилу $\sigma_\mu(w_{1i}) = w_{1i}$, $\sigma_\mu(w_{2i}) = w_{2i}$, $i \in \{1, \dots, n\}$, и заменим в 2-разметке η' метки $\eta'_1(x_1^{(t)}), \dots, \eta'_1(x_n^{(t)})$ и $\eta'_2(x_1^{(t)}), \dots, \eta'_2(x_n^{(t)})$ на w_{11}, \dots, w_{1n} и w_{21}, \dots, w_{2n} соответственно. Таким образом, мы построили 2-разметку η'' сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$, при этом для любой правильной непротиворечивой 2-разметки μ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ существует отображение σ_μ , удовлетворяющее условию $\sigma_\mu: \eta'' \rightarrow \mu$.

Проведём процедуру устранения противоречий в 2-разметке η'' с уточнениями:

- если при устранении противоречия требуется отождествить метки v_{ri} и y_{sj} , то будем заменять метку y_{sj} на v_{ri} ;
- если при устранении противоречия требуется отождествить метки w_{ri} и y_{sj} , то будем заменять метку y_{sj} на w_{ri} .

Пусть η — правильная непротиворечивая 2-разметка сети Σ , полученная из 2-разметки η'' устранением противоречий. Тогда, согласно уточнениям, все метки $\eta_1(x_1^{(0)}), \eta_2(x_1^{(0)}), \dots, \eta_1(x_n^{(0)}), \eta_2(x_n^{(0)}), \eta_1(x_1^{(t)}), \eta_2(x_1^{(t)}), \dots, \eta_1(x_n^{(t)}), \eta_2(x_n^{(t)})$ содержатся в множестве $\{v_{11}, v_{21}, \dots, v_{1n}, v_{2n}, w_{11}, w_{21}, \dots, w_{1n}, w_{2n}\}$. При этом, согласно лемме 1, для любой правильной непротиворечивой 2-разметки μ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ соответствующее отображение σ_μ удовлетворяет условию $\sigma_\mu: \eta \rightarrow \mu$. Значит, 2-разметка η является свободной 2-разметкой сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$. ■

Следующая теорема фактически оправдывает название свободной разметки с ограничениями.

Теорема 9. Пусть имеются свободная 2-разметка η сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$, а также правильная непротиворечивая 2-разметка μ сети Σ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}$ и $\begin{pmatrix} \bar{v}_{21} & \dots & \bar{v}_{2n} \\ \bar{w}_{21} & \dots & \bar{w}_{2n} \end{pmatrix}$, при которых возможно определить отображение σ_μ по правилу

$$\sigma_\mu(v_{1i}) = \bar{v}_{1i}, \quad \sigma_\mu(w_{1i}) = \bar{w}_{1i}, \quad \sigma_\mu(v_{2i}) = \bar{v}_{2i}, \quad \sigma_\mu(w_{2i}) = \bar{w}_{2i}, \quad i \in \{1, \dots, n\}.$$

Тогда отображение σ_μ допускает такое продолжение, что $\sigma_\mu: \eta \rightarrow \mu$.

Доказательство. Пусть $\mathbb{N} \setminus \{v_{11}, v_{21}, \dots, v_{1n}, v_{2n}, w_{11}, w_{21}, \dots, w_{1n}, w_{2n}\} = \{y_{11}, y_{21}, y_{12}, y_{22}, \dots\}$ и свободная 2-разметка $\eta' = (\eta'_1, \eta'_2)$ сети Σ получена в результате свободного продолжения начальной 2-разметки $\eta'_1(x_1^{(0)}) = v_{11}, \dots, \eta'_1(x_n^{(0)}) = v_{1n}$ и

$\eta'_2(x_1^{(0)}) = v_{21}, \dots, \eta'_2(x_n^{(0)}) = v_{2n}$ с использованием меток $y_{11}, y_{21}, y_{12}, y_{22}, \dots$. Тогда, согласно теореме 6, для любой правильной непротиворечивой 2-разметки μ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}$ и $\begin{pmatrix} \bar{v}_{21} & \dots & \bar{v}_{2n} \\ \bar{w}_{21} & \dots & \bar{w}_{2n} \end{pmatrix}$, при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_{1i}) = \bar{v}_{1i}, \sigma_\mu(v_{2i}) = \bar{v}_{2i}, \sigma_\mu(w_{1i}) = \bar{w}_{1i}, \sigma_\mu(w_{2i}) = \bar{w}_{2i}, i \in \{1, \dots, n\}$, отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu: \eta' \rightarrow \mu$. Заменив в 2-разметке η' метки $\eta'_1(x_1^{(t)}), \dots, \eta'_1(x_n^{(t)})$ и $\eta'_2(x_1^{(t)}), \dots, \eta'_2(x_n^{(t)})$ на w_{11}, \dots, w_{1n} и w_{21}, \dots, w_{2n} соответственно, получим 2-разметку η'' сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$, при этом для любой правильной непротиворечивой 2-разметки μ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}$ и $\begin{pmatrix} \bar{v}_{21} & \dots & \bar{v}_{2n} \\ \bar{w}_{21} & \dots & \bar{w}_{2n} \end{pmatrix}$, при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_{1i}) = \bar{v}_{1i}, \sigma_\mu(v_{2i}) = \bar{v}_{2i}, \sigma_\mu(w_{1i}) = \bar{w}_{1i}, \sigma_\mu(w_{2i}) = \bar{w}_{2i}, i \in \{1, \dots, n\}$, отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu: \eta'' \rightarrow \mu$.

Проведём процедуру устранения противоречий в 2-разметке η'' с уточнениями:

- если при устранении противоречия требуется отождествить метки v_{ri} и y_{sj} , то будем заменять метку y_{sj} на v_{ri} ;
- если при устранении противоречия требуется отождествить метки w_{ri} и y_{sj} , то будем заменять метку y_{sj} на w_{ri} .

В доказательстве теоремы 8 показано, что правильная непротиворечивая 2-разметка сети Σ , полученная из 2-разметки η'' устранением противоречий, является свободной 2-разметкой сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$. Не ограничивая общности, можно считать, что при устранении противоречий в разметке η'' получается свободная разметка η . Поскольку для любой правильной непротиворечивой разметки μ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}$ и $\begin{pmatrix} \bar{v}_{21} & \dots & \bar{v}_{2n} \\ \bar{w}_{21} & \dots & \bar{w}_{2n} \end{pmatrix}$, при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_{1i}) = \bar{v}_{1i}, \sigma_\mu(v_{2i}) = \bar{v}_{2i}, \sigma_\mu(w_{1i}) = \bar{w}_{1i}, \sigma_\mu(w_{2i}) = \bar{w}_{2i}, i \in \{1, \dots, n\}$, отображение σ_μ допускает продолжение, удовлетворяющее условию $\sigma_\mu: \eta'' \rightarrow \mu$, то, согласно лемме 1, данное продолжение также удовлетворяет условию $\sigma_\mu: \eta \rightarrow \mu$. ■

Следствие 5. В условиях теоремы 9 если G и F — минимальные правила разметок η и μ соответственно, то при всех допустимых $z_i, z_j \in \mathbb{N}$ выполняется равенство $\sigma_\mu(G(z_i, z_j)) = F(\sigma_\mu(z_i), \sigma_\mu(z_j))$.

Достаточным условием для существования правильных непротиворечивых 2-разметок сети Σ при всех возможных ограничениях из \mathbb{N} является существование правильных непротиворечивых 2-разметок сети Σ при всех возможных ограничениях из Ω_{4n} . Однако, как и в случае с простыми разметками, справедливо более сильное утверждение.

Теорема 10. Сеть Σ допускает нетривиальные правильные непротиворечивые 2-разметки при всех возможных ограничениях $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ из \mathbb{N} в том и только в том случае, когда сеть Σ допускает нетривиальные правильные

непротиворечивые 2-разметки при всех возможных ограничениях $\begin{pmatrix} \bar{v}_{11} & \cdots & \bar{v}_{1n} \\ \bar{w}_{11} & \cdots & \bar{w}_{1n} \end{pmatrix}$ и $\begin{pmatrix} \bar{v}_{21} & \cdots & \bar{v}_{2n} \\ \bar{w}_{21} & \cdots & \bar{w}_{2n} \end{pmatrix}$ из Ω_3 .

Доказательство. Необходимость очевидна, докажем достаточность. Пусть $\mathbb{N} \setminus \{v_{11}, v_{21}, \dots, v_{1n}, v_{2n}, w_{11}, w_{21}, \dots, w_{1n}, w_{2n}\} = \{y_{11}, y_{21}, y_{12}, y_{22}, \dots\}$ и свободная 2-разметка $\eta' = (\eta'_1, \eta'_2)$ сети Σ получена в результате свободного продолжения начальной 2-разметки $\eta'_1(x_1^{(0)}) = v_{11}, \dots, \eta'_1(x_n^{(0)}) = v_{1n}$ и $\eta'_2(x_1^{(0)}) = v_{21}, \dots, \eta'_2(x_n^{(0)}) = v_{2n}$ с использованием меток $y_{11}, y_{21}, y_{12}, y_{22}, \dots$. Тогда, согласно теореме 6, для любой правильной непротиворечивой 2-разметки μ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \cdots & \bar{v}_{1n} \\ \bar{w}_{11} & \cdots & \bar{w}_{1n} \end{pmatrix}$ и $\begin{pmatrix} \bar{v}_{21} & \cdots & \bar{v}_{2n} \\ \bar{w}_{21} & \cdots & \bar{w}_{2n} \end{pmatrix}$ из Ω_3 , при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_{1i}) = \bar{v}_{1i}$, $\sigma_\mu(v_{2i}) = \bar{v}_{2i}$, $\sigma_\mu(w_{1i}) = \bar{w}_{1i}$, $\sigma_\mu(w_{2i}) = \bar{w}_{2i}$, $i \in \{1, \dots, n\}$, отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu: \eta' \rightarrow \mu$. Заменяя в 2-разметке η' метки $\eta'_1(x_1^{(t)}), \dots, \eta'_1(x_n^{(t)})$ и $\eta'_2(x_1^{(t)}), \dots, \eta'_2(x_n^{(t)})$ на w_{11}, \dots, w_{1n} и w_{21}, \dots, w_{2n} соответственно, получим 2-разметку η'' сети Σ с ограничениями $\begin{pmatrix} v_{11} & \cdots & v_{1n} \\ w_{11} & \cdots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \cdots & v_{2n} \\ w_{21} & \cdots & w_{2n} \end{pmatrix}$, при этом для любой правильной непротиворечивой 2-разметки μ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \cdots & \bar{v}_{1n} \\ \bar{w}_{11} & \cdots & \bar{w}_{1n} \end{pmatrix}$ и $\begin{pmatrix} \bar{v}_{21} & \cdots & \bar{v}_{2n} \\ \bar{w}_{21} & \cdots & \bar{w}_{2n} \end{pmatrix}$, при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_{1i}) = \bar{v}_{1i}$, $\sigma_\mu(v_{2i}) = \bar{v}_{2i}$, $\sigma_\mu(w_{1i}) = \bar{w}_{1i}$, $\sigma_\mu(w_{2i}) = \bar{w}_{2i}$, $i \in \{1, \dots, n\}$, отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu: \eta'' \rightarrow \mu$.

Проведём процедуру устранения противоречий в 2-разметке η'' с уточнениями:

- если при устранении противоречия требуется отождествить метки v_{ri} и y_{sj} , то будем заменять метку y_{sj} на v_{ri} ;
- если при устранении противоречия требуется отождествить метки w_{ri} и y_{sj} , то будем заменять метку y_{sj} на w_{ri} .

Пусть η — правильная непротиворечивая 2-разметка сети Σ , полученная из 2-разметки η'' устранением противоречий. Тогда, согласно сделанным уточнениям, все метки $\eta_1(x_1^{(0)}), \eta_2(x_1^{(0)}), \dots, \eta_1(x_n^{(0)}), \eta_2(x_n^{(0)}), \eta_1(x_1^{(t)}), \eta_2(x_1^{(t)}), \dots, \eta_1(x_n^{(t)}), \eta_2(x_n^{(t)})$ содержатся в множестве $\{v_{11}, v_{21}, \dots, v_{1n}, v_{2n}, w_{11}, w_{21}, \dots, w_{1n}, w_{2n}\}$. При этом, согласно лемме 1, для любой правильной непротиворечивой 2-разметки μ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \cdots & \bar{v}_{1n} \\ \bar{w}_{11} & \cdots & \bar{w}_{1n} \end{pmatrix}$ и $\begin{pmatrix} \bar{v}_{21} & \cdots & \bar{v}_{2n} \\ \bar{w}_{21} & \cdots & \bar{w}_{2n} \end{pmatrix}$ из Ω_3 , при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_{1i}) = \bar{v}_{1i}$, $\sigma_\mu(v_{2i}) = \bar{v}_{2i}$, $\sigma_\mu(w_{1i}) = \bar{w}_{1i}$, $\sigma_\mu(w_{2i}) = \bar{w}_{2i}$, $i \in \{1, \dots, n\}$, отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu: \eta \rightarrow \mu$.

Методом от противного покажем, что правильная непротиворечивая 2-разметка η является 2-разметкой с ограничениями $\begin{pmatrix} v_{11} & \cdots & v_{1n} \\ w_{11} & \cdots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \cdots & v_{2n} \\ w_{21} & \cdots & w_{2n} \end{pmatrix}$. Согласно уточнениям, в разметке η могли появиться противоречия только следующих типов:

- $\eta_r(x_i^{(0)}) = v_{sj} \neq v_{ri}$;
- $\eta_r(x_i^{(0)}) = w_{sj} \neq v_{ri}$;
- $\eta_r(x_i^{(t)}) = v_{sj} \neq w_{ri}$;
- $\eta_r(x_i^{(t)}) = w_{sj} \neq w_{ri}$.

Разберём подслучай $\eta_1(x_i^{(0)}) = v_{1j} \neq v_{1i}$, относящийся к первому типу противоречий. Не ограничивая общности, будем считать, что ограничения $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ не содержат элементов из Ω_3 . Поскольку $(v_{11}, \dots, v_{1n}) \neq (v_{21}, \dots, v_{2n})$ и $(w_{11}, \dots, w_{1n}) \neq (w_{21}, \dots, w_{2n})$, то существуют такие $i_1, i_2 \in \{1, \dots, n\}$, что $v_{1i_1} \neq v_{2i_1}$ и $w_{1i_2} \neq w_{2i_2}$. Заменяем в ограничениях $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \dots & v_{2n} \\ w_{21} & \dots & w_{2n} \end{pmatrix}$ все вхождения элемента v_{1i} на 1 и все вхождения элемента v_{1j} на 2. Полученные ограничения будем обозначать $\begin{pmatrix} v'_{11} & \dots & v'_{1n} \\ w'_{11} & \dots & w'_{1n} \end{pmatrix}$ и $\begin{pmatrix} v'_{21} & \dots & v'_{2n} \\ w'_{21} & \dots & w'_{2n} \end{pmatrix}$. Нетрудно понять, что $v'_{1i_1} \neq v'_{2i_1}$ и $w'_{1i_2} \neq w'_{2i_2}$, и, не ограничивая общности, достаточно рассмотреть следующие варианты:

- 1) если $v'_{1i_1}, v'_{2i_1}, w'_{1i_2} \in \Omega_2$ и $w'_{2i_2} \notin \Omega_2$, то заменим в ограничениях $\begin{pmatrix} v'_{11} & \dots & v'_{1n} \\ w'_{11} & \dots & w'_{1n} \end{pmatrix}$ и $\begin{pmatrix} v'_{21} & \dots & v'_{2n} \\ w'_{21} & \dots & w'_{2n} \end{pmatrix}$ все вхождения элемента w'_{2i_2} на $(3 - w'_{1i_2})$;
- 2) если $v'_{1i_1}, v'_{2i_1} \in \Omega_2$ и $w'_{1i_2}, w'_{2i_2} \notin \Omega_2$, то заменим в ограничениях $\begin{pmatrix} v'_{11} & \dots & v'_{1n} \\ w'_{11} & \dots & w'_{1n} \end{pmatrix}$ и $\begin{pmatrix} v'_{21} & \dots & v'_{2n} \\ w'_{21} & \dots & w'_{2n} \end{pmatrix}$ все вхождения элемента w'_{1i_2} на 1, все вхождения w'_{2i_2} на 2;
- 3) если $v'_{1i_1}, w'_{1i_2} \in \Omega_2$ и $v'_{2i_1} = w'_{2i_2} \notin \Omega_2$, то заменим в ограничениях $\begin{pmatrix} v'_{11} & \dots & v'_{1n} \\ w'_{11} & \dots & w'_{1n} \end{pmatrix}$ и $\begin{pmatrix} v'_{21} & \dots & v'_{2n} \\ w'_{21} & \dots & w'_{2n} \end{pmatrix}$ все вхождения элементов v'_{2i_1} и w'_{2i_2} на 3;
- 4) если $v'_{1i_1}, w'_{1i_2} \in \Omega_2$ и $v'_{2i_1}, w'_{2i_2} \notin \Omega_2$, $v'_{2i_1} \neq w'_{2i_2}$, то заменим в ограничениях $\begin{pmatrix} v'_{11} & \dots & v'_{1n} \\ w'_{11} & \dots & w'_{1n} \end{pmatrix}$ и $\begin{pmatrix} v'_{21} & \dots & v'_{2n} \\ w'_{21} & \dots & w'_{2n} \end{pmatrix}$ все вхождения элементов v'_{2i_1} на $(3 - v'_{1i_1})$ и все вхождения элементов w'_{2i_2} на $(3 - w'_{1i_2})$;
- 5) если $v'_{1i_1}, w'_{2i_2} \in \Omega_2$ и $v'_{2i_1} = w'_{1i_2} \notin \Omega_2$, то заменим в ограничениях $\begin{pmatrix} v'_{11} & \dots & v'_{1n} \\ w'_{11} & \dots & w'_{1n} \end{pmatrix}$ и $\begin{pmatrix} v'_{21} & \dots & v'_{2n} \\ w'_{21} & \dots & w'_{2n} \end{pmatrix}$ все вхождения элементов v'_{2i_1} и w'_{1i_2} на 3;
- 6) если $v'_{1i_1}, w'_{2i_2} \in \Omega_2$ и $v'_{2i_1}, w'_{1i_2} \notin \Omega_2$, $v'_{2i_1} \neq w'_{1i_2}$, то заменим в ограничениях $\begin{pmatrix} v'_{11} & \dots & v'_{1n} \\ w'_{11} & \dots & w'_{1n} \end{pmatrix}$ и $\begin{pmatrix} v'_{21} & \dots & v'_{2n} \\ w'_{21} & \dots & w'_{2n} \end{pmatrix}$ все вхождения элементов v'_{2i_1} на $(3 - v'_{1i_1})$ и все вхождения элементов w'_{1i_2} на $(3 - w'_{2i_2})$;
- 7) если $v'_{1i_1} \in \Omega_2$ и $v'_{2i_1}, w'_{1i_2}, w'_{2i_2} \notin \Omega_2$, то заменим в ограничениях $\begin{pmatrix} v'_{11} & \dots & v'_{1n} \\ w'_{11} & \dots & w'_{1n} \end{pmatrix}$ и $\begin{pmatrix} v'_{21} & \dots & v'_{2n} \\ w'_{21} & \dots & w'_{2n} \end{pmatrix}$ все вхождения элемента v'_{2i_1} на $(3 - v'_{1i_1})$, после чего всё аналогично первому или второму варианту;
- 8) если $v'_{1i_1}, v'_{2i_1}, w'_{1i_2}, w'_{2i_2} \notin \Omega_2$, то заменим в ограничениях $\begin{pmatrix} v'_{11} & \dots & v'_{1n} \\ w'_{11} & \dots & w'_{1n} \end{pmatrix}$ и $\begin{pmatrix} v'_{21} & \dots & v'_{2n} \\ w'_{21} & \dots & w'_{2n} \end{pmatrix}$ все вхождения элемента v'_{1i_1} на 1, а все вхождения элемента v'_{2i_1} на 2, после чего всё аналогично первому или второму варианту.

При $v'_{1i_1}, v'_{2i_1}, w'_{1i_2}, w'_{2i_2} \in \Omega_2$, а также в заключение случаев 1, 2, 4, 6, 7, 8 следует заменить в имеющихся ограничениях все элементы, отличные от 1 и 2, на 1, а в заключение случаев 3, 5 — все элементы, отличные от 1, 2 и 3, на 1. В результате

будут получены ограничения $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}$ и $\begin{pmatrix} \bar{v}_{21} & \dots & \bar{v}_{2n} \\ \bar{w}_{21} & \dots & \bar{w}_{2n} \end{pmatrix}$ из Ω_3 . Согласно условию теоремы, существует правильная непротиворечивая 2-разметка μ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}$ и $\begin{pmatrix} \bar{v}_{21} & \dots & \bar{v}_{2n} \\ \bar{w}_{21} & \dots & \bar{w}_{2n} \end{pmatrix}$ из Ω_3 , при этом отображение σ_μ , определённое по правилу $\sigma_\mu(v_{1i}) = \bar{v}_{1i}$, $\sigma_\mu(v_{2i}) = \bar{v}_{2i}$, $\sigma_\mu(w_{1i}) = \bar{w}_{1i}$, $\sigma_\mu(w_{2i}) = \bar{w}_{2i}$, $i \in \{1, \dots, n\}$, продолжается таким образом, что удовлетворяет условию $\sigma_\mu: \eta \rightarrow \mu$. Получили противоречие, поскольку $\sigma_\mu(\eta_1(x_i^{(0)})) = \sigma_\mu(v_{1j}) = 2 \neq 1 = \bar{v}_{1i} = \mu(x_i^{(0)})$. Отсутствие противоречий всех остальных типов устанавливается аналогичным образом. ■

Замечание 4. Интересным представляется вопрос о том, останется ли верным утверждение теоремы 10, если в нём заменить множество Ω_3 на Ω_2 . Нетрудно видеть, что имеющееся доказательство теоремы 10 существенным образом использует условие $|\Omega| \geq 3$. Так, согласно доказательству теоремы 10, если Σ — биективная сеть ширины 2, то существование правильных непротиворечивых разметок сети Σ при всех возможных ограничениях $\begin{pmatrix} \bar{v}_{11} & \bar{v}_{12} \\ \bar{w}_{11} & \bar{w}_{12} \end{pmatrix}$ и $\begin{pmatrix} \bar{v}_{21} & \bar{v}_{22} \\ \bar{w}_{21} & \bar{w}_{22} \end{pmatrix}$ из Ω_2 не может гарантировать существование правильной непротиворечивой разметки сети Σ с ограничениями $\begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix}$ и $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$.

В дальнейшем нам потребуются естественные обобщения понятия свободной 2-разметки с ограничениями.

Определение 25. Правильную непротиворечивую 2-разметку η сети Σ будем называть *свободной 2-разметкой сети Σ с условиями*

$$\begin{aligned} \eta_1(x_1^{(0)}) = v_{11}, \dots, \eta_1(x_n^{(0)}) = v_{1n}, \quad \eta_2(x_1^{(0)}) = v_{21}, \dots, \eta_2(x_n^{(0)}) = v_{2n}, \\ \eta_1(x_{i_1}^{(t)}) = w_{1i_1}, \dots, \eta_1(x_{i_k}^{(t)}) = w_{1i_k}, \quad \eta_2(x_{j_1}^{(t)}) = w_{2j_1}, \dots, \eta_2(x_{j_l}^{(t)}) = w_{2j_l}, \end{aligned}$$

если для любой правильной непротиворечивой 2-разметки μ сети Σ с аналогичными условиями

$$\begin{aligned} \mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}, \quad \mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}, \\ \mu_1(x_{i_1}^{(t)}) = w_{1i_1}, \dots, \mu_1(x_{i_k}^{(t)}) = w_{1i_k}, \quad \mu_2(x_{j_1}^{(t)}) = w_{2j_1}, \dots, \mu_2(x_{j_l}^{(t)}) = w_{2j_l} \end{aligned}$$

существует такое отображение σ_μ , что $\sigma_\mu: \eta \rightarrow \mu$.

Теорема 11. Если существует правильная непротиворечивая 2-разметка μ сети Σ с условиями

$$\begin{aligned} \mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}, \quad \mu_2(x_1^{(0)}) = v_{21}, \dots, \mu_2(x_n^{(0)}) = v_{2n}, \\ \mu_1(x_{i_1}^{(t)}) = w_{1i_1}, \dots, \mu_1(x_{i_k}^{(t)}) = w_{1i_k}, \quad \mu_2(x_{j_1}^{(t)}) = w_{2j_1}, \dots, \mu_2(x_{j_l}^{(t)}) = w_{2j_l}, \end{aligned}$$

то существует единственная, с точностью до переобозначений, свободная разметка η сети Σ с теми же условиями.

Теорема 12. Пусть для свободной 2-разметки η с условиями

$$\begin{aligned} \eta_1(x_1^{(0)}) = v_{11}, \dots, \eta_1(x_n^{(0)}) = v_{1n}, \quad \eta_2(x_1^{(0)}) = v_{21}, \dots, \eta_2(x_n^{(0)}) = v_{2n}, \\ \eta_1(x_{i_1}^{(t)}) = w_{1i_1}, \dots, \eta_1(x_{i_k}^{(t)}) = w_{1i_k}, \quad \eta_2(x_{j_1}^{(t)}) = w_{2j_1}, \dots, \eta_2(x_{j_l}^{(t)}) = w_{2j_l} \end{aligned}$$

и правильной непротиворечивой разметки μ с условиями

$$\begin{aligned} \mu_1(x_1^{(0)}) = \bar{v}_{11}, \dots, \mu_1(x_n^{(0)}) = \bar{v}_{1n}, \quad \mu_2(x_1^{(0)}) = \bar{v}_{21}, \dots, \mu_2(x_n^{(0)}) = \bar{v}_{2n}, \\ \mu_1(x_{i_1}^{(t)}) = \bar{w}_{1i_1}, \dots, \mu_1(x_{i_k}^{(t)}) = \bar{w}_{1i_k}, \quad \mu_2(x_{j_1}^{(t)}) = \bar{w}_{2j_1}, \dots, \mu_2(x_{j_l}^{(t)}) = \bar{w}_{2j_l} \end{aligned}$$

возможно определить отображение σ_μ по правилу $\sigma_\mu(v_{1i}) = \bar{v}_{1i}$, $\sigma_\mu(v_{2i}) = \bar{v}_{2i}$, $\sigma_\mu(w_{1i_s}) = \bar{w}_{1i_s}$, $\sigma_\mu(w_{2j_t}) = \bar{w}_{2j_t}$. Тогда отображение σ_μ допускает продолжение, удовлетворяющее условию $\sigma_\mu: \eta \rightarrow \mu$.

Следствие 6. В условиях теоремы 12, если G и F — минимальные правила 2-разметок η и μ соответственно, то при всех допустимых $z_i, z_j \in \mathbb{N}$ выполняется равенство $\sigma_\mu(G(z_i, z_j)) = F(\sigma_\mu(z_i), \sigma_\mu(z_j))$. В частности, если метка $\mu_r(x_i^{(s)})$ не содержится в области определения F , то метка $\eta_r(x_i^{(s)})$ не содержится в области определения G .

3. Построение 2-транзитивных сетей

Приведём алгоритм 1 модификации произвольной биективной сети Σ до биективной сети $\widehat{\Sigma}$, которая является 2-транзитивной для всех достаточно больших множеств.

Алгоритм 1. Построение 2-транзитивной сети

Вход: произвольная биективная сеть $\Sigma = (\Sigma_{11} \dots \Sigma_{1k_1}) \dots (\Sigma_{n1} \dots \Sigma_{nk_n})$.

1: Для всех $s = 1, 2, \dots, n - 1$:

Пусть первые $(s - 1)$ слоев канонического представления сети Σ уже модифицированы, $\widehat{\Sigma}_{s-1} = (\Sigma_{11} \dots \Sigma_{1\widehat{k}_1}) \dots (\Sigma_{(s-1)1} \dots \Sigma_{(s-1)\widehat{k}_{s-1}})$, μ — свободная разметка сети $\widehat{\Sigma}_{s-1}$ с условиями $\mu(x_1^{(0)}) = v_1, \dots, \mu(x_n^{(0)}) = v_1, \mu(x_1^{(\widehat{k}_1)}) = v_1, \dots, \mu(x_{s-1}^{(\widehat{k}_1 + \dots + \widehat{k}_{s-1})}) = v_1$. Продолжим разметку μ сети $\widehat{\Sigma}_{s-1}$ свободным образом до разметки сети $\widehat{\Sigma}'_s = (\Sigma_{11} \dots \Sigma_{1\widehat{k}_1}) \dots (\Sigma_{(s-1)1} \dots \Sigma_{(s-1)\widehat{k}_{s-1}}) \cdot (\Sigma_{s1} \dots \Sigma_{sk_s})$ и выберем такую вершину $x_j^{(\widehat{k}_1 + \dots + \widehat{k}_{s-1} + k_s)}$, метка которой $\mu(x_j^{(\widehat{k}_1 + \dots + \widehat{k}_{s-1} + k_s)})$ не содержится в области определения $F_{\widehat{\Sigma}'_s}$ — минимального правила разметки μ сети $\widehat{\Sigma}'_s$.

2: Если $s \leq n - 2$ и $j = s$, то

3: выберем произвольные $l, m \in \{s + 1, \dots, n\}$, $l \neq m$, и модифицируем s -й слой $\Sigma_{s1} \dots \Sigma_{sk_s}$ следующим образом: $\Sigma_{s1} \dots \Sigma_{s\widehat{k}_s} = \Sigma_{s1} \dots \Sigma_{sk_s} \cdot \Sigma_l^{(s,l)} \cdot \Sigma_s^{(s,l)} \cdot \Sigma_m^{(l,m)}$.

4: Если $s \leq n - 2$ и $j \neq s$, то

5: выберем произвольный $m \in \{s + 1, \dots, n\}$, $m \neq j$, и модифицируем s -й слой $\Sigma_{s1} \dots \Sigma_{sk_s}$ следующим образом: $\Sigma_{s1} \dots \Sigma_{s\widehat{k}_s} = \Sigma_{s1} \dots \Sigma_{sk_s} \cdot \Sigma_s^{(s,j)} \cdot \Sigma_m^{(s,m)} \cdot \Sigma_s^{(j,s)}$.

6: Если $s = n - 1$ и $j = n - 1$, то

7: модифицируем $(n - 1)$ -й слой $\Sigma_{(n-1)1} \dots \Sigma_{(n-1)k_{n-1}}$ следующим образом:

$$\Sigma_{(n-1)1} \dots \Sigma_{(n-1)\widehat{k}_{n-1}} = \Sigma_{(n-1)1} \dots \Sigma_{(n-1)k_{n-1}} \cdot (\Sigma_n^{(n-1,n)} \cdot \Sigma_{n-1}^{(n-1,n)}) \cdot (\Sigma_n^{(n,n-1)} \dots \Sigma_n^{(n,1)}).$$

8: Если $s = n - 1$ и $j = n$, то

9: модифицируем $(n - 1)$ -й слой $\Sigma_{(n-1)1} \dots \Sigma_{(n-1)k_{n-1}}$ следующим образом:

$$\Sigma_{(n-1)1} \dots \Sigma_{(n-1)\widehat{k}_{n-1}} = \Sigma_{(n-1)1} \dots \Sigma_{(n-1)k_{n-1}} \cdot (\Sigma_{n-1}^{(n,n-1)} \cdot \Sigma_n^{(n,n-1)}) \cdot (\Sigma_{n-1}^{(n-1,n)} \dots \Sigma_{n-1}^{(n-1,1)}).$$

Выход: $(\Sigma_{11} \dots \Sigma_{1\widehat{k}_1}) \dots (\Sigma_{(n-1)1} \dots \Sigma_{(n-1)\widehat{k}_{n-1}})$ — «почти» каноническое представление биективной сети $\widehat{\Sigma}$, сложность которой не превосходит $\|\Sigma\| + 6n - 7$.

Не ограничивая общности, всюду далее будем считать, что произвольная биективная сеть Σ совпадает со своим каноническим представлением

$$(\Sigma_{11} \cdot \dots \cdot \Sigma_{1k_1}) \cdot \dots \cdot (\Sigma_{n1} \cdot \dots \cdot \Sigma_{nk_n})$$

с множеством вершин $X_0 \cup X_{11} \cup \dots \cup X_{1k_1} \cup \dots \cup X_{n1} \cup \dots \cup X_{nk_n}$ и что первый слой имеет вид $\Sigma_{11} \cdot \dots \cdot \Sigma_{1k_1} = \Sigma_{11} \cdot \dots \cdot \Sigma_{1l} \cdot \Sigma_2^{(1,2)} \cdot \dots \cdot \Sigma_n^{(n-1,n)} \cdot \Sigma_{n-1}^{(n-1,n)} \cdot \dots \cdot \Sigma_1^{(1,2)}$, в противном случае приведём его к такому виду, добавив не более $2(n-1)$ соответствующих элементарных сетей.

Теорема 13. Пусть Σ — произвольная биективная сеть ширины n . Тогда её модификация $\widehat{\Sigma}$ является 2-транзитивной для любого множества Ω , мощность которого больше чем $2\|\Sigma\| + 14n - 14$.

Доказательство. Всяду далее будем полагать, что каждая свободная 2-разметка получена при помощи параллельного свободного продолжения соответствующих начальных условий с использованием пары множеств $Y_1 = \{y_{11}, y_{12}, \dots\}$ и $Y_2 = \{y_{21}, y_{22}, \dots\}$. Описание корректности действий, выполняемых на шагах 0 и 1, сформулируем в виде следующей леммы.

Лемма 2. Для любой свободной 2-разметки $\eta = (\eta_1, \eta_2)$ сети

$$\widehat{\Sigma}'_1 = \Sigma_{11} \cdot \dots \cdot \Sigma_{1l} \cdot \Sigma_2^{(1,2)} \cdot \dots \cdot \Sigma_n^{(n-1,n)} \cdot \Sigma_{n-1}^{(n-1,n)} \cdot \dots \cdot \Sigma_1^{(1,2)}$$

с начальными условиями $(v_{11}, \dots, v_{1n}) \neq (v_{21}, \dots, v_{2n})$ справедливо, что

- $\eta_1(x_1^{(k_1)}), \dots, \eta_1(x_n^{(k_1)}) \in Y_1$ и $\eta_2(x_1^{(k_1)}), \dots, \eta_2(x_n^{(k_1)}) \in Y_2$;
- метки $\eta_1(x_1^{(k_1)})$ и $\eta_2(x_1^{(k_1)})$ (и только они), независимо от условий (v_{11}, \dots, v_{1n}) и (v_{21}, \dots, v_{2n}) , не содержатся в области определения $G_{\widehat{\Sigma}'_1}$ — минимального правила 2-разметки η .

Доказательство. Введём понятие уровня метки в свободной 2-разметке η произвольной сети $\Sigma = \Sigma_1 \cdot \dots \cdot \Sigma_t$. Для меток $\eta_1(x_1^{(0)}), \dots, \eta_1(x_n^{(0)}); \eta_2(x_1^{(0)}), \dots, \eta_2(x_n^{(0)})$ уровень $h(\eta_*(x_i^{(0)}))$ полагаем равным нулю. Если метка z_s удовлетворяет соотношению $G_\Sigma(z_l, z_r) = z_s$ для минимального правила G_Σ 2-разметки η сети Σ , то полагаем $h(z_s) = \max\{h(z_l), h(z_r)\} + 1$. Такое определение корректно, поскольку минимальное правило G_Σ свободной 2-разметки η удовлетворяет условию

$$(G_{\Sigma, \eta}(z_1, z_2) = G_{\Sigma, \eta}(z_3, z_4)) \implies ((z_1, z_2) = (z_3, z_4))$$

при всех допустимых $z_1, z_2, z_3, z_4 \in \mathbb{N}$.

Индукцией по длине произведения элементарных сетей $\Sigma_1 \cdot \dots \cdot \Sigma_t$ нетрудно показать, что для любой вершины $x_i^{(s)}$ сети Σ уровни меток $\eta_1(x_i^{(s)})$ и $\eta_2(x_i^{(s)})$ совпадают.

Пусть η — свободная 2-разметка сети

$$\Sigma_{11} \cdot \dots \cdot \Sigma_{1l} \cdot \Sigma_2^{(1,2)} \cdot \dots \cdot \Sigma_n^{(n-1,n)},$$

полученная в результате параллельного свободного продолжения начальных условий $(v_{11}, \dots, v_{1n}) \neq (v_{21}, \dots, v_{2n})$ с использованием пары множеств $Y_1 = \{y_{11}, y_{12}, \dots\}$ и $Y_2 = \{y_{21}, y_{22}, \dots\}$. Тогда в 2-разметке η ровно две метки имеют максимальный уровень — это $\eta_1(x_n^{(l+n-1)})$ и $\eta_2(x_n^{(l+n-1)})$. Из максимальной уровня следует, что обе метки $\eta_1(x_n^{(l+n-1)})$ и $\eta_2(x_n^{(l+n-1)})$ не содержатся в области определения минимального правила 2-разметки η .

Соотношение $\eta_1(x_n^{(l+n-1)}) \in Y_1$ выполняется по построению. Предположим, что $\eta_2(x_n^{(l+n-1)}) \in Y_1$. Тогда, согласно построению, метка $\eta_2(x_n^{(l+n-1)})$ впервые появилась в разметке η_1 и, следовательно, в силу максимальности её уровня, должна совпадать единственно с $\eta_1(x_n^{(l+n-1)})$. Но в таком случае, пользуясь конструктивной особенностью свободной 2-разметки, нетрудно показать совпадение $\eta_1(x_1^{(l)}) = \eta_2(x_1^{(l)})$, \dots , $\eta_1(x_n^{(l)}) = \eta_2(x_n^{(l)})$, которое противоречит начальному условию $(v_{11}, \dots, v_{1n}) \neq (v_{21}, \dots, v_{2n})$. Таким образом, $\eta_2(x_n^{(l+n-1)}) \in Y_2$.

Ввиду того, что обе метки $\eta_1(x_n^{(l+n-1)})$ и $\eta_2(x_n^{(l+n-1)})$ не содержатся в области определения минимального правила 2-разметки η , при продолжении свободной 2-разметки η свободным образом (с использованием пары множеств $Y_1 = \{y_{11}, y_{12}, \dots\}$ и $Y_2 = \{y_{21}, y_{22}, \dots\}$) до свободной разметки сети

$$\widehat{\Sigma}'_1 = \Sigma_{11} \cdot \dots \cdot \Sigma_{1l} \cdot \Sigma_2^{(1,2)} \cdot \dots \cdot \Sigma_n^{(n-1,n)} \cdot \Sigma_{n-1}^{(n-1,n)} \cdot \dots \cdot \Sigma_1^{(1,2)}$$

будут выполнены следующие условия:

- $\eta_1(x_1^{(k_1)}), \dots, \eta_1(x_n^{(k_1)}) \in Y_1$ и $\eta_2(x_1^{(k_1)}), \dots, \eta_2(x_n^{(k_1)}) \in Y_2$;
- метки $\eta_1(x_1^{(k_1)})$ и $\eta_2(x_1^{(k_1)})$ (и только они), независимо от условий (v_{11}, \dots, v_{1n}) и (v_{21}, \dots, v_{2n}) , не содержатся в области определения $G_{\widehat{\Sigma}'_1}$ — минимального правила 2-разметки η .

Лемма доказана. ■

Ввиду леммы 2, для свободной разметки μ сети $\widehat{\Sigma}'_1 = \Sigma_{11} \cdot \dots \cdot \Sigma_{1k_1}$ с начальным условием (v_1, \dots, v_1) метка $\mu(x_1^{(k_1)})$ (и только она) не содержится в области определения $F_{\widehat{\Sigma}'_1}$ — минимального правила разметки μ и при модификации слоя $\widehat{\Sigma}'_1 = \Sigma_{11} \cdot \dots \cdot \Sigma_{1k_1}$ до сети $\widehat{\Sigma}_1 = \Sigma_{11} \cdot \dots \cdot \Sigma_{1k_1} \cdot \Sigma_l^{(1,l)} \cdot \Sigma_1^{(1,l)} \cdot \Sigma_m^{(l,m)}$; свободная 2-разметка η сети $\widehat{\Sigma}'_1$ с начальными условиями (v_{11}, \dots, v_{1n}) и (v_{21}, \dots, v_{2n}) свободным образом продолжается до свободной 2-разметки η сети $\widehat{\Sigma}_1$ с любыми условиями $\eta_1(x_1^{(k_1)}) = w_{11}$ и $\eta_2(x_1^{(k_1)}) = w_{21}$, при этом:

- $\eta_1(x_2^{(\widehat{k}_1)}), \dots, \eta_1(x_n^{(\widehat{k}_1)}) \in Y_1$ и $\eta_2(x_2^{(\widehat{k}_1)}), \dots, \eta_2(x_n^{(\widehat{k}_1)}) \in Y_2$;
- метки $\eta_1(x_m^{(\widehat{k}_1)})$ и $\eta_2(x_m^{(\widehat{k}_1)})$, независимо от условий $\eta_1(x_1^{(\widehat{k}_1)}) = w_{11}$ и $\eta_2(x_1^{(\widehat{k}_1)}) = w_{21}$, не содержатся в области определения $G_{\widehat{\Sigma}_1}$ — минимального правила 2-разметки η сети $\widehat{\Sigma}_1$.

Докажем корректность действий, выполняемых на шаге с номером $s \in \{2, \dots, n-2\}$. Пусть первые $(s-1)$ слоёв канонического представления сети Σ уже модифицированы таким образом, что сеть

$$\widehat{\Sigma}_{s-1} = (\Sigma_{11} \cdot \dots \cdot \Sigma_{1\widehat{k}_1}) \cdot \dots \cdot (\Sigma_{(s-1)1} \cdot \dots \cdot \Sigma_{(s-1)\widehat{k}_{s-1}})$$

допускает свободную 2-разметку $\eta = (\eta_1, \eta_2)$ при любых условиях

$$\eta_1(x_1^{(0)}) = v_{11}, \dots, \eta_1(x_n^{(0)}) = v_{1n}, \eta_1(x_1^{(\widehat{k}_1)}) = w_{11}, \dots, \eta_1(x_{s-1}^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})}) = w_{1s-1}, \quad (2)$$

$$\eta_2(x_1^{(0)}) = v_{21}, \dots, \eta_2(x_n^{(0)}) = v_{2n}, \eta_2(x_1^{(\widehat{k}_1)}) = w_{21}, \dots, \eta_2(x_{s-1}^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})}) = w_{2s-1}, \quad (3)$$

при этом:

- $\eta_1(x_s^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})}), \dots, \eta_1(x_n^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})}) \in Y_1$ и $\eta_2(x_s^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})}), \dots, \eta_2(x_n^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})}) \in Y_2$;
- среди $x_s^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})}, \dots, x_n^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})}$ существует вершина $x_i^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})}$, метки которой $\eta_1(x_i^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})})$ и $\eta_2(x_i^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})})$, независимо от условий (2) и (3), не содержатся в области определения $G_{\widehat{\Sigma}_{s-1}}$ — минимального правила 2-разметки η сети $\widehat{\Sigma}_{s-1}$.

Тогда для свободной разметки μ сети $\widehat{\Sigma}_{s-1}$ с условиями

$$\mu(x_1^{(0)}) = v_1, \dots, \mu(x_n^{(0)}) = v_1, \mu(x_1^{(\widehat{k}_1)}) = v_1, \dots, \mu(x_{s-1}^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})}) = v_1$$

метка $\mu(x_i^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})})$ также не содержится в области определения минимального правила $F_{\widehat{\Sigma}_{s-1}}$ и при продолжении разметки μ сети $\widehat{\Sigma}_{s-1}$ свободным образом до разметки сети

$$\widehat{\Sigma}'_s = (\Sigma_{11} \cdot \dots \cdot \Sigma_{1\widehat{k}_1}) \cdot \dots \cdot (\Sigma_{(s-1)1} \cdot \dots \cdot \Sigma_{(s-1)\widehat{k}_{s-1}}) \cdot (\Sigma_{s1} \cdot \dots \cdot \Sigma_{s\widehat{k}_s}),$$

согласно лемме 5 из [1], среди вершин $x_s^{(\widehat{k}_1+\dots+\widehat{k}_{s-1}+k_s)}, \dots, x_n^{(\widehat{k}_1+\dots+\widehat{k}_{s-1}+k_s)}$ существует такая вершина $x_j^{(\widehat{k}_1+\dots+\widehat{k}_{s-1}+k_s)}$, что метка $\mu(x_j^{(\widehat{k}_1+\dots+\widehat{k}_{s-1}+k_s)})$ не содержится в области определения $F_{\widehat{\Sigma}'_s}$ — минимального правила разметки μ сети $\widehat{\Sigma}'_s$.

Поскольку разметка μ по построению является свободной разметкой сети $\widehat{\Sigma}'_s$ с условиями

$$\mu(x_1^{(0)}) = v_1, \dots, \mu(x_n^{(0)}) = v_1, \mu(x_1^{(\widehat{k}_1)}) = v_1, \dots, \mu(x_{s-1}^{(\widehat{k}_1+\dots+\widehat{k}_{s-1})}) = v_1,$$

то, согласно следствию 12 из [1], для продолжения свободной 2-разметки η сети $\widehat{\Sigma}'_s$ с условиями (2) и (3) выполнены следующие условия:

- $\eta_1(x_s^{(\widehat{k}_1+\dots+\widehat{k}_{s-1}+k_s)}), \dots, \eta_1(x_n^{(\widehat{k}_1+\dots+\widehat{k}_{s-1}+k_s)}) \in Y_1$;
- $\eta_2(x_s^{(\widehat{k}_1+\dots+\widehat{k}_{s-1}+k_s)}), \dots, \eta_2(x_n^{(\widehat{k}_1+\dots+\widehat{k}_{s-1}+k_s)}) \in Y_2$;
- метки $\eta_1(x_j^{(\widehat{k}_1+\dots+\widehat{k}_{s-1}+k_s)})$ и $\eta_2(x_j^{(\widehat{k}_1+\dots+\widehat{k}_{s-1}+k_s)})$, независимо от условий (2) и (3), не содержатся в области определения $G_{\widehat{\Sigma}'_s}$ — минимального правила 2-разметки η сети $\widehat{\Sigma}'_s$.

В каждом из возможных вариантов модификации сети $\widehat{\Sigma}'_s$ свободная 2-разметка η сети $\widehat{\Sigma}'_s$ с условиями (2) и (3) свободным образом продолжается до свободной 2-разметки η сети

$$\widehat{\Sigma}_s = (\Sigma_{11} \cdot \dots \cdot \Sigma_{1\widehat{k}_1}) \cdot \dots \cdot (\Sigma_{(s-1)1} \cdot \dots \cdot \Sigma_{(s-1)\widehat{k}_{s-1}}) \cdot (\Sigma_{s1} \cdot \dots \cdot \Sigma_{s\widehat{k}_s})$$

с любыми условиями $\eta_1(x_s^{(\widehat{k}_1+\dots+\widehat{k}_s)}) = w_{1s}$ и $\eta_2(x_s^{(\widehat{k}_1+\dots+\widehat{k}_s)}) = w_{2s}$, при этом:

- $\eta_1(x_{s+1}^{(\widehat{k}_1+\dots+\widehat{k}_s)}), \dots, \eta_1(x_n^{(\widehat{k}_1+\dots+\widehat{k}_s)}) \in Y_1$ и $\eta_2(x_{s+1}^{(\widehat{k}_1+\dots+\widehat{k}_s)}), \dots, \eta_2(x_n^{(\widehat{k}_1+\dots+\widehat{k}_s)}) \in Y_2$;
- метки $\eta_1(x_m^{(\widehat{k}_1+\dots+\widehat{k}_s)})$ и $\eta_2(x_m^{(\widehat{k}_1+\dots+\widehat{k}_s)})$, независимо от условий $\eta_1(x_s^{(\widehat{k}_1+\dots+\widehat{k}_s)}) = w_{1s}$ и $\eta_2(x_s^{(\widehat{k}_1+\dots+\widehat{k}_s)}) = w_{2s}$, не содержатся в области определения $G_{\widehat{\Sigma}_s}$ — минимального правила 2-разметки η сети $\widehat{\Sigma}_s$.

В завершение проведем обоснование корректности действий, выполняемых на шаге с номером $(n-1)$. Пусть первые $(n-2)$ слоев канонического представления сети Σ уже модифицированы таким образом, что сеть

$$\widehat{\Sigma}_{n-2} = (\Sigma_{11} \cdot \dots \cdot \Sigma_{1\widehat{k}_1}) \cdot \dots \cdot (\Sigma_{(n-2)1} \cdot \dots \cdot \Sigma_{(n-2)\widehat{k}_{n-2}}),$$

допускает свободную 2-разметку $\eta = (\eta_1, \eta_2)$ при любых условиях

$$\eta_1(x_1^{(0)}) = v_{11}, \dots, \eta_1(x_n^{(0)}) = v_{1n}, \eta_1(x_1^{(\widehat{k}_1)}) = w_{11}, \dots, \eta_1(x_{n-2}^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})}) = w_{1n-2}, \quad (4)$$

$$\eta_2(x_1^{(0)}) = v_{21}, \dots, \eta_2(x_n^{(0)}) = v_{2n}, \eta_2(x_1^{(\widehat{k}_1)}) = w_{21}, \dots, \eta_2(x_{n-2}^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})}) = w_{2n-2}, \quad (5)$$

при этом:

- $\eta_1(x_{n-1}^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})}), \eta_1(x_n^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})}) \in Y_1$ и $\eta_2(x_{n-1}^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})}), \eta_2(x_n^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})}) \in Y_2$;
- среди $x_{n-1}^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})}, x_n^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})}$ существует вершина $x_i^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})}$, метки которой $\eta_1(x_i^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})})$ и $\eta_2(x_i^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})})$, независимо от условий (4) и (5), не содержатся в области определения $G_{\widehat{\Sigma}_{n-2}}$ — минимального правила 2-разметки η сети $\widehat{\Sigma}_{n-2}$.

Тогда для свободной разметки μ сети $\widehat{\Sigma}_{n-2}$ с условиями

$$\mu(x_1^{(0)}) = v_1, \dots, \mu(x_n^{(0)}) = v_1, \mu(x_1^{(\widehat{k}_1)}) = v_1, \dots, \mu(x_{n-2}^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})}) = v_1,$$

согласно сделанному предположению, метка $\mu(x_i^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})})$ не содержится в области определения минимального правила $F_{\widehat{\Sigma}_{n-2}}$ и при продолжении разметки μ сети $\widehat{\Sigma}_{n-2}$ свободным образом до разметки сети

$$\widehat{\Sigma}'_{n-1} = (\Sigma_{11} \cdot \dots \cdot \Sigma_{1\widehat{k}_1}) \cdot \dots \cdot (\Sigma_{(n-2)1} \cdot \dots \cdot \Sigma_{(n-2)\widehat{k}_{n-2}}) \cdot (\Sigma_{(n-1)1} \cdot \dots \cdot \Sigma_{(n-1)k_{n-1}}),$$

согласно лемме 5 из [1], среди вершин $x_{n-1}^{(\widehat{k}_1+\dots+\widehat{k}_{n-2}+k_{n-1})}, x_n^{(\widehat{k}_1+\dots+\widehat{k}_{n-2}+k_{n-1})}$ существует такая вершина $x_j^{(\widehat{k}_1+\dots+\widehat{k}_{n-2}+k_{n-1})}$, что метка $\mu(x_j^{(\widehat{k}_1+\dots+\widehat{k}_{n-2}+k_{n-1})})$ не содержится в области определения минимального правила $F_{\widehat{\Sigma}'_{n-1}}$.

Поскольку разметка μ по построению является свободной разметкой сети $\widehat{\Sigma}'_{n-1}$ с условиями

$$\mu(x_1^{(0)}) = v_1, \dots, \mu(x_n^{(0)}) = v_1, \mu(x_1^{(\widehat{k}_1)}) = v_1, \dots, \mu(x_{n-2}^{(\widehat{k}_1+\dots+\widehat{k}_{n-2})}) = v_1,$$

то, согласно следствию 12 из [1], для продолжения свободной 2-разметки η сети $\widehat{\Sigma}'_{n-1}$ с условиями (4) и (5) выполнены следующие условия:

- $\eta_1(x_{n-1}^{(\widehat{k}_1+\dots+\widehat{k}_{n-1}+k_{n-1})}), \eta_1(x_n^{(\widehat{k}_1+\dots+\widehat{k}_{n-2}+k_{n-1})}) \in Y_1$;
- $\eta_2(x_{n-1}^{(\widehat{k}_1+\dots+\widehat{k}_{n-2}+k_{n-1})}), \eta_2(x_n^{(\widehat{k}_1+\dots+\widehat{k}_{n-2}+k_{n-1})}) \in Y_2$;
- метки $\eta_1(x_j^{(\widehat{k}_1+\dots+\widehat{k}_{n-2}+k_{n-1})})$ и $\eta_2(x_j^{(\widehat{k}_1+\dots+\widehat{k}_{n-2}+k_{n-1})})$, независимо от условий (4) и (5), не содержатся в области определения $G_{\widehat{\Sigma}'_{n-1}}$ — минимального правила 2-разметки η сети $\widehat{\Sigma}'_{n-1}$.

В каждом из возможных вариантов модификации сети $\widehat{\Sigma}'_{n-1}$ свободная 2-разметка η сети $\widehat{\Sigma}'_{n-1}$ с условиями (4) и (5) свободным образом продолжается до свободной 2-разметки η сети

$$\widehat{\Sigma} = (\Sigma_{11} \cdot \dots \cdot \Sigma_{1\widehat{k}_1}) \cdot \dots \cdot (\Sigma_{(n-2)1} \cdot \dots \cdot \Sigma_{(n-2)\widehat{k}_{n-2}}) \cdot (\Sigma_{(n-1)1} \cdot \dots \cdot \Sigma_{(n-1)\widehat{k}_{n-1}})$$

с произвольными условиями $\eta_1(x_{n-1}^{(\widehat{k}_1+\dots+\widehat{k}_{n-1})}) = w_{1n-1}$, $\eta_1(x_n^{(\widehat{k}_1+\dots+\widehat{k}_{n-1})}) = w_{1n}$ и $\eta_2(x_{n-1}^{(\widehat{k}_1+\dots+\widehat{k}_{n-1})}) = w_{2n-1}$, $\eta_2(x_n^{(\widehat{k}_1+\dots+\widehat{k}_{n-1})}) = w_{2n}$.

Таким образом, в результате работы алгоритма каноническое представление исходной сети Σ модифицировано до «почти» канонического представления

$$(\Sigma_{11} \cdot \dots \cdot \Sigma_{1\widehat{k}_1}) \cdot \dots \cdot (\Sigma_{(n-1)1} \cdot \dots \cdot \Sigma_{(n-1)\widehat{k}_{n-1}})$$

новой биективной сети $\widehat{\Sigma}$, сложность которой не превосходит $\|\Sigma\| + 6n - 7$. При этом построенная сеть $\widehat{\Sigma}$ допускает свободную 2-разметку с произвольными ограничениями

$\begin{pmatrix} v_{11} & \cdots & v_{1n} \\ w_{11} & \cdots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \cdots & v_{2n} \\ w_{21} & \cdots & w_{2n} \end{pmatrix}$ из \mathbb{N} . Поскольку $\|\widehat{\Sigma}\| \leq \|\Sigma\| + 6n - 7$, для проведения свободной 2-разметки сети $\widehat{\Sigma}$ с произвольными ограничениями $\begin{pmatrix} v_{11} & \cdots & v_{1n} \\ w_{11} & \cdots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \cdots & v_{2n} \\ w_{21} & \cdots & w_{2n} \end{pmatrix}$ из \mathbb{N} потребуется не более чем $2\|\Sigma\| + 14n - 14$ различных меток. Значит, при выборе любого множества Ω , мощность которого больше $2\|\Sigma\| + 14n - 14$, можно считать, что сеть $\widehat{\Sigma}$ допускает свободную 2-разметку элементами множества Ω при произвольных ограничениях $\begin{pmatrix} v_{11} & \cdots & v_{1n} \\ w_{11} & \cdots & w_{1n} \end{pmatrix}$ и $\begin{pmatrix} v_{21} & \cdots & v_{2n} \\ w_{21} & \cdots & w_{2n} \end{pmatrix}$ из Ω . Последнее утверждение, согласно теореме 7, равносильно 2-транзитивности сети $\widehat{\Sigma}$ для множества Ω . ■

Следствие 7. Для любого $n \geq 2$ существует сеть $\widehat{\Sigma}$ ширины n и веса $6n - 7$, которая 2-транзитивна для всех множеств, мощность которых больше чем $14n - 14$.

4. k -Транзитивность сетей

Определение 26. Биактивную сеть Σ будем называть k -транзитивной для множества Ω , если множество отображений $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ является k -транзитивным.

Как уже было отмечено, аппарат разметки сетей позволяет проверять не только транзитивность сети, но и более сложное свойство k -транзитивности при $k \geq 2$. При этом понятия аппарата k -разметки биактивных сетей и основные результаты, полученные с их помощью, являются очевидным обобщением соответствующих понятий и результатов для 2-разметки. Поэтому далее приведены только необходимые определения и точные формулировки основных результатов.

Определение 27. Произвольный набор $\mu = (\mu_1, \dots, \mu_k)$ разметок сети Σ будем называть k -разметкой сети Σ . При этом метки разметок μ_1, \dots, μ_k будем называть метками k -разметки μ . Пусть $F : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ — частично определённое отображение. Тогда k -разметку $\mu = (\mu_1, \dots, \mu_k)$ сети Σ будем называть *правильной относительно F* , если каждая из разметок μ_1, \dots, μ_k является правильной относительно F , а отображение F будем называть *правилом k -разметки μ* .

Определение 28. Пусть μ — k -разметка сети Σ с правилом F и при этом никакое сужение частичного отображения F не является правилом k -разметки μ . Тогда будем говорить, что F является *минимальным правилом k -разметки μ* . Правильную k -разметку μ будем называть *непротиворечивой*, если её минимальное правило является непротиворечивым отображением.

Пусть $\eta = (\eta_1, \dots, \eta_k)$ и $\mu = (\mu_1, \dots, \mu_k)$ — k -разметки сети Σ и для отображения σ выполняются соотношения $\sigma \circ \eta_1 = \mu_1, \dots, \sigma \circ \eta_k = \mu_k$. Будем обозначать это условие как $\sigma : \eta \rightarrow \mu$.

Определение 29. Правильную k -разметку η сети Σ с начальным условием $(v_{11}, \dots, v_{1n}), \dots, (v_{k1}, \dots, v_{kn})$ будем называть *свободной*, если для любой правильной k -разметки μ сети Σ с начальным условием $(v_{11}, \dots, v_{1n}), \dots, (v_{k1}, \dots, v_{kn})$ существует отображение σ_μ , удовлетворяющее условию $\sigma_\mu : \eta \rightarrow \mu$.

При $k \geq 2$ естественным образом определяются *процедуры последовательного и параллельного свободного продолжения разметки*, относительно которых сохраняются основные результаты.

Теорема 14. Пусть k -разметка μ' получена в результате последовательного свободного продолжения начальной k -разметки

$$\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}, \dots, \mu_k(x_1^{(0)}) = v_{k1}, \dots, \mu_k(x_n^{(0)}) = v_{kn}$$

и правила F относительно сети Σ , а k -разметка μ'' получена в результате параллельного свободного продолжения начальной k -разметки

$$\mu_1(x_1^{(0)}) = v_{11}, \dots, \mu_1(x_n^{(0)}) = v_{1n}, \dots, \mu_k(x_1^{(0)}) = v_{k1}, \dots, \mu_k(x_n^{(0)}) = v_{kn}$$

и правила F относительно сети Σ . Тогда k -разметки μ' и μ'' отличаются только обратимой заменой меток.

Теорема 15. Пусть k -разметка η получена в результате параллельного свободного продолжения начальной k -разметки

$$\eta_1(x_1^{(0)}) = v_{11}, \dots, \eta_1(x_n^{(0)}) = v_{1n}, \dots, \eta_k(x_1^{(0)}) = v_{k1}, \dots, \eta_k(x_n^{(0)}) = v_{kn}$$

и пустого правила G относительно сети Σ . Тогда η — свободная k -разметка сети Σ , а отображение $G_{\Sigma, \eta}$ — её минимальное правило.

Теорема 16. Пусть η — свободная k -разметка сети Σ , μ — правильная k -разметка сети Σ и возможно определить отображение σ_μ по правилу

$$\sigma_\mu(\eta_i(x_i^{(0)})) = \mu_i(x_i^{(0)}), \dots, \sigma_\mu(\eta_k(x_i^{(0)})) = \mu_k(x_i^{(0)}), \quad i \in \{1, \dots, n\}.$$

Тогда отображение σ_μ допускает продолжение, удовлетворяющее условию $\sigma_\mu: \eta \rightarrow \mu$.

Следствие 8. В условиях теоремы, если G и F — минимальные правила k -разметок η и μ соответственно, то при всех допустимых $z_i, z_j \in \mathbb{N}$ выполняется равенство $\sigma_\mu(G(z_i, z_j)) = F(\sigma_\mu(z_i), \sigma_\mu(z_j))$.

Определение 30. k -разметку $\mu = (\mu_1, \dots, \mu_k)$ сети Σ будем называть k -разметкой сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$, если μ_1, \dots, μ_k — разметки сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$ соответственно. При этом будем говорить, что сеть Σ допускает k -разметку μ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$. Нетривиальной будем называть такую k -разметку $\mu = (\mu_1, \dots, \mu_k)$ сети Σ , у которой μ_1, \dots, μ_k — различные разметки сети Σ .

Теорема 17. Пусть Σ — биективная сеть ширины n и Ω — множество мощности строго больше чем $k\|\Sigma\|$. Тогда следующие утверждения эквивалентны:

- 1) сеть Σ является k -транзитивной для множества Ω ;
- 2) сеть Σ допускает нетривиальную правильную непротиворечивую k -разметку элементами из Ω при любых ограничениях $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$.

Следствие 9. Для биективной сети Σ следующие утверждения эквивалентны:

- 1) сеть Σ является k -транзитивной для некоторого множества, мощность которого строго больше чем $k\|\Sigma\| + kn$;
- 2) сеть Σ является k -транзитивной для произвольного множества, мощность которого строго больше чем $k\|\Sigma\| + kn$.

Для k -разметки аналогичным образом определяется процедура устранения противоречий, относительно которой сохраняется основной результат.

Лемма 3. Пусть η — произвольная k -разметка сети Σ , μ — правильная непротиворечивая k -разметка сети Σ и при этом существует отображение $\sigma_\mu: \eta \rightarrow \mu$. Тогда для любой k -разметки $\tilde{\eta}$, полученной из k -разметки η устранением противоречий, также выполняется условие $\sigma_\mu: \tilde{\eta} \rightarrow \mu$.

Следствие 10. Пусть η — произвольная 2-разметка сети Σ . Тогда правильная непротиворечивая k -разметка $\tilde{\eta}$, полученная из k -разметки η устранением противоречий, определена однозначно с точностью до обратимого переобозначения меток.

Определение 31. Правильную непротиворечивую k -разметку η сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$ будем называть *свободной k -разметкой* сети Σ с ограничениями, если для любой правильной непротиворечивой k -разметки μ сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$ существует отображение $\sigma_\mu: \eta \rightarrow \mu$.

Теорема 18. Если сеть Σ допускает нетривиальную правильную непротиворечивую k -разметку с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$, то существует свободная k -разметка сети Σ с указанными ограничениями.

Теорема 19. Пусть имеются свободная k -разметка η сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$, а также правильная непротиворечивая 2-разметка μ сети Σ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}, \dots, \begin{pmatrix} \bar{v}_{k1} & \dots & \bar{v}_{kn} \\ \bar{w}_{k1} & \dots & \bar{w}_{kn} \end{pmatrix}$, при которых возможно определить отображение σ_μ по правилу

$$\sigma_\mu(v_{1i}) = \bar{v}_{1i}, \sigma_\mu(w_{1i}) = \bar{w}_{1i}, \dots, \sigma_\mu(v_{ki}) = \bar{v}_{ki}, \sigma_\mu(w_{ki}) = \bar{w}_{ki}, i \in \{1, \dots, n\}.$$

Тогда отображение σ_μ допускает продолжение, удовлетворяющее условию $\sigma_\mu: \eta \rightarrow \mu$.

Следствие 11. В условиях теоремы, если G и F — минимальные правила разметок η и μ соответственно, то при всех допустимых $z_i, z_j \in \mathbb{N}$ выполняется равенство $\sigma_\mu(G(z_i, z_j)) = F(\sigma_\mu(z_i), \sigma_\mu(z_j))$.

Сформулируем и докажем один из немногих результатов из области k -разметки, доказательство которого существенным образом отличается от соответствующего доказательства в случае 2-разметки.

Теорема 20. Сеть Σ допускает нетривиальные правильные непротиворечивые k -разметки при всех возможных ограничениях $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$ из \mathbb{N} в том и только в том случае, когда сеть Σ допускает нетривиальные правильные непротиворечивые k -разметки при всех возможных ограничениях $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}, \dots, \begin{pmatrix} \bar{v}_{k1} & \dots & \bar{v}_{kn} \\ \bar{w}_{k1} & \dots & \bar{w}_{kn} \end{pmatrix}$ из Ω_{k+1} .

Доказательство. Необходимость очевидна. Докажем достаточность. Будем считать, что $\mathbb{N} \setminus \{v_{11}, \dots, v_{k1}, \dots, v_{kn}, w_{11}, \dots, w_{k1}, \dots, w_{kn}\} = \{y_{11}, \dots, y_{k1}, y_{12}, \dots, y_{k2}, \dots\}$. Пусть свободная k -разметка $\eta' = (\eta'_1, \dots, \eta'_k)$ сети Σ получена в результате свободного продолжения начальной k -разметки $\eta'_1(x_1^{(0)}) = v_{11}, \dots, \eta'_1(x_n^{(0)}) = v_{1n}, \dots, \eta'_k(x_1^{(0)}) = v_{k1},$

$\dots, \eta'_k(x_n^{(0)}) = v_{kn}$ с использованием меток $y_{11}, \dots, y_{k1}, y_{12}, \dots, y_{k2}, \dots$. Тогда, согласно теореме 19, для любой правильной непротиворечивой k -разметки μ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}, \dots, \begin{pmatrix} \bar{v}_{k1} & \dots & \bar{v}_{kn} \\ \bar{w}_{k1} & \dots & \bar{w}_{kn} \end{pmatrix}$ из Ω_{k+1} , при которых возможно определить отображение σ_μ по правилу

$$\sigma_\mu(v_{1i}) = \bar{v}_{1i}, \dots, \sigma_\mu(v_{ki}) = \bar{v}_{ki}, \sigma_\mu(w_{1i}) = \bar{w}_{1i}, \dots, \sigma_\mu(w_{ki}) = \bar{w}_{ki}, i \in \{1, \dots, n\},$$

отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu: \eta' \rightarrow \mu$. Заменяя в k -разметке η' метки $\eta'_1(x_1^{(t)}), \dots, \eta'_1(x_n^{(t)}), \dots, \eta'_k(x_1^{(t)}), \dots, \eta'_k(x_n^{(t)})$ на $w_{11}, \dots, w_{1n}, \dots, w_{k1}, \dots, w_{kn}$ соответственно, получим k -разметку η'' сети Σ с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$, и при этом для любой правильной непротиворечивой k -разметки μ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}, \dots, \begin{pmatrix} \bar{v}_{k1} & \dots & \bar{v}_{kn} \\ \bar{w}_{k1} & \dots & \bar{w}_{kn} \end{pmatrix}$, при которых возможно определить отображение σ_μ по правилу

$$\sigma_\mu(v_{1i}) = \bar{v}_{1i}, \dots, \sigma_\mu(v_{ki}) = \bar{v}_{ki}, \sigma_\mu(w_{1i}) = \bar{w}_{1i}, \dots, \sigma_\mu(w_{ki}) = \bar{w}_{ki}, i \in \{1, \dots, n\},$$

отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu: \eta'' \rightarrow \mu$. Проведём процедуру устранения противоречий в k -разметке η'' с уточнениями:

- если при устранении противоречия требуется отождествить метки v_{ri} и y_{sj} , то будем заменять метку y_{sj} на v_{ri} ;
- если при устранении противоречия требуется отождествить метки w_{ri} и y_{sj} , то будем заменять метку y_{sj} на w_{ri} .

Пусть η — правильная непротиворечивая k -разметка сети Σ , полученная из k -разметки η'' устранением противоречий. Тогда, согласно сделанным уточнениям, метки $\eta_1(x_1^{(0)}), \dots, \eta_k(x_1^{(0)}), \dots, \eta_1(x_n^{(0)}), \dots, \eta_k(x_n^{(0)}), \eta_1(x_1^{(t)}), \dots, \eta_k(x_1^{(t)}), \dots, \eta_1(x_n^{(t)}), \dots, \eta_k(x_n^{(t)})$ содержатся в множестве $\{v_{11}, \dots, v_{k1}, \dots, v_{1n}, \dots, v_{kn}, w_{11}, \dots, w_{k1}, \dots, w_{1n}, \dots, w_{kn}\}$. При этом, согласно лемме 3, для любой правильной непротиворечивой 2-разметки μ с ограничениями $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}, \dots, \begin{pmatrix} \bar{v}_{k1} & \dots & \bar{v}_{kn} \\ \bar{w}_{k1} & \dots & \bar{w}_{kn} \end{pmatrix}$ из Ω_{k+1} , при которых возможно определить отображение σ_μ по правилу

$$\sigma_\mu(v_{1i}) = \bar{v}_{1i}, \dots, \sigma_\mu(v_{ki}) = \bar{v}_{ki}, \sigma_\mu(w_{1i}) = \bar{w}_{1i}, \dots, \sigma_\mu(w_{ki}) = \bar{w}_{ki}, i \in \{1, \dots, n\},$$

отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu: \eta \rightarrow \mu$.

Методом от противного покажем, что правильная непротиворечивая k -разметка η будет k -разметкой с ограничениями $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$. Согласно уточнениям, в разметке η могли появиться противоречия только следующих типов:

- $\eta_r(x_i^{(0)}) = v_{sj} \neq v_{ri}$;
- $\eta_r(x_i^{(0)}) = w_{sj} \neq v_{ri}$;
- $\eta_r(x_i^{(t)}) = v_{sj} \neq w_{ri}$;
- $\eta_r(x_i^{(t)}) = w_{sj} \neq w_{ri}$.

Разберём подслучай $\eta_1(x_i^{(0)}) = v_{1j} \neq v_{1i}$, относящийся к первому типу противоречий. Предварительно докажем два вспомогательных утверждения.

Лемма 4. Пусть $\mathbf{v}_1 = (v_{11}, \dots, v_{1n}), \dots, \mathbf{v}_k = (v_{k1}, \dots, v_{kn}) \in \mathbb{N}^n$ — различные векторы. Тогда существует такое отображение $\sigma: \mathbb{N} \rightarrow \Omega_k$, при котором все векторы $\sigma(\mathbf{v}_1) = (\sigma(v_{11}), \dots, \sigma(v_{1n})), \dots, \sigma(\mathbf{v}_k) = (\sigma(v_{k1}), \dots, \sigma(v_{kn}))$ различны.

Доказательство. Не ограничивая общности, будем считать, что векторы $\mathbf{v}_1, \dots, \mathbf{v}_k$ не содержат элементов из множества Ω_k .

Докажем индукцией по s существование отображения $\sigma_s: \mathbb{N} \rightarrow \Omega_s$, при котором множество $\{\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_k)\}$ содержит не менее s различных векторов.

База при $\sigma_1: \mathbb{N} \rightarrow \{1\}$ очевидна — множество векторов $\{\sigma_1(\mathbf{v}_1), \dots, \sigma_1(\mathbf{v}_k)\}$ является одноэлементным.

Предположим, что построено такое отображение $\sigma_s: \mathbb{N} \rightarrow \Omega_s$, при котором множество $\{\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_k)\}$ содержит $r \geq s$ различных векторов. Не ограничивая общности, будем считать, что $\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_r)$ — все различные элементы множества $\{\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_k)\}$. Если $r = k$, то утверждение доказано, в противном случае, не ограничивая общности, можно считать, что $\sigma_s(\mathbf{v}_r) = \sigma_s(\mathbf{v}_{r+1})$. Поскольку $\mathbf{v}_r \neq \mathbf{v}_{r+1}$, то $v_{rj} \neq v_{r+1j}$ для некоторого j ; полагая

$$\sigma_{s+1}(x) = \begin{cases} s+1, & \text{если } x = v_{r+1j}, \\ \sigma_s(x), & \text{если } x \neq v_{r+1j}, \end{cases}$$

получим, что множество $\{\sigma_{s+1}(\mathbf{v}_1), \dots, \sigma_{s+1}(\mathbf{v}_k)\}$ содержит как минимум $(r+1) \geq (s+1)$ различных векторов: $\sigma_{s+1}(\mathbf{v}_1), \dots, \sigma_{s+1}(\mathbf{v}_{r+1})$. ■

Лемма 5. Пусть $\mathbf{v}_1 = (v_{11}, \dots, v_{1n}), \dots, \mathbf{v}_k = (v_{k1}, \dots, v_{kn}) \in \mathbb{N}^n$ — различные векторы и $\mathbf{w}_1 = (w_{11}, \dots, w_{1n}), \dots, \mathbf{w}_k = (w_{k1}, \dots, w_{kn}) \in \mathbb{N}^n$ — различные векторы. Тогда существует такое отображение $\sigma: \mathbb{N} \rightarrow \Omega_k$, при котором все векторы $\sigma(\mathbf{v}_1), \dots, \sigma(\mathbf{v}_k)$ различны и все векторы $\sigma(\mathbf{w}_1), \dots, \sigma(\mathbf{w}_k)$ различны.

Доказательство. Не ограничивая общности, будем считать, что векторы $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{w}_1, \dots, \mathbf{w}_k$ не содержат элементов из множества Ω_k .

Докажем индукцией по s существование отображения $\sigma_s: \mathbb{N} \rightarrow \Omega_s$, при котором каждое из множеств $\{\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_k)\}$ и $\{\sigma_s(\mathbf{w}_1), \dots, \sigma_s(\mathbf{w}_k)\}$ содержит не менее s различных векторов.

База при $\sigma_1: \mathbb{N} \rightarrow \{1\}$ очевидна — множества векторов $\{\sigma_1(\mathbf{v}_1), \dots, \sigma_1(\mathbf{v}_k)\}$ и $\{\sigma_1(\mathbf{w}_1), \dots, \sigma_1(\mathbf{w}_k)\}$ являются одноэлементными.

Предположим, что построено такое отображение $\sigma_s: \mathbb{N} \rightarrow \Omega_s$, при котором $|\{\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_k)\}| \geq s$ и $|\{\sigma_s(\mathbf{w}_1), \dots, \sigma_s(\mathbf{w}_k)\}| \geq s$. Если $|\{\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_k)\}| = k$ или $|\{\sigma_s(\mathbf{w}_1), \dots, \sigma_s(\mathbf{w}_k)\}| = k$, то дальнейшее построение отображения σ осуществляется согласно доказательству леммы 4, в частности, если $|\{\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_k)\}| = |\{\sigma_s(\mathbf{w}_1), \dots, \sigma_s(\mathbf{w}_k)\}| = k$, то утверждение доказано. В противном случае достаточно рассмотреть два возможных случая.

С л у ч а й 1. Если существует такой $a \in \mathbb{N}$, при котором для отображения

$$\sigma_{s+1}(x) = \begin{cases} s+1, & \text{если } x = a, \\ \sigma_s(x), & \text{если } x \neq a, \end{cases}$$

выполняются оба неравенства $|\{\sigma_{s+1}(\mathbf{v}_1), \dots, \sigma_{s+1}(\mathbf{v}_k)\}| > |\{\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_k)\}|$ и $|\{\sigma_{s+1}(\mathbf{w}_1), \dots, \sigma_{s+1}(\mathbf{w}_k)\}| > |\{\sigma_s(\mathbf{w}_1), \dots, \sigma_s(\mathbf{w}_k)\}|$, то отображение σ_{s+1} является искомым.

С л у ч а й 2. Если при любом $a \in \mathbb{N}$ для отображения

$$\sigma_{s+1}(x) = \begin{cases} s+1, & \text{если } x = a, \\ \sigma_s(x), & \text{если } x \neq a, \end{cases}$$

выполняется одно из равенств $|\{\sigma_{s+1}(\mathbf{v}_1), \dots, \sigma_{s+1}(\mathbf{v}_k)\}| = |\{\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_k)\}|$ или $|\{\sigma_{s+1}(\mathbf{w}_1), \dots, \sigma_{s+1}(\mathbf{w}_k)\}| = |\{\sigma_s(\mathbf{w}_1), \dots, \sigma_s(\mathbf{w}_k)\}|$, то выберем произвольное $a \in \mathbb{N}$, для которого выполняется неравенство

$$|\{\sigma_{s+1}(\mathbf{v}_1), \dots, \sigma_{s+1}(\mathbf{v}_k)\}| > |\{\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_k)\}|,$$

и произвольное $b \in \mathbb{N}$, для которого выполняется неравенство

$$|\{\sigma_{s+1}(\mathbf{w}_1), \dots, \sigma_{s+1}(\mathbf{w}_k)\}| > |\{\sigma_s(\mathbf{w}_1), \dots, \sigma_s(\mathbf{w}_k)\}|.$$

В таком случае искомым является отображение

$$\sigma_{s+1}(x) = \begin{cases} s+1, & \text{если } x = a \text{ или } x = b, \\ \sigma_s(x), & \text{если } x \neq a, b, \end{cases}$$

для которого выполняются оба неравенства $|\{\sigma_{s+1}(\mathbf{v}_1), \dots, \sigma_{s+1}(\mathbf{v}_k)\}| > |\{\sigma_s(\mathbf{v}_1), \dots, \sigma_s(\mathbf{v}_k)\}|$ и $|\{\sigma_{s+1}(\mathbf{w}_1), \dots, \sigma_{s+1}(\mathbf{w}_k)\}| > |\{\sigma_s(\mathbf{w}_1), \dots, \sigma_s(\mathbf{w}_k)\}|$. ■

Из последнего результата следует, что для произвольных нетривиальных ограничений $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$ существует отображение $\sigma: \mathbb{N} \rightarrow \Omega_k$, при

котором $\begin{pmatrix} \sigma(v_{11}) & \dots & \sigma(v_{1n}) \\ \sigma(w_{11}) & \dots & \sigma(w_{1n}) \end{pmatrix}, \dots, \begin{pmatrix} \sigma(v_{k1}) & \dots & \sigma(v_{kn}) \\ \sigma(w_{k1}) & \dots & \sigma(w_{kn}) \end{pmatrix}$ — нетривиальные ограничения из Ω_k . Если $\sigma(v_{1j}) = \sigma(v_{1i})$, то корректно переопределить σ так, что $\sigma(v_{1i}) = k+1$.

В результате будут получены нетривиальные ограничения $\begin{pmatrix} \sigma(v_{11}) & \dots & \sigma(v_{1n}) \\ \sigma(w_{11}) & \dots & \sigma(w_{1n}) \end{pmatrix},$

$\dots, \begin{pmatrix} \sigma(v_{k1}) & \dots & \sigma(v_{kn}) \\ \sigma(w_{k1}) & \dots & \sigma(w_{kn}) \end{pmatrix}$ из Ω_{k+1} . Согласно условию теоремы, существует правильная

непротиворечивая k -разметка μ с ограничениями $\begin{pmatrix} \sigma(v_{11}) & \dots & \sigma(v_{1n}) \\ \sigma(w_{11}) & \dots & \sigma(w_{1n}) \end{pmatrix}, \dots,$

$\begin{pmatrix} \sigma(v_{k1}) & \dots & \sigma(v_{kn}) \\ \sigma(w_{k1}) & \dots & \sigma(w_{kn}) \end{pmatrix}$ из Ω_{k+1} , и при этом отображение σ_μ , определённое по правилу

$\sigma_\mu(v_{1i}) = \bar{v}_{1i}, \sigma_\mu(v_{2i}) = \bar{v}_{2i}, \sigma_\mu(w_{1i}) = \bar{w}_{1i}, \sigma_\mu(w_{2i}) = \bar{w}_{2i}, i \in \{1, \dots, n\}$, продолжается таким образом, что удовлетворяет условию $\sigma_\mu: \eta \rightarrow \mu$. Получили противоречие, поскольку $\sigma_\mu(\eta_1(x_i^{(0)})) = \sigma(v_{1j}) \neq \sigma(v_{1i}) = \mu(x_i^{(0)})$. Отсутствие противоречий всех остальных типов устанавливается аналогичным образом. ■

Следствие 12. Пусть Σ — биективная сеть ширины n и Ω — множество мощности не менее чем $k\|\Sigma\| + kn$. Тогда следующие утверждения эквивалентны:

- 1) сеть Σ является k -транзитивной для множества Ω ;
- 2) сеть Σ допускает нетривиальную правильную непротиворечивую k -разметку элементами множества Ω при любых ограничениях $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ w_{11} & \dots & w_{1n} \end{pmatrix}, \dots, \begin{pmatrix} v_{k1} & \dots & v_{kn} \\ w_{k1} & \dots & w_{kn} \end{pmatrix}$ из множества Ω ;
- 3) сеть Σ допускает нетривиальную правильную непротиворечивую k -разметку элементами множества Ω при любых ограничениях $\begin{pmatrix} \bar{v}_{11} & \dots & \bar{v}_{1n} \\ \bar{w}_{11} & \dots & \bar{w}_{1n} \end{pmatrix}, \dots, \begin{pmatrix} \bar{v}_{k1} & \dots & \bar{v}_{kn} \\ \bar{w}_{k1} & \dots & \bar{w}_{kn} \end{pmatrix}$ из множества $\Omega_{k+1} \subset \Omega$;

- 4) множество преобразований $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ действует транзитивным образом на подмножестве $\Omega_{k+1}^n \subset \Omega^n$.

Замечание 5. Вспомогательные леммы 4 и 5 из доказательства теоремы 20, по всей видимости, сами по себе являются интересными результатами, допускающими эквивалентные формулировки в разных областях дискретной математики. Так, например, лемму 4 можно переформулировать на языке теории графов следующим образом: «хроматическое число графа, содержащего не более $k(k-1)/2$ рёбер, не превышает k ».

В заключение отметим, что приведённый в работе алгоритм построения 2-транзитивной сети на самом деле строит сеть, которая k -транзитивна для всех достаточно больших множеств.

Теорема 21. Пусть Σ — произвольная биективная сеть ширины n . Тогда её модификация $\widehat{\Sigma}$ является k -транзитивной для любого множества Ω , мощность которого больше чем $k\|\Sigma\| + 7k(n-1)$.

Следствие 13. Для любого $n \geq 2$ существует сеть $\widehat{\Sigma}$ ширины n и веса $6n-7$, которая k -транзитивна для всех множеств, мощность которых больше чем $7k(n-1)$.

Автор выражает благодарность А. В. Черемушкину за постановку задачи и внимание к проводимым исследованиям.

ЛИТЕРАТУРА

1. *Чередник И. В.* Один подход к построению транзитивного множества блочных преобразований // Прикладная дискретная математика. 2017. № 38. С. 5–34.
2. *Белоусов В. Д.* Основы теории квазигрупп и луп. М.: Наука, 1967.

REFERENCES

1. *Cherednik I. V.* Odin podhod k postroeniyu tranzitivnogo mnozhestva blochnyh preobrazovanij [One approach to constructing a transitive class of block transformations]. *Prikladnaya Diskretnaya Matematika*, 2017, no. 38, pp. 5–34. (in Russian)
2. *Belousov V. D.* *Osnovy teorii kvazigrupp i lup* [Foundations of the Quasigroups and Loops Theory]. Moscow, Nauka Publ., 1967. (in Russian)

МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

УДК 519.7

КРИПТОАНАЛИЗ ДВУХКАСКАДНОГО КОНЕЧНО-АВТОМАТНОГО ГЕНЕРАТОРА С ФУНКЦИОНАЛЬНЫМ КЛЮЧОМ¹

И. В. Боровкова, И. А. Панкратова, Е. В. Семенова

Национальный исследовательский Томский государственный университет, г. Томск, Россия

Рассматривается криптографический генератор $G = A_1 \cdot A_2$, представляющий собой последовательное соединение двух абстрактных конечных автоматов A_1 и A_2 над полем \mathbb{F}_2 . Ключом генератора является функция f_1 выходов автомата A_1 и, возможно, начальные состояния автоматов. Задача криптоанализа генератора G состоит в определении его ключа по заданному отрезку $\gamma = z(1)z(2)\dots z(l)$ его выходной последовательности. Описаны алгоритмы анализа автомата A_2 в общем случае и для конечно-автоматного генератора (δ, τ) -шагов, позволяющие найти поступающий на вход автомата A_2 прообраз $u(1)\dots u(l)$ последовательности γ . Значения $u(t)$ суть значения функции f_1 на наборах $x(t)$, $t = 1, 2, \dots, l$, где $x(t)$ — состояние автомата A_1 в момент времени t . Если начальное состояние $x(1)$ и класс функций C_1 , которому принадлежит f_1 , известны, то задача поиска функции f_1 сводится к доопределению частичной булевой функции до функции в классе C_1 .

Ключевые слова: *конечный автомат, криптографический генератор, генератор (δ, τ) -шагов, криптоанализ, метод DSS.*

DOI 10.17223/20710410/42/3

CRYPTANALYSIS OF 2-CASCADE FINITE AUTOMATA GENERATOR WITH FUNCTIONAL KEY

I. V. Borovkova, I. A. Pankratova, E. V. Semenova

National Research Tomsk State University, Tomsk, Russia

E-mail: iborovkova95@gmail.com, pank@mail.tsu.ru, katrinevs@mail.ru

A cryptographic generator under consideration is a serial connection $G = A_1 \cdot A_2$ of two finite state machines (finite automata) $A_1 = (\mathbb{F}_2^n, \mathbb{F}_2, g_1, f_1)$ (it is autonomous) and $A_2 = (\mathbb{F}_2, \mathbb{F}_2^m, \mathbb{F}_2, g_2, f_2)$. The key of the generator is the function f_1 and possibly the initial states $x(1), y(1)$ of the automata A_1, A_2 . The cryptanalysis problem for G is the following: given an output sequence $\gamma = z(1)z(2)\dots z(l)$, find the generator's key. Two algorithms for analysis of A_2 are presented, they allow to find a preimage $u(1)\dots u(l)$ of γ in general case and in the case when A_2 is the Moore automaton with the transition function $g_2(u, y) = \neg u g^\delta(y) + u g^\tau(y)$ for some $g : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^m$ and $\delta, \tau \in \mathbb{N}$. This preimage is an input to A_2 and an output from A_1 . The values $u(t)$ equal the values $f_1(x(t))$ where $x(t)$ is the state of A_1 at a time t , $t = 1, 2, \dots, l$. If the

¹Работа поддержана грантом РФФИ, проект № 17-01-00354.

initial state $x(1)$ and a function class C_1 containing f_1 are known, then f_1 can be determined by its specifying in the class C_1 .

Keywords: *finite automaton, cryptographic generator, (δ, τ) -step generator, cryptanalysis, DSS method.*

1. Определение генератора

Рассматривается двухкаскадный конечно-автоматный криптографический генератор $G = A_1 \cdot A_2$, схема которого показана на рис. 1. Генератор представляет собой последовательное соединение автономного автомата $A_1 = (\mathbb{F}_2^n, \mathbb{F}_2, g_1, f_1)$ (с функцией переходов $g_1 : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ и функцией выходов $f_1 : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$) и автомата $A_2 = (\mathbb{F}_2, \mathbb{F}_2^m, \mathbb{F}_2, g_2, f_2)$ (с функцией переходов $g_2 : \mathbb{F}_2 \times \mathbb{F}_2^m \rightarrow \mathbb{F}_2^m$ и функцией выходов $f_2 : \mathbb{F}_2 \times \mathbb{F}_2^m \rightarrow \mathbb{F}_2$).

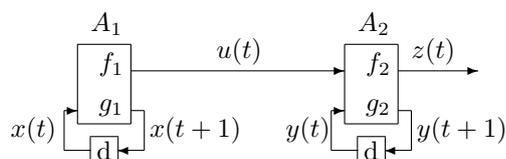


Рис. 1. Схема генератора G

Генератор функционирует в дискретном времени $t = 1, 2, \dots$, в каждый момент t которого автомат A_1 , находясь в состоянии $x(t) \in \mathbb{F}_2^n$, выдаёт выходной символ $u(t) = f_1(x(t))$ и переходит в следующее состояние $x(t+1) = g_1(x(t))$, а автомат A_2 , находясь в состоянии $y(t) \in \mathbb{F}_2^m$, принимает от A_1 символ $u(t)$, выдаёт на выход генератора выходной символ $z(t) = f_2(u(t), y(t))$ и переходит в следующее состояние $y(t+1) = g_2(u(t), y(t))$. Последовательность $u(1) \dots u(l)$, $l \in \mathbb{N}$, выходных символов автомата A_1 называется управляющей последовательностью автомата A_2 , а последовательность $z(1) \dots z(l)$ выходных символов автомата A_2 — выходной последовательностью генератора G . Ключом генератора может быть любое непустое подмножество множества $\{x(1), y(1), f_1, g_1, f_2, g_2\}$.

2. Криптоанализ генератора G

2.1. Основная задача

Задача криптоанализа состоит в определении ключа генератора по его выходной последовательности. Рассмотрим сначала случай, когда ключом служит только функция f_1 , все остальные параметры известны. Как правило, известен ещё и класс функций, которому принадлежит f_1 , потому что выходные функции автоматов в генераторе должны обладать определёнными свойствами: иметь ограниченную сложность задания, полиномиальную вычислимость, достаточную криптографическую стойкость и т. д. Целью данной работы является решение ряда вспомогательных задач для следующей основной задачи.

Задача 1

Дано: $\gamma = z(1) \dots z(l)$ — выходная последовательность генератора; $x(1), y(1)$ — начальные состояния автоматов A_1, A_2 ; g_1 — функция переходов автомата A_1 ; C_1 — класс функций, которому принадлежит функция выходов автомата A_1 ; g_2 и f_2 — функции соответственно переходов и выходов автомата A_2 .

Найти: функцию выходов $f_1 \in C_1$, такую, что $z(t) = f_2(f_1(x(t)), y(t))$ при $x(t+1) = g_1(x(t))$ и $y(t+1) = g_2(f_1(x(t)), y(t))$ для $t = 1, \dots, l$.

Поскольку функция f_1 является ключом генератора, криптоаналитику неизвестна управляющая последовательность $u(1)u(2) \dots$. Знание этой последовательности упрощает решение задачи 1, давая информацию о некоторых значениях функции f_1 , а именно

$$u(t) = f_1(g_1^{t-1}(x(1))), \quad (1)$$

где $g_1^0(x) = x$; $g_1^t(x) = g_1(g_1^{t-1}(x))$, $t = 1, \dots, l$. В связи с этим основная задача 1 распадается на две вспомогательные задачи:

- 1) анализ автомата A_2 — по выходной последовательности генератора G найти управляющие последовательности автомата A_2 ;
- 2) анализ автомата A_1 — по управляющей последовательности на выходе автомата A_1 найти его функцию выходов f_1 .

2.2. Анализ автомата A_2

Обозначим $U(\gamma, y(1))$ множество всех управляющих последовательностей $u(1) \dots u(l)$, отображаемых автоматом $A_2 = (\mathbb{F}_2, \mathbb{F}_2^m, \mathbb{F}_2, g_2, f_2)$ в начальном состоянии $y(1)$ в выходную последовательность $\gamma = z(1) \dots z(l)$, т. е. таких, что

$$f_2(u(t), y(t)) = z(t), \quad y(t+1) = g_2(u(t), y(t)), \quad t = 1, \dots, l. \quad (2)$$

Задача анализа автомата A_2 ставится следующим образом.

Дано: γ — выходная последовательность автомата A_2 ; $y(1)$ — его начальное состояние; g_2, f_2 — функции переходов и выходов.

Найти: множество $U(\gamma, y(1))$.

Для решения этой задачи построим граф, вершины которого расположены по ярусам с номерами $t \in \{1, 2, \dots, l, l+1\}$ и помечены состояниями автомата A_2 , дуги помечены значениями 0 и 1. На первом ярусе — вершина с меткой $y(1)$. Для каждой вершины v с меткой q построенного яруса t , $t = 1, \dots, l$, составляем уравнение $z(t) = f_2(u, q)$ относительно $u \in \{0, 1\}$ и добавляем к вершине v столько потомков на ярусе $t+1$, сколько решений имеет это уравнение. Для каждого пути в графе от вершины первого яруса к вершине $(l+1)$ -го яруса выписываем последовательность меток $u(1) \dots u(l)$ дуг этого пути. По сути, это есть реализация метода DSS (Devide, Solve and Substitute) [1–3]. Более подробно действия описаны в алгоритме 1.

Корректность алгоритма. Пусть c_1, \dots, c_l — последовательность меток дуг некоторого пути от первого к $(l+1)$ -му ярусу, а q_1, \dots, q_{l+1} — последовательность меток вершин этого пути. По построению $q_{t+1} = g_2(c_t, q_t)$, $f_2(c_t, q_t) = z(t)$, т. е. выполнены условия (2) при $u(t) = c_t$, $y(t) = q_t$. Следовательно, $c_1 \dots c_l \in U(\gamma, y(1))$.

Полнота алгоритма. Пусть $u(1) \dots u(l) \in U(\gamma, y(1))$, т. е. выполнены условия (2). Тогда $f_2(u(1), y(1)) = z(1)$ и по построению на ярусе 2 есть вершина v с меткой $q = g_2(u(1), y(1))$, соединённая с вершиной первого яруса дугой с меткой $u(1)$. Положим $y(2) = q$; ввиду того, что $f_2(u(2), y(2)) = z(2)$, на ярусе 3 есть вершина, соединённая с v дугой, помеченной $u(2)$, и т. д. до яруса $(l+1)$. Значит, $u(1) \dots u(l)$ есть последовательность дуг некоторого пути от первого до $(l+1)$ -го яруса.

Алгоритм 1 реализован на языке ЛЯПАС-Т [4, 5]. Граф в программе представляется логическим комплексом L , элементы которого соответствуют вершинам и хранят их метки; для элемента $L[i]$, $i = 0, 1, \dots$, потомками являются элемент $L[2i+1]$ (дуга к нему от $L[i]$ помечена знаком 0) и элемент $L[2i+2]$ (дуга помечена знаком 1); если вершина отсутствует или удаляется, то элементу присваивается специальное значение (-1) .

Алгоритм 1. Анализ автомата A_2

Вход: $\gamma = z(1) \dots z(l)$ — выходная последовательность автомата A_2 ; $y(1)$ — его начальное состояние; g_2, f_2 — функции переходов и выходов.

Выход: множество $U(\gamma, y(1))$.

- 1: На ярусе 1 — одна вершина с меткой $y(1)$.
- 2: Для $t = 1, 2, \dots, l$ строим ярус $t + 1$ по следующим правилам.
- 3: **Если** вершин на ярусе t нет, **то**
- 4: выход из алгоритма с ответом « $y(1)$ не может быть начальным состоянием автомата A_2 ».
- 5: Рассматриваем каждую вершину v на ярусе t ; пусть q — метка вершины v .
- 6: **Если** $z(t) = f_2(0, q)$, **то**
- 7: к вершине v добавляем потомка с меткой $g_2(0, q)$, соединяем v с потомком дугой с меткой 0.
- 8: **Если** $z(t) = f_2(1, q)$, **то**
- 9: к вершине v добавляем потомка с меткой $g_2(1, q)$; соединяем v с потомком дугой с меткой 1.
- 10: **Если** $z(t) \neq f_2(0, q) = f_2(1, q)$, **то**
- 11: потомков у вершины v нет; удаляем вершину v и дуги, ведущие в неё; поднимаемся по ярусам вверх, удаляя по пути все вершины, не имеющие потомков, и дуги, ведущие в них. Если граф стал пустым, то выход с ответом « $y(1)$ не может быть начальным состоянием автомата A_2 ».
- 12: Вершины яруса $t + 1$, имеющие одинаковые метки, отождествляем.
- 13: Выполняем обход построенного графа в глубину. Последовательность меток дуг каждого пути, идущего из вершины 1-го яруса к вершинам $(l + 1)$ -го яруса, задаёт возможную управляющую последовательность; включаем её в множество $U(\gamma, y(1))$.

2.3. Анализ автомата A_1

Перейдём ко второй вспомогательной задаче: по множеству $U(\gamma, y(1))$ найти функцию выходов f_1 автомата A_1 . Пусть $\beta = u(1) \dots u(l)$ — произвольная последовательность из $U(\gamma, y(1))$. Положив $h_\beta(g_1^{t-1}(x(1))) = u(t)$, $t = 1, \dots, l$, получим частично определённую булеву функцию $h_\beta(x)$. В соответствии с формулами (1) функция f_1 является доопределением функции h_β для некоторой $\beta \in U(\gamma, y(1))$. И наоборот: из описания работы генератора G следует, что если f_1' — любое доопределение функции h_β , то автомат $A_1' = (\mathbb{F}_2^n, \mathbb{F}_2, g_1, f_1')$ в состоянии $x(1)$ за l тактов работы выдаст управляющую последовательность β , а генератор $G' = A_1' \cdot A_2$ — выходную последовательность γ .

Обозначим $F(\gamma, x(1), y(1))$ множество всех булевых функций f , таких, что автомат $A = (\mathbb{F}_2^n, \mathbb{F}_2, g_1, f)$ в состоянии $x(1)$ за l тактов работы выдаёт управляющую последовательность из множества $U(\gamma, y(1))$, а именно:

$$f(x) \in F(\gamma, x(1), y(1)) \Leftrightarrow u(1) \dots u(l) \in U(\gamma, y(1)), \text{ где } u(t) = f(g_1^{t-1}(x(1))), t = 1, \dots, l.$$

Тогда любая функция из множества $F(\gamma, x(1), y(1)) \cap C_1$ является решением основной задачи 1. Получаем следующую постановку задачи.

Дано: множество $U(\gamma, y(1))$; $x(1)$ — начальное состояние автомата A_1 ; g_1 — его функция переходов; C_1 — класс функций, которому принадлежит f_1 .

Найти: множество $F(\gamma, x(1), y(1)) \cap C_1$.

Необходимые действия описаны в алгоритме 2.

Алгоритм 2. Анализ автомата A_1

Вход: множество $U(\gamma, y(1))$; $x(1)$ — начальное состояние автомата A_1 ; g_1 — его функция переходов; C_1 — класс функций, которому принадлежит f_1 .

Выход: множество $M = F(\gamma, x(1), y(1)) \cap C_1$ возможных функций выходов автомата A_1 .

1: Положим $M := \emptyset$.

2: Для каждой последовательности $u(1) \dots u(l) \in U(\gamma, y(1))$

3: находим частично определённую функцию h , полагая $h(g_1^{t-1}(x(1))) = u(t)$, $t = 1, \dots, l$;

4: находим все доопределения функции h в классе C_1 , добавляем их в множество M .

Способ выполнения шага 4 алгоритма 2 зависит от конкретного класса C_1 . В связи с этим актуальны следующие задачи: поиск условий существования (несуществования) доопределения заданной частично определённой булевой функции в данном классе, условий единственности такого доопределения, метода построения всех её доопределений и др. Примеры их решения в случае, когда классом C_1 является множество функций с заданным или ограниченным числом существенных переменных, можно найти в [6, 7].

Количество повторений шагов 3, 4 алгоритма 2 зависит от мощности множества $U(\gamma, y(1))$. Компьютерные эксперименты показывают, что эта мощность сильно меняется даже при незначительном изменении параметров генератора G , например при изменении только начальных состояний автоматов A_1 и A_2 .

В частном случае, когда функция $f_2(u, y)$ зависит от u линейно, всегда получим $|U(\gamma, y(1))| = 1$: при $f_2(u, y) = u \oplus \varphi(y)$ уравнение $z(t) = f_2(u, y)$ имеет единственное решение $u = z(t) \oplus \varphi(y)$, следовательно, в алгоритме 1 каждая вершина графа имеет одного потомка и путь от первого до последнего яруса единственный.

Для решения задачи 1 достаточно последовательно применить алгоритмы 1 и 2.

2.4. Некоторые обобщения основной задачи

Рассмотрим случаи, когда начальные состояния автоматов A_1 и/или A_2 входят в ключ генератора вместе с функцией f_1 .

Задача 2

Пусть ключом генератора является пара $(y(1), f_1)$. Задача криптоанализа ставится так же, как задача 1, за исключением того, что $y(1)$ неизвестно, его надо найти вместе с функцией f_1 . Решениями будут все пары (y, f) , такие, что $U(\gamma, y) \neq \emptyset$ и $f \in F(\gamma, x(1), y) \cap C_1$.

В самом деле, ввиду $f \in F(\gamma, x(1), y)$, автомат $A = (\mathbb{F}_2^n, \mathbb{F}_2, g_1, f)$ в состоянии $x(1)$ за l тактов работы выдаёт управляющую последовательность из множества $U(\gamma, y)$, которая отображается автоматом $A_2 = (\mathbb{F}_2, \mathbb{F}_2^m, \mathbb{F}_2, g_2, f_2)$ в начальном состоянии y в выходную последовательность γ .

Для поиска решений можно применить такой метод: поочерёдно для всех $y \in \mathbb{F}_2^m$ полагаем $y(1) = y$, выполняем алгоритм 1 и если $U(\gamma, y) \neq \emptyset$, то алгоритм 2.

Аналогично рассуждая, получаем следующие задачи и их решения.

Задача 3

Если ключом генератора является пара $(x(1), f_1)$, то решениями задачи криптоанализа будут все пары (x, f) , такие, что $F(\gamma, x, y(1)) \cap C_1 = M \neq \emptyset$ и $f \in M$. Для их

нахождения применяем алгоритм 1 (заметим, что при «правильном» $y(1)$ множество $U(\gamma, y(1))$ всегда непусто). Затем для каждого $x(1) = x \in \mathbb{F}_2^n$ выполняем алгоритм 2.

Задача 4

Наконец, если ключ — тройка $(x(1), y(1), f_1)$, то решение задачи криптоанализа имеет следующий вид: $\{(x, y, f) : U(\gamma, y) \neq \emptyset \ \& \ F(\gamma, x, y) \cap C_1 = M \neq \emptyset \ \& \ f \in M\}$. Для его нахождения перебираем все $x(1) = x \in \mathbb{F}_2^n$ и решаем для них задачу 2, или перебираем все $y(1) = y \in \mathbb{F}_2^m$ и решаем для них задачу 3.

Предложенные решения задач 2–4, скорее всего, не являются лучшими и даже приемлемыми; поиск более эффективных методов составляет предмет дальнейших исследований.

Компьютерные эксперименты с задачей 2 в случае, когда никаких ограничений на функцию f_1 не накладывается (класс C_1 содержит все булевы функции от n переменных), дали следующие результаты. Пусть $Y = \{y \in \mathbb{F}_2^m : U(\gamma, y) \neq \emptyset\}$ — множество начальных состояний автомата A_2 , в которых он отображает хотя бы одну управляющую последовательность в выходную последовательность γ . Будем оценивать среднее значение $|Y|$ при случайном выборе параметров генератора. Как и ожидалось, оно уменьшается с ростом длины l выходной последовательности γ и стабилизируется в некотором значении $|Y|_{\text{ср}}$ при некотором l (разном для разных n и m). В частности, $|Y|_{\text{ср}} \approx 2$, если функция $f_2(u, y)$ не зависит от u , $|Y|_{\text{ср}} \approx 2^{m-1}$, если она имеет вид $u \vee \varphi(y)$ или $u \wedge \varphi(y)$. Исключение составляет случай, когда функция f_2 зависит от u линейно — $f_2(u, y) = u \oplus \varphi(y)$; в этом случае всегда $|Y| = 2^m$ (т.е. $Y = \mathbb{F}_2^m$), потому что автомат A_2 в любом начальном состоянии $y(1)$ отображает последовательность $u(1) \dots u(l)$ в последовательность $\gamma = z(1) \dots z(l)$, если взять $u(t) = z(t) \oplus \varphi(y(t))$, $t = 1, \dots, l$.

3. Криптоанализ конечно-автоматного генератора (δ, τ) -шагов

Рассмотрим важный частный случай [2, 3] конечно-автоматного генератора $G = A_1 \cdot A_2$, в котором автомат A_2 является автоматом Мура, т.е. его функция выходов не зависит от u , а именно: $f_2(u, y) = f(y)$ для некоторой функции $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2$; функция переходов автомата A_2 имеет вид $g_2(u, y) = \neg u g^\delta(y) + u g^\tau(y)$ для некоторых $g : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^m$ и $\delta, \tau \in \mathbb{N}$. По аналогии с известным генератором (δ, τ) -шагов на регистрах сдвига с линейной обратной связью [8] будем называть его конечно-автоматным генератором (δ, τ) -шагов.

Обозначим выходную последовательность автомата A_2 через $\gamma = z(1)z(2) \dots z(l)$, где $z(t) = f(y(t))$; $y(t+1) = g_2(u(t), y(t))$, $t = 1, 2, \dots, l$; введём в рассмотрение ещё одну последовательность $S = s_1 s_2 \dots$, где $s_i = f(q(i))$; $q(1) = y(1)$; $q(i+1) = g(q(i))$, $i = 1, 2, \dots$. Если ключом генератора является только функция f_1 выходов первого автомата, т.е. во всей описанной схеме, кроме f_1 , криптоаналитику неизвестна только управляющая последовательность $u(1)u(2) \dots u(l)$, то последовательность S можно вычислить заранее, до атаки. Задача анализа автомата A_2 сводится в этом случае к ещё одной вспомогательной задаче — классической задаче поиска в последовательности S такой подпоследовательности $s_{i_1} s_{i_2} \dots s_{i_l}$, что $s_{i_t} = z(t)$, $t = 1, \dots, l$, но со следующим ограничением: $i_{t+1} - i_t \in \{\delta, \tau\}$ для всех $t = 1, \dots, l-1$. Будем называть последовательность индексов $i_1 \dots i_l$, для которой выполнены указанные условия, допустимой; допустимых последовательностей для одних и тех же γ, S, δ, τ может быть несколько.

Для каждой допустимой последовательности индексов $i_1 \dots i_l$ полагаем $u(t) = 0$, если $i_{t+1} - i_t = \delta$, и $u(t) = 1$, если $i_{t+1} - i_t = \tau$. Заметим, что ввиду $z(1) = f(y(1)) = f(q(1)) = s_1$ всегда $i_1 = 1$ и что по выходной последовательности длины l можно

найти только $(l - 1)$ символов управляющей последовательности — от $u(l)$ значения $z(1), \dots, z(l)$ не зависят. Задача анализа автомата A_2 для конечно-автоматного генератора (δ, τ) -шагов ставится следующим образом.

Дано: $\gamma = z(1) \dots z(l)$ — выходная последовательность автомата A_2 ; $\delta, \tau \in \mathbb{N}$; последовательность $S = s_1 s_2 \dots$

Найти: множество $U(\gamma, y(1))$.

Необходимые действия описаны в алгоритме 3. В процессе работы алгоритм строит таблицу T — двумерную таблицу с l строками переменной длины, t -я строка которой содержит возможные значения i_t в допустимых последовательностях индексов. Этот шаг аналогичен построению очередного яруса в алгоритме 1. Этап просеивания соответствует удалению из графа вершин, не имеющих потомков. На этапе 3 рекурсивно строятся все допустимые последовательности индексов; множество $M[t]$, $t = 1, \dots, l-1$, содержит допустимые последовательности для префикса $z(1) \dots z(t)$, каждая из которых может быть продолжена одним или двумя способами; эти продолжения записываются в $M[t+1]$. На этапе 4 для каждой допустимой последовательности индексов в $M[l]$ строится соответствующая ей управляющая последовательность. Алгоритм 3 и метод решения задачи 2 на его основе реализованы на языке C++; компьютерные эксперименты дали результаты, аналогичные полученным для алгоритма 1.

Алгоритм 3. Анализ автомата A_2 для конечно-автоматного генератора (δ, τ) -шагов

Вход: $\gamma = z(1) \dots z(l)$ — выходная последовательность автомата A_2 ; $\delta, \tau \in \mathbb{N}$; последовательность $S = s_1 s_2 \dots$

Выход: множество $U(\gamma, y(1))$.

Э т а п 1. Построение таблицы

1: $T[1] := \{1\}$.

2: Для $t = 2, \dots, l$

$T[t] := \{k + \delta : k \in T[t-1] \ \& \ s_{k+\delta} = z(t)\} \cup \{k + \tau : k \in T[t-1] \ \& \ s_{k+\tau} = z(t)\}$.

Э т а п 2. Просеивание

3: Для $t = l, \dots, 2$

из $T[t-1]$ удаляем все элементы j , такие, что $\forall k \in T[t]$ ($j \neq k - \delta$ & $j \neq k - \tau$).

Э т а п 3. Построение допустимых последовательностей индексов

4: $M[1] := \{(1)\}$.

5: Для $t = 2, \dots, l$

6: $M[t] := \emptyset$.

7: Для всех $(i_1 \dots i_{t-1}) \in M[t-1]$

8: Если $j = i_{t-1} + \delta \in T[t]$, то

9: $M[t] := M[t] \cup \{(i_1 \dots i_{t-1}j)\}$.

10: Если $k = i_{t-1} + \tau \in T[t]$, то

11: $M[t] := M[t] \cup \{(i_1 \dots i_{t-1}k)\}$.

Э т а п 4. Построение множества $U(\gamma, y(1))$

12: Для всех $(i_1 \dots i_l) \in M[l]$

13: полагаем $u(t) = 0$, если $i_{t+1} - i_t = \delta$, и $u(t) = 1$, если $i_{t+1} - i_t = \tau$, $t = 1, \dots, l-1$; включаем последовательность $u(1) \dots u(l-1)$ в множество $U(\gamma, y(1))$.

Заключение

Рассмотрен двухкаскадный конечно-автоматный генератор в общем случае и с ограничениями на автомат второго каскада; в обоих случаях представлены алго-

ритмы решения задачи криптоанализа генератора с функцией выходов f_1 автомата первого каскада в роли ключа. Предложены способы криптоанализа в случае, когда в ключ вместе с функцией f_1 входят начальные состояния обоих или одного из автоматов. В дальнейшем предполагается исследование генераторов с ключами, содержащими и другие их параметры.

Авторы выражают глубокую признательность Геннадию Петровичу Агибалову за постановку задачи и помощь в работе.

ЛИТЕРАТУРА

1. Агибалов Г. П. Криптоавтоматы с функциональными ключами // Прикладная дискретная математика. 2017. № 36. С. 59–72.
2. Агибалов Г. П., Панкратова И. А. О двухкаскадных конечно-автоматных криптографических генераторах и методах их криптоанализа // Прикладная дискретная математика. 2017. № 35. С. 38–47.
3. Агибалов Г. П., Панкратова И. А. К криптоанализу двухкаскадных конечно-автоматных криптографических генераторов // Прикладная дискретная математика. Приложение. 2016. № 9. С. 41–43.
4. Торопов Н. Р. Язык программирования ЛЯПАС // Прикладная дискретная математика. 2009. № 2(4). С. 9–25.
5. Агибалов Г. П., Липский В. Б., Панкратова И. А. О криптографическом расширении и его реализации для русского языка программирования // Прикладная дискретная математика. 2013. № 3(21). С. 93–104.
6. Агибалов Г. П. О некоторых доопределениях частичной булевой функции // Труды Сибирского физико-технического института. 1970. Вып. 49. С. 12–19.
7. Агибалов Г. П., Сунгурова О. Г. Криптоанализ конечно-автоматного генератора ключевого потока с функцией выходов в качестве ключа // Вестник Томского государственного университета. Приложение. Август 2006. № 17. С. 104–108.
8. Фомичёв В. М. Методы дискретной математики в криптологии. М.: ДИАЛОГ-МИФИ, 2010. 424 с.

REFERENCES

1. Agibalov G. P. Kriptoavtomaty s funktsional'nymi klyuchami [Cryptautomata with functional keys]. Prikladnaya Diskretnaya Matematika, 2017, no. 36, pp. 59–72.
2. Agibalov G. P. and Pankratova I. A. O dvukhkaskadnykh konechno-avtomatnykh kriptograficheskikh generatorakh i metodakh ikh kriptoolnaliza [About 2-cascade finite automata cryptographic generators and their cryptanalysis]. Prikladnaya Diskretnaya Matematika, 2017, no. 35, pp. 38–47.
3. Agibalov G. P. and Pankratova I. A. K kriptoolnalizu dvukhkaskadnykh konechno-avtomatnykh kriptograficheskikh generatorov [To cryptanalysis of 2-cascade finite automata cryptographic generators]. Prikladnaya Diskretnaya Matematika. Prilozhenie, 2016, no. 9, pp. 41–43.
4. Toropov N. R. Yazik programmirovaniya LYaPAS [Programming language LYaPAS]. Prikladnaya Diskretnaya Matematika, 2009, no. 2(4), pp. 9–25. (in Russian)
5. Agibalov G. P., Lipskiy V. B., and Pankratova I. A. O kriptograficheskom rasshirenii i ego realizatsii dlya russkogo yazyka programmirovaniya [Cryptographic extension and its implementation for Russian programming language]. Prikladnaya Diskretnaya Matematika, 2013, no. 3(21), pp. 93–104. (in Russian)
6. Agibalov G. P. O nekotorykh doopredeleniyakh chastichnoy bulevoy funktsii [Some completions of partial Boolean function]. Trudy SPhTI, 1970, iss. 49, pp. 12–19. (in Russian)

7. *Agibalov G. P. and Sungurova O. G.* Kriptoanaliz konechno-avtomatnogo generatora klyuchevogo potoka s funktsiey vykhodov v kachestve klyucha [Cryptanalysis of a finite-state key-stream generator with an output function as a key]. *Vestnik TSU. Prilozhenie*, 2006, no. 17, pp. 104–108. (in Russian)
8. *Fomichev V. M.* *Metody diskretnoy matematiki v kriptologii* [Methods of Discrete Mathematics in Cryptology]. Moscow, DIALOG-MEPHI Publ., 2010. 424 p. (in Russian)

UDC 519.7

DOI 10.17223/20710410/42/4

ELGAMAL CRYPTOSYSTEMS ON BOOLEAN FUNCTIONS¹

G. P. Agibalov

*National Research Tomsk State University, Tomsk, Russia***E-mail:** agibalov@mail.tsu.ru

Here is a description of ElGamal public-key encryption and digital signature schemes constructed on the base of bijective systems of Boolean functions. The description is illustrated with a simple example in which the used Boolean functions are written in logical notation. In our encryption and signature schemes on Boolean functions, every one ciphertext or message signature is a pair of values, as in the basic ElGamal cryptosystem on a group. In our case, these values are Boolean vectors. Each vector in the pair depends on the value of a function on a plaintext or on a message, and this function is typically obtained from a given bijective vector Boolean function g by applying some random and secret negation and permutation operations on the sets of variables and coordinate functions of g . For the pair of vectors in the ciphertext or in the message signature, the decryption algorithm produces the plaintext, and the signature verification algorithm accepts the signature, performing some computation on this pair. The signature is accepted for a message if and only if the computation results in this message. All the computations in the processes of encryption, decryption, signing and verification are logical and performed for Boolean values, promising their implementation efficiency to be more high than in the basic ElGamal schemes on groups.

Keywords: *bijective vector Boolean functions, permutation and negation operations, ElGamal encryption, ElGamal signature.*

Introduction

The ElGamal cryptosystems, including the basic encryption and signature schemes as well as their multiple generalizations and variations [1], are typically defined on the base of some groups in which the group operation is easily to apply and the discrete logarithm problem is computationally infeasible. The multiplicative groups \mathbb{Z}_p^* , $\mathbb{F}_{2^m}^*$ and additive group of points on elliptic curve over \mathbb{F}_q have received the most attention [1]. It is known that the public-key cryptosystems based on similar groups are particularly susceptible to quantum attacks. The ElGamal cryptosystems are not excluded from this family.

In this paper, we try to propose an alternative mathematical background for constructing ElGamal cryptosystems, namely the algebra of bijective vector Boolean functions with the negation and permutation operations on the sets of their variables and coordinate functions. Section 2 of the paper is a collection of the basic elements of this background that we use in the description of our ElGamal encryption and signature schemes in Sections 3 and 5 respectively and of an illustrative example in Section 4. For any of operations encryption and signature, we consider different variations of the scheme and describe each of them in the form of the corresponding basic ElGamal scheme (encryption or signature). For reader's convenience, Section 1 recalls the basic ElGamal encryption and signature schemes in this form from [1].

¹The author was supported by the RFBR-grant no. 17-01-00354.

1. Basic ElGamal cryptosystem

1.1. Basic ElGamal encryption scheme

Parameters: p is a large random prime, α is a generator of the multiplicative group \mathbb{Z}_p^* , a is a random integer, $1 \leq a \leq p-2$, m is a plaintext, $m \in \mathbb{Z}_p$.

Public key is (p, α, α^a) , *private key* is a .

Encryption: $k \in_R \{1, 2, \dots, p-2\}$ (here and further, the symbol \in_R means “to be randomly chosen”), $\gamma = \alpha^k \bmod p$, $\delta = m(\alpha^a)^k \bmod p$, (γ, δ) is the *ciphertext*.

Decryption: $\gamma^{-a}\delta (= \alpha^{-ak}m\alpha^{ak}) = m \bmod p$.

1.2. Basic ElGamal signature scheme

Parameters: p is a large random prime, α is a generator of the multiplicative group \mathbb{Z}_p^* , a is a random integer, $1 \leq a \leq p-2$, $\beta = \alpha^a$, m is a message (or its hash value), $m \in \mathbb{Z}_p$.

Public key is (p, α, β) , *private key* is a .

Signing: $k \in_R \{1, 2, \dots, p-2\}$, $(k, p-1) = 1$, $\gamma = \alpha^k \bmod p$, $\delta = k^{-1}(m - a\gamma) \bmod (p-1)$, signature for m is the pair (γ, δ) .

Verification: if $\gamma \leq 1$ or $\gamma > p-1$, then reject the signature (γ, δ) , otherwise accept the signature (γ, δ) if and only if $\beta^\gamma \gamma^\delta = \alpha^m \bmod p$.

2. Algebra of bijective vector Boolean functions

First of all, we note that earlier some elements of this algebra were used in constructing and cryptanalysis of cryptographic systems with functional keys, namely in [2] — for symmetric block ciphers, in [3] — for public-key encryption and signature schemes.

2.1. Permutation and negation operations

We begin with the notions of the permutation and negation operations over Boolean vectors. Let n be an integer, $n \geq 2$, and \mathbb{S}_n be the set of all permutations of the row $(12\dots n)$, that is, $\mathbb{S}_n = \{(i_1 i_2 \dots i_n) : i_j \in \{1, 2, \dots, n\}, j \neq r \Rightarrow i_j \neq i_r, j, r \in \{1, \dots, n\}\}$. A permutation $\pi = (i_1 i_2 \dots i_n) \in \mathbb{S}_n$ is called a *permutation operation* on \mathbb{F}_2^n if the result of its application to any vector $w = w_1 w_2 \dots w_n$ in \mathbb{F}_2^n is the vector $\pi(w) = w_{i_1} w_{i_2} \dots w_{i_n}$. A Boolean vector $\sigma = b_1 b_2 \dots b_n \in \mathbb{F}_2^n$ is called a *negation operation* on \mathbb{F}_2^n if the result of its application to a vector $\alpha = a_1 a_2 \dots a_n$ in \mathbb{F}_2^n is the vector $\alpha^\sigma = a_1^{b_1} a_2^{b_2} \dots a_n^{b_n}$, where for a and b in \mathbb{F}_2 , we have $a^b = a$ if $b = 1$ and $a^b = \neg a$ if $b = 0$. Both of these operations are invertible. The inversions for them are denoted in the usual manner, namely π^{-1} and σ^{-1} . By the definition, if $\pi = (i_1 i_2 \dots i_n)$, $s(k) = i_k$, and $\pi^{-1} = (j_1 j_2 \dots j_n)$, then $s^{-1}(i_k) = k$, $s^{-1}(k) = j_k$, and $j_k = s^{-1}(s^{-1}(i_k))$, $k \in \{1, 2, \dots, n\}$. The permutation and negation operations π and σ are called *identity* and denoted by 1 if $\pi = (12\dots n)$ and $\sigma = 11\dots 1$ respectively. So $1(w) = w$ and $a^1 = a$.

2.2. Combinatorial and algebraic notations

Let $x = (x_1, x_2, \dots, x_n)$ be a string of n different Boolean variables, $g : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ be a n -dimensional vector Boolean function $g(x)$, and $g_i : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$, $i \in \{1, 2, \dots, n\}$, be the coordinate functions of g . That is, $g(x) = g_1(x)g_2(x)\dots g_n(x)$. Let π_1, π_2 and σ_1, σ_2 be the symbols of variables with the values, respectively, of permutation operations in \mathbb{S}_n and of negation operations in \mathbb{F}_2^n , namely σ_1, π_1 — over the variables in x and σ_2, π_2 — over the coordinates in $g(x)$. Let also $I = \{\sigma_1, \pi_1, \sigma_2, \pi_2\}$, $J \subseteq I$, V_J be the set of all strings of values for the variables in I in which (strings) the value of each variable from $I \setminus J$ is equal to 1, i.e. $V_J = \{(s_1 p_1 s_2 p_2) : s_i = 1 \text{ if } \sigma_i \in I \setminus J \text{ and } p_i = 1 \text{ if } \pi_i \in I \setminus J; s_i \in \mathbb{F}_2^n \text{ if } \sigma_i \in J \text{ and } p_i \in \mathbb{S}_n \text{ if } \pi_i \in J; i \in \{1, 2\}\}$,

$$\pi_i^J = \begin{cases} 1, & \text{if } \pi_i \in I \setminus J, \\ \pi_i, & \text{if } \pi_i \in J, \end{cases} \quad \sigma_i^J = \begin{cases} 1, & \text{if } \sigma_i \in I \setminus J, \\ \sigma_i, & \text{if } \sigma_i \in J, \end{cases} \quad i \in \{1, 2\},$$

and $g^J(x)$ be the formula $\pi_2^J(g^{\sigma_2^J}(\pi_1^J(x^{\sigma_1^J})))$. Particularly, for any $a = (s_1 p_1 s_2 p_2) \in V_J$, a formula $g^a(x)$ is defined too as $g^a(x) = p_2(g^{s_2}(p_1(x^{s_1})))$. In fact, $g^J(x)$ is a subformula of $g^I(x) = \pi_2(g^{\sigma_2}(\pi_1(x^{\sigma_1})))$ with the negation and permutation operations from a subset $J \subseteq I$. For example, if $J = \{\sigma_1, \pi_2\}$, then $\pi_1^J = 1$, $\sigma_2^J = 1$, and $g^J(x) = \pi_2(g(x^{\sigma_1}))$.

The formulas $g^J(x)$ for all possible J are given in the Table 1:

Table 1

| | | | | | | | | |
|----------|-------------|-------------------------------------|---------------------------------|--|-------------------------------------|---------------------------------|------------------------------|--------------------------|
| J | \emptyset | $\{\sigma_1\}$ | $\{\pi_1\}$ | $\{\sigma_2\}$ | $\{\pi_2\}$ | $\{\sigma_1, \pi_1\}$ | $\{\sigma_1, \sigma_2\}$ | $\{\sigma_1, \pi_2\}$ |
| $g^J(x)$ | $g(x)$ | $g(x^{\sigma_1})$ | $g(\pi_1(x))$ | $g^{\sigma_2}(x)$ | $\pi_2(g(x))$ | $g(\pi_1(x^{\sigma_1}))$ | $g^{\sigma_2}(x^{\sigma_1})$ | $\pi_2(g(x^{\sigma_1}))$ |
| | | $\{\pi_1, \sigma_2\}$ | $\{\pi_1, \pi_2\}$ | $\{\sigma_2, \pi_2\}$ | $\{\sigma_1, \pi_1, \sigma_2\}$ | $\{\sigma_1, \pi_1, \pi_2\}$ | | |
| | | $g^{\sigma_2}(\pi_1(x))$ | $\pi_2(g(\pi_1(x)))$ | $\pi_2(g^{\sigma_2}(x))$ | $g^{\sigma_2}(\pi_1(x^{\sigma_1}))$ | $\pi_2(g(\pi_1(x^{\sigma_1})))$ | | |
| | | $\{\sigma_1, \sigma_2, \pi_2\}$ | $\{\pi_1, \sigma_2, \pi_2\}$ | $\{\sigma_1, \pi_1, \sigma_2, \pi_2\}$ | | | | |
| | | $\pi_2(g^{\sigma_2}(x^{\sigma_1}))$ | $\pi_2(g^{\sigma_2}(\pi_1(x)))$ | $\pi_2(g^{\sigma_2}(\pi_1(x^{\sigma_1})))$ | | | | |

To make distinction between signs of kinds $g^J(x)$ and $g^{\sigma_2^J}(x)$ as well as between signs of kinds $g^a(x)$ and $g^{s_2}(x)$, we often write $(g(x))^{\sigma_2^J}$ and $(g(x))^{s_2}$ instead of $g^{\sigma_2^J}(x)$ and $g^{s_2}(x)$ respectively. So, $g^J(x) = \pi_2^J(g(\pi_1^J(x^{\sigma_1^J})))^{\sigma_2^J}$ and $g^a(x) = p_2(g(p_1(x^{s_1})))^{s_2}$.

For any vector-columns a, σ in \mathbb{F}_2^n and a permutation $\pi = (i_1 i_2 \dots i_n) \in \mathbb{S}_n$, if $c = \neg\sigma$, $T = (t_{kj})$ is a permutation matrix of order n over \mathbb{F}_2 where $t_{kj} = 1 \Leftrightarrow j = i_k$ for all $k, j \in \{1, 2, \dots, n\}$ (we call it *matrix of π*), then $a^\sigma = a \oplus c$ and $\pi(a) = Ta$. This allows us to introduce the more simple notation in which A and D are the matrices of permutations π_1 and π_2 respectively and b and d are the vector-columns $\neg\sigma_1$ and $\neg\sigma_2$ respectively, to use the symbols of variables A, D, b, d instead of symbols of operations $\pi_1, \pi_2, \sigma_1, \sigma_2$ respectively in the sets I, J as well as in the formulas for $f(x), f^{-1}(x)$ and to apply linear algebra methods in solving the equations $y = f(x)$ and $x = f^{-1}(y)$ with regard to unknown key parameters. Further, the fact of such replacement is denoted by the sign \simeq . For example, $\{\pi_1, \sigma_1, \sigma_2\} \simeq \{A, b, d\}$, $g^I(x) = \pi_2(g(\pi_1(x^{\sigma_1})))^{\sigma_2} \simeq D(g(A(x \oplus b)) \oplus d)$. The formulas under consideration with symbols of permutation and negation operations $\sigma_1, \pi_1, \sigma_2, \pi_2$ are said to be ones in *combinatorial notation* and the formulas where the operations are represented by symbols b, A, d, D of matrices and vectors are formulas in *algebraic notation*.

All the formulas $g^J(x)$ in algebraic notation are given in the Table 2:

Table 2

| | | | | | | | | |
|----------|-------------|-----------------------------|---------------------|--------------------------------|-----------------------------|---------------------|--------------------------|------------------|
| J | \emptyset | $\{b\}$ | $\{A\}$ | $\{d\}$ | $\{D\}$ | $\{b, A\}$ | $\{b, d\}$ | $\{b, D\}$ |
| $g^J(x)$ | $g(x)$ | $g(x \oplus b)$ | $g(Ax)$ | $g(x \oplus d)$ | $Dg(x)$ | $g(A(x \oplus b))$ | $g(x \oplus b) \oplus d$ | $Dg(x \oplus b)$ |
| | | $\{A, d\}$ | $\{A, D\}$ | $\{d, D\}$ | $\{b, A, d\}$ | $\{b, A, D\}$ | | |
| | | $g(Ax) \oplus d$ | $Dg(Ax)$ | $D(g(x \oplus d))$ | $g(A(x \oplus b)) \oplus d$ | $Dg(A(x \oplus b))$ | | |
| | | $\{b, d, D\}$ | $\{A, d, D\}$ | $\{b, A, d, D\}$ | | | | |
| | | $D(g(x \oplus b) \oplus d)$ | $D(g(Ax) \oplus d)$ | $D(g(A(x \oplus b)) \oplus d)$ | | | | |

2.3. Permutation-negation compositions

There are two kinds of composition for permutation-negation operations — multiplicative and serial. We begin with the first one.

Multiplicative composition

For any subsets $J, L \subseteq I$, define it as

$$g^{J^L}(x) = \pi_2^L(g^J(\pi_1^L(x^{\sigma_1^L})))^{\sigma_2^L}.$$

Particularly, this means that for any $a = (s_1 p_1 s_2 p_2) \in V_J$ and $k = (r_1 q_1 r_2 q_2) \in V_L$, the value $g^{a^k}(x)$ is defined as

$$g^{a^k}(x) = q_2(g^a(q_1(x^{r_1})))^{r_2},$$

where $g^a(x) = p_2(g(p_1(x^{s_1})))^{s_2}$, therefore

$$g^{a^k}(x) = q_2(p_2(g(p_1((q_1(x^{r_1}))^{s_1})))^{s_2})^{r_2}.$$

By the definition, we should write $(g^J)^L$ and $(g^a)^k$ instead of g^{J^L} and g^{a^k} respectively, but for simplicity we remove the parentheses.

Let $b^J = \neg\sigma_1^J, b^L = \neg\sigma_1^L, d^J = \neg\sigma_2^J, d^L = \neg\sigma_2^L$, and A^J, A^L, D^J, D^L denote the matrices of $\pi_1^J, \pi_1^L, \pi_2^J, \pi_2^L$ respectively. We have $g^{J^{\sigma_2^L}}(x) = g^J(x) \oplus d^L = D^J(g(A^J(x \oplus b^J)) \oplus d^J) \oplus d^L$ and $\pi_1^L(x^{\sigma_1^L}) = A^L(x \oplus b^L)$. Hence, $g^{J^{\sigma_2^L}}(\pi_1^L(x^{\sigma_1^L})) = D^J(g(A^J(A^L(x \oplus b^L) \oplus b^J)) \oplus d^J) \oplus d^L$ and

$$g^{J^L}(x) = D^L(D^J(g(A^J(A^L(x \oplus b^L) \oplus b^J)) \oplus d^J) \oplus d^L).$$

Particularly,

$$g^{a^k}(x) = D'(D(g(A(A'(x \oplus b') \oplus b)) \oplus d) \oplus d'),$$

where $b = \neg s_1, b' = \neg r_1, d = \neg s_2, d' = \neg r_2$, and A, A', D, D' are the matrices of permutations p_1, q_1, p_2, q_2 respectively.

Serial composition

For the subsets $J, L \subseteq I$, it is defined as follows

$$\begin{aligned} g^L(g^J(x)) &= \pi_2^L(g(\pi_1^L((g^J(x))^{\sigma_1^L}))^{\sigma_2^L}) = \pi_2^L(g(\pi_1^L((\pi_2^J(g(\pi_1^J(x^{\sigma_1^J})))^{\sigma_2^J})^{\sigma_1^L}))^{\sigma_2^L}) = \\ &= D^L(g(A^L((D^J(g(A^J(x \oplus b^J)) \oplus d^J) \oplus b^L)) \oplus d^L), \end{aligned}$$

and for the permutation-negation operations $a = (s_1 p_1 s_2 p_2) \in V_J$ and $k = (r_1 q_1 r_2 q_2) \in V_L$ — in the following way

$$\begin{aligned} g^k(g^a(x)) &= q_2(g(q_1((g^a(x))^{r_1})))^{r_2} = q_2(g(q_1((p_2(g(p_1(x^{s_1})))^{s_2})^{r_1})))^{r_2} = \\ &= D'(g(A'((D(g(A(x \oplus b)) \oplus d) \oplus b') \oplus d')). \end{aligned}$$

2.4. Derived functions

The order of operation performing in $g^J(x)$ is determined by the parentheses and the following additional agreement: in a subformula $g^\sigma(u)$, the value of $g(u)$ is calculated before performing the operation σ . So, the operations in $g^J(x)$, including the function g , are performed in the order $\sigma_1^J, \pi_1^J, g, \sigma_2^J, \pi_2^J$. Under particular operations s_1, p_1, s_2, p_2 as possible values for variables $\sigma_1^J, \pi_1^J, \sigma_2^J, \pi_2^J$ respectively, for particular function g and a value α of x , the value of $g^J(\alpha)$ is sequentially computed as follows: $v_1(\alpha) = \alpha^{s_1}, v_2(\alpha) = p_1(v_1(\alpha)), v_3(\alpha) = g(v_2(\alpha)), v_4(\alpha) = v_3^{s_2}(\alpha), g^J(\alpha) = p_2(v_4(\alpha))$. This defines a function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ such that $f(x) = p_2(v_4(x))$. By the definition, $f(x)$ is uniquely determined by the function $g(x)$ and negation and permutation transformations of its variables and coordinates. For $a = (s_1, p_1, s_2, p_2)$, we denote it $g^a(x)$ and call it a *derived* function (derived from g by the transformation a). Thus, $g^a(x) = p_2(g^{s_2}(p_1(x^{s_1}))) = p_2(g(p_1(x^{s_1})))^{s_2}$. The second of these expressions for $g^a(x)$ explicitly shows the order of applying operations in the process of computing $g^a(x)$. Schematically, the computation according to it can be expressed with the following chain:

$$x \xrightarrow{s_1} x^{s_1} \xrightarrow{p_1} p_1(x^{s_1}) \xrightarrow{g} g(p_1(x^{s_1})) \xrightarrow{s_2} g^{s_2}(p_1(x^{s_1})) \xrightarrow{p_2} g^a(x).$$

In every case when $g(x)$ is a bijective vector Boolean function on \mathbb{F}_2^n , so should be the function $g^a(x)$. Its inverse $g^{a^{-1}}(x)$ satisfies the identity relation $g^{a^{-1}}(g^a(x)) = x$ and can be performed in the following way: if $y = g^a(x)$, then $x = g^{a^{-1}}(y) = [p_1^{-1}(g^{-1}((p_2^{-1}(y))^{s_2}))]^{s_1}$. Schematically, the computation according to this formula can be expressed with the following chain:

$$y \xrightarrow{p_2^{-1}} g^{s_2}(p_1(x^{s_1})) \xrightarrow{s_2} g(p_1(x^{s_1})) \xrightarrow{g^{-1}} p_1(x^{s_1}) \xrightarrow{p_1^{-1}} x^{s_1} \xrightarrow{s_1} x.$$

Computational complexities of function $g(x)$ and its derived functions are of the same order. In particular, if $g(x)$ is of a polynomial complexity, then $g^a(x)$ with known g and a is of a polynomial complexity too what we can not say about $g^{a^{-1}}$.

3. ElGamal encryption on Boolean functions

We need to say that in reality we can construct on Boolean functions very many different variations of ElGamal encryption schemes which can differ each other in public and private keys definitions and in encryption and decryption equations. The following variation seems to have the most simple expression and insufficiently strong private key.

3.1. Encryption scheme $\mathcal{E}1$

Parameters: n is an integer, $n \geq 2$; $g(x) = g_1(x)g_2(x) \dots g_n(x)$ is a bijective vector Boolean function with the coordinate functions $g_1(x), \dots, g_n(x)$ specified in a constructive way and computed with a polynomial (in n) time complexity, $g : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$; $\emptyset \neq J, L \subseteq I = \{\sigma_1, \pi_1, \sigma_2, \pi_2\}$, where π_1, π_2 and σ_1, σ_2 are the symbols of variables with the values, respectively, of permutation operations in \mathbb{S}_n and of negation operations in \mathbb{F}_2^n ; $a = (s_1 p_1 s_2 p_2) \in_R V_J$ and $g^a(x) = p_2(g^{s_2}(p_1(x^{s_1})))$.

Public key is $(g(x), g^a(x))$, *private key* is $g^{a^{-1}}(x)$, *secret* parameter is a .

Encryption: m is a *plaintext*, $m \in \mathbb{F}_2^n$; k is a *randomization* parameter, $k = (r_1, q_1, r_2, q_2) \in_R V_L$; $\gamma(m) = g^k(m) = q_2(g^{r_2}(q_1(m^{r_1})))$, $\delta(m) = g^k(m) \oplus g^a(m)$; $(\gamma(m), \delta(m))$ is the *ciphertext*.

Decryption: $m = g^{a^{-1}}(\gamma(m) \oplus \delta(m))$.

Proof that decryption works: $g^{a^{-1}}(\gamma(m) \oplus \delta(m)) = g^{a^{-1}}(g^k(m) \oplus g^k(m) \oplus g^a(m)) = g^{a^{-1}}(g^a(m)) = m$.

3.2. Encryption scheme $\mathcal{E}2$

Public key is $g^a(x)$, *private key* is $g^{a^{-1}}(x)$, *secret* parameter is a .

Encryption: m is a *plaintext*, $m \in \mathbb{F}_2^n$; k is a *randomization* parameter, $k = (r_1, q_1, r_2, q_2) \in_R V_L$; $\gamma(m) = g^{a^k}(m) = q_2(g^a(q_1(m^{r_1})))^{r_2}$, $\delta(m) = g^{a^k}(m) \oplus g^a(m)$; $(\gamma(m), \delta(m))$ is the *ciphertext*.

Decryption: $m = g^{a^{-1}}(\gamma(m) \oplus \delta(m))$.

Proof that decryption works: $g^{a^{-1}}(\gamma(m) \oplus \delta(m)) = g^{a^{-1}}(g^{a^k}(m) \oplus g^{a^k}(m) \oplus g^a(m)) = g^{a^{-1}}(g^a(m)) = m$.

3.3. Encryption scheme $\mathcal{E}3$

This variation is proposed by V. A. Roman'kov.

Public key is $g^a(x)$, *private key* is $g^{a^{-1}}(x)$, *secret* parameter is a .

Encryption: m is a *plaintext*, $m \in \mathbb{F}_2^n$; (k, u) are *randomization* parameters, $k = (r_1, q_1, r_2, q_2) \in_R V_L$, $u \in_R \mathbb{F}_2^n$; $\gamma = g^a(g^k(u))$, $\delta = g^k(u) \oplus m$; (γ, δ) is the *ciphertext*.

Decryption: $m = g^{a^{-1}}(\gamma) \oplus \delta$.

Proof that decryption works: $g^{a^{-1}}(\gamma) \oplus \delta = g^{a^{-1}}(g^a(g^k(u))) \oplus g^k(u) \oplus m = g^k(u) \oplus g^k(u) \oplus m = m$.

3.4. Encryption scheme $\mathcal{E}4$

This variation is proposed by I. A. Pankratova.

Public key is $g^a(x)$, *private key* is $g^{a^{-1}}(x)$, *secret parameter* is a .

Encryption: m is a *plaintext*, $m \in \mathbb{F}_2^n$; u is a *randomization parameter*, $u \in_R \mathbb{F}_2^n$; $\gamma = g^a(u)$, $\delta = u \oplus m$; (γ, δ) is the *ciphertext*.

Decryption: $m = g^{a^{-1}}(\gamma) \oplus \delta$.

Proof that decryption works: $g^{a^{-1}}(\gamma) \oplus \delta = g^{a^{-1}}(g^a(u)) \oplus u \oplus m = u \oplus u \oplus m = m$.

4. Example

Here, we illustrate the ElGamal encryption on Boolean functions effectively represented in an analytical form (not by tables).

Let $n = 4$, $x = x_1x_2x_3x_4$, $g : \mathbb{F}_2^4 \rightarrow \mathbb{F}_2^4$, $g(x) = g_1(x)g_2(x)g_3(x)g_4(x)$,

$$g_1(x) = x_1 \oplus x_2 \oplus x_3 \oplus x_4, \quad g_2(x) = x_1x_2 \vee \bar{x}_1\bar{x}_2, \quad g_3(x) = x_4, \quad g_4(x) = x_2\bar{x}_3 \vee x_1x_3,$$

$g^{-1} : \mathbb{F}_2^4 \rightarrow \mathbb{F}_2^4$, $g^{-1}(x) = g'_1(x)g'_2(x)g'_3(x)g'_4(x)$. We have

$$\begin{aligned} g'_1(x) &= x_2x_4 \vee \bar{x}_1\bar{x}_3x_4 \vee x_1\bar{x}_2\bar{x}_3\bar{x}_4 \vee \bar{x}_1\bar{x}_2x_3\bar{x}_4 \vee x_1x_3x_4, \\ g'_2(x) &= x_2x_4 \vee \bar{x}_1x_3x_4 \vee x_1\bar{x}_3x_4 \vee \bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4 \vee x_1\bar{x}_2x_3\bar{x}_4, \\ g'_3(x) &= \bar{x}_1\bar{x}_2\bar{x}_3 \vee \bar{x}_1x_2x_3 \vee x_1x_2\bar{x}_3 \vee x_1\bar{x}_2x_3, \quad g'_4(x) = x_3. \end{aligned}$$

Let also $J = L = I$, $V_J = V_L = \{(s_1, p_1, s_2, p_2) : p_1, p_2 \in \mathbb{S}_4; s_1, s_2 \in \mathbb{F}_2^4\}$;

$$a = (s_1, p_1, s_2, p_2) \in V_J, \quad p_1 = 2341, \quad p_2 = 4123, \quad s_1 = 1001, \quad s_2 = 0111;$$

$$k = (r_1, q_1, r_2, q_2) \in V_L, \quad q_1 = 4321, \quad q_2 = 3412, \quad r_1 = 0001, \quad r_2 = 1000.$$

We have that

$$\begin{aligned} x^{s_1} &= x_1\bar{x}_2\bar{x}_3x_4, \quad p_1(x^{s_1}) = \bar{x}_2\bar{x}_3x_4x_1, \\ g^{s_2}(x) &= \bar{g}_1(x)g_2(x)g_3(x)g_4(x), \quad p_2(g^{s_2}(x)) = g_4(x)\bar{g}_1(x)g_2(x)g_3(x); \\ g^a(x) &= p_2(g^{s_2}(p_1(x^{s_1}))) = (g_4(\bar{x}_2\bar{x}_3x_4x_1), \bar{g}_1(\bar{x}_2\bar{x}_3x_4x_1), g_2(\bar{x}_2\bar{x}_3x_4x_1), g_3(\bar{x}_2\bar{x}_3x_4x_1)) = \\ &= ((\bar{x}_3\bar{x}_4 \vee \bar{x}_2x_4), \neg(\bar{x}_2 \oplus \bar{x}_3 \oplus x_4 \oplus x_1), (\bar{x}_2\bar{x}_3 \vee x_2x_3), (x_1)); \\ y &= y_1y_2y_3y_4, \quad p_2^{-1}(y) = y_2y_3y_4y_1, \quad (p_2^{-1}(y))^{s_2} = \bar{y}_2\bar{y}_3\bar{y}_4\bar{y}_1, \\ p_1^{-1}(x) &= x_4x_1x_2x_3, \quad (p_1^{-1}(x))^{s_1} = x_4\bar{x}_1\bar{x}_2x_3; \\ g^{a^{-1}}(y) &= [p_1^{-1}(g^{-1}((p_2^{-1}(y))^{s_2}))]^{s_1} = [p_1^{-1}(g^{-1}(\bar{y}_2\bar{y}_3\bar{y}_4\bar{y}_1))]^{s_1} = [p_1^{-1}(g'_1(\bar{y}_2\bar{y}_3\bar{y}_4\bar{y}_1), g'_2(\bar{y}_2\bar{y}_3\bar{y}_4\bar{y}_1), \\ &g'_3(\bar{y}_2\bar{y}_3\bar{y}_4\bar{y}_1), g'_4(\bar{y}_2\bar{y}_3\bar{y}_4\bar{y}_1))]^{s_1} = [g'_4(\bar{y}_2\bar{y}_3\bar{y}_4\bar{y}_1), \bar{g}'_1(\bar{y}_2\bar{y}_3\bar{y}_4\bar{y}_1), \bar{g}'_2(\bar{y}_2\bar{y}_3\bar{y}_4\bar{y}_1), \\ &g'_3(\bar{y}_2\bar{y}_3\bar{y}_4\bar{y}_1)] = [y_4, \neg(y_1y_3 \vee y_1y_2\bar{y}_4 \vee \bar{y}_1\bar{y}_2\bar{y}_3\bar{y}_4 \vee \bar{y}_1y_2\bar{y}_3y_4 \vee y_1\bar{y}_2y_4), \\ &\neg(y_1y_3 \vee y_1y_2y_4 \vee y_1\bar{y}_2\bar{y}_4 \vee \bar{y}_1y_2\bar{y}_3\bar{y}_4 \vee \bar{y}_1\bar{y}_2\bar{y}_3y_4), y_2\bar{y}_3\bar{y}_4 \vee y_2y_3y_4 \vee \bar{y}_2y_3\bar{y}_4 \vee \bar{y}_2\bar{y}_3y_4]; \\ x^{r_1} &= \bar{x}_1\bar{x}_2\bar{x}_3x_4, \quad q_1(x^{r_1}) = x_4\bar{x}_3\bar{x}_2\bar{x}_1, \quad g^{r_2}(x) = g_1(x)\bar{g}_2(x)\bar{g}_3(x)\bar{g}_4(x), \\ q_2(g^{r_2}(x)) &= \bar{g}_3(x)\bar{g}_4(x)g_1(x)\bar{g}_2(x); \\ g^k(x) &= y = y_1y_2y_3y_4 = q_2(g^{r_2}(q_1(x^{r_1}))) = \\ &= (\bar{g}_3(x_4\bar{x}_3\bar{x}_2\bar{x}_1), \bar{g}_4(x_4\bar{x}_3\bar{x}_2\bar{x}_1), g_1(x_4\bar{x}_3\bar{x}_2\bar{x}_1), \bar{g}_2(x_4\bar{x}_3\bar{x}_2\bar{x}_1)) = \\ &= (x_1, \neg(x_2\bar{x}_3 \vee \bar{x}_2x_4), x_4 \oplus \bar{x}_3 \oplus \bar{x}_2 \oplus \bar{x}_1, \neg(\bar{x}_3x_4 \vee x_3\bar{x}_4)); \\ q_2^{-1}(y) &= y_3y_4y_1y_2, \quad (q_2^{-1}(y))^{r_2} = y_3\bar{y}_4\bar{y}_1\bar{y}_2, \quad q_1^{-1}(x) = x_4x_3x_2x_1, \quad (q_1^{-1}(x))^{r_1} = \bar{x}_4\bar{x}_3\bar{x}_2x_1; \end{aligned}$$

$$\begin{aligned}
g^{k^{-1}}(y) &= [q_1^{-1}(g^{-1}((q_2^{-1}(y))^{r_2}))]^{r_1} = [q_1^{-1}(g^{-1}(y_3\bar{y}_4\bar{y}_1\bar{y}_2))]^{r_1} = \\
&= [q_1^{-1}(g'_1(y_3\bar{y}_4\bar{y}_1\bar{y}_2), g'_2(y_3\bar{y}_4\bar{y}_1\bar{y}_2), g'_3(y_3\bar{y}_4\bar{y}_1\bar{y}_2), g'_4(y_3\bar{y}_4\bar{y}_1\bar{y}_2))]^{r_1} = \\
&= [\bar{g}'_4(y_3\bar{y}_4\bar{y}_1\bar{y}_2), \bar{g}'_3(y_3\bar{y}_4\bar{y}_1\bar{y}_2), \bar{g}'_2(y_3\bar{y}_4\bar{y}_1\bar{y}_2), \bar{g}'_1(y_3\bar{y}_4\bar{y}_1\bar{y}_2)] = \\
&= [y_1, \neg(y_1\bar{y}_3y_4 \vee \bar{y}_1\bar{y}_3\bar{y}_4 \vee y_1y_3\bar{y}_4 \vee \bar{y}_1y_3y_4), \neg(\bar{y}_2\bar{y}_4 \vee \bar{y}_1\bar{y}_2\bar{y}_3 \vee y_1\bar{y}_2y_3 \vee y_1y_2\bar{y}_3y_4 \vee \bar{y}_1y_2y_3y_4), \\
&\quad \bar{y}_2\bar{y}_4 \vee y_1\bar{y}_2\bar{y}_3 \vee y_1y_2y_3y_4 \vee \bar{y}_1y_2\bar{y}_3y_4 \vee \bar{y}_1\bar{y}_2y_3]; \\
g^{k^k}(x) &= q_2(g^a(q_1(x^{r_1})))^{r_2} = q_2(g^a(x_4\bar{x}_3\bar{x}_2\bar{x}_1))^{r_2} = \\
&= q_2((x_1x_2 \vee \bar{x}_1x_3), \neg(x_4 \oplus \bar{x}_3 \oplus \bar{x}_2 \oplus \bar{x}_1), (x_2x_3 \oplus \bar{x}_2\bar{x}_3), x_4)^{r_2} = \\
&= q_2((x_1x_2 \vee \bar{x}_1x_3), \neg(x_1 \oplus x_2 \oplus x_3 \oplus x_4), \neg(x_2x_3 \oplus \bar{x}_2\bar{x}_3), \bar{x}_4) = \\
&= (\bar{x}_2x_3 \vee x_2\bar{x}_3, \bar{x}_4, x_1x_2 \vee \bar{x}_1x_3, \neg(x_1 \oplus x_2 \oplus x_3 \oplus x_4)).
\end{aligned}$$

Suppose, we want to encrypt the plaintext $m = x_1x_2x_3x_4 = 1010$, applying the scheme $\mathcal{E}1$. We compute $\gamma(m) = \gamma(1010) = g^k(1010) = 1110$, $g^a(m) = g^a(1010) = 0101$, $\delta(m) = \delta(1010) = g^k(1010) \oplus g^a(1010) = 1110 \oplus 0101 = 1011$ and obtain the ciphertext $(\gamma(m), \delta(m)) = (1110, 1011)$. To decrypt this ciphertext, we compute $g^{a^{-1}}(\gamma(m) \oplus \delta(m)) = g^{a^{-1}}(1110 \oplus 1011) = g^{a^{-1}}(0101) = 1010 = m$.

Suppose, we also want to encrypt the same plaintext $m = 1010$, applying the scheme $\mathcal{E}2$. In this case, we compute $\gamma(m) = g^{a^k}(1010) = 1101$, $g^a(m) = 0101$, $\delta(m) = g^{a^k}(1010) \oplus g^a(1010) = 1101 \oplus 0101 = 1000$ and obtain the ciphertext $(\gamma(m), \delta(m)) = (1101, 1000)$. To decrypt this ciphertext, we compute $g^{a^{-1}}(\gamma(m) \oplus \delta(m)) = g^{a^{-1}}(1101 \oplus 1000) = g^{a^{-1}}(0101) = 1010 = m$.

Now, by applying to $m = 1010$ the encryption scheme $\mathcal{E}3$ under $u = 1100$, we obtain $g^k(u) = 1011$, $\gamma = g^a(g^k(u)) = 1001$, $\delta = g^k(u) \oplus m = 0001$, $g^{a^{-1}}(\gamma) \oplus \delta = 1011 \oplus 0001 = 1010 = m$.

At last, by applying to $m = 1010$ the encryption scheme $\mathcal{E}4$ under $u = 1100$, we obtain $\gamma = g^a(u) = 1101$, $\delta = u \oplus m = 0110$, $g^{a^{-1}}(\gamma) \oplus \delta = 1100 \oplus 0110 = 1010 = m$.

5. ElGamal signature scheme on Boolean functions

The ElGamal signature schemes are all randomized ones, as are all ElGamal encryption schemes. This means that there are many valid signatures for any given message, as are many ciphertexts for any given plaintext. It is known (see, for instance, [4]) there is a method by which an adversary can sign a random message m without knowing the private key by choosing (γ, δ) and m simultaneously. Any adversary knowing a valid signature (γ, δ) for a message m can also sign various other messages [4]. Both of these methods for producing the valid forged signatures do not “enable an opponent to forge a signature on a message of his own choosing”. The ElGamal signature schemes on Boolean functions described in this paper below enable an adversary, knowing a valid signature (γ, δ) for a message m , to produce valid forged signatures (γ', δ') for the same message m and do not seem to represent a threat to the security of our ElGamal signature schemes, as do not these methods to the security of the ElGamal signature schemes on groups.

Each of encryption schemes $\mathcal{E}1$ – $\mathcal{E}4$ becomes a signature scheme with appendix after appointing keys and equations to play the proper roles in it. So we obtain the following ElGamal signature schemes on Boolean functions. In the description of them, the terms that are not explained once more have the former meanings.

5.1. Signature scheme $\mathcal{S}1$

Private key (for signing) is $\{g(x), a\}$, *public key* (for verifying) is $g^{a^{-1}}(x)$.

Signing: m is a message, $m \in \mathbb{F}_2^n$; $\gamma(m) = g^k(m)$, $\delta(m) = g^k(m) \oplus g^a(m)$, $k \in_R V_L$; $(\gamma(m), \delta(m))$ is the signature.

Verification: accept the signature iff $g^{a^{-1}}(\gamma(m) \oplus \delta(m)) = m$.

5.2. Signature scheme $\mathcal{S}2$

Private key (for signing) is $g^a(x)$, *public key* (for verifying) is $g^{a^{-1}}(x)$, *secret parameter* is a .

Signing: m is a message, $m \in \mathbb{F}_2^n$; $\gamma(m) = g^{a^k}(m)$, $\delta(m) = g^{a^k}(m) \oplus g^a(m)$, $k \in_R V_L$; $(\gamma(m), \delta(m))$ is the signature.

Verification: accept the signature iff $g^{a^{-1}}(\gamma(m) \oplus \delta(m)) = m$.

5.3. Signature scheme $\mathcal{S}3$

Private key (for signing) is $\{g(x), a\}$, *public key* (for verifying) is $g^{a^{-1}}(x)$.

Signing: m is a message, $m \in \mathbb{F}_2^n$; $k \in_R V_L$, $u \in_R \mathbb{F}_2^n$; $\gamma = g^a(g^k(u))$, $\delta = g^k(u) \oplus g^a(m)$; (γ, δ) is the signature.

Verification: accept the signature iff $g^{a^{-1}}(g^{a^{-1}}(\gamma) \oplus \delta) = m$.

5.4. Signature scheme $\mathcal{S}4$

Private key (for signing) is $g^a(x)$, *public key* (for verifying) is $g^{a^{-1}}(x)$, *secret parameter* is a .

Signing: m is a message, $m \in \mathbb{F}_2^n$; $u \in_R \mathbb{F}_2^n$; $\gamma = g^a(u)$, $\delta = u \oplus g^a(m)$; (γ, δ) is the signature.

Verification: accept the signature iff $g^{a^{-1}}(g^{a^{-1}}(\gamma) \oplus \delta) = m$.

5.5. Signature scheme $\mathcal{S}5$

Private key (for signing) is $g^a(x)$, *public key* (for verifying) is $g^{a^{-1}}(x)$, *secret parameter* is a .

Signing: m is a message, $m \in \mathbb{F}_2^n$; $u \in_R \mathbb{F}_2^n$; $\gamma = u$, $\delta = u \oplus g^a(m)$; (γ, δ) is the signature.

Verification: accept the signature iff $g^{a^{-1}}(\gamma \oplus \delta) = m$.

Conclusion

We should say that the paper doesn't provide a solution of a research problem. We have only described a new approach to constructing ElGamal encryption and signature schemes by using the algebra of bijective vector Boolean functions with the negation and permutation operations on the sets of variables and coordinate functions in them. We are not really sure whether the given schemes are secure or not. Naturally this approach has begot quite a large number of new problems for a subsequent research. These problems are directly related to the cryptanalysis of new ElGamal cryptographic schemes described (or not yet) in the paper, to constructing ElGamal signature schemes on Boolean functions with message recovery, and to the development of the used algebra. Computational methods and estimates of their complexity are the most important subject in researching the last.

Acknowledgements

I would like to thank my colleagues Irina A. Pankratova for reading and editing the manuscript and for suggesting me the encryption scheme $\mathcal{E}4$, and Vitaliy A. Romankov for suggesting me the encryption scheme $\mathcal{E}3$.

REFERENCES

1. *Menezes A., van Oorshot P., and Vanstone S.* Handbook of Applied Cryptography. CRC Press Inc., 1997. 661 p.
2. *Agibalov G. P.* Substitution block ciphers with functional keys. *Prikladnaya Diskretnaya Matematika*, 2017, no. 38, pp. 57–65.
3. *Agibalov G. P. and Pankratova I. A.* Asymmetric cryptosystems on Boolean functions. *Prikladnaya Diskretnaya Matematika*, 2018, no. 40, pp. 23–33.
4. *Stinson D. R.* Cryptography: Theory and Practice. CRC Press Inc., 1995. 434 p.

ПРИКЛАДНАЯ ТЕОРИЯ ГРАФОВ

УДК 519.1

ОБ ОДНОЙ ЗАДАЧЕ КЛАСТЕРИЗАЦИИ ГРАФА
С ЧАСТИЧНЫМ ОБУЧЕНИЕМ¹А. В. Ильев^{*,**}, В. П. Ильев^{*,***}

^{*} Омский государственный технический университет, г. Омск, Россия,

^{**} Институт математики им. С. Л. Соболева СО РАН, г. Омск, Россия,

^{***} Омский государственный университет им. Ф. М. Достоевского, г. Омск, Россия

В задачах кластеризации требуется разбить данное множество объектов на несколько подмножеств (кластеров) только на основе сходства объектов друг с другом. Рассматривается вариант задачи кластеризации графа, являющийся одной из формализаций задачи кластеризации с частичным обучением. Доказано, что эта задача является NP-трудной. Для одного варианта задачи предложен полиномиальный 3-приближённый алгоритм.

Ключевые слова: граф, кластер, кластеризация с частичным обучением.

DOI 10.17223/20710410/42/5

ON A SEMI-SUPERVISED GRAPH CLUSTERING PROBLEM

A. V. Il'ev^{*,**}, V. P. Il'ev^{*,***}

^{*} Omsk State Technical University, Omsk, Russia

^{**} Sobolev Institute of Mathematics SB RAS, Omsk, Russia

^{***} Dostoevsky Omsk State University, Omsk, Russia

E-mail: artyom_iljev@mail.ru, iljev@mail.ru

In the clustering problems one has to partition a given set of objects into some subsets (called clusters) taking into consideration only similarity of the objects. In this paper we study a version of the graph correlation clustering that can be considered as a formalization of the semi-supervised clustering. We prove that the problem under consideration is NP-hard and propose a polynomial-time 3-approximation algorithm for a version of the problem.

Keywords: graph, cluster, semi-supervised clustering.

Введение

Важный раздел теории распознавания образов и машинного обучения составляют методы решения задач кластеризации [1]. В задаче кластеризации требуется разбить данное множество объектов на несколько подмножеств (кластеров) только на основе сходства объектов друг с другом. Одной из наиболее наглядных формализаций

¹Работа поддержана грантом РФФ №17-11-01117.

задач кластеризации является *задача кластеризации графа* [2, 3]. В ней отношение сходства объектов задаётся посредством неориентированного графа, вершины которого взаимно однозначно соответствуют объектам, а рёбра соединяют похожие объекты, обладающие достаточным количеством одинаковых признаков. Требуется разбить множество исходных объектов на попарно непересекающиеся группы (кластеры) так, чтобы минимизировать число рёбер между кластерами и число недостающих рёбер внутри кластеров. Количество кластеров может быть задано, ограничено или заранее не определено.

В машинном обучении задачи кластеризации относят к разделу *обучения без учителя*. Наряду с этим рассматриваются также *задачи кластеризации с частичным обучением*, в которых часть объектов (как правило, небольшая) изначально распределена по кластерам [4, 5]. Задачи кластеризации графов имеют многочисленные приложения. Задачи кластеризации графов с частичным обучением тесно связаны с исследованием систем уравнений над графами [6].

В п. 1 настоящей работы рассматриваются три варианта задачи кластеризации графа, приводится краткий обзор известных результатов. В п. 2 рассматривается постановка задачи кластеризации графа, которая является одной из формализаций задачи кластеризации с частичным обучением. В этой задаче дано множество, состоящее из n объектов, которые необходимо распределить по k кластерам, $k < n$. Среди заданных объектов выделены k объектов, никакие два из которых не должны принадлежать одному и тому же кластеру. Отношение сходства объектов задано с помощью неориентированного графа. В работе доказано, что рассматриваемая задача является NP-трудной. Для случая $k = 2$ предложен полиномиальный 3-приближённый алгоритм.

1. Задачи кластеризации графов

1.1. Постановки задач

Будут рассматриваться только *обыкновенные* графы, т. е. графы без петель и кратных рёбер. Обыкновенный граф называется *кластерным*, если каждая его компонента связности является полным графом [7].

Обозначим через $\mathcal{M}(V)$ множество всех кластерных графов на множестве вершин V , $\mathcal{M}_k(V)$ — множество всех кластерных графов на множестве вершин V , имеющих ровно k непустых компонент связности, $\mathcal{M}_{1,k}(V)$ — множество всех кластерных графов на множестве V , имеющих не более k компонент связности, $2 \leq k \leq |V|$.

Если $G_1 = (V, E_1)$ и $G_2 = (V, E_2)$ — обыкновенные графы на одном и том же множестве вершин V , то *расстояние* $\rho(G_1, G_2)$ между ними определяется как

$$\rho(G_1, G_2) = |E_1 \Delta E_2| = |E_1 \setminus E_2| + |E_2 \setminus E_1|,$$

т. е. $\rho(G_1, G_2)$ — число несовпадающих рёбер в графах G_1 и G_2 .

В 60–80-е годы XX века в литературе изучались следующие три варианта задачи кластеризации графа, которые в то время принято было называть *задачами аппроксимации графов*. Постановки и различные интерпретации задач аппроксимации графов можно найти в [8–11]. В дальнейшем задачи аппроксимации графов неоднократно переоткрывались и независимо изучались под разными названиями (Correlation Clustering [12], Cluster Editing [7, 13] и др.).

Задача \mathbf{GC} . Дан обыкновенный граф $G = (V, E)$. Найти такой граф $M^* \in \mathcal{M}(V)$, что

$$\rho(G, M^*) = \min_{M \in \mathcal{M}(V)} \rho(G, M).$$

Задача \mathbf{GC}_k . Дан обыкновенный граф $G = (V, E)$ и целое число k , $2 \leq k \leq |V|$. Найти такой граф $M^* \in \mathcal{M}_k(V)$, что

$$\rho(G, M^*) = \min_{M \in \mathcal{M}_k(V)} \rho(G, M).$$

Задача $\mathbf{GC}_{1,k}$. Дан обыкновенный граф $G = (V, E)$ и целое число k , $2 \leq k \leq |V|$. Найти такой граф $M^* \in \mathcal{M}_{1,k}(V)$, что

$$\rho(G, M^*) = \min_{M \in \mathcal{M}_{1,k}(V)} \rho(G, M).$$

Рассматривались также варианты задач кластеризации графов, в которых ограничения накладываются на размеры кластеров [14].

1.2. Вычислительная сложность и приближённые алгоритмы

В 1986 г. М. Крживанек и Дж. Моравек показали [15], что задача \mathbf{GC} является NP-трудной, однако их работа осталась незамеченной. В 2004 г. в [7, 12] доказана NP-трудность задачи \mathbf{GC} . В [7] доказано также, что задача \mathbf{GC}_k NP-трудна при любом фиксированном $k \geq 2$; в [16] приведено более простое доказательство этого же результата. В [17] доказано, что задачи \mathbf{GC}_2 и $\mathbf{GC}_{1,2}$ NP-трудны уже на кубических графах, откуда следует, что все упомянутые ранее варианты задачи кластеризации графа являются NP-трудными, включая и задачу $\mathbf{GC}_{1,k}$.

В 2004 г. в [12] предложен полиномиальный 3-приближённый алгоритм для задачи $\mathbf{GC}_{1,2}$. В 2006 г. в [17] доказано существование рандомизированной полиномиальной приближённой схемы для задачи $\mathbf{GC}_{1,2}$, а в [16] предложена рандомизированная полиномиальная приближённая схема для задачи \mathbf{GC}_k (для любого фиксированного $k \geq 2$). Указав, что сложность полиномиальной приближённой схемы из [16] лишает её перспективы практического использования, авторы [18] предложили 2-приближённый алгоритм для задачи $\mathbf{GC}_{1,2}$, применив процедуру локального поиска к допустимому решению, полученному с помощью 3-приближённого алгоритма из [12]. Для задачи \mathbf{GC}_2 в работе [19] предложен $(3 - 6/n)$ -приближённый алгоритм с достижимой гарантированной оценкой точности. В 2005 г. авторы [20] показали, что задача \mathbf{GC} является APX-трудной и разработали для неё 4-приближённый алгоритм. В 2008 г. в [21] предложен 2,5-приближённый алгоритм для задачи \mathbf{GC} .

Более подробный обзор результатов, относящихся к задачам \mathbf{GC} , \mathbf{GC}_k , $\mathbf{GC}_{1,k}$ и их взвешенным вариантам, можно найти в [22].

2. Задача кластеризации с частичным обучением

2.1. Вычислительная сложность

Рассмотрим задачу кластеризации графа с частичным обучением, самая общая постановка которой такова.

Дан обыкновенный n -вершинный граф $G = (V, E)$ и целое число k , $2 \leq k \leq n$. Выделено множество попарно различных вершин $X = \{x_1, \dots, x_k\} \subset V$, на котором задано бинарное симметричное иррефлексивное отношение P . Требуется найти ближайший к G граф $M^* \in \mathcal{M}(V)$, такой, что никакие две вершины множества X , связанные

отношением P , не принадлежат одной и той же компоненте связности (т. е. одному кластеру) графа M^* .

Более подробно изучим следующий вариант задачи кластеризации графа с частичным обучением.

Задача \mathbf{SGC}_k . Дан обыкновенный n -вершинный граф $G = (V, E)$ и целое число k , $2 \leq k \leq n$. Выделено множество попарно различных вершин $X = \{x_1, \dots, x_k\} \subset V$. Требуется найти такой граф $M^* \in \mathcal{M}_k(V)$, что

$$\rho(G, M^*) = \min_{M \in \mathcal{M}_k(V)} \rho(G, M),$$

причём минимум берётся по всем кластерным графам $M = (V, E_M) \in \mathcal{M}_k(V)$, в которых $x_i, x_j \notin E_M$ для любых $i, j \in \{1, \dots, k\}$; другими словами, никакие две вершины множества $X = \{x_1, \dots, x_k\}$ не принадлежат одной и той же компоненте связности (т. е. одному кластеру) графа M .

Несложно свести по Тьюрингу задачу \mathbf{GC}_k к \mathbf{SGC}_k и тем самым показать, что задача \mathbf{SGC}_k NP-трудна. Действительно, рассмотрим произвольный граф $G = (V, E)$ — вход задачи \mathbf{GC}_k — и фиксируем целое число k и произвольный набор $\{x_1, \dots, x_k\}$, состоящий из k попарно различных вершин графа G . Имея оптимальное решение $M(x_1, \dots, x_k)$ задачи \mathbf{SGC}_k для любого такого набора $\{x_1, \dots, x_k\} \subset V$ и выбрав среди них ближайший к графу G кластерный граф

$$M^*(x_1, \dots, x_k) = \arg \min_{\{x_1, \dots, x_k\} \subset V} \rho(G, M(x_1, \dots, x_k)),$$

мы, очевидно, получим оптимальное решение исходной задачи \mathbf{GC}_k . Легко видеть, что при фиксированном k построение всех C_n^k входов задачи \mathbf{SGC}_k и получение оптимального решения исходной задачи \mathbf{GC}_k можно выполнить за время $O(n^k)$, где $n = |V|$. Таким образом, справедлива следующая теорема.

Теорема 1. Задача \mathbf{SGC}_k NP-трудна при любом фиксированном $k \geq 2$.

2.2. Задача \mathbf{SGC}_2

Рассмотрим частный случай задачи \mathbf{SGC}_k , когда $k = 2$.

Пусть $G_1 = (V, E_1)$ и $G_2 = (V, E_2)$ — n -вершинные графы на одном и том же множестве вершин V . Обозначим через $D(G_1, G_2)$ граф на множестве вершин V с множеством рёбер $E_1 \Delta E_2$. Через $d_D(u)$ обозначим степень вершины u в графе $D = D(G_1, G_2)$.

Нетрудно заметить, что $\rho(G_1, G_2)$ равно числу рёбер в графе D , которое по лемме о рукопожатиях равно половине суммы степеней вершин графа D . Отсюда получаем следующее утверждение.

Лемма 1 [19]. Пусть $d_{\min} = \min_{u \in V} d_D(u)$ — минимум степеней вершин в графе $D = D(G_1, G_2)$. Тогда

$$\rho(G_1, G_2) \geq \frac{nd_{\min}}{2}.$$

Введём следующие обозначения. Для произвольного графа $G = (V, E)$ и $u \in V$ обозначим $N_G(u) = \{v \in V : uv \in E\}$, $\overline{N}_G(u) = V \setminus (N_G(u) \cup \{u\})$. Таким образом, $|N_G(u)| = d_G(u)$ — степень вершины u в графе G .

Для множеств $V_1, V_2 \subseteq V$, таких, что $V_1 \cap V_2 = \emptyset$ и $V_1 \cup V_2 = V$, обозначим через $M(V_1, V_2)$ кластерный граф из класса $\mathcal{M}_2(V)$ с компонентами связности, порождёнными множествами V_1, V_2 . Нам понадобится следующая простая лемма.

Лемма 2. Пусть $G = (V, E)$ — n -вершинный граф, $u \in V$ — произвольная вершина, $M, M' \in \mathcal{M}_2(V)$ — кластерные графы на множестве V , такие, что M' получен из M путём переноса вершины u в другую компоненту связности. Тогда

$$\rho(G, M') - \rho(G, M) \leq n - 1.$$

Доказательство. Пусть $M = M(V_1, V_2)$, где V_1 — множество, содержащее вершину u . Тогда $M' = M(V_1 \setminus \{u\}, V_2 \cup \{u\})$. Заметим, что графы $D = D(G, M)$ и $D' = D(G, M')$ отличаются только рёбрами вида uv , $v \in V \setminus \{u\}$, а именно: u, v смежны в D тогда и только тогда, когда они не смежны в D' . Остальные рёбра графов D и D' совпадают. Следовательно,

$$\rho(G, M') - \rho(G, M) = |N_{D'}(u)| - |N_D(u)| = d_{D'}(u) - d_D(u).$$

Очевидно, $d_D(u) \geq 0$, $d_{D'}(u) \leq n - 1$, поэтому

$$\rho(G, M') - \rho(G, M) = d_{D'}(u) - d_D(u) \leq n - 1,$$

что и требовалось. ■

2.3. Приближённый алгоритм для задачи \mathbf{SGC}_2

Пусть $x_1, x_2 \in V$ — фиксированные вершины графа $G = (V, E)$, $x_1 \neq x_2$, $v \in V$ — произвольная вершина в G . Построим граф $M_v \in \mathcal{M}_2(V)$ по следующим правилам:

- (а) Если в графе G вершина v смежна ровно с одной из вершин x_1, x_2 и не совпадает с другой, то полагаем $M_v = M(V_1, V_2)$, где $V_1 = \{v\} \cup N_G(v)$, $V_2 = V \setminus V_1$.
- (б) Если в графе G вершина v смежна с обеими вершинами x_1, x_2 , то полагаем $M'_v = M(V'_1, V'_2)$. Здесь $V'_1 = (\{v\} \cup N_G(v)) \setminus \{x_2\}$, $V'_2 = V \setminus V'_1$, $M''_v = M(V''_1, V''_2)$, где $V''_1 = (\{v\} \cup N_G(v)) \setminus \{x_1\}$, $V''_2 = V \setminus V''_1$. Если при этом $\rho(G, M'_v) \leq \rho(G, M''_v)$, то полагаем $M_v = M'_v$, в противном случае $M_v = M''_v$.
- (в) Если в графе G вершина v не смежна и не совпадает ни с одной из вершин x_1, x_2 , то полагаем $M'_v = M(V'_1, V'_2)$. Здесь $V'_1 = \{v\} \cup N_G(v) \cup \{x_1\}$, $V'_2 = V \setminus V'_1$, $M''_v = M(V''_1, V''_2)$, где $V''_1 = \{v\} \cup N_G(v) \cup \{x_2\}$, $V''_2 = V \setminus V''_1$. Если при этом $\rho(G, M'_v) \leq \rho(G, M''_v)$, то полагаем $M_v = M'_v$, в противном случае $M_v = M''_v$.
- (г) Если вершина v совпадает с одной из вершин x_1, x_2 , то полагаем $M_v = M(V_1, V_2)$ для $V_1 = (\{v\} \cup N_G(v)) \setminus \{x\}$, $V_2 = V \setminus V_1$, где $x = x_1$, если $v = x_2$, и $x = x_2$, если $v = x_1$.

Очевидно, что построенный по этим правилам кластерный граф M_v является допустимым решением задачи \mathbf{SGC}_2 для графа G .

Лемма 3. Пусть $M^* \in \mathcal{M}_2(V)$ — оптимальное решение задачи \mathbf{SGC}_2 на графе $G = (V, E)$, $D = D(G, M^*)$, $v \in V$ — вершина минимальной степени в графе D : $d_D(v) = \min_{u \in V} d_D(u) = d_{\min}$. Тогда

$$\rho(G, M_v) \leq \rho(G, M^*) + d_{\min}(n - 1), \quad (1)$$

где $M_v = M(V_1, V_2)$ — кластерный граф, построенный по правилам (а)–(г).

Доказательство. Пусть $M^* = M(V_1^*, V_2^*)$, где V_1^* — множество, содержащее вершину v . Тогда из определения графа D вытекает, что

$$V_1^* = \{v\} \cup \left(N_G(v) \setminus N_D(v) \right) \cup \left(\overline{N_G(v)} \cap N_D(v) \right). \quad (2)$$

С л у ч а й 1. Пусть в графе $G = (V, E)$ вершина v смежна ровно с одной из вершин x_1, x_2 и не совпадает с другой. Покажем, что граф M_v , построенный по правилу (а), может быть получен из графа M^* путем переноса d_{\min} вершин в другую компоненту. Для этого достаточно оценить, как сильно отличаются множества V_1 и V_1^* . По построению графа M_v имеем

$$V_1 = \{v\} \cup N_G(v) = \{v\} \cup \left(N_G(v) \setminus N_D(v) \right) \cup \left(N_G(v) \cap N_D(v) \right).$$

Учитывая (2), получаем

$$V_1^* \Delta V_1 = (V_1^* \setminus V_1) \cup (V_1 \setminus V_1^*) = \left(\overline{N_G}(v) \cap N_D(v) \right) \cup \left(N_G(v) \cap N_D(v) \right) = N_D(v).$$

Таким образом, $|V_1^* \Delta V_1| = |N_D(v)| = d_D(v) = d_{\min}$, поэтому граф M_v может быть получен из графа M^* путём переноса d_{\min} вершин множества $N_D(v)$ в другую компоненту связности.

Нетрудно заметить, что если в графе M^* последовательно переносить все вершины множества $N_D(v)$ в другую компоненту, то после каждого переноса для вновь получаемых кластерных графов будут выполнены условия леммы 2. Так как граф M_v получен из графа M^* путём переноса d_{\min} вершин множества $N_D(v)$ в другую компоненту, то, применяя d_{\min} раз лемму 2, получаем $\rho(G, M_v) - \rho(G, M^*) \leq d_{\min}(n - 1)$, т. е. $\rho(G, M_v) \leq \rho(G, M^*) + d_{\min}(n - 1)$. Итак, неравенство (1) выполнено.

С л у ч а й 2. Пусть в графе $G = (V, E)$ вершина v смежна с обеими вершинами x_1, x_2 , т. е. $x_1, x_2 \in N_G(v)$. В этом случае граф M_v строится по правилу (б). Так как вершины x_1 и x_2 находятся в разных компонентах связности графа M^* и $x_1, x_2 \in N_G(v)$, то $d_{\min} \geq 1$. Покажем, что один из графов M'_v, M''_v может быть получен из графа M^* путём переноса $d_{\min} - 1$ вершин в другую компоненту связности.

Рассмотрим случай, когда $x_1 \in V_1^*$ (следовательно, $x_2 \in V_2^*$). Докажем, что в этом случае граф M'_v может быть получен из графа M^* путём переноса $d_{\min} - 1$ вершин в другую компоненту связности. По построению графа M'_v справедливо

$$V'_1 = \left(\{v\} \cup N_G(v) \right) \setminus \{x_2\} = \left(\{v\} \cup \left(N_G(v) \setminus N_D(v) \right) \cup \left(N_G(v) \cap N_D(v) \right) \right) \setminus \{x_2\}.$$

В силу (2), учитывая, что $x_2 \in N_G(v) \cap N_D(v)$, получаем

$$V_1^* \Delta V'_1 = \left(\left(\overline{N_G}(v) \cap N_D(v) \right) \cup \left(N_G(v) \cap N_D(v) \right) \right) \setminus \{x_2\} = N_D(v) \setminus \{x_2\},$$

т. е. $|V_1^* \Delta V'_1| = |N_D(v)| - 1 = d_{\min} - 1$, поэтому граф M'_v может быть получен из графа M^* путём переноса $d_{\min} - 1$ вершин множества $N_D(v) \setminus \{x_2\}$ в другую компоненту связности.

Применяя $d_{\min} - 1$ раз лемму 2, получаем $\rho(G, M'_v) - \rho(G, M^*) \leq (d_{\min} - 1)(n - 1)$, т. е. $\rho(G, M'_v) \leq \rho(G, M^*) + (d_{\min} - 1)(n - 1)$.

В случае $x_2 \in V_1^*$ (значит, $x_1 \in V_2^*$) с помощью аналогичных рассуждений доказывается неравенство $\rho(G, M''_v) \leq \rho(G, M^*) + (d_{\min} - 1)(n - 1)$. Таким образом,

$$\begin{aligned} \rho(G, M_v) &= \min \left(\rho(G, M'_v), \rho(G, M''_v) \right) \leq \rho(G, M^*) + (d_{\min} - 1)(n - 1) < \\ &< \rho(G, M^*) + d_{\min}(n - 1), \end{aligned}$$

т. е. неравенство (1) выполнено.

С л у ч а й 3. Пусть в графе $G = (V, E)$ вершина v не смежна и не совпадает ни с одной из вершин x_1, x_2 , т.е. $x_1, x_2 \in \overline{N_G(v)}$. В этом случае граф M_v строится по правилу (в). Так как вершины x_1 и x_2 находятся в разных компонентах связности графа M^* и $x_1, x_2 \in \overline{N_G(v)}$, то $d_{\min} \geq 1$. Покажем, что один из графов M'_v, M''_v может быть получен из графа M^* путём переноса $d_{\min} - 1$ вершин в другую компоненту связности.

Рассмотрим случай, когда $x_1 \in V_1^*$ (следовательно, $x_2 \in V_2^*$). Докажем, что в этом случае граф M'_v может быть получен из графа M^* путем переноса $d_{\min} - 1$ вершин в другую компоненту связности. По построению графа M'_v ,

$$V'_1 = \{v\} \cup N_G(v) \cup \{x_1\} = \{v\} \cup \left(N_G(v) \setminus N_D(v) \right) \cup \left(N_G(v) \cap N_D(v) \right) \cup \{x_1\}.$$

В силу (2), учитывая, что $x_1 \in \overline{N_G(v)} \cap N_D(v)$, получаем

$$V_1^* \Delta V'_1 = \left(\left(\overline{N_G(v)} \cap N_D(v) \right) \setminus \{x_1\} \right) \cup \left(N_G(v) \cap N_D(v) \right) = N_D(v) \setminus \{x_1\},$$

т.е. $|V_1^* \Delta V'_1| = |N_D(v)| - 1 = d_{\min} - 1$, поэтому граф M'_v может быть получен из графа M^* путём переноса $d_{\min} - 1$ вершин множества $N_D(v) \setminus \{x_1\}$ в другую компоненту. Рассуждая, как в случае 2, получаем $\rho(G, M'_v) \leq \rho(G, M^*) + (d_{\min} - 1)(n - 1)$.

В случае $x_2 \in V_1^*$ (значит, $x_1 \in V_2^*$) аналогично доказывается неравенство $\rho(G, M''_v) \leq \rho(G, M^*) + (d_{\min} - 1)(n - 1)$. Таким образом,

$$\rho(G, M_v) = \min \left(\rho(G, M'_v), \rho(G, M''_v) \right) < \rho(G, M^*) + d_{\min}(n - 1),$$

что и требовалось доказать.

С л у ч а й 4. Пусть в графе $G = (V, E)$ вершина v совпадает с одной из вершин x_1, x_2 . В этом случае граф M_v строится по правилу (г). Без ограничения общности будем считать, что $v = x_1$. Тогда $V_1 = \left(\{v\} \cup N_G(v) \right) \setminus \{x_2\}$. Если $x_2 \notin N_G(v)$, то $V_1 = \{v\} \cup N_G(v)$ и неравенство (1) доказывается, как в случае 1, а если $x_2 \in N_G(v)$, то как в случае 2.

Итак, в каждом из четырёх случаев справедливо неравенство (1). ■

Рассмотрим следующий алгоритм приближённого решения задачи **SGC₂**.

Алгоритм 1. Приближённое решение задачи **SGC₂**

- 1: Для каждой вершины $v \in V$ определить кластерный граф $M_v \in \mathcal{M}_2(V)$ по правилам (а)–(г).
 - 2: Среди всех графов M_v выбрать такой граф M , что $\rho(G, M) = \min_{v \in V} \rho(G, M_v)$.
-

Вычислительная сложность алгоритма 1 может быть оценена как $O(n^3)$, где n — число вершин графа G .

Справедлива следующая гарантированная оценка точности алгоритма 1.

Теорема 2. Для любого n -вершинного графа $G = (V, E)$ и любого множества $X = \{x_1, x_2\} \subset V$ ($x_1 \neq x_2$) алгоритм 1 находит такой кластерный граф $M \in \mathcal{M}_2(V)$ — допустимое решение задачи **SGC₂**, что

$$\rho(G, M) \leq \left(3 - \frac{2}{n} \right) \rho(G, M^*),$$

где $M^* \in \mathcal{M}_2(V)$ — оптимальное решение задачи **SGC₂** на графе G .

Доказательство. Очевидно, среди всех вершин на шаге 1 будет выбрана такая вершина $v \in V$, что $d_D(v) = d_{\min}$, поэтому $\rho(G, M) \leq \rho(G, M_v)$. Оценим величину $\rho(G, M_v)$ с учётом неравенства (1) и леммы 1:

$$\rho(G, M_v) \leq \rho(G, M^*) + d_{\min}(n-1) \leq \rho(G, M^*) + 2\rho(G, M^*) \left(1 - \frac{1}{n}\right) = \left(3 - \frac{2}{n}\right)\rho(G, M^*).$$

Отсюда получаем гарантированную оценку точности алгоритма 1:

$$\rho(G, M) \leq \rho(G, M_v) \leq \left(3 - \frac{2}{n}\right)\rho(G, M^*).$$

Теорема 2 доказана. ■

ЛИТЕРАТУРА

1. Журавлев Ю. И., Рязанов В. В., Сенько О. В. Распознавание. Математические методы. Программная система. Практические применения. М.: ФАЗИС, 2005.
2. Kulis B., Basu S., Dhillon I., and Mooney R. Semi-supervised graph clustering: a kernel approach // Mach. Learn. 2009. V. 74. No. 1. P. 1–22.
3. Schaeffer S. E. Graph clustering // Comput. Sci. Rev. 2005. V. 1. No. 1. P. 27–64.
4. Bair E. Semi-supervised clustering methods // Wiley Interdisciplinary Reviews: Computational Statistics. 2013. V. 5. No. 5. P. 349–361.
5. Chapelle O., Schölkopf B., and Zein A. Semi-supervised Learning. Cambridge, Massachusetts: MIT Press, 2006.
6. Ильев А. В., Ремесленников В. Н. Исследование совместности систем уравнений над графами и нахождение их общих решений // Вестник Омского университета. 2017. № 4(86). С. 26–32.
7. Shamir R., Sharan R., and Tsur D. Cluster graph modification problems // Discrete Appl. Math. 2004. V. 144. No. 1–2. P. 173–182.
8. Zahn C. T. Approximating symmetric relations by equivalence relations // J. Soc. Indust. Appl. Math. 1964. V. 12. No. 4. P. 840–847.
9. Tomescu I. La reduction minimale d'un graphe à une reunion de cliques // Discrete Math. 1974. V. 10. No. 1–2. P. 173–179.
10. Фридман Г. Ш. Исследование одной задачи классификации на графах // Методы моделирования и обработка информации. Новосибирск: Наука, 1976. С. 147–177.
11. Ильев В. П., Фридман Г. Ш. К задаче аппроксимации графами с фиксированным числом компонент // Докл. АН СССР. 1982. Т. 264. № 3. С. 533–538.
12. Bansal N., Blum A., and Chawla S. Correlation clustering // Mach. Learn. 2004. V. 56. No. 1–3. P. 89–113.
13. Ben-Dor A., Shamir R., and Yakhimi Z. Clustering gene expression patterns // J. Comput. Biol. 1999. V. 6. No. 3–4. P. 281–297.
14. Ильев В. П., Навроцкая А. А. Вычислительная сложность задачи аппроксимации графами с компонентами связности ограниченного размера // Прикладная дискретная математика. 2011. № 3(13). С. 80–84.
15. Křivánek M. and Morávek J. NP-hard problems in hierarchical-tree clustering // Acta Inform. 1986. V. 23. P. 311–323.
16. Giotis I. and Guruswami V. Correlation clustering with a fixed number of clusters // Theory Comput. 2006. V. 2. No. 1. P. 249–266.
17. Агеев А. А., Ильев В. П., Кононов А. В., Талевнин А. С. Вычислительная сложность задачи аппроксимации графов // Дискрет. анализ и исслед. операций. Сер. 1. 2006. Т. 13. № 1. С. 3–11.

18. *Coleman T., Saunderson J., and Wirth A.* A local-search 2-approximation for 2-correlation clustering // ESA 2008. LNCS. 2008. V. 5193. P. 308–319.
19. *Ильев В. П., Ильева С. Д., Навроцкая А. А.* Приближенные алгоритмы для задач аппроксимации графов // Дискрет. анализ и исслед. операций. 2011. Т. 18. № 1. С. 41–60.
20. *Charikar M., Guruswami V., and Wirth A.* Clustering with qualitative information // J. Comput. Syst. Sci. 2005. V. 71. No. 3. P. 360–383.
21. *Ailon N., Charikar M., and Newman A.* Aggregating inconsistent information: Ranking and clustering // J. ACM. 2008. V. 55. No. 5. P. 1–27.
22. *Ильев В., Ильева С., Кононов А.* Short survey on graph correlation clustering with minimization criteria // DOOR 2016. LNCS. 2016. V. 9869. P. 25–36.

REFERENCES

1. *Zhuravlev Yu. I., Ryazanov V. V., and Sen'ko O. V.* Raspoznavanie. Matematicheskie metody. Programmnyaya sistema. Prakticheskie primeneniya [Recognition. Mathematical methods. Program system. Practical applications]. Moscow, FAZIS Publ., 2005. (in Russian)
2. *Kulis B., Basu S., Dhillon I., and Mooney R.* Semi-supervised graph clustering: a kernel approach. Mach. Learn., 2009, vol. 74, no. 1, pp. 1–2.
3. *Schaeffer S. E.* Graph clustering. Comput. Sci. Rev., 2005, vol. 1, no. 1, pp. 27–64.
4. *Bair E.* Semi-supervised clustering methods. Wiley Interdisciplinary Reviews: Computational Statistics. 2013, vol. 5, no. 5, pp. 349–361.
5. *Chapelle O., Schölkopf B., and Zein A.* Semi-supervised Learning. Cambridge, Massachusetts, MIT Press, 2006.
6. *Ильев А. В. and Ремесленников В. Н.* Исследование совместности систем уравнений над графами и нахождение их общих решений [Study of the compatibility of systems of equations over graphs and finding their general solutions]. Vestnik Omskogo Universiteta, 2017, no. 4(86), pp. 26–32. (in Russian)
7. *Shamir R., Sharan R., and Tsur D.* Cluster graph modification problems. Discrete Appl. Math., 2004, vol. 144, no. 1–2, pp. 173–182.
8. *Zahn C. T.* Approximating symmetric relations by equivalence relations. J. Soc. Indust. Appl. Math., 1964, vol. 12, no. 4, pp. 840–847.
9. *Tomescu I.* La reduction minimale d'un graphe à une reunion de cliques. Discrete Math., 1974, vol. 10, no. 1–2, pp. 173–179.
10. *Fridman G. Sh.* Исследование одной задачи классификации на графах [Investigation of a classifying problem on graphs]. In: Metody Modelirovaniya i Obrabotka Informatsii, Novosibirsk, Nauka Publ., 1976, pp. 147–177. (in Russian)
11. *Ильев В. П. and Fridman G. Sh.* On the problem of approximation by graphs with a fixed number of components. Sov. Math. Dokl., 1982, vol. 25, pp. 666–670.
12. *Bansal N., Blum A., and Chawla S.* Correlation clustering. Mach. Learn., 2004, vol. 56, no. 1–3, pp. 89–113.
13. *Ben-Dor A., Shamir R., and Yakhimi Z.* Clustering gene expression patterns. J. Comput. Biol., 1999, vol. 6, no. 3–4, pp. 281–297.
14. *Ильев В. П. and Навроцкая А. А.* Выхислительная сложность задачи аппроксимации графами с компонентами связности ограниченного размера [Computational complexity of the problem of approximation by graphs with connected components of bounded size]. Prikladnaya Diskretnaya Matematika, 2011, no. 3(13), pp. 80–84. (in Russian)
15. *Křivánek M. and Morávek J.* NP-hard problems in hierarchical-tree clustering. Acta inform., 1986, vol. 23, pp. 311–323.

16. *Giotis I. and Guruswami V.* Correlation clustering with a fixed number of clusters. *Theory Comput.*, 2006, vol. 2, no. 1, pp. 249–266.
17. *Ageev A. A., Il'ev V. P., Kononov A. V., and Talevnin A. S.* Computational complexity of the graph approximation problem. *J. Appl. Indust. Math.*, 2007, vol. 1, no. 1, pp. 1–8.
18. *Coleman T., Saunderson J., and Wirth A.* A local-search 2-approximation for 2-correlation clustering. *ESA 2008, LNCS*, 2008, vol. 5193, pp. 308–319.
19. *Il'ev V. P., Il'eva S. D., and Navrotskaya A. A.* Approximation algorithms for graph approximation problems. *J. Appl. Indust. Math.*, 2011, vol. 5, no. 4, pp. 569–581.
20. *Charikar M., Guruswami V., and Wirth A.* Clustering with qualitative information. *J. Comput. Syst. Sci.*, 2005, vol. 71, no. 3, pp. 360–383.
21. *Ailon N., Charikar M., and Newman A.* Aggregating inconsistent information: Ranking and clustering. *J. ACM*, 2008, vol. 55, no. 5, pp. 1–27.
22. *Il'ev V., Il'eva S., and Kononov A.* Short survey on graph correlation clustering with minimization criteria. *DOOR 2016, LNCS*, 2016, vol. 9869, pp. 25–36.

УДК 519.17

**ДЕТЕРМИНИРОВАННЫЕ МЕТОДЫ ПОСТРОЕНИЯ ГРАФОВ
РАМАНУДЖАНА, ПРЕДНАЗНАЧЕННЫХ ДЛЯ ПРИМЕНЕНИЯ
В КРИПТОГРАФИЧЕСКИХ АЛГОРИТМАХ, ОСНОВАННЫХ
НА ОБОБЩЁННЫХ КЛЕТОЧНЫХ АВТОМАТАХ¹**

П. Г. Ключарёв

*Московский государственный технический университет им. Н. Э. Баумана, г. Москва,
Россия*

Рассматриваются детерминированные методы построения графов Рамануджана в контексте их применения в качестве графов обобщённых клеточных автоматов, предназначенных для использования в криптографии. Изучены два семейства графов Любоцкого — Филиппса — Сарнака ($X^{p,q}$ и $Y^{p,q}$), семейство графов Пайзера и семейство графов Моргенштерна. Сделан вывод, что для применения в указанном качестве подходят графы Пайзера и графы $Y^{p,q}$. Приведены значения параметров графов из этих семейств, полученные численно.

Ключевые слова: *расширяющий граф, граф Рамануджана.*

DOI 10.17223/20710410/42/6

**DETERMINISTIC METHODS OF RAMANUJAN GRAPH
CONSTRUCTION FOR USE IN CRYPTOGRAPHIC ALGORITHMS
BASED ON GENERALIZED CELLULAR AUTOMATA**

P. G. Klyucharev

*Bauman Moscow State Technical University, Moscow, Russia***E-mail:** pk.iu8@yandex.ru

Earlier, the author proposed a number of methods for constructing symmetric cryptographic algorithms based on generalized cellular automata. In order to make such automata to be cryptographically strong, their graphs must satisfy a number of requirements. In particular, they must be regular not bipartite graphs with a small diameter, a small degree (but not less than 4) and the amount of graphs in the family with the number of vertices from dozens to several thousand must be large enough (it would be desirable to have at least several dozens of graphs with a number of vertices more or less uniformly distributed in the given range). Some of Ramanujan graphs satisfy these requirements. There are two ways to construct relatively small Ramanujan graph: the random way and the deterministic way. In this paper, the deterministic methods for Ramanujan graphs construction in the context of their application in generalized cellular automata being a base of cryptographic algorithms are considered. Each method can be identified with the family of graphs generated by it. Among them are two families of graphs constructed by Lubotzky, Philips and Sarnak — $X^{p,q}$ and $Y^{p,q}$, the family of graphs constructed by Pizer, and the family of graphs constructed by Morgenstern. Values of parameters of graphs from these families are numerically computed. After research, we came to conclusion that Pizer

¹Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 16-07-00542.

graphs (based on isogenies of elliptic curves over finite fields) and the $Y^{p,q}$ Lubotzky — Philips — Sarnak graphs (based on projective transformations of a projective line over a finite field) are suitable for the purposes under consideration, because, according to literature review, they meet all the necessary requirements, in particular, they are not bipartite, and among them there are sufficiently large amount of relatively small graphs with small degrees (all Ramanujan graphs are regular and have a small diameter). At the same time, the $X^{p,q}$ Lubotzky — Philips — Sarnak graphs and Morgenstern graphs are not suitable for considered purposes, because among them there are too few not bipartite graphs with a small degree and with a number of vertices in the desired range.

Keywords: *expander graph, Ramanujan graph.*

Введение

Ранее автором в ряде работ (в том числе в [1, 2]) предложены методы построения симметричных шифров и криптографических хэш-функций, основанных на обобщённых клеточных автоматах. Криптоалгоритмы, полученные с помощью этих методов, обладают рядом ценных свойств, в частности высокой производительностью при аппаратной реализации. Для того чтобы такие алгоритмы были криптостойкими, графы обобщённых клеточных автоматов должны удовлетворять ряду требований. Данная работа посвящена детерминированным методам построения таких графов.

1. Основные понятия

Будем использовать термин «граф», допуская наличие петель и кратных рёбер.

Обобщённым клеточным автоматом называется ориентированный граф (*граф обобщённого клеточного автомата*) с множеством вершин $V = \{v_1, \dots, v_N\}$, с каждой вершиной v_i которого ассоциированы:

- булева переменная m_i , которая называется *ячейкой*;
- булева функция $f_i(x_1, \dots, x_{d_i})$, которая называется локальной функцией связи вершины v_i (d_i — степень вершины v_i).

При этом каждой паре (v, e) , где $v \in V$ — вершина, e — входящее в неё ребро, соответствует номер аргумента локальной функции связи, вычисляемой в вершине v . Будем называть его *номером ребра e относительно вершины v* . Работа обобщённого клеточного автомата происходит следующим образом. В начальный момент времени каждая ячейка m_i , $i = 1, \dots, N$, принимает некоторое начальное значение $m_i(0)$. Автомат работает пошагово, значения ячеек на шаге номер t вычисляются по формуле

$$m_i(t) = f_i(m_{\eta(i,1)}(t-1), m_{\eta(i,2)}(t-1), \dots, m_{\eta(i,d_i)}(t-1)),$$

где $\eta(i, j)$ — номер вершины, из которой исходит ребро, входящее в вершину v_i и имеющее относительно этой вершины номер j .

Будем рассматривать только *неориентированные* обобщённые клеточные автоматы, т. е. такие, что для каждого ребра (u, v) в графе автомата существует и ребро (v, u) . Такой граф можно рассматривать как неориентированный, для чего достаточно заменить каждую пару ориентированных рёбер (u, v) и (v, u) на неориентированное ребро $\{u, v\}$. Будем рассматривать только обобщённые клеточные автоматы, графы которых являются регулярными ($d_1 = \dots = d_N = d$).

Обобщённый клеточный автомат представляет собой обобщение клеточного автомата [3], предложенного Дж. фон Нейманом. Подобные обобщения под разными на-

званиями использовались в различных областях (см., например, [4, 5]). Методы применения обобщённых клеточных автоматов в криптографии развиты в [1, 2].

2. Требования к графам

Как обосновано в предыдущих работах автора, для построения обобщённых клеточных автоматов, применяемых в составе криптографических алгоритмов, требуются связные неориентированные графы, обладающие следующими свойствами:

- граф должен иметь свойства, близкие к свойствам случайного графа;
- диаметр графа должен быть близок к минимально возможному;
- число петель и кратных рёбер в графе должно быть как можно меньшим;
- граф должен являться регулярным;
- степень графа должна быть как можно меньшей (для повышения эффективности аппаратной реализации обобщённого клеточного автомата);
- степень графа должна быть не меньше четырёх;
- граф не должен являться двудольным;
- количество графов с числом вершин от нескольких десятков до нескольких тысяч в семействе графов должно быть достаточно велико, чтобы в семействе существовали графы, годящиеся для построения обобщённых клеточных автоматов с различным числом ячеек (хотелось бы иметь хотя бы несколько десятков графов с числом вершин, более или менее равномерно распределённым в данном диапазоне).

Заметим, что в этом ряду требований словосочетание «как можно меньшее» следует понимать в приближённом смысле — достаточно лишь близость к минимально возможным значениям. Приведённым требованиям удовлетворяют графы Рамануджана [6–8].

3. Расширяющие графы

Графы Рамануджана изучаются в теории расширяющих графов. Это сравнительно молодая область дискретной математики, нашедшая много приложений в различных теоретических и прикладных областях математики и компьютерных наук. Ей посвящено большое количество работ, в том числе [6, 7, 9–12]. Приведём некоторые сведения из этой теории, необходимые для дальнейшего изложения.

Коэффициентом рёберного расширения неориентированного d -регулярного графа G с множеством вершин V называется величина

$$h(G) = \min_{\{S \subset V: 0 < |S| \leq |V|/2\}} \frac{|\partial S|}{|S|},$$

где $|\partial S|$ — число рёбер, каждое из которых соединяет вершину из множества S с вершиной из множества $V \setminus S$.

Расширяющим графом (expander graph) называется неориентированный регулярный граф G , для которого $h(G) \geq c$, где c — некоторая наперёд заданная положительная константа.

Коэффициент рёберного расширения графа связан с его спектральными свойствами. Спектр неориентированного графа — это набор собственных значений его матрицы смежности, отсортированный по невозрастанию: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$. Здесь и далее N — число вершин графа.

Как известно из спектральной теории графов, для d -регулярных графов $\lambda_1 = d$ и справедливо следующее *неравенство Чигера*:

$$\frac{1}{2}(d - \lambda_2) \leq h(G) \leq \sqrt{2d(d - \lambda_2)}.$$

Пусть $\lambda = \lambda(G) = \max_{|\lambda_i| < d} |\lambda_i|$. Графом Рамануджана называется связный d -регулярный неориентированный граф G , для которого справедливо следующее неравенство:

$$\lambda(G) \leq 2\sqrt{d-1}. \quad (1)$$

Для диаметра $D(G)$ графов Рамануджана справедливо соотношение [13]

$$D(G) \leq 2 \log_{d-1} N + O(1).$$

Этот диаметр достаточно близок к границе Мура.

Известны следующие подходы к построению небольших графов Рамануджана:

- 1) построение при помощи известного детерминированного метода;
- 2) случайная генерация с последующей проверкой значения λ .

В данной работе мы рассмотрим известные детерминированные методы построения графов Рамануджана, чтобы выбрать подходящие для использования сгенерированных с их помощью графов в качестве графов обобщённых клеточных автоматов с расчётом на их применение в криптографии. Такой метод должен давать достаточную свободу в выборе числа вершин графа. Важным является вопрос о диаметре таких графов. Для некоторых приложений важен также обхват графа (обхватом графа называется длина наименьшего содержащегося в нём цикла).

Одним из понятий, используемых в явных конструкциях графов Рамануджана, являются графы Кэли. Пусть H — конечная группа, S — её подмножество, такое, что $1 \notin S$ и $x^{-1} \in S$ для любого $x \in S$. Графом Кэли $\mathfrak{G}(H, S)$ группы H по множеству S называется неориентированный граф, вершинами которого являются элементы группы H . Вершины $u \in H$ и $v \in H$ соединены ребром тогда и только тогда, когда существует такое $s \in S$, что $v = us$. Такой граф, очевидно, является $|S|$ -регулярным, в нём отсутствуют петли и кратные ребра.

4. Семейство графов $X^{p,q}$ Любоцкого — Филиппа — Сарнака

Пожалуй, наиболее известным семейством графов Рамануджана является так называемое семейство $X^{p,q}$ Любоцкого — Филиппа — Сарнака [8, 10, 14]. Чтобы описать графы из этого семейства, напомним стандартные определения некоторых теоретико-групповых конструкций [15–19].

Пусть K — коммутативное кольцо с единицей, \mathbb{F} — поле. Полной линейной группой $GL(n, K)$ называется мультипликативная группа всех обратимых матриц размера $n \times n$ над кольцом K . Её подгруппа $SL(n, K)$ матриц с определителем, равным единице, называется специальной линейной группой.

Проективной полной линейной группой $PGL(n, \mathbb{F})$ называется фактор-группа полной линейной группы $GL(n, \mathbb{F})$ по её центру — подгруппе ненулевых скалярных матриц. Проективной специальной линейной группой $PSL(n, \mathbb{F})$ называется фактор-группа специальной линейной группы $SL(n, \mathbb{F})$ по её центру. Элементами групп $PGL(n, \mathbb{F})$ и $PSL(n, \mathbb{F})$ являются смежные классы, которые будем обозначать их представителями.

Для удобства изложения через $PSL'(2, \mathbb{F}_m)$ обозначим подгруппу группы $PGL(2, \mathbb{F}_m)$, содержащую те и только те смежные классы, определители элементов которых являются квадратами. Группа $PSL'(2, \mathbb{F}_m)$ изоморфна группе $PSL(2, \mathbb{F}_m)$.

Перейдём к описанию метода построения рассматриваемых графов [8, 10, 14].

Пусть p и q — различные простые числа, такие, что $p \equiv 1 \pmod{4}$ и $q \equiv 1 \pmod{4}$. Пусть $i \in \mathbb{F}_q$ такое, что $i^2 + 1 = 0$.

Рассмотрим все наборы из четырёх элементов $(a_0, a_1, a_2, a_3) \in \mathbb{Z}^4$, для которых выполняются следующие условия:

- 1) a_0 — нечётное положительное число;
- 2) a_1, a_2, a_3 — чётные числа;
- 3) $a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$.

Таких наборов ровно $p + 1$. Каждому такому набору $\alpha = (a_0, a_1, a_2, a_3)$ сопоставим $s_\alpha \in \text{PGL}(2, \mathbb{F}_q)$:

$$s_\alpha = \begin{bmatrix} a_0 + ia_1 & a_2 + ia_3 \\ -a_2 + ia_3 & a_0 - ia_1 \end{bmatrix}.$$

Пусть множество S состоит из всех таких s_α . Его мощность равна $p + 1$.

Отметим, что если p является квадратичным вычетов по модулю q , то все s_α лежат в подгруппе $\text{PSL}'(2, \mathbb{F}_q)$, изоморфной группе $\text{PSL}(2, \mathbb{F}_q)$.

Семейство графов Любоцкого — Филиппа — Сарнака $X^{p,q}$ определяется следующим образом:

$$X^{p,q} = \begin{cases} \mathfrak{G}(\text{PGL}(2, \mathbb{F}_q), S), & \text{если } \left(\frac{p}{q}\right) = -1, \\ \mathfrak{G}(\text{PSL}'(2, \mathbb{F}_q), S), & \text{если } \left(\frac{p}{q}\right) = 1, \end{cases}$$

где $\left(\frac{p}{q}\right)$ — символ Лежандра. Такие графы являются $(p + 1)$ -регулярными. В них отсутствуют петли и кратные рёбра. Число вершин у этих графов равно

$$N = \begin{cases} q(q^2 - 1), & \text{если } \left(\frac{p}{q}\right) = -1, \\ \frac{q(q^2 - 1)}{2}, & \text{если } \left(\frac{p}{q}\right) = 1. \end{cases}$$

Для графов $X^{p,q}$ доказано [8], что они являются графами Рамануджана, для их диаметра выполняется неравенство

$$D(X^{p,q}) \leq 2(\log_p N + \log_p 2) + 1,$$

а обхват таких графов имеет следующую оценку:

$$\text{girth}(X^{p,q}) \geq \begin{cases} 2(2 \log_p q - \log_p 2), & \text{если } \left(\frac{p}{q}\right) = -1, \\ 2 \log_p q, & \text{если } \left(\frac{p}{q}\right) = 1. \end{cases}$$

Если $\left(\frac{p}{q}\right) = -1$, то граф $X^{p,q}$ является двудольным, а если $\left(\frac{p}{q}\right) = 1$ — недвудольным. К сожалению, недвудольных графов небольшой степени с небольшим числом вершин из этого семейства очень мало. Так, если рассматривать только графы степени $d \leq 20$ с числом вершин $N \leq 20000$, то таких графов всего четыре (их параметры — числа p и q , степень d и число вершин N — приведены в табл. 1). Такая особенность делает эти графы непригодными для использования в рассматриваемых целях.

Т а б л и ц а 1

| p | q | d | N |
|-----|-----|-----|-------|
| 5 | 29 | 6 | 12180 |
| 13 | 17 | 14 | 2448 |
| 13 | 29 | 14 | 12180 |
| 17 | 13 | 18 | 1092 |

5. Семейство графов $Y^{p,q}$ Любоцкого — Филиппа — Сарнака

Рассмотрим другое семейство графов Рамануджана — семейство $Y^{p,q}$, также предложенное Любоцким, Филиппом и Сарнаком [10]. Графы, принадлежащие этому семейству, являются недвудольными. Опишем метод построения этих графов.

Выберем простые числа p и q , для которых выполняются следующие условия:

$$p \equiv 1 \pmod{4}, \quad q \equiv 1 \pmod{4}, \quad p \neq q, \quad \left(\frac{p}{q}\right) = 1.$$

Построим граф $Y^{p,q}$. Множеством V его вершин является проективная прямая над конечным полем \mathbb{F}_q , т. е. $V = \mathbb{F}_q \cup \{\infty\}$. Каждая вершина $u \in V$ соединена ребром с вершиной v , определяемой по формулам

$$v = \begin{cases} \frac{(a_0 + ia_1)u + (a_2 + ia_3)}{(-a_2 + ia_3)u + (a_0 - ia_1)}, & \text{если } (a_2 - ia_3)u \neq a_0 - ia_1 \text{ и } u \neq \infty, \\ \infty, & \text{если } (a_2 - ia_3)u = a_0 - ia_1 \text{ и } u \neq \infty, \\ \frac{ia_1 + a_0}{ia_3 - a_2}, & \text{если } ia_3 \neq a_2 \text{ и } u = \infty, \\ \infty, & \text{если } ia_3 = a_2 \text{ и } u = \infty, \end{cases}$$

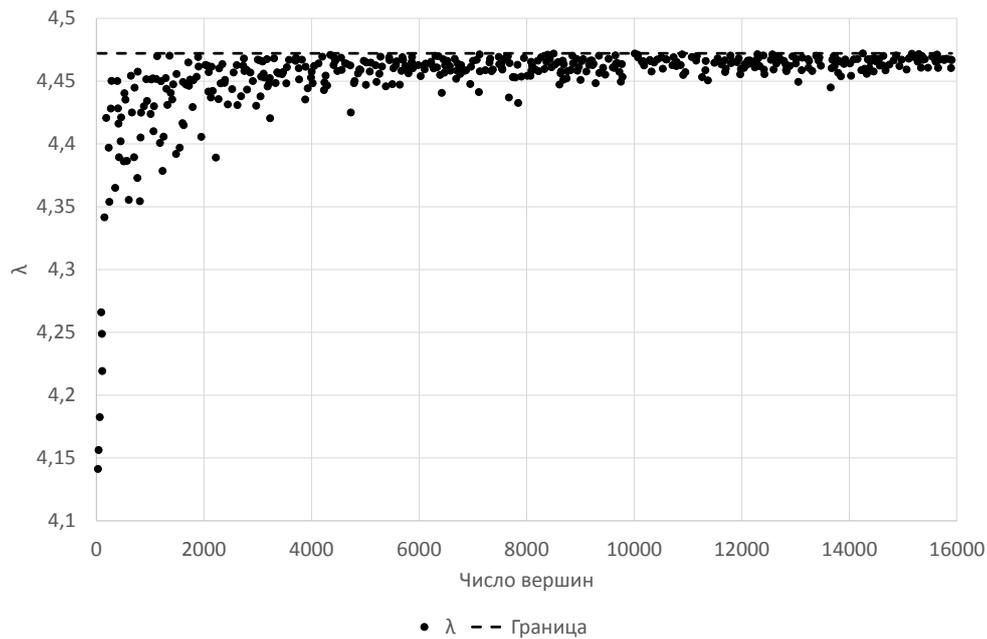
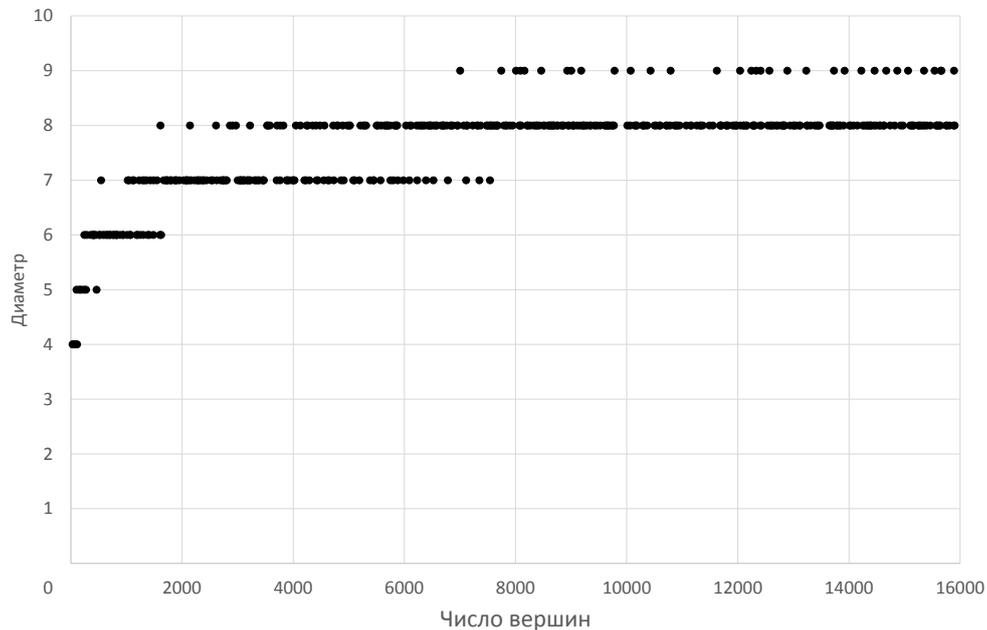
для каждой четвёрки $(a_0, a_1, a_2, a_3) \in \mathbb{Z}^4$, такой, что выполняются следующие условия:

- 1) a_0 — нечётное положительное число;
- 2) a_1, a_2, a_3 — чётные числа;
- 3) $a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$.

Здесь $i \in \mathbb{F}_q$ такое, что $i^2 + 1 = 0$. Граф имеет степень, равную количеству таких четвёрок $(p + 1)$. В построенном графе могут присутствовать кратные рёбра и петли. Для графов $Y^{p,q}$ доказано [10], что они являются графами Рамануджана.

Различных графов из семейства $Y^{p,q}$ с числом вершин от 100 до 10000 для $p = 5$ (т. е. 6-регулярных) насчитывается 298, а для $p = 13$ (т. е. 14-регулярных) — 301. При этом они достаточно равномерно распределены в указанном диапазоне.

В целях проведения вычислительных экспериментов разработано программное обеспечение на языке Python. С помощью него построено 448 таких графов степени 6 с числом вершин от 30 до 15902. Их параметры λ приведены на рис. 1, а диаметры — на рис. 2. Количество петель у каждого из построенных графов равно 6, а количество пар кратных рёбер — либо 0, либо 12. Заметим, что существенно сократить число петель и кратных рёбер позволяет описанный в п. 8 способ коррекции.

Рис. 1. Значения параметра λ графов $Y^{p,q}$ степени 6 и граница (1)Рис. 2. Диаметры графов $Y^{p,q}$ степени 6

6. Семейство графов Моргенштерна

Другое семейство графов Рамануджана предложено Моргенштерном в работе [20]. Опишем метод построения таких графов [20–22].

Графы Моргенштерна представляют собой графы Кэли для группы PSL' или PGL над полем Галуа по некоторому множеству.

Сначала рассмотрим случай поля чётной характеристики. В этом случае граф Моргенштерна недвудольный. Пусть $q = 2^\tau$ (для некоторого τ). Выберем такое $\varepsilon \in \mathbb{F}_q$, что

многочлен $f(x) = x^2 + x + \varepsilon$ неприводим над \mathbb{F}_q , и неприводимый многочлен $P_n(X) \in \mathbb{F}_q[X]$ степени n , где n — чётное. Будем использовать поле $\mathbb{F}_{q^n} = \mathbb{F}_q[X]/(P_n(X))$.

Пусть $i \in \mathbb{F}_{q^n}$ — корень многочлена $f(x)$ и $S_{\text{Morg}} = \{s_0, s_1, \dots, s_q\}$, где

$$s_j = \begin{bmatrix} 1 & a_j + b_j i \\ (a_j + b_j i + b_j)X & 1 \end{bmatrix}.$$

Здесь a_j, b_j — все лежащие в поле \mathbb{F}_q решения уравнения

$$a_j^2 + a_j b_j + b_j^2 \varepsilon = 1.$$

В случае поля чётной характеристики граф Моргенштерна представляет собой граф Кэли

$$G_{\text{Morg}}^{q,n} = \mathfrak{G}(\text{PSL}'(2, \mathbb{F}_{q^n}), S_{\text{Morg}}).$$

Рассмотрим теперь случай поля нечётной характеристики. Пусть p — нечётное простое, $q = p^\tau$ (для некоторого τ). Выберем $\varepsilon \in \mathbb{F}_q$, такое, что оно не является квадратом в поле \mathbb{F}_q ; неприводимый многочлен $P_n(X) \in \mathbb{F}_q[X]$ степени n , где n — чётное; $i \in \mathbb{F}_{q^n}$, такое, что $i^2 = \varepsilon$. Будем использовать поле $\mathbb{F}_{q^n} = \mathbb{F}_q[X]/(P_n(X))$. Определим множество $S'_{\text{Morg}} = \{s'_0, s'_1, \dots, s'_q\}$, где

$$s'_j = \begin{bmatrix} 1 & a_j - b_j i \\ (a_j + b_j i)(X - 1) & 1 \end{bmatrix}.$$

Здесь a_j, b_j — все лежащие в поле \mathbb{F}_q решения уравнения

$$b_j^2 \varepsilon - a_j^2 = 1.$$

В случае поля нечётной характеристики граф Моргенштерна представляет собой граф Кэли

$$G_{\text{Morg}}^{q,n} = \begin{cases} \mathfrak{G}(\text{PSL}'(2, \mathbb{F}_{q^n}), S'_{\text{Morg}}), & \text{если } \left(\frac{X}{P_n(X)}\right) = 1, \\ \mathfrak{G}(\text{PGL}(2, \mathbb{F}_{q^n}), S'_{\text{Morg}}), & \text{если } \left(\frac{X}{P_n(X)}\right) = -1, \end{cases}$$

где $\left(\frac{a}{P_n(X)}\right)$ — символ Лежандра в поле \mathbb{F}_{q^n} .

Графы Моргенштерна являются $(q+1)$ -регулярными графами Рамануджана без петель и кратных рёбер с числом вершин

$$N = \begin{cases} q^n(q^{2n} - 1) & \text{для группы } \text{PGL}(2, \mathbb{F}_{q^n}), \\ q^n(q^{2n} - 1)/2 & \text{для группы } \text{PSL}'(2, \mathbb{F}_{q^n}). \end{cases}$$

Такой граф является недвудольным при использовании группы $\text{PSL}'(2, \mathbb{F}_{q^n})$ и двудольным при использовании группы $\text{PGL}(2, \mathbb{F}_{q^n})$. Для таких графов доказано [20], что они имеют малый диаметр и большой обхват:

$$\begin{aligned} D(G_{\text{Morg}}^{q,n}) &\leq 2 \log_q(N) + 2, \\ \text{girth}(G_{\text{Morg}}^{q,n}) &\geq \begin{cases} \frac{4}{3} \log_q(N) & \text{для группы } \text{PGL}(2, \mathbb{F}_{q^n}), \\ \frac{2}{3} \log_q(N) + 1 & \text{для группы } \text{PSL}'(2, \mathbb{F}_{q^n}). \end{cases} \end{aligned}$$

К сожалению, недвудольных графов Моргенштерна малых степеней с небольшим числом вершин очень мало. Параметры таких графов приведены в табл. 2, из которой видно, что для обсуждаемых целей это семейство графов не подходит.

Таблица 2

**Параметры недвудольных графов Моргенштерна
малых степеней (для небольших значений параметра n)**

| Параметры | | | | Число вершин N графа Моргенштерна $G_{\text{Morg}}^{q,n}$ | | | | |
|-----------|-----|--------|-----|---|----------------------|----------------------|----------------------|----------------------|
| d | p | τ | q | $n = 2$ | $n = 4$ | $n = 6$ | $n = 8$ | $n = 10$ |
| 4 | 3 | 1 | 3 | 360 | 265680 | $1,94 \cdot 10^8$ | $1,41 \cdot 10^{11}$ | $1,03 \cdot 10^{14}$ |
| 5 | 2 | 2 | 4 | 2040 | 8388480 | $3,44 \cdot 10^{10}$ | $1,41 \cdot 10^{14}$ | $5,76 \cdot 10^{17}$ |
| 6 | 5 | 1 | 5 | 7800 | 122070000 | $1,91 \cdot 10^{12}$ | $2,98 \cdot 10^{16}$ | $4,66 \cdot 10^{20}$ |
| 8 | 7 | 1 | 7 | 58800 | 6920642400 | $8,14 \cdot 10^{14}$ | $9,58 \cdot 10^{19}$ | $1,13 \cdot 10^{25}$ |
| 9 | 2 | 3 | 8 | 131040 | 34359736320 | $9,01 \cdot 10^{15}$ | $2,36 \cdot 10^{21}$ | $6,19 \cdot 10^{26}$ |
| 10 | 3 | 2 | 9 | 265680 | $1,41 \cdot 10^{11}$ | $7,5 \cdot 10^{16}$ | $3,99 \cdot 10^{22}$ | $2,12 \cdot 10^{28}$ |

7. Семейство графов Пайзера

Опишем ещё одно семейство графов Рамануджана — графы Пайзера [21, 23, 24]. Они основаны на эллиптических кривых. Эллиптическим кривым и их применению в криптографии посвящено большое количество работ, например [25–28].

Напомним, что эллиптической кривой над полем \mathbb{F} называется гладкая (т. е. не имеющая особых точек) алгебраическая кривая над этим полем, задаваемая уравнением

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \quad (2)$$

вместе с точкой в бесконечности O . Если эллиптическая кривая определена над полем, характеристика которого отлична от 2 и 3, то линейным преобразованием координат её уравнение сводится к форме Вейерштрасса

$$y^2 = x^3 + ax + b. \quad (3)$$

Пусть p и l — простые числа, причём $p \equiv 1 \pmod{12}$, а l является квадратичным вычетом по модулю p . Множеством вершин графа Пайзера является множество классов изоморфизма суперсингулярных эллиптических кривых над полем \mathbb{F}_{p^2} . Вершины такого графа удобно задавать с помощью j -инвариантов соответствующих эллиптических кривых. Будем писать «вершина j », имея в виду вершину, соответствующую классу изоморфизма эллиптических кривых с j -инвариантом j . Представителя такого класса будем обозначать E_j . Для $j \notin \{0, 1728\}$ эллиптическую кривую E_j над полем характеристики, отличной от 2 и 3, можно задать, например, уравнением вида

$$y^2 = x^3 - \frac{3j}{j - 1728}x + \frac{2j}{j - 1728}.$$

Вершины j_1 и j_2 являются смежными, если существует l -изогения между E_{j_1} и E_{j_2} . Более подробно построение рёбер графа описано далее. Напомним, что если E' и E'' — эллиптические кривые, то изогенией из E' в E'' называется морфизм $\varphi : E' \rightarrow E''$, для которого $\varphi(O) = O$. Изогения является гомоморфизмом групп точек эллиптических кривых. Степенью изогении называется мощность её ядра. Изогению степени l часто называют l -изогенией.

Полученный граф Пайзера, который будем обозначать $\Pi_{l,p}$, имеет $N = \lfloor p/12 \rfloor$ вершин. Он является недвудольным $(l + 1)$ -регулярным неориентированным графом и графом Рамануджана. Для его диаметра справедливо неравенство

$$D(\Pi_{l,p}) \leq 2 \log_l N + 2,$$

а для обхвата — неравенство

$$\text{girth}(\Pi_{l,p}) \geq \lceil \log_l p - \log_l 4 \rceil.$$

Важным является понятие группы l -кручения эллиптической кривой. Группа l -кручения $E[l]$ эллиптической кривой E , определённой над полем \mathbb{F} , представляет собой множество точек, которое задается следующим образом:

$$E[l] = \{Q \in E(\overline{\mathbb{F}}) : lQ = O\}, \quad (4)$$

где $E(\overline{\mathbb{F}})$ — группа точек эллиптической кривой E с координатами из алгебраического замыкания поля \mathbb{F} . Известно, что если характеристика поля \mathbb{F} не делит l , то имеет место изоморфизм $E[l] \cong \mathbb{Z}_l \oplus \mathbb{Z}_l$. В этом случае $|E[l]| = l^2$.

Нас будет интересовать случай, когда число l — простое. Тогда группа l -кручения имеет $l + 1$ подгруппу порядка l , а x -координаты точек, входящих в неё, являются корнями некоторого специального полинома ψ_l . Для эллиптических кривых в форме Вейерштрасса (над полями характеристики, отличной от 2 и 3) этот полином вычисляется в соответствии с рекуррентными формулами

$$\psi_{2m+1} = \psi_{m+2}\psi_m^3 - \psi_{m-1}\psi_{m+1}^3 \quad \text{для } m \geq 2; \quad (5)$$

$$\psi_{2m} = \frac{\psi_m}{2y}(\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2) \quad \text{для } m \geq 3, \quad (6)$$

при этом $\psi_1 = 1$; $\psi_2 = 2y$; $\psi_3 = 3x^4 + 6ax^2 + 12bx - a^2$; $\psi_4 = 4y(x^6 + 5ax^4 + 20bx^3 - 5a^2x^2 - 4abx - 8b^2 - a^3)$.

Рёбра графа Пайзера соответствуют изогениям. Для построения всех рёбер, инцидентных вершине z , следует найти j -инварианты образов всех l -изогений из эллиптической кривой E_z . Для построения изогении нужно найти её ядро. Ядро l -изогении является подгруппой порядка l группы l -кручения эллиптической кривой. Таким образом, каждой подгруппе порядка l группы $E_z[l]$ соответствует инцидентное вершине z ребро.

Пусть для эллиптической кривой вида (2) изогения задана ядром C . Образ изогении может быть найден с помощью формул Велу, предложенных в [29]. Приведём их (формулы приводятся на основе работы [30]).

Для точки $P = (x_P, y_P) \in C$, не являющейся точкой в бесконечности, определим следующие параметры:

$$g_P^x = 3x_P^2 + 2a_2x_P - a_1y_P + a_4, \\ g_P^y = -2y_P - a_3 - a_1x_P,$$

$$r_P = \begin{cases} g_P^x, & \text{если } 2P = O, \\ 2g_P^x - a_1g_P^y, & \text{если } 2P \neq O, \end{cases} \\ u_P = (g_P^y)^2.$$

Пусть $C_2 \subseteq C$ — множество точек порядка 2, содержащихся в множестве C . Выберем такое множество $R \subset C$, что C является объединением четырёх попарно непересекающихся множеств:

$$C = \{O\} \cup C_2 \cup R \cup (-R),$$

где $(-R) = \{-Q : Q \in R\}$. Вычислим параметры r и w следующим образом:

$$r = \sum_{Q \in RUC_2} r_Q; \quad w = \sum_{Q \in RUC_2} (x_Q r_Q + u_Q).$$

Тогда образом изогении будет являться эллиптическая кривая

$$y^2 + a'_1 xy + a'_3 y = x^3 + a'_2 x^2 + a'_4 x + a'_6,$$

где $a'_1 = a_1$; $a'_2 = a_2$; $a'_3 = a_3$; $a'_4 = a_4 - 5r$; $a'_6 = a_6 - (a_1^2 + 4a_2)r - 7w$. Сама изогения задаётся формулами

$$x' = x + \sum_{Q \in RUC_2} \left(\frac{r_Q}{x - x_Q} + \frac{u_Q}{(x - x_Q)^2} \right),$$

$$y' = y - \sum_{Q \in RUC_2} \left(u_Q \frac{2y + a_1 x + a_3}{(x - x_Q)^3} + r_Q \frac{y - y_Q + a_1(x - x_Q)}{(x - x_Q)^2} + \frac{a_1 u_Q - g_Q^x g_Q^y}{(x - x_Q)^2} \right).$$

Чтобы узнать, какой вершине графа Пайзера соответствует образ изогении (ребро, соответствующее изогении, инцидентно этой вершине), остаётся вычислить его j -инвариант.

Для эллиптических кривых в форме Вейерштрасса (3) вышеприведенные формулы Велу упрощаются:

$$g_P^x = 3x_P^2 + a,$$

$$g_P^y = -2y_P,$$

$$r_P = \begin{cases} g_P^x, & \text{если } 2P = O, \\ 2g_P^x, & \text{если } 2P \neq O, \end{cases}$$

$$u_P = (g_P^y)^2,$$

$$r = \sum_{Q \in RUC_2} r_Q, \quad w = \sum_{Q \in RUC_2} (x_Q r_Q + u_Q).$$

Образом изогении в этом случае является эллиптическая кривая, заданная уравнением

$$y^2 = x^3 + (a - 5r)x + (b - 7w).$$

Для построения всех образов изогений для данной эллиптической кривой можно предложить следующую методику:

- 1) Найти полином ψ_l по формулам (5), (6).
- 2) Найти корни полинома ψ_l . Множество точек, соответствующих этим корням, вместе с точкой в бесконечности составляет группу l -кручения эллиптической кривой.
- 3) Найти все подгруппы порядка l группы l -кручения. Их число равно $l + 1$.
- 4) Каждая подгруппа порядка l группы l -кручения образует ядро одной из l -изогений. Найти образы всех l -изогений с помощью формул Велу.

Итак, построение графа Пайзера $\Pi_{l,p}$ происходит следующим образом:

- 1) выбирается j -инвариант, соответствующий суперсингулярной эллиптической кривой над полем \mathbb{F}_{p^2} ;
- 2) производится обход графа в ширину с построением образов l -изогений и вычислением их j -инвариантов. Получившийся граф является графом Пайзера $\Pi_{l,p}$.

Графов Пайзера существует достаточно много. Так, число графов Пайзера степени 6 с числом вершин от 100 до 10000 составляет 1360. Данные по количеству графов Пайзера различных степеней с числом вершин в этом диапазоне приведены в табл. 3.

Т а б л и ц а 3

| Степень | Параметр l | Число графов Пайзера |
|---------|--------------|----------------------|
| 4 | 3 | 2749 |
| 6 | 5 | 1360 |
| 8 | 7 | 1361 |
| 12 | 11 | 1377 |
| 14 | 13 | 1352 |
| 18 | 17 | 1360 |
| 20 | 19 | 1351 |

Некоторые графы Пайзера (степеней 4, 6, 8) были явно построены с целью проведения вычислительных экспериментов. Всего построено 56 графов степени 6, 57 графов степени 4 и 21 граф степени 8. Программа построения графов разработана автором для системы компьютерной алгебры Магма [31, 32]. Параметры λ построенных графов приведены на рис. 3–5, а значения их диаметров — на рис. 6.

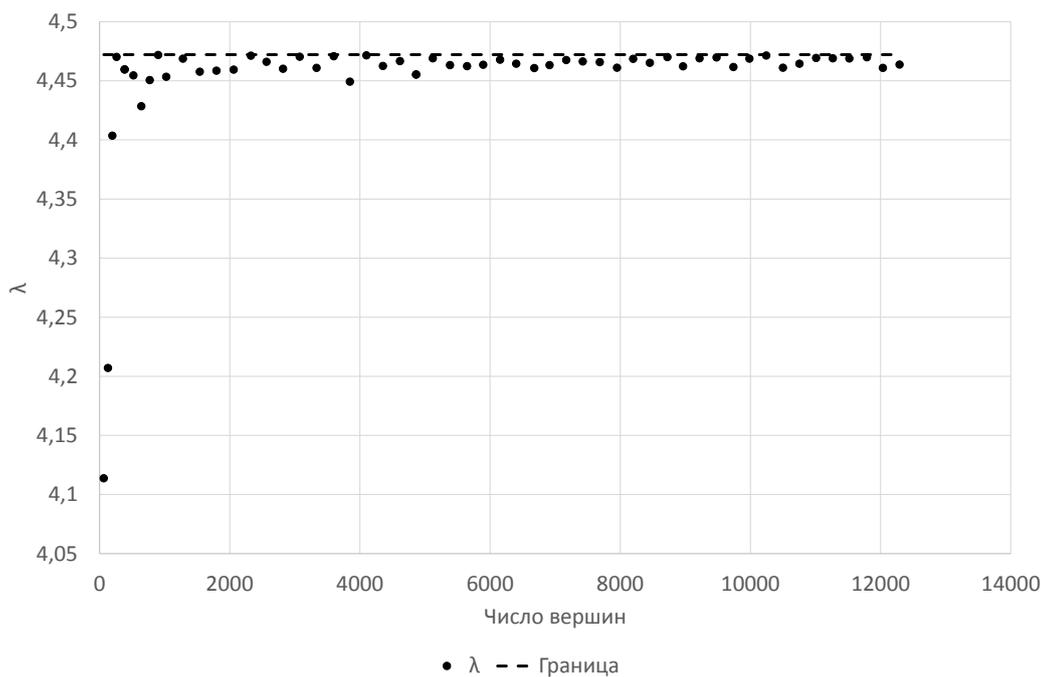
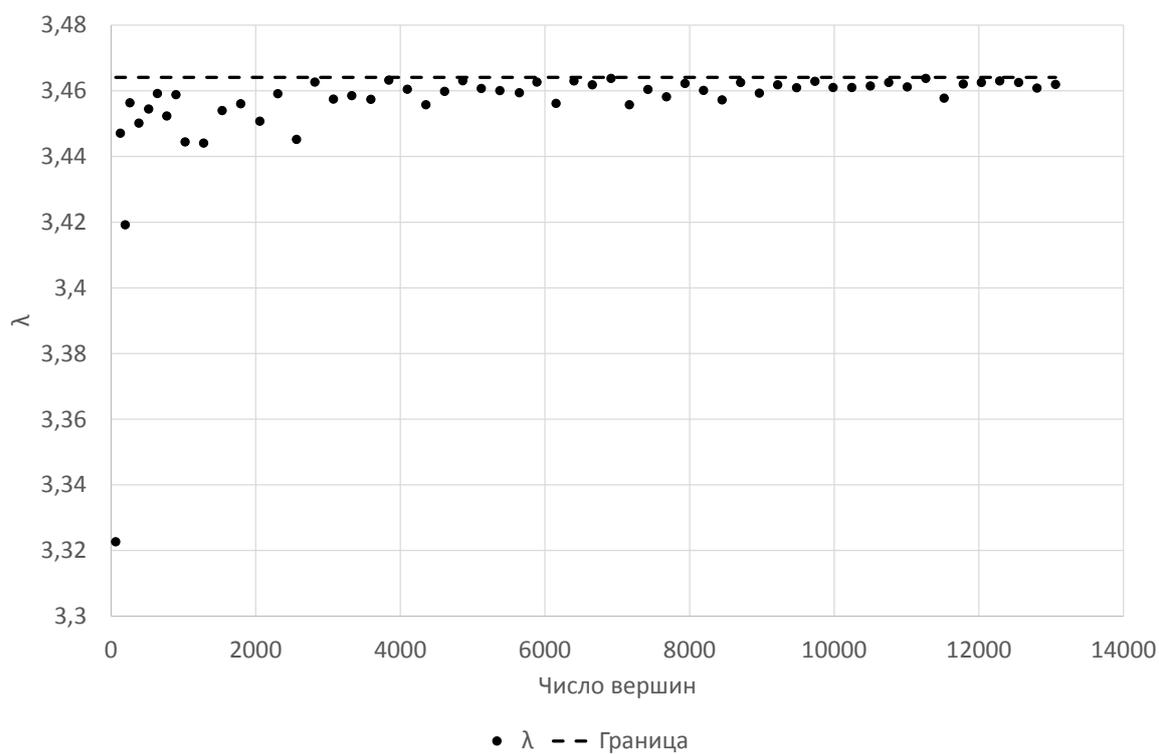
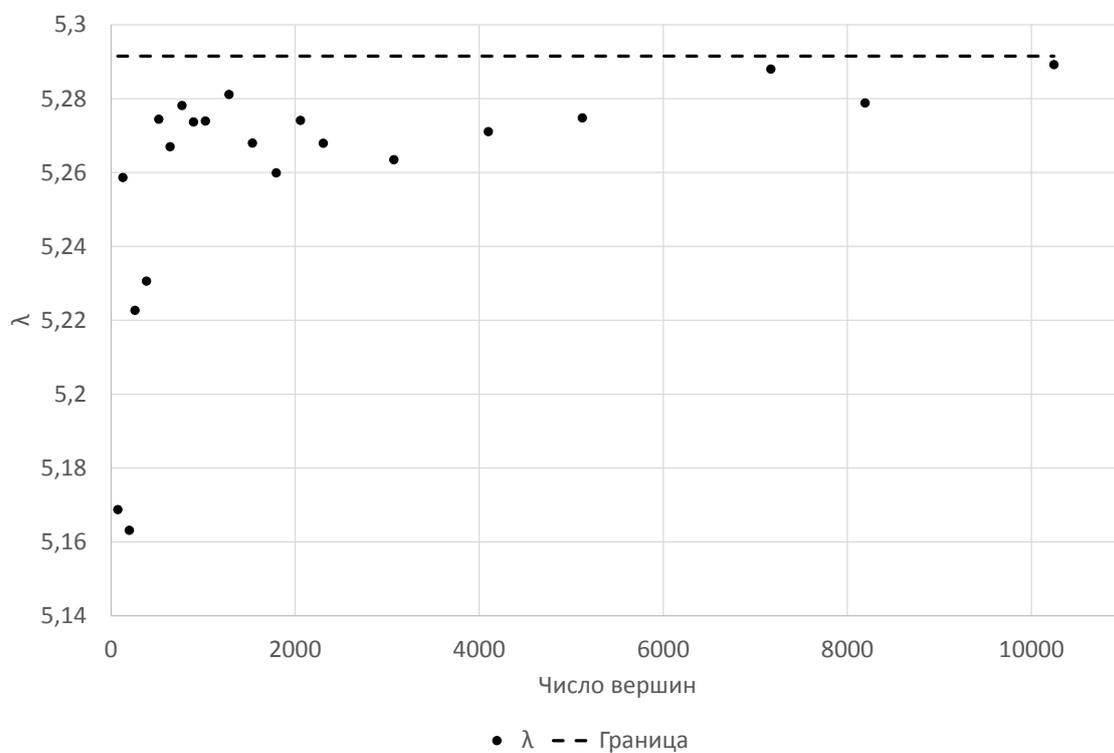


Рис. 3. Значения параметра λ графов Пайзера степени 6 и граница (1)

Рис. 4. Значения параметра λ графов Пайзера степени 4 и граница (1)Рис. 5. Значения параметра λ графов Пайзера степени 8 и граница (1)

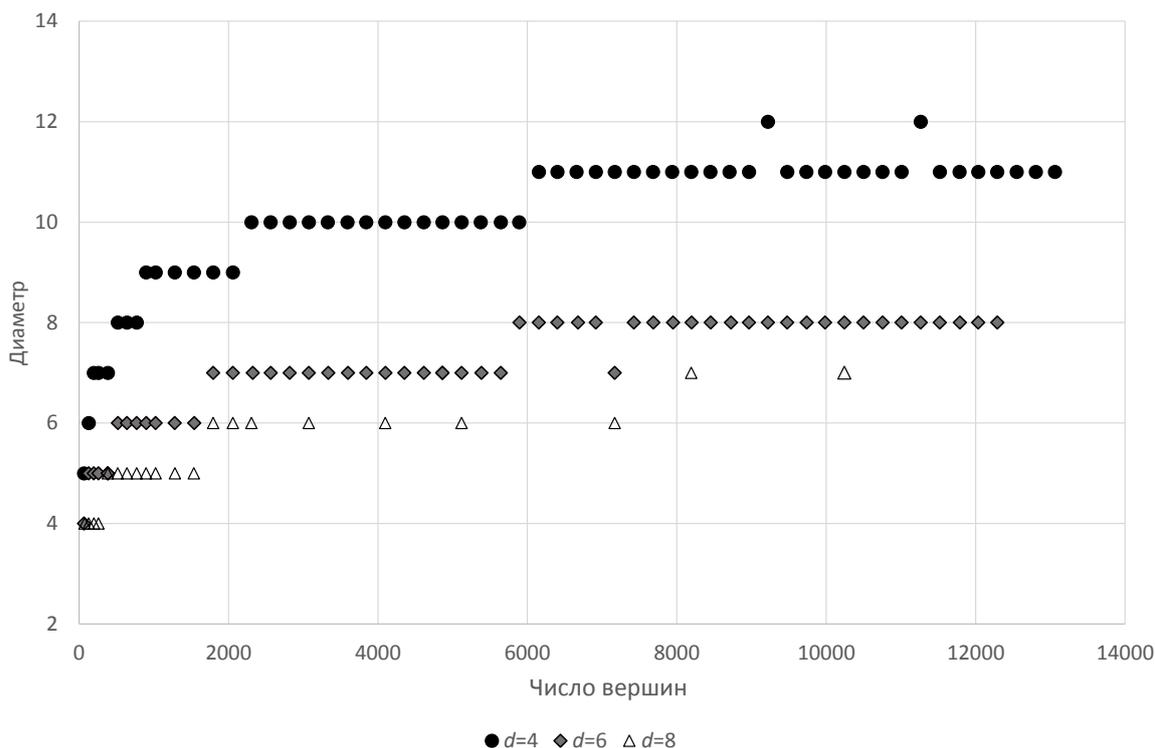


Рис. 6. Диаметры графов Пайзера степени d

8. Удаление петель и кратных рёбер

Наличие в графе петель и кратных рёбер может ухудшить криптографические свойства обобщённого клеточного автомата, поэтому необходимо уменьшить их количество. Сделать это нужно так, чтобы граф оставался регулярным. Приведём простой алгоритм, вариант которого использовался автором ранее [33].

Обработаем кратные рёбра следующим образом. Рассмотрим две пары кратных рёбер: пару рёбер, соединяющих вершины u_1, v_1 , и пару рёбер, соединяющих вершины u_2, v_2 (здесь $u_1, v_1, u_2, v_2 \in V$ — попарно различные вершины графа, V — множество вершин графа). Удалим одно ребро $\{u_1, v_1\}$ и одно ребро $\{u_2, v_2\}$, после чего добавим рёбра $\{u_1, u_2\}$ и $\{v_1, v_2\}$. Обработаем так все пары кратных рёбер в графе (если их количество нечётное, то все, кроме одной).

Петли обработаем следующим образом. Пусть в графе имеются петли $\{u_1, u_1\}$, $\{u_2, u_2\}$, ..., $\{u_t, u_t\}$, где $u_1, \dots, u_t \in V$. Удалим эти петли из графа, после чего добавим рёбра $\{u_1, u_2\}$, $\{u_2, u_3\}$, ..., $\{u_{t-1}, u_t\}$, $\{u_t, u_1\}$ (заметим, что если при этом появятся кратные рёбра, то порядок вершин последовательности u_1, u_2, \dots, u_t следует, если возможно, изменить так, чтобы кратных рёбер не появлялось).

Если после выполнения этих процедур петли и кратные рёбра остаются, можно произвести следующие действия. Пусть для некоторых попарно различных вершин $u, v, w \in V$ имеется пара кратных рёбер, соединяющих вершины u, v , и петля $\{w, w\}$. Тогда можно удалить одно ребро из этой пары и петлю, после чего добавить рёбра $\{u, w\}$ и $\{v, w\}$.

Приведённый алгоритм позволяет существенно уменьшить количество кратных рёбер и петель, не влияя на регулярность графа и не приводя к увеличению его диаметра.

Для большинства сгенерированных в рамках настоящей работы графов его применение не привело к существенному изменению параметра λ .

Заключение

Для применения в качестве графов обобщённых клеточных автоматов графы Моргенштерна и графы $X^{p,q}$, очевидно, не подходят, поскольку таких графов небольшой степени с небольшим числом вершин существует очень мало. Вместе с тем для рассматриваемых целей подходят графы $Y^{p,q}$ и графы Пайзера, так как они удовлетворяют всем необходимым требованиям. Следует заметить, что число петель и кратных рёбер в графе может быть при необходимости сведено к минимуму способом, изложенным в п. 8.

ЛИТЕРАТУРА

1. Ключарёв П. Г. Блочные шифры, основанные на обобщённых клеточных автоматах // Наука и образование. МГТУ им. Н. Э. Баумана. Электрон. журн. 2012. № 12. С. 361–374.
2. Ключарёв П. Г. Криптографические хэш-функции, основанные на обобщённых клеточных автоматах // Наука и образование. МГТУ им. Н. Э. Баумана. Электрон. журн. 2013. № 1. С. 161–172.
3. Тоффоли Т., Марголюс Н. Машины клеточных автоматов: пер. с англ. М.: Мир, 1991. 280 с.
4. Kauffman S. A. Metabolic stability and epigenesis in randomly constructed genetic net // J. Theor. Biol. 1969. No. 22. P. 437–467.
5. Сухинин Б. М. Разработка генераторов псевдослучайных двоичных последовательностей на основе клеточных автоматов // Наука и образование. МГТУ им. Н. Э. Баумана. Электрон. журн. 2010. № 9. <http://engineering-science.ru/doc/159714.html>
6. Davidoff G., Sarnak P., and Valette A. Elementary Number Theory, Group Theory and Ramanujan Graphs. Cambridge: Cambridge University Press, 2003. V. 55. 144 p.
7. Hoory S., Linial N., and Wigderson A. Expander graphs and their applications // Bull. Amer. Math. Soc. 2006. V. 43. No. 4. P. 439–562.
8. Lubotzky A., Phillips R., and Sarnak P. Ramanujan graphs // Combinatorica. 1988. V. 8. No. 3. P. 261–277.
9. Krebs M. and Shaheen A. Expander Families and Cayley Graphs: A Beginner's Guide. Oxford: Oxford University Press, 2011. 258 p.
10. Sarnak P. Some Applications of Modular Forms. Cambridge: Cambridge University Press, 1990. V. 99. 111 p.
11. Chung F. Spectral Graph Theory. Amer. Math. Soc., 1997. 207 p.
12. Sarnak P. What is ... an expander? // Notices Amer. Math. Soc. 2004. V. 51. P. 762–770.
13. Lubotzky A. Discrete Groups, Expanding Graphs and Invariant Measures. Springer Science & Business Media, 2010. 196 p.
14. Lubotzky A., Phillips R., and Sarnak P. Explicit expanders and the Ramanujan conjectures // Proc. 18th Ann. ACM Symp. on Theory of Computing. ACM, 1986. P. 240–246.
15. Grove L. Classical Groups and Geometric Algebra. Fields Institute Communications. Amer. Math. Soc., 2002. 169 p.
16. Humphreys J. A Course in Group Theory. Oxford Graduate Texts in Mathematics. Oxford: Oxford University Press, 1996. 279 p.
17. James G. and Liebeck M. Representations and Characters of Groups. Cambridge Mathematical Textbooks. Cambridge: Cambridge University Press, 2001. 458 p.
18. Lanski C. Concepts in Abstract Algebra. Amer. Math. Soc., 2005. 545 p.

19. *Каргаполов М. И., Мерзляков Ю. И.* Основы теории групп. М.: Наука, Физматлит, 1996. 287 с.
20. *Morgenstern M.* Existence and explicit constructions of $q + 1$ regular Ramanujan graphs for every prime power q // *J. Combinatorial Theory. Ser. B.* 1994. V. 62. No. 1. P. 44–62.
21. *Petit C.* On Graph-Based Cryptographic Hash Functions. PhD Thesis. Catholic University of Louvain, 2009. 286 p. www0.cs.ucl.ac.uk/staff/c.petit/files/thesis.pdf
22. *Nikkel T.* Ramanujan Graphs. Master's Thesis. University of Manitoba, 2007. 112 p. <http://mspace.lib.umanitoba.ca/bitstream/handle/1993/9146/thesis.pdf>
23. *Pizer A. K.* Ramanujan graphs and Hecke operators // *Bull. Amer. Math. Soc.* 1990. V. 23. No. 1. P. 127–137.
24. *Charles D. X., Lauter K. E., and Goren E. Z.* Cryptographic hash functions from expander graphs // *J. Cryptology.* 2009. V. 22. No. 1. P. 93–113.
25. *Silverman J.* Advanced Topics in the Arithmetic of Elliptic Curves. Graduate Texts in Mathematics. N.Y.: Springer, 2013. 528 p.
26. *Blake I., Seroussi G., and Smart N.* Elliptic Curves in Cryptography. Lecture Note Series. Cambridge: Cambridge University Press, 1999. 204 p.
27. *Silverman J.* The Arithmetic of Elliptic Curves. Graduate Texts in Mathematics. N.Y.: Springer, 2009. 402 p.
28. *Washington L.* Elliptic Curves: Number Theory and Cryptography. 2nd Ed. Discrete Mathematics and its Applications. Boca Raton: CRC Press, 2008. 536 p.
29. *Vélu J.* Isogénies entre courbes elliptiques // *CR Acad. Sci. Paris Sér. AB.* 1971. V. 273. P. A238–A241.
30. *Shumow D.* Isogenies of Elliptic Curves: A Computational Approach. Master's Thesis. University of Washington, 2009. 78 p. <https://arxiv.org/abs/0910.5370>
31. *Bosma W., Cannon J., and Playoust C.* The Magma algebra system. I. The user language // *J. Symbolic Comput.* 1997. V. 24. No. 3–4. P. 235–265.
32. Handbook of Magma Functions. Edition 2.20 / eds. W. Bosma, J. Cannon, C. Fieker, and A. Steel. 2014. 5583 p. <https://www.math.uzh.ch/sepp/magma-2.19.8-cr/Handbook.pdf>
33. *Ключарёв П. Г.* Построение псевдослучайных функций на основе обобщенных клеточных автоматов // *Наука и образование. МГТУ им. Н.Э. Баумана. Электрон. журн.* 2012. № 10. С. 263–274.

REFERENCES

1. *Klyucharev P. G.* Blochnye shifry, osnovannye na obobshchyonnykh kletochnykh avtomatah [Block ciphers based on generalized cellular automata]. *Science and Education of the Bauman MSTU*, 2012, no. 2, pp. 361–374. (in Russian)
2. *Klyucharev P. G.* Kriptograficheskie hesh-funkcii, osnovannye na obobshchyonnykh kletochnykh avtomatah [Cryptographic hash functions based on generalized cellular automata]. *Science and Education of the Bauman MSTU*, 2013, no. 1, pp. 161–172. (in Russian)
3. *Toffoli T. and Margolus N.* Cellular Automata Machines. MIT Press, 1987. 276 p.
4. *Kauffman S. A.* Metabolic stability and epigenesis in randomly constructed genetic net. *J. Theor. Biol.*, 1969, no. 22, pp. 437–467.
5. *Suhinin B. M.* Razrabotka generatorov psevdosluchajnykh dvoichnykh posledovatelnoyey na osnove kletochnykh avtomatov [Construction of pseudorandom binary sequence generators based on cellular automata]. *Science and Education of the Bauman MSTU*, 2010, no. 9. <http://engineering-science.ru/doc/159714.html>. (in Russian)
6. *Davidoff G., Sarnak P., and Valette A.* Elementary Number Theory, Group Theory and Ramanujan Graphs. Cambridge, Cambridge University Press, 2003, vol. 55. 144 p.

7. Hoory S., Linial N., and Wigderson A. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 2006, vol. 43, no. 4. pp. 439–562.
8. Lubotzky A., Phillips R., and Sarnak P. Ramanujan graphs. *Combinatorica*, 1988, vol. 8, no. 3, pp. 261–277.
9. Krebs M. and Shaheen A. Expander Families and Cayley Graphs: A Beginner’s Guide. Oxford, Oxford University Press, 2011. 258 p.
10. Sarnak P. Some Applications of Modular Forms. Cambridge, Cambridge University Press, 1990, vol. 99. 111 p.
11. Chung F. Spectral Graph Theory. *Amer. Math. Soc.*, 1997, 207 p.
12. Sarnak P. What is ... an expander? *Notices Amer. Math. Soc.*, 2004, vol. 51, pp. 762–770.
13. Lubotzky A. Discrete Groups, Expanding Graphs and Invariant Measures. Springer Science & Business Media, 2010. 196 p.
14. Lubotzky A., Phillips R., and Sarnak P. Explicit expanders and the Ramanujan conjectures. *Proc. 18th Ann. ACM Symp. on Theory of Computing, ACM*, 1986, pp. 240–246.
15. Grove L. Classical Groups and Geometric Algebra. Fields Institute Communications. *Amer. Math. Soc.*, 2002. 169 p.
16. Humphreys J. A Course in Group Theory. Oxford Graduate Texts in Mathematics. Oxford, Oxford University Press, 1996. 279 p.
17. James G. and Liebeck M. Representations and Characters of Groups. Cambridge Mathematical Textbooks. Cambridge, Cambridge University Press, 2001. 458 p.
18. Lanski C. Concepts in Abstract Algebra. *Amer. Math. Soc.*, 2005. 545 p.
19. Kargapolov M. I. and Merzlyakov Yu. I. Osnovy Teorii Grupp [Foundations of Group Theory]. Moscow, Nauka, Fizmatlit Publ., 1996. 287 p. (in Russian)
20. Morgenstern M. Existence and explicit constructions of $q + 1$ regular Ramanujan graphs for every prime power q . *J. Combinatorial Theory, Ser. B*, 1994, vol. 62, no. 1, pp. 44–62.
21. Petit C. On Graph-Based Cryptographic Hash Functions. PhD Thesis, Catholic University of Louvain, 2009. 286 p. www0.cs.ucl.ac.uk/staff/c.petit/files/thesis.pdf
22. Nikkel T. Ramanujan Graphs. Master’s Thesis, University of Manitoba, 2007. 112 p. <http://mpace.lib.umanitoba.ca/bitstream/handle/1993/9146/thesis.pdf>
23. Pizer A. K. Ramanujan graphs and Hecke operators. *Bull. Amer. Math. Soc.*, 1990, vol. 23, no. 1, pp. 127–137.
24. Charles D. X., Lauter K. E., and Goren E. Z. Cryptographic hash functions from expander graphs. *J. Cryptology*, 2009, vol. 22, no. 1, pp. 93–113.
25. Silverman J. Advanced Topics in the Arithmetic of Elliptic Curves. Graduate Texts in Mathematics. N.Y., Springer, 2013. 528 p.
26. Blake I., Seroussi G., and Smart N. Elliptic Curves in Cryptography. Lecture Note Series. Cambridge, Cambridge University Press, 1999. 204 p.
27. Silverman J. The Arithmetic of Elliptic Curves. Graduate Texts in Mathematics. N.Y., Springer, 2009. 402 p.
28. Washington L. Elliptic Curves: Number Theory and Cryptography, 2nd Edition. Discrete Mathematics and Its Applications. Boca Raton, CRC Press, 2008. 536 p.
29. Vélú J. Isogénies entre courbes elliptiques. *CR Acad. Sci. Paris Sér. AB*, 1971, vol. 273, pp. A238–A241.
30. Shumow D. Isogenies of Elliptic Curves: A Computational Approach. Master’s Thesis, University of Washington, 2009. 78 p. <https://arxiv.org/abs/0910.5370>
31. Bosma W., Cannon J., and Playoust C. The Magma algebra system. I. The user language. *J. Symbolic Comput.*, 1997, vol. 24, no. 3–4, pp. 235–265.

32. Handbook of Magma Functions. Edition 2.20 / eds. W. Bosma, J. Cannon, C. Fieker, and A. Steel. 2014. 5583 p. <https://www.math.uzh.ch/sepp/magma-2.19.8-cr/Handbook.pdf>
33. *Klyucharev P. G.* Postroenie psevdosluchajnyh funkcij na osnove obobshchennyh kletochnyh avtomatov [Construction of pseudorandom functions based on generalized cellular automata]. Science and Education of the Bauman MSTU, 2012, no. 10, pp. 263–274. (in Russian)

ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ В ДИСКРЕТНОЙ МАТЕМАТИКЕ

УДК 512.54

РЕСУРСНО-ЭФФЕКТИВНЫЙ АЛГОРИТМ ДЛЯ ИССЛЕДОВАНИЯ РОСТА В КОНЕЧНЫХ ДВУПОРОЖДЁННЫХ ГРУППАХ ПЕРИОДА 5¹

А. А. Кузнецов*, А. С. Кузнецова**

* *Сибирский государственный университет науки и технологий имени академика
М.Ф. Решетнева, г. Красноярск, Россия*

** *Красноярский государственный аграрный университет, г. Красноярск, Россия*

Пусть $B_0(2, 5) = \langle a_1, a_2 \rangle$ — наибольшая конечная двупорождённая бернсайдова группа периода 5, порядок которой равен 5^{34} . Для каждого элемента данной группы существует уникальное коммутаторное представление вида $a_1^{\alpha_1} \cdot a_2^{\alpha_2} \cdot \dots \cdot a_{34}^{\alpha_{34}}$, где $\alpha_i \in \mathbb{Z}_5$, $i = 1, 2, \dots, 34$. Здесь a_1 и a_2 — порождающие элементы $B_0(2, 5)$; a_3, \dots, a_{34} — коммутаторы, которые вычисляются рекурсивно через a_1 и a_2 . Определим фактор-группу группы $B_0(2, 5)$ следующего вида: $B_k = B_0(2, 5) / \langle a_{k+1}, \dots, a_{34} \rangle$. Очевидно, что $|B_k| = 5^k$. В работе представлен ресурсно-эффективный алгоритм для исследования роста в конечных группах. Цель — минимизировать пространственную сложность алгоритма, сохранив при этом вычислительную сложность на приемлемом уровне. При помощи нового алгоритма вычислены функции роста группы B_{18} для минимального $A_2 = \{a_1, a_2\}$ и симметричного $A_4 = \{a_1, a_1^{-1}, a_2, a_2^{-1}\}$ порождающих множеств, а для группы B_{19} — только для A_4 . На основе полученных данных сформулирована гипотеза о значениях диаметров графов Кэли группы $B_0(2, 5)$ для указанных порождающих множеств.

Ключевые слова: *функция роста, группа Бернсайда, граф Кэли.*

DOI 10.17223/20710410/42/7

A RESOURCE-EFFICIENT ALGORITHM FOR STUDY THE GROWTH IN FINITE TWO-GENERATOR GROUPS OF EXPONENT 5

A. A. Kuznetsov*, A. S. Kuznetsova**

* *Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia*

** *Krasnoyarsk State Agrarian University, Krasnoyarsk, Russia*

E-mail: alex_kuznetsov80@mail.ru

For studying the growth in finite groups, we present a resource-efficient algorithm which is a modification of our early algorithm. The purpose of the modification is to

¹Исследование выполнено при финансовой поддержке РФФИ, Правительства Красноярского края, Красноярского краевого фонда поддержки научной и научно-технической деятельности в рамках научного проекта № 17-47-240318.

minimize the space complexity of the algorithm and to save its time complexity at an acceptable level. The main idea of the modified algorithm is to take in the given group G a suitable subgroup N such that $|N| \ll |G|$, to calculate growth functions for all cosets gN independently of each other, to summarize these functions and to obtain the growth function for the group G . By using this algorithm, we calculate the growth functions for the group B_{18} with two generators a_1 and a_2 and for the groups B_{18} , B_{19} with four generators a_1, a_1^{-1}, a_2 and a_2^{-1} , where $B_k = B_0(2, 5)/\langle a_{k+1}, \dots, a_{34} \rangle$ is a quotient of the group $B_0(2, 5) = \langle a_1, a_2 \rangle$ which is the largest two-generator Burnside group of exponent 5 (its order is 5^{34}), a_1 and a_2 are generators of $B_0(2, 5)$ and a_3, \dots, a_{34} are the commutators of $B_0(2, 5)$, so each element in $B_0(2, 5)$ can be represented as $a_1^{\alpha_1} \cdot a_2^{\alpha_2} \cdot \dots \cdot a_{34}^{\alpha_{34}}$, $\alpha_i \in \mathbb{Z}_5, i = 1, 2, \dots, 34$. Based on these data, we formulate a hypothesis about the diameters of Cayley graphs of the group $B_0(2, 5)$ with generating sets $A_2 = \{a_1, a_2\}$ and $A_4 = \{a_1, a_1^{-1}, a_2, a_2^{-1}\}$, namely, $D_{A_2}(B_0(2, 5)) \approx 105$ and $D_{A_4}(B_0(2, 5)) \approx 69$.

Keywords: *Burnside group, the Cayley graph, the growth function.*

Введение

Настоящая работа продолжает исследования, начатые в [1], и посвящена разработке ресурсно-эффективного алгоритма для исследования роста в конечных группах, в частности в двупорождённых группах периода 5. В [1] упор сделан на создании алгоритмов минимальной вычислительной сложности, что позволило получить ряд новых результатов о росте в указанных группах. Однако при работе с группой, состоящей из $5^{18} \approx 4 \cdot 10^{12}$ элементов, возникла принципиальная проблема — нехватка объёма памяти, несмотря на то, что для вычислений было задействовано значительное количество ресурсов: 1,2 Тбайт оперативной и 10 Тбайт дисковой памяти. По этой причине была поставлена цель — существенно снизить пространственную сложность алгоритма, сохранив при этом вычислительную сложность на приемлемом уровне.

Напомним основные определения, используемые в [1]. Пусть $G = \langle X \rangle$. Шаром K_s радиуса s группы G будем называть множество всех её элементов, которые могут быть представлены в алфавите X в виде несократимых групповых слов длины не больше s . Соответственно все элементы одинаковой длины i образуют сферу P_i радиуса i . Единица группы e является пустым словом, длина которого равна нулю. Согласно данным определениям,

$$K_s(G, X) = \bigcup_{i=0}^s P_i(G, X).$$

Для каждого целого неотрицательного i можно определить (сферическую) функцию роста группы F_i , которая равна числу элементов в сфере P_i :

$$F_i(G, X) = |P_i(G, X)|.$$

Если из контекста ясно, о какой группе $G = \langle X \rangle$ идёт речь, то для краткости вместо $K_i(G, X)$, $P_i(G, X)$ и $F_i(G, X)$ будем писать K_i , P_i и F_i , соответственно.

Обратим внимание, что при вычислении функции роста группы мы одновременно выясняем характеристики ассоциированного с группой графа Кэли, например, такие, как диаметр и средний диаметр [2]. Пусть $F_{s_0} > 0$, но $F_{s_0+1} = 0$, тогда s_0 является диаметром графа Кэли группы G в алфавите порождающих X , который будем обозначать $D_X(G)$. Соответственно средний диаметр $\bar{D}_X(G)$ равен $\frac{1}{|G|} \sum_{s=0}^{s_0} s \cdot F_s$.

К сожалению, вычисление функции роста большой конечной группы является хотя и разрешимой, но весьма сложной проблемой. Это связано с тем, что в общем случае задача по определению минимального слова элемента группы, как показали С. Ивен и О. Голдрейх [3], является NP-трудной. Поэтому для эффективного решения указанной задачи необходимо создание параллельных алгоритмов, адаптированных к использованию на многопроцессорных вычислительных системах.

Далее представлены алгоритмы, на основе которых при помощи компьютерных вычислений получены новые результаты о росте в конечных бернсайдовых двупорождённых группах периода 5.

1. Алгоритм 1

Пусть X — конечное порождающее множество произвольной конечной группы G . Базовый алгоритм 1, который вычисляет шар $K_s(G, X)$ фиксированного радиуса s , представляет собой ограниченную версию алгоритма A-I из [1].

Алгоритм 1. $K_s = \text{Ball}(G, X, s)$

Вход: X — порождающее множество группы G , радиус s

Выход: шар K_s группы G радиуса s

- 1: $K_s := \bigcup_{i=0}^s P_i$, где $P_i := \emptyset$ — сферы радиуса i
 - 2: $P_0 := \{e\}$
 - 3: **Для всех** $i = 1, 2, \dots, s$
 - 4: **Для всех** $x \in X$ и $p \in P_{i-1}$
 - 5: $g := x \cdot p$
 - 6: **Если** $g \notin K_s$, **то**
 - 7: добавить g в $P_i \subset K_s$
 - 8: **Если** $|P_i| = 0$, **то**
 - 9: переход в п. 10
 - 10: **Вернуть** K_s
-

Лемма 1. Алгоритм 1 корректен, т.е. он за конечное число шагов вычисляет шар K_s фиксированного радиуса s произвольной конечной группы G , заданной порождающим множеством X .

Доказательство. По построению алгоритм 1 выражает каждый элемент группы G в виде группового слова наименьшей длины в алфавите X . После каждого перехода от п. 3 до п. 7 множество K_s представляет собой шар радиуса i группы G относительно X . Конечность G гарантирует остановку при некотором $i \leq s$. ■

Для оценки пространственной и вычислительной сложности алгоритмов воспользуемся асимптотическим анализом [4]. Введём следующие обозначения:

- $T_i(G, X, s)$ — вычислительная сложность (i — номер алгоритма);
- $M_i(G, X, s)$ — пространственная сложность;
- $O(f(G, X, s))$ — верхняя асимптотическая оценка сложности;
- $\Theta(f(G, X, s))$ — одновременно верхняя и нижняя оценка сложности.

Здесь и далее нас будет интересовать случай $|X| \ll |G|$.

Лемма 2. $T_1 \in \Theta(|K_s|^2)$ и $M_1 \in \Theta(|K_s|)$.

Доказательство. Алгоритм 1 является ограниченной по радиусу s версией алгоритма А-I из [1], поэтому, согласно [1], $T_1 \in \Theta(|K_s|^2)$.

Для того чтобы получить асимптотическую оценку пространственной сложности, следует брать во внимание только множество K_s . Поскольку $|K_s| < M_1 < 2|K_s|$, то $M_1 \in \Theta(|K_s|)$. ■

2. Алгоритм 2

В случае достаточно большой конечной группы G в базовом алгоритме 1 для хранения одного элемента группы необходимо по крайней мере несколько байт. Это означает, что при наличии одного Тбайта памяти мы сможем построить шар группы, состоящий не более чем из 10^{12} элементов. Однако если удастся найти подходящую фактор-группу Q группы G , то можно значительно увеличить первоначально допустимый предел вычислений.

Пусть φ — гомоморфизм G на группу Q и N — ядро φ , т. е. $Q = G/N$. По аналогии с группой, для каждого смежного класса qN определим сферу $P_i(q)$, шар $K_s(q)$ и функцию роста $F_i(q)$:

$$P_i(q) = \{g : g \in P_i \text{ и } \varphi(g) = q\}, \quad K_s(q) = \bigcup_{i=0}^s P_i(q), \quad F_i(q) = |P_i(q)|.$$

Если Q — сравнительно большая группа, то множество $K_{2s}(q)$ будет значительно меньше, чем $K_{2s}(G)$. Данный факт взят за основу построения алгоритма 2, который, получив на входе шар K_s группы G радиуса s , фактор-группу $Q = G/N$ и некоторый элемент $q \in Q$, возвращает функцию роста $F(q)$ для шара $K_{2s}(q)$ смежного класса qN радиуса $2s$.

Алгоритм 2. $F(q) = \text{QuotientGrowthFunction}(K_s, Q, q)$

Вход: Шар $K_s = \bigcup P_i$ группы G радиуса s , фактор-группа $Q = G/N$, $q \in Q$

Выход: Функция роста $F(q)$ шара смежного класса $qN \subset G$ радиуса $2s$

- 1: $F(q) := (0, \dots, 0)$ — нулевой вектор размерности $2s + 1$
 - 2: **Для всех** $w \in Q$ и $i = 0, 1, \dots, s$
 - 3: $P_i(w) := \{g \mid g \in P_i \text{ и } \varphi(g) = w\}$
 - 4: $F_i(q) := |P_i(q)|$
 - 5: $K_s(q) := \bigcup_{i=0}^s P_i(q)$
 - 6: **Для всех** $i = s + 1, s + 2, \dots, 2s$
 - 7: $K_i(q) := K_{i-1}(q)$
 - 8: **Для всех** $v \in Q$
 - 9: $u := q \circ v^{-1}$
 - 10: **Для всех** $g_1 \in P_{i-s}(u)$ и $g_2 \in P_s(v)$
 - 11: $g := g_1 \cdot g_2$
 - 12: **Если** $g \notin K_i(q)$, **то**
 - 13: добавить g в $K_i(q)$
 - 14: $F_i(q) := F_i(q) + 1$
 - 15: **Если** $|K_i(q)| = |N|$, **то**
 - 16: переход в п. 17
 - 17: **Вернуть** $F(q)$
-

Лемма 3. Алгоритм 2 корректен, т.е. он, получив на входе шар K_s конечной группы G радиуса s , фактор-группу $Q = G/N$ и некоторый элемент $q \in Q$, за конечное число шагов вычислит функцию роста $F(q)$ для шара $K_{2s}(q)$ смежного класса qN радиуса $2s$.

Доказательство. Алгоритм 2 получает на входе шар $K_s = \bigcup_{i=0}^s P_i$ и в результате шагов 2–5 преобразует его к виду $K_s = \bigcup_{q \in Q} \bigcup_{i=0}^s P_i(q)$, а также вычисляет $K_s(q) = \bigcup_{i=0}^s P_i(q)$ и $F_i(q)$ для $0 \leq i \leq s$.

Отличительной особенностью алгоритма 2 является то, что при $i > s$ не строятся шары K_i всей группы. Вместо этого вычисляются шары $K_i(q)$ смежного класса qN (пп. 6–17 алгоритма).

Заметим, что при $i > s$ мы не сможем по аналогии с алгоритмом 1 вычислить $P_i(q)$ путём умножения элементов порождающего множества на $P_{i-1}(q)$. Вместо этого будем получать элементы $g \in P_i(q)$ путём выбора таких $u, v \in Q$, что $u \circ v = q$, на основе которых вычислим $g = g_1 \cdot g_2$, где $g_1 \in P_{i-s}(u)$ и $g_2 \in P_s(v)$.

По построению алгоритм 2 представляет каждый элемент $K_i(q)$ в виде слова минимальной длины. Конечность K_s гарантирует построение $K_{2s}(q)$ через конечное число шагов. Кроме того, если возникает ситуация, при которой $|K_i(q)| = |N|$, то это означает, что шар смежного класса qN вычислен. В результате построения шара $K_{2s}(q)$ получим искомую функцию роста $F(q)$. ■

Лемма 4. $T_2 \in O(|N| \cdot |K_s|^2)$ и $M_2 \in \Theta(|K_s| + |Q| + |N|)$.

Доказательство. Вычислительная сложность пп. 1–5 алгоритма 2 линейно зависит от $|K_s|$. Наиболее трудоёмкий участок — пп. 6–16. Очевидно, что число элементов g , которые необходимо вычислить, не превышает $|K_s|^2$. Для проверки $g \notin K_{2s}(q)$ необходимо не более $|N|$ элементарных операций, поэтому $T_2 \in O(|N| \cdot |K_s|^2)$.

Для анализа пространственной сложности необходимо принять во внимание только множества K_s , Q и $K_{2s}(q)$, откуда получим $M_2 \in \Theta(|K_s| + |Q| + |N|)$. ■

При реализации алгоритма 2 возникает естественный вопрос: какой порядок $|Q|_{\text{opt}}$ должен быть у группы Q , чтобы M_2 приняла наименьшее значение? Ответ на поставленный вопрос дает следующая

Лемма 5. $\min_{|Q|} (M_2) \in \Theta(\max(|K_s|, |G|^{1/2}))$, где $|Q|_{\text{opt}} = |G|^{1/2}$.

Доказательство. Пусть $f(Q) = |K_s| + |Q| + |N|$. Так как $|N| = |G|/|Q|$, то $f(Q) = |K_s| + |Q| + |G|/|Q|$. Нетрудно показать, что $f(Q)$ принимает минимальное значение при $|Q| = |G|^{1/2}$. Согласно лемме 4, $\min_{|Q|} (M_2) \in \Theta(|K_s| + 2|G|^{1/2}) = \Theta(\max(|K_s|, |G|^{1/2}))$.

Лемма доказана. ■

Замечание 1. Если мы располагаем быстрым способом нумерации элементов $K_{2s}(q)$, например как в [1], то $K_{2s}(q)$ можно представить в виде булева вектора размерности $|N|$, в котором на i -м месте стоит единица, если элемент с указанным номером лежит в $K_{2s}(q)$, и ноль в противном случае. Теперь для хранения элемента достаточно одного бита. Кроме того, по номеру элемента легко осуществить проверку: встречался ли ранее данный элемент группы? Сложность этой операции $O(1)$. В этом случае, взяв во внимание лемму 4, получим $T'_2 \in O(|K_s|^2)$.

3. Алгоритм 3

Объединив алгоритмы 1 и 2, получим алгоритм 3, который вычисляет функцию роста $F(G)$ шара K_{2s} фиксированного радиуса $2s$ произвольной конечной группы G , заданной порождающим множеством X .

Алгоритм 3. $F(G) = \text{GrowthFunction}(G, X, Q, s)$

Вход: X – порождающее множество группы G , фактор-группа $Q = G/N$, радиус s

Выход: Функция роста $F(G)$ шара K_{2s} радиуса $2s$

- 1: $F(G) := (0, \dots, 0)$ – нулевой вектор размерности $2s + 1$
 - 2: $K_s := \text{Ball}(G, X, s)$
 - 3: **Для всех** $q \in Q$
 - 4: $F(q) := \text{QuotientGrowthFunction}(K_s, Q, q)$
 - 5: $F(G) := F(G) + F(q)$
 - 6: **Вернуть** $F(G)$
-

Теорема 1. Алгоритм 3 корректен, т. е. он за конечное число шагов вычисляет функцию роста $F(G)$ шара K_{2s} фиксированного радиуса $2s$ произвольной конечной группы G , заданной порождающим множеством X . Кроме того, верны следующие оценки вычислительной и пространственной сложности данного алгоритма:

- 1) $T_3 \in O(|G| \cdot |K_s|^2)$;
- 2) $M_3 \in \Theta(|K_s| + |Q| + |N|)$;
- 3) $\min_{|Q|} (M_3) \in \Theta(\max(|K_s|, |G|^{1/2}))$, где $|Q|_{\text{opt}} = |G|^{1/2}$.

Доказательство. Следует из лемм 1–5 и формулы $K_{2s} = \bigcup_{q \in Q} K_{2s}(q)$. ■

Замечание 2. Эксперименты в различных группах по алгоритму 3 показали, что для вычисления функции роста всей группы G достаточно подобрать радиус s таким, чтобы выполнялось $|K_s| \sim |G|^{2/3}$. При этом $|Q| \sim |G|^{1/2}$ и, следовательно, $T_3 \in O(|G|^{7/3})$ и $M_3 \in \Theta(|G|^{2/3})$. Для сравнения: $T_0 \in \Theta(|G|^2)$ и $M_0 \in \Theta(|G|)$ – оценки сложности алгоритма А-I из [1]. Для больших групп получим $M_3 \ll M_0$.

Если выполнено условие замечания 1, то $T'_3 \in O(|Q| \cdot |K_s|^2) = O(|G|^{11/6})$. Для модифицированного алгоритма А-I показано [1], что $T'_0 \in \Theta(|G|)$.

Замечание 3. Для снижения вычислительной сложности алгоритма 3 воспользуемся остроумным трюком, предложенным Ч. Симсом [5].

Пусть A – группа автоморфизмов $G = \langle X \rangle$, элементы которой оставляют неизменной длину любого группового слова G . В частности, это автоморфизмы, которые фиксируют порождающее множество X . Кроме того, если $x^{-1} \in X$ для любого $x \in X$, то автоморфизм, преобразующий элементы $g \in G$ в обратные, сохраняет длину g . Предположим, что ядро N инвариантно под действием A . Это означает, что A также действует на Q . Выбрав в каждой A -орбите группы Q по одному представителю, получим множество Q_0 . Для всех $q \in Q_0$ определим размер орбиты $l(q)$. Теперь достаточно вычислить шары $K_{2s}(q)$ только для $q \in Q_0$. В результате пп. 3–5 алгоритма 3 примут следующий вид:

- 3: **Для всех** $q \in Q_0$ // параллельно
- 4: $F(q) := \text{QuotientGrowthFunction}(K_s, Q, q)$
- 5: $F(G) := F(G) + l(q) \cdot F(q)$

Отметим, что цикл по переменной q легко распараллеливается, что позволяет значительно ускорить вычисления.

4. Компьютерные вычисления в двупорождённых группах периода 5

Пусть $B_0(2, 5) = \langle a_1, a_2 \rangle$ — максимальная конечная двупорождённая бернсайдова группа периода 5, порядок которой равен 5^{34} [6]. Используя систему компьютерной алгебры GAP, несложно получить pc -представление (*Power Commutator presentation*) данной группы [8, 7]. В этом случае каждый элемент $g \in B_0(2, 5)$ может быть однозначно записан в следующем виде:

$$g = a_1^{\alpha_1} \cdot a_2^{\alpha_2} \cdot \dots \cdot a_{34}^{\alpha_{34}}, \quad \alpha_i \in \mathbb{Z}_5, \quad i = 1, 2, \dots, 34.$$

Здесь a_1 и a_2 — порождающие элементы $B_0(2, 5)$; a_3, \dots, a_{34} — коммутаторы, которые вычисляются рекурсивно через a_1 и a_2 [6].

Обозначим через B_k фактор-группу $B_0(2, 5)$ следующего вида:

$$B_k = B_0(2, 5) / \langle a_{k+1}, \dots, a_{34} \rangle.$$

Очевидно, что $|B_k| = 5^k$ и для всех $g \in B_k$

$$g = a_1^{\alpha_1} \cdot a_2^{\alpha_2} \cdot \dots \cdot a_k^{\alpha_k}.$$

Пусть $A_2 = \{a_1, a_2\}$ и $A_4 = \{a_1, a_1^{-1}, a_2, a_2^{-1}\}$ — минимальное и симметричное порождающие множества групп B_k соответственно.

Вычислить функцию роста $B_0(2, 5)$ относительно порождающего множества A_2 или A_4 в настоящее время едва ли возможно, поскольку количество её элементов очень велико:

$$5^{34} = 582076609134674072265625 \approx 5 \cdot 10^{23}.$$

Отметим, что на данный момент при помощи компьютерных вычислений удалось получить функции роста групп B_k , порядок которых не превышает 5^{17} [1].

Как уже было сказано, попытка применения алгоритма А-Г из [1] для исследования роста группы B_{18} , состоящей из $5^{18} \approx 4 \cdot 10^{12}$ элементов, потерпела неудачу: возникла принципиальная проблема — нехватка объёма памяти. В связи с этим была осуществлена попытка применить алгоритм 3, который, согласно замечанию 2, имеет значительно меньшую пространственную сложность.

Алгоритм 3 был реализован на языке C++. Для снижения вычислительной сложности (см. замечание 1) элементы смежных классов qN нумеровались аналогичным с [1] способом. Учитывая замечание 3, при помощи GAP для каждого случая найдены A , Q , Q_0 и $l(q)$. Для эффективного умножения элементов применялись полиномы Холла [9]. В качестве инструмента распараллеливания использована библиотека OpenMP. Для вычислений был задействован компьютер, имеющий два 16-ядерных процессора и 64 Гбайта оперативной памяти, на котором установлена операционная система Linux. Трансляция программ осуществлялась встроенным в систему компилятором gcc.

На рис. 1–3 представлены графики функций роста групп $B_{18} = \langle A_2 \rangle$, $B_{18} = \langle A_4 \rangle$ и $B_{19} = \langle A_4 \rangle$. Для наглядности на каждом рисунке проведена аппроксимирующая гауссова кривая. В таблице указаны диаметры D и средние диаметры \bar{D} соответствующих графов Кэли, параметры алгоритма, время вычислений, а также объём используемой оперативной памяти.

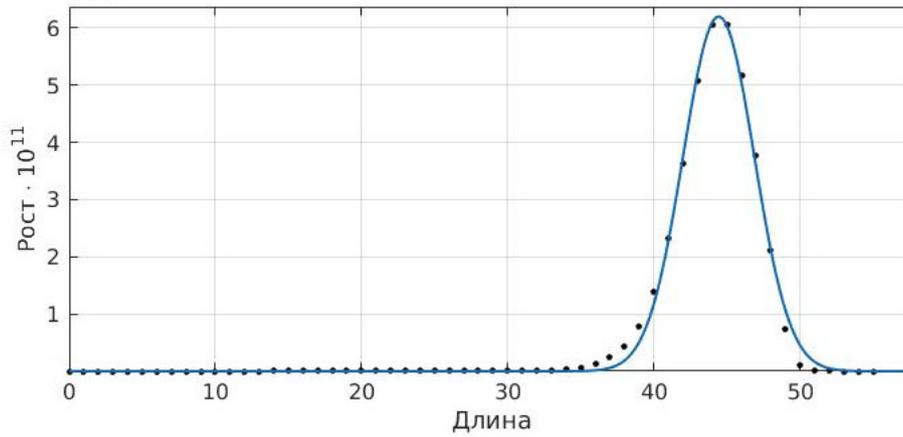


Рис. 1. График функции роста группы $B_{18} = \langle A_2 \rangle$

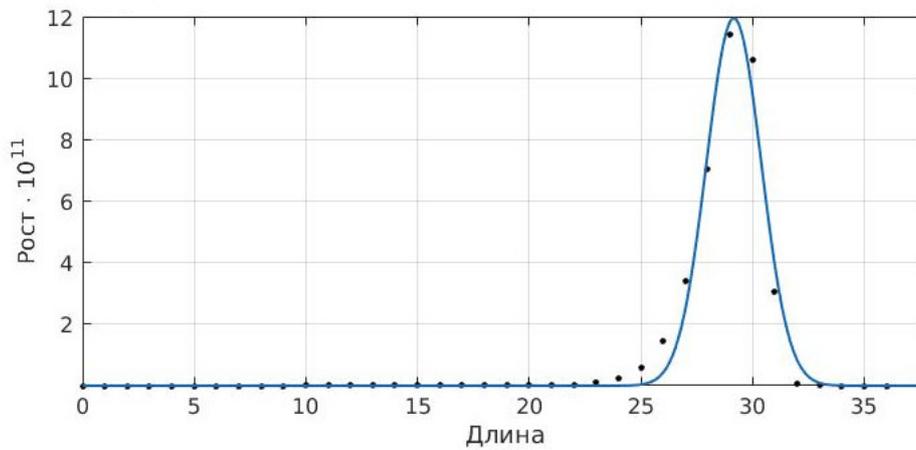


Рис. 2. График функции роста группы $B_{18} = \langle A_4 \rangle$

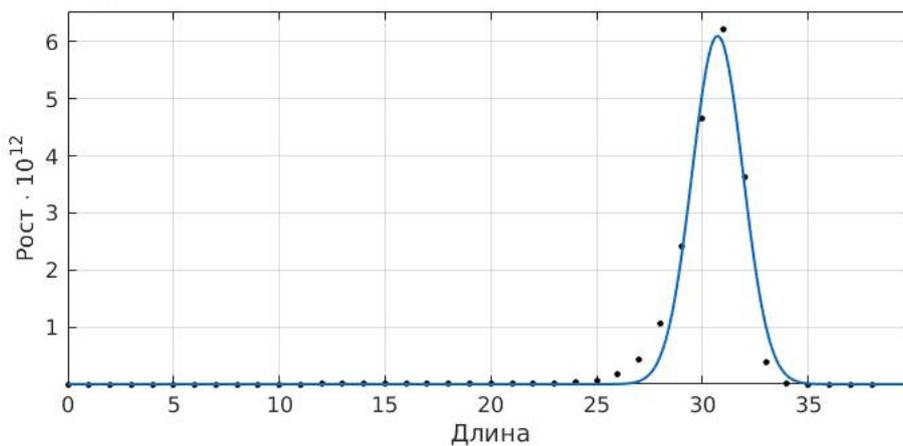


Рис. 3. График функции роста группы $B_{19} = \langle A_4 \rangle$

| Группа | D | \bar{D} | Параметры алгоритма | Время | Память |
|--------------------------------|-----|-----------|--|--------|----------|
| $B_{18} = \langle A_4 \rangle$ | 36 | 29 | $ A = 16, Q = 5^8, Q_0 = 25311$ | 62 ч | 8 Гбайт |
| $B_{18} = \langle A_2 \rangle$ | 55 | 44 | $ A = 2, Q = 5^8, Q_0 = 195375$ | 20 сут | 11 Гбайт |
| $B_{19} = \langle A_4 \rangle$ | 38 | 31 | $ A = 8, Q = 5^{10}, Q_0 = 1226797$ | 44 сут | 25 Гбайт |

На рис. 4 и 5 приведены графики известных диаметров графов Кэли $D_{A_2}(B_k)$ и $D_{A_4}(B_k)$, а также их линейные аппроксимирующие функции, на основе которых можно сделать следующее предположение.

Гипотеза 1. $D_{A_2}(B_0(2, 5)) \approx 105$ и $D_{A_4}(B_0(2, 5)) \approx 69$.

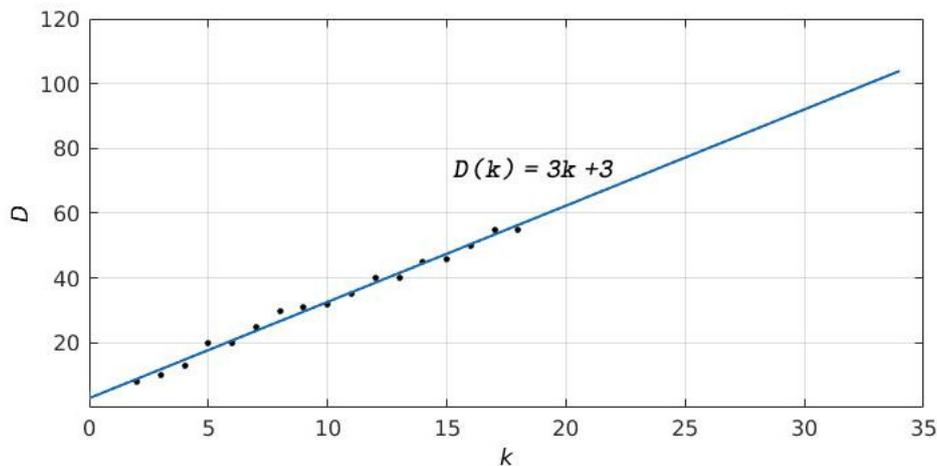


Рис. 4. График $D_{A_2}(B_k)$

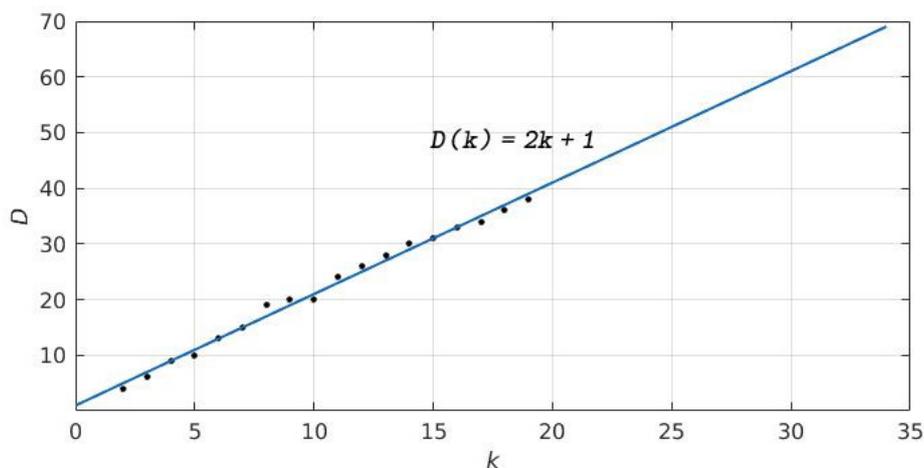


Рис. 5. График $D_{A_4}(B_k)$

ЛИТЕРАТУРА

1. Кузнецов А. А. Об одном алгоритме вычисления функций роста в конечных двупорождённых группах периода 5 // Прикладная дискретная математика. 2016. № 3(33). С. 116–125.
2. Кузнецов А. А., Кузнецова А. С. Параллельный алгоритм для исследования графов Кэли групп подстановок // Вестник СибГАУ. 2014. № 1. С. 34–39.
3. Even S. and Goldreich O. The Minimum Length Generator Sequence is NP-Hard // J. Algorithms. 1981. No. 2. P. 311–313.
4. Skiena S. The Algorithm Design Manual. London: Springer Science+Business Media, 2008. 730 p.
5. Sims C. Fast multiplication and growth in groups // Proc. 1998 Intern. Symp. Symbolic Algebraic Computation. 1998. P. 165–170.

6. *Havas G., Wall G., and Wamsley J.* The two generator restricted Burnside group of exponent five // Bull. Austral. Math. Soc. 1974. No. 10. P. 459–470.
7. *Sims C.* Computation with Finitely Presented Groups. Cambridge: Cambridge University Press, 1994. 628 p.
8. *Holt D., Eick B., and O'Brien E.* Handbook of Computational Group Theory. Boca Raton: Chapman & Hall/CRC Press, 2005. 514 p.
9. *Кузнецов А. А. Кузнецова А. С.* Быстрое умножение элементов в конечных двупорождённых группах периода пять // Прикладная дискретная математика. 2013. № 1. С. 110–116.

REFERENCES

1. *Kuznetsov A. A.* Ob odnom algoritme vychisleniya funkciy rosta v konechnykh dvuporozhdyonnykh gruppakh perioda 5 [An algorithm for computation of the growth functions in finite two-generated groups of exponent 5]. Prikladnaya Diskretnaya Matematika, 2016, no. 3(33), pp. 116–125. (in Russian)
2. *Kuznetsov A. A. and Kuznetsova A. S.* Parallelnyy algoritm dlya issledovaniya grafov Keli grupp podstanovok [A parallel algorithm for study of the Cayley graphs of permutation groups]. Vestnik SibSAU, 2014, no. 1, pp. 34–39. (in Russian)
3. *Even S. and Goldreich O.* The Minimum Length Generator Sequence is NP-Hard. J. Algorithms, 1981, no. 2, pp. 311–313.
4. *Skiena S.* The Algorithm Design Manual. London, Springer Science+Business Media, 2008. 730 p.
5. *Sims C.* Fast multiplication and growth in groups. Proc. 1998 Intern. Symp. Symbolic Algebraic Computation, 1998, pp. 165–170.
6. *Havas G., Wall G., and Wamsley J.* The two generator restricted Burnside group of exponent five. Bull. Austral. Math. Soc., 1974, no. 10, pp. 459–470.
7. *Sims C.* Computation with Finitely Presented Groups. Cambridge, Cambridge University Press, 1994. 628 p.
8. *Holt D., Eick B., and O'Brien E.* Handbook of Computational Group Theory. Boca Raton, Chapman & Hall/CRC Press, 2005. 514 p.
9. *Kuznetsov A. A. and Kuznetsova A. S.* Bystroe umnozhenie elementov v konechnykh dvuporozhdennykh gruppakh perioda pyat' [Fast multiplication in finite two-generated groups of exponent five]. Prikladnaya Diskretnaya Matematika, 2013, no. 1, pp. 110–116. (in Russian)

ДИСКРЕТНЫЕ МОДЕЛИ РЕАЛЬНЫХ ПРОЦЕССОВ

УДК 004.942

**МОДЕЛИРОВАНИЕ ДВИЖЕНИЯ ОДНОКЛЕТОЧНОГО
МИКРООРГАНИЗМА «АМОЕВА PROTEUS»
МЕТОДОМ ПОДВИЖНЫХ КЛЕТОЧНЫХ АВТОМАТОВ**

Е. П. Газдюк*, В. В. Жихаревич*, О. М. Никитина**, С. Э. Остапов*

** Черновицкий национальный университет имени Юрия Федьковича, г. Черновцы, Украина**** Черновицкий факультет Национального технического университета
«Харьковский политехнический институт», г. Черновцы, Украина*

Работа посвящена моделированию амебоидной подвижности. В качестве объекта выбран одноклеточный микроорганизм «Amoeba Proteus», рассмотрены основные принципы локомоции, на основании которых построена модель. В качестве метода моделирования задействован аппарат подвижных клеточных автоматов. Получена компьютерная модель, имитирующая амебоидную локомоцию.

Ключевые слова: *подвижные клеточные автоматы, амебоидная подвижность, компьютерное моделирование, принцип соседства.*

DOI 10.17223/20710410/42/8

**THE UNICELLULAR MICROORGANISMS “АМОЕВА PROTEUS”
LOCOMOTION SIMULATION WITH THE USE OF MOVABLE
CELLULAR AUTOMATA METHOD**

Ye. P. Hazdiuk*, V. V. Zhikharevich*, O. M. Nikitina**, S. E. Ostapov*

** Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine**** Chernivtsi Department of National Technical University “Kharkiv Polytechnic Institute”,
Chernivtsi, Ukraine***E-mail:** kateryna.gazdyik@gmail.com

In this work, the method of movable cellular automata is applied to the modeling of amoeba-like locomotion. A significant advantage of this method is the possibility of transition from a static grid to the concept of neighbors. A unicellular biological organism “Amoeba Proteus” was chosen as an object. The basic principles of locomotion, namely the movement of the amoeba on the basis of cytoskeletal transformations inside the cell, are considered. This approach most accurately describes the process of locomotion in the living cell. The rules of cellular automata interactions were found for the constructed model according to the concept of neighbors. As a result, a computer model imitating amoeboid locomotion was obtained.

Keywords: *movable cellular automata, amoeba-like movement, computer simulation, neighborhood principle.*

Введение

По мере накопления знаний о жизнедеятельности организмов становится очевидным, сколько ещё не изученных вопросов содержит данная предметная область. Поведение, анатомия, клеточные функции, нейронные системы, последовательности генов — все эти компоненты, составляющие живое существо, являются субъектами массовых исследований, над которыми работают тысячи людей во всем мире, чтобы разгадать сложные тайны, установленные природой.

Одна из неотъемлемых особенностей всего живого на Земле — способность двигаться. Коллективное движение возникает в живых клетках, играя важную функциональную роль. Все движения, на которые способны биологические системы, основываются на биологической активности клеток или их совокупности. Подвижность на клеточном уровне проявляется в самых разнообразных конфигурациях — начиная от сокращений различных типов мышечных клеток, приводящих к перемещению биологических организмов, и заканчивая внутриклеточным движением, играющим важную роль в осуществлении обмена и распределении веществ внутри клетки.

В основе перемещения в пространстве, необходимого как отдельным свободноживущим клеткам, так и клеткам, составляющим живой организм, лежит непрерывный процесс самоорганизации движения и самого двигательного аппарата. В основе процесса самоорганизации лежит отсутствие единого управляющего центра. В отличие от подхода централизованного управления, каждый элемент самоорганизующихся систем действует сам по себе, взаимодействуя только с небольшим количеством соседних элементов. Этого оказывается достаточно для того, чтобы упорядочить хаотические структуры.

В перспективе изучение и решение поставленных проблем самоорганизации биологических организмов даст ключ к основополагающим принципам, лежащим в основе всей жизни на Земле, вплоть до остановки движения и роста раковых клеток и замедления процессов старения.

Целью работы является поиск методов моделирования амебоидной подвижности, поскольку многие тканевые клетки животных (например, лейкоциты крови) имеют амебоидный тип подвижности и исследование закономерностей их локомоции является важной как теоретической, так и практической задачей.

1. Принципы амебоидной подвижности

Амебоидная локомоция — это общий термин, описывающий подвижность, которая присуща адгезивным эукариотическим клеткам, движущимся путём распространения псевдоподий, потоковой передачи цитоплазмы и изменения их формы. Амебоидная локомоция проявляется многими типами клеток, включая свободноживущие амебы и клетки млекопитающих. В последнем случае амебоидная локомоция имеет жизненно важное значение для развития эмбриона, заживления ран, работы иммунной системы и т. д. Она также отвечает за распространение опухолей. Физические теории амебоидной локомоции развивались с середины XIX века [1]. Новые методы позволили измерить силы, действующие и испытываемые клетками, а также силы и динамику отдельных молекул цитоскелета [2]. Крупные свободноживущие амебы, такие как *Amoeba Proteus* и *Chaos*, демонстрируют псевдоподии с четким различием между слоями эктоплазматического геля и эндоплазматического золья, которые свободно перетекают по мере движения амебы (рис. 1).

Важную роль в процессе локомоции выполняет цитоскелет [3]. Общим для элементов цитоскелета является то, что все они представляют собой белковые неветвящиеся

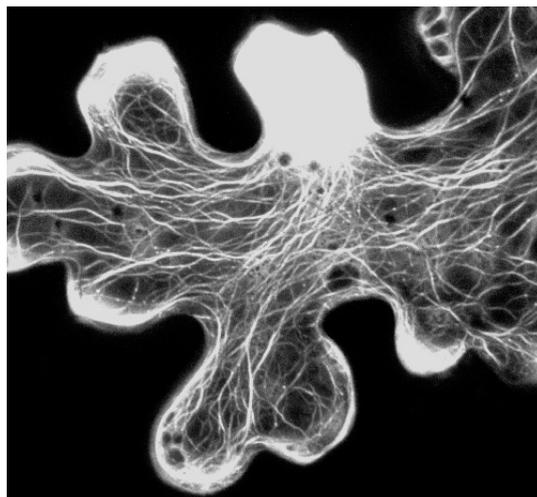


Рис. 1. Амoeba Proteus — вид с цитоскелетом

фибрилярные полимеры, нестабильные, способные к полимеризации и деполимеризации. При взаимодействии с другими специальными белками-транслокаторами (или моторными белками) они могут участвовать в разнообразных клеточных движениях. Прежде всего, цитоскелет придает клетке форму и механическую устойчивость к деформации, а также обеспечивает связь между мембраной и органеллами. Он, как каркас, представляет собой динамичную структуру, которая постоянно обновляется по мере изменения внешних условий и состояния клетки.

Цитоскелет также может сокращаться, тем самым деформируя клетку и её среду и позволяя ей двигаться. Он организует движение органоидов в цитоплазме (т. н. течение протоплазмы), лежащее в основе амeбoидного движения. Компоненты цитоскелета определяют направление и координируют движение, деление, изменение формы клеток в процессе роста, перемещение органелл, движение цитоплазмы. Он служит в качестве «рельсов» для транспортировки органелл и других крупных комплексов внутри клетки.

2. Обзор существующих исследований

На сегодняшний день достаточно активно создаются и исследуются компьютерные модели развития и поведения различных элементарных микроорганизмов (вольвоксов, нематод, гидр, амeб и т. д.). Основной целью таких исследований является определение общих механизмов и принципов самоорганизации потоков сигналов, управляющих ими. Исследования в данной отрасли начались во второй половине XX столетия [2, 4] с изучения эффектов самоорганизации в физических системах и родственных им явлений в более сложных биологических системах, а также конкретных механизмов возникновения и совершенствования организации [5].

В работе [6] разработан агентный подход для моделирования поведения и процессов в клетках. Он представляет собой программное обеспечение высокого уровня, которое обеспечивает удобный и мощный способ описания абстрактной модели с точки зрения её поведения в контекстной вычислительной среде. Преимуществом этого метода является то, что возникающие явления могут быть смоделированы с помощью простых правил, регулирующих поведение каждого агента. Кроме того, динамические структуры биологических систем могут быть интуитивно представлены и эффективно реализованы в агентно-ориентированных симуляторах. Авторами разработана струк-

тура, которая может быть использована для моделирования различных типов клеток и клеточных процессов, обучения и адаптивного поведения открытых систем.

В [7] исследована роль автоколебаний в биологической подвижности, разработана модель автоколебательных движений плазмодия *Physarum Polycerphalum*. Разработанная математическая модель в некотором приближении может описать периоды колебаний скорости протоплазмы как реакцию на лазерное освещение, а также стимулировать новые опыты по исследованию фотореакции амeboидных клеток.

Цель сборника научных трудов [8], которые касаются построения теоретических моделей и разработки нелинейных аналитических методов, заключается в том, чтобы предоставить достаточно информации, позволяющей не только специалисту, но и технически ориентированному читателю следовать основным теоретическим рассуждениям. Использование концепций нелинейной биологической динамики, или, вкратце, биодинамики, для постановки и решения критических исследовательских вопросов быстро расширяется по многим биологическим дисциплинам от клеточной и молекулярной биологии до неврологии. Этот обзор демонстрирует, что концепции нелинейной динамики в целом и активных возбуждающих колебаний в частности являются незаменимыми для описания и анализа биологических ритмов.

Математическая модель амeboидных клеток разработана японскими учёными [9]. Эта модель описывает одностороннее движение клеточных и внутриклеточных реакций активатора, ингибитора и актиновых филаментов. Показано, что скорость производства активатора является ключевым фактором для определения того, станет ли форма ячейки типом нейтрофилов или кератоцитарным типом. В дальнейших работах [10, 11] для выяснения физического механизма амeboидной динамики разработана вычислительная модель, которая выделяет группу ингибирующих молекул для полимеризации актина. Основываясь на этой модели, авторы предлагают гипотезу о том, что тормозящие молекулы отправляются в заднюю часть движущейся клетки, чтобы там накапливаться. Последовательность переключения режимов между персистентными движениями и случайными поворотами приводит к супердиффузионной миграции при отсутствии внешнего наводящего сигнала.

Отдельное внимание стоит уделить проектам по реализации реальных механических прототипов, принцип движения которых основан на амeboидной подвижности. Целью исследования японских учёных [12, 13] является понимание основополагающего механизма поведенческого разнообразия живых организмов, а затем использование результатов для построения адаптивных роботов. Одним из организмов, которые привлекли их внимание, является плазмодий (*Physarum Polycerphalum*), демонстрирующий принципы амeboидной подвижности. Разработана математическая модель и настоящий физический робот, в который включены два децентрализованных контроллера. Численные и экспериментальные результаты показывают, что, объединяя контроллеры с разными временными константами, робот может использовать предложенную модель для успешного согласования «узкого прохода» путём деформирования его формы тела динамически.

Вдохновленные простой амeboй, исследователи из Virginia Tech разработали новую форму локомоции для робототехники [14], основанную на способе передвижения одноклеточной амeboй. В отличие от любых других роботов, технология Virginia Tech предназначена для использования всей своей внешней кожи как средства движения. Тороидальные по форме, немного похожие на удлиненные цилиндрические пончики, роботы этой новой породы отличаются от колесных, гусеничных или ножных ботов тем, что они перемещаются, постоянно поворачиваясь изнутри. Сжимая кольца в зад-

ней части робота и расширяя их по направлению к фронту, они могут генерировать движение. Это очень похоже на принцип псевдоподий, используемый одноклеточными организмами, такими как амёбы.

Стоит отметить, что, несмотря на широкий диапазон исследований в этой отрасли биоинженерии, подход к моделированию локомоции одноклеточного организма под воздействием цитоскелетных преобразований до сих пор не был рассмотрен. Новизна наших исследований заключается в использовании этого существенно нового подхода к моделированию локомоции одноклеточного организма «Amoeba Proteus», который максимально точно имитирует процесс локомоции в живой клетке. Это даёт возможность прогнозировать перемещение клеток амёбоидного типа для дальнейшего изучения их взаимодействий.

3. Подход и метод исследования

Классические инженерные системы состоят из ряда уникальных гетерогенных компонентов, собранных очень тонкими и сложными способами. Ожидается, что они будут работать определённо и предсказуемо, следуя спецификациям, заданным их разработчиками. Напротив, самоорганизация в естественных системах (физическая, биологическая, экологическая, социальная) часто опирается на мириады идентичных агентов и, по существу, подчиняется стохастической динамике. Здесь нетривиальные шаблоны и коллективное поведение могут возникать из относительно простых правил для отдельных агентов — факт, который характеризует сложные системы. В работе [15] предложен детальный обзор подходов и методов исследования процессов самоорганизации, в котором основное внимание уделяется строгим архитектурным и сложным функциональным свойствам систем и тому, как эти свойства могут быть реализованы или запрограммированы на микроуровне.

Учитывая вышеизложенное, мы предлагаем метод подвижных клеточных автоматов (ПКА) для исследования модельного микроорганизма одного из видов наипростейших — амёбы (лат. Amoeba Proteus). Это одноклеточный микроорганизм, передвигающийся за счёт роста мембраны в виде так называемых псевдоножек. При этом некоторые другие части мембраны уменьшаются. Изменения в мембране обусловлены обратимыми преобразованиями «золь» ↔ «гель». Руководит этими процессами цитоскелет, структура которого постоянно меняется аналогично динамике активных нейронных сетей в нервных системах высших организмов. Результатом сложных мембранно-цитоскелетных взаимодействий является гармоничная динамика движения микроорганизма.

Модель амёбы построена на основе ПКА. Согласно классификации, предложенной в [16], модель является асинхронным стохастическим ПКА с символьным алфавитом, что соответствует множеству состояний ПКА. Заметим, что, в отличие от классического подхода, метод ПКА позволяет перейти от статической сеточной концепции к концепции соседей [17–19]. Описанные в рассмотренных работах модели ПКА являются детерминированными синхронными и используются для моделирования физико-механических процессов в твёрдых деформируемых телах. Предложенный в данной работе вариант подвижных клеточных автоматов является дальнейшим развитием вышеупомянутых моделей и воплощает такие основные идеи ПКА, как подвижность элементов и многочастичность их взаимодействия. Автоматы имеют возможность менять своих соседей путём переключения состояния пар. Как известно, основное отличие ПКА от классических клеточных автоматов состоит в способности ПКА принимать произвольные непрерывные значения координат в пространстве. В связи с этим

возникает необходимость выделения дополнительных параметров в структурах данных, хранящих информацию о состоянии всего множества ПКА. Мы использовали динамический индексный массив, позволяющий не только идентифицировать любой ПКА в системе, но и содержащий индексы его ближайших соседей (рис. 2).



Рис. 2. Схематическое представление индексного массива ПКА[N, P]:
 N — количество ПКА; P — количество параметров

В качестве параметров выберем: тип ПКА, определяющий его свойства; координаты (для простоты и наглядности остановимся на двумерной модели); индексы ближайших соседей. При моделировании использована гексагональная схема соседства ($M = 6$), предусматривающая слабую и жёсткую связность пар. Кроме того, для уменьшения количества соседей будем вносить в ячейки индексных соседей нулевые значения.

Метод поиска ближайших соседей в условиях их фиксированной величины описан в [20]. В этой работе критерием поиска соседних ячеек является минимальное расстояние между ними, а поиск осуществляется путём введения вспомогательного индексного массива. Предлагаемый алгоритм показал хорошие результаты как для равномерного, так и для случайного распределения ячеек в пространстве. Существенным преимуществом является скорость и низкая потребность в вычислительных ресурсах, что важно для разработки и функционирования процессов, связанных с эволюцией динамических автономных систем. Пренебрежение инерционными свойствами тела амёбы позволяет отказаться от содержания в индексном массиве компонент векторов скорости ПКА (рис. 3).

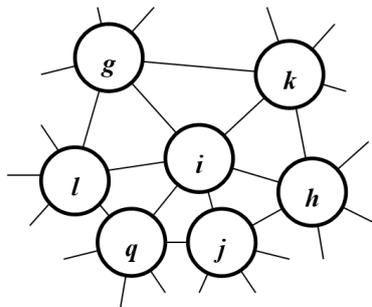


Рис. 3. Фрагмент множества ПКА

Основная суть работы алгоритма — итерационная модификация содержимого индексного массива. Для этого реализован асинхронный подход, при котором равновероятным образом выбираются один из N ПКА и один из M его ближайших соседей:

```
for i = random(1...N)
j = MCA[i, P - random(1...M) + 1]
endfor
```

Здесь i — индекс выбранного ПКА; j — индекс его соседа; МСА (Movable Cellular Automata) — индексный массив, схематически изображённый на рис. 3. По индексам двух выбранных ПКА определяются их типы: $a_s^i = \text{МСА}[i, 1]$, $a_s^j = \text{МСА}[j, 1]$, где s — тип (состояние) i -го и j -го ПКА. Выбор асинхронного подхода при разработке алгоритма ПКА обусловлен тем, что позволяет избежать коллизий, то есть удовлетворить критерию корректности (не будет ни одной попытки изменить состояние одной и той же клетки более одного раза в один и тот же момент времени t).

Клеточно-автоматная модель допускает шесть возможных состояний подвижных клеточных автоматов, которые описываются при помощи алфавита $A = \{a_1, a_2, \dots, a_6\}$ и имеют разное назначение и функционал, для различия которых используются круги радиусов r_i , $i = 1, 2, 3$, $r_1 < r_2 < r_3$.

Примем следующие соответствия. ПКА типа a_1 , моделирующие центральную вытянутую часть жидкости («золь»), обозначены кругами зелёного цвета радиуса r_2 . Они могут вступать во взаимодействие с такими же ПКА и ПКА мембраны (типа a_2) — «геля» — более жёсткого слоя, окружающего организм. Эти ПКА обозначены красными кругами радиуса r_2 .

ПКА радиуса r_3 типа a_3 — аналог centrosомы. Он также является неким управляющим центром, из которого растут нити цитоскелета и периодически спонтанно саморазрушаются.

ПКА, которые отвечают за нити цитоскелета (типов a_4, a_5, a_6), обозначены соответственно синими, лиловыми и бирюзовыми кругами радиуса r_1 в зависимости от функций, которые они определяют. Так, синим кругом радиуса r_1 обозначены нити цитоскелета, которые задают направление роста псевдоподий (побуждают переход клетки типа «золь» в тип «гель»). Направление роста нити может отвечать градиенту наибольшей концентрации полезных веществ, а сам рост нити происходит из «центросомы». Лиловым кругом радиуса r_1 обозначены нити цитоскелета, катализирующие перенос «геля» в «золь». Этот процесс является обратным предыдущему и моделирует рост уроподий на другой стороне амёбы, противоположной к градиенту движения. Бирюзовым кругом радиуса r_1 маркированы клетки, соответствующие фрагментам цитоскелета, которые предназначены для образования тела амёбы.

Описанная типизация — следствие стремления построить, для начала, довольно простую модель амёбы. Как известно, процесс формообразования тела амёбы управляется цитоскелетом. При этом постоянно происходит самосборка и разрушение фрагментов цитоскелета. Самосборка обычно осуществляется в «центросоме». Можно выделить разные типы нитей цитоскелета (филаментов или микротрубочек), которые по-разному влияют на превращение в мембране (актиновые и миозиновые филаменты). В одних местах при взаимодействии с соответствующими типами окончаний нитей филаментов происходит переход из состояния «гель» в состояние «золь» и мембрана стягивается, в других протекает обратный процесс и мембрана растёт в виде ложноножек. После выбора двух соседних ПКА реализуется функция взаимодействий. Данная

функция зависит от типов взаимодействующих ПКА, что удобно представить в матричном виде (рис. 4). Естественно, матрица должна быть симметричной. При этом в некоторых случаях, описанных ниже, взаимодействия осуществляются с участием более двух ПКА.

| | | Соседний ПКА j | | | | | |
|-------------------|-------|------------------|-------|-------|-------|-------|-------|
| | | a_1 | a_2 | a_3 | a_4 | a_5 | a_6 |
| Выбранный ПКА i | a_1 | F_2 | F_2 | F_0 | F_0 | F_0 | F_0 |
| | a_2 | F_2 | F_3 | F_1 | F_4 | F_5 | F_1 |
| | a_3 | F_0 | F_1 | F_1 | F_7 | F_7 | F_7 |
| | a_4 | F_0 | F_4 | F_7 | F_6 | F_0 | F_0 |
| | a_5 | F_0 | F_5 | F_7 | F_0 | F_6 | F_0 |
| | a_6 | F_0 | F_1 | F_7 | F_0 | F_0 | F_6 |

Рис. 4. Матричное представление функции взаимодействий ПКА

Каждая из функций F — композиция отдельных элементарных операций, осуществляемых при взаимодействиях соответствующих типов. Опишем их:

$$F_0 = \{f_0\};$$

$$F_1 = \{f_0, f_1\};$$

$$F_2 = \{f_0, f_1, f_2\};$$

$$F_3 = \{f_0, f_1, f_2, f_3\};$$

$$F_4 = \{f_4, f_7, f_9\}, \text{ если есть терминальный ПКА, } F_4 = F_1 \text{ в противном случае;}$$

$$F_5 = \{f_5, f_6, f_9\}, \text{ если есть терминальный ПКА, } F_5 = F_1 \text{ в противном случае;}$$

$$F_6 = F_1, \text{ если есть терминальный ПКА, } F_6 = F_3 \text{ в противном случае;}$$

$$F_7 = \{f_0, f_1, f_2, f_8\}.$$

Предназначения функций следующие: F_0 — имитация тепловых колебаний; F_1 — отталкивание двух ПКА при их взаимопроникновении; F_2 — отталкивание при сближении и притягивание при удалении (имитация конденсированного состояния среды); F_3 — выравнивание ПКА (для моделирования «упругих» свойств нитей цитоскелета и поверхности мембраны); F_4 — катализ преобразования «золь» \leftrightarrow «гель», при котором ПКА внутренней среды a_1 преобразуются в ПКА мембраны a_2 ; F_5 — катализ обратного преобразования «гель» \leftrightarrow «золь»; F_6 — аналог F_3 , но если ПКА является окончанием нити цитоскелета (терминальным), то ведёт себя как F_1 при взаимодействии с однотипными ПКА; F_7 — имитация роста нитей из центросомы. При этом действуют следующие ограничения, учитываемые при взаимодействиях:

- 1) ПКА мембраны a_2 не могут иметь более двух соседей такого же типа a_2 . Таким образом обеспечивается целостность мембраны и предотвращается её «слипание». В дальнейшем при моделировании процессов захвата и переваривания пищи, а также деления амёб это ограничение будет модифицировано.
- 2) ПКА нитей цитоскелета a_4 , a_5 и a_6 , если они не терминальные, вообще не могут иметь более двух соседей, что обеспечивается путём занесения четырёх нулевых значений в поле индексов соседей массива МСА.
- 3) ПКА, соседствующие с a_3 (клеткой центросомы), считаются с ней связанными. Таким образом, ПКА центросомы может иметь произвольное количество связанных с ней нитей цитоскелета. В дальнейшем при моделировании процесса деления амёб будет предусмотрена возможность деления ПКА центросомы,

что, в свою очередь, будет инициировать процесс разделения тела амёбы на два отдельных одноклеточных организма.

На рис. 5–9 схематически показаны основные клеточно-автоматные взаимодействия, которые приводят к амёбной подвижности и являются следствием композиции соответствующих элементарных операций.

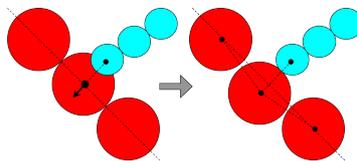


Рис. 5. ПКА, соответствующие фрагментам цитоскелета, предназначенным для образования тела амёбы, при взаимодействии с ПКА «геля» удерживают их на определённом расстоянии

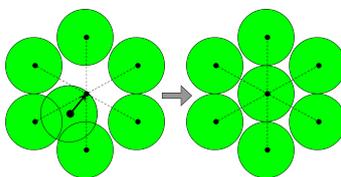


Рис. 6. ПКА типа «золь» и «гель», стремящиеся расположиться равномерно относительно своих ближайших соседей

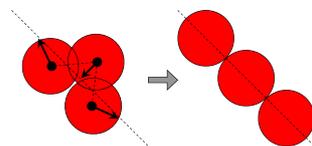


Рис. 7. Клетки, отвечающие мембране амёбы (слой «геля»), а также клетки фрагментов цитоскелета, стремящиеся выровняться по одной линии

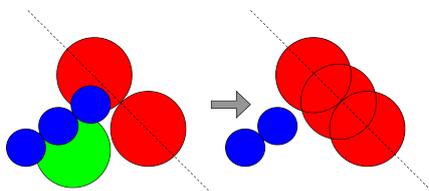


Рис. 8. ПКА, отвечающие фрагментам цитоскелета, инициирующим процесс превращения «золь» → «гель», при взаимодействии с ПКА «геля» добавляют новый элемент к мембране и удаляют ближайший элемент «золя»

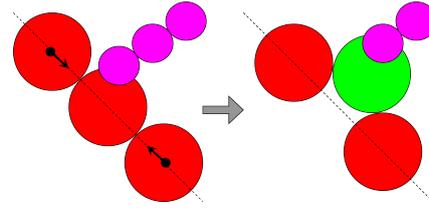


Рис. 9. ПКА, отвечающие фрагментам цитоскелета, инициирующим процесс превращения «гель» → «золь», при взаимодействии с ПКА «геля» удаляют ближайший элемент мембраны и добавляют новый элемент «золя»

Распишем реализацию отдельных элементарных операций, из которых состоят приведённые функции ПКА-взаимодействий.

- f_0 — простейшая операция, осуществляющая имитацию тепловых колебаний ПКА, то есть случайное смещение координат ПКА в произвольном направлении:

```
alpha = 2*pi*random(360)/360
r = rmax*random(100)/100
xi = MCA[i, 2]
yi = MCA[i, 3]
xi = xi + r*cos(alpha)
yi = yi + r*sin(alpha)
```

Здесь α — угол смещения; r — величина смещения; r_{\max} — максимально возможное расстояние смещения при тепловых колебаниях; x_i и y_i — координаты i -го ПКА.

- f_1 — операция, осуществляющая отталкивание двух ПКА при их перекрытии. При этом имитируется несжимаемость среды. Каждому типу a_i поставлен в соответствие радиус r_i , и при сближении на расстояние, меньшее $(r_i + r_j)$, ПКА необходимо оттолкнуть:

```
d = sqrt((xi - xj)^2 + (yi - yj)^2)
delta = (ri + rj) - d
if delta > 0 then
    cos(alpha) = (xi - xj)/d
    sin(alpha) = (yi - yj)/d
    xi = xi + delta*rj/(ri + rj)*cos(alpha)
    yi = yi + delta*rj/(ri + rj)*sin(alpha)
    xj = xj - delta*ri/(ri + rj)*cos(alpha)
    yj = yj - delta*ri/(ri + rj)*sin(alpha)
endif
```

- f_2 — операция, осуществляющая притягивание двух ПКА. При этом имитируется конденсированное состояние среды и предотвращается появление пустот. Таким образом, вместе с операцией f_1 обеспечивается постоянство объёма текучей внутренней среды тела амебы. Кроме того, операция f_2 используется в мембранных и цитоскелетных взаимодействиях для обеспечения их целостности. При отдалении на расстояние, большее $(r_i + r_j) + \delta_{\max}$, где $\delta_{\max} \geq 0$ — максимально возможное отдаление двух ПКА, их необходимо сблизить:

```
d = sqrt((xi - xj)^2 + (yi - yj)^2)
```

```

delta = d - (ri + rj + deltamax)
if delta>0 then
  cos(alpha) = (xi - xj)/d
  sin(alpha) = (yi - yj)/d
  xi = xi - delta*rj/(ri + rj)*cos(alpha)
  yi = yi - delta*rj/(ri + rj)*sin(alpha)
  xj = xj + delta*ri/(ri + rj)*cos(alpha)
  yj = yj + delta*ri/(ri + rj)*sin(alpha)
endif

```

• f_3 — операция, осуществляющая выравнивание некоторого ПКА относительно однотипных соседей. При этом имитируется локальное формообразование фрагментов нитей цитоскелета и поверхности мембраны. Без такого выравнивания нити и мембрана хаотически скручивались бы в произвольную случайную структуру. Выравнивание некоторого i -го ПКА относительно двух соседей с индексами j и k происходит следующим образом:

$$x_i = (x_j + x_k)/2$$

$$y_i = (y_j + y_k)/2$$

• f_4 — операция деления ПКА мембраны. При этом происходит поиск двух ближайших соседних ПКА типа a_2 (j и k) и создаётся новый i -й ПКА типа a_2 с соответствующими координатами:

$$x_i = (x_j + x_k)/2$$

$$y_i = (y_j + y_k)/2$$

Индексный массив МСА увеличивается ($N_{(t+1)} = N_{(t)} + 1 = i$), а поля массива заполняются соответствующими значениями. Корректируются также индексы соседей для j -го и k -го ПКА, поскольку появляется новый сосед.

• f_5 — операция удаления ПКА мембраны. При этом выбранный для удаления i -й ПКА заменяется N -м. Индексный массив МСА уменьшается ($N_{(t+1)} = N_{(t)} - 1$), а поля корректируются соответственно новым значениям индексов (вместо N -го соседа — i -й). Для ПКА, у которых был i -й сосед, производится поиск нового ближайшего соседа.

• f_6 — операция деления ПКА внутренней среды (цитоплазмы) — аналогична операции f_4 , но координаты вновь созданного ПКА могут генерироваться случайным образом, как это осуществлялось для f_0 .

• f_7 — операция удаления ПКА внутренней среды (цитоплазмы) — аналогична операции f_5 .

• f_8 — операция деления ПКА нити цитоскелета при взаимодействии с центросомой — аналогична операции f_4 .

• f_9 — операция удаления терминального ПКА нити цитоскелета — аналогична операции f_5 , но при этом происходит переопределение соседнего связанного ПКА нити в состоянии терминального путём замены нулевых значений индексов соседей индексами ближайших ПКА.

На фоне описанных операций реализовано также два вероятностных процесса: зарождение новой нити цитоскелета и спонтанное её саморазрушение. При имитации зарождения и разрушения нитей выбирается некое среднее количество итерационных циклов, в течение которых происходят соответствующие процессы. Естественно, это количество должно быть пропорционально полному количеству ПКА, моделирующих

тело амебы. Коэффициент пропорциональности подбирается эмпирическим путём исходя из наблюдений за динамикой модели амебы. В нашем случае

$$c = \text{random}(1000 * N)$$

Если $c = 1$, то зарождается нить типа a_4 ; $c = 2$ — зарождается нить типа a_5 ; $c = 3$ — зарождается нить типа a_6 ; $c = 4$ — разрушается нить типа a_4 ; $c = 5$ — разрушается нить типа a_5 ; $c = 6$ — разрушается нить типа a_6 .

Процесс зарождения нити имитируется путём создания нового ПКА соответствующего типа и размещения его в окрестности центрисомы a_3 в произвольном направлении. Далее запоминаются в отдельной структуре данных направления нитей разных типов, что даёт возможность порождать нити типов a_4 и a_5 в противоположных направлениях. В таком случае амеба начинает двигаться направленно, иначе деформации мембраны приводят лишь к колебаниям на одном месте. В дальнейшем при моделировании процесса хемотаксиса (движения амебы вдоль градиента химических веществ) алгоритм выбора направления роста нитей будет соответствующим образом модифицирован.

Процесс разрушения нити имитируется путём цепного удаления всех ПКА, связанных в единую нить. Описанные процессы сопровождаются соответствующей коррекцией содержимого индексного массива МСА.

4. Результаты моделирования

Согласно сформулированным правилам клеточно-автоматных взаимодействий, разработан алгоритм, позволяющий визуализировать амебоидное движение, а также его программная реализация, предназначенная для исследования закономерностей и механизмов амебоидной локомоции. В результате программного эксперимента наблюдается хаотический рост нитей цитоскелета, который задаёт разнородную динамику движения модели. Фрагменты этого движения приведены на рис. 10.

На верхнем рисунке наблюдается форма модели при усиленном росте нитей цитоскелета, катализирующих процесс роста двух псевдоподий. Форма модели на среднем рисунке определяется разницей в длине нитей цитоскелета, катализирующих рост уropодий, и нитей цитоскелета, отвечающих за образование формы тела. На нижнем рисунке наблюдается фрагмент движения модели «Amoeba Proteus» при одновременном усиленном росте псевдоподий и замедленном росте уropодий.

Входными параметрами модели является количество ПКА (N) и соотношение радиусов ПКА, моделирующих «золь» и «гель» (r_2), и радиусов ПКА, моделирующих нити цитоскелета (r_1). При $N < 100$ процент корректной работы модели составляет $\approx 10\%$. При $N \geq 300$ результаты моделирования позволяют наблюдать амебоидную локомоцию примерно в 90% случаев. При $N > 700$ скорость вычислительного процесса стремительно снижается по причине медленного роста нитей цитоскелета и примерно в 50% случаев не выдаёт ожидаемого результата, поскольку нити исчезают раньше, чем происходит взаимодействие с мембраной. Выбор $N = 300$ сделан эмпирическим путём.

На корректность модели влияет также соотношение радиусов r_2 и r_1 . Результат моделирования корректен при $1 < r_2/r_1 < 4$. При $r_2/r_1 \geq 4$ нити цитоскелета начинают образовывать витки, что приводит к некорректной работе модели. Случай $r_2/r_1 \leq 1$ не рассматривался ввиду того, что он не естественен с точки зрения соотношения соответствующих фрагментов в живой клетке. В нашей модели выбор $r_2/r_1 = 2$ сделан эмпирическим путём.

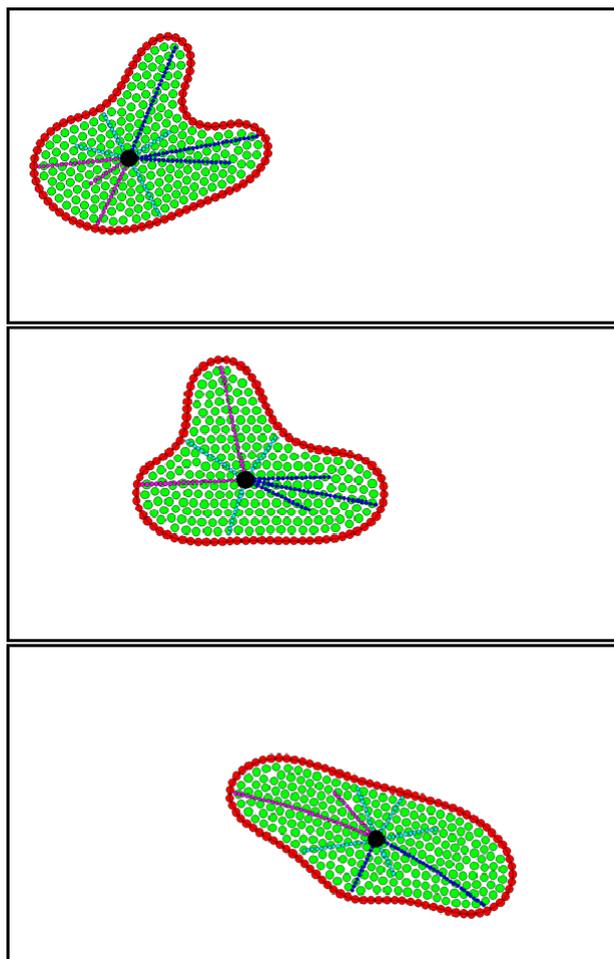


Рис. 10. Фрагменты движения модели «Amoeba Proteus»

Отметим, что число ПКА N и соотношение радиусов ПКА r_2/r_1 взаимосвязаны. При изменении одного из параметров следует повторно подбирать другой.

Заключение

В работе применён метод подвижных клеточных автоматов к моделированию амебодной подвижности, существенным преимуществом которого является возможность перехода от статической сетки к концепции соседей. В качестве объекта выбран одноклеточный биологический организм «Amoeba Proteus», рассмотрены базовые принципы локомоции, на основании которых построена модель. Найдены правила клеточно-автоматных взаимодействий согласно концепции соседства. В результате получена компьютерная модель, имитирующая амебодную локомоцию. Следует отметить, что модель является качественной, а не количественной, и позволяет продемонстрировать принципиальную возможность ПКА для моделирования амебодной локомоции.

Новизна исследований заключается в использовании существенно нового подхода к моделированию локомоции одноклеточного организма «Amoeba Proteus» — моделирования передвижения амебы на основе цитоскелетных преобразований внутри клетки. Этот подход наиболее точно описывает процесс локомоции в живой клетке.

В дальнейшем планируется моделировать как простые, так и более сложные многоклеточные организмы, собирать данные и изучать их поведение, эмбриогенез, самоорганизацию, саморепликацию, их нервную систему, мышечную систему и т. д. Конечно,

это возможно только при более подробном биологическом анализе предметной области и с развитием программной инженерии.

ЛИТЕРАТУРА

1. *De Bruyn P. P. H.* Theories of amoeboid movement // *Quarterly Rev. Biology.* 1947. V. 22. No. 1. P. 1–24.
2. *Howard J.* *Mechanics of Motor Proteins and the Cytoskeleton.* Massachusetts: Sinauer Associates, 2001. 384 p.
3. *Shih Y.-L. and Rothfield L.* The bacterial cytoskeleton // *Microbiology and Molecular Biology Rev.* 2006. V. 70. No. 3. P. 729–754.
4. *Романовский Ю. М., Степанова Н. В., Чернавский Д. С.* Математическое моделирование в биофизике. М.: Наука, 1975. 344 с.
5. *Чернавский Д. С.* Проблема происхождения жизни и мышления с точки зрения современной физики // *УФН.* 2000. Т. 170. № 2. С. 157–183. <http://ufn.ru/ru/articles/2000/2/c/>
6. *De Cerqueira Gatti MA. and de Lucena C. J. P.* *Cell Simulation: an Agent-based Software Engineering Approach.* Rio de Janeiro: Monografias em Ciencia da Computacao, 2008. No. 18/08. 17 p.
7. *Романовский Ю. М., Тепло В. А.* Физические основы клеточного движения. Механизмы самоорганизации амебодной подвижности // *УФН.* 1995. Т. 165. № 5. С. 555–578.
8. *Schlick T.* *Molecular Modeling and Simulation: An Interdisciplinary Guide.* Ed. 2. N.Y.: Springer Science and Business Media, 2010. 723 p.
9. *Nishimura S. I. and Sasai M.* Modulation of the reaction rate of regulating protein induces large morphological and motional change of amoebic cell // *J. Theor. Biol.* 2007. No. 245. P. 230–237.
10. *Nishimura S. I., Ueda M., and Sasai M.* Non-Brownian dynamics and strategy of amoeboid cell locomotion // *Phys. Rev. E.* 2012. V. 85. <http://www.biomedsearch.com/nih/Non-Brownian-dynamics-strategy-amoeboid/22680500.html>
11. *Nishimura S. I., Ueda M., and Sasai M.* Cortical factor feedback model for cellular locomotion and cytofission // *PLoS Comput Biol.* 2009. No. 5(3): e1000310.
12. *Umedachi T., Ito K., and Ishiguro A.* Soft-bodied amoeba-inspired robot that switches between qualitatively different behaviors with decentralized stiffness control // *Adaptive Behavior.* 2015. V. 23. P. 97–108.
13. *Umedachi T., Horikiri S., Kobayashi R., and Ishiguro A.* Enhancing adaptability of amoeboid robot by synergetically coupling two decentralized controllers inspired by true slime mold // *Adaptive Behavior.* 2015. V. 23. P. 109–21.
14. *Graham-Rowe D.* Amoebalike robots for search and rescue // *MIT Technology Rev.* March 29, 2007. www.technologyreview.com/s/407603/amoebalike-robots-for-search-and-rescue/
15. *Doursat R., Sayama H., and Michel O.* A review of morphogenetic engineering // *Natural Computing.* 2013. V. 12. P. 517–535.
16. *Бандман О. Л.* Клеточно-автоматные модели естественных процессов и их реализация на современных компьютерах // *Прикладная дискретная математика.* 2017. № 35. С. 102–121.
17. *Psakhie S. G., Ostermeyer G. P., Dmitriev A. I., et al.* Method of movable cellular automata as a new trend of discrete computational mechanics. I. Theoretical description // *Phys. Mesomechanics.* 2000. No. 3(2). P. 5–12.

18. *Psakhie S. G., Horie Y., Ostermeyer G. P., et al.* Movable cellular automata method for simulating materials with mesostructure // *Theor. Appl. Fracture Mechanics*. 2001. V.37. No. 1–3. P. 311–334.
19. *Shilko E. V., Psakhie S. G., Schmauder S., et al.* Overcoming the limitations of distinct element method for multiscale modeling of materials with multimodal internal structure // *Comput. Mater. Sci.* 2015. V. 102. P. 267–285.
20. *Жихаревич В. В., Газдюк К. П.* Алгоритм определения соседних элементов множества подвижных клеточных автоматов при условии фиксированного количества соседей // *Вестник Национального технического университета Харьковский политехнический институт. Сер. Информатика и моделирование*. 2015. № 33. С.75–82.

REFERENCES

1. *De Bruyn P. P. H.* Theories of amoeboid movement. *Quarterly Rev. Biology*, 1947, vol. 22, no. 1, pp. 1–24.
2. *BibAuthorHoward J.* *Mechanics of Motor Proteins and the Cytoskeleton*. Massachusetts, Sinauer Associates, 2001. 384 с.
3. *Shih Y.-L. and Rothfield L.* The bacterial cytoskeleton. *Microbiology and Molecular Biology Rev.*, 2006, vol. 70, no. 3, pp. 729–754.
4. *Romanovskiy Yu. M., Stepanova N. V., and Chernavskiy D. S.* *Matematicheskoye modelirovaniye v biofizike [Mathematical Modeling in Biophysics]*. Moscow, Nauka Publ., 1975. 344 p. (in Russian)
5. *Chernavskiy D. S.* Problema proiskhozhdeniya zhizni i myshleniya s toчки zreniya sovremennoy fiziki [The problem of origin of life and thinking from the point of view of modern physics]. *Uspekhi Fizicheskikh Nauk*, 2000, vol. 170, no. 2, pp. 157–183. <http://ufn.ru/ru/articles/2000/2/c/> (in Russian)
6. *De Cerqueira Gatti MA. and de Lucena C. J. P.* *Cell Simulation: an Agent-based Software Engineering Approach*. Rio de Janeiro, Monografias em Ciencia da Computacao, 2008, no. 18/08. 17 p.
7. *Romanovskiy Yu. M. and Teplov V. A.* Fizicheskiye osnovy kletochnogo dvizheniya. Mekhanizmy samoorganizatsii ameboidnoy podvizhnosti [The physical basis of cell movement. Mechanisms of amoeboid locomotion self-organization]. *Uspekhi Fizicheskikh Nauk*, 1995, vol. 165, no. 5, pp. 555–578. (in Russian)
8. *Schlick T.* *Molecular Modeling and Simulation: An Interdisciplinary Guide*. Ed 2. N.Y., Springer Science and Business Media, 2010. 723 p.
9. *Nishimura S. I. and Sasai M.* Modulation of the reaction rate of regulating protein induces large morphological and motional change of amoebic cell. *J. Theor. Biol.*, 2007, no. 245, pp. 230–237.
10. *Nishimura S. I., Ueda M., and Sasai M.* Non-Brownian dynamics and strategy of amoeboid cell locomotion. *Phys. Rev. E*, 2012, vol. 85. <http://www.biomedsearch.com/nih/Non-Brownian-dynamics-strategy-amoeboid/22680500.html>
11. *Nishimura S. I., Ueda M., and Sasai M.* Cortical factor feedback model for cellular locomotion and cytofission. *PLoS Comput Biol.*, 2009, no. 5(3): e1000310.
12. *Umedachi T., Ito K., and Ishiguro A.* Soft-bodied amoeba-inspired robot that switches between qualitatively different behaviors with decentralized stiffness control. *Adaptive Behavior*, 2015, vol. 23, pp. 97–108.
13. *Umedachi T., Horikiri S., Kobayashi R., and Ishiguro A.* Enhancing adaptability of amoeboid robot by synergetically coupling two decentralized controllers inspired by true slime mold. *Adaptive Behavior*, 2015, vol. 23, pp. 109–21.

14. *Graham-Rowe D.* Amoebalike robots for search and rescue. MIT Technology Rev., March 29, 2007. www.technologyreview.com/s/407603/amoebalike-robots-for-search-and-rescue/
15. *Doursat R., Sayama H., and Michel O.* A review of morphogenetic engineering. Natural Computing, 2013, vol. 12, pp. 517–535.
16. *Bandman O. L.* Kletочно-автоматные модели естественных процессов и их реализация на современных компьютерах [Cellular-Automata models of natural processes, implementation on supercomputers]. Прикладная Дискретная Математика, 2017, no. 35, pp. 102–121. (in Russian)
17. *Psakhie S. G., Ostermeyer G. P., Dmitriev A. I., et al.* Method of movable cellular automata as a new trend of discrete computational mechanics. I. Theoretical description. Phys. Mesomechanics, 2000, no. 3(2), pp. 5–12.
18. *Psakhie S. G., Horie Y., Ostermeyer G. P., et al.* Movable cellular automata method for simulating materials with mesostructure. Theor. Appl. Fracture Mechanics, 2001, vol. 37, no. 1–3, pp. 311–334.
19. *Shilko E. V., Psakhie S. G., Schmauder S., et al.* Overcoming the limitations of distinct element method for multiscale modeling of materials with multimodal internal structure. Comput. Mater. Sci., 2015, vol. 102, pp. 267–285.
20. *Zhykharevych V. V. and Hazdiuk K. P.* Алгоритм определения соседей элементов множества подвижных клеточных автоматов при условии фиксированного количества соседей [Algorithm for determining the neighboring elements of a set of movable cellular automata under the condition of a fixed number of neighbors]. Bull. National Technical University Kharkov Polytechnic Institute. Ser. Computer Science and Modeling, 2015, no. 33, pp. 75–82. (in Russian)

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

УДК 519.81

ПОСТРОЕНИЕ АГРЕГИРОВАННОГО ОТНОШЕНИЯ, МИНИМАЛЬНО УДАЛЁННОГО ОТ ЭКСПЕРТНЫХ ПРЕДПОЧТЕНИЙ

В. Н. Нефёдов, С. О. Смерчинская, Н. П. Яшина

Московский авиационный институт, г. Москва, Россия

Рассматривается задача группового выбора. Профиль индивидуальных предпочтений экспертов на множестве альтернатив может быть задан бинарными отношениями или численными оценками альтернатив. Предлагаются способы построения матриц предпочтений для различных типов экспертной информации, а также методы формирования агрегированного отношения, удовлетворяющего условию минимальности суммарного расстояния до экспертных предпочтений. Вид агрегированного отношения предпочтения зависит от выбора формулы для определения расстояния между матрицами предпочтений. Разработанная методика может быть использована и при решении многокритериальных задач.

Ключевые слова: *групповой выбор, агрегированное отношение, профиль индивидуальных предпочтений экспертов, минимальное расстояние от предпочтений, мажоритарный граф.*

DOI 10.17223/20710410/42/9

CONSTRUCTING AN AGGREGATED RELATION WITH A MINIMUM DISTANCE FROM THE EXPERT PREFERENCES

V. N. Nefedov, S. O. Smerchinskaya, N. P. Yashina

*Moscow Aviation Institute, Moscow, Russia***E-mail:** svetlana_os@mail.ru

The paper considers the problem of collective choice. Profile of experts' individual preferences on the set of alternatives can be given by binary relations or numerical evaluations of alternatives. Methods for constructing preferences matrices for various types of expert information are proposed, as well as methods for forming an aggregated relation that satisfies the condition of minimizing the total distance to expert preferences. Let expert preferences in the form of binary relations $\rho_1, \rho_2, \dots, \rho_m$ are given by the vertex adjacency matrices R^1, R^2, \dots, R^m of the corresponding digraphs. The distance between the relations is defined as the Hamming distance. We prove that the aggregate relation, which is built according to the rule of "the majority of experts", satisfies the condition of minimum distance from expert preferences. In the case, when the profile of expert preferences is given by relations of a strict order and the number of experts is odd, the aggregated relation is unique. Let the estimates of the alternatives a_1, a_2, \dots, a_n by the t -th expert be given in the form of a vector

$h^t = \langle h_1^t, h_2^t, \dots, h_n^t \rangle$ with positive real components. Then the elements of the preference matrix $R^t = \|r_{ij}^t\|$ have the form $r_{ij}^t = \frac{h_i^t}{h_i^t + h_j^t}$, if the values of estimates on the scale are maximized. We prove that the aggregated preference relation depends on the choice of the formula for determining the distance between the preferences matrices. When you specify the distance through the module of the difference of the preference matrices elements, then the total distance is minimal if all the elements of the aggregated matrix are equal to the medians of the corresponding elements of expert matrices. When you specify the distance through the square of the elements difference, then the distance is minimal if all the elements of the aggregated matrix are equal to the arithmetic means of the corresponding elements of expert matrices. The developed technique can be used to solve multi-criteria problems in the assessment of alternatives in the scales of relations.

Keywords: *collective choice, aggregate relation, profile of experts' individual preferences, minimum distance from preferences, majority graph.*

Введение

При принятии сложных решений часто используются опыт и знания группы экспертов — специалистов в данной предметной области. К групповому решению обычно предъявляется ряд требований [1–5], важнейшими из которых являются непротиворечивость [2–5] и наиболее полный учёт индивидуальных предпочтений экспертов [1, 3, 5]. Непротиворечивость обычно характеризуется отсутствием контуров в графе агрегированного отношения предпочтения; приближённость к экспертным предпочтениям — числом голосов экспертов, отданных за построенное упорядочение. При разработке алгоритмов группового выбора основная трудность состоит в поиске разумного компромисса в выполнении этих условий.

Ещё Кондорсе пытался построить групповое ранжирование, разрушив контур по слабому звену — дуге с минимальной разностью в предпочтениях экспертов [3]. Построенное ранжирование должно было удовлетворять условию максимальности суммарных голосов экспертов, соответствующих данному упорядочению. Но Янг привел пример, в котором уже для четырёх альтернатив это условие не выполнялось. Алгоритм ранжирования, основанный на поиске пути, удовлетворяющего условию максимальности голосов экспертов, был предложен также Шульце. Одним из наиболее удачных, математически обоснованных методов нахождения коллективного ранжирования является построение медианы Кемени [2]. В качестве критерия приближённости к экспертным предпочтениям он берёт суммарное расстояние между исходными ранжированиями и групповым. Однако этот алгоритм имеет экспоненциальную вычислительную сложность и даже его эвристические модификации лишь незначительно её уменьшают, при этом не строя все возможные ранжирования.

В качестве агрегированного отношения предпочтения мажоритарный граф, построенный по правилу большинства, удобен в силу того, что минимально удалён от экспертных предпочтений и выполняются условия Эрроу (универсальность, полнота, независимость, монотонность, ненавязанность, Парето) [1, 2]. Но Эрроу формулировал свою теорему для ранжирований, а мажоритарный граф с большой вероятностью не является ранжированием и может содержать противоречивые контуры [5].

В данной работе проводится анализ единственности агрегированного отношения предпочтения, построенного по правилу большинства и, следовательно, удовлетворяющего условию минимальности суммарного расстояния до экспертных предпочтений.

Профиль экспертных предпочтений может также быть задан численными оценками альтернатив или информацией о том, во сколько раз одна альтернатива предпочтительнее другой. В работе представлен способ построения матриц предпочтений в случае задания числовой экспертной информации. Предлагаются методы построения агрегированного отношения, удовлетворяющего условию минимальности суммарного расстояния до экспертных предпочтений. Вид группового решения зависит от выбора формулы для нахождения расстояния между предпочтениями. Разработанная методика может быть использована и при решении многокритериальных задач [6].

1. Агрегирование экспертных предпочтений, заданных бинарными отношениями

Рассмотрим задачу построения агрегированного отношения в случае, когда профиль экспертных предпочтений задан бинарными отношениями.

Дано множество альтернатив $A = \{a_1, a_2, \dots, a_n\}$ и множество экспертов $E = \{E_1, E_2, \dots, E_m\}$. Профиль индивидуальных предпочтений экспертов на множестве A задан бинарными отношениями предпочтения $\rho_1, \rho_2, \dots, \rho_m$. Требуется построить агрегированное отношение $\hat{\rho}$, максимально согласованное с экспертными предпочтениями $\rho_1, \rho_2, \dots, \rho_m$.

На основе агрегированного отношения предпочтения обычно также требуется ранжировать альтернативы и/или выбрать наилучшие из них. Способы выбора наилучших альтернатив подробно описаны в работах [1–4] и основываются на стандартных процедурах на графах: нахождения внутренне и внешне устойчивых подмножеств, ядра графа, разбиения графа на уровни [7].

Индивидуальные предпочтения экспертов могут быть заданы отношениями строгого порядка (асимметричное и транзитивное), квазипорядка (рефлексивное и транзитивное), а также произвольными бинарными отношениями. В частности, предпочтения задаются строгим или нестрогим ранжированием альтернатив. Строгое ранжирование соответствует отношению строгого линейного порядка, нестрогое — отношению квазипорядка, в котором все альтернативы попарно сравнимы. Будем полагать, что $\langle a_i, a_j \rangle \in \rho_t$ ($t = 1, \dots, m$; $a_i, a_j \in A$), если элемент a_i не более предпочтителен, чем элемент a_j (отношения ρ_t можно выбрать и по-другому: $\langle a_i, a_j \rangle \in \rho_t$, если элемент a_i не менее предпочтителен, чем a_j , но стандартные процедуры на графах, используемые для выбора наилучших альтернатив, удобнее реализовывать для отношения «не более предпочтителен»).

Введём понятие расстояния между бинарными отношениями. Поставим в соответствие отношению ρ орграф $G = (A, \rho)$ с множеством вершин-альтернатив A , множеством дуг ρ и матрицей смежности $R = \|r_{ij}\|$. Экспертные предпочтения $\rho_1, \rho_2, \dots, \rho_m$ будем задавать матрицами смежности R^1, R^2, \dots, R^m соответствующих орграфов.

Определение 1. Расстоянием между двумя отношениями ρ_k и ρ_t назовём величину $d(\rho_k, \rho_t)$, определяемую по формуле

$$d(\rho_k, \rho_t) = \sum_{i=1}^n \sum_{j=1}^n |r_{ij}^k - r_{ij}^t|.$$

Фактически $d(\rho_k, \rho_t)$ равно числу несовпадений элементов r_{ij}^k и r_{ij}^t матриц смежности этих отношений R^k и R^t ($i, j = 1, \dots, n$) и, следовательно, является расстоянием Хэмминга, удовлетворяющим аксиомам метрики.

Для построения отношения, наиболее полно отражающего предпочтения экспертов, необходимо, чтобы сумма расстояний между агрегированным отношением $\widehat{\rho}$ и отношениями $\rho_1, \rho_2, \dots, \rho_m$ была минимальной.

Определение 2. Суммарное расстояние $D(\widehat{\rho})$ от отношения $\widehat{\rho}$ до отношений $\rho_1, \rho_2, \dots, \rho_m$ минимально, если выполняется

$$D(\widehat{\rho}) = \sum_{t=1}^m d(\widehat{\rho}, \rho_t) \rightarrow \min.$$

Пусть экспертные предпочтения $\rho_1, \rho_2, \dots, \rho_m$ заданы матрицами смежности R^1, R^2, \dots, R^m и q — произвольное отношение с матрицей смежности $Q = \|q_{ij}\|$.

Теорема 1. Суммарное расстояние $D(q) = \sum_{t=1}^m d(q, \rho_t)$ минимально, если $q_{ij} = 1 \Leftrightarrow r_{ij}^t = 1$ не менее чем для половины экспертов ($i, j = 1, \dots, n; t \in \{1, \dots, m\}$). В случае $\sum_{t=1}^m r_{ij}^t = m/2$ (при чётном m) $D(q)$ остается минимальным и при выборе $q_{ij} = 0$.

Доказательство. Запишем

$$D(q) = \sum_{t=1}^m d(q, \rho_t) = \sum_{t=1}^m \sum_{i=1}^n \sum_{j=1}^n |q_{ij} - r_{ij}^t| = \sum_{i=1}^n \sum_{j=1}^n \sum_{t=1}^m |q_{ij} - r_{ij}^t|.$$

Проведём декомпозицию по i, j . Тогда задача сводится к нахождению минимума величин $\sum_{t=1}^m |q_{ij} - r_{ij}^t|$ ($i, j = 1, \dots, n$). Очевидно, что минимум достигается при $q_{ij} = 1 \Leftrightarrow \sum_{t=1}^m r_{ij}^t \geq m/2 \Leftrightarrow r_{ij}^t = 1$ не менее, чем для половины экспертов. При этом в случае $\sum_{t=1}^m r_{ij}^t = m/2$ (что возможно только при чётном m) минимум достигается при любом значении $q_{ij} \in \{0, 1\}$. ■

Пусть $\rho(A)$ — множество всех бинарных отношений, заданных на множестве A . Обозначим $\text{Argmin}_{q \in \rho(A)} D(q)$ множество отношений, имеющих минимальное суммарное расстояние до экспертных предпочтений.

Полученный результат совпадает с результатом, доказанным в [1] для отношения-медианы по правилу большинства ρ_Σ , допускающему равенство числа экспертов: $\rho_\Sigma \in \text{Argmin}_{q \in \rho(A)} D(q)$.

Другим примером агрегированного отношения, минимально удалённого от экспертных предпочтений, может служить медиана Кемени. В отличие от медианы по правилу большинства, процедура Кемени позволяет построить транзитивное и, следовательно, непротиворечивое агрегированное предпочтение. Медиана Кемени строится на основе экспертных ранжирований и сама является ранжированием альтернатив.

К сожалению, в силу неоднозначности построения отношения, для которого суммарное расстояние до экспертных предпочтений минимально, выбор его в качестве агрегированного не всегда целесообразен. Сложный переборный алгоритм Кемени позволяет найти лишь одно из возможных медианных ранжирований. Например, в случае,

когда отношения $\rho_1, \rho_2, q_1, q_2, q_3, q_4$ имеют матрицы смежности

$$R^1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad R^2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

$$Q_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Q_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Q_4 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

получим $D(q_1) = D(q_2) = D(q_3) = D(q_4) = 2$, т. е. все четыре отношения q_1, q_2, q_3, q_4 являются минимальными, причём два из них (q_3, q_4) — строгие ранжирования.

Из теоремы 1 следует, что неоднозначность минимального отношения происходит только в случае, когда число экспертов чётное.

Следствие 1. Пусть профиль экспертных предпочтений $\rho_1, \rho_2, \dots, \rho_m$ задан произвольными бинарными отношениями. Тогда для всех $\rho_* \in \rho(A) \setminus \{\rho_\Sigma\}$ при нечётном числе экспертов m выполняется $D(\rho_\Sigma) < D(\rho_*)$.

Из следствия 1 в силу строгого неравенства следует и обратное утверждение.

Следствие 2. Пусть суммарное расстояние $D(\rho_*)$ от экспертных предпочтений до отношения ρ_* , заданного на множестве альтернатив A , при нечётном числе экспертов является минимальным. Тогда $\rho_\Sigma = \rho_*$.

Полученные утверждения показывают, что для мажоритарного графа (графа отношения ρ_Σ [1]) выполняется условие минимальности суммарного расстояния от агрегированного отношения до экспертных предпочтений. При нечётном числе экспертов минимальное отношение единственное, что даёт основание взять его в качестве агрегированного отношения. Напомним, что отношение ρ_Σ в общем случае не транзитивно (в частности, может содержать контуры), что затрудняет выбор наилучших альтернатив, а тем более их ранжирование.

Следующие два примера демонстрируют различие в построении минимального отношения для чётного и нечётного числа экспертов.

Пример 1. Профиль предпочтений четырёх экспертов задан ранжированием альтернатив (строгими и нестрогими). Сравним агрегированное отношение, построенное по правилу большинства, с медианой Кемени.

Экспертные ранжирования представлены в следующей таблице (наилучшие альтернативы в верхней строке):

| ρ_1 | ρ_2 | ρ_3 | ρ_4 |
|----------|----------|-------------|-------------|
| a_3 | a_2 | — | — |
| a_2 | a_3 | a_1 | a_4 |
| a_4 | a_4 | $a_3 - a_4$ | a_3 |
| a_1 | a_1 | a_2 | $a_1 - a_2$ |

Построим агрегированное отношение предпочтения по правилу большинства. Матрицы смежности экспертных предпочтений имеют следующий вид:

$$R^1 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}, \quad R^2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix},$$

$$R^3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}, \quad R^4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Единицы на диагоналях матриц смежности объясняются тем, что нестрогое ранжирование соответствует отношению квазипорядка, которое является рефлексивным.

Согласно теореме 1, существует шестнадцать отношений, минимально удалённых от экспертных предпочтений (суммарное расстояние $D(\hat{\rho}) = 16$), со следующими матрицами смежности:

$$Q^{\min} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 \vee 1 & 1 & 1 & 0 \vee 1 \\ 0 & 0 & 1 & 0 \vee 1 \\ 0 & 0 \vee 1 & 1 & 1 \end{pmatrix}.$$

Среди них пять ранжирований:

$$\begin{pmatrix} a_3 \\ a_4 \\ a_1 - a_2 \end{pmatrix}, \quad \begin{pmatrix} a_3 \\ a_2 - a_4 \\ a_1 \end{pmatrix}, \quad \begin{pmatrix} a_3 - a_4 \\ a_1 - a_2 \end{pmatrix}, \quad \begin{pmatrix} a_3 \\ a_4 \\ a_2 \\ a_1 \end{pmatrix}, \quad \begin{pmatrix} a_3 \\ a_2 \\ a_4 \\ a_1 \end{pmatrix}.$$

Отношение ρ_Σ , построенное по правилу большинства, является одним из шестнадцати отношений, минимально удалённых от экспертных предпочтений. Матрица смежности этого отношения имеет вид

$$R_\Sigma = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}.$$

Здесь ρ_Σ не транзитивно и, следовательно, не является ранжированием. Построенная на основе данного профиля экспертных предпочтений медиана Кемени (ранжирование, минимально удалённое от экспертных предпочтений) [8] является одним из пяти возможных ранжирований:

$$K = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} - \begin{pmatrix} a_3 \\ a_2 \\ a_4 \\ a_1 \end{pmatrix}.$$

Пример 2. Пусть на множестве альтернатив $A = \{a_1, a_2, a_3, a_4\}$ профиль предпочтений трёх экспертов ($m = 3$) задан отношениями строгого линейного порядка (строгим ранжированием):

| ρ_1 | ρ_2 | ρ_3 |
|----------|----------|----------|
| a_1 | a_2 | a_3 |
| a_2 | a_3 | a_4 |
| a_3 | a_4 | a_1 |
| a_4 | a_1 | a_2 |

Матрицы смежности соответствующих графов имеют следующий вид:

$$R^1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \quad R^2 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad R^3 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

По следствию 1 в этом случае существует одно отношение, минимально удалённое от экспертных предпочтений ($D = 10$), и оно совпадает с ρ_Σ :

$$R_\Sigma = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

Мажоритарный граф $G = (A, \rho_\Sigma)$ изображен на рис. 1. Соответствующее ему отношение ρ_Σ является минимально удалённым от экспертных предпочтений и при этом единственным. Граф отношения ρ_Σ содержит контуры, что затрудняет выбор наилучших альтернатив. Действительно, граф имеет три доминирующих (внешне устойчивых) подмножества: $\{a_1, a_2\}$, $\{a_2, a_3\}$, $\{a_2, a_4\}$ [7]. Для нахождения наилучших альтернатив в этом случае можно воспользоваться процедурой Коупленда [2]: каждой вершине графа ставится в соответствие число, равное разности количеств входящих и исходящих из неё дуг (сумма единиц соответствующего столбца минус сумма единиц строки матрицы смежности). Согласно этому алгоритму, получим следующие индексы альтернатив: $a_1(-1)$, $a_2(1)$, $a_3(1)$, $a_4(-1)$. Наилучшие альтернативы, имеющие максимальный индекс, — a_2 , a_3 .

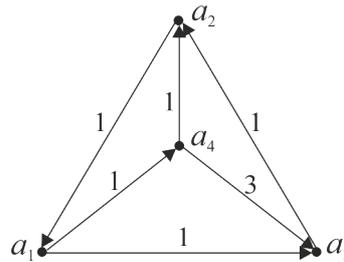


Рис. 1

Дополнительной характеристикой мажоритарного графа являются веса на дугах, равные разности чисел экспертов, для которых альтернатива a_i не более предпочтительна a_j и a_j не более предпочтительна a_i . Эту разность легко найти, введя матрицу суммарных предпочтений $P = \|p_{ij}\|$, где $p_{ij} = \sum_{t=1}^m r_{ij}^t$. Тогда вес дуги $\langle a_i, a_j \rangle \in \rho_\Sigma$ равен $p_{ij} - p_{ji}$.

Веса на дугах мажоритарного графа (рис. 1) вычислены с помощью матрицы $P = R^1 + R^2 + R^3$:

$$P = \begin{pmatrix} 0 & 1 & 2 & 2 \\ 2 & 0 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 2 & 3 & 0 \end{pmatrix}.$$

Модифицируем алгоритм Коупленда с учётом весов на дугах, поставив каждой альтернативе a_i в соответствие индекс равный $\sum_{j=1}^n (p_{ji} - p_{ij})$. Получим $a_1(-1)$, $a_2(1)$, $a_3(3)$, $a_4(-3)$. В этом случае имеем строгое ранжирование альтернатив. Наилучшей оказалась альтернатива a_3 .

В качестве индекса альтернативы можно взять также величину, равную не разности, а отношению сумм элементов столбца и строки матрицы суммарных предпочтений P , соответствующих данной альтернативе. Если сумма элементов строки равна

нулю, то альтернатива является наилучшей и ей надо поставить в соответствие индекс, заведомо больший, чем у альтернатив с ненулевой суммой строки. Этим методом (его естественно назвать методом «отношения сумм») для рассматриваемого примера получим следующие индексы альтернатив: $a_1(1/2)$, $a_2(2)$, $a_3(4)$, $a_4(1/4)$. Полученное таким образом ранжирование совпадает с результатом применения модифицированной процедуры Коупленда.

Легко убедиться, что у всех построенных ранжирований суммарное расстояние до экспертных предпочтений больше минимального.

С помощью процедур нахождения ранжирования альтернатив на основе мажоритарного графа, содержащего контуры, не всегда возможно получить агрегированное отношение, минимально удалённое от экспертных предпочтений. В частности, если профиль экспертных предпочтений — строгие ранжирования и число экспертов нечётно, то единственное отношение с минимальным суммарным расстоянием соответствует мажоритарному графу. При этом, если граф содержит контуры, он не соответствует строгому ранжированию альтернатив. Тогда можно рекомендовать лицу, принимающему решения, выбирать ранжирование с суммарным расстоянием, ближайшим к минимальному.

2. Агрегирование предпочтений, заданных численными оценками альтернатив

Рассмотрим задачу нахождения минимального агрегированного отношения в случае, когда профиль индивидуальных предпочтений экспертов задан численными оценками альтернатив. Предложим способ формирования матриц экспертных предпочтений и алгоритм построения минимального агрегированного предпочтения.

Пусть предпочтения экспертов на множестве альтернатив $A = \{a_1, a_2, \dots, a_n\}$ заданы векторами h^1, h^2, \dots, h^m , где $h^t = \langle h_1^t, h_2^t, \dots, h_n^t \rangle$ — вектор с компонентами $h_i^t \in \mathbb{R}^+$, равными численным оценкам альтернатив ($i \in \{1, \dots, n\}$, $t \in \{1, \dots, m\}$). Предполагается, что эксперты могут оценивать альтернативы в разных шкалах, что затрудняет возможность построения суммарной оценки для каждой альтернативы.

Матрицы экспертных предпочтений R^1, R^2, \dots, R^m построим на основе векторов h^1, h^2, \dots, h^m следующим образом: $R^t = \|r_{ij}^t\|$ — квадратная матрица порядка n (n — число альтернатив, m — число экспертов) с элементами

$$r_{ij}^t = \begin{cases} \frac{h_j^t}{h_i^t + h_j^t}, & \text{если значения оценок } t\text{-го эксперта максимизируются,} \\ \frac{h_i^t}{h_i^t + h_j^t}, & \text{если значения оценок } t\text{-го эксперта минимизируются,} \end{cases} \quad (i \neq j);$$

$$r_{ii}^t = 1; \quad i, j = 1, \dots, n, \quad t = 1, \dots, m.$$

Заметим, что для элементов матрицы предпочтений R^t выполняется:

- 1) $\frac{r_{ij}^t}{r_{ji}^t} = \begin{cases} h_j^t/h_i^t, & \text{если значения оценок } t\text{-го эксперта максимизируются,} \\ h_i^t/h_j^t, & \text{если значения оценок } t\text{-го эксперта минимизируются} \end{cases}$
(сохраняется информация о том, во сколько раз альтернатива a_j предпочтительнее альтернативы a_i);
- 2) $r_{ij}^t + r_{ji}^t = 1$ ($i, j = 1, \dots, n, i \neq j$), что фактически заменяет процедуру приведения шкал экспертных предпочтений к однородным.

Для формирования матриц экспертных предпочтений необязательно задавать численные оценки альтернатив. Достаточно задать информацию о том, во сколько раз

одна альтернатива предпочтительнее другой. Если альтернатива a_i предпочтительнее a_j в α_{ij} раз, то элементы r_{ij}^t и r_{ji}^t ($i \neq j$) матрицы предпочтений R^t вычисляются следующим образом:

$$r_{ij}^t = \frac{1}{1 + \alpha_{ij}}, \quad r_{ji}^t = \frac{\alpha_{ij}}{1 + \alpha_{ij}}.$$

Следует отметить, что для получения полной и непротиворечивой информации достаточно сравнить между собой $n-1$ пару альтернатив, например, a_1 со всеми остальными альтернативами.

Найдём отношение q , имеющее минимальное суммарное расстояние до экспертных предпочтений $\rho_1, \rho_2, \dots, \rho_m$, заданных матрицами R^1, R^2, \dots, R^m , для элементов которых выполняются условия $r_{ij}^t \in [0; 1]$ и $r_{ji}^t = 1 - r_{ij}^t$ ($t = 1, \dots, m$). Пусть q — бинарное отношение с матрицей предпочтения $Q = \|q_{ij}\|$, $q_{ij} \in [0, 1]$. Суммарное расстояние до экспертных предпочтений вычисляется по формуле

$$D(q) = \sum_{t=1}^m d(q, \rho_t) = \sum_{t=1}^m \sum_{i=1}^n \sum_{j=1}^n |q_{ij} - r_{ij}^t| = \sum_{i=1}^n \sum_{j=1}^n \sum_{t=1}^m |q_{ij} - r_{ij}^t|. \quad (1)$$

Проведём декомпозицию по i, j . Тогда задача сводится к нахождению минимума величин $\sum_{t=1}^m |q_{ij} - r_{ij}^t|$ ($i, j = 1, \dots, n$). Будем для простоты обозначений считать, что для данных i, j выполняется $r_{ij}^{t_1} \leq r_{ij}^{t_2} \leq \dots \leq r_{ij}^{t_m}$. Рассмотрим случай, когда число экспертов нечётно. Покажем, что если $m = 2k + 1$, минимум достигается при $q_{ij} = r_{ij}^{t_{k+1}}$. Найдём сумму модулей

$$\begin{aligned} & |q_{ij} - r_{ij}^{t_1}| + |q_{ij} - r_{ij}^{t_2}| + \dots + |q_{ij} - r_{ij}^{t_{2k}}| + |q_{ij} - r_{ij}^{t_{2k+1}}| = \\ & = \left(|q_{ij} - r_{ij}^{t_1}| + |q_{ij} - r_{ij}^{t_{2k+1}}| \right) + \dots + \left(|q_{ij} - r_{ij}^{t_k}| + |q_{ij} - r_{ij}^{t_{k+2}}| \right) + |q_{ij} - r_{ij}^{t_{k+1}}| \geq \\ & \geq |r_{ij}^{t_1} - r_{ij}^{t_{2k+1}}| + \dots + |r_{ij}^{t_k} - r_{ij}^{t_{k+2}}| + |q_{ij} - r_{ij}^{t_{k+1}}|. \end{aligned}$$

Последнее соотношение следует из неравенства треугольников, причём величина $q_{ij} = r_{ij}^{t_{k+1}}$ — медиана, она находится внутри каждого из полученных интервалов, следовательно, при $q_{ij} = r_{ij}^{t_{k+1}}$ неравенство обращается в равенство и достигается наименьшее значение искомой суммы. В противном случае получим строгое неравенство.

Так как по определению матриц предпочтения $r_{ji}^t = 1 - r_{ij}^t$, элементы r_{ji}^t упорядочиваются в обратном порядке, а наименьшее значение достигается при $q_{ij} = 1 - r_{ij}^{t_{k+1}}$. Таким образом, для элементов матрицы Q , как и для матриц экспертных предпочтений, выполняется $q_{ij} + q_{ji} = 1$.

Аналогично можно доказать, что для чётного числа экспертов $m = 2k$ наибольшее значение сумма модулей принимает при любом значении $q_{ij} \in [r_{ij}^{t_k}, r_{ij}^{t_{k+1}}]$, т. е. матрица предпочтений Q отношения q , доставляющего минимальное значение величины $D(q)$, в этом случае строится неоднозначно. Очевидно, что условие $q_{ij} + q_{ji} = 1$ может не выполняться.

Из приведённых рассуждений следует справедливость следующей теоремы.

Теорема 2. При нечётном числе экспертов суммарное расстояние $D(q)$, заданное по формуле (1), минимально, если все элементы q_{ij} матрицы предпочтений Q равны медиане соответствующих элементов матриц экспертных предпочтений $r_{ij}^1, \dots, r_{ij}^m$ ($i, j = 1, \dots, n$).

Отношение ρ_Σ , соответствующее матрице суммарных предпочтений $P = \sum_{k=1}^m R^k$ и, следовательно, матрице $\frac{1}{m}P$, не совпадает с минимальным отношением q из теоремы 2. Элементы матрицы $\frac{1}{m}P$ равны средним арифметическим значениям элементов матриц экспертных предпочтений, а матрицы Q — медианным. Как известно, значение медианы заданных чисел в общем случае не совпадает с их средним арифметическим.

Покажем, что матрица предпочтений минимального суммарного отношения будет равна средним арифметическим значениям соответствующих элементов матриц экспертных предпочтений, если расстояние между отношениями задать по формуле

$$d(\rho_k, \rho_t) = d(R^k, R^t) = \sum_{i=1}^n \sum_{j=1}^n (r_{ij}^k - r_{ij}^t)^2.$$

Тогда минимальное суммарное расстояние до экспертных предпочтений

$$D(q) = \sum_{t=1}^m d(q, \rho_t) = \sum_{t=1}^m \sum_{i=1}^n \sum_{j=1}^n (q_{ij} - r_{ij}^t)^2 = \sum_{i=1}^n \sum_{j=1}^n \sum_{t=1}^m (q_{ij} - r_{ij}^t)^2 \quad (2)$$

достигается для матрицы Q со средними арифметическими значениями элементов $r_{ij}^1, \dots, r_{ij}^m$ ($i, j = 1, \dots, n$). Действительно, минимум сильно выпуклой функции

$$f(q_{ij}) = (q_{ij} - r_{ij}^1)^2 + \dots + (q_{ij} - r_{ij}^m)^2$$

достигается при $q_{ij} = (r_{ij}^1 + \dots + r_{ij}^m)/m$, так как в этом (и только в этом) случае выполняется

$$f'(q_{ij}) = 2(q_{ij} - r_{ij}^1) + \dots + 2(q_{ij} - r_{ij}^m) = 0.$$

Заметим, что для булевых матриц

$$\sum_{i=1}^n \sum_{j=1}^n (r_{ij}^k - r_{ij}^t)^2 = \sum_{i=1}^n \sum_{j=1}^n |r_{ij}^k - r_{ij}^t|.$$

Теорема 3. Суммарное расстояние $D(q)$, заданное по формуле (2), минимально, если все элементы q_{ij} матрицы предпочтений Q равны среднему арифметическому значению соответствующих элементов матриц экспертных предпочтений $r_{ij}^1, \dots, r_{ij}^m$, т. е. при $q_{ij} = (r_{ij}^1 + \dots + r_{ij}^m)/m$ ($i, j = 1, \dots, n$).

Матрица средних арифметических значений соответствующих элементов R^1, R^2, \dots, R^m вычисляется через матрицу суммарных предпочтений $P = \sum_{t=1}^m R^t$ и равна $\frac{1}{m}P$.

В случае, когда заданы весовые коэффициенты k_1, k_2, \dots, k_m , оценивающие компетентность экспертов [9], матрицы экспертных предпочтений фактически примут вид $k_1 R^1, \dots, k_m R^m$. Тогда элементы q_{ij} матрицы Q будут в соответствии с введённым расстоянием равны медианному или среднему арифметическому значению элементов $k_1 r_{ij}^1, \dots, k_m r_{ij}^m$ ($i, j = 1, \dots, n$).

Рассмотрим пример, в котором агрегированные отношения строятся на основе медианных и средних арифметических значений элементов матриц экспертных предпочтений.

Пример 3. Пусть множество альтернатив $A = \{a_1, a_2, a_3, a_4\}$. Профиль предпочтений трёх экспертов задан численными оценками альтернатив (значения по шкалам экспертных оценок максимизируются):

| | h^1 | h^2 | h^3 |
|-------|-------|-------|-------|
| a_1 | 3 | 4 | 2 |
| a_2 | 2 | 3 | 4 |
| a_3 | 3 | 3 | 2 |
| a_4 | 2 | 4 | 4 |

Матрицы парных сравнений альтернатив для каждого эксперта соответственно имеют вид

$$R^1 = \begin{pmatrix} 1 & 2/5 & 1/2 & 2/5 \\ 3/5 & 1 & 3/5 & 1/2 \\ 1/2 & 2/5 & 1 & 2/5 \\ 3/5 & 1/2 & 3/5 & 1 \end{pmatrix}, R^2 = \begin{pmatrix} 1 & 3/7 & 3/7 & 1/2 \\ 4/7 & 1 & 1/2 & 4/7 \\ 4/7 & 1/2 & 1 & 4/7 \\ 1/2 & 3/7 & 3/7 & 1 \end{pmatrix}, R^3 = \begin{pmatrix} 1 & 2/3 & 1/2 & 2/3 \\ 1/3 & 1 & 1/3 & 1/2 \\ 1/2 & 2/3 & 1 & 2/3 \\ 1/3 & 1/2 & 1/3 & 1 \end{pmatrix}.$$

Вычислим матрицу суммарных предпочтений P :

$$P = \begin{pmatrix} 3 & 157/105 & 10/7 & 47/30 \\ 158/105 & 3 & 43/30 & 11/7 \\ 11/7 & 47/30 & 3 & 172/105 \\ 43/30 & 10/7 & 143/105 & 3 \end{pmatrix}.$$

Агрегированное предпочтение строим на основе матрицы, содержащей средние арифметические значения элементов матриц экспертных предпочтений и равной $\frac{1}{4}P$. Матрица смежности соответствующего мажоритарного графа имеет следующий вид:

$$R_\Sigma = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Граф отношения ρ_Σ изображён на рис. 2 (для лучшего восприятия петли на рисунке отсутствуют).

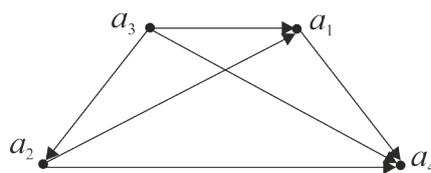


Рис. 2

Граф (без учёта петель) не содержит контуров — его можно разбить на уровни с помощью алгоритма Демукрона [7] (рис. 3).

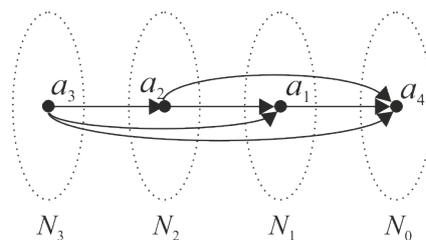


Рис. 3

Получим ранжирование всех альтернатив, начиная с наилучшей: $a_4 - a_1 - a_2 - a_3$. Процедуры Коупленда (простая и модифицированная) дают то же самое упорядочение альтернатив.

Построим минимальное отношение на основе медианных значений соответствующих элементов матриц R^1, R^2, R^3 . Матрица Q медианных значений имеет вид

$$Q = \begin{pmatrix} 1 & 3/7 & 1/2 & 1/2 \\ 4/7 & 1 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1 & 4/7 \\ 1/2 & 1/2 & 3/7 & 1 \end{pmatrix}.$$

Получим матрицу смежности $R(Q)$ агрегированного предпочтения:

$$R(Q) = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}.$$

Соответствующий граф изображён на рис. 4. Граф содержит контуры. Для выбора наилучших альтернатив можно воспользоваться процедурой Коупленда. Получим $a_1(1), a_2(-1), a_3(-1), a_4(1)$. Наилучшие альтернативы по разности входящих в вершину и исходящих из неё дуг a_1 и a_4 . Модифицированная процедура Коупленда с учётом весов на дугах и процедура отношения весов дуг дают тот же результат. Индексы альтернативы a_i ($i = 1, \dots, 4$) в процедуре отношения весов дуг можно вычислить, разделив сумму элементов i -го столбца на сумму элементов i -й строки матрицы Q (без учёта диагональных элементов, не несущих в данном случае какой-либо существенной информации). Отношения полученных индексов показывают (в каком-то смысле), во сколько раз одна альтернатива более (или менее) предпочтительнее другой.

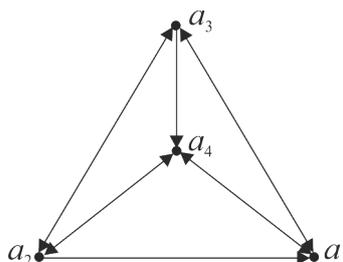


Рис. 4

Агрегированные отношения, построенные для средних арифметических и для медианных значений элементов матриц экспертных предпочтений, не совпадают, но наиболее предпочтительные альтернативы и в том и в другом случае — a_1 и a_4 .

Заключение

Разработаны алгоритмы построения агрегированного отношения предпочтения, удовлетворяющего требованию минимальности суммарного расстояния до экспертных предпочтений. Профиль экспертных предпочтений может быть задан бинарными отношениями на множестве альтернатив или численными оценками альтернатив. Предложен способ построения матриц предпочтения в случае задания численных оценок альтернатив или информации о том, во сколько раз одна альтернатива превосходит

другую. Вид агрегированного отношения, минимально удалённого от экспертных предпочтений, зависит от типа экспертной информации и формы её представления, а также от выбора формулы для нахождения расстояния между матрицами предпочтений. Проведён сравнительный анализ предложенных алгоритмов, который позволит лицу, принимающему решение, найти разумный компромисс между непротиворечивостью агрегированного отношения и максимальным учётом экспертной информации.

ЛИТЕРАТУРА

1. *Миркин Б. Г.* Проблема группового выбора. М.: Наука, 1974. 256 с.
2. *Петровский А. Б.* Теория принятия решений. М.: Академия, 2009. 400 с.
3. *Мулен Э.* Кооперативное принятие решений: Аксиомы и модели. М.: Мир, 1991. 464 с.
4. *Осипова В. А., Подиновский В. В., Яшина Н. П.* О непротиворечивом расширении отношений предпочтения в задачах принятия решений // Журнал вычисл. матем. и матем. физики. 1984. № 6. С. 831–839.
5. *Нефёдов В. Н., Осипова В. А., Смерчинская С. О., Яшина Н. П.* Непротиворечивое агрегирование отношений строгого порядка // Изв. вузов. Математика. 2018. № 5. С. 71–85.
6. *Smerchinskaya S. O. and Yashina N. P.* On an algorithm for pairwise comparison of alternatives in multi-criteria problems // Intern. J. Modeling, Simulation, and Scientific Computing. 2018. V. 9. No. 1.
7. *Нефёдов В. Н., Осипова В. А.* Курс дискретной математики. М.: Изд-во МАИ, 1992. 262 с.
8. <http://helpiks.org/3-61482.html> — Электронный учебник по теории принятия решений.
9. *Смерчинская С. О., Яшина Н. П.* Анализ компетентности экспертов в задачах группового выбора // Информационные и телекоммуникационные технологии. 2012. № 15. С. 103–115.

REFERENCES

1. *Mirkin B. G.* Group Choice. V. H. Winston & Sons Publ., 1979. 252 p.
2. *Petrovskiy A. B.* Teoriya prinyatiya resheniy [Decision Making Theory]. Moscow, Akademiya Publ., 2009. 400 p. (in Russian)
3. *Moulin H.* Axioms of Cooperative Decision Making. Cambridge, Cambridge University Press, 1988.
4. *Osipova V. A., Podinovski V. V. and Yashina N. P.* On non-contradictory extension of preference relations in decision making problems. USSR Comput. Math. and Math. Physics, 1984, vol. 24, pp. 128–134.
5. *Nefyodov V. N., Osipova V. A., Smerchinskaya S. O., and Yashina N. P.* Non-contradictory aggregation of strict order relations. Russian Mathematics, 2018, vol. 62, pp. 61–73.
6. *Smerchinskaya S. O. and Yashina N. P.* On an algorithm for pairwise comparison of alternatives in multi-criteria problems. Intern. J. of Modeling, Simulation, and Scientific Computing, 2018, vol. 9, no. 1.
7. *Nefedov V. N. and Osipova V. A.* Kurs diskretnoy matematiki [Discrete Mathematics Course]. Moscow, MAI Publ., 1992. 262 p. (in Russian)
8. <http://helpiks.org/3-61482.html> — Electronic textbook on decision making theory.
9. *Smerchinskaya S. O. and Yashina N. P.* Analiz kompetentnosti ekspertov v zadachakh gruppovogo vybora [Analysis of the competence of experts in the problems of group selection]. Inform. Telecomm. Technologies, 2012, no. 15, pp. 103–115. (in Russian)

СВЕДЕНИЯ ОБ АВТОРАХ

АГИБАЛОВ Геннадий Петрович — доктор технических наук, профессор, главный научный сотрудник лаборатории компьютерной криптографии Национального исследовательского Томского государственного университета, г. Томск.

E-mail: agibalov@mail.tsu.ru

БОРОВКОВА Ирина Вячеславовна — студентка кафедры защиты информации и криптографии Национального исследовательского Томского государственного университета, г. Томск. E-mail: iborovkova95@gmail.com

ГАЗДЮК Екатерина Петровна — аспирантка Черновицкого национального университета имени Юрия Федыковича, г. Черновцы.

E-mail: kateryna.gazdyik@gmail.com

ЖИХАРЕВИЧ Владимир Викторович — кандидат физико-математических наук, доцент, доцент Черновицкого национального университета имени Юрия Федыковича, г. Черновцы. E-mail: vzhikhar81@gmail.com

ИЛЬЕВ Артем Викторович — младший научный сотрудник Омского государственного технического университета, младший научный сотрудник Института математики им. С. Л. Соболева СО РАН, г. Омск. E-mail: artyom_iljev@mail.ru

ИЛЬЕВ Виктор Петрович — доктор физико-математических наук, профессор, главный научный сотрудник Омского государственного технического университета, профессор Омского государственного университета им. Ф. М. Достоевского, г. Омск.

E-mail: iljev@mail.ru

КЛЮЧАРЁВ Петр Георгиевич — кандидат технических наук, доцент кафедры «Информационная безопасность» Московского государственного технического университета им. Н. Э. Баумана, г. Москва. E-mail: pk.iu8@yandex.ru

КУЗНЕЦОВ Александр Алексеевич — доктор физико-математических наук, профессор, директор Института космических исследований и высоких технологий Сибирского государственного университета науки и технологий имени академика М. Ф. Решетнева, г. Красноярск. E-mail: alex_kuznetsov80@mail.ru

КУЗНЕЦОВА Александра Сергеевна — кандидат физико-математических наук, доцент кафедры информационных технологий и математического обеспечения информационных систем Красноярского государственного аграрного университета, г. Красноярск. E-mail: alexakuznetsova85@gmail.com

МИРОНКИН Владимир Олегович — старший преподаватель кафедры компьютерной безопасности Московского института электроники и математики им. А. Н. Тихонова Национального исследовательского университета «Высшая школа экономики», г. Москва. E-mail: mironkin.v@mail.ru

НЕФЕДОВ Виктор Николаевич — кандидат физико-математических наук, доцент, доцент Московского авиационного института, г. Москва.

E-mail: nefedovvn54@yandex.ru

НИКИТИНА Ольга Михайловна — кандидат физико-математических наук, доцент Черновицкого факультета Национального технического университета «Харьковский политехнический институт», г. Черновцы. E-mail: o.nikitina.chv@gmail.com

ОСТАПОВ Сергей Эдуардович — доктор физико-математических наук, профессор, заведующий кафедрой программного обеспечения компьютерных систем Черновицкого национального университета имени Юрия Федьковича, г. Черновцы. E-mail: sergey.ostapov@gmail.com

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

СЕМЕНОВА Екатерина Вадимовна — студентка кафедры защиты информации и криптографии Национального исследовательского Томского государственного университета, г. Томск. E-mail: katrinevs@mail.ru

СМЕРЧИНСКАЯ Светлана Олеговна — старший преподаватель Московского авиационного института, г. Москва. E-mail: svetlana_os@mail.ru

ЧЕРЕДНИК Игорь Владимирович — преподаватель Российского технологического университета (МИРЭА), г. Москва. E-mail: p.n.v.k.s@mail.ru

ЯШИНА Нина Павловна — кандидат физико-математических наук, доцент, доцент Московского авиационного института, г. Москва. E-mail: nina_p_yashina@mail.ru