



№4-2020

ISSN 1999-9429

ИЗВЕСТИЯ ЮФУ

ТЕХНИЧЕСКИЕ НАУКИ

- Искусственный интеллект и нечеткие системы
- Автоматизация проектирования
- Системы управления и нелинейная динамика
- Информационная безопасность

ИЗВЕСТИЯ ЮФУ. ТЕХНИЧЕСКИЕ НАУКИ IZVESTIYA SFedU. ENGINEERING SCIENCES

Свидетельство о регистрации средства массовой информации
ПИ № ФС77-28889 от 12.07.2007

Научно-технический и прикладной журнал
Издается с 1995 года, до середины 2007 года под названием «Известия ТРТУ»
Подписной индекс 41970

№ 4 (214). 2020 г.

Тематический выпуск

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Журнал включен в «Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук».

Редакционный совет

Каляев И.А. (председатель); Курейчик В.В. (зам. председателя); Курейчик В.М. (зам. председателя); Бородинский И.М. (ученый секретарь); Абрамов С.М.; Агеев О.А.; Бабенко Л.К.; Веселов Г.Е.; Гонкальвес Ж.; Колесников А.А.; Коноплев Б.Г.; Левин И.И.; Макаревич О.Б.; Маркович И.И.; Микрин Е.А.; Никитов С.А.; Обуховец В.А.; Осипов Г.С.; Панатов Г.С.; Панич А.Е.; Петров В.В.; Петровский А.Б.; Пшихопов В.Х.; Редько В.Г.; Румянцев К.Е.; Саламах М.; Солдатов А.В.; Стемповский А.Л.; Сухинов А.И.; Сысоев В.В.; Тарасов С.П.; Фрадков А.Л.; Хашемипур М.; Чаплыгин Ю.А.; Чередниченко Д.И.; Четверушкин Б.Н.; Чичков Б.Н.

Учредитель Южный федеральный университет.

Издатель Южный федеральный университет.

Ответственный за выпуск Курейчик В.В.

Технический редактор Ярошевич Н.В.

Оригинал-макет выполнен Ярошевич Н.В.

Подписано к печати 26.11.2020. Формат 70×108 $\frac{1}{16}$. Бумага офсетная.

Офсетная печать. Усл. печ. л. – 20,1. Уч.-изд. л. – 17,9.

Заказ № 7779. Тираж 250 экз.

Адрес издателя: 344091, г. Ростов-на-Дону, пр. Стачки, 200/1. Тел. 8(863)2478051.

Адрес типографии: Отпечатано в отделе полиграфической, корпоративной и сувенирной продукции Издательско-полиграфического комплекса КИБИ МЕДИА ЦЕНТРА ЮФУ. 344090, г. Ростов-на-Дону, пр. Стачки, 200/1, тел. 8 (863) 247-80-51.

Адрес редколлегии: 347922, г. Таганрог, ул. Чехова, 22, ЮФУ, тел. +7 (928) 909-57-82, e-mail: iborodyanskiy@sfedu.ru, <http://izv-tn.tti.sfedu.ru/>.

16+

Цена свободная

ISSN 1999-9429 (Print)

ISSN 2311-3103 (Online)

© Южный федеральный университет, 2020

СОДЕРЖАНИЕ

РАЗДЕЛ I. ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И НЕЧЕТКИЕ СИСТЕМЫ

В.В. Курейчик, В.В. Бова, Ю.А. Кравченко МЕТОД ПОИСКА ПОСЛЕДОВАТЕЛЬНЫХ ПАТТЕРНОВ ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЕЙ В ИНТЕРНЕТ-ПРОСТРАНСТВЕ	6
Е.М. Герасименко РЕШЕНИЕ ЗАДАЧИ МОДЕЛИРОВАНИЯ ЧАСТИЧНО РЕВЕРСИВНОГО ПОТОКА МИНИМАЛЬНОЙ СТОИМОСТИ В НЕЧЕТКИХ УСЛОВИЯХ	21
В.А. Деркачев, В.В. Бахчевников, А.Н. Бакуменко КЛАССИФИКАТОР ИЗОБРАЖЕНИЙ СЕМЯН СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ.....	32
Е.М. Герасименко, В.В. Стеценко ИНТЕЛЛЕКТУАЛЬНЫЙ МЕТОД ИЗВЛЕЧЕНИЯ ЗНАНИЙ НА ОСНОВЕ ОПРЕДЕЛЕНИЯ ТОНАЛЬНОСТИ ОТЗЫВОВ	39
Л.А. Гладков, Н.В. Гладкова ЭВОЛЮЦИОНИРУЮЩИЕ МНОГОАГЕНТНЫЕ СИСТЕМЫ И ЭВОЛЮЦИОННОЕ ПРОЕКТИРОВАНИЕ	48
А.Н. Каркищенко, В.Б. Мнухин МЕТОД ДЕТЕКЦИИ ХАРАКТЕРНЫХ ТОЧЕК ИЗОБРАЖЕНИЯ С ПОМОЩЬЮ ЗНАКОВОГО ПРЕДСТАВЛЕНИЯ	59
В.М. Курейчик, И.Б. Сафроненкова ОНТОЛОГИЧЕСКИЙ ПОДХОД К РЕАЛИЗАЦИИ ТЕХНОЛОГИЙ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ В СЕТИ ИНТЕРНЕТ	71
Б.К. Лебедев, В.Б. Лебедев, О.Б. Лебедев ПОПУЛЯЦИОННЫЙ АЛГОРИТМ ПОСТРОЕНИЯ ДЕРЕВА РЕШЕНИЙ МЕТОДОМ КРИСТАЛЛИЗАЦИИ РОССЫПИ АЛЬТЕРНАТИВ	82
А.А. Сорокин, И.М. Бородянский, А.В. Дагаев СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ВОССТАНОВЛЕНИЯ ПРОПУЩЕННЫХ ДАННЫХ	93
Г.Ф. Филаретов, З. Бучаала НЕПАРАМЕТРИЧЕСКИЙ МЕТОД ОБНАРУЖЕНИЯ РАЗЛАДКИ ВРЕМЕННОГО РЯДА С ИСПОЛЬЗОВАНИЕМ МЕХАНИЗМА СЛУЧАЙНЫХ БЛУЖДАНИЙ.....	107

РАЗДЕЛ II. АВТОМАТИЗАЦИЯ ПРОЕКТИРОВАНИЯ

Н.О. Васильев, М.А. Заплетина, Г.А. Иванова, А.Н. Щелоков ЛОГИЧЕСКИЙ РЕСИНТЕЗ КОМБИНАЦИОННЫХ СХЕМ ДЛЯ ПОВЫШЕНИЯ СБОЕУСТОЙЧИВОСТИ	118
Б.К. Лебедев, В.Б. Лебедев, О.Б. Лебедев ПОИСКОВЫЙ ПОПУЛЯЦИОННЫЙ АЛГОРИТМ РАЗМЕЩЕНИЯ ЭЛЕМЕНТОВ СБИС.....	125
Н.О. Васильев, П.И. Фролова, Г.А. Иванова, А.Н. Щелоков МЕТОДЫ ЛОГИЧЕСКОГО РЕСИНТЕЗА ДЛЯ ТОПОЛОГИЧЕСКОГО ПРОЕКТИРОВАНИЯ МИКРОЭЛЕКТРОННЫХ СХЕМ	137
Л.А. Гладков, Н.В. Гладкова, М.Д. Ясир ГИБРИДНЫЙ ПОДХОД К СОВМЕСТНОМУ РЕШЕНИЮ ЗАДАЧ РАЗМЕЩЕНИЯ И ТРАССИРОВКИ	144
В.И. Данильченко, Е.В. Данильченко, В.М. Курейчик АВТОМАТИЗИРОВАННЫЙ СТРУКТУРНО-ПАРАМЕТРИЧЕСКИЙ СИНТЕЗ СТУПЕНЧАТОГО НАПРАВЛЕННОГО ОТВЕТВИТЕЛЯ НА СВЯЗАННЫХ ЛИНИЯХ НА ОСНОВЕ ГЕНЕТИЧЕСКОГО АЛГОРИТМА	157
Д.В. Тельпухов, Т.Д. Жукова, А.Н. Щелоков АНАЛИЗ ХАРАКТЕРИСТИК СФК НА ОСНОВЕ МЕТОДОВ ИЗБЫТОЧНОГО КОДИРОВАНИЯ	165

РАЗДЕЛ III. СИСТЕМЫ УПРАВЛЕНИЯ И НЕЛИНЕЙНАЯ ДИНАМИКА

А.Г. Клово, Г.В. Куповых, И.А. Ляпунова МАТЕМАТИЧЕСКАЯ ЗАДАЧА ОБ ОПТИМАЛЬНОМ УПРАВЛЕНИИ СТРУНОЙ.....	178
А.С. Черепанцев СТАТИСТИЧЕСКИЕ МЕТОДЫ ОЦЕНКИ СВЯЗНОСТИ ДИНАМИЧЕСКИХ СИСТЕМ ПО ОТДЕЛЬНОЙ ВРЕМЕННОЙ ПРОЕКЦИИ	192
Е.Н. Обухова ПРИМЕНЕНИЕ МЕТОДА ИНТЕГРАЛЬНОЙ АДАПТАЦИИ ДЛЯ СИНТЕЗА АДАПТИВНЫХ ЗАКОНОВ УПРАВЛЕНИЯ ПНЕВМОПРИВОДОМ В УСЛОВИЯХ ГАРМОНИЧЕСКОГО ВОЗМУЩЕНИЯ.....	200

РАЗДЕЛ IV. ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

Л.К. Бабенко, И.Д. Русаловский МЕТОД РЕАЛИЗАЦИИ ГОМОМОРФНОГО ДЕЛЕНИЯ.....	212
Я.К. Миронов, П.Д. Миронова, К.Е. Румянцев ВЕРОЯТНОСТНЫЕ ХАРАКТЕРИСТИКИ ПОРОГОВОГО АЛГОРИТМА ОБНАРУЖЕНИЯ СИНХРОИМПУЛЬСОВ В СИСТЕМЕ КВАНТОВОГО РАСПРЕДЕЛЕНИЯ КЛЮЧА НА ОСНОВЕ ИНФОРМАЦИИ СО СМЕЖНОЙ ПАРЫ ВРЕМЕННЫХ СЕГМЕНТОВ.....	221

CONTENT

SECTION I. ARTIFICIAL INTELLIGENCE AND FUZZY SYSTEMS

V.V. Kureychik, V.V. Bova, Y.A. Kravchenko METHOD FOR SEARCHING SEQUENTIAL PATTERNS OF USER'S BEHAVIOR ON THE INTERNET.....	7
E.M. Gerasimenko SOLUTION OF THE PARTIALLY REVERSAL MODELLING TASK OF THE MINIMUM COST FLOW FINDING IN FUZZY CONDITIONS.....	21
V.A. Derkachev, V.V. Bakhchevnikov, A.N. Bakumenko CLASSIFIER OF IMAGES OF AGRICULTURAL CROPS SEEDS USING A CONVOLUTION NEURAL NETWORK.....	32
E.M. Gerasimenko, V.V. Stetsenko INTELLIGENT METHOD OF KNOWLEDGE EXTRACTION BASED ON SENTIMENT ANALYSIS.....	40
L.A. Gladkov, N.V. Gladkova EVOLUTIONARY DESIGN AS A TOOL FOR DEVELOPING MULTI-AGENT SYSTEMS.....	48
A.N. Karkishchenko, V.B. Mnukhin METHOD FOR DETECTING FEATURE POINTS OF AN IMAGE USING A SIGN REPRESENTATIONS.....	60
V.M. Kureichik, I.B. Safronenkova AN ONTOLOGICAL APPROACH TO DISTRIBUTED COMPUTING TECHNOLOGIES IMPLEMENTATION ON THE INTERNET.....	71
B.K. Lebedev, V.B. Lebedev, O.B. Lebedev POPULATION ALGORITHM FOR CONSTRUCTING A TREE OF SOLUTIONS BY METHOD OF CRYSTALLIZATION OF ALTERNATIVES FIELD.....	82
A.A. Sorokin, I.M. Borodyansky, A.V. Dagaev COMPARATIVE ANALYSIS OF MISSING DATA RECOVERY METHODS.....	94
G.F. Filaretov, Z. Bouchaala NON-PARAMETRIC METHOD FOR DETECTING BREAKDOWN OF TIME SERIES USING THE RANDOM WALKS THEORY MECHANISM.....	108

SECTION II. DESIGN AUTOMATION

N.O. Vasilyev, M.A. Zapletina, G.A. Ivanona, A.N. Schelokov LOGICAL RESYNTHESIS OF COMBINATIONAL CIRCUITS FOR RELIABILITY INCREASE.....	118
B.K. Lebedev, V.B. Lebedev, O.B. Lebedev SEARCH POPULATION ALGORITHM FOR VLSI ELEMENTS PLACEMENT.....	126
N.O. Vasilyev, P.I. Frolova, G.A. Ivanona, A.N. Schelokov LOGIC RESYNTHESIS METHODS FOR LAYOUT DESIGN OF MICROELECTRONIC CIRCUITS.....	137
L.A. Gladkov, N.V. Gladkova, M.J. Yasir HYBRID APPROACH THE JOINT SOLUTION OF PLACEMENT AND TRACING PROBLEMS.....	144
V.I. Danilchenko, Y.V. Danilchenko, V.M. Kureichik AUTOMATED STRUCTURAL-PARAMETRIC SYNTHESIS OF A STEPPED DIRECTIONAL RESPONDER ON CONNECTED LINES BASED ON A GENETIC ALGORITHM.....	157
D.V. Telpukhov, A.N. Schelokov, T.D. Zhukova ANALYSIS OF CHARACTERISTIC FOR CED CIRCUITS BASED ON REDUNDANT ENCODING METHODS.....	165

SECTION III. CONTROL SYSTEMS AND NONLINEAR DYNAMICS

A.G. Klovo, G.V. Kupovykh, I.A. Lyapunova THE MATHEMATICAL PROBLEM OF OPTIMAL CONTROL OF THE STRING	178
A.S. Cherepantsev STATISTICAL METHODS FOR EVALUATING THE CONNECTIVITY OF DYNAMIC SYSTEMS BASED ON A SEPARATE TIME PROJECTION	192
E.N. Obukhova APPLICATION OF THE METHOD OF INTEGRAL ADAPTATION FOR SYNTHESIS OF ADAPTIVE LAWS OF CONTROL IN A PNEUMATIC DRIVE UNDER HARMONIC DISTURBANCES	201

SECTION IV. INFORMATION SECURITY

L.K. Babenko, I.D. Rusalovsky METHOD OF IMPLEMENTING HOMOMORPHIC DIVISION.....	212
Ya.K. Mironov, P.D. Mironova, K.E. Rumyantsev PROBABILISTIC CHARACTERISTICS OF THE THRESHOLD ALGORITHM FOR DETECTING SYNCHRONIZING PULSES IN THE QUANTUM KEY DISTRIBUTION SYSTEM BASED ON INFORMATION FROM AN ADJACENT PAIR OF TIME SEGMENTS	222

Раздел I. Искусственный интеллект и нечеткие системы

УДК 004.853

DOI 10.18522/2311-3103-2020-4-6-21

В.В. Курейчик, В.В. Бова, Ю.А. Кравченко

МЕТОД ПОИСКА ПОСЛЕДОВАТЕЛЬНЫХ ПАТТЕРНОВ ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЕЙ В ИНТЕРНЕТ-ПРОСТРАНСТВЕ*

Одной из важных задач интеллектуального анализа данных является выделение закономерностей и обнаружение связанных событий в последовательных данных на основе анализа последовательных паттернов. В статье исследуются возможность применения последовательных паттернов для анализа событий поисково-познавательной деятельности пользователей при взаимодействии с Интернет-ресурсами открытой информационно-образовательной среды. Поиск последовательных паттернов является сложной вычислительной задачей, цель которой состоит в извлечении всех частых последовательностей, отражающих потенциальные связи внутри элементов из транзакционной базы данных последовательностей событий поисковой активности при заданной минимальной поддержке. Для ее решения в статье предлагается метод поиска закономерностей в последовательностях событий для обнаружения скрытых закономерностей, указывающих с возможные уровни уязвимости при выполнении задач информационного поиска в Интернет-пространстве. Описана математическая модель поведения пользователей в поисковой сессии, основанная на теории последовательных паттернов. Для повышения вычислительной эффективности метода разработан модифицированный алгоритм генерации последовательных паттернов, на первом этапе которого выполняется *AprioriAll*, формирующий частые последовательности-кандидаты всевозможных длин, а на втором – генетический алгоритм оптимизации входных параметров признакового пространства сгенерированного множества для поиска максимальных паттернов. Проведены серии вычислительных экспериментов на тестовых данных корпуса MSNBS, библиотеки интеллектуального анализа данных с открытым исходным кодом SPMF. Сравнительный анализ проводился с алгоритмами VMSP и GSP. Результаты исследований подтвердили эффективность поиска максимальных последовательных паттернов предложенным алгоритмом с точки зрения времени выполнения и количества извлеченных паттернов. Результаты проведенных экспериментальных исследований метода показали, что для увеличения стабильности и точности работы размер выборки, полученной в результате работы ГА, позволит сократить необходимое число сканирований базы данных паттернов, обеспечивая приемлемые вычислительные затраты, сопоставимые с алгоритмом VMSP и превосходящий по времени поиска последовательных паттернов алгоритм GSP в среднем более чем на 150 %.

Поиск последовательных паттернов; секвенциальный анализ; генетический алгоритм; транзакционная база данных; информационный поиск.

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-29-22019.

V.V. Kureychik, V.V. Bova, Y.A. Kravchenko

METHOD FOR SEARCHING SEQUENTIAL PATTERNS OF USER'S BEHAVIOR ON THE INTERNET

One of the important tasks of data mining is to isolate patterns and detect related events in sequential data based on the analysis of sequential patterns. The article examines the possibility of using sequential patterns to analyze the events of search and cognitive activity of users when interacting with Internet resources of an open information and educational environment. Searching for sequential patterns is a complex computational task whose goal is to retrieve all frequent sequences representing potential relationships within elements from a transactional database of sequences of search activity events with a given minimum support. To solve it, the article proposes a method for searching for patterns in sequences of events to detect hidden patterns that indicate possible levels of vulnerability when performing information search tasks in the Internet space. A mathematical model of user behavior in a search session based on the theory of sequential patterns is described. To improve the computational efficiency of the method, a modified algorithm for generating sequential patterns has been developed, at the first stage of which AprioriAll is performed, which forms frequent candidate sequences of all possible lengths, and at the second stage, a genetic algorithm for optimizing the input parameters of the feature space of the generated set to search for maximum patterns. A series of computational experiments were carried out on test data from the MSNBC corpus, the SPMF open source data mining library. The comparative analysis was carried out with the VMSP and GSP algorithms. The research results confirmed the efficiency of the search for maximum sequential patterns by the proposed algorithm in terms of the execution time and the number of extracted patterns. The results of the experimental studies of the method showed that to increase the stability and accuracy of the work, the sample size obtained as a result of the GA operation will reduce the required number of scans of the pattern database, providing acceptable computational costs comparable to the VMSP algorithm and the GSP algorithm that exceeds the search time for sequential patterns. an average of more than 150 %.

Search for sequential patterns; sequential analysis; genetic algorithm; transactional database; information search.

Введение. В настоящее время для поиска закономерностей при решении задач, связанных с необходимостью извлечения новых знаний при обработке данных об активности пользователей при взаимодействии с Интернет-ресурсами открытой информационно-образовательной среды, получили методы машинного обучения: ассоциативные правила и последовательные паттерны [1–3]. Задача поиска закономерностей в таких данных связана с необходимостью обнаружения и анализа возможных ассоциативных последовательностей событий поисковой сессии при выполнении поставленных перед пользователем задач [3–6]. Исследуемые объекты данных описываются числовыми признаками, поэтому для поиска ассоциативных зависимостей и построения на их основе последовательных паттернов необходимо осуществлять предварительную обработку таких атрибутов с учетом значимости признаков [4, 7].

Для решения задач повышения эффективности взаимодействия пользователей с информационными поисковыми системами ряд исследователей предлагают анализировать паттерны (шаблоны) активности пользователей, т.е. повторяющиеся последовательности действий [8–10]. Однако существующие методы анализа ориентированы на поиск лишь предопределенных исследователями шаблонов, основанных на заранее собранных рекомендациях и гипотезах относительно того, как пользователь взаимодействует с Интернет-пространством [6, 11–13]. Данные методы не используют существующие методы интеллектуального анализа ассоциативных правил и последовательных паттернов для поиска новых (ранее неизвестных) зависимостей в данных о сетевой активности пользователей.

При решении задачи извлечения новых знаний об исследуемых зависимостях в наборах транзакционных данных об активности пользователей возникают проблемы нахождения всех наборов элементов в том числе и неявных, позволяющих выявлять косвенные ассоциации в последовательности действий. Главная проблема при их поиске – большое число частых последовательностей, возникающих при исследовании больших бинарных контекстов длинных транзакций, масштабируемость во времени и пространстве для которых становится критичной. Это существенно усложняет экспертный анализ выявленных последовательных паттернов. Для решения этой проблемы используются различные меры значимости, такие как поддержка (support) и достоверность (confidence) [5, 7]. С их помощью найденные ассоциативные связи фильтруются, и для анализа предъявляются только те, для которых значения мер значимости превышают заданные пороговые значения. Подобная фильтрация способствует снижению количества сгенерированных паттернов, но не решает проблему размерности полностью. Это приводит к значительному возрастанию размерности решаемой задачи и повышает требования к вычислительным ресурсам.

Другой проблемой при поиске ассоциативных зависимостей в данных о сетевой активности пользователей, является распознавание событий, связанных с возможными уровнями уязвимости при выполнении задач информационного поиска. Поэтому требуется понимание семантики паттернов конкретных действий поисковой сессии [11, 14]. Для решения рассмотренных проблем авторами предлагается метод поиска последовательных паттернов и математическая модель поведения пользователей, для обнаружения скрытых закономерностей, указывающих на возникновение рисков поисково-познавательной деятельности при взаимодействии пользователей с Интернет-пространством.

Постановка задачи. Введем несколько основных понятий, необходимых для формализации задачи поиска последовательных паттернов (SPM) поведения пользователей при организации поисковой деятельности, определяемой следующими показателями [15]:

- ◆ эффективность – точность и полнота, с которой пользователи достигают поставленных целей (процент достигнутых целей и процент пользователей, успешно завершивших задачу);
- ◆ продуктивность – отношение израсходованных ресурсов к точности и полноте, с которой пользователи достигают поставленных целей (время необходимое на завершение задачи и задачи, выполненные в единицу времени);
- ◆ удовлетворенность – комфорт и приемлемость использования. Оценивается как восприятие пользователем таких показателей, как достоверность, полезность или легкость в нахождении необходимой информации.

Событие – факт, зафиксированный в определённый момент времени при взаимодействии определенного пользователя на определенном устройстве с информационной поисковой системой [8]. Событие обладает непустым уникальным набором атрибутов: пользователь, устройство, время, тип события (например, событие действия, командное событие) и специальные атрибуты, зависящие от типа события. Обозначим E множество всех зафиксированных событий: $E = \{e_1, \dots, e_n\}$, где $\{e_i\}$, $i = \overline{1, n}$ – отдельные события; n – мощность множества E .

Сессия – зафиксированный временной промежуток, в течение которого пользователь взаимодействовал с поисковой системой [9]. В рамках сессии все события накапливаются и хранятся в хронологическом порядке. Важно отметить, что каждое событие является уникальным и может быть включено только в одну сессию. Обозначим S множество всех зафиксированных сессий: $S = \{s_1, \dots, s_m\}$, где $\{s_i\}$, $i = \overline{1, m}$, – отдельные сессии; m – мощность множества S .

Сессия представляет собой размещение элементов множества E без повторений: $s_i = \langle e_{i1}, \dots, e_{ij} \rangle$, где $\{e_{ij}\}, i = \overline{1, m}, j = \overline{1, l_i}$ – отдельное событие i -й сессии; l_i – мощность размещения s_i .

Множество E формируется в результате объединения всех множеств сессий, полученных из данных активности пользователей. При этом любое событие принадлежит хотя бы одной сессии: $\forall e \in E, \exists s \in S, e \in s$.

Размещение – упорядоченный набор элементов множества либо с повторениями, либо без повторений в соответствии с критериями ассоциативных правил классификации [4]. Заметим, что в размещении без повторений, которым изначально является сессия, невозможно существование часто повторяющихся паттернов как упорядоченных подмножеств. Поэтому требуется предварительная классификация всех событий [7].

Классом событий назовем произвольную совокупность событий, обладающих определенным свойством или признаком. Обозначим C_E множество всех определенных классов событий: $C_E = \{c_1, \dots, c_k\}, \{c_i\}$, где $i = \overline{1, k}$, – отдельный класс; k – мощность множества C_E .

Паттерн представляет собой размещение элементов множества C_E с повторениями: $p_i = \langle c_{i1}, \dots, c_{ij} \rangle$, где $\{c_{ij}\}, i = \overline{1, r}, j = \overline{1, q_i}, c_{ij} \in C_E$ – отдельное событие i -го паттерна; r – мощность множества P ; q_i – мощность размещения p_i , то есть совокупность событий в паттерне.

Каждое событие описывается непустым множеством признаков M (набором атрибутов), присущим объекту множества E . Предполагаем, что все объекты в E и признаки в M различны. Пусть задано отношение $I \mid E \times M$ инцидентности между множествами E и M . Существование в I пары $(e, m), e \in E$ и $m \in M$, означает, что объект e имеет признак m и, наоборот, признак m характерен для объекта e . Тройку $K=(E, M, I)$ принято называть контекстом поиска ассоциативных правил [5, 7].

Ассоциативное правило $AR=(X, Y)$ на множестве признаков контекста $K=(E, M, I)$ имеет вид $X \Rightarrow Y$ и количественно характеризуется с помощью двух числовых функций: $sup(X \Rightarrow Y)$ – поддержка, $conf(X \Rightarrow Y)$ – достоверность. Поддержка и достоверность ассоциативного правила определяются через понятие поддержки множества признаков M .

Задача поиска последовательных паттернов заключается в обнаружении максимальных последовательностей, имеющих поддержку выше заданного порога. Требуется найти для заданного контекста $K_p=(E, M, I)$ множество AR всех ассоциативных зависимостей последовательных шаблонов P . Заметим, что искомым набор правил AR параметризован относительно пороговых значений sup и $conf$. Для классификации событий вычисляется целевая функция (ЦФ) достоверности правила согласно:

$$conf(X \Rightarrow Y) = \max f(E \rightarrow C_E). \quad (1)$$

2. Метод поиска закономерностей в последовательностях событий. Секвенциальный анализ (sequential pattern mining, поиск/добыча последовательных паттернов) – это метод интеллектуального анализа данных (data mining), объектом которого является база последовательностей – кортежа из наборов элементов (itemsets) – непустых множеств одновременно встречающихся элементов [2, 10]. Применительно к задаче поисковой деятельности пользователей, набор элементов будет соответствовать содержимому одного запроса, а вся последовательность будет представлять собой совокупность Интернет-ресурсов, просмотренных пользователем за всё время наблюдения. Пример последовательности с указанием её компонентного состава показан на рис. 1. Целью секвенциального анализа является

ся получение часто встречающихся подпоследовательностей классов событий C_E в заданной сессии S , удовлетворяющие ограничению минимальной поддержки, которые называются последовательными паттернами P [2, 16].

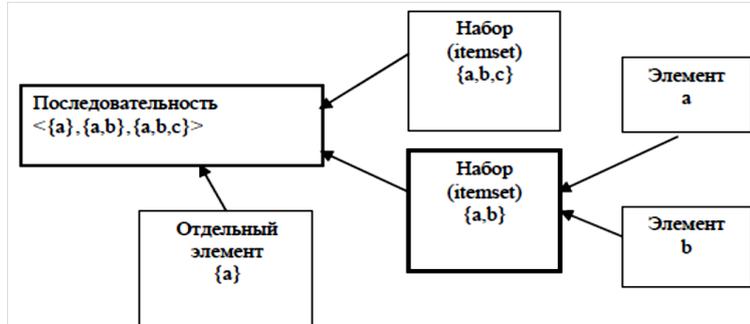


Рис. 1. Структура последовательного паттерна

Рассмотрим формализацию задачи поиска последовательных паттернов на основе модели поведения пользователя в поисковой сессии [13].

Пусть C_E множество всех определенных классов событий: $C_E = \{c_1, \dots, c_k\}, \{c_i\}$, где $i = \overline{1, k}$, – отдельный класс; k – мощность множества C_E .

Обозначим через S' множество всех зафиксированных сессий, после классификации и фильтрации $S' = \{s'_1, \dots, s'_m\}$, где $\{s'_i\}$, $i = \overline{1, m}$, – отдельная сессия; m – мощность множества S' . Сессия после классификации представляет собой размещение элементов множества C_E с повторениями: $s'_i = \langle c_{i1}, \dots, c_{ij} \rangle$, где $\{c_{ij}\}$, $i = \overline{1, m}, j = \overline{1, l'_i}$ – отдельное событие i -й сессии; l'_i – мощность размещения s'_i .

Введем n' суммарное количество событий в сессиях после классификации и фильтрации: $n' = \sum_{i=1}^m l'_i$, где m – мощность множества S' ; l'_i – мощность размещения s'_i .

Обозначим P множество кандидатов последовательных паттернов: $P = \{p_1, \dots, p_r\}$, где $\{p_i\}$, $i = \overline{1, r}$ – отдельный паттерн; r – мощность множества P . Будем считать длиной последовательного паттерна или сессии количество событий в указанном размещении, а также то, что паттерн p входит в сессию s' , если все элементы p содержатся в s' , при этом порядок элементов в подмножестве из s' соответствует порядку элементов p .

Примем за $\mu_p^{s'}$ значение функции принадлежности i -го ассоциативного признака паттерна, вычисленное как количество непересекающихся упорядоченных вхождений p в сессию s' . Тогда паттерн p называется поддерживаемым сессией s' , и как следствие, поддерживаемым пользователем, если количество вхождений $\mu_p^{s'} > 0$.

Выразим $\text{sup}(X \Rightarrow Y)$ через $\lambda_p^{s'}$ поддержку p сессией s' , рассчитываемую как:

$$\lambda_p^{s'} = \frac{\mu_p^{s'} * q}{l}, 0 \leq \lambda_p^{s'} \in \mathbb{R} \leq 1, \quad (2)$$

где $\lambda_p^{s'}$ – количество вхождений p в сессию s' ; q – длина паттерна p ; l – длина сессии s' .

Таким образом, для одной сессии можно описать значение поддержки как долю содержания паттерна в сессии. Данное условие необходимо для сравнения разных паттернов по степени влияния на процесс генерации максимальных паттернов взаимодействия пользователя с Интернет-ресурсами [17].

Число сессий пользователя может быть различным, поэтому необходимо агрегировать значение поддержки, сохранив их семантику. Рассчитаем общую поддержку паттерна как взвешенную среднюю арифметическую согласно формуле:

$$\lambda_p^{s'} = \sum_{i=1}^m \left(\frac{\mu_p^{s'_i} * q}{l_i} * \frac{l_i}{n'} \right) \leq \lambda_p^{s'} \in \mathbb{R} \leq 1, \quad (3)$$

где S' – множество сессий после классификации и фильтрации; m – мощность множества S' ; $\mu_p^{s'_i}$ – количество вхождений паттерна p в сессию s'_i ; q – длина паттерна p ; l – длина сессии s'_i ; n' – суммарное количество событий в сессиях после классификации и фильтрации.

Формулы (2) и средняя ЦФ (3) позволяют рассчитывать значение поддержки различных последовательных паттернов активности пользователей. Сравнение числовых значений поддержки позволяет ранжировать паттерны по степени приоритета для дальнейшего расчета ЦФ достоверности.

Для оценки эффективности поисковой деятельности пользователя необходимо рассчитать затраченные ресурсы пользователя относительно выбранных паттернов, которые содержат не исходные события, а классы, не имеющие информации о реальной длительности временных интервалов между событиями.

Для преобразования классов в затрачиваемое время вводится скалярная функция t , которая может быть записана следующим образом: $t: C_E \rightarrow \mathbb{R}_+$, где C_E – множество всех определенных классов событий; \mathbb{R}_+ – множество положительных вещественных чисел. На данный момент авторы предлагают, чтобы функцию t определял эксперт на основе знаний о возможном времени исполнения поисковой сессии.

Представим математическую модель поискового поведения пользователей с учетом затраченного времени как:

$$M_{User} = \{E, C_E, f, S, S', P, \Lambda, t\}, \quad (4)$$

где $E = \{e_i\}$, $i = \overline{1, n}$ – множество событий с атрибутами; $C_E = \{c_i\}$, $i = \overline{1, k}$ – множество классов событий; $f: E \rightarrow C_E$ – функция классификации событий; $S = \{s_i\}$, $i = \overline{1, m}$ – множество сессий до классификации, $s_i = \langle e_{i1}, \dots, e_{il} \rangle$, $i = \overline{1, m}$ – сессия до классификации, l – длина i -й сессии; $S' = \{s'_i\}$, $i = \overline{1, m}$ – множество сессий после классификации и фильтрации, $s'_i = \langle c_{i1}, \dots, c_{il} \rangle$, $i = \overline{1, m}$ – сессия после классификации и фильтрации, l' – длина i -й сессии после классификации и фильтрации; $P = \{p_i\}$, $i = \overline{1, r}$ – множество последовательных паттернов, $p_i = \langle c_{i1}, \dots, c_{iq} \rangle$, $i = \overline{1, r}$ – последовательный паттерн; $\Lambda = \{\lambda_p^{s'} \in \mathbb{R} | 0 \leq \lambda_p^{s'} \leq 1, p_i \in P, i = \overline{1, r}\}$ – множество значений поддержки последовательных шаблонов; $t: C_E \rightarrow \mathbb{R}_+$ – функция преобразования класса событий в затрачиваемое время.

Алгоритм поиска максимальных последовательных паттернов на основе генетической оптимизации. Согласно постановке задачи, база данных (БД) о транзакциях пользователей имеет следующие поля: идентификатор пользователя, время транзакции события поиска и посещенные Интернет-ресурсы. Каждая транзакция соответствует набору элементов, упорядоченных по времени транзакции для формирования последовательности поведения пользователей. Поддержка последовательности – это количество клиентских последовательностей, которые ее содержат. Если поддержка последовательности больше, чем минимальная поддержка, указанная пользователем, мы называем это частой последовательностью. Алгоритм анализа последовательного паттерна обнаружит частые последователь-

ности, называемые последовательными паттернами, среди всех транзакций. Таким образом, цель последовательного исследования паттернов - обнаружить все частые последовательности как последовательные паттерны, которые отражают потенциальные связи внутри элементов из базы данных последовательностей при заданной минимальной поддержке.

В последние десятилетия было предложено много алгоритмов [14, 18–20] для широкого спектра реальных приложений, таких как анализ поведения клиентов [9, 16] и извлечение информации [21] и др. Для повышения вычислительной эффективности метода разработан двухэтапный алгоритм генерации последовательных паттернов, на первом этапе которого выполняется AprioriAll [20], формирующий частые последовательности-кандидаты всех возможных длин, а на втором - генетический алгоритм (ГА) [22] оптимизации входных параметров признакового пространства сгенерированного множества для поиска максимальных паттернов.

В отличие от почти всех существующих алгоритмов SPM, которые итеративно чередуют расширение набора элементов и расширение последовательности, предложенный алгоритм позволяет проводить поиск лучших решений в глубину на основе:

1) анализа закрытого набора элементов на основе алгоритма AprioriAll. Все частые наборы элементов извлекаются, чтобы получить начальный набор последовательностей размера l ;

2) генерации максимальных последовательных паттернов на основе генетического поиска. Новые последовательности формируются путем анализа значения достоверности (приспособленности) сформированных на 1 этапе последовательностей с иерархической связью без поиска дополнительных часто встречающихся наборов элементов.

1 этап: работа алгоритма AprioriAll основана на подходе генерации кандидатов, который сканирует базу данных для создания и подсчета последовательных паттернов кандидатов и удаляет те, которые редко встречаются. Обобщенная схема алгоритма AprioriAll с введенными ранее обозначениями, представлена на рис. 2, где S_l – множество всех частых последовательностей, l – длина последовательности, P_l – множество последовательностей кандидатов длины l , p_i – последовательности-кандидаты, входящие в P_l , *Sup* - оператор вычисления поддержки.

Рассмотрим подробнее работу алгоритма. В качестве входных данных определены: база данных последовательностей (SDB), определяемый пользователем порог минимальной поддержки *minsupp* (значение в диапазоне [0, 1], представленное в процентах), максимальная длина последовательности l . Сначала сканируется входная база данных последовательностей один раз, чтобы построить вертикальное представление данных и множества частых последовательностей S_l . Каждая итерация алгоритма предусматривает «проход» по исходному набору данных, в котором каждый частый элемент $p_i \in S_l$, вызывает процедуру поиска [17].

Во время первой итерации вычисляется поддержка для каждого объекта (одноэлементной последовательности P_1) и выполняется фильтрация. В результате исходные данные для следующей итерации алгоритма формируются из последовательностей, чья поддержка равна, либо превышает пороговое значение *minsupp*. Последовательности, не имеющие достаточного уровня поддержки, исключаются. Далее генерируются более длинные последовательности-кандидаты, снова подсчитывается их поддержка и производится фильтрация множества кандидатов последовательных паттернов и их размещение во множестве частых последовательностей S_l зафиксированной сессии, результаты которой послужат исходными данными для работы генетического алгоритма. Число элементов последовательностей-кандидатов на каждой итерации алгоритма одинаково.

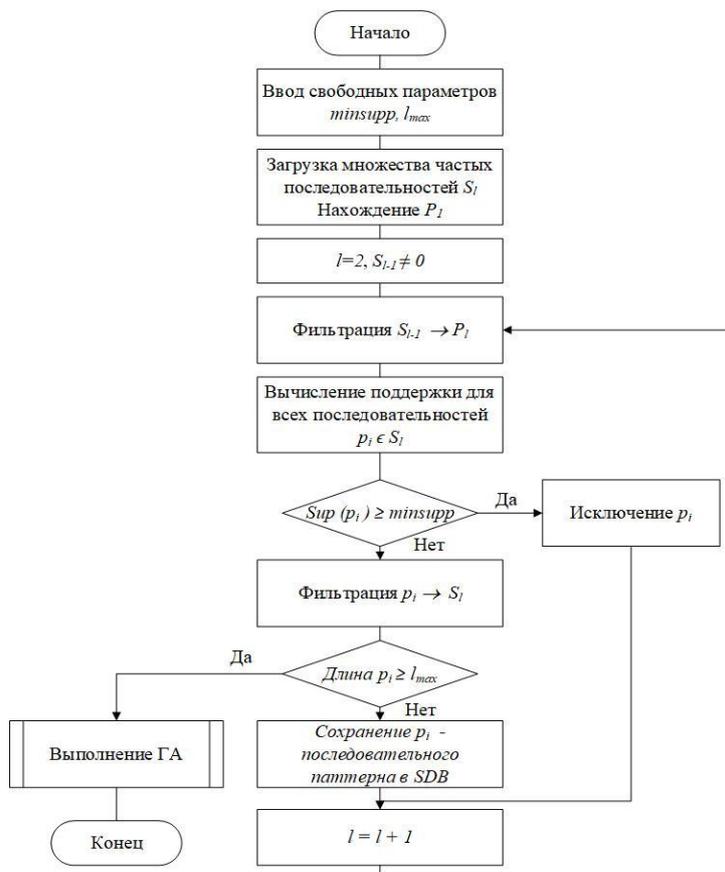


Рис. 2. Обобщенная схема алгоритма на основе AprioriAll

Работа 1 этапа алгоритма завершается тогда, когда не найдено ни одной новой последовательности с достаточным уровнем поддержки в конце очередного шага или, когда невозможно сформировать новых кандидатов.

На рис. 3 представлен фрагмент возможной структуры множества событий S_l -сессии, где красным цветом выделены элементы сформированного последовательного паттерна на 1 этапе алгоритма.

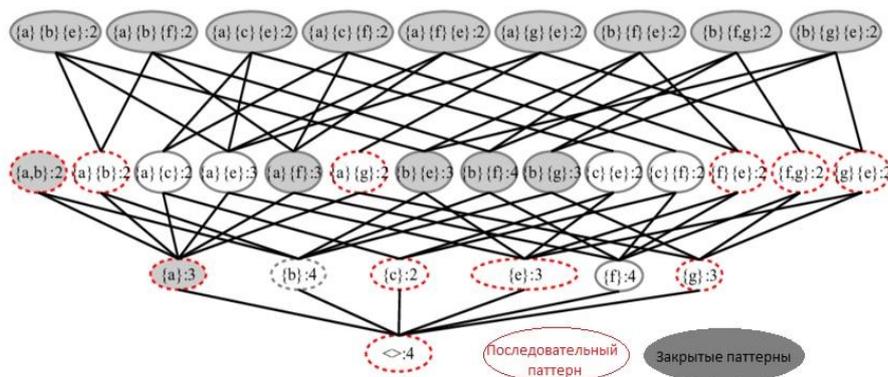


Рис. 3. Фрагмент структуры сформированного паттерна

2 этап: выполнение ГА. Для эффективной настройки параметров решаемой задачи генерации максимальных паттернов, авторы предлагают модифицированный алгоритм генетической оптимизации, структурная схема которого приведена на рис. 4 [22]. Рассмотрим данную структурную схему и опишем назначение каждого ее блока более подробно.

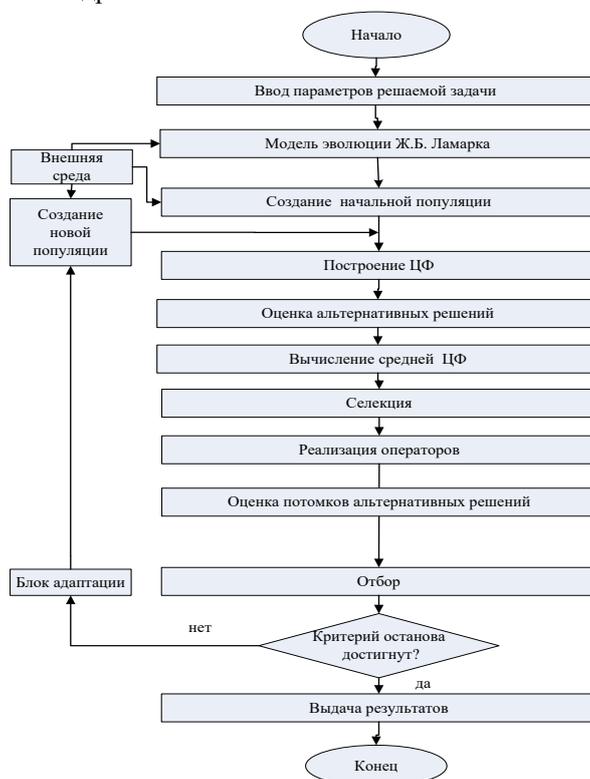


Рис. 4. Структурная схема алгоритма генетической оптимизации

Сначала на предварительном этапе производится ввод свободных параметров решаемой задачи генерации максимальных паттернов: количество сформированных последовательностей на 1 этапе алгоритма в базе данных транзакций – мощность сессии S , M ; длина максимального паттерна, q ; количество итераций, N . Затем происходит ввод таких параметров ГА, как: применяемая модель эволюции и селекции; виды и вероятности операторов поиска; выбор критерия останова алгоритма. В качестве модели эволюции по которому будет реализовываться генетическая оптимизация, авторы предлагают использовать модель эволюции Ж.Б. Ламарка – «наследование приобретенных признаков» [23].

Далее производится построение целевой функции и создание начальной популяции решаемой задачи на основе известного принципа «дробовика» [22]. Затем на основе построенной согласно формуле (1) ЦФ производится оценка приспособленности альтернативных решений и вычисляется среднее значение ЦФ популяции всех сгенерированных альтернативных решений по формуле (3).

В следующем блоке формируются родительские пары на основе выполнения турнирной селекции для реализации с заданной вероятностью различных операторов генетической оптимизации [24]. После получения набора потомков производится их оценка.

Далее на основе используемой модели эволюции производится отбор решений для последующей итерации поиска. Затем производится проверка достижения критерия останова, и если он не достигнут, то решения передаются в блок эволюционной адаптации. В этом блоке происходит выработка решений о перестройки текущей популяции и возможности изменения ее размера на каждой итерации поиска, что позволяет алгоритму избегать локальные оптимумы.

Затем решения передаются в блок создания новой популяции альтернативных решений и поиск продолжается итерационно до получения набора квазиоптимальных решений решаемой задачи. При достижении критерия останова – количества итераций алгоритма, производится «выдача результата».

Для управления всем процессом генетической оптимизации авторами введен блок внешней среды. Отметим, что предложенный алгоритм генетической оптимизации позволяет быстро получать более эффективные решения за счет концентрации поиска, т.е. отбирать те наследуемые признаки, которые в последующем будут переданы потомкам.

Экспериментальные исследования. Продемонстрируем результаты экспериментов по секвенциальному анализу поведения пользователей при взаимодействии с информационно-поисковым пространством Интернет с последующим построением максимальных паттернов. Поисковая деятельность демонстрируют типовые последовательности действий в поведении пользователя в процессе поисковой сессии, которую можно представить в виде транзакционной БД, состоящей из последовательностей событий при просмотре Интернет-ресурсов, сохраненных в лог-файлах. Анализ последовательности действий, которую можно обнаружить в логах будем проводить с помощью поиска последовательных паттернов (рис. 5). Каждый лог-файл будем считать единичным набором элементов в последовательностях-кандидатах потоков событий поисковой сессии.

Для экспериментов использовалась коллекция наборов данных MSNBC, библиотеки интеллектуального анализа данных с открытым исходным кодом SPMF [25], содержащая 989 818 последовательностей о потоках событий с веб-сайта MSNBC, преобразованных из исходных данных из репозитория UCI, со средним значением длины последовательности – 14 элементов.

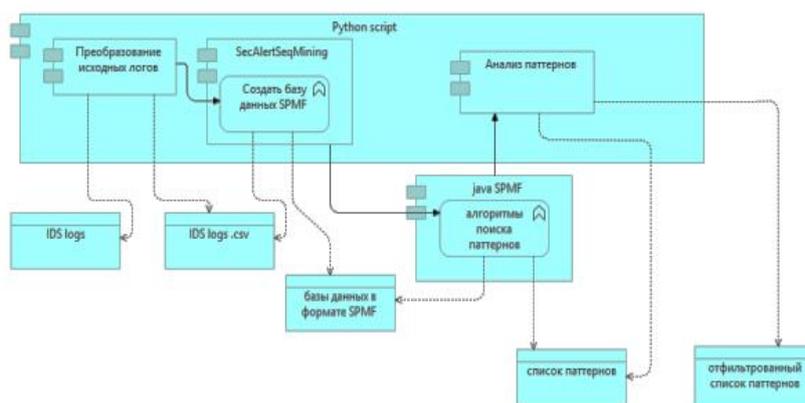


Рис. 5. Схема получения и анализа паттернов

Для анализа были выбраны алгоритмы поиска максимальных последовательных паттернов Generalized Sequential Pattern (GSP) и Vertical Mining of Maximal Sequential Patterns (VMSP). В алгоритме GSP используется горизонтальное внутреннее представление исходных данных и он обладает приемлемой скоростью работы [17]. VMSP – это быстрый алгоритм для обнаружения максимальных после-

довательных паттернов в базах данных последовательностей с вертикальным представлением данных, извлекающий максимальные последовательные паттерны с однократным сканированием базы данных [14].

Для тестирования эффективности алгоритмы сравнивались по числу сгенерированных максимальных паттернов – N и времени работы – t . В экспериментах оценивались максимальные наборы, для которых вычисляется поддержка $sup(\{p_1, p_2, \dots, p_n\}) \geq minsupp$ и для $\forall C_E$ осуществлялась классификация элементов по правилу $sup(\{p_1, p_2, \dots, p_n, C_E\}) > minsupp$. Для рассматриваемых алгоритмов тестовая серия производилась на разных значениях коэффициента уровня поддержки от 0,01 до 1. Численные результаты экспериментов для алгоритмов по критериям сравнения приведены в таблице, а также на рис. 6 и 7.

Таблица

Результаты сравнения алгоритмов для задачи поиска последовательных паттернов

Значения параметра поддержки	Методы					
	Apriori + ГА		GSP		VMSP	
$minsupp, \%$	t	N	t	N	t	N
1	2	90	5	69	2	80
0,5	2,5	166	6	106	3	154
0,3	3	172	8,5	125	4	163
0,2	4	285	9	130	4	186
0,1	5	297	12	162	6	180
0,05	6	305	17	170	8	232
0,03	8	380	25	292	12	304
0,02	9	580	32	518	14	419

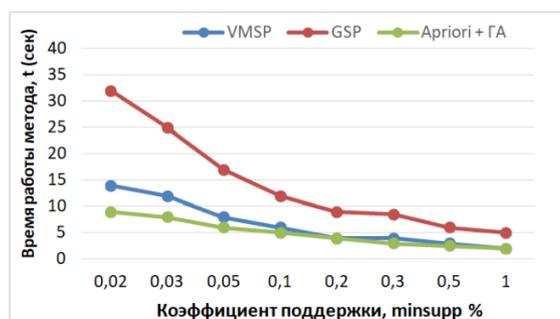


Рис. 6. График зависимости времени извлечения максимальных паттернов от значений уровня поддержки

Методика предобработки данных алгоритмом AprioriAll и ГА позволяет добиться лучших результатов извлечения максимальных паттернов в среднем на 20–40 % по сравнению с алгоритмами GSP и VMSP соответственно. На сравнительных графиках зависимости найденных паттернов и времени работы метода при различных значениях $minsupp$, представленных на рис. 6 и 7 можно оценить, что наилучшие результаты работы метода достигаются при значениях $minsupp = 0,02$. Предложенный метод с оптимизацией на основе ГА по времени работы сопоставим с алгоритмом VMSP и позволяет эффективно идентифицировать максимальные паттерны из большого набора данных при минимальном значении поддержки.

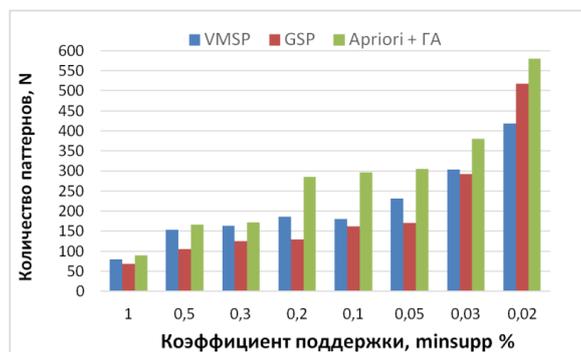


Рис. 7. График зависимости количества сгенерированных максимальных паттернов от значений уровня поддержки

Для определения эффективности разработанного алгоритма были проведены исследования количества выявленных последовательных паттернов на наборах тестовых примеров, различающихся количеством элементов выборки из транзакционной БД при варьировании параметра поддержки sup . На рис. 8 показано распределение паттернов, обнаруженных на DataSet1 и DataSet2, содержащих 29248 и 40100 транзакцию событий соответственно. В ходе вычислительного эксперимента было определено лучшее значение длины сформированных последовательных паттернов $l=5$, при котором число паттернов достигает максимального значения при различных варьируемых показателях уровня поддержки. Снижение данного показателя при $l > 5$, обусловлено большим количеством различных элементов в множестве часто встречающихся последовательностей. Расчетные данные результатов экспериментов показали, что количество сгенерированных закономерностей (рис. 8) увеличивается с при минимальных значениях sup . Так, например, в DS1 для $sup=0,0125$ количество $P=550$. Это объясняется тем, что по мере уменьшения уровня поддержки общее число связей увеличивается и количество несвязанных элементов в транзакциях уменьшается.

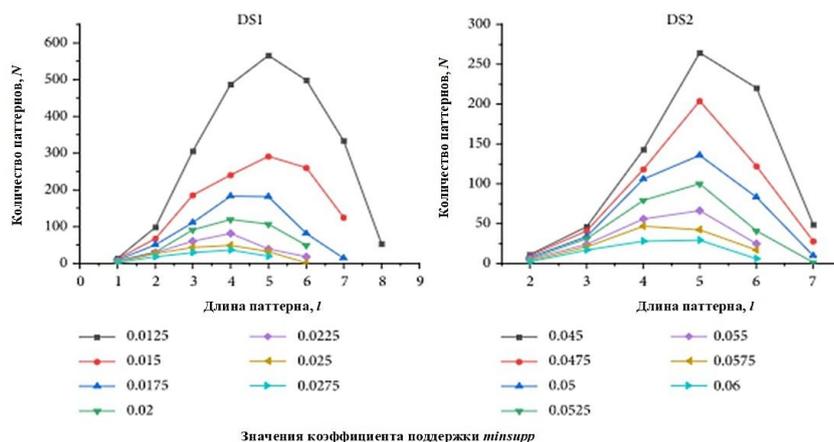


Рис. 8. Графики распределения количества сгенерированных максимальных паттернов от значений уровня поддержки на различных наборах данных

Результаты, полученные в ходе экспериментов доказывают, что стратегия на основе генетической оптимизации снижает время выполнения, так как она значительно сокращает пространство для поиска в тестовых наборах данных. Кроме того, результаты также показывают, что предложенный алгоритм позволяет эффективно извлечь последовательности при низком значении *sup*.

Заключение. Предложен метод получения максимальных последовательных паттернов, использующий методы секвенциального анализа и генетической оптимизации. Метод позволяет эффективно работать с категорийной входной информацией в процессе извлечения ассоциативных зависимостей при меньших временных затратах и избежать затратной процедуры предобработки транзакционной базы данных событий поисковой активности пользователей. Результаты проведенных экспериментальных исследований метода показали, что для увеличения стабильности и точности работы размер выборки, полученной в результате работы ГА, позволит сократить необходимое число сканирований базы данных паттернов, обеспечивая приемлемые вычислительные затраты, сопоставимые с алгоритмом VMSP и превосходящий по времени поиска максимальных паттернов алгоритм GSP более чем на 150 %. Так же результаты экспериментов показали, что ГА может эффективно идентифицировать максимальные паттерны из большого набора данных с низкими значениями показателя поддержки и позволяет извлекать на 20–40 % больше максимальных последовательностей чем алгоритмы VMSP и GSP соответственно. Предложенный метод можно использовать для исследования и анализа данных о поисковом поведении пользователей, выявления аномалий и угроз в информационных событиях их сетевой активности и распознавания возможных рисков информационно-образовательной деятельности в интернет-пространстве.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Tingting Z., Chen L.Y., Liang-Hsien T.* Understanding user motivation for evaluating online content: a self-determination theory perspective // *J. Behaviour and Information Technology*. – 2015. – Vol. 34. – P. 479-491.
2. *Gupta M., Han J.* Approaches for pattern discovery using sequential data mining. *Pattern Discovery Using Sequence Data Mining: Applications and Studies* // IGI Global. – 2012. – P. 137-154.
3. *Jalalirad A., Tjalkens T.* Using feature-based models with complexity penalization for selecting features // *J. Signal Processing Systems*. – 2018. – Vol. 90, Issue 2. – P. 201-210.
4. *Зайко Т.А., Олейник А.А., Субботин С.А.* Извлечение численных ассоциативных правил с учетом значимости признаков // *Восточно-Европейский журнал передовых технологий*. – 2013. – Т. 5, № 4 (65). – С. 28-34.
5. *Бова В.В., Щеглов С.Н., Лецанов Д.В.* Modified Approach to Problems of Associative Rules Processing based on Genetic Search // *International Russian Automation Conference (RusAutoCon)*. – 2019. – № 8867675.
6. *Bova V., Kravchenko Yu., Rodzin S., Kuliev E.* Hybrid method for prediction of users' information behavior in the Internet based on bioinspired search // *J. of Physics: Conference Series*. – 2019. – DOI: 10.1088/1742-6596/1333/3/032008.
7. *Wedyan S.* Review and Comparison of Associative Classification Data Mining Approaches // *International Journal of Computer, Information, Systems and Control Engineering*. – 2014. – Vol. 8. – P. 34-45.
8. *Оболонный В.И.* Обнаружение последовательных паттернов в событиях безопасности системы детекции вторжений // *Молодой ученый*. – 2018. – № 23 (209). – С. 181-187.
9. *Jingjun Zhu GG, Wu Haiyan.* An efficient method of web sequential pattern mining based on session filter and transaction identification // *J. Netw.* – 2010. – № 5 (9). – P. 1017-1024.
10. *Wang J.-Z., Huang J.-L.* On efficiently mining high utility sequential patterns // *Knowledge and Information Systems*. – 2016. – Vol. 49, No. 2. – P. 597-627.
11. *Кравченко Ю.А.* Модель фильтра знаний для задач семантической идентификации // *Известия ЮФУ. Технические науки*. – 2018. – № 4 (165). – С. 197-211.

12. Bova V.V., Nuzhnov E.V., Kureichik V.V. The combined method of semantic similarity estimation of problem oriented knowledge on the basis of evolutionary procedures // *Advances in Intelligent Systems and Computing*. – 2017. – Vol. 573. – P. 74-83.
13. Бова В.В., Кравченко Ю.А. Биоинспирированный подход к решению задачи классификации профилей поведения пользователей в интеллектуальных интернет-сервисах // *Известия ЮФУ*. – 2019. – № 4 (206). – С. 89-102.
14. Fournier Viger P., Cheng-W., Gomariz A., Tseng V. VMSP: Efficient Vertical Mining of Maximal Sequential Patterns. – 2014. – DOI: 10.1007/978-3-319-06483-3_8.
15. Лызь Н.А., Истратова О.Н. Информационно-образовательная деятельность в интернет-пространстве: виды, факторы, риски // *Педагогика*. – 2019. – № 4. – С. 16-26.
16. Truong-Chi T. Fournier-Viger P. A survey of high utility sequential pattern mining // *High-Utility Pattern Mining: Theory, Algorithms and Applications*. – Springer, 2019. – P. 97-129.
17. Fournier-Viger P., Wu C.-W., Tseng V.-S. Mining Maximal Sequential Patterns without Candidate Maintenance // *Proc. 9th Intern. Conference on Advanced Data Mining and Applications*. Springer. LNAI 8346. – 2013. – P. 169-180.
18. Gan W., Lin J. C.-W., Fournier-Viger P., Chao H.-C., Hong T.-P. A survey of incremental high-utility itemset mining // *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. – 2018. – Vol. 8, No.2. Art. e1242.
19. Gomariz A., Campos M., Marin R., Goethals B. ClaSP: An Efficient Algorithm for Mining Frequent Closed Sequences // *Proc. 17th Pacific-Asia Conf. Knowledge Discovery and Data Mining*. – Springer, 2013. – P. 50-61.
20. Singh J., Ram H. Improving Efficiency of Apriori Algorithm Using // *Journal of Scientific and Research Publications*. – 2013. – Vol. 3. – P. 1-4.
21. Лежебоков А.А., Кулиев Э.В. Технологии визуализации для прикладных задач интеллектуального анализа данных // *Известия Кабардино-Балкарского научного центра РАН*. – 2019. – № 4 (90). – С. 14-23.
22. Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы: учебник. – М.: Физматлит. 2010. – 368 с.
23. Курейчик В.В., Курейчик В.М., Сороколетов П.В. Анализ и обзор моделей эволюции // *Известия РАН. Теория и системы управления*. – 2007. – № 5. – С. 114-126.
24. Kureichik V., Zaporozhets D., Zaruba D. Generation of bioinspired search procedures for optimization problems // *Application of Information and Communication Technologies, AICT 2016 - 10*. – 2016. – P. 7991822.
25. SPMF: an open-source data mining library. – <https://www.philippe-fournier-viger.com/spmf/index.php?link=datasets.php>.

REFERENCES

1. Tingting Z., Chen L.Y., Liang-Hsien T. Understanding user motivation for evaluating online content: a self-determination theory perspective, *J. Behaviour and Information Technology*, 2015, Vol. 34, pp. 479-491.
2. Gupta M., Han J. Approaches for pattern discovery using sequential data mining. *Pattern Discovery Using Sequence Data Mining: Applications and Studies, IGI Global*, 2012, pp. 137-154.
3. Jalalirad A., Tjalkens T. Using feature-based models with complexity penalization for selecting features, *J. Signal Processing Systems*, 2018, Vol. 90, Issue 2, pp. 201-210.
4. Zayko T.A., Oleynik A.A., Subbotin S.A. Izvlechenie chislennykh associativnykh pravil s uchetom znachimosti priznakov [Extracting numeric Association rules taking into account the importance of the signs], *Vostochno-Evropeyskiy zhurnal peredovykh tekhnologiy* [East European journal of advanced technologies], 2013, Vol. 5, No. 4 (65), pp. 28-34.
5. Bova V.V., Shcheglov S.N., Leshchanov D.V. Modified Approach to Problems of Associative Rules Processing based on Genetic Search, *International Russian Automation Conference (RusAutoCon)*, 2019, No. 8867675.
6. Bova V., Kravchenko Yu., Rodzin S., Kuliev E. Hybrid method for prediction of users' information behavior in the Internet based on bioinspired search, *J. of Physics: Conference Series*, 2019. DOI: 10.1088/1742-6596/1333/3/032008.
7. Wedyan S. Review and Comparison of Associative Classification Data Mining Approaches, *International Journal of Computer, Information, Systems and Control Engineering*, 2014, Vol. 8, pp. 34-45.

8. *Obolonnyy V.I.* Obnaruzhenie posledovatel'nostnykh patternov v sobyitiyakh bezopasnosti sistemy detekcii vtorzheniy [Detection of sequential patterns in security events of the intrusion detection system], *Molodoy uchenyy* [Young scientist], 2018, No. 23 (209), pp. 181-187.
9. *Jingjun Zhu GG, Wu Haiyan.* An efficient method of web sequential pattern mining based on session filter and transaction identification, *J. Netw.*, 2010, No. 5 (9), pp. 1017-1024.
10. *Wang J.-Z., Huang J.-L.* On efficiently mining high utility sequential patterns, *Knowledge and Information Systems*, 2016, Vol. 49, No. 2, pp. 597-627.
11. *Kravchenko Yu.A.* Model' fil'tra znaniy dlya zadach semanticheskoy identifikatsii [Knowledge filter model for semantic identification tasks], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2018, No. 4 (165), pp. 197-211.
12. *Bova V.V., Nuzhnov E.V., Kureichik V.V.* The combined method of semantic similarity estimation of problem oriented knowledge on the basis of evolutionary procedures, *Advances in Intelligent Systems and Computing*, 2017, Vol. 573, pp. 74-83.
13. *Bova V.V., Kravchenko Yu.A.* Bioinspirirovannyi podkhod k resheniyu zadachi klassifikatsii profilye povedeniya pol'zovateley v intellektual'nykh internet-servisakh [Bioinspired approach to solving the problem of classifying user behavior profiles in intelligent Internet services], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2019, No. 4 (206), pp. 89-102.
14. *Fournier Viger P., Cheng-W., Gomariz A., Tseng V.* VMSP: Efficient Vertical Mining of Maximal Sequential Patterns, 2014. DOI: 10.1007/978-3-319-06483-3_8.
15. *Lyz' N.A., Istratova O.N.* Informacionno-obrazovatel'naya deyatelnost' v internet-prostranstve: vidy, faktory, riski [Information and educational activities in the Internet space: types, factors, risks], *Pedagogika* [Pedagogy], 2019, No. 4, pp. 16-26.
16. *Truong-Chi T. Fournier-Viger P.* A survey of high utility sequential pattern mining, *High-Utility Pattern Mining: Theory, Algorithms and Applications*. Springer, 2019, pp. 97-129.
17. *Fournier-Viger P., Wu C.-W., Tseng V.-S.* Mining Maximal Sequential Patterns without Candidate Maintenance, *Proc. 9th Intern. Conference on Advanced Data Mining and Applications*. Springer. LNAI 8346, 2013, pp. 169-180.
18. *Gan W., Lin J. C.-W., Fournier-Viger P., Chao H.-C., Hong T.-P.* A survey of incremental high-utility itemset mining, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2018, Vol. 8, No.2. Art. e1242.
19. *Gomariz A., Campos M., Marin R., Goethals B.* ClaSP: An Efficient Algorithm for Mining Frequent Closed Sequences, *Proc. 17th Pacific-Asia Conf. Knowledge Discovery and Data Mining*. Springer, 2013, pp. 50-61.
20. *Singh J., Ram H.* Improving Efficiency of Apriori Algorithm Using, *Journal of Scientific and Research Publications*, 2013, Vol. 3, pp. 1-4.
21. *Lezhebokov A.A., Kuliev E.V.* Tekhnologii vizualizatsii dlya prikladnykh zadach intellektual'nogo analiza dannykh [Visualization technologies for data mining applications], *Izvestiya Kabardino-Balkarskogo nauchnogo centra RAN* [Izvestiya Kabardino-Balkar scientific center of the Russian Academy of Sciences], 2019, No. 4 (90), pp. 14-23.
22. *Gladkov L.A., Kureychik V.V., Kureychik V.M.* Geneticheskie algoritmy: uchebnik [Genetic algorithms: textbook]. Moscow: Fizmatlit. 2010, 368 p.
23. *Kureychik V.V., Kureychik V.M., Sorokoletov P.V.* Analiz i obzor modeley evolyutsii [Analysis and review of evolution models], *Izvestiya RAN. Teoriya i sistemy upravleniya* [Izvestiya RAS. Theory and control systems], 2007, No. 5, pp. 114-126.
24. *Kureichik V., Zaporozhets D., Zaruba D.* Generation of bioinspired search procedures for optimization problems, *Application of Information and Communication Technologies, AICT 2016 – 10*, 2016, pp. 799-822.
25. SPMF: an open-source data mining library. Available at: <https://www.philippe-fournier-viger.com/spmf/index.php?link=datasets.php>.

Статью рекомендовал к опубликованию д.т.н., профессор Ю.А. Гатчин.

Курейчик Владимир Викторович – Южный федеральный университет; e-mail: vkur@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371651; кафедра систем автоматизированного проектирования; зав. кафедрой; профессор.

Бова Виктория Викторовна – e-mail: vvbova@sfedu.ru; кафедра систем автоматизированного проектирования; доцент.

Кравченко Юрий Алексеевич – e-mail: yakravchenko@sfedu.ru; кафедра систем автоматизированного проектирования; доцент.

Kureychik Vladimir Victorovich – Southern Federal University; e-mail: vkur@sfedu.ru; 44, Nekrasovskiy lane, Taganrog, 347928, Russia; phone: +78634371651; the department of computer aided design; head of department; professor.

Bova Victoria Victorovna – e-mail: vvbova@sfedu.ru; the department of computer aided design; associate professor.

Kravchenko Yury Alekseevich – e-mail: yakravchenko@sfedu.ru; the department of computer aided design; associate professor.

УДК 519.178

DOI 10.18522/2311-3103-2020-4-21-31

Е.М. Герасименко

РЕШЕНИЕ ЗАДАЧИ МОДЕЛИРОВАНИЯ ЧАСТИЧНО РЕВЕРСИВНОГО ПОТОКА МИНИМАЛЬНОЙ СТОИМОСТИ В НЕЧЕТКИХ УСЛОВИЯХ*

Данная статья посвящена разработке алгоритма решения задачи моделирования частично реверсивного потока минимальной стоимости в нечеткой транспортной сети. Задача нахождения потока минимальной стоимости является центральной задачей при планировании перевозок и эвакуационном моделировании. Актуальность такого рода задач обусловлена необходимостью поиска оптимальных с точки зрения стоимости маршрутов перевозок и передачи по ним максимального потока. Данная статья посвящена решению данной задачи в нечетких условиях, так как аппарат теории нечетких множеств позволяет задавать параметры сети, такие как пропускные способности участков дорог, стоимости перевозок в нечетком виде. Такой способ представления удобен в ситуациях, когда имеет место нехватка данных о моделируемом объекте, их лингвистический характер, погрешности в измерениях и пр. В задачах эвакуационного моделирования, которые происходят спонтанно, также наблюдается нехватка точной информации о пропускных способностях и стоимостях перевозок. Концепция контрпотока, используемая в статье, используется для увеличения суммарной пропускной способности путем реверсирования движения. Техника реверсирования движения является современной методикой увеличения передаваемого потока путем увеличения выходной пропускной способности сети. Применение реверсирования движения позволяет освободить загруженные участки дороги и перераспределить движение в сторону пустых дорог, устраняя заторы и «пробки» на дорогах. Предложен метод оперирования нечеткими числами, не приводящий к «размытию» границ результирующего числа и позволяющий оперировать нечеткими границами на последних итерациях, в то время как на остальных предшествующих итерациях производятся вычисления только с центрами нечетких чисел. Рассмотрен численный пример, который иллюстрирует работу предложенного алгоритма.

Частично реверсивный поток, нечеткий поток минимальной стоимости, транспортные сети.

E.M. Gerasimenko

SOLUTION OF THE PARTIALLY REVERSAL MODELLING TASK OF THE MINIMUM COST FLOW FINDING IN FUZZY CONDITIONS

This article is devoted to the development of an algorithm for solving the problem of modeling a partially reversal flow of minimum cost in a fuzzy transportation network. The minimum cost flow problem is a central problem in transportation planning and evacuation modelling. The relevance of these tasks is due to necessity to find optimal transportation routes in terms of cost and

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-01-00559 а.

transfer the maximum flow along them. This article is devoted to solving this problem in fuzzy conditions, since the apparatus of the theory of fuzzy sets allows you to set network parameters, such as the capacity of road sections, the cost of transportation in a fuzzy form. This method of assignment is convenient in situations where there is a lack of data on the modeled object, linguistic nature of data, measurement errors, etc. In the problems of evacuation modelling, which occur spontaneously, there is also a lack of accurate information about the capacity and cost of transportation. The contraflow concept, which was used in the paper, allows increasing the total flow by reversing traffic. The lane reversal technique is a modern technique for increasing the transmitted traffic by increasing the network output capacity. The use of traffic reversal allows releasing congested sections of the road and redistributing traffic towards unloaded roads, eliminating congestion and "traffic jams" on the roads. A method of operating with fuzzy numbers is proposed, which does not lead to "blurring" of the boundaries of the resulting number and allows operating with fuzzy boundaries at the last iterations, while at the rest of the previous iterations, calculations are performed only with the centers of fuzzy numbers. A numerical example is considered that illustrates the operation of the proposed algorithm.

Partially reversal flow, fuzzy minimum cost flow, transportation networks.

Введение. Задача нахождения потока минимальной стоимости является одной из важнейших в теории потоков, поскольку позволяет оптимизировать стоимость перевозимого груза. Особое значение она имеет в задачах, где необходимо найти оптимальные пути перевозки заданного или максимального количества груза. В основе задачи лежит поиск так называемого кратчайшего пути (если расчетным параметром является время), или пути минимальной стоимости (если требуется минимизировать стоимость перевозки). Поиск пути осуществляется с помощью различных алгоритмов, среди которых наиболее известными и широко используемыми являются алгоритмы авторов Э. Дейкстра, Р. Беллман и Л. Форд, Р. Флойд и С. Уоршалл, Б. Левит и Д. Джонсон [1]. Алгоритмы отличаются временной сложностью, возможностью работы с отрицательными весами ребер, поиска кратчайшего пути между каждой парой вершин графа и пр. В задачах эвакуационного моделирования эксперт часто сталкивается с задачами необходимости перемещении максимального количества эвакуированных из опасных зон в безопасные укрытия за минимальное возможное время. Принимая время передачи потока за стоимость, исследователь приходит к классической задаче нахождения потока минимальной стоимости.

При моделировании потоков и эвакуационном моделировании нередко возникает необходимость решения потоковых задач в условиях неполноты и неточности информации. Данные факторы обусловлены нехваткой данных о моделируемых объектах и их параметрах, уникальностью постановок задач, отсутствием статистических данных, погрешностями и ошибками в измерениях. Иногда можно только приблизительно оценить пропускные способности и стоимости перевозок. В задачах эвакуационного моделирования данные о параметрах сети могут отсутствовать, и эксперту необходимо принимать решение на основе имеющейся неполной информации. В связи с этим некоторые задачи нахождения минимального потока в ряде случаев следует решать с использованием аппарата теории нечетких множеств [2–3]. Представляя параметры сети в виде нечетких чисел различной формы, лица, принимающие решения, могут получать более корректные и чувствительные к изменениям окружающей среды решения.

В задачах эвакуационного моделирования, основанных на теории потоков, имеет место техника реверсирования движения, которая представляет собой усовершенствованный метод, позволяющий увеличивать значения транспортируемого потока за счет увеличения исходящей пропускной способности сети. Когда движение реверсируется, это позволяет освободить загруженные участки дороги и перераспределить движение, направив его в сторону незагруженных участков дорог, тем самым устраняя заторы на дорогах. Концепция маршрутизации эвакуации по полосам была предложена в [4–5].

Комплексный обзор потоковых задач, основанный на концепции встречного потока, был проведен в [4]. Авторы Dhamala и Ryakurel [6] разработали методы нахождения контрпотока и включили свои разработки в механизм поиска классических потоков в транспортных сетях с целью увеличения общей исходящей пропускной способности. Кроме того, ими была введена концепция частичного встречного потока для изменения дуг моделируемой сети до требуемой пропускной способности [6–9]. Этот метод решает проблему с «пробками» позволяя освободить участки высокой загруженности.

Задачи нахождения потока минимальной стоимости в нечётких условиях описаны в литературе авторами [10–14], однако отсутствуют методы решения подобных задач с учетом реверсирования движения. Авторами [15–16] рассмотрены методы нахождения нечёткого лексикографического потока и потока наискорейшего прибытия. Таким образом, в данной статье разработан метод нахождения потока минимальной стоимости в нечёткой транспортной сети с частичным реверсированием движения.

Постановка задачи и разработанный алгоритм. Рассмотрим постановку задачи нахождения потока минимальной стоимости в нечётких условиях и в условиях частичного реверсирования движения (1)–(3).

$$\sum_{(x_i, x_j)} \tilde{c}_{ij} \tilde{\xi}_{ij} \rightarrow \min, \quad (1)$$

$$\sum_{x_j \in \Gamma(x_i)} \tilde{\xi}_{ij} = \sum_{x_k \in \Gamma^{-1}(x_i)} \tilde{\xi}_{ki} = \begin{cases} \tilde{\rho}, x_i = s, \\ -\tilde{\rho}, x_i = t, \\ \tilde{0}, x_i \neq s, t. \end{cases} \quad (2)$$

$$0 \leq \tilde{\xi}_{ij} \leq \tilde{u}_{ij}, \forall (x_i, x_j) \in \tilde{A}. \quad (3)$$

Представим разработанный алгоритм поиска потока минимальной стоимости в нечетких условиях.

Этап 1. Выполняем построение модифицированной сети для реверсирования дуг.

Нечеткая модифицированная транспортная сеть с учетом частичного реверсирования движения – это такая сеть, в которой исходные пропускные способности дуг графа \tilde{u}_{ij} , \tilde{u}_{ji} заменяются новой пропускной способностью $\tilde{u}_{ij}^{\mu} = \tilde{u}_{ij} + \tilde{u}_{ji}$, а стоимость прохождения единицы потока по дуге равна:

$$\tilde{c}_{ij}^{\mu} = \begin{cases} \tilde{c}_{ij}, (i, j) \in \tilde{A}, \\ \tilde{c}_{ji}, \text{ в противном случае.} \end{cases}$$

Этап 2. Строим остаточную сеть для модифицированной сети.

Модифицированная остаточная сеть $\tilde{G}^{\mu r}$ с учетом частичного реверсирования движения по дугам строится следующим образом:

2.1. если для потока выполняется условие $\tilde{\xi}_{ij}^{\mu} < \tilde{u}_{ij}^{\mu}$, то остаточная пропускная способность для прямой дуги определяется как $\tilde{u}^{\mu r} = \tilde{u}_{ij}^{\mu} - \tilde{\xi}_{ij}^{\mu}$.

2.2. если для потока выполняется условие $\tilde{\xi}_{ij}^{\mu} > \tilde{0}$, то остаточная пропускная способность противоположной дуги определяется как $\tilde{u}^{\mu r} = \tilde{\xi}_{ij}^{\mu}$.

Этап 3. Поиск пути.

Ищем путь минимальной стоимости $\tilde{p}^{\mu r}$ в остаточной модифицированной сети $\tilde{G}^{\mu r}$ [11–12].

а). Пометка вершины-источника = 0, пометки остальных вершин считаются равными бесконечности.

Все вершины сети помечаем как непосещенные.

б). Если все вершины сети посещены, алгоритм заканчивает свою работу.

В противном случае, вершина x с минимальной пометкой выбирается из не-посещенных вершин.

Мы рассматриваем все возможные маршруты, в которых x – предпоследняя вершина. Вершины, к которым ведут ребра из x , являются соседними для вершины x . Для каждого соседа вершины x , кроме тех, которые помечены как посещенные, рассмотрим новый путь, равный сумме значений текущей пометки вершины x и стоимости ребра, соединяющего x с соседними вершинами.

Если полученное значение стоимости меньше, чем пометка соседней вершины, заменяем ее пометку на полученное значение стоимости. Рассмотрев все соседние вершины, отмечаем x как помеченную и повторяем шаг алгоритма.

3.1. Если путь $\tilde{p}^{\mu r}$ найден, переходим к шагу 4.

3.2. Если путь $\tilde{p}^{\mu r}$ не найден, максимальный поток минимальной стоимости получен, переход к этапу 6.

Этап 4. Передача потока.

Передаем поток в соответствии с минимальным значением из пропускных способностей минимального пути, равный пропускной способности дуги в модифицированной остаточной сети $\tilde{\delta}^{\mu r} = \min[\tilde{u}_{ij}^{\mu r}]$.

Этап 5. Обновление потока.

Пересчитываем значения потоков и переходим к шагу 2.

1) для дуг в противоположных направлениях между вершинами (x_j, x_i) модифицируем значение потока $\tilde{\xi}_{ij}^{\mu}$ и уменьшаем его на значение $\tilde{\delta}^{\mu}$. Финальное значение потока $\tilde{\xi}_{ij}^{\mu} - \tilde{\delta}^{\mu}$.

2) для дуг между вершинами (x_i, x_j) модифицируем значение потока $\tilde{\xi}_{ij}^{\mu}$ и увеличиваем его на значение $\tilde{\delta}^{\mu}$. Финальное значение потока $\tilde{\xi}_{ij}^{\mu} + \tilde{\delta}^{\mu}$.

Этап 6. Реверсирование дуг.

6.1. Изменяем направление дуги $(x_j, x_i) \in \tilde{A}$ на противоположное до пропускной способности $\tilde{\xi}_{ij} - \tilde{u}_{ij}$, если $\tilde{\xi}_{ij} > \tilde{u}_{ij}$ заменяя \tilde{u}_{ij} на $\tilde{\delta}$, если $(x_i, x_j) \notin \tilde{A}$.

6.2. Для каждой дуги $(x_i, x_j) \in \tilde{A}$, если дуга (x_i, x_j) реверсируется, $\tilde{s}_{ij} = \tilde{u}_{ij}^{\mu} - \tilde{\xi}_{ji}$, $\tilde{s}_{ji} = \tilde{\delta}$. Если \tilde{s}_{ij} и \tilde{s}_{ji} не реверсируются, то $\tilde{s}_{ij} = \tilde{u}_{ij} - \tilde{\xi}_{ij}$, где \tilde{s}_{ij} – сохраненная пропускная способность.

Численный пример. Представим численный пример, иллюстрирующий работу алгоритма нахождения потока минимальной стоимости с частичным реверсированием движения. Задача заключается в поиске максимального значения потока минимальной стоимости, который может пройти между источником и стоком в нечеткой транспортной сети, при этом учитывая технологию частичного изменения направления движения по полосам дороги. Рис. 1 иллюстрирует исходную нечеткую транспортную сеть, дугам которой приписаны значения нечетких пропускных способностей и стоимостей перевозок. Для эффективного оперирования нечеткими числами будем использовать метод, основанный на «размытии» границ нечетких чисел на последнем шаге, тем самым на протяжении всего алгоритма в вычислениях будут использоваться только центры нечетких чисел.

Этап 1. Преобразуем исходную нечеткую транспортную сеть на рис. 1 в модифицированную сеть для возможности частичного реверсирования движения, как показано на рис. 2.

Этап 2. Строим остаточную сеть для модифицированной сети. На данном этапе, так как величины потоков, текущие по дугам, равны 0, остаточная сеть совпадает с исходной.

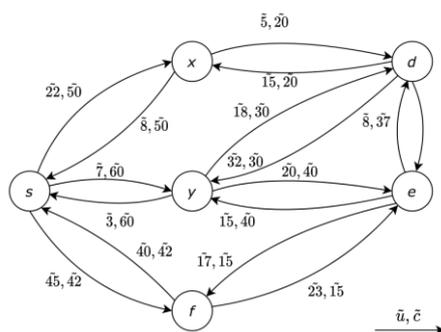


Рис. 1. Исходная транспортная сеть

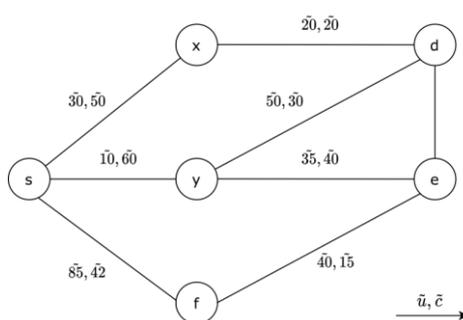


Рис. 2. Модифицированная транспортная сеть для реверсирования движения

Этап 3. Ищем путь минимальной стоимости в построенной сети. Согласно алгоритму Дейкстры [17] – это путь $s \rightarrow x \rightarrow d$ стоимости $\bar{70}$ единиц.

Этап 4. Передаем по найденному пути $s \rightarrow x \rightarrow d$ минимальную из пропускных способностей дуг пути, равную $\bar{20}$ единицам потока, как показано на рис. 3.

Этап 5. Пересчитываем значение потока и строим остаточную сеть.

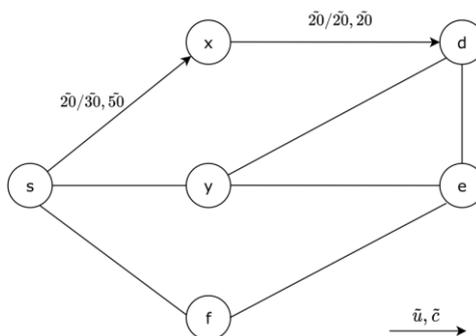


Рис. 3. Сеть с потоком $\bar{20}$ единиц

Этап 3. Находим путь минимальной стоимости в остаточной сети. Согласно алгоритму Дейкстры – это путь $s \rightarrow y \rightarrow d$ стоимости $\bar{90}$ единиц.

Этап 4. Передаем по найденному пути $s \rightarrow y \rightarrow d$ минимальную из пропускных способностей дуг пути, равную $\bar{10}$ единиц потока (рис. 4).

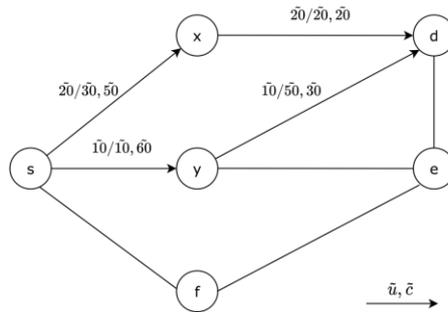


Рис. 4. Сеть с потоком $\overline{30}$ единиц

Этап 5. Пересчитываем значение потока и строим остаточную сеть.

Этап 3. Ищем путь минимальной стоимости в построенной сети. Согласно алгоритму Дейкстры – это путь $s \rightarrow f \rightarrow e \rightarrow d$ стоимости $\overline{94}$ единиц.

Этап 4. Передаем по найденному пути $s \rightarrow f \rightarrow e \rightarrow d$ минимальную из пропускных способностей дуг пути, равную $\overline{18}$ единиц потока (рис. 5).

Этап 5. Пересчитываем значение потока и строим остаточную сеть.

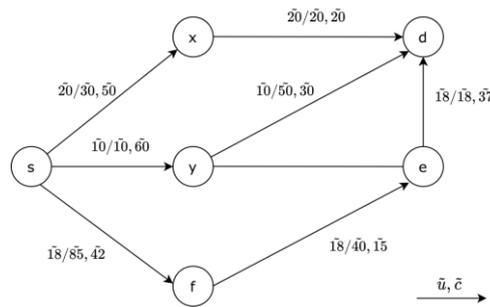


Рис. 5. Сеть с потоком $\overline{48}$ единиц

Этап 3. Ищем путь минимальной стоимости в построенной сети. Согласно алгоритму Дейкстры – это путь $s \rightarrow f \rightarrow e \rightarrow y \rightarrow d$ стоимости $\overline{127}$ единиц.

Этап 4. Передаем по найденному пути путь $s \rightarrow f \rightarrow e \rightarrow y \rightarrow d$ минимальную из пропускных способностей дуг пут, равную $\overline{22}$ единиц потока (рис. 6).

Этап 5. Пересчитываем значение потока и строим остаточную сеть.

Этап 3. Так как не удалось найти пути в остаточной сети, получен максимальный поток минимальной стоимости. Максимальный поток составляет $\overline{70}$ единиц суммарной стоимости $\overline{6786}$ условных единиц. Переходим к этапу 6.

Этап 6. Вычисляем значения сохраненных пропускных способностей в соответствии с частичным реверсированием движения по дугам. Итоговое распределение сохраненных пропускных способностей показано на рис. 7.

1) Дуги, которые реверсируются полностью: $(y \rightarrow s), (d \rightarrow x), (e \rightarrow f), (d \rightarrow e)$.

2) Дуги, которые частично реверсируются $(x \rightarrow s)$ до пропускной способности $\overline{10}$; $(d \rightarrow y)$ до пропускной способности $\overline{18}$, $(y \rightarrow e)$ до пропускной способности $\overline{13}$.

3) Список дуг с сохраненными пропускными способностями:

Сохраненная пропускная способность до $\tilde{40}$ units: $(f \rightarrow s)$;

Сохраненная пропускная способность до $\tilde{5}$ units: $(s \rightarrow f)$;

Сохраненная пропускная способность до $\tilde{13}$ units: $(y \rightarrow e)$;

Сохраненная пропускная способность $\tilde{18}$ units: $(d \rightarrow y)$;

Сохраненная пропускная способность $\tilde{10}$ units: $(x \rightarrow s)$.

С помощью предложенного метода удалось вычислить значения максимального потока минимальной стоимости в нечеткой сети, а также значения сохраненных пропускных способностей, следовательно, следующим шагом является нахождение результирующего нечеткого треугольного числа, учитывая степени размытия границ числа [18–19].

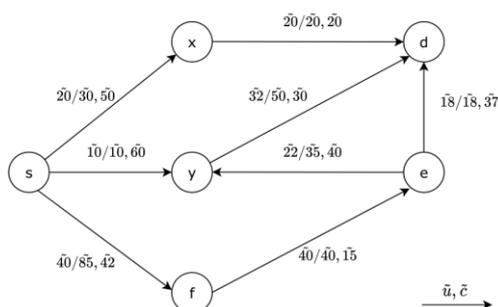


Рис. 6. Сеть с потоком $\tilde{70}$ единиц

Применим метод, описанный в [1]. Суть данного метода заключается в том, что у нас нет необходимости оперировать границами отклонений нечетких чисел на первом шаге осуществления вычислений: достаточно оперировать центральными значениями треугольных чисел, «размывая» их границы на последнем шаге с помощью заданных экспертами базовых значений. Преимуществом данного подхода является сохранение результирующим значением потока своего значения без сильного размытия границ числа.

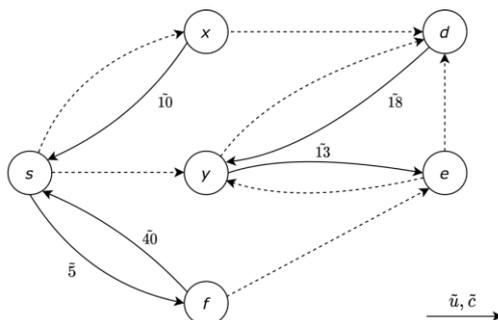


Рис. 7. Сеть с сохраненными пропускными способностями

Рис. 8 иллюстрирует методику размытия чисел. Таким образом, рис. 9 и 10 показывают границы отклонений базовых значений нечетких пропускных способностей и стоимостей перевозок, заданных экспертами.

Границы результирующего числа определяются по формуле (4):

$$l^L = \frac{(a_2 - a')}{(a_2 - a_1)} \times l_1^L + \left(1 - \frac{(a_2 - a')}{(a_2 - a_1)}\right) \times l_2^L, \quad l^R = \frac{(a_2 - a')}{(a_2 - a_1)} \times l_1^R + \left(1 - \frac{(a_2 - a')}{(a_2 - a_1)}\right) \times l_2^R. \quad (4)$$

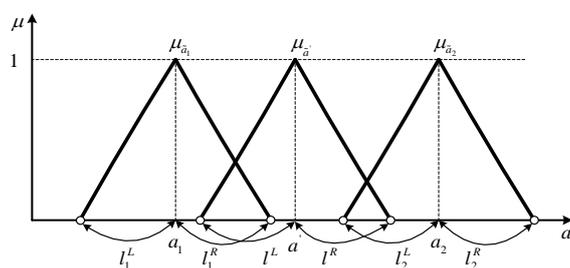


Рис. 8. Определение функции принадлежности $\mu_{a'}(a)$

Нечеткий поток «около 70» расположен между двумя соседними базовыми значениями пропускных способностей: 54 с левой границей отклонения $l_1^L = 9$, правой границей отклонения – $l_1^R = 11$ и 80 с левой границей отклонения $l_1^L = 13$, правой границей – $l_1^R = 15$. В итоге, получаем конечные границы отклонений: $l^L \approx 11, l^R \approx 13$. Таким образом, найденный максимальный суммарный поток минимальной стоимости с учетом реверсирования движения может быть представлен нечетким треугольным числом (59, 70, 83) единиц.

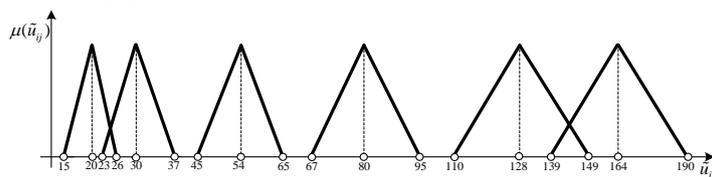


Рис. 9. Функции принадлежности базовых значений пропускных способностей ребер исходной сети

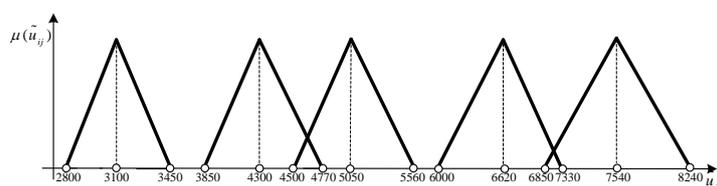


Рис. 10. Функции принадлежности базовых значений стоимостей ребер исходной сети

Нечеткая стоимость «около 6786» расположена между двумя соседними базовыми значениями стоимостей перевозок: 6620 с левой границей отклонения $l_1^L = 620$, правой границей отклонения – $l_1^R = 710$ и 7540 с левой границей отклонения $l_1^L = 690$, правой границей – $l_1^R = 700$. В итоге, получаем конечные границы отклонений: $l^L \approx 632, l^R \approx 708$. Таким образом, найденный максимальный суммарный поток минимальной стоимости с учетом реверсирования движения может быть представлен нечетким треугольным числом (6154, 6786, 7494) единиц.

Временная сложность алгоритма. Для вычисления временной сложности алгоритма была разработана программная реализация на языке C#. Оценка временной сложности разработанного алгоритма производилась с помощью измерения времени работы его программной реализации для различных входных данных. При этом оценка временная сложность проводилась по количеству ребер графа, учитывающему связи между элементами графа.

На рис. 11 приведен график зависимости времени работы алгоритма от количества рёбер в графе.

При построении графика на рис. 11 в качестве времени работы алгоритма был взят 90-й перцентиль для пятидесяти итераций.



Рис. 11. Временная сложность алгоритма

Заключение. Статья иллюстрирует подход к нахождению потока минимальной стоимости в нечеткой транспортной сети с учетом реверсирования движения по дугам графа. Модель представляет собой нечеткую транспортную сеть с заданными нечеткими пропускными способностями и стоимостями перевозок. Техника реверсирования движения применяется для увеличения результирующего потока путем увеличения исходных пропускных способностей участков пути для разрешения движения в обратную сторону. Правило оперирования нечеткими значениями пропускных способностей и стоимостей, которое не приводит к сильному размытию границ чисел обеспечивает более реалистичные и чувствительные к изменениям окружающей среды решения. Предложенный подход имеет важное практическое применение при транспортном планировании и эвакуационном моделировании и оптимизации перевозок, когда необходимо перевозить грузы по оптимальным с точки зрения стоимости маршрутам или при моделировании эвакуационного сценария, когда необходимо перевести пострадавших из опасных зон в безопасные наиболее быстро, принимая параметры времени в качестве стоимостей перевозок. В будущем планируется разработка алгоритмов нахождения динамических нечетких потоков, допускающих частичное реверсирование движения в транспортных сетях.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Bozhenyuk A., Gerasimenko E., Kacprzyk J., Rozenberg I.* Flows in networks under fuzzy conditions // *Studies in Fuzziness and Soft Computing*, – Heidelberg: Springer-Verlag, 2017. – Vol. 346. – P. 269-291.
2. *Zadeh L.A.* Fuzzy sets // *Information and Control*. – 1965. – Vol. 8 (3). – P. 338-353.
3. *Novák V., Perfilieva I., Močkoř J.* Mathematical principles of fuzzy logic. – Dordrecht: Kluwer Academic, 1999.
4. *Rebennack S., Arulselvan A., Elefteriadou L., Pardalos P.M.* Complexity analysis of maximum flow problem with arc reversal // *Journal of Combinatorial Optimization*. – 2010. – Vol. 29. – P. 200-216.
5. *Cova T., Johnson J.P.* A network flow model for lane-based evacuation routing // *Transp. Res. Part Policy Pract.* – 2003. – Vol. 37. – P. 579-604.
6. *Pyakurel U., Nath H.N., Dempe S., Dhamala T.N.* Efficient Dynamic Flow Algorithms for Evacuation Planning Problems with Partial Lane Reversal // *Mathematics*. – 2019. – Vol. 7 (10). – P. 1-29.
7. *Pyakurel U., Dhamala T.N., Dempe S.* Efficient continuous contraflow algorithms for evacuation planning problems // *Ann. Oper. Res.* – 2017. – Vol. 254. – P. 335-364.

8. Köhler E., Langkau, K. Skutella, M. Time expanded graphs for flow depended transit times // European Symposium on Algorithms. – Springer: Berlin/Heidelberg, Germany, 2002. – P. 599-611.
9. Dhamala T.N. A survey on models and algorithms for discrete evacuation planning network problems // J. Ind. Manag. Optim. – 2015. – Vol. 11. – P. 265-289.
10. Боженюк А. Герасименко Е. Розенберг И. Алгоритм нахождения нечеткого потока в транспортной сети с нечеткими стоимостями и пропускными способностями // Известия ЮФУ. Технические науки. – 2012. – № 5 (130). – С. 118-122.
11. Герасименко Е. Нахождение потока минимальной стоимости в транспортной сети методом ранжирования математического ожидания нечетких функций стоимостей // Известия ЮФУ. Технические науки. – 2012. – № 4 (129). – С. 247-251.
12. Kureichik V., Gerasimenko E. Approach to the Minimum Cost Flow Determining in Fuzzy Terms Considering Vitality Degree. In Silhavy R., Senkerik R., Kominkova Oplatkova Z., Prokopova Z., Silhavy P. (eds) Artificial Intelligence Trends in Intelligent Systems. CSOC 2017. Advances in Intelligent Systems and Computing. – Springer, Cham, 2017. – Vol. 573. – P. 200-209.
13. Боженюк А.В., Герасименко Е.М., Розенберг И.Н. Разработка метода определения потока минимальной стоимости в транспортной сети с нечеткими пропускными способностями и стоимостями с помощью метода потенциалов // Известия ЮФУ. Технические науки. – 2015. – № 6 (167). – С. 138-149.
14. Bozhenyuk A., Gerasimenko E., Rozenberg I. Determining the Minimum Cost Flow in Fuzzy Dynamic Network with GIS «ObjectLand» // Proceedings of 9th International Conference on Application of Information and Communication Technologies, AICT 2015, Rostov on Don, Russia. – P. 294-298.
15. Gerasimenko E., Rozenberg I. Earliest arrival dynamic flow model for emergency evacuation in fuzzy conditions // IOP Conference Series: Materials Science and Engineering. – 2020. – Vol. 734 (1). – P. 1-6.
16. Gerasimenko E., Kureichik V. The Maximum Lexicographic Contraflow Finding in a Fuzzy Dynamic Network // Advances in Intelligent Systems and Computing. – 2021. – Vol. 1197. – P. 981-989.
17. Dijkstra E. A note on two problems in connection with graphs // Numerische Mathematik. – 1959. – Vol. 1. – P. 269-271.
18. Ganesan K., Veeramani P. Fuzzy Linear Programs with Trapezoidal Fuzzy Numbers // Ann Oper Res. – 2006. – P. 305-315.
19. Kumar A., Kaur J., Singh P. Fuzzy Optimal Solution of Fully Fuzzy Linear Programming Problems with Inequality Constraints // International Journal of Mathematical and Computer Sciences. – 2010. – 6:1. – P. 37-41.

REFERENCES

1. Bozhenyuk A., Gerasimenko E., Kacprzyk J., Rozenberg I. Flows in networks under fuzzy conditions, *Studies in Fuzziness and Soft Computing*. Heidelberg: Springer-Verlag, 2017, Vol. 346, pp. 269-291.
2. Zadeh L.A. Fuzzy sets, *Information and Control*, 1965, Vol. 8 (3), pp. 338-353.
3. Novák V., Perfilieva I., Močkoř J. Mathematical principles of fuzzy logic. Dordrecht: Kluwer Academic, 1999.
4. Rebennack S., Arulselman A., Elefteriadou L., Pardalos P.M. Complexity analysis of maximum flow problem with arc reversal, *Journal of Combinatorial Optimization*, 2010, Vol. 29, pp. 200-216.
5. Cova T., Johnson J.P. A network flow model for lane-based evacuation routing, *Transp. Res. Part Policy Pract.*, 2003, Vol. 37, pp. 579-604.
6. Pyakurel U., Nath H.N., Dempe S., Dhamala T.N. Efficient Dynamic Flow Algorithms for Evacuation Planning Problems with Partial Lane Reversal, *Mathematics*, 2019, Vol. 7 (10), pp. 1-29.
7. Pyakurel U. Dhamala T.N. Dempe S. Efficient continuous contraflow algorithms for evacuation planning problems, *Ann. Oper. Res.*, 2017, Vol. 254, pp. 335-364.
8. Köhler E., Langkau, K. Skutella, M. Time expanded graphs for flow depended transit times, *European Symposium on Algorithms*. Springer: Berlin/Heidelberg, Germany, 2002, pp. 599-611.

9. *Dhamala T.N.* A survey on models and algorithms for discrete evacuation planning network problems, *J. Ind. Manag. Optim.*, 2015, Vol. 11, pp. 265-289.
10. *Bozhenyuk A., Gerasimenko E., Rozenberg I.* Algoritm nakhozhdeniya nechetkogo potoka v transportnoy seti s nechetkimi stoimostyami i propusknyimi sposobnostyami [Algorithm of the fuzzy flow finding in the transportation network with fuzzy costs and arc capacities], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2012, No. 5 (130), pp. 118-122.
11. *Gerasimenko E.* Nakhozhdenie potoka minimal'noy stoimosti v transportnoy seti metodom ranzhirovaniya matematicheskogo ozhidaniya nechetkikh funktsiy stoimostey [Minimum cost flow finding in the network by the method of expectation ranking of fuzzy cost functions], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2012, No. 4 (129), pp. 247-251.
12. *Kureichik V., Gerasimenko E.* Approach to the Minimum Cost Flow Determining in Fuzzy Terms Considering Vitality Degree. In Silhavy R., Senkerik R., Kominkova Oplatkova Z., Prokopova Z., Silhavy P. (eds) *Artificial Intelligence Trends in Intelligent Systems. CSOC 2017. Advances in Intelligent Systems and Computing*. Springer, Cham, 2017, Vol. 573, pp. 200-209.
13. *Bozhenyuk A.V., Gerasimenko E.M., Rozenberg I.N.* Razrabotka metoda opredeleniya potoka minimal'noy stoimosti v transportnoy seti s nechetkimi propusknyimi sposobnostyami i stoimostyami s pomoshch'yu metoda potentsialov [Development of the minimum cost flow method with fuzzy arc capacities and costs by the potential method], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, No. 6 (167), pp. 138-149.
14. *Bozhenyuk A., Gerasimenko E., Rozenberg I.* Determining the Minimum Cost Flow in Fuzzy Dynamic Network with GIS «ObjectLand», *Proceedings of 9th International Conference on Application of Information and Communication Technologies, AICT 2015, Rostov on Don, Russia*, pp. 294-298.
15. *Gerasimenko E., Rozenberg I.* Earliest arrival dynamic flow model for emergency evacuation in fuzzy conditions, *IOP Conference Series: Materials Science and Engineering*, 2020, Vol. 734 (1), pp. 1-6.
16. *Gerasimenko E., Kureichik V.* The Maximum Lexicographic Contraflow Finding in a Fuzzy Dynamic Network, *Advances in Intelligent Systems and Computing*, 2021, Vol. 1197, pp. 981-989.
17. *Dijkstra E.* A note on two problems in connection with graphs, *Numerische Mathematik*, 1959, Vol. 1, pp. 269-271.
18. *Ganesan K., Veeramani P.* Fuzzy Linear Programs with Trapezoidal Fuzzy Numbers, *Ann Oper Res.*, 2006, pp. 305-315.
19. *Kumar A., Kaur J., Singh P.* Fuzzy Optimal Solution of Fully Fuzzy Linear Programming Problems with Inequality Constraints, *International Journal of Mathematical and Computer Sciences*, 2010, 6:1, pp. 37-41.

Статью рекомендовал к опубликованию д.т.н., профессор А.В. Боженюк.

Герасименко Евгения Михайловна – Южный федеральный университет; e-mail: egerasimenko@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371651; кафедра систем автоматизированного проектирования; доцент.

Gerasimenko Evgeniya Michailovna – Southern Federal University; e-mail: egerasimenko@sfedu.ru; 44, Nekrasovskiy lane, Taganrog, 347928, Russia, phone: +78634371621; the department of computer-aided design; associate professor.

В.А. Деркачев, В.В. Бахчевников, А.Н. Бакуменко

**КЛАССИФИКАТОР ИЗОБРАЖЕНИЙ СЕМЯН
СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР С ИСПОЛЬЗОВАНИЕМ
СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ**

В настоящей статье рассматривается создание архитектуры сверточной нейронной сети, классифицирующей изображения сельскохозяйственных культур (в частности пшеницы) для последующего применения в оптическом сепараторе семян (фотосепараторе). Интерес к проектированию нейронных сетей классификации изображений в последнее время сильно возрос, что связано как с развитием теории глубоких нейронных сетей, так и возросшей вычислительной мощностью настольных компьютеров, а также переносом вычислений на графические процессоры. Целью статьи является разработка архитектуры нейронной сети позволяющей осуществить разделение входного потока семян пшеницы на два класса: «хорошие» семена и «плохие» (с изъянами по форме и цвету) семена. Архитектура полученной нейронной сети является сверточной, так как в отличие от полносвязной, данный класс нейронных сетей в определенных пределах невосприимчив к изменению масштаба и угла поворота объектов во входных данных. В работе для формирования обучающей, валидационной и тестовой выборки использовались изображения семян, полученные с использованием бытовой фотокамеры, что негативно сказалось на результатах обучения и тестирования нейронной сети относительно возможного результата применения в реальном фотосепараторе. Архитектура разработанной нейронной сети предварительно оптимизирована для использования на ПЛИС, однако, в рассмотренном случае не осуществлен переход от значений весовых коэффициентов из типа данных с плавающей запятой к целочисленному типу, что может привести к снижению точности работы нейронной сети, при этом позволив значительно уменьшить объем ресурсов ПЛИС. Применение предложенной архитектуры позволяет получить достаточно точную оценку классифицируемых семян пшеницы по верификационным и тестовым наборам данных.

Нейронные сети; сверточные нейронные сети; классификация изображений; классификация семян пшеницы.

V.A. Derkachev, V.V. Bakhchevnikov, A.N. Bakumenko

**CLASSIFIER OF IMAGES OF AGRICULTURAL CROPS SEEDS USING
A CONVOLUTION NEURAL NETWORK**

This article discusses the creation of a convolutional neural network architecture that classifies images of crops (in particular wheat) for subsequent use in an optical seed separator (photo separator). Interest in the design of neural networks for classifying images has recently increased significantly, which is associated both with the development of the theory of deep neural networks and the increased computing power of desktop computers, as well as the transfer of computing to graphic processors. The aim of the article is to develop the architecture of a neural network that allows the separation of the input flow of wheat seeds into two classes: “good” seeds and “bad” (with defects in shape and color) seeds. The architecture of the resulting neural network is convolutional, because, unlike a fully connected one, this class of neural networks is within certain limits immune to changes in the scale and angle of rotation of objects in the input data. In the work, for the formation of training, validation and test samples, seed images obtained using a household camera were used, which negatively affected the results of training and testing the neural network regarding the possible result of application in a real photo separator. The architecture of the developed neural network is preliminarily optimized for use on FPGAs, however, in the considered case, the transition from the values of weighting factors from the data type from a floating point to an integer type has not been made, which can lead to a decrease in the accuracy of the neural network, while significantly reducing the amount of resources FPGA. Application of the proposed architecture allows one to obtain a fairly accurate estimate of classified wheat seeds from verification and test data sets.

Neural networks; convolutional neural networks; image classification; wheat seed classification.

Введение. Основной проблемой классификации изображений является выявление особенностей, с помощью которых возможно устойчивое определение класса изображений. Одним из методов классификации изображений является корреляционный метод [1]. Так же в качестве примера классического классификатора можно упомянуть статистически оптимальные байесовские классификаторы. В таких классификаторах все наблюдаемые объекты, описываются с применением специального набора признаков с использованием байесовского правила классификации [2, 3]. Классические подходы требуют создание признаков вручную для каждого класса объектов, что весьма трудозатратно, особенно при разработке классификаторов сложных объектов с большим количеством признаков. Помимо трудностей ручного задания признаков существует проблема малой скорости подстройки классификаторов под конкретные задачи. Данная проблема связана с наличием большого числа вариантов признаков объектов, которые необходимо учесть для правильной классификации объектов. Перечисленные проблемы увеличивают финансово-временные затраты, связанные с разработкой традиционных классификаторов под каждый из возможных классов объектов.

Использование нейронных сетей для классификации изображений является многообещающим направлением относительно классических методов обладая следующими преимуществами: возможность работать с большими или неточными данными, выделяя особенности, которые не очевидны для человека.

Развитие сверточных нейронных сетей со времени работ Ю. Лекуна [4] в последнее десятилетие сильно продвинулось в реализуемой точности и сложности моделей [5, 6], что значительно расширило область применения этой технологии.

Целью данной работы является разработка нейросетевой архитектуры классификатора изображений семян сельскохозяйственных культур (в данном случае, пшеницы) для дальнейшего применения в оптическом сортировочном устройстве (фотосепараторе). Из-за различий в ориентации семян и их размера во входных оптических изображениях предлагается глубокая сверточная нейронная сеть в качестве архитектуры классификатора. Данная архитектура построения нейронной сети уменьшает негативные эффекты, возникающие из-за возможного низкого качества входных изображений, их ориентации и масштаба.

Архитектура сверточной нейронной сети. В этой статье была использована сверточная нейронная сеть. Структура нейронной сети подразумевает наличие различного типа слоев: входного слоя, нескольких скрытых слоев и выходного слоя. Охарактеризуем используемые виды слоев:

- ◆ **Входной слой.** В этом слое каждому нейрону соответствует отдельная точка изображения с учетом количества цветовых каналов.

- ◆ **Сверточный слой.** Осуществляет двумерную свертку значений из предыдущего уровня, используя набор ядер свертки (фильтров). Обычно в каждом слое содержится определенный набор различных ядер свертки, что необходимо для выделения нескольких признаков в одном слое. Значения коэффициентов фильтров перед обучением устанавливаются случайным образом или с использованием специальных алгоритмов (обычно алгоритмов Ксавье [7] или Хе [8]), что важно для возможности достижения глобального минимума в процессе обучения. Далее каждое из ядер свертки обучается для выделения определенных признаков из массива данных, полученных из предыдущего слоя. На выходе сверточного слоя, в случае отсутствия дополнения краевыми точками, размерность массива данных уменьшается в соответствии со следующим правилом: $(n - k + 1)$, где n – является размерность массива данных на входе, k – размерность используемого ядра свертки.

- ◆ **Слой субдискретизации.** В данном слое происходит уменьшение размерности данных (обобщение) путем объединения значений входной матрицы данных в рамках определенной области в одно значение. Этот слой позволяет увеличить

устойчивость нейронной сети к изменению масштаба изображений, а также уменьшает число требуемых вычислений в последующих слоях нейронной сети. Существует несколько возможных вариантов субдискретизации: по максимуму [9, 10] (самый популярный вариант в настоящее время), по среднему значению [11] и промежуточные варианты [12]. В данной архитектуре применена субдискретизация по максимальному значению.

◆ Полносвязный слой. Нейроны в данном слое связаны с каждым из предыдущего и осуществляют конечное разделение на классы.

В случае классификатора семян сельскохозяйственных культур можно предложить следующую архитектуру сети:

1. Входное изображение размерностью $3 \times 91 \times 91$ (3 канала цвета при разрешении 91×91);

2. Сверточный слой с ядрами свертки 5×5 в количестве 10 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность $10 \times 87 \times 87$, функция активации ReLU;

3. Слой субдискретизации с окном 2×2 и шагом 2. Выходная размерность слоя составила $10 \times 43 \times 43$;

4. Сверточный слой с ядрами свертки 5×5 в количестве 10 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность $10 \times 39 \times 39$, функция активации ReLU;

5. Слой субдискретизации с окном 2×2 и шагом 2. Выходная размерность слоя составила $10 \times 19 \times 19$;

6. Сверточный слой с ядрами свертки 3×3 в количестве 20 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность $20 \times 17 \times 17$, функция активации ReLU;

7. Сверточный слой с ядрами свертки 3×3 в количестве 20 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность $20 \times 15 \times 15$, функция активации ReLU;

8. Слой субдискретизации с окном 2×2 и шагом 2. Выходная размерность слоя составила $20 \times 7 \times 7$;

9. Сверточный слой с ядрами свертки 3×3 в количестве 20 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность $20 \times 5 \times 5$, функция активации ReLU;

10. Сверточный слой с ядрами свертки 3×3 в количестве 20 штук. Так как изображение не дополняется по краям дополнительно, выходная карта свертки имеет размерность $20 \times 3 \times 3$, функция активации ReLU;

11. Полносвязный слой с числом нейронов 128, функция активации ReLU;

12. Полносвязный слой с числом нейронов 32, функция активации ReLU;

13. Выходной слой с двумя нейронами, соответствующими двум классифицируемыми классам, функция активации гиперболический тангенс.

Полученная архитектура имеет 43 312 весовых коэффициентов. На рис. 1 схематично показана итоговая архитектура нейронной сети.

Параметры модели и обучения. Одним из важнейших параметров нейросети является функция активации нейронов. В данной работе используется два вида функции активации: в большинстве слоев функция «линейный выпрямитель» ReLU (rectified linear unit) и гиперболический тангенс для выходного слоя. Выбор функции активации ReLU в качестве основной был сделан исходя из следующих соображений: данная функция ускоряет обучение нейронной сети [13], приводит к лучшим решениям [14, 15], чем обычные сигмоидоподобные функции, при этом в отличие от гиперболического тангенса и сигмоиды имеет малую вычислительную сложность, что принципиально важно для дальнейшей реализации разрабатываемой

мой нейросети на ПЛИС. Гиперболический тангенс, выбранный в качестве функции активации выбран из-за подходящего диапазона значений (-1, 1) и хороших результатов в применении в нейронных сетях [16, 17].

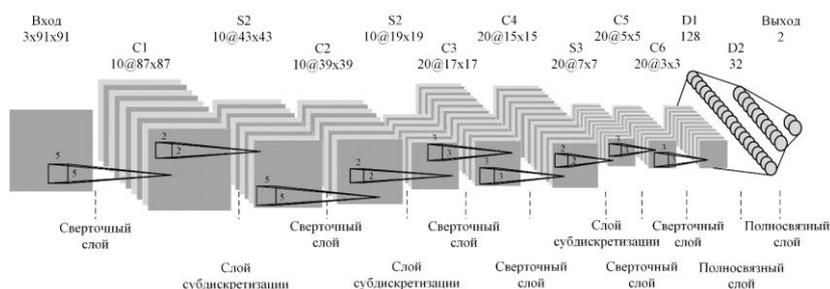


Рис. 1. Архитектура нейросети классификации семян

Инициализация весов в данной нейронной сети выполнена с применением метода Хе, который достаточно хорошо подходит для использования с функцией активации ReLU [8]. Для решения проблемы переобучения применена регуляризация L_2 [18] (частный случай регуляризации Тихонова [19]), со значением коэффициента регуляризации равно $\lambda=10^{-4}$.

В качестве метода градиентной оптимизации стохастических целевых функций, применяемого в качестве обучающего нейронную сеть, используется алгоритм Adam, основанный на адаптивных оценках моментов более низкого порядка [20].

Тестирование. Для обучения и тестирования полученной архитектуры нейронной сети был создан набор из фотографий семян пшеницы разрешением 91x91 пикселей, разделенный на два класса «хорошие» и «плохие» семена. Изображения для данной предварительной выборки получены путем фотографирования отобранных вручную семян с использованием бытовой камеры. Пример характерных изображения семян из каждого классифицируемых наборов данных показаны на рисунках ниже (рис. 2, 3).



Рис. 2. Пример обучающей выборки «хороших» семян



Рис. 3. Пример обучающей выборки «плохих» семян

В обоих классах были выделены следующие подклассы: обучающий подкласс, состоящий из 400 изображений, верификационный подкласс, состоящий из 68 изображений и проверочный, состоящий из 50 изображений. Всего получаем следующий набор данных: тренировочные изображения – 800 штук, верификационные изображения – 136 штук и тестовые изображения – 100 штук. Все используемые данные для каждого этапа обучения берутся в случайном порядке, за исключением тестовых данных, которые сохраняют исходную последовательность для корректного последующего сравнения результатов обучения нейронной сети.

Обучение и тестирование нейронной сети произведено в Tensorflow версии 2.1 установленной в системе Ubuntu 18.04 LTS. Для обучения нейронной сети применялась система с графическим процессором Nvidia 2070 Super, что позволило осуществлять обучение нейронной сети за 25-27 секунд. Графики точности нейронной сети и функции потерь в процессе обучения показаны на рис. 4.

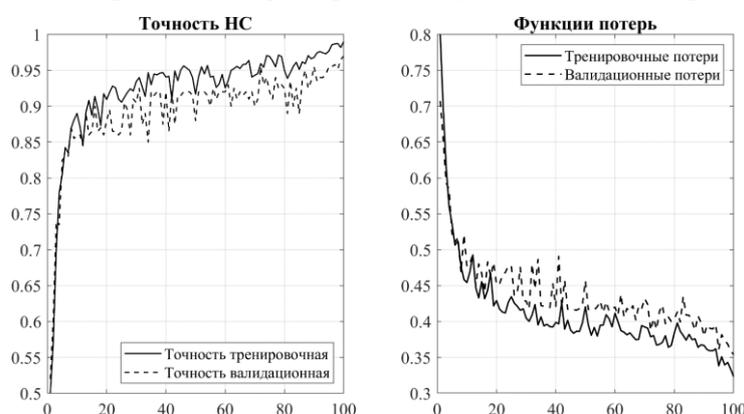


Рис. 4. Графики процесса обучения нейронной сети

Результаты обучения (100 этапов по 8 изображений) показаны в табл. 1.

Таблица 1

Параметр	Потери тренировочные	Потери валидационные	Точность тренировочная	Точность валидационная	Точность тестовая
Значение	0,3233	0,3540	0,9900	0,9600	0,9700

Выводы. Исходя из результатов показанных в табл. 1, точность классификации на основе предложенной архитектуры глубокой сверточной нейронной сети на тестовой выборке достигла 97 %. Однако, следует учесть тот факт, что обучающая и тестовая выборки достаточно сильно ограничены в объеме (800 и 100 соответственно). Данные ограничения, возможно, могут привести к некоторой ошибке оценки правильности классификации изображений. Помимо малого объема данной выборки следует отметить, что при реальном применении данного алгоритма в фотосепараторе ожидается увеличение вероятности правильной классификации из-за отсутствия теней на изображении с семенами, а также отсутствием фона. Результаты получены с использованием типа данных с плавающей запятой, для применения результатов на ПЛИС желательно осуществить переход к целочисленному типу данных (INT8) [21], так как полученная нейронная сеть обладает значительным количеством весовых коэффициентов (43 312), что может затруднить реализацию даже на самых больших ПЛИС настоящего времени.

Научные исследования выполнены в рамках проекта «Создание высокотехнологического производства аппаратно-программных комплексов обработки сельхозсырья на основе СВЧ излучения» (Соглашение с Минобрнауки РФ № 075-11-2019-083 от 20.12.2019 г., Договор ЮФУ № 18 от 20.09.2019 г., номер работы в ЮФУ № ХД/19-25-РТ).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений: пер. с англ. – М.: Техносфера, 2005. – 1073 с.
2. Anderson T.W., Goodman L.A. Statistical inference about Markov chains // The Annals of Mathematical Statistics. – 1957. – P. 89-110.
3. Вагис А.А. Процедура распознавания на байесовских сетях // Компьютерная математика. – 2010. – № 2. – С. 124-130.
4. LeCun Y. et al. Gradient-based learning applied to document recognition // Proceedings of the IEEE. – 1998. – Vol. 86, Issue 11. – P. 2278-2324.
5. Gu J. et al. Recent advances in convolutional neural networks // Pattern Recognition. – 2018. – Vol. 77. – P. 354-377.
6. Srinivas S. et al. A taxonomy of deep convolutional neural nets for computer vision // Frontiers in Robotics and AI. – 2016. – Vol. 2. – P. 36.
7. Glorot X., Bengio Y. Understanding the difficulty of training deep feedforward neural networks // Proceedings of the thirteenth international conference on artificial intelligence and statistics. – 2010. – P. 249-256.
8. He K. et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification // Proceedings of the IEEE international conference on computer vision. – 2015. – P. 1026-1034.
9. LeCun Y. et al. Backpropagation applied to handwritten zip code recognition // Neural computation. – 1989. – Vol. 1, Issue 4. – P. 541-551.
10. Boureau Y.L. et al. Learning mid-level features for recognition // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – IEEE, 2010. – P. 2559-2566.
11. Scherer D., Müller A., Behnke S. Evaluation of pooling operations in convolutional architectures for object recognition // International conference on artificial neural networks. – Springer, Berlin, Heidelberg, 2010. – P. 92-101.
12. Boureau Y.L., Ponce J., LeCun Y. A theoretical analysis of feature pooling in visual recognition // Proceedings of the 27th international conference on machine learning (ICML-10). – 2010. – P. 111-118.
13. Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks // Advances in neural information processing systems. – 2012. – P. 1097-1105.
14. Maas A.L., Hannun A.Y., Ng A.Y. Rectifier nonlinearities improve neural network acoustic models // Proc. icml. – 2013. – Vol. 30, Issue No. 1. – P. 3.
15. Zeiler M.D. et al. On rectified linear units for speech processing // 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. – IEEE, 2013. – P. 3517-3521.
16. Kalman B.L., Kwasny S.C. Why tanh: choosing a sigmoidal function // [Proceedings 1992] IJCNN International Joint Conference on Neural Networks. – IEEE, 1992. – Vol. 4. – P. 578-581.
17. Marra S., Iachino M.A., Morabito F.C. Tanh-like activation function implementation for high-performance digital neural systems // 2006 Ph. D. Research in Microelectronics and Electronics. – IEEE, 2006. – P. 237-240.
18. Cortes C., Mohri M., Rostamizadeh A. L2 regularization for learning kernels // arXiv preprint arXiv:1205.2653. – 2012.
19. Тихонов А.Н. О некорректных задачах линейной алгебры и устойчивом методе их решения // ДАН СССР. – 1965. – Т. 163, № 3. – С. 591-594.
20. Kingma D.P., Ba J. Adam: A method for stochastic optimization // arXiv preprint arXiv:1412.6980. – 2014.
21. Wang J. et al. Design flow of accelerating hybrid extremely low bit-width neural network in embedded FPGA // 2018 28th International Conference on Field Programmable Logic and Applications (FPL). – IEEE, 2018. – P. 163-1636.

REFERENCES

1. *Gonzalez R., Vuds R.* TSifrovaya obrabotka izobrazheniy [Digital image processing]: transl. from engl. Moscow: Tekhnosfera, 2005, 1073 p.
2. *Anderson T.W., Goodman L.A.* Statistical inference about Markov chains, *The Annals of Mathematical Statistics*, 1957, pp. 89-110.
3. *Vagis A.A.* Protsedura raspoznavaniya na bayesovskikh setyakh, *Komp'yuternaya matematika*, 2010, No. 2, pp. 124-130.
4. *LeCun Y. et al.* Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, 1998, Vol. 86, Issue 11, pp. 2278-2324.
5. *Gu J. et al.* Recent advances in convolutional neural networks, *Pattern Recognition*, 2018, Vol. 77, pp. 354-377.
6. *Srinivas S. et al.* A taxonomy of deep convolutional neural nets for computer vision, *Frontiers in Robotics and AI*, 2016, Vol. 2, pp. 36.
7. *Glorot X., Bengio Y.* Understanding the difficulty of training deep feedforward neural networks, *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249-256.
8. *He K. et al.* Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026-1034.
9. *LeCun Y. et al.* Backpropagation applied to handwritten zip code recognition, *Neural computation*, 1989, Vol. 1, Issue 4, pp. 541-551.
10. *Boureau Y.L. et al.* Learning mid-level features for recognition, *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2010, pp. 2559-2566.
11. *Scherer D., Müller A., Behnke S.* Evaluation of pooling operations in convolutional architectures for object recognition, *International conference on artificial neural networks*. Springer, Berlin, Heidelberg, 2010, pp. 92-101.
12. *Boureau Y.L., Ponce J., LeCun Y.* A theoretical analysis of feature pooling in visual recognition, *Proceedings of the 27th international conference on machine learning (ICML-10)*, 2010, pp. 111-118.
13. *Krizhevsky A., Sutskever I., Hinton G.E.* Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, 2012, pp. 1097-1105.
14. *Maas A.L., Hannun A.Y., Ng A.Y.* Rectifier nonlinearities improve neural network acoustic models, *Proc. icml.*, 2013, Vol. 30, Issue No. 1, pp. 3.
15. *Zeiler M.D. et al.* On rectified linear units for speech processing, *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 3517-3521.
16. *Kalman B.L., Kwasy S.C.* Why tanh: choosing a sigmoidal function, *[Proceedings 1992] IJCNN International Joint Conference on Neural Networks*. IEEE, 1992, Vol. 4, pp. 578-581.
17. *Marra S., Iachino M.A., Morabito F.C.* Tanh-like activation function implementation for high-performance digital neural systems, *2006 Ph. D. Research in Microelectronics and Electronics*. IEEE, 2006, pp. 237-240.
18. *Cortes C., Mohri M., Rostamizadeh A.* L2 regularization for learning kernels, *arXiv preprint arXiv:1205.2653*, 2012.
19. *Tikhonov A.N.* O nekorrektnykh zadachakh lineynoy algebrы i ustoychivom metode ikh resheniya [On ill – posed linear algebra problems and a stable method for solving them], *DAN SSSR [Reports of the Academy of Sciences of USSR]*, 1965, Vol. 163, No. 3, pp. 591-594.
20. *Kingma D.P., Ba J.* Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*, 2014.
21. *Wang J. et al.* Design flow of accelerating hybrid extremely low bit-width neural network in embedded FPGA, *2018 28th International Conference on Field Programmable Logic and Applications (FPL)*. IEEE, 2018, pp. 163-1636.

Статью рекомендовал к опубликованию д.т.н. А.В. Ковалев.

Деркачев Владимир Александрович – Южный федеральный университет; e-mail: vderkachev@sfedu.ru; 347922, г. Таганрог, ул. Шевченко, 2; тел.: +79614154733; кафедра радиотехнических и телекоммуникационных систем; ассистент.

Бахчевников Валентин Владимирович – e-mail: bahchevnikov@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский; 44; тел.: +79518289271; кафедра радиотехнических и телекоммуникационных систем; ассистент.

Бакуменко Алексей Николаевич – e-mail: baku@sfedu.ru; 347900, г. Таганрог, ул. Петровская, 81; тел.: +79886031853; Инжиниринговый центр приборостроения радио и микроэлектроники; инженер.

Derkachev Vladimir Aleksandrovich – Southern Federal University; e-mail: vderkachev@sfedu.ru; 2, Shevchenko street, Taganrog, 347922, Russia; phone +79614154733; the department of radio engineering & telecommunication systems; assistant.

Bakhchevnikov Valentin Vladimirovich – e-mail: bahchevnikov@sfedu.ru; 44, Nekrasovsky lane, Taganrog, 347928, Russia; phone: +79518289271; the department of radio engineering & telecommunication systems; assistant.

Bakumenko Alexey Nikolaevich – e-mail: baku@sfedu.ru; 81, Petrovskaya street, Taganrog, 347900, Russia; phone: +79886031853; Engineering center of instrument making, radio- and microelectronics; engineer.

УДК 0004.912

DOI 10.18522/2311-3103-2020-4-39-48

Е.М. Герасименко, В.В. Степенко

ИНТЕЛЛЕКТУАЛЬНЫЙ МЕТОД ИЗВЛЕЧЕНИЯ ЗНАНИЙ НА ОСНОВЕ ОПРЕДЕЛЕНИЯ ТОНАЛЬНОСТИ ОТЗЫВОВ*

В этой работе исследуется влияние возраста и пола при анализе тональности отзывов, поскольку эти данные могут помочь ретейлерам электронной коммерции увеличить продажи, ориентируясь на определенные демографические группы, а также увеличить удовлетворение потребностей людей разных возрастных и гендерных групп. Используемый набор данных сформирован путем сбора отзывов о книгах. Был создан вопросник, содержащий информацию о предпочтениях книжных носителей (мнения пользователей об электронных книгах, книгах в мягкой и твердой обложках, изображениях и аудиокнигах), а также данные о возрастной группе и гендерной принадлежности. Помимо этого, вопросник также содержит информацию о положительном либо отрицательном мнении касательно предпочтений, которая послужила основой достоверности для классификаторов. В результате, было получено 900 анкет, которые были разделены на группы по половому признаку и возрасту. Каждая конкретная группа данных была разделена на обучающую и тестовую. Были проанализированы сегментированные данные на предмет настроений в зависимости от каждой возрастной группы и пола. Возрастная группа «старше 50 лет» продемонстрировала лучшие результаты по сравнению со всеми другими возрастными группами во всех классификаторах; данные в женской группе показали более высокую точность по сравнению с данными из групп без информации о гендерной принадлежности. Высокие результаты, показанные этими группами, показывают, что подходы к анализу тональности способны предсказать настроения в этих группах лучше, чем в других. Анализ тональности проводился с использованием различных подходов машинного обучения (ML), включая максимальную энтропию, метод опорных векторов, сверточную нейронную сеть и долгую краткосрочную память.

Анализ тональности; машинное обучение; сверточная нейронная сеть; долгая краткосрочная память.

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-07-00099 а.

E.M. Gerasimenko, V.V. Stetsenko

INTELLIGENT METHOD OF KNOWLEDGE EXTRACTION BASED ON SENTIMENT ANALYSIS

The paper explores the impact of age and gender in sentiment analysis, as this data can help e-commerce retailers increase sales by targeting specific demographic groups. The data set used was created by collecting book reviews. A questionnaire was created containing questions about preferences in books, as well as age groups and gender information. The article analyzes segmented data on the subject of moods depending on each age group and gender. Sentiment analysis was performed using various machine learning (ML) approaches, including maximum entropy, support vector method, convolutional neural network, and long short-term memory. This paper investigates the impact of age and gender in sentiment analysis, because this data can help e-commerce retailers to increase sales by targeting specific demographic groups, as well as increase the satisfaction of the needs of people of different age and gender groups. The dataset used is generated by collecting book reviews. A questionnaire was created containing questions about preferences in books (user opinions of e-books, paperbacks, hardbacks, images and audiobooks), as well as data on age group and gender. In addition, the questionnaire also contains information on a positive or negative opinion regarding preferences, which served as the basis for reliability for the classifiers. As a result, 900 questionnaires were received, which were divided into groups according to gender and age. Each specific group of data was divided into training and test one. Segmented data were analyzed for sentiment analysis depending on age group and gender. The age group "over 50 years old" showed the best results in comparison with all other age groups in all classifiers; data in the female group performed higher accuracy compared to data from the groups without gender information. The high scores shown by these groups indicate that sentiment analysis approaches are able to predict moods in these groups better than in others. Sentiment analysis was performed using a variety of machine learning (ML) approaches, including maximum entropy, support vector machines, convolutional neural networks, and long short term memory.

Sentiment analysis; machine learning; convolutional neural network; long short-term memory.

Введение. Рост популярности Интернета привел к огромному притоку данных, содержащих обширную и ценную информацию об общественном мнении. Каждый пользователь интернета, выражающий свое мнение, становится частью информационной цепи, где другие пользователи, изучая публичные обзоры, могут принимать обоснованные решения. Для определения общественного мнения (положительного, отрицательного или нейтрального) используются данные (обзоры, посты, комментарии), собранные из различных социальных сетей, таких как Facebook, Twitter, Amazon, Goodreads, IMDb или блогов. Процесс определения полярности общественного мнения называется анализом тональности. Анализ тональности обычно проводится на основе обзоров фильмов [1, 2], обзоров ресторанов или продуктов питания [3, 4], а также данных из микроблогов [5, 6], что помогает различным организациям улучшить бизнес-стратегии за счет привлечения новых клиентов. Разделение клиентов по возрасту и полу позволяет получить важную информацию, использование которой поможет эффективнее удовлетворить потребности людей разных возрастных и гендерных групп.

Анализ тональности как направление интеллектуального анализа данных [7–10] развивался в течение многих лет, при этом использовались различные методики на основе словарей и машинного обучения. С появлением методов глубокого обучения [11–13] предшествующая информация также сыграла большую роль в адекватном выражении полярности мнений.

Ли и соавторы [14] предложили структуру, обеспечивающую определение полярности мнений при помощи анализа тональности. В предложенном подходе авторы приняли во внимание субъективность мнения и доверие пользователей. Локенхоф и соавторы [15] проанализировали, как разные возрастные группы вы-

ражают свои эмоции. Авторы обнаружили, что пожилые люди описывают положительные эмоции лучше, чем отрицательные, по сравнению с более молодыми людьми. Циммерман и соавторы [16] исследовали тональность в зависимости от возраста и пола [17]. Результаты показали, что люди в подростковом возрасте демонстрировали наименьшую эмоциональную регуляцию по сравнению с другими возрастными группами. Гендерные различия проявлялись как при занижении, так и при переоценке конкретной эмоции. Ох и соавторы [18] изучили сегментацию рынка по полу и возрасту пользователей в целях определения потенциальных направлений путешествий для разных групп в зависимости от их доходов и свободного времени. Кешари и соавторы [19] проанализировали эффективность рекламных объявлений для разных гендерных и возрастных групп на основе того, как потребители реагируют на эти рекламные объявления. Таким образом, основываясь на этих установленных психологических различиях между людьми разного возраста и пола, мы постараемся выяснить, существуют ли эти различия в мнениях, которые люди выражают на онлайн-платформах.

На рис. 1 показана основная блок-схема структуры, используемой для анализа тональности отзывов. Отзывы, собранные из социальных сетей, используются для составления двух наборов данных на основе возраста и пола. Затем в полученных наборах данных выполняется анализ тональности с применением различных подходов анализа тональности.

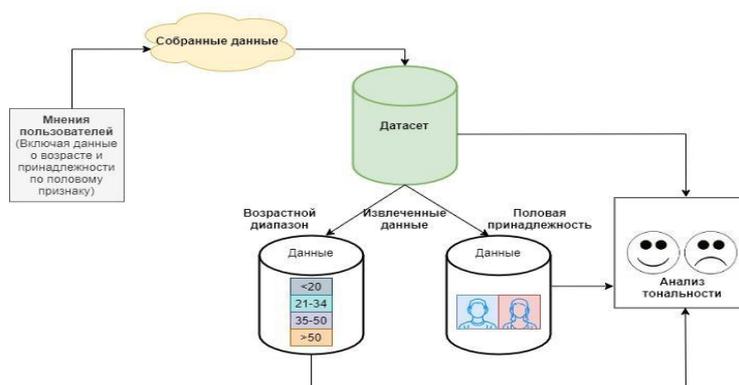


Рис. 1. Блок-схема структуры анализа

Обзор источников. До появления интернета было проведено множество исследований, посвященных тому, как разные люди справляются с разными эмоциями и выражают их. Авторы [20] исследовали гендерные различия при проведении исследования 400 студентов из колледжей, разделенных на пять возрастных групп: от дошкольников до взрослых. Исследование было связано со стереотипами половой и возрастной эмоциональной выразительности. Стоунер и соавторы [21] рассматривали людей обоих полов и разных возрастных групп с целью изучения их способности выражать гнев. Исследование показало, что в группе молодых людей гнев был выражен больше, чем в старших возрастных группах. В данном исследовании автор не обнаружил больших различий по половому признаку в этом аспекте.

Исследование Дэвиса [22] по гендерным различиям в выражении отрицательных эмоций показало, что мальчики проявляли больше отрицательных эмоций, когда были разочарованы, по сравнению с девочками. Броди и соавторы [23] более детально исследовали гендерную и эмоциональную экспрессию и показали, что гендерные различия в эмоциональной выразительности были культурно специфичны среди азиатских иностранных студентов. В исследовании Кринга и соав-

торов [24] группе студентов были продемонстрированы эмоциональные видео, и ученые получили подтверждение того факта, что женщины, как правило, более выразительны, чем мужчины, даже в случае пережитых эмоций. В исследовании Бердитта [25] изучались возрастные и гендерные различия при описании эмоциональных реакций. В эксперименте приняли участие 185 человек (из них – 85 мужчин и 100 женщин) в возрасте от 13 до 99 лет. Результат исследования показал, что подростки и молодые взрослые чаще сообщали о гневе и давали более интенсивные отрицательные ответы, чем группа взрослых мужчин

Методология. Для обработки обзоров выполняются следующие шаги, отображенные на рис. 2.



Рис. 2. Блок-схема, представляющая шаги, предпринятые для анализа тональности

Изначальный набор данных разделен на два подмножества, называемых датасетами (от англ. data set – набор данных), а затем каждый из них дополнительно разделен на группы, в зависимости от конкретного возраста и пола. Каждая конкретная группа данных разделяется на обучающие данные и данные для тестирования. Алгоритм классификации реализуется в конце фазы обучения после предварительной обработки данных и извлечения характеристик и используется на фазе тестирования для получения окончательных результатов. Обзоры пользователей проходят предварительную обработку и извлечение характеристик на этапе тестирования, а затем проходят этап классификации.

Обзоры предварительно обрабатываются с целью удаления ненужной информации, которая не влияет на полярность предложения [26], с помощью этапов, представленных ниже:

1. Удаление знаков препинания, цифр и символов, поскольку они не несут смысловой нагрузки.
2. Удаление стоп-слов, поскольку они не влияют на оценку настроения высказанного мнения.
3. Замена аббревиатур слова реальным словом.
4. Преобразование текста в нижний регистр.
5. Замена смайликов эмоциями, которые выражает смайлик.
6. Токенизация.

Далее происходит извлечение характеристик. И, наконец, алгоритм классификатора предсказывает метку, которая при сравнении с основополагающей достоверностью определяет точность классификатора.

Были собраны данные о предпочтениях людей в отношении книг (твердый переплет, электронные книги или аудиокниги), а также информация о возрасте и поле. В подходе реализуются различные алгоритмы анализа тональности для каждого набора данных в отдельности, а затем результаты сравниваются, и выявляются соответствующие различия среди групп. Также, в собранном наборе данных был реализован подход анализа тональности на основе словаря.

Для извлечения характеристик применяется «мешок слов» [27] (используется в методах NB, ME и SVM), а word2vec создает вектор признаков, используя либо непрерывный пакет слов, либо модель Skip-gram, которая в дальнейшем используется в LSTM и CNN. Методы описаны ниже.

1. «Мешок слов» (*Bag of Words*)

Bag of Words – очень гибкая и простая модель, используемая для извлечения характеристик. Эта модель отслеживает количество вхождений, также называемых частотой каждого слова, используемого в предложении. Кроме того, каждому слову в предложении присваивается конкретная оценка субъективности. Общая оценка тональности предложения складывается из суммы оценок каждого слова.

2. *Word2Vec*

Модель Word2Vec используется для формирования встраивания слов. Эта модель очень полезна, поскольку обнаруживает сходство слов в векторной форме, а не в текстовом формате. Эти сходства обнаруживаются на основе значения слова, угаданного по его прошлым появлениям и ассоциации с другими словами.

3. *Словарь на основе классификатора*

VADER (Valence Aware Dictionary and sEntiment Reasoner) – это лексикон и инструмент анализа настроений на основе правил, который специально настроен на чувства, выраженные в социальных сетях. Каждому слову, присутствующему в предложении, присваивается оценка по значению этого слова в словаре. Рассчитывается итоговая сложная оценка предложения, которая варьируется от -1 до 1. Эта оценка показывает, является ли предложение положительным или отрицательным. Составной балл для каждого предложения в наборе данных объединяется, и для всего документа анализируется средний балл. Чтобы сравнить его с другими подходами машинного обучения, мы преобразуем среднюю оценку в точность, разделив оценку всего документа на общее количество проверок в этом конкретном наборе данных.

Классификаторы на основе машинного обучения. Мы применили пять алгоритмов, основанных на машинном обучении, для определения точности набора данных.

1. Наивный байесовский метод.
2. Максимальная энтропия.
3. Метод опорных векторов (SVM).
4. Долгая краткосрочная память (LSTM).
5. Сверточная нейронная сеть (CNN).

Остановимся подробнее на LSTM и сверточной нейронной сети.

На первом этапе определяется информация, которая будет удалена из ячейки памяти (уравнение 1). Сигмовидный слой выполняет это решение после просмотра предыдущей информации i_{t-1} и текущего ввода c_t . Этот сигмовидный слой выводит число от 0 до 1, которое определяет количество информации, которое необходимо сохранить, основываясь на весе W_0 . 0_t представляет собой выходной сигнал текущей ячейки, а b_0 представляет собой смещение для этой конкретной ячейки.

$$0_t = \sigma(W_0 * [i_{t-1}, c_t] + b_0). \quad (1)$$

Затем определяется новая информация, которая должна быть обновлена в ячейку памяти. Этот процесс происходит в два этапа: сначала задействуется сигмовидный слой для определения значений для обновления и затем коричневый слой для создания вектора новых значений (уравнение 2). n_t обозначает информацию, которая должна быть обновлена на основе веса W_n и смещения b_n , а \tilde{V}_t – это данные, которые должны быть включены в информацию о текущем состоянии (уравнение 3):

$$n_t = \sigma(W_n * [i_{t-1}, c_t] + b_n), \quad (2)$$

$$\tilde{V}_t = \tanh(W_V * [i_{t-1}, c_t] + b_V), \quad (3)$$

Далее эта информация обновляется в следующую ячейку V_t путем умножения старого состояния на 0_t согласно (4).

$$V_t = 0_t * V_{t-1} + n_t * \tilde{V}_t, \quad (4)$$

На заключительном шаге мы снова используется сигмовидный слой, чтобы найти f_t , которая обозначает информацию, которая будет предоставлена в качестве выходных данных на основе веса W_f и смещения b_f (уравнение 5). Слой \tanh обновляет необходимые части и выдает их в качестве выходных данных ячейки (уравнение 6).

$$f_t = \sigma(W_f * [i_{t-1}, c_t] + b_f), \quad (5)$$

$$i_t = f_t * \tanh(V_t), \quad (6)$$

Окончательный вывод из этой ячейки будет служить в качестве предварительной информации для следующей ячейки, чтобы определить ее последующее состояние.

Мы обучили модель CNN [28] с использованием одного слоя на характеристиках, извлеченных из предложений с использованием модели word2vec. Первый слой – это сверточный слой, в котором мы перемещаем несколько фильтров разных размеров по 128 измерениям вложения слов, чтобы создать карту объектов на основе конкретного фильтра. Затем идет слой максимального пуллинга, который сворачивает результаты предыдущего слоя в один длинный вектор признаков. Слой максимального пуллинга находит наиболее заметный вектор объектов из карты объектов, принадлежащий каждому фильтру, который затем полностью передается подключенному слою softmax. Регуляризация отсева выполняется от того, как мы используем слой softmax для классификации результата. Регуляризация случайным образом выбрасывает некоторые скрытые единицы из слоя, чтобы предотвратить совместную адаптацию обучающих данных, которая может привести к перегрузке.

Ход исследования и результаты. В качестве набора данных для этого исследования были использованы обзоры 900 человек. Обзоры представлены в виде вопросника, содержащего несколько вопросов, касающихся предпочтений книжного носителя (мнения пользователя об электронных книгах, книгах в мягкой и твердой обложках, изображениях и аудиокнигах). Кроме того, в вопроснике обсуждается, считает ли пользователь, что цифровые носители, такие как Kindle или электронные книги, могли бы заменить книги в твердом переплете или в мягкой обложке. Вопросы уточняли, нравятся ли пользователю аудиокниги больше других форматов, а также предлагалось кратко изложить собственное мнение. Вопросник содержит мнение пользователя, а также пол и возрастную группу опрошиваемого. Наряду с мнением пользователей, они также содержат информацию о положительном / отрицательном мнении касательно предпочтений, которая служит основой достоверности для классификаторов.

60 % обзоров являются положительными, а остальные 40% – отрицательными. Этот набор данных мы разделили на отдельные группы: сначала по полу (70 % принадлежит женщинам, 30 % – мужчинам), затем по возрастному признаку (40 % принадлежат к возрастной группе до 20 лет, 30 % – возрастной группе 21–34 года, 20 % – возрастной группе 35–50 лет, 10 % – возрастной группе – старше 50 лет).

Мы получили результаты с использованием классификаторов на основе машинного обучения и основанных на словаре. Результаты этих классификаторов выражены в терминах точности [29].

$$\text{Точность} = \frac{\text{количество правильно предсказанных наблюдений}}{\text{общее количество наблюдений}}$$

Влияние возрастного фактора

Извлеченный набор данных, основанный на возрасте, подразделяется следующие группы: первая группа – младше 20 лет, вторая – от 21 до 34 лет, третья – от 35 до 50 лет и последняя – старше 50 лет. Таким образом, всего создано четыре группы, содержащие положительные и отрицательные отзывы от людей каждой конкретной возрастной группы. Другая группа (без информации о возрасте), содержащая обзоры из всех возрастных групп, сформирована для сравнения ее результатов с другими группами, как показано на рис. 3.

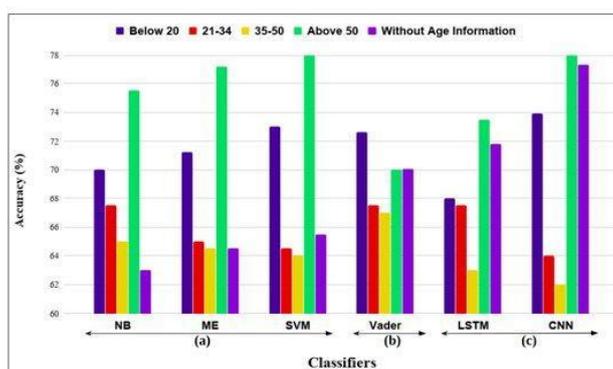


Рис. 3. Сравнение различных методов извлечения характеристик на основе возраста: ((a) – классификаторы, основанные на машинном обучении; (b) – классификаторы на основе словаря; (c) – подходы машинного обучения с использованием метода извлечения признаков word2vec)

Возрастная группа «старше 50 лет» демонстрирует лучшие результаты по сравнению со всеми другими возрастными группами во всех классификаторах с максимальной точностью 78 % в классификаторах CNN и SVM. Возрастная группа «до 20 лет» имеет более высокую точность по сравнению с двумя другими группами среднего возраста, где возрастная группа «21–4» показывает лучший результат, чем другие возрастные группы во всех случаях, хотя разница между этими двумя возрастными группами незначительна. Высокие показатели старших возрастных групп показывают, что подходы анализа тональности способны предсказать настроения в этой возрастной группе лучше, чем в других группах. Группа данных без какой-либо информации о возрасте работает лучше в LSTM и CNN.

Влияние фактора гендерной принадлежности. Исходный набор данных разделен на две группы в зависимости от пола (мужскую и женскую) и содержит положительные и отрицательные отзывы. Данные также проходят этапы предварительной обработки, извлечения характеристик и применения классификаторов. Результаты представлены на рис. 4.

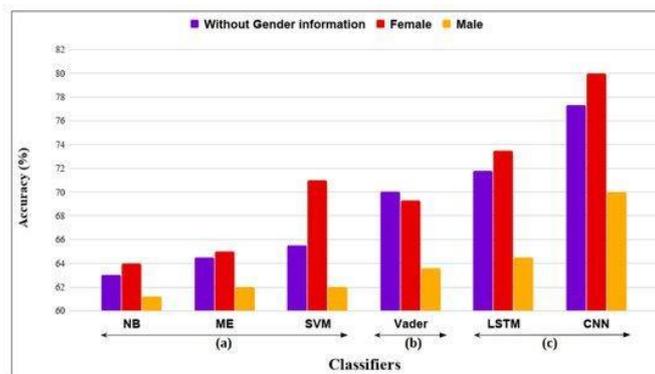


Рис. 4. Сравнение результатов по половому признаку среди различных методов извлечения характеристик и классификаторов (a) – классификаторы, основанные на машинном обучении; (b) – классификаторы на основе словаря; (c) – подходы машинного обучения с использованием метода извлечения признаков word2vec

Хорошо видно, что данные в женской группе имеют более высокую точность по сравнению с данными из групп без гендерной информации и мужской. Данные из женской группы имеют лучшую точность в классификаторе CNN – 80 %, нежели в других классификаторах. Этот результат согласуется с психологическими исследованиями, согласно которым, женщины выражают свое мнение лучше, чем мужчины. Настроения в женских группах легче предсказать, что дает большую точность результата.

Заключение. В этой статье мы сравнили различные методы анализа настроений в наборе данных, состоящим из 900 анкет, содержащих информацию о возрасте и поле пользователя. Чтобы проанализировать влияние возраста и пола на то, как пользователи выражают свое мнение, набор данных был разделен на четыре группы. Для определения тональности отзывов использовались методы машинного обучения и словарные технические приемы. В итоге, женская группа показала лучшую точность по сравнению с мужской, а возрастная группа старше 50 лет имеет более высокую точность, по сравнению со всеми другими возрастными группами.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Manek A.S., Shenoy P.D., Mohan M.C., Venugopal K. Aspect term extraction for sentiment analysis in large movie reviews using gini index feature selection method and svm classifier, *Worldw. Web.*, 2017, No. 20, pp. 135-154.
2. Dos Santos C., Gatti M. Deep convolutional neural networks for sentiment analysis of short texts, *Proceedings of the COLING, the 25th International Conference on Computational Linguistics: Technical Papers, Dublin, Ireland, 23–29 August 2014*, pp. 69-78.
3. Kiritchenko S., Zhu, X., Cherry C., Mohammad S. Nrc-canada-2014: Detecting aspects and sentiment in customer reviews, *Proceedings of the 8th International Workshop on Semantic Evaluation, Dublin, Ireland, 23–24 August 2014*, pp. 437-442.
4. Pontiki M., Galanis D., Papageorgiou H., Manandhar S. Androutopoulos I. Semeval-2015 task 12: Aspect based sentiment analysis, *Proceedings of the 9th International Workshop on Semantic Evaluation, Denver, CO, USA, 4–5 June 2015*, pp. 486-495.
5. Cao D., Ji R., Lin D., Li S. A cross-media public sentiment analysis system for microblog, *Multim. Syst.*, 2016, No. 22, pp. 479-486.
6. Ghosh R., Zhang L., Dekhil M.E., Liu B. Performing sentiment analysis on microblogging data, including identifying a new opinion term therein. US Patent 9,275,041, 1 March 2016.

7. Ullah M.A., Islam M.M., Azman N.B., Zaki Z.M. An overview of multimodal sentiment analysis research: Opportunities and difficulties, *Proceedings of the 2017 IEEE International Conference on Imaging, Vision & Pattern Recognition, Himeji, Japan, 1–3 September 2017*, pp. 1-6.
8. Cambria E. Affective computing and sentiment analysis, *IEEE Intell. Syst.*, 2016, No. 31, pp. 102-107.
9. Liu B. Sentiment analysis and subjectivity, *Handb. Nat. Lang. Proc.*, 2010, No. 2, pp. 627-666.
10. Kumar S., Yadava M., Roy P.P. Fusion of eeg response and sentiment analysis of products review to predict customer satisfaction, *Inf. Fus.*, 2019, No. 52, pp. 41–52.
11. Kim J.H., Kim B.G., Roy P.P. Jeong D.M. Efficient facial expression recognition algorithm based on hierarchical deep neural network structure, *IEEE Access*, 2019, No. 7, pp. 41273-41285.
12. Yoo S.M., Cho C., Lee K.H., Park J., Jin, S., Lee Y., Kim B.G. Structure of deep learning inference engines for embedded systems, *Proceedings of the IEEE 2019 International Conference on Information and Communication Technology Convergence, Kuala Lumpur, Malaysia, 24–26 July 2019*, pp. 920-922.
13. Kim J.H., Hong G.S., Kim B.G., Dogra D.P. Deepgesture: Deep learning-based gesture recognition scheme using motion sensors, *Displays*, 2018, No. 55, pp. 38-45.
14. Li Y.M., Li T.Y. Deriving market intelligence from microblogs, *Decis. Support Syst.*, 2013, No. 55, pp. 206-217.
15. Lockenhoff C.E., Costa P.T., Lane R.D. Age differences in descriptions of emotional experiences in oneself and others, *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.*, 2008, No. 63, pp. 92-99.
16. Zimmermann P., Iwanski A. Emotion regulation from early adolescence to emerging adulthood and middle adulthood, *Int. J. Behav. Dev.*, 2014, No. 38, pp. 182-194.
17. Kaur B., Singh D., Roy P.P. Age and gender classification using brain–computer interface, *Neural Comput. Appl.*, 2019, No. 31, pp. 5887-5900.
18. Oh H., Parks S.C., Demicco F.J. Age-and gender-based market segmentation: A structural understanding, *Int. J. Hosp. Tour. Adm.*, 2002, No. 3, pp. 1-20.
19. Keshari P., Jain S. Effect of age and gender on consumer response to advertising appeals, *Paradigm*, 2016, No. 20, pp. 69-82.
20. Fabes R.A., Martin C.L. Gender and age stereotypes of emotionality, *Personal. Soc. Psychol. Bull.*, 1991, No. 17, pp. 532-540.
21. Stoner S.B., Spencer W.B. Age and gender differences with the anger expression scale, *Educ. Psychol. Meas.*, 1987, pp. 487-492.
22. Davis T.L. Gender differences in masking negative emotions: Ability or motivation?, *Dev. Psychol.*, 1995, No. 31, pp. 660-667.
23. Brody L.R. Gender and emotion: Beyond stereotypes, *J. Soc. Issues*, 2010, No. 53, pp. 369-393.
24. Kring A.M., Gordon A.H. Sex differences in emotion: Expression, experience, and physiology, *J. Personal. Soc. Psychol.*, 1998, Vol. 74, pp. 686-703.
25. Birditt K.S., Fingerman K.L. Age and gender differences in adults' descriptions of emotional reactions to interpersonal problems, *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.*, 2003, No. 58, pp. 237-245.
26. Kharde V.A., Sonawane S.S. Sentiment analysis of Twitter data: A survey of techniques, *Int. J. Comput. Appl.*, 2016, No. 139, pp. 5–15.
27. Saini R., Kumar P., Roy P.P., Pal U. Trajectory classification using feature selection by genetic algorithm, *Proceedings of the 3rd International Conference on Computer Vision and Image Processing*. Springer: Berlin/Heidelberg, Germany, 2020, pp. 377-388.
28. Kim Y. Convolutional neural networks for sentence classification, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014*, pp. 1746-1751.
29. Saito T., Rehmsmeier M. The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets, *PLoS ONE*, 2015, No.10. e0118432.

Статью рекомендовал к опубликованию д.т.н., профессор А.В.-Боженюк.

Герасименко Евгения Михайловна – Южный федеральный университет; e-mail: egerasimenko@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371651; кафедра систем автоматизированного проектирования; доцент.

Стеценко Валентина Витальевна – e-mail: stecenکو@sfedu.ru; кафедра систем автоматизированного проектирования; магистрант.

Gerasimenko Evgeniya Michailovna – Southern Federal University; e-mail: egerasimenko@sfedu.ru; 44, Nekrasovskiy lane, Taganrog, 347928, Russia, phone: +78634371621; the department of computer-aided design; associate professor.

Stetsenko Valentina Vitalievna – e-mail: stecenکو@sfedu.ru; the department of computer-aided design; master student.

УДК 519.712.2

DOI 10.18522/2311-3103-2020-4-48-59

Л.А. Гладков, Н.В. Гладкова**ЭВОЛЮЦИОНИРУЮЩИЕ МНОГОАГЕНТНЫЕ СИСТЕМЫ
И ЭВОЛЮЦИОННОЕ ПРОЕКТИРОВАНИЕ***

Статья посвящена обсуждению проблем построения эволюционирующих мультиагентных систем на основе использования принципов эволюционного проектирования и гибридных моделей. Рассмотрено понятие агента. Представлен набор базовых свойств агента. Рассмотрены аналогии между многоагентными и эволюционными системами. Рассмотрены принципы построения и организации мультиагентных систем. Отмечены сходства между основными определениями теории агентов и теории эволюции. Отмечено, что основные модели эволюции и эволюционные алгоритмы, могут быть с успехом использованы при проектировании многоагентных систем. Проведен анализ существующих методов и методологий проектирования агентов и многоагентных систем. Отмечены существующие различия в подходах к проектированию многоагентных систем. Описаны основные типы моделей и приведены их важнейшие характеристики. Представлена модель взаимодействия агентов, включающая описание услуг (сервисов), взаимосвязей и обязательств, существующих между агентами. Описана модель отношений (контактов), которая задает коммуникационные связи между агентами. Отмечена важность и перспективность использования агентно-ориентированного подхода к проектированию многоагентных систем. Предложена концепция проектирования агентов и многоагентных систем, согласно которой процесс проектирования включает в себя базовые компоненты самоорганизации, в том числе процессы взаимодействия, скрещивания, адаптации к среде и т.д. Рассмотрены различные подходы к эволюционному проектированию искусственных систем. Предложена Эволюционная модель формирования агентов и агентств, как основной компонент эволюционного проектирования. Предложены модифицированные эволюционные операторы кроссинговера для реализации процесса проектирования агентов.

Мультиагентная система; агент; свойства агентов; агентство; эволюционное проектирование; эволюционные операторы; методология проектирования; гибридные методы.

L.A. Gladkov, N.V. Gladkova**EVOLUTIONARY DESIGN AS A TOOL FOR DEVELOPING MULTI-AGENT
SYSTEMS**

The article is devoted to the discussion of the problems of constructing evolving multi-agent systems based on the use of the principles of evolutionary design and hybrid models. The concept of an agent is considered. A set of basic properties of the agent is presented. The analogies between multi-agent and evolutionary systems are considered. The principles of construction and organization of multi-agent systems are considered. The similarities between the main definitions of the theory of agents and the theory of evolution are noted. It that the main evolution models and evolutionary algorithms can be successfully used in the design of multi-agent systems is noted. The analysis of existing methods and

* Исследование выполнено при поддержке РФФИ в рамках научного проекта № 19-01-00715.

methodologies for designing agents and multi-agent systems is carried out. The existing differences in approaches to the design of multi-agent systems are noted. The main types of models are described and their most important characteristics are given. A model of agent interaction, including a description of services (services), relationships and obligations existing between agents is presented. The model of relations (contacts), which defines communication links between agents is described. The importance and prospects of using the agent-based approach to the design of multi-agent systems are noted. The concept of designing agents and multi-agent systems, according to which the design process includes the basic components of self-organization, including the processes of interaction, crossing, adaptation to the environment, etc is proposed. Various approaches to the evolutionary design of artificial systems are considered. An evolutionary model of the formation of agents and agencies as the main component of evolutionary design is proposed. Modified evolutionary crossing-over operators to implement the agent design process are proposed.

Multiagent system; agent; properties of agents; agents; evolutionary design; evolutionary operators; design methodology; hybrid methods.

Введение. Многоагентные системы являются в настоящее время одним из наиболее активно развивающихся научных направлений. Различные подходы и методы построения искусственных агентов и многоагентных систем известны (МАС) [1–3].

С одной стороны агент может пониматься как «квазисубъект», способный в некотором смысле замещать другого субъекта (своего «хозяина»), имеющий определенные обязательства перед ним и действующий по его поручению. С другой стороны, агентом считается «активный объект» или метаобъект, способный формировать собственное поведение.

Формирование поведения включает восприятие данных из внешней среды посредством сенсоров с ограниченным разрешением, позволяющее получить информацию о событиях, происходящих в среде, интерпретацию этой информации и действие на среду посредством эффекторов [4, 5]. Здесь выделены четыре исходных агентообразующих фактора: среда, восприятие, интерпретация, действие.

Характер среды накладывает существенные требования на модель агента, при этом в процессе проектирования структуру многоагентной системы, индивидуальных и совместных действий агентов необходимо учитывать различные модели поведения, а также особенности взаимодействия агентов.

Агент должен обладать определенным набором качеств в базовом варианте. К ним можно отнести следующие свойства.

Автономность – это способность функционировать без вмешательства людей или других агентов, сохраняя контроль над своими действиями и внутренними состояниями.

Реактивность – способность оценивать текущее состояние окружающей среды, воспринимать происходящие изменения и адекватно на них реагировать.

Активность – способность не реагировать на воздействия, поступающие из внешней среды и сохранять собственное целенаправленное поведение.

Общественное поведение (Social Ability) – способность решать свои задачи совместно с другими агентами при наличии возможностей коммуникации.

Агентов можно также разделить на интеллектуальные (когнитивные) и реактивные агенты [1, 6]. В первом случае агенты способны на основе модели внешней среды выполнять рассуждения и принимать на их основе собственные действия направленные на ее изменение. Во втором агенты не имеют подобной возможности и вынуждены полагаться на получение знаний извне.

Определенные аналогии между принципами построения интеллектуальных организаций и эволюционными алгоритмами, теорией эволюции становятся очевидными в ходе изучения теории агентов и мультиагентных систем. Например, «индивид» – агент, хромосома – набор свойств агента, семья – агентство, популяция – эволюционирующая многоагентная система и т.д.

Холланд [7] определял агента как искусственный организм, развивающийся в популяции себе подобных, стремящийся обучаться и адаптироваться к внешней среде, чтобы выжить в ней (и победить конкурентов). Эта интерпретация агента опирается на теоретические подходы и модели искусственной эволюции (мутации и модификации агентов, их борьба за существование, естественный отбор) и ключевые принципы искусственной жизни (самовоспроизведение, самосохранение, самоопределение, саморегуляция агентов и пр.) [8, 9].

Простейший пример – это муравьи, которые являются естественными реактивными агентами. Отдельный муравей представляет собой примитивного агента, все действия которого, по сути, сводятся к элементарным инстинктивным реакциям на окружающую обстановку и действиям других муравьев [10, 11]. Однако собираясь вместе муравьи образуют сложную систему коллективного поведения которую называют роевой интеллект (Swarm Intelligence) [12, 13].

Анализ методов проектирования агентов и мультиагентных систем. Методология агентно-ориентированного проектирования может строиться на основе объектно-ориентированного подхода с дополнением его элементами теории агентов. Оба подхода используют обмен сообщениями для взаимодействия, а также принципы наследования и агрегирования для определения архитектуры. Таким образом, элементы сходства объектно-ориентированных и агентно-ориентированных подходов здесь используются. Различные варианты реализации методологии существуют [1, 14].

Построение множества моделей, которые определяют спецификацию многоагентной системы предусматривается в классическом варианте. Каждая модель состоит из компонентов и взаимоотношений между ними. Разрабатываемые модели разделяются на внешние и внутренние. Внешние модели относятся к системному уровню описания: основными компонентами в них являются сами агенты, взаимодействия между которыми описываются с использованием отношений наследования, агрегации, и т.п. Таким образом, абстрактные структуры агентов разрабатываются. Внутренние модели предлагаются для каждого отдельного класса агентов и описывают внутренние структуры агентов: их мнения, цели, планы и пр.

Обычно выделяются два основных вида внешних моделей: модель агентов и модель взаимодействий, определяющая способы связи (коммуникации) между агентами. Модель агентов разделяется на модель классов агентов и модель экземпляров агентов. Эти две модели определяют классы агентов и их возможные реализации, связанные между собой отношениями наследования, агрегации и означивания. Классы агентов определяют различные атрибуты агентов, включая атрибуты, задающие мнения, цели и планы агента.

Внутренняя модель, представляющая мнения, цели и планы конкретного класса агентов, является непосредственным расширением объектно-ориентированных моделей (мнения и цели) и динамических моделей (планы).

Кроме указанных видов моделей, могут строиться также:

- ◆ модели задач, описывающие задачи, которые могут выполняться агентами (исходные цели, варианты их декомпозиции, методы решения задач, и пр.);
- ◆ модели организации (например, описание социальной организации сообщества агентов или характеристики организации, куда должна внедряться данная МАС);
- ◆ модели коммуникации, которые уточняют характеристики партнерского интерфейса человека с компьютером.

Практической ценностью методологии агентно-ориентированного проектирования (АОП) является акцент на использование абстрактных классов как средства объединения ролей в процессе анализа системы. Это позволяет перенести окончательное определение ограничений конкретных агентов на более поздние стадии процесса проектирования.

Данная методология позволяет разработчику МАС переходить от перечисления требований к предельно детализированной разработке. Что в свою очередь, позволяет в дальнейшем образом создавать легко реализуемые модели. Происходит постепенный переход от более абстрактных к более конкретным структурам.

Таким образом, методологии разработки МАС помогают объединить в единую цепочку (сеть) различные процессы, необходимые для реализации многоагентной системы.

Методология Gaia является одной из самых детально проработанных методологий проектирования многоагентных систем [15–17]. Согласно этой методологии, многоагентную систему следует рассматривать как набор ролей, находящихся между собой в определенном отношении и взаимодействующих друг с другом. Исходя из этого, организационная модель многоагентной системы распадается на две части: модель ролей и модель взаимодействий (рис. 1).

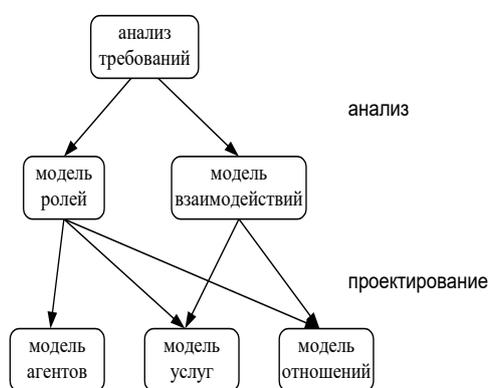


Рис. 1. Связь между моделями

Модель ролей описывает ключевые роли в системе. Здесь роль может рассматриваться как абстрактное описание функционального свойства агента. Другими словами, роль аналогична, в некотором смысле, понятию должность (например, «директор» или «старший менеджер»). Любые роли (должности) можно описать с помощью атрибутов 3-х типов:

- ◆ ответственность;
- ◆ разрешения – «права» агента, связанные с ролью;
- ◆ протоколы (рис. 2).

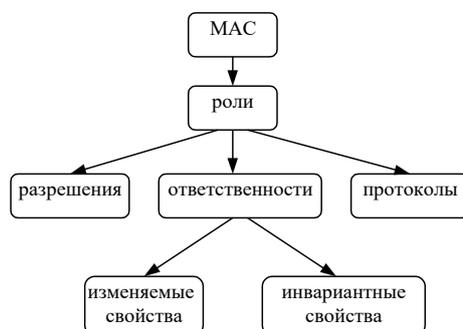


Рис. 2. Общая концепция объектно-ориентированного анализа МАС

Назначение модели агентов состоит в описании различных типов агентов, существующих в системе. Типы агентов определяются множеством ролей. Поэтому разработчик может предложить объединить несколько сходных ролей в один тип агентов. Главным критерием на этой стадии является эффективность реализации: проектировщик прежде всего стремится к оптимизации решений, и объединение нескольких ролей в один тип есть один из способов достижения этой эффективности.

Ситуация, в которой вычислительные ресурсы, требуемые для каждого агента, очень велики может служить примером, когда принимается подобное решение. В этом случае число агентов желательно сократить, поэтому принимается решение о возложении нескольких ролей на одного агента. Разумный компромисс между простотой понимания функциональных характеристик агента и эффективностью его реализации надо соблюдать здесь.

Модель взаимодействия агентов включает в себя описание услуг (сервисов), взаимосвязей и обязательств, существующих между агентами. Она состоит из множества протоколов, определяемых для каждого межролевого взаимодействия. Здесь протокол может пониматься как схема взаимодействия. Общее определение протокола состоит из следующего набора атрибутов:

- ◆ назначение: краткое описание смысла взаимодействия (например, «запрос информации», «выдача задания»);
- ◆ инициатор: роль, ответственная за инициирование взаимодействия;
- ◆ респондент: роль(и), с которой(ыми) осуществляется взаимодействие;
- ◆ входы: информация, используемая инициатором для начала взаимодействия;
- ◆ выходы: информация, предоставляемая респондентом в ходе взаимодействия.

Здесь предполагается, что реализация протокола будет вызывать серию взаимодействий.

Данная схема определяется формально, абстрагируясь от конкретной схемы реализации (непосредственной последовательности шагов). Подобное рассмотрение взаимодействий означает, что основное внимание уделяется природе и назначению взаимодействия, а не точной схеме обмена сообщениями.

Наконец, модель отношений (контактов) задает коммуникационные связи между агентами. При этом не уточняется, когда и какое сообщение проходит от одного агента к другому, а просто указывается, что существует канал коммуникации между агентами. В частности, назначение модели отношений состоит в нахождении узких мест в системе коммуникаций между агентами. Решение о возврате к стадии анализа и перепроектировании МАС для устранения этих узких мест может быть принято на основании анализа и разработки модели отношений.

Модель контактов представляет собой граф, в котором узлы соответствуют типам агентов, а дуги – коммуникационным путям (рис. 3). Обычно граф модели является ориентированным, так что $a \rightarrow b$ означает, что агент a может связываться с агентом b , и совсем не обязательно, что при этом b может связываться с a . Модель отношений строится на основе моделей ролей, протоколов и агентов.

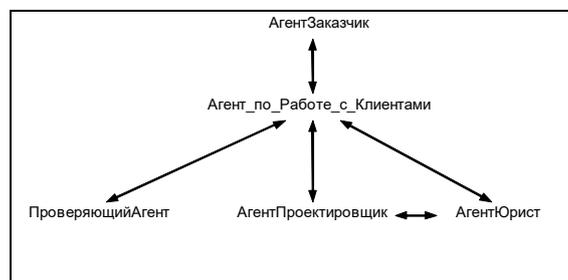


Рис. 3. Модель контактов

Итак, стадии разработки МАС включают процессы анализа и проектирования в рамках данной методологии.

Предварительная идентификация ролей происходит на стадии анализа, затем соответствующие протоколы определяются и документируются и, наконец, окончательная модель ролей строится.

Задачей процесса проектирования является преобразование абстрактных моделей высокого уровня, разработанных на стадии анализа, в модели более низкого уровня, обеспечивающих достаточно простую реализацию. Модель агента создается на стадии проектирования, т.е. роли агрегируются в типы агентов, иерархия типов формируется, примеры каждого типа документируются. Модель услуг (сервисов) и модель основных, наиболее тесных отношений (контактов) разрабатываются далее.

Агентно-ориентированный анализ приводит к пониманию того, как сообщество агентов выполняет совместные действия при решении различных задач, и какие действия требуются от конкретного агента.

Таким образом, программная технология агентов и агентно-ориентированный подход в целом понимаются как естественное развитие идей объектно-ориентированного программирования (ООП). При этом агент представляет собой самодостаточный программный процесс, включающий в себя некоторое состояние и имеющий возможность взаимодействовать с другими агентами через обмен сообщениями. Соответственно, агентно-ориентированное программирование (АОП) [17] – это новая парадигма программирования, основанная на «социальном взгляде» на вычисления.

Эволюционное проектирование агентов. Эволюционное проектирование агентов и мультиагентные системы находятся на стыке нескольких научных теорий и методологий проектирования. Различные подходы и модели эволюционного проектирования, по сути, реализуют «пограничную стратегию» между проектированием агентов и их самоорганизацией.

Разные подходы к эволюционному проектированию агентов и мультиагентных систем возможны на основе различных моделей эволюции [18]. Эволюционное проектирование может рассматриваться как процесс эволюционной адаптации системы к внешней среде. В этом случае внешняя среда является движущей силой эволюции системы, а главным направлением ее развития является адаптация к текущим условиям среды.

Формально проблему эволюционного проектирования (ЭП) искусственных систем можно представить в виде [19, 20]:

$$ED = \langle E, K, O, Q \rangle,$$

где E – множество моделей эволюции; K – множество критериев ЭП; O – множество объектов ЭП; Q – множество процедур ЭП.

Два вида эволюционного проектирования выделим: параметрическое и структурное. Параметрическое эволюционное проектирование является наиболее распространенным типом эволюционного проектирования. В этом случае речь идет о формировании агента – потомка на основе параметров агентов-родителей.

Структурное эволюционное проектирование включает в себя формирование структур агентств и единиц эволюционных МАС из агентов-родителей и агентов-потомков. Здесь базовым понятием является «семья» (агентство), состоящая, по крайней мере, из двух агентов-родителей и одного агента-потомка. Два этих уровня проектирования МАС тесно связаны между собой.

Эволюционная модель формирования агентов и агентств, выступает как основной компонент эволюционного проектирования [21–23]. Она предназначена для формирования агента – потомка на основе анализа возможных видов взаимо-

действия агентов – родителей. Популяция агентов рассматривается в рамках данной модели как эволюционирующая многоагентная система (ЭМАС) с определенным набором параметров. В качестве инструмента построения модели используется эволюционный алгоритм. Эволюционный алгоритм выполняет роль вышестоящего координатора, накладывающего ограничения на деятельность всей популяции агентов. Это позволяет накапливать в популяции положительные свойства и формировать наиболее подходящие под конкретные условия структуры агентств и потомков. Таким образом, использование эволюционного алгоритма направлено на создание эффективных методов взаимодействия между агентами и получение агента-потомка.

Основной проблемой при использовании эволюционных алгоритмов является подбор методики кодирования исходной информации [24]. Агент выступает в качестве исходного объекта в модели. Он представляется в виде хромосомы, состоящей из набора генов. Простейший случай, когда количество генов в хромосоме для всех агентов одинаково и равно двум. Гены принимают нормализованные числовые значения из интервала [0, 1]. Каждая позиция (локус) в хромосоме отвечает за определенную характеристику, поэтому перестановка генов внутри одной хромосомы невозможна. Таким образом, каждый агент в ЭМАС характеризуется своей хромосомой с конкретными значениями генов.

Выбор используемых генетических операторов непосредственно зависит от видов взаимодействий между агентами. В предлагаемой модели использовались четыре возможных вида операторов скрещивания [21–23]:

Ассоциация (*As*) – агенты-родители имеют примерно одинаковый общий ресурс и на паритетных началах участвуют в создании нового агента; стимуляция к скрещиванию у обоих агентов является внутренней, а ресурс вновь образованного агента-потомка меньше ресурса каждого из родителей;

Комбинация (*Comb*) – один из агентов-родителей имеет приоритет и обладает более значительным ресурсом, ресурс агента-потомка находится в диапазоне между ресурсами агентов-родителей; при этом для «слабого» родителя стимуляция к скрещиванию может являться внешней;

Селективное объединение (*SelU*) – используются элитные агенты-родители, в результате чего после формирования потомка у них остается достаточно ресурсов для выживания;

Слияние (*Mer*) – ресурс агента-потомка оказывается больше, чем у каждого из агентов-родителей в результате выполнения операции, а остаточный ресурс исходных агентов-родителей оказывается ниже прожиточного минимума; при этом стимуляция к скрещиванию является внешней.

Необходимые условия срабатывания и механизм оценки перспективности использования различных пар родителей задаются для каждого случая.

Процесс порождения потомка определяется следующими факторами:

Характер отношений между родителями, определяемый соотношением их общих ресурсов: Равноправные/Неравноправные;

Используемые при производстве потомка ресурсы: Равномерно/Неравномерно;

Выживание родителей: Да/Нет;

Тип стимуляции: Внутренняя/Внешняя.

Мы будем рассматривать относительные величины ресурсных параметров, предполагая, что конкретные показатели агента были преобразованы к виду [0,1] с помощью экспертов или путем нормализации.

Относительная величина общего ресурса определяется как,

$$a = (RES_{gmax} - RES_{gmin}) / (RES_{gmax} - RES_{gmin}),$$

где RES_{gtek} – общий ресурс агента на текущем поколении G; RES_{gmin} – минимальный общий ресурс агента для выживания (для каждого агента параметр одинаковый); RES_{gmax} – максимальный общий ресурс агента на все время его жизни (для каждого агента параметр разный).

Относительная величина «вкладываемого» ресурса определяется как

$$b = (RES_{btek} - RES_{bmin}) / (RES_{bmax} - RES_{bmin}),$$

где RES_{btek} – «вкладываемый» ресурс агента-родителя на текущем поколении G; RES_{bmin} – минимальный «вкладываемый» ресурс агента-родителя для образования агента-потомка (для каждого агента параметр одинаковый); RES_{bmax} – максимальный «вкладываемый» ресурс агента для образования агента-потомка на все время жизни агента-родителя (для каждого агента параметр разный).

Первый ген a соответствует относительной величине общего ресурса, которым обладает агент.

Второй ген b соответствует относительной величине ресурса, который агент-родитель готов отдать на создание агента-потомка. Таким образом, каждый ген может принимать значения из интервала $[0,1]$.

Генотип агента состоит из двух генов: a, b (табл. 1).

Таблица 1

Структура генотипа агента

a	b
0,7	0,6

Следующие обозначения для описания модели взаимодействия агентов-родителей введем:

- ◆ A_i, A_j – агенты, участвующие в схеме (агенты - родители);
- ◆ $\{a_i, b_i\}$ – гены агента A_i ;
- ◆ $\{a_i^*, b_i^*\}$ – гены агента A_i после осуществления схемы взаимодействия;
- ◆ A_n – агент, возникающий в результате осуществления схемы (потомок);
- ◆ E – эффективность пары, зависящая от параметров e_i , значения которых определяются в зависимости от удовлетворения одному из условий срабатывания схемы.

Эффективность пары – это характеристика, которая показывает, насколько пара взаимодействующих агентов это характеристика, которая показывает, насколько пара взаимодействующих агентов близка к наиболее предпочтительным хромосомам для конкретной схемы.

Изменения в генах родителей в результате взаимодействия определяются следующим образом:

$$a_i^* = a_i - b_i.$$

Данное выражение показывает, что в результате взаимодействия ресурс агента-родителя уменьшается на величину b , которая является вкладываемым ресурсом для образования агента-потомка.

Далее полагаем, что ресурс агентов, идущий на размножение, уменьшается вдвое после каждого взаимодействия, т.е.

$$b_i^* = b_i / 2.$$

Генетический оператор (ГО) в модели рассматривается как черный ящик, на вход которому подается популяция родителей, а на выходе получаем популяцию потомков. В данной модели выделены 4 класса ГО:

- ◆ генетический оператор создания начальной популяции. После его выполнения создается популяция агентов (случайное создание);

◆ генетический оператор мутации. После их выполнение создаются агенты-родители, отличающиеся от исходных агентов на какую дельта-величину.

◆ генетический оператор выбора родительской пары и скрещивания. В результате их выполнения из популяции выбираются два агента – родителя, которые скрещиваются и образуют потомка.

◆ генетический оператор отбора (селекции). В результате их действия популяция потомков – это популяция, содержащая часть агентов-родителей и агентов-потомков. Отметим, что ГО данного класса действуют по стратегии $(\mu + \lambda)$, т.е. при селекции используются и агенты-родители и агенты-потомки.

В данной модели рассматриваются три типа условий эволюционного проектирования агентов:

1) Условия взаимодействия агентов-родителей, которые определяют различные варианты (схемы) взаимодействий – паритетные (равноправные) и иерархические. Эти условия определяются параметром общего ресурса агента (точнее соотношением между ресурсами).

◆ Если $a_i \approx a_j$, то отношения паритетные.

◆ Если $a_i \gg a_j$, то отношения иерархические.

2) Условия формирования агента – потомка

Пусть a_n – это общий ресурс потомка. Тогда:

а) $a_n \leq T(b_i, b_j)$ – это условие образования потомка при взаимодействии родителей по сценарию «Ассоциация», где T – треугольная норма, т.е. общий ресурс агента-потомка меньше каждого из вкладываемых ресурсов родителей (наследование с «высоким налогом»), причем $b_i + b_j \leq \min(a_i^*, a_j^*)$;

б) $T(b_i, b_j) \leq a_n \leq S(b_i, b_j)$ – условие образования потомка при взаимодействии родителей по сценарию «Комбинация», где T – треугольная норма, S – треугольная конорма, т.е. общий ресурс агента-потомка меньше одного из вкладываемых ресурсов родителей и больше другого из вкладываемых ресурсов родителей (наследование со «средним налогом»), при этом $\min(a_i^*, a_j^*) \leq b_i + b_j \leq \max(a_i^*, a_j^*)$;

в) $a_n \geq S(b_i, b_j)$ – это условие образования потомка при взаимодействии по сценарию «Селективное объединение» или «Слияние», где S – треугольная конорма, т.е. общий ресурс агента-потомка больше каждого из вкладываемых ресурсов родителей (наследование с «низким налогом»), при этом $b_i + b_j \geq \max(a_i^*, a_j^*)$;

3) Условия выживания агентов.

◆ оставшийся общий ресурс у родителя должен быть выше некоторого порога $a_i^* \geq a_0$;

◆ образованный потомок должен иметь общий ресурс выше некоторого порога $a_n \geq a_0$.

Заключение. Общая методика эволюционного проектирования искусственных агентов и многоагентных систем построена. При ее разработке использованы понятия генотипа и фенотипа агентов: генотип связывается с наследованием ресурса агентов-родителей, а фенотип с условиями (правилами) их взаимодействия. Три типа условий нечеткого эволюционного проектирования агентов сформулированы: условия взаимодействия агентов-родителей (условия скрещивания).

Исследования предложенной модели проектирования многоагентных систем проводятся в настоящее время, определяются условия формирования агентов-потомков с использованием аппарата треугольных норм и конорм, необходимые условия выживания. Гибридные нечетко-эволюционные модели формирования агентов и агентств в контексте разработки эволюционирующих многоагентных систем предполагается сформировать в результате.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Tarasov B.B.* От многоагентных систем к интеллектуальным организациям. – М.: Эдиториал УРСС, 2002.
2. *Russel S.J., Norvig P.* Artificial Intelligence. A modern Approach. – Prentice Hall, 2003.
3. *Luger G.F.* Artificial Intelligence. Structures and Strategies for Complex Problem Solving. – 6th ed. Addison Wesley, Boston MA, 2009.
4. *Wooldridge M.* An Introduction to Multi-Agent Systems. – 2nd Ed. – New York: John Wiley and Sons, 2009.
5. *Wooldridge M., Jennings N.* Agent Theories, Architectures and Languages: a Survey // Intelligent Agents: ECAI-94 Workshop on Agent Theories, Architectures and Languages / ed. by M. Wooldridge, N. Jennings. – Berlin: Springer Verlag, 1995.
6. *Brooks R.* Intelligence Without Representation // Artificial Intelligence. – 1991. – Vol. 47. – P. 139-159.
7. *Holland, J.H.* Adaptation in Natural and Artificial Systems. – Ann Arbor: The University of Michigan Press, 1975.
8. *Редько В.Г.* Моделирование когнитивной эволюции. На пути к теории эволюционного происхождения мышления. – М.: Изд-во УРСС, 2015.
9. *Langton C.* (Ed.). Artificial Life. – New York: Addison-Wesley, 1988.
10. *Colomi A., Dorigo M., Maniezzo V.* Distributed Optimization by Ant Colonies // Proceedings of the First European Conference on Artificial Life, Paris, France, F. Varela and P. Bourguine (Eds.). – Elsevier Publishing, 1991. – P. 134-142.
11. *Colomi A., Dorigo M., Maniezzo V.* The Ant System: Optimization by a colony of cooperating agents. Tech.Rep.IRIDIA/94-28, Université Libre de Bruxelles, Belgium, 1996.
12. *Bonabeau E., Dorigo M., Theraulaz G.* Swarm Intelligence: From Natural to Artificial Systems. – New York: Oxford University Press, 1999.
13. *Гладков Л.А., Курейчик В.М., Курейчик В.В., Сороколетов П.В.* Биоинспирированные методы в оптимизации. – М.: Физматлит, 2009.
14. *Tarasov B.B.* Агенты, многоагентные системы, виртуальные сообщества: стратегическое направление в информатике и искусственном интеллекте // Новости искусственного интеллекта. – 1998. – № 2. – С. 55-63.
15. *Tarasov B.B.* Восходящее и нисходящее проектирование многоагентных систем // Проблемы управления и моделирования в сложных системах. – Самара: Самарский научный центр РАН, 1999. – С. 268-274.
16. *Wooldridge M., Jennings N.R., Kinny D.* The Gaia Methodology for Agent-Oriented Analysis and Design // Autonomous Agents and Multi-Agent Systems. – Dordrecht: Kluwer Academic Publishers, 2000. – Vol. 3. – P. 285-312.
17. *Shoham Y.* Agent Oriented Programming // Artificial Intelligence. – 1993. – Vol. 60, No. 1. – P. 51-92.
18. *Tarasov B.B., Голубин А.В.* Эволюционное проектирование: на границе между проектированием и самоорганизацией // Известия ТРТУ. – 2006. – № 8 (63). – С. 77-82.
19. *Прангишвили И.В.* Системный подход и общесистемные закономерности. – М.: СИНТЕГ, 2000.
20. *Борисов В.В., Круглов В.В., Федюлов А.С.* Нечеткие модели и сети. – М.: Горячая линия – Телеком, 2007.
21. *Gladkov L.A., Gladkova N.V., Gusev N.Y., Semushina N.S.* Integrated approach to the solution of computer-aided design problems // Proceedings of the 4th International Scientific Conference “Intelligent Information Technologies for Industry” (ITI’19). Advances in Intelligent Systems and Computing. – Vol. 875. – Springer, Cham, 2020. – P. 246-257.
22. *Gladkov L.A., Gladkova N.V., Gromov S.A.* Hybrid models of solving optimization tasks on the basis of integrating evolutionary design and multiagent technologies // Advances in Intelligent Systems and Computing. – Vol. 985. Artificial Intelligence Methods on Intelligent Algorithms. Proceeding of 8th Computer Science On-line Conference CSOC 2019. – Vol. 2. – Springer Nature Switzerland AG 2019. – P. 381-391.
23. *Gladkov L.A., Gladkova N.V., Dmitrienko N.A.* Integrated Model for Constructing Evolving Multi-Agent Subsystems // Proceedings of International Russian Automation Conference “RusAutoCon 2019”.
24. *Гладков Л.А., Курейчик В.В., Курейчик В.М.* Генетические алгоритмы. – М.: Физматлит, 2010.

REFERENCES

1. *Tarasov B.B.* Ot mnogoagentnykh sistem k intellektual'nym organizatsiyam [From multi-agent systems to intelligent organizations]. Moscow: Editorial URSS, 2002.
2. *Russel S.J., Norvig P.* Artificial Intelligence. A modern Approach. Prentice Hall, 2003.
3. *Luger G.F.* Artificial Intelligence. Structures and Strategies for Complex Problem Solving. 6th ed. Addison Wesley, Boston MA, 2009.
4. *Wooldridge M.* An Introduction to Multi-Agent Systems. 2nd ed. New York: John Wiley and Sons, 2009.
5. *Wooldridge M., Jennings N.* Agent Theories, Architectures and Languages: a Survey, *Intelligent Agents: ECAI-94 Workshop on Agent Theories, Architectures and Languages*, ed. by M. Wooldridge, N. Jennings. Berlin: Springer Verlag, 1995.
6. *Brooks R.* Intelligence Without Representation, *Artificial Intelligence*, 1991, Vol. 47, pp. 139-159.
7. *Holland, J.H.* Adaptation in Natural and Artificial Systems. Ann Arbor: The University of Michigan Press, 1975.
8. *Red'ko V.G.* Modelirovanie kognitivnoy evolyutsii. Na puti k teorii evolyutsionnogo proiskhozhdeniya myshleniya [Modeling cognitive evolution. On the way to the theory of the evolutionary origin of thinking]. Moscow: Izd-vo URSS, 2015.
9. *Langton C.* (Ed.). Artificial Life. New York: Addison-Wesley, 1988.
10. *Colomi A., Dorigo M., Maniezzo V.* Distributed Optimization by Ant Colonies, *Proceedings of the First European Conference on Artificial Life, Paris, France*, F. Varela and P. Bourguine (Eds.). Elsevier Publishing, 1991, pp. 134-142.
11. *Colomi A., Dorigo M., Maniezzo V.* The Ant System: Optimization by a colony of cooperating agents. Tech.Rep.IRIDIA/94-28, Université Libre de Bruxelles, Belgium, 1996.
12. *Bonabeau E., Dorigo M., Theraulaz G.* Swarm Intelligence: From Natural to Artificial Systems. New York: Oxford University Press, 1999.
13. *Gladkov L.A., Kureychik V.M., Kureychik V.V., Sorokoletov P.V.* Bioinspirirovannye metody v optimizatsii [Bioinspired methods in optimization]. Moscow: Fizmatlit, 2009.
14. *Tarasov B.B.* Agenty, mnogoagentnye sistemy, virtual'nye soobshchestva: strategicheskoe napravlenie v informatike i iskusstvennom intellekte [Agents, multi-agent systems, virtual communities: strategic direction in computer science and artificial intelligence], *Novosti iskusstvennogo intellekta* [Artificial intelligence news], 1998, No. 2, pp. 55-63.
15. *Tarasov B.B.* Voskhodyashchee i niskhodyashchee proektirovanie mnogoagentnykh sistem [Upstream and downstream design of multi-agent systems], *Problemy upravleniya i modelirovaniya v slozhnykh sistemakh* [Problems of management and modeling in complex systems]. Samara: Samarskiy nauchnyy tsentr RAN, 1999, pp. 268-274.
16. *Wooldridge M., Jennings N.R., Kinny D.* The Gaia Methodology for Agent-Oriented Analysis and Design, *Autonomous Agents and Multi-Agent Systems*. Dordrecht: Kluwer Academic Publishers, 2000, Vol. 3, pp. 285-312.
17. *Shoham Y.* Agent Oriented Programming, *Artificial Intelligence*, 1993, Vol. 60, No. 1, pp. 51-92.
18. *Tarasov B.B., Golubin A.V.* Evolyutsionnoe proektirovanie: na granitse mezhdu proektirovaniem i samoorganizatsiyey [Evolutionary design: on the border between design and self-organization], *Izvestiya TRTU* [Izvestiya TSURE], 2006, No. 8 (63), pp. 77-82.
19. *Prangishvili I.V.* Sistemnyy podkhod i obshchesistmenye zakonomernosti [System approach and system-wide regularities]. Moscow: SINTEG, 2000.
20. *Borisov V.V., Kruglov V.V., Fedulov A.S.* Nechetkie modeli i seti [Fuzzy models and networks]. Moscow: Goryachaya liniya – Telekom, 2007.
21. *Gladkov L.A., Gladkova N.V., Gusev N.Y., Semushina N.S.* Integrated approach to the solution of computer-aided design problems, *Proceedings of the 4th International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'19). Advances in Intelligent Systems and Computing*. Vol. 875. Springer, Cham, 2020, pp. 246-257.
22. *Gladkov L.A., Gladkova N.V., Gromov S.A.* Hybrid models of solving optimization tasks on the basis of integrating evolutionary design and multiagent technologies, *Advances in Intelligent Systems and Computing*. Vol. 985. *Artificial Intelligence Methods on Intelligent Algorithms. Proceeding of 8th Computer Science On-line Conference CSOC 2019*, Vol. 2. Springer Nature Switzerland AG 2019, pp. 381-391.

23. Gladkov L.A., Gladkova N.V., Dmitrienko N.A. Integrated Model for Constructing Evolving Multi-Agent Subsystems, *Proceedings of International Russian Automation Conference "RusAutoCon 2019"*.
24. Gladkov L.A., Kureychik V.V., Kureychik V.M. Geneticheskie algoritmy [Genetic algorithm]. Moscow: Fizmatlit, 2010.

Статью рекомендовал к опубликованию к.т.н., профессор В.Б. Тарасов.

Гладков Леонид Анатольевич – Южный федеральный университет; e-mail: leo_gladkov@mail.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371625; кафедра САПР; доцент.

Гладкова Надежда Викторовна – e-mail: nadyusha.gladkova77@mail.ru; кафедра САПР; старший преподаватель.

Gladkov Leonid Anatol'evich – Southern Federal University; e-mail: leo_gladkov@mail.ru; 44, Nekrasovskiy lane, Taganrog, 347928, Russia; phone: +78634371625; CAD department; associate professor.

Gladkova Nadezhda Viktorovna – e-mail: nadyusha.gladkova77@mail.ru; CAD department; senior teacher.

УДК 004.932.72

DOI 10.18522/2311-3103-2020-4-59-70

А.Н. Каркищенко, В.Б. Мнухин

МЕТОД ДЕТЕКЦИИ ХАРАКТЕРНЫХ ТОЧЕК ИЗОБРАЖЕНИЯ С ПОМОЩЬЮ ЗНАКОВОГО ПРЕДСТАВЛЕНИЯ*

Целью исследования является разработка метода детекции характерных точек цифрового изображения, обладающего устойчивостью по отношению к определенному классу преобразований яркости. Необходимость в подобном методе обусловлена потребностями выделения ключевых точек изображений в системах видеонаблюдения и распознавания лиц, зачастую работающих в условиях меняющейся освещенности. Особенностью предлагаемого метода, отличающего его от ряда известных подходов к проблеме выделения характерных точек, является использование так называемого знакового представления изображений. В отличие от обычного задания цифрового изображения дискретной функцией яркости, при знаковом представлении изображение задается в виде ориентированного графа, соответствующего бинарному отношению увеличения яркости на множестве пикселей. Тем самым, знаковое представление определяет не единственное изображение, а множество изображений, функции яркости которых связаны строго монотонными преобразованиями яркости. Именно это свойство знакового представления определяет его эффективность для решения задач, обусловленных поставленной выше целью. Особенностью рассматриваемого метода является особый подход к интерпретации характерных точек изображения. Это понятие в теории обработки изображений не является строго определенным; можно сказать, что характерная точка отличается повышенной «сложностью» структуры изображения в её окрестности. Поскольку знаковое представление изображения может быть представлено в виде ориентированного графа, в данной работе для оценки меры сложности локальной окрестности его вершин предложено использовать известный в спектральной теории графов метод ранжирования, основанный на теореме Перрона-Фробениуса. Его суть состоит в том, что в качестве меры сложности вершины выступает значение компоненты так называемого перроновского собственного вектора матрицы смежностей данного графа. Для проведения экспериментальных исследо-

* Работа выполнена при финансовой поддержке РФФИ, проект № 19-07-00873.

ваний предложенного подхода был разработан комплекс программ, результаты работы которых подтверждают работоспособность метода и демонстрируют, что с его помощью удастся на модельных примерах получать близкие к ожидаемым результаты. В работе предложен также ряд рекомендаций по применению данного метода.

Цифровое изображение; обработка изображений; графы; знаковые представления; характерные точки; группа; функция яркости; монотонные преобразования; спектр; собственный вектор.

A.N. Karkishchenko, V.B. Mnukhin

METHOD FOR DETECTING FEATURE POINTS OF AN IMAGE USING A SIGN REPRESENTATIONS

The aim of the study is to develop a method for detecting feature points of a digital image that is stable with respect to a certain class of brightness transformations. The need for such a method is due to the needs of detecting feature points of images in video surveillance systems and face recognition, often working in a changing light environment. A feature of the proposed method that distinguishes it from a number of well-known approaches to the problem of distinguishing characteristic points is the use of the so-called sign representation of images. In contrast to the usual defining of a digital image by a discrete brightness function, with a sign representation, the image is set in the form of an oriented graph corresponding to the binary relation of the increase in brightness on a set of pixels. Thus, the sign representation determines not a single image, but a set of images, the brightness functions of which are connected by strictly monotonic brightness transformations. It is this property of the sign representation that determines its effectiveness for solving the problems caused by the goal set above. A feature of the method under consideration is a special approach to the interpretation of the characteristic points of the image. This concept in image processing theory is not strictly defined; we can say that the characteristic point is characterized by increased "complexity" of the image structure in its vicinity. Since the sign representation of the image can be represented in the form of a directed graph, in this paper, to evaluate the complexity measure of the local neighborhood of its vertices, it is proposed to use the ranking method known in the spectral theory of graphs based on the Perron-Frobenius theorem. Its essence lies in the fact that the value of the component of the so-called Perron eigenvector of the adjacency matrix of this graph acts as a measure of the complexity of the vertex. To conduct experimental studies of the proposed approach, a set of programs was developed, the results of which confirm the efficiency of the method and demonstrate that with its help it is possible to obtain results close to the expected ones on model examples. The paper also offers a number of recommendations on the use of this method.

Digital image; image processing; graphs; sign representations; feature points; group; brightness function; monotonous transformations; spectrum; eigenvector.

Введение. Одной из наиболее распространенных задач обработки визуальной информации является задача классификации. В этом случае, как правило, требуется установить, принадлежит ли анализируемое изображение или его часть некоторому классу изображений. Примеры таких задач появляются при необходимости автоматического поиска, детекции и локализации объектов на изображении или в видеоряде. В частности, большой класс представляют задачи распознавания людей, которые зачастую обусловлены необходимостью наблюдения за доступом в контролируемые зоны, предупреждения о появлении разыскиваемых персон или опознании личности. Иными словами, целью разработки систем технического зрения является автоматическое распознавание за эффективно приемлемое время.

Как правило, разрабатываемые программно-технические системы распознавания построены на комбинации различных принципов, что делает задачу классификации методов распознавания непростой. Между тем, условно можно выделить следующие категории методов.

Методы целостного соответствия. Эти методы в качестве входных данных для алгоритма распознавания используют яркости точек всего изображения. Для распознавания людей одним из распространенных в данной категории является метод собственных лиц [1, 2], основанный на методе главных компонент.

Методы структурного соответствия. Обычно в этих методах сначала выделяются характерные объекты, после чего для распознавания используются характеристики их геометрического расположения и локальные статистики.

Гибридные методы. Методы данной категории представляют собой комбинации различных методик. Внутри каждой из этих категорий возможна более детальная классификация. Например, существует много методов распознавания лиц, основанных на методе главных компонент: метод собственных лиц (eigenfaces) [3], основанный на методе классификации «ближайший сосед»; линейный дискриминантный анализ (Fisherfaces) [4–7]; байесовские методы, использующие расстояние Махаланобиса [8, 9]. В работе [9] показано, что метод анализа независимых компонент является более представительным, чем метод главных компонент, а значит, может быть более качественным для задачи распознавания.

В области распознавания лиц наряду с этими классическими методами, существуют также подходы, основанные на применении нейронных сетей [7], которые потенциально могут давать более общее решение при наличии достаточно большой обучающей базы. Многие из подобных методов относятся к категории структурного соответствия. Эти методы в качестве характерных признаков при распознавании лиц используют антропометрические показатели [10–12]. Впоследствии разрабатывались также методы, основанные на комбинации различных расстояний [13]. К данной категории относятся и так называемые скрытые марковские модели, в которых для анализа используются фиксированные прямоугольные области лица [10]. К гибридным методам можно отнести модульный метод собственных лиц [12], гибридный метод локальных особенностей, основанный на методе главных компонент и локальном анализе особенностей [13, 14], методы, основанные на анализе формы лица, а также трехмерные модели лица [15]. В [5] предлагается подход, основанный на совмещении стандартного метода собственных лиц и спектральных разложений отдельных частей лица. Несмотря на то, что эксперименты демонстрируют лишь незначительное улучшение по сравнению со стандартным методом собственных лиц, эти методы считаются перспективными и предполагают развитие дальнейших исследований.

В работах [16, 17] был предложен метод, относящийся к категории структурного соответствия. В его основе лежит модель активных контуров (Active Shape Model), основанная на применении расстояния Махаланобиса для поиска расстояния между профилями точки. Базовыми элементами в данном методе являются *характерные точки (key points)*, которые представляют собой четко различимые ориентиры на рассматриваемых изображениях. Примером характерной точки на изображении лица может быть зрачок левого глаза, кончик носа или правый угол рта, однако в методе нет ограничений на определение точек.

В общем случае под характерной (особой, ключевой, опорной) точкой можно понимать любую точку на изображении рассматриваемого объекта, обладающую локальной особенностью. Это значит, что окрестность данной точки имеет некоторые признаки, существенно отличающие ее от остальных точек изображения. При этом желательно, чтобы упомянутые признаки таких точек были инварианты относительно изменения масштаба, ракурса съемки, освещения, поворота и других преобразований изображения, встречающихся в практических задачах, а также быть относительно устойчивыми к статистическим шумам. Описание признаков характерной точки представляется в виде так называемого дескриптора -

вектора числовых характеристик окрестности этой точки. При этом важно, чтобы дескриптор имел не очень высокую размерность, чтобы гарантировать эффективную вычислимость. Вместе с тем, низкоразмерные дескрипторы приводят к снижению точности распознавания.

В настоящей работе предлагается метод, основанный на использовании *знакового представления изображения*, суть которого заключается в описании изображения с помощью отношений квазипорядка на множестве пикселей изображения. Особенностью знакового представления является его инвариантность относительно группы строго монотонных преобразований функции яркости. Вместе с тем оно полностью сохраняет структурные особенности изображения, что обеспечивает эффективность его применения в системах технического зрения.

В частности, в работах [18–20] рассматривается применение знакового представления изображений в задачах детекции и идентификации лиц на изображениях. Применение знакового представления в модели активных контуров [21] позволяет вычислительно эффективно локализовывать антропометрические особенности лица. Исследованию устойчивости знаковых представлений к шумам посвящены работы [22, 23]. В работе [24] рассматриваются связи знакового представления с локальными бинарными шаблонами и понятием *формы* в морфологии Ю.П. Пытьева [25]

Ниже в работе понятие знакового представления изображения рассматривается более подробно; в частности, дается его интерпретация в виде ориентированного графа ([26], гл. 4). На основе данного представления предлагается метод детекции характерных точек изображений путем построения дескриптора, основанного на анализе спектральных свойств графа его знакового представления с помощью теоремы Перрона-Фробениуса. Суть метода состоит в том, что в качестве меры сложности вершины указанного графа (пикселя изображения) выступает значение компоненты так называемого перроновского собственного вектора матрицы смежностей данного графа.

Знаковое представление. Пусть $\Omega = \{1, \dots, M\} \times \{1, \dots, N\} \subset \mathbb{N}^2$ – целочисленная сетка размера $M \times N$, а $f = f(x_1, x_2)$ – функция, заданная в узлах этой сетки и принимающая значения из интервала $[0, 1]$, т.е. $f : \Omega \rightarrow [0, 1]$. Под *изображением* условимся понимать пару $\langle f, \Omega \rangle$. Пару чисел $\mathbf{x} = (x_1, x_2)$ будем называть *пикселем* или *точкой*, а значение функции f в точке \mathbf{x} – *яркостью* соответствующего пикселя. Функцию f будем называть *функцией яркости* и для краткости отождествлять с изображением. Множество всех изображений $f : \Omega \rightarrow [0, 1]$ обозначим через \mathcal{F} .

Определение 1. Отношение $\tau \subseteq \Omega \times \Omega$ будем называть *знаковым представлением* изображения $f \in \mathcal{F}$, если выполняются следующие условия:

1. $(\mathbf{x}, \mathbf{y}) \in \tau$, если $f(\mathbf{x}) \leq f(\mathbf{y})$;
2. $(\mathbf{x}, \mathbf{y}) \in \tau$, $(\mathbf{y}, \mathbf{x}) \notin \tau$, если $f(\mathbf{x}) < f(\mathbf{y})$.

Из определения следует, что знаковое представление описывает отношение яркостей в общем случае не для всех пар пикселей, а лишь для некоторого подмножества. Таким образом, по одному и тому же изображению можно построить ряд его различных знаковых представлений. В частности, под полным знаковым представлением понимается отношение, обладающее свойством связности, т.е. содержащее все пары пикселей изображения:

$$\tau = \{(\mathbf{x}, \mathbf{y}) \in \Omega^2 : f(\mathbf{x}) \leq f(\mathbf{y})\}.$$

Отметим, что полное знаковое представление заданного изображения однозначно определяется условием связности отношения.

Другим важным примером является *оконное знаковое представление* – компактный вариант знакового представления, когда учитываются лишь отношения на достаточно близко расположенных пикселях:

$$\tau = \{(\mathbf{x}, \mathbf{y}) \in \Omega^2 : f(\mathbf{x}) \leq f(\mathbf{y}), \mathbf{y} \in O_\varepsilon(\mathbf{x})\},$$

где $O_\varepsilon(\mathbf{x})$ – некоторая окрестность точки \mathbf{x} , например,

$$O_\varepsilon(\mathbf{x}) = \{\mathbf{y} \in \Omega : \|\mathbf{x} - \mathbf{y}\| < \varepsilon\},$$

где под нормой понимается манхеттенское расстояние между пикселями

$$\|\mathbf{x} - \mathbf{y}\| = |x_1 - x_2| + |y_1 - y_2|.$$

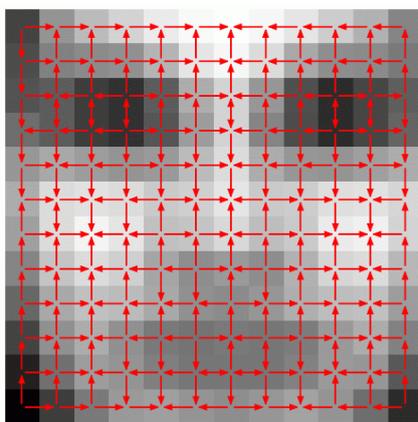


Рис. 1. Пример оконного знакового представления для $\varepsilon = 2$. Петли не показаны

Пример такого оконного представления для $\varepsilon = 2$ показан на рис. 1. В этом случае пиксели, принадлежащие одной окрестности, являются смежными в обычном смысле, т.е. непосредственно прилегающими друг к другу сторонами. Отметим, что для прикладных задач именно оконные представления с небольшим ε представляет наибольший интерес.

Очевидно, что поскольку знаковое представление характеризует лишь соотношение яркости некоторых элементов изображения, то разные изображения могут иметь одно и то же знаковое представление. Действительно, пусть $\varphi: [0, 1] \rightarrow [0, 1]$ – произвольная функция, изменяющая яркость. Тогда, для всякого изображения $f: \Omega \rightarrow [0, 1]$ композиция $\varphi(f)$ также является изображением, поэтому φ можно считать преобразованием $\varphi: \mathcal{F} \rightarrow \mathcal{F}$. Знаковое представление τ будем называть *инвариантным* относительно этого преобразования, если для любого изображения $f \in \mathcal{F}$ и изображения $g = \varphi(f)$ выполняется равенство $\tau(g) = \tau(f)$.

Понятно, что преобразование яркости φ не является, вообще говоря, обратимым, а его применение к изображению может привести к потере содержащейся в нем информации. Поэтому рассмотрим множество $\Phi = \{\varphi\}$ всех строго монотонно возрастающих преобразований $\varphi: [0, 1] \rightarrow [0, 1]$ таких, что $\varphi(0) = 0$ и $\varphi(1) = 1$. Как очевидно, Φ является группой относительно операции суперпозиции, где роль единичного элемента играет тождественное преобразование.

Как показано в [22], справедливо следующее утверждение.

Теорема 1. Знаковые представления являются инвариантными относительно преобразований группы Φ .

Примерами строго монотонно возрастающих преобразований яркости могут служить функции

$$\varphi = \begin{cases} \frac{1-(1-2z)^\alpha}{2} & \text{при } 0 \leq z \leq 0.5, \\ \frac{1+(2z-1)^\alpha}{2} & \text{при } 0.5 \leq z \leq 1, \end{cases}$$

где $\alpha \geq 0$ – некоторый параметр. При $\alpha = 1$ получаем тождественное преобразование $\varphi(z) = z$. Графики преобразования $\varphi(z)$ для различных значений α показаны на рисунке ниже.

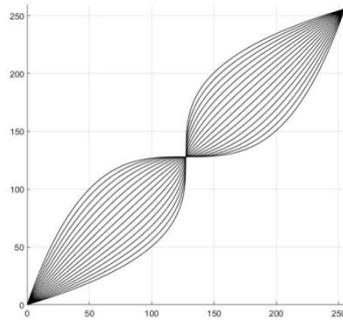


Рис. 2. Графики монотонно возрастающих преобразований яркости

На рис. 3 показаны результаты преобразования изображения человеческого лица. Параметр преобразования был равен $\alpha = a^m$, где $a = 1.2$, а $m = -8, -4, 0, 4, 8$. Исходное (не преобразованное) изображение находится в центре, отрицательные значения m соответствуют усилению контрастности, а положительные – её уменьшению. Все пять изображений имеют одно и то же знаковое представление.



Рис. 3. Преобразованные изображения лица для $m = -8, -4, 0, 4, 8$.
(Изображение лица взято из открытых источников)

Алгоритм выделения характерных точек. Напомним некоторые результаты теории спектральной теории графов [27].

Определение 2. Суммой двух графов $G_1 = \langle V, E_1 \rangle$ и $G_2 = \langle V, E_2 \rangle$ называется граф $G_1 + G_2 = \langle W, E \rangle$, где $W = V \times U$ и вершина $(v_1, u_1) \in W$ смежна с $(v_2, u_2) \in W$ тогда и только тогда, когда либо $v_1 = v_2$ и $(u_1, u_2) \in E_2$, либо же когда $u_1 = u_2$ и $(v_1, v_2) \in E_1$.

Если $A(G)$ – матрица смежности графа G , то

$$A(G_1 + G_2) = A(G_1) \otimes I_N + I_M \otimes A(G_2),$$

где \otimes – кронекерово произведение матриц, M и N – числа вершин графов G_1 и G_2 соответственно, а I_M и I_N – единичные матрицы соответствующих размеров.

Как известно [27], спектром графа G называется множество $\text{Spec}(G)$ всех собственных значений его матрицы смежности. Как показано в [27, с.70], если $\text{Spec}(G_1) = \{\lambda_1, \lambda_2, \dots, \lambda_M\}$, а $\text{Spec}(G_2) = \{\mu_1, \mu_2, \dots, \mu_N\}$, то

$$\text{Spec}(G_1 + G_2) = \{\lambda_i + \mu_j : 1 \leq i \leq M, 1 \leq j \leq N\},$$

что и объясняет название операции $G_1 + G_2$.

В частности, если P_M – простая неориентированная цепь длины M , то граф $P_M + P_N$ представляет собой прямоугольную решетку размера $M \times N$, см. рис. 4. Поскольку [27 с.73] характеристический многочлен графа P_M представляет собой многочлен Чебышева второго рода порядка M , то

$$\text{Spec}(P_M) = \left\{ 2 \cos \frac{\pi i}{M+1} : 1 \leq i \leq M \right\}$$

и

$$\text{Spec}(P_M + P_N) = \left\{ 2 \cos \frac{\pi i}{M+1} + 2 \cos \frac{\pi j}{N+1} : 1 \leq i \leq M, 1 \leq j \leq N \right\}.$$

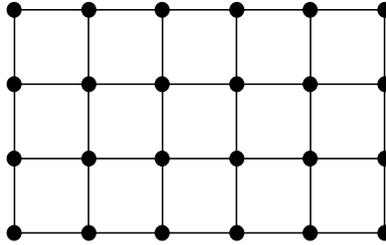


Рис. 4. Граф $P_4 + P_6$

Таким образом, наибольшее собственное значение графа $P_M + P_N$ равно

$$2 \cos \frac{\pi}{M+1} + 2 \cos \frac{\pi}{N+1} < 4. \quad (1)$$

Спектр ориентированного графа (ввиду несимметричности его матрицы смежности) может включать в себя комплексные числа. Однако из хорошо известной теоремы Перрона-Фробениуса [28, с. 334] следует, что если орграф G сильно связан (т.е. в G существует путь из любой вершины в любую другую), то его наибольшее по модулю собственное значение единственно, действительно и положительно. Его называют перроновским собственным значением (или индексом графа) и обозначают $\lambda(G) > 0$. Соответствующий нормированный собственный вектор $\mathbf{p}(G) = (p_1, p_2, \dots, p_N)$ называют перроновским собственным вектором графа G . Все его компоненты положительны, $0 < p_1 \leq 1$, и допускают [29, с. 590] следующую комбинаторную интерпретацию.

Обозначим через $\sigma_i(k)$ число замкнутых маршрутов длины $k > 2$, начинающихся и заканчивающихся в вершине v_i орграфа G , т.е.

$$(\sigma_1(k), \sigma_2(k), \dots, \sigma_N(k)) = \text{diag}(A^k(G)).$$

Тогда

$$p_i = \lim_{k \rightarrow \infty} \sqrt[k]{\frac{\sigma_i(k)}{\lambda^k}}, \quad \text{где } \lambda = \lambda(G). \quad (2)$$

Таким образом, величина p_i в некотором смысле характеризует «сложность строения» окрестностей вершины v_i графа G , что обуславливает широкое применение перроновских векторов в различных задачах, сводящихся к ранжированию элементов сетей.

Мы воспользуемся перроновскими векторами для выделения характерных точек изображения. Для этого свяжем знаковые представления с графами. Заметим, что знаковому представлению τ изображения f естественно ставится в соответствие ориентированный граф $G_\tau = \langle \Omega, E_\tau \rangle$, множеством вершин которого является множество пикселей Ω , а множество дуг E_τ совпадает с множеством τ ; при этом, если $(\mathbf{x}, \mathbf{y}) \in \tau$, то соответствующая дуга графа направлена от вершины \mathbf{x} к вершине \mathbf{y} . Понятно, что построенный таким образом орграф G_τ имеет петли на каждой вершине. Далее условимся рассматривать G_τ без петель, сохраняя за ним прежнее обозначение. Кроме того, будем считать пару противоположно направленных дуг между двумя вершинами одним неориентированным ребром.

Пусть далее τ – оконное знаковое представление и $\varepsilon = 2$. Если f – тривиальное изображение, все пиксели которого имеют одинаковую яркость, т.е. $f(\mathbf{x}) = \text{const}$ для всех $\mathbf{x} \in \Omega$, то $G_\tau = P_M + P_N$. Следовательно, граф нетривиального изображения получается из $P_M + P_N$ заданием ориентации на некоторых ребрах, или другими словами, удалением одной из дуг в парах, соответствующих ребрам. Поскольку удаление дуги не может увеличить индекс графа, из (1) немедленно вытекает, что

$$\lambda(G_\tau) \leq 2 \cos \frac{\pi}{M+1} + 2 \cos \frac{\pi}{N+1} < 4.$$

Таким образом, граф изображения размера $M \times N$ имеет MN вершин, причем информация об изображении кодируется в виде «отклонения матрицы $A(G_\tau)$ от симметричности»: парам смежных пикселей одинаковой яркости соответствуют единицы, симметричные относительно главной диагонали матрицы, а различной яркости – симметрично расположенные 0 и 1.

Идея предлагаемого подхода к выделению характерных точек изображения основана на интерпретации таких точек как максимальных компонент перроновского вектора некоторой матрицы, ассоциированной с G_τ . При этом использование $A(G_\tau)$ нецелесообразно ввиду того, что вершины G_τ , соответствующие пикселям однотонных областей, инцидентны неориентированным ребрам. Это значительно увеличивает число замкнутых маршрутов, проходящих через такие вершины и, в соответствии с (2), максимизирует соответствующую компоненту перроновского вектора. В то же время интуитивно очевидно, что характерные точки должны соответствовать областям значительного изменения яркости. Это приводит нас к следующему определению.

Определение 3. Пусть $A(G_\tau) = (a_{ij})$, где $1 \leq i, j \leq MN$. Квадратную матрицу $S = S_\tau = (|a_{ij} - a_{ji}|)$ будем называть *градиентной матрицей* знакового представления τ изображения f .

Как очевидно, S – симметрическая (0,1)-матрица, единицы которой соответствуют парам пикселей различной яркости. Граф, матрицей смежности которого является S_τ будем обозначать H_τ .

Разумеется, при обработке реальных изображений f нет необходимости в работе одновременно со всей матрицей S_τ . Это связано как со значительными размерами этой матрицы, так и с тем, что понятие характерной точки по самой своей сути является локальным. Предлагаемый алгоритм предусматривает разбиение анализируемого изображения на небольшие прямоугольные фрагменты (с возможностью их перекрытия). Размеры этих фрагментов ограничены следующими ограничениями:

- ◆ максимальный размер фрагмента ограничен вычислительной сложностью нахождения перроновского вектора (действительно, даже для фрагмента изображения с размерами 32×32 приходится работать с матрицей S_τ размера 1024×1024 . В то же время возникающие матрицы сильно разрежены, что позволяет использовать для работы с ними известные эффективные алгоритмы;

- ◆ фрагмент должен быть достаточно большим для того, чтобы находящиеся в нем потенциальные характерные точки могли себя «проявить»;

- ◆ фрагмент не может быть слишком маленьким, поскольку при его уменьшении растёт вероятность несвязности графа H_τ , что нарушает условие применимости теоремы Перрона-Фробениуса. (Разумеется, эта вероятность всегда остается ненулевой, однако опыт применения алгоритма позволяет считать её пренебрежимо малой уже для фрагментов порядка 10×10).

Кроме того, при построении знакового представления представляется целесообразным не учитывать незначительных изменений яркости (порядка 1%-2% от наибольшего), вызываемых неизбежными шумами.

На рис. 5 приведен пример работы алгоритма для изображения лица на рис. 3. Анализируемое изображение имело размер 330×330 , используемое окно – 9×9 , параметр $\varepsilon = 3$.

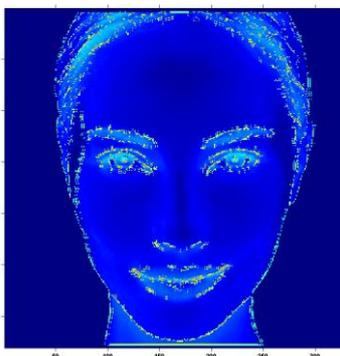


Рис. 5. Пример построения характерных точек на основе предлагаемого алгоритма

Заключение. Резюмируя изложенное, можно сделать следующие выводы. В работе рассмотрен метод построения характерных точек цифрового изображения, заданного своим знаковым представлением. Особенностью метода является его инвариантность относительно группы строго монотонных преобразований функции яркости. Метод основан на интерпретации характерных точек как максимальных компонент перроновского вектора некоторой матрицы, ассоциированной с графом знакового представления. Приведен алгоритм выделения характерных точек и рассмотрены условия и ограничения его применимости. Приведен пример работы алгоритма.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Kirby M., Sirovich L. Application of the Karhunen-Loeve procedure for characterization of human faces // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1990. – Vol. 12. – P. 23-35.
2. Sirovich L., Kirby M. Low-dimensional procedure for characterization of human face // Journal of the Optical Society of America. – 1987. – Vol. 4. – P. 519-524.
3. Turk M., Pentland A. Eigenfaces for recognition // Journal of Cognitive Neuroscience. – 1991. – Vol. 3. – P. 72-86.
4. Liu C., Wechsler H. A shape- and texture-based enhanced fisher classifier for face recognition // IEEE Transactions Image Processing. – 2001. – Vol. 10. – P. 598-608.
5. Swets D.L., Weng J. Using discriminant eigenfeatures for image retrieval // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1996. – Vol. 18. – P. 831-836.
6. Moghaddam B., Pentland A. Probabilistic visual learning for object representation // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1997. – Vol. 19. – P. 696-710.
7. Lin S.H., Kung S.Y., Lin L.J. Face recognition/detection by probabilistic decision-based neural network // Transactions on Neural Network. – 1997. – Vol. 8. – P. 114-132.
8. Liu C., Wechsler H. Evolutionary pursuit and its application to face recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2000. – Vol. 22. – P. 570-582.
9. Cox I.J., Ghosn J., Yianilos P.N. Feature-based face recognition using mixture distance // Computer Vision and Pattern Recognition. – 1996. – P. 209-216.
10. Nefian A.V., Hayes M.H. Hidden Markov models for face recognition // Acoustics, Speech and Signal Processing. – 1998. – Vol. 5. – P. 2721-2724.
11. Samaria F., Young S. HMM-based architecture for face identification // Image and Vision Computing. – 1994. – Vol. 12. – P. 537-543.
12. Pentland A., Moghaddam B., Starner T. View-based and modular eigenspaces for face recognition // Computer Vision and Pattern Recognition. – 1994. – P. 84-91.
13. Penev P., Atick J. Local feature analysis: A general statistical theory for object representation // Computation in Neural Systems. – 1996. – No. 7. – P. 477-500.
14. Lanitis A., Taylor C.J., Cootes T.F. Automatic face identification system using flexible appearance models // Image and Vision Computing. – 1995. – Vol. 13. – P. 393-401.
15. Weyrauch B., Heisele B., Huang J., Blanz V. Component-Based Face Recognition with 3D Morphable Models // Computer Vision and Pattern Recognition Workshop. – 2004. – P. 85-93.
16. Гончаров А.В., Каркищенко А.Н. Влияние освещенности на качество распознавания фронтальных лиц // Известия ЮФУ. Технические науки. – 2008. – № 4(81). – С. 82-92.
17. Goncharov A., Gubarev V. Comparison of high-level and low-level face recognition methods // Proc. Pattern recognition and image analysis: new information technologies (PRIA-9-2008). – 2008. – P. 178-181.
18. Гончаров А.В. Исследование свойств знакового представления изображений в задачах распознавания образов // Известия ЮФУ. Технические науки. – 2009. – № 8 (97). – С. 178-188.
19. Броневич А.Г., Гончаров А.В. Знаковое представление изображений и его информативность // 14-я Всероссийская конф. Математические методы распознавания образов. – М.: МАКС Пресс, 2009. – С. 309-312.
20. Гончаров А.В. Распознавание лиц на основе многомасштабного знакового представления изображений // Цифровая обработка сигналов. – 2010. – Т. 1. – С. 10-13.

21. Гречухин И.А., Каркищенко А.Н. О методе построения характерных точек на основе вариации яркостного профиля // Тр. ОАО «НИИАС». Сб. научных трудов. Вып. 9. – М.: Изд-во ООО «Группа ИДТ». – 2011. – С. 203-220.
22. Каркищенко А.Н., Гончаров А.В. Геометрия знакового представления изображений и её приложение к исследованию устойчивости к шумам // Международная конференция Интеллектуализация обработки информации (ИОИ-8): Сб. докладов. – М.: МАКС Пресс, 2010. – С. 335-339.
23. Каркищенко А.Н., Гончаров А.В. Исследование устойчивости знакового представления изображений // Автоматика и телемеханика. – 2010. – Т. 9. – С. 57-69.
24. Мясников В.В. Описание изображений с использованием конфигурационного отношения эквивалентности // Компьютерная оптика. – 2018. – Т. 42, № 6. – С. 998-1007.
25. Пытьев Ю.П., Чуличков А.И. Методы морфологического анализа изображений. – М.: Физматлит, 2010. – 336 с.
26. Броневич А.Г., Каркищенко А.Н., Ленский А.Е. Анализ неопределенности выделения информативных признаков и представлений изображений. – М.: Физматлит, 2013. – 308 с.
27. Svetković D.M., Doob M., Sachs H. Spectra of Graphs – Theory and Application. – Berlin, 1980. – 368 p.
28. Гантмахер Ф.П. Теория матриц. – М.: Наука, 1988. – 552 с.
29. Хорн Р., Джонсон Ч. Матричный анализ. – М.: Мир, 1989. – 655 с.

REFERENCES

1. Kirby M., Sirovich L. Application of the Karhunen-Loeve procedure for characterization of human faces, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, Vol. 12, pp. 23-35.
2. Sirovich L., Kirby M. Low-dimensional procedure for characterization of human face, *Journal of the Optical Society of America*, 1987, Vol. 4, pp. 519-524.
3. Turk M., Pentland A. Eigenfaces for recognition, *Journal of Cognitive Neuroscience*, 1991, Vol. 3, pp. 72-86.
4. Liu C., Wechsler H. A shape- and texture-based enhanced fisher classifier for face recognition, *IEEE Transactions Image Processing*, 2001, Vol. 10, pp. 598-608.
5. Swets D.L., Weng J. Using discriminant eigenfeatures for image retrieval, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996, Vol. 18, pp. 831-836.
6. Moghaddam B., Pentland A. Probabilistic visual learning for object representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, Vol. 19, pp. 696-710.
7. Lin S.H., Kung S.Y., Lin L.J. Face recognition/detection by probabilistic decision-based neural network, *Transactions on Neural Network*, 1997, Vol. 8, pp. 114-132.
8. Liu C., Wechsler H. Evolutionary pursuit and its application to face recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, Vol. 22, pp. 570-582.
9. Cox I.J., Ghosn J., Yianilos P.N. Feature-based face recognition using mixture distance, *Computer Vision and Pattern Recognition*, 1996, pp. 209-216.
10. Nefian A.V., Hayes M.H. Hidden Markov models for face recognition, *Acoustics, Speech and Signal Processing*, 1998, Vol. 5, pp. 2721-2724.
11. Samaria F., Young S. HMM-based architecture for face identification, *Image and Vision Computing*, 1994, Vol. 12, pp. 537-543.
12. Pentland A., Moghaddam B., Starner T. View-based and modular eigenspaces for face recognition, *Computer Vision and Pattern Recognition*, 1994, pp. 84-91.
13. Penev P., Atick J. Local feature analysis: A general statistical theory for object representation, *Computation in Neural Systems*, 1996, No. 7, pp. 477-500.
14. Lanitis A., Taylor C.J., Cootes T.F. Automatic face identification system using flexible appearance models, *Image and Vision Computing*, 1995, Vol. 13, pp. 393-401.
15. Weyrauch B., Heisele B., Huang J., Blanz V. Component-Based Face Recognition with 3D Morphable Models, *Computer Vision and Pattern Recognition Workshop*, 2004, pp. 85-93.
16. Goncharov A.V., Karkishchenko A.N. Vliyaniye osveshchennosti na kachestvo raspoznavaniya frontal'nykh lits [Effects of illumination and quality of frontal faces recognition], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2008, No. 4 (81), pp. 82-92.

17. *Goncharov A., Gubarev V.* Comparison of high-level and low-level face recognition methods, *Proc. Pattern recognition and image analysis: new information technologies (PRIA-9-2008)*, 2008, pp. 178-181.
18. *Goncharov A.V.* Issledovanie svoystv znakovogo predstavleniya izobrazheniy v zadachakh raspoznavaniya obrazov [Investigation of properties of images sign representation in the pattern recognition problems], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2009, No. 8 (97), pp. 178-188.
19. *Bronevich A.G., Goncharov A.V.* Znakovoe predstavlenie izobrazheniy i ego informativnost' [Sign representations of images and its informativity], *14-ya Vserossiyskaya konf. Matematicheskie metody raspoznavaniya obrazov* [Proceedings of 14th All-Russia scientific conference of mathematical methods for images recognition]. Moscow: MAKS Press, 2009, pp. 309-312.
20. *Goncharov A.V.* Raspoznavanie lits na osnove mnogomasshtabnogo znakovogo predstavleniya izobrazheniy [On face recognition based on multiscale sign representations of images], *Tsifrovaya obrabotka signalov* [Digital signal processing], 2010, Vol. 1, pp. 10-13.
21. *Grechukhin I.A., Karkishchenko A.N.* O metode postroeniya kharakternykh toчек na osnove variatsii yarkostnogo profilya [On a method for key points construction based on brightness cross-section variations], *Tr. OAO «NILAS». Sb. nauchnykh trudov* [Proceedings of JSC NILAS]. Issue 9. Moscow: Izd-vo OOO «Gruppa IDT», 2011, pp. 203-220.
22. *Karkishchenko A.N., Goncharov A.B.* Geometriya znakovogo predstavleniya izobrazheniy i ee prilozhenie k issledovaniyu ustoychivosti k shumam [Sign representations geometry applied for noise stability investigations], *Mezhdunarodnaya konferentsiya Intellektualizatsiya obrabotki informatsii (IOI-8): Sb. dokladov* [Proceedings of International conference on intellectualization of images processing (IIP-8)]. Moscow: MAKS Press, 2010, pp. 335-339.
23. *Karkishchenko A.N., Goncharov A.V.* Issledovanie ustoychivosti znakovogo predstavleniya izobrazheniy [Stability investigation of the sign representation of images], *Avtomatika i telemekhanika* [Automation and Remote Control], 2010, Vol. 9, pp. 57-69.
24. *Myasnikov V.V.* Opisanie izobrazheniy s ispol'zovaniem konfiguratsionnogo otnosheniya ekvivalentnosti [Description of images based on configurational equivalence relation], *Komp'yuternaya optika* [Computer Optics], 2018, Vol. 42, No. 6, pp. 998-1007.
25. *Pyt'ev Yu.P., Chulichkov A.I.* Metody morfologicheskogo analiza izobrazheniy [Methods for Morphological Analysis of Images]. Moscow: Fizmatlit, 2010, 336 p.
26. *Bronevich A.G., Karkishchenko A.N., Lepskiy A.E.* Analiz neopredelennosti vydeleniya informativnykh priznakov i predstavleniy izobrazheniy [Uncertainty analysis of informational features selection and images representations]. Moscow: Fizmatlit, 2013, 308 p.
27. *Cvetković D.M., Doob M., Sachs H.* Spectra of Graphs – Theory and Application. Berlin, 1980, 368 p.
28. *Gantmakher F.R.* Teoriya matrits [Theory of Matrices]. Moscow: Nauka, 1988, 552 p.
29. *Khorn R., Dzhonson Ch.* Matrichnyy analiz [Matrix Analysis]. Moscow: Mir, 1989, 655 p.

Статью рекомендовал к опубликованию д.ф.-м.н. Г.В. Куповых.

Каркищенко Александр Николаевич – Научно-исследовательский институт робототехники и процессов управления ЮФУ, e-mail: karkishalex@gmail.com; 347928, г. Таганрог, ул. Шевченко, 2; тел.: +78634371694; д.ф.-м.н.; профессор; ведущий научный сотрудник.

Мнухин Валерий Борисович – Институт компьютерных технологий и информационной безопасности ЮФУ; e-mail: mnukhin.valeriy@mail.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: +78634371606; к.ф.-м.н.; доцент.

Karkishchenko Alexander Nikolayevich – Research and Development Institute of Robotics and Control Systems of the Southern Federal University; e-mail: karkishalex@gmail.com; 2 Shevchenko street, Taganrog, 347928 Russia; phone: +78634371694; dr. of phys. and math. sc.; professor; leading researcher.

Mnukhin Valeriy Borisovich – Institute of computer technology and information security of the Southern Federal University; e-mail: mnukhin.valeriy@mail.ru; 44 Nekrasovskiy lane, Taganrog, 347928, Russia; phone: +78634371606; cand. of phys. and math. sc.; associate professor.

В.М. Курейчик, И.Б. Сафроненкова**ОНТОЛОГИЧЕСКИЙ ПОДХОД К РЕАЛИЗАЦИИ ТЕХНОЛОГИЙ
РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ В СЕТИ ИНТЕРНЕТ**

Развитие технологий распределенных вычислений позволило объединить географически-распределенные ресурсы, таким образом, предоставив возможности для эффективного решения ресурсоемких задач в различных областях науки и техники. Наряду с этим актуализировался ряд задач, который требует к своему решению новых подходов, учитывающих особенности реализации современных Интернет технологий. В настоящей работе рассмотрена проблема, связанная с переносом вычислительной нагрузки в распределенной системе автоматизированного проектирования (САПР), функционирующей в «туманной» среде. Целью данной работы является разработка онтологического подхода к решению задачи переноса вычислительной нагрузки в САПР с учетом особенностей «туманной» среды. Онтологический подход заключается в проведении процедуры онтологического анализа, которая позволяет «отсеивать» узлы-кандидаты, не отвечающие ресурсным требованиям для переноса части нагрузки. Научная новизна работы заключается в использовании моделей онтологии для решения задачи переноса вычислительной нагрузки в САПР. Это позволяет сократить число узлов-кандидатов в «тумане» для переноса нагрузки, тем самым сократить время моделирования процессов размещения и, соответственно, общее время решения задачи переноса нагрузки. Принципиальным отличием данного подхода является использование знаний о предметной области, отраженных в модели онтологии, для решения задачи переноса вычислительной нагрузки. Проведенные в работе вычислительные эксперименты доказали целесообразность использования онтологического анализа для решения задачи переноса вычислительной нагрузки.

Распределенная САПР; «туманные» вычисления; задача переноса вычислительной нагрузки; онтология.

V.M. Kureichik, I.B. Safronenkova**AN ONTOLOGICAL APPROACH TO DISTRIBUTED COMPUTING
TECHNOLOGIES IMPLEMENTATION ON THE INTERNET**

Distributed computing technologies development has allowed uniting geographically distributed resources and has provided an opportunity for effective resource intensive problem-solving in various fields of science and technology. At the same time a set of problems, which demands the development of new approaches, taking into account contemporary Internet technologies implementation, has risen. In this paper a problem of workload relocation in distributed computer-aided design system (DCAD) operating in the “fog” environment was considered. The goal of this paper is ontological approach development to workload relocation problem-solving in DCAD taking into account some “fog” environment features. The ontological approach involves an ontological procedure implementation, which allows “filtering” the candidate-nodes, which have insufficient resources for workload relocation. The scientific novelty of this paper is ontological models using for workload relocation problem-solving in DCAD. It allows reducing the number of candidate-nodes in the “fog” for workload relocation, thereby contributing to reduce the time of location process modeling and, consequently, the total time of workload relocation problem-solving is also reduced. The fundamental difference of presented approach is domain knowledge, represented in ontological model, applying for workload relocation problem-solving. The experimental study results have shown the expediency of ontological analysis for workload relocation problem-solving.

Distributed computer-aided design system; “fog”- computing; workload relocation problem; ontology.

Введение. Задача формирования расписания не нова, имеет различные постановки и решалась многократно множеством ученых [1, 2]. Однако данная проблема не теряет актуальности и в настоящее время, приобретая новые особенности в рамках развития технологий распределенных вычислений и Интернет-технологий [3–5].

Стоит отметить, что проектирование современных СБИС даже посредством распределенных систем автоматизированного проектирования (РСАПР) требует наличия значительных вычислительных и временных ресурсов. В этой связи быстроедействие РСАПР является определяющим фактором конкурентоспособности производства в целом [6].

Тот факт, что современные РСАПР успешно функционируют на базе Интернет-технологий, выполняющих задачи коммуникации и информационного обмена между приложениями [7], подтверждается появлением на рынке множества продуктов автоматизированных систем, реализованных на базе «облачных» и «туманных» концепций [8, 9].

Системы, в которых применяется «безсерверная архитектура» (serverless architecture) обеспечивают выполнение исходного кода тысяч и миллионов пользователей внутри вычислительной среды, не заботясь о масштабировании ресурсов. Это, в свою очередь, делает возможным обеспечить доступ к информационным ресурсам распределенных баз данных и организацию дистанционного взаимодействия распределенных коллективов пользователей САПР, что гарантирует высокий технический уровень проектов, уменьшение сроков их реализации и сведение к минимуму числа ошибок [10].

Туманные вычисления – это высоко-виртуализированная платформа, предоставляющая вычислительные, запоминающие и сетевые сервисы между конечными устройствами и центрами данных облачных вычислений, и которая обычно, но не всегда, располагается на границах сети [11]. «Туманную» среду можно рассматривать как расширение «облачных» вычислений до границ сети, которая дополняет централизованную «облачную» модель. В этом случае часть функций по обработке и хранению данных передается граничным узлам сети, что позволяет решать многие задачи, возникающие при экспоненциальном росте числа подключенных к сети устройств. В связи с этим, главной характеристикой «туманных» вычислений является возможность переноса части вычислительно нагрузки, что, в условиях «туманной» среды, предполагает решение оптимизационной задачи с определенными критериями и ограничениями, которые частично формируются требованиями самой вычислительной системы [12, 13].

В настоящей работе описан онтологический подход к решению задачи переноса вычислительной нагрузки в РСАПР, функционирующей на базе «туманных» вычислений.

Результаты проведенных вычислительных экспериментов подтвердили целесообразность использования онтологического подхода к реализации технологий распределенных вычислений.

Постановка задачи переноса вычислительной нагрузки в «туманной» среде. Пусть дан граф вычислительной задачи G , с некоторой вычислительной сложностью подзадачи x_i и объемом информации w_i , передаваемой между подзадачами. Граф G разделен на два подграфа G' и G'' . Необходимо разместить подграф G' на вычислительные устройства (ВУ) сегмента сети P' «туманного» слоя, в то время, как вычислительные задачи подграфа G'' продолжает исполняться на сегменте сети P'' (рис. 1).

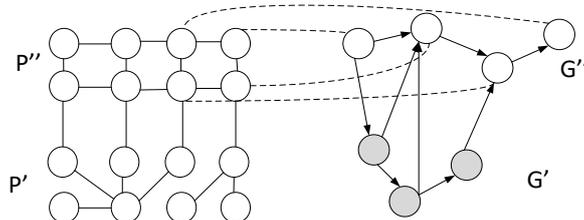


Рис. 1. Распределение вычислительных подзадач при переносе части вычислительной нагрузки

Будем считать, что основным критерием при решении оптимизационной задачи переноса вычислительной нагрузки является надежность системы.

Рассмотрим графовое описание набора задач

$$G = \{ \langle i, x_i, w_i \rangle \}, \quad (1)$$

где i – уникальный идентификатор вычислительной подзадачи;

x_i – вычислительная сложность i -ой подзадачи;

w_i – объем информации, передаваемый i -ой подзадачей в коммуникационную среду.

Вычислительные подзадачи графа G связаны с узлами вычислительных устройств из множества P . Причем P описывается графой структурой

$$P = \{ \langle j, p_j \rangle \text{ list} \}, \quad (2)$$

где j – идентификатор узла;

p_j – производительность узла;

$list$ – матрица пропускной способности каналов связи между инцидентными узлами сети.

Теперь рассмотрим подграф вычислительных подзадач G' , которые требуется перенести в то время, как вычислительные подзадачи подграфа G'' продолжают свое исполнение. Между данными подграфами существует несколько информационных потоков. Потоки, связанные с подграфом G' могут быть описаны набором кортежей:

$Flow_in = \{ \langle id_{out}, id_{in}, w_{out_in} \rangle \}$ – кортеж, описывающий количество информации, передаваемой между узлом $G'' id_{out}$ и узлом $G' id_{in}$ от G'' к G' .

$Flow_out = \{ \langle id_{in}, id_{out}, w_{in_out} \rangle \}$ – кортеж, описывающий количество информации, передаваемой между узлом $G' id_{out}$ и узлом $G'' id_{in}$ от G' к G'' .

Рассмотрим задачу размещения нагрузки. Пусть имеется подграф вычислительных задач G'' , связанный с P'' , и $Flow_in, Flow_out$. Необходимо разместить вычислительные подзадачи подграфа G' на множестве устройств сети P' таким образом, что общее время выполнения вычислительных подзадач G меньше заданного времени T с учетом выполнения критерия надежности системы.

Решением поставленной задачи является установление связи между вычислительными подзадачами G' и вычислительными узлами P' , которое может быть описано матрицей A :

$$A = \begin{pmatrix} \langle t_o^{ij}, u_{ij} \rangle & & & \\ & \dots & & \\ & & & \langle t_o^{NM}, u_{NM} \rangle \end{pmatrix}, \quad (3)$$

где t_o^{ij} – момент времени, когда начинается вычисление i -ой подзадачи j -м узлом;
 u^{ij} – доля общей производительности p_j заданного j -го узла для выполнения i -ой подзадачи.

Для дальнейшей разработки модели необходимо рассмотреть следующие параметры:

- ◆ $L_p(A)$ – нагрузка узла, порождаемая переносом вычислительной подзадачи на узел;
- ◆ $L_{dist}(A, Flow_in, Flow_out)$ – нагрузка узла, порождаемая обменом информации между подграфами G' и G'' ;
- ◆ $L_{tr}(A, Flow_in, Flow_out)$ – нагрузка узла, порождаемая передачей информации через узел;
- ◆ D_{lk} – список ребер графа P , который определяет маршрут между узлами l и k ;
- ◆ $ListD_{lk}$ – матрица, описывающая пропускную способность каналов связи между узлами l и k .

Рассмотрим целевую функцию как набор значений функций надежности определенных узлов F_j :

$$F_j = e^{-\lambda_j t}, \quad (4)$$

где λ_j – интенсивность отказов j -го узла, t – время функционирования устройства.

Поскольку

$$\lambda = \lambda_0 \cdot 2^{\Delta T / 10}, \quad (5)$$

$$\Delta T = kL, \quad (6)$$

где L – нагрузка устройства, k – коэффициент, который зависит от типа устройства.

Тогда зависимость между функцией надежности и нагрузкой описывается следующей формулой:

$$F_j = e^{-\lambda_{j0} \cdot 2^{kL/10}}. \quad (7)$$

Загрузка устройства зависит от распределения вычислительных задач по ВУ, которые описываются матрицей A . Рассмотренные выше параметры должны быть включены в модель задачи: $L_p(A)$, $L_{dist}(A, Flow_in, Flow_out)$, $L_{tr}(A, Flow_in, Flow_out)$.

Таким образом, полная нагрузка j -го ВУ описывается следующей формулой:

$$L_j = L_{pj}(A) + L_{distj}(A, Flow_in, Flow_out) + L_{trj}(A). \quad (8)$$

Главным ограничением для этой задачи является время исполнения T графа вычислительной задачи G , т.е. для:

$$G = G' \cup G'', \forall i \in G: \frac{x_i}{p_j u_{ij}} + t_{dist}(i) < T. \quad (9)$$

где $t_{dist}(i)$ – максимальное время доставки информации от i -ой подзадачи к подзадаче-получателю данных.

Поскольку модель рассматривает маршруты потока информации, $t_{dist}(i)$ описывается функцией:

$$t_{dist}(i) = \xi(A, G, P). \quad (10)$$

Более точно время доставки информации может быть вычислено на основе полной информации о подзадачах, назначаемых на вычислительные узлы с учетом параметров D_{ik} и $ListD_{ik}$ [14].

Онтологический подход к решению задачи переноса вычислительной нагрузки. Как известно, онтологии широко применяются для концептуального описания различных предметных областей [15–17].

В основе онтологического подхода к решению задачи переноса вычислительной нагрузки лежит онтологический анализ.

Введем понятие «узел-лидер». Условимся, что узел-лидер выбирается с использованием алгоритма избрания лидера из числа вычислительных узлов, на которых изначально выполнялась задача, и выполняет функцию сбора информации об узлах-кандидатах, расположенных в «тумане», потенциально готовых принять часть вычислительной нагрузки.

Общая схема процедуры онтологического анализа изображена на рис. 2 и включает в себя выполнение следующих этапов:

1. Входные данные, т.е. исходные данные параллельного популяционного алгоритма, реализуемого в РСАПР.
2. Классификация совокупности входных данных в соответствии с классами предметной онтологии;
3. Применение системы продукционных правил к классам онтологии с целью ограничения множества узлов, которое было получено в результате сбора сведений об имеющихся ресурсах узлом-лидером.
4. Принятие решения о выборе ограниченного множества узлов для переноса вычислительной нагрузки.



Рис. 2. Схема процедуры онтологического анализа

Ключевым элементом онтологического анализа является модель онтологии, которая позволяет структурировать знания об аспектах, являющихся ключевыми для решения задачи переноса вычислительной нагрузки, а именно:

Исходные данные переносимой задачи:

- ◆ классы параллельных алгоритмов, используемые в РСАПР (алгоритмы, вдохновленные живой и неживой природой, алгоритмы, инспирированные человеческим обществом и прочие) [18];

- ◆ модели распараллеливания алгоритмов (глобальная модель, островная модель) [19];
 - ◆ изменяемые параметры данных моделей (длительность сезона);
 - ◆ способы разбиения моделей распараллеливания (будем считать число способов разбиений конечно);
 - ◆ регламентированное время выполнения алгоритма;
- Информация о вычислительных ресурсах узлов-кандидатов:
- ◆ исходное размещение вычислительных узлов («туман», «облако»);
 - ◆ требования к вычислительным ресурсам (производительность узла-кандидата, удаленность узла-кандидата от узла-лидера, загруженность узла-кандидата).

Фрагмент разработанной модели онтологии изображен на рис. 3.

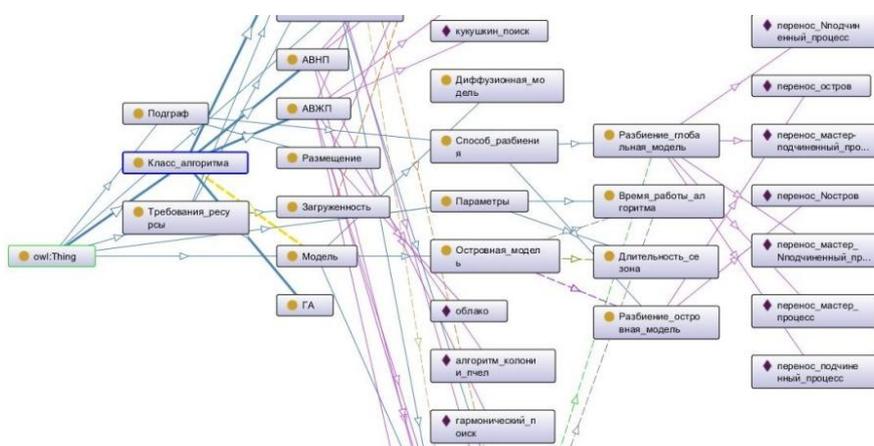


Рис. 3. Фрагмент разработанной онтологии

Таким образом, онтологический подход к решению задачи переноса вычислительной нагрузки, в основе которого лежит онтологический анализ позволяет «отсеять» узлы-кандидаты, заведомо не удовлетворяющие требованиям, необходимым для переноса части вычислительной нагрузки. Это позволяет сократить время, требуемое для моделирования процессов размещения, и, тем самым сократить время решения задачи переноса вычислительной нагрузки.

Реализация предложенной процедуры онтологического анализа производилась в среде программирования Python 3.8.3 для работы в операционной среде семейства MS Windows [20].

Вычислительный эксперимент. Для демонстрации эффективности работы процедуры онтологического анализа для распределенных систем автоматизированного проектирования, проведем две серии экспериментов в различных условиях. Для этого введем некоторые понятия.

1. В качестве критерия эффективности выберем время, необходимое для решения задачи переноса вычислительной нагрузки.

2. Эффективность работы модуля, реализующего процедуру онтологического анализа, будем определять по формуле:

$$\Delta = \frac{F_n - F_k}{F_n} \times 100\% , \quad (1)$$

где F_n – начальное значение целевой функции;

F_k – конечное значение целевой функции;

Δ – оценочная функция изменения времени, необходимого для решения задачи переноса вычислительной нагрузки в РСАПР.

3. Под «глубиной туманного» слоя будем понимать максимальное расстояние в хопах от узла-лидера до узла-кандидата, расположенного в «туманном» слое.

Первая серия вычислительных экспериментов заключалась в проведении экспериментов по реализации процедуры онтологического анализа на множестве доступных вычислительных узлов «туманного» слоя, глубина которого лежит в интервале от 2 до 5 хопов, доступное количество узлов в «туманном» слое $F=100$ и $F=1000$.

Вторая серия вычислительных экспериментов заключалась в проведении экспериментов по реализации процедуры онтологического анализа на множестве доступных вычислительных узлов «туманного» слоя, глубина которого лежит в интервале от 2 до 30 хопов, доступное количество узлов в «туманном» слое $F=100$ и $F=1000$.

Результаты двух серий вычислительных экспериментов отражены в табл. 1, 2 и на рис. 4, 5.

Таблица 1

Сравнение эффективности процедуры онтологического анализа при глубине «туманного» слоя 2 до 5 хопов ($\Delta\{2;5\}$) и от 2 до 30 хопов ($\Delta\{2;30\}$) в процентах, $F=100$

Номер эксперимента	Количество доступных вычислительных узлов, F	Количество узлов после процедуры онтологического анализа, F_s для глубины «туманного слоя» {2;5}	Количество узлов после процедуры онтологического анализа, F_s для глубины «туманного слоя» {2;30}	$\Delta\{2;5\}$, %	$\Delta\{2;30\}$, %
1	100	14	3	86	97
2	100	11	1	89	99
3	100	15	3	85	97
4	100	8	2	92	98
5	100	17	3	83	97
6	100	14	3	86	97
7	100	15	2	85	98
8	100	10	1	90	99
9	100	11	2	89	98
10	100	21	2	79	98

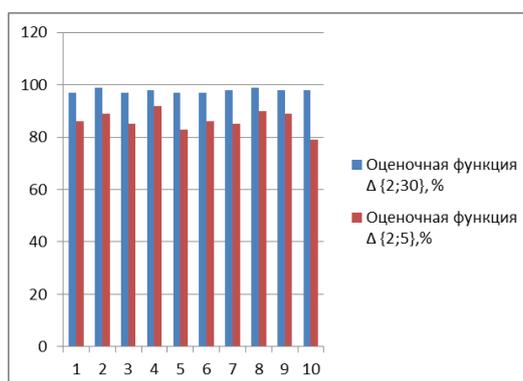


Рис. 4. Гистограмма сравнения эффективности процедуры онтологического анализа при глубине «туманного» слоя 2 до 5 хопов ($\Delta\{2;5\}$) и от 2 до 30 хопов ($\Delta\{2;30\}$) в процентах, $F=100$

Как видно из табл. 1 и рис. 4, наилучший результат улучшения целевой функции для глубины «туманного» слоя от 2 до 5 хопов составил 92 %, для глубины «туманного» слоя от 2 до 30 хопов – 99 % при количестве доступных узлов в «туманном» слое $F=100$.

Таблица 2

Сравнение эффективности процедуры онтологического анализа при глубине «туманного» слоя 2 до 5 хопов ($\Delta\{2;5\}$) и от 2 до 30 хопов ($\Delta\{2;30\}$) в процентах, $F=1000$

Номер эксперимента	Количество доступных вычислительных узлов, F	Количество узлов после процедуры онтологического анализа, F_s для глубины «туманного» слоя» {2;5}	Количество узлов после процедуры онтологического анализа, F_s для глубины «туманного» слоя» {2;30}	$\Delta\{2;5\}$, %	$\Delta\{2;30\}$, %
1	1000	138	15	86,2	98,5
2	1000	135	16	86,5	98,4
3	1000	119	13	88,1	98,7
4	1000	144	21	85,6	97,9
5	1000	149	10	85,1	99
6	1000	142	16	85,8	98,4
7	1000	139	15	86,1	98,5
8	1000	160	14	84	98,6
9	1000	152	28	84,8	97,2
10	1000	141	19	85,9	98,1

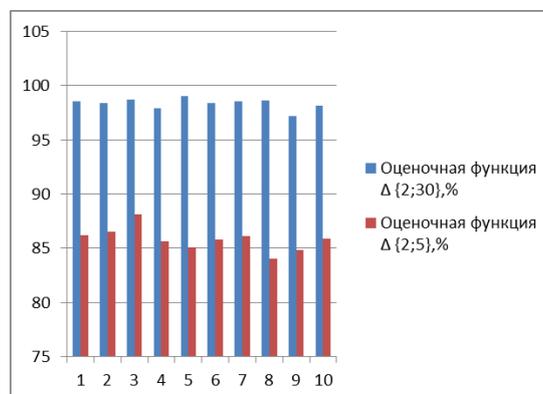


Рис. 5. Гистограмма сравнения эффективности процедуры онтологического анализа при глубине «туманного» слоя 2 до 5 хопов ($\Delta\{2;5\}$) и от 2 до 30 хопов ($\Delta\{2;30\}$) в процентах, $F=1000$

Как видно из табл. 2 и рис. 5, наилучший результат улучшения целевой функции для глубины «туманного» слоя от 2 до 5 хопов составил 88,1 %, для глубины «туманного» слоя от 2 до 30 хопов – 99 % при количестве доступных узлов в «туманном» слое $F=1000$.

Проведя анализ полученных результатов (табл. 1, 2 рис. 4, 5) можно сделать следующие выводы:

1. Наилучшие результаты получены при глубине «туманного» слоя в интервале от 2 до 30 хопов, при этом достигается улучшение целевой функции на 99 %.

2. Эффективность процедуры онтологического анализа зависит от глубины «туманного» слоя: чем глубже «туман», тем меньше времени требуется на решение задачи переноса вычислительной нагрузки в РСАПР.

Заключение. Актуальность разработки новых подходов и методов решения задачи переноса вычислительной нагрузки обусловлено развитием технологий распределенных вычислений, в том числе «туманной» концепции. В работе приведена постановка задачи переноса вычислительной нагрузки с учетом особенностей «туманной» среды. Предложен онтологический подход к решению данной задачи, основой которого является онтологический анализ. Разработана предметная онтология как ключевой компонент онтологического анализа, на основе которого разработан программный модуль в среде Python 3.8.3 и проведены вычислительные эксперименты.

Научная новизна предложенного подхода заключается в использовании онтологического анализа для решения задачи переноса вычислительной нагрузки в РСАПР, функционирующей в «туманной» среде. Это позволяет «отсеять» узлы-кандидаты, принадлежащие «туманному» слою, заведомо не отвечающие требованиям к ресурсам, необходимым для размещения вычислительной нагрузки. Принципиальным отличием предложенного подхода является использование знаний об используемых в РСАПР параллельных алгоритмах для решения задачи переноса вычислительной нагрузки.

Вычислительные эксперименты показали возможность улучшения по времени в случаях с различной глубиной «туманного» слоя: для интервала от 2 до 5 хопов улучшение целевой функции составило 92 % при $F=100$ и 88,1 % при $F=1000$; для интервала от 2 до 30 хопов улучшение целевой функции составило 99 % в обоих случаях.

Проведенный вычислительный эксперимент показал, что проведение процедуры онтологического анализа позволяет улучшить эффективность решения задачи переноса вычислительной нагрузки на 99%.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Кузюрин Н.Н., Грушин Д.А., Фомин С.А. Проблемы двумерной упаковки и задачи оптимизации в распределенных вычислительных системах // Тр. ИСП РАН. – 2014. – № 1. – С.483-501.
2. Жук. С.Н. О построении расписаний выполнения параллельных задач на группах кластеров с различной производительностью // Тр. Института системного программирования РАН. – 2012. – Т. 23. – С. 447-454.
3. Мельник Э.В., Клименко А.Б., Иванов Д.Я. Модель задачи распределения вычислительной нагрузки для информационно-управляющих систем на базе концепции туманных вычислений // Известия ТулГУ. Технические науки. – 2018. – № 2. – С. 174-187.
4. Melnik E.V., Klimenko A.B. A Novel Approach to the Reconfigurable Distributed Information and Control Systems Load-Balancing Improvement // 11th International Conference on Application of Information and Communication Technologies (AICT) – 2017. – P. 355-359.
5. Cisco Fog Computing Solutions: Unleash the Power of the Internet of Things. – Режим доступа: <https://docplayer.net/19735338-Unleashing-the-power-of-the-internet-of-things.html> (дата обращения: 16.07.2020).
6. Глушань В.М., Лаврик П.В. Распределенные САПР. Архитектура и возможности. – Старый Оскол: ТНТ, 2015. – 187 с.
7. Дмитриевич Г.Д., Анисимов Д.А. Построение систем автоматизированного проектирования на основе сервис-ориентированной архитектуры // Известия СПбГЭТУ «ЛЭТИ». – 2014. – № 2. – С. 14-18.
8. Облачные и интерактивные САПР. – Режим доступа: <https://www.autodesk.ru/solutions/cloud-based-online-cad-software> (дата обращения: 16.07.2020).
9. AWS IoT Greengrass. – Режим доступа: <https://aws.amazon.com/ru/greengrass/> (дата обращения: 16.07.2020).

10. Глушань В.М., Иванько Р.В. Анализ эффективности распределенных САПР // Известия ТРТУ. – 2006. – № 8 (63). – С. 115-120.
11. Chiang M., Balasubramanian B., Bonomi F. Fog for 5G and IoT. – Wiley, 2017. – 305 p.
12. Kalyaev I., Melnik E., Klimenko A. A Technique of Adaptation of the Workload Distribution Problem Model for the Fog-Computing Environment // Advances in Intelligent Systems and Computing. – 2019. – Vol. 986.
13. Гузик В.Ф., Каляев И.А., Левин И.И. Реконфигурируемые вычислительные системы: учебное пособие / под общ. ред. И.А. Каляева. – Таганрог: Изд-во ЮФУ, 2016. – 472 с.
14. Melnik E.V., Klimenko A.B., Ivanov D.Ya. Модель задачи формирования сообществ устройств информационно-управляющих систем в средах туманных вычислений // XIII Всероссийское совещание по проблемам управления ВСПУ-2019. – Тула: Тульский государственный университет, 2018. – Вып. 2.
15. Цуканова Н.И. Онтологическая модель представления и организации знаний: учеб. пособие для вузов. – М.: Горячая линия – Телеком, 2014. – 272 с.
16. Курейчик В.М., Сафроненкова И.Б. Создание онтологической модели систем автоматизированного проектирования в среде Protege 4.2 // Проблемы разработки перспективных микро- и наноэлектронных систем – 2016: Сб. трудов / под общ. ред. Академика РАН А.Л. Стемповского. – М.: ИПИМ РАН, 2016. – Ч. III. – С. 240-246.
17. Kureichik V., Safronenkova I.: Ontology-Based Decision Support System for the Choice of Problem-Solving Procedure of Commutation Circuit Partitioning // Creativity in Intelligent Technologies and Data Science. – 2017. – Vol. 754.
18. Гладков Л.А., Курейчик В.В., Курейчик В.М., Сороколетов П.В. Биоинспирированные методы в оптимизации. – М.: Физматлит, 2009. – 384 с.
19. Карпенко А.П. Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой: учеб. пособие. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2014. – 446 с.
20. Свидетельство о государственной регистрации программы для ЭВМ 2020616523 Российская Федерация. Модуль онтологического анализа для распределенной системы автоматизированного проектирования / И.Б. Сафроненкова, Э.В. Мельник, В.М. Курейчик; заявители и правообладатели Федеральное государственное бюджетное учреждение науки «Федеральный исследовательский центр Южный научный центр Российской академии наук» (ЮНЦ РАН), Сафроненкова Ирина Борисовна. – No 2020614892; заявл. 18.05.2020; опубл 18.06.2020. – 1 с.

REFERENCES

1. Kuzuryn N.N., Grushin D.A., Fomin S.A. Problemy dvumernoy upakovki i zadachi optimizatsii v raspredelennykh vychislitel'nykh sistemakh [Two-dimensional packing problems and optimization in distributed computing systems], *Tr. ISP RAN* [Proceedings of the Institute for System Programming of the RAS], 2014, No. 1, pp. 483-501.
2. Zhuk. S.N. O postroenii raspisaniy vypolneniya parallel'nykh zadach na gruppakh klasterov s razlichnoy proizvoditel'nost'yu [On-line algorithm for scheduling parallel tasks on a group of related clusters], *Tr. Instituta sistemnogo programirovaniya RAN* [Proceedings of the Institute for System Programming of the RAS], 2012, Vol. 23, pp. 447-454.
3. Mel'nik E.V., Klimenko A.B., Ivanov D.Ya. Model' zadachi raspredeleniya vychislitel'noy nagruzki dlya informatsionno-upravlyayushchikh sistem na baze kontseptsii tumannykh vychisleniy [Model of the problem of distribution of computing load for information-control systems on the basis of the concept of miscalculation], *Izvestiya TulGU. Tekhnicheskie nauki* [Izvestiya Tula State University], 2018, No. 2, pp. 174-187.
4. Melnik E.V., Klimenko A.B. A Novel Approach to the Reconfigurable Distributed Information and Control Systems Load-Balancing Improvement, *11th International Conference on Application of Information and Communication Technologies (AICT) – 2017*, pp. 355-359.
5. Cisco Fog Computing Solutions: Unleash the Power of the Internet of Things. Available at: <https://docplayer.net/19735338-Unleashing-the-power-of-the-internet-of-things.html> (accessed 16 July 2020).

6. *Glushan' V.M., Lavrik P.V.* Raspredelelynye SAPR. Arkhitektura i vozmozhnosti [Distributed CAD systems. Architecture and possibilities]. Staryy Oskol: TNT, 2015, 187 p.
7. *Dmitrevich G.D., Anisimov D.A.* Postroenie sistem avtomatizirovannogo proektirovaniya na osnove servis-orientirovannoy arkhitektury [Computer-aided design systems engineering based on service-oriented architecture], *Izvestiya SPbGETU «LETI»* [Izvestia SPbETU "LETI"], 2014, No. 2, pp. 14-18.
8. Oblachnye i interaktivnye SAPR [Cloud and interactive CAD systems]. Available at: <https://www.autodesk.ru/solutions/cloud-based-online-cad-software> (accessed 16 July 2020).
9. AWS IoT Greengrass. Available at: <https://aws.amazon.com/ru/greengrass/> (accessed 16 July 2020).
10. *Glushan' V.M., Ivan'ko R.V.* Analiz effektivnosti raspredelelynykh SAPR [The analysis of CAD system effectiveness], *Izvestiya TRTU* [Izvestiya TSURE], 2006, No. 8 (63), pp. 115-120.
11. *Chiang M., Balasubramanian B., Bonomi F.* Fog for 5G and IoT. Wiley, 2017, 305 p.
12. *Kalyaev I., Melnik E., Klimenko A.* A Technique of Adaptation of the Workload Distribution Problem Model for the Fog-Computing Environment, *Advances in Intelligent Systems and Computing*, 2019, Vol. 986.
13. *Guzik V.F., Kalyaev I.A., Levin I.I.* Rekonfiguriruemye vychislitel'nye sistemy: uchebnoe posobie [Reconfigurable computer systems], ed. by I.A. Kalyaeva. Taganrog: Izd-vo YuFU, 2016, 472 p.
14. *Melnik E.V., Klimenko A.B., Ivanov D.Ya.* Model' zadachi formirovaniya soobshchestv ustroystv informatsionno-upravlyayushchikh sistem v sredakh tumannykh vychisleniy [A model of local group device of information and management system forming in the fog-environmental], *XIII Vserossiyskoe soveshchanie po problemam upravleniya VSPU-2019* [Proceedings of the XIII VSPU-2019.]. Tula: Tul'skiy gosudarstvennyy universitet, 2018, Issue 2.
15. *Tsukanova N.I.* Ontologicheskaya model' predstavleniya i organizatsii znaniy: ucheb. posobie dlya vuzov [An ontological model of knowledge representation and management: study guide for universities]. Moscow: Goryachaya liniya – Telekom, 2014, 272 p.
16. *Kureychik V.M., Safronenkova I.B.* Sozdanie ontologicheskoy modeli sistem avtomatizirovannogo proektirovaniya v srede Protege 4.2 [Creation of CAD-systems ontology using Protege 4.2] *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem – 2016: Sb. trudov* [All-Russia Science & Technology Conference "Problems of Advanced Micro- and Nanoelectronic Systems Development"], ed. by Akademiya RAN A.L. Stempovskogo. Moscow: IPPM RAN, 2016, Part III, pp. 240-246.
17. *Kureichik V., Safronenkova I.*: Ontology-Based Decision Support System for the Choice of Problem-Solving Procedure of Commutation Circuit Partitioning, *Creativity in Intelligent Technologies and Data Science*, 2017, Vol. 754.
18. *Gladkov L.A., Kureychik V.V., Kureychik V.M., Sorokoletov P.V.* Bioinspirirovannyye metody v optimizatsii [Bio-inspired methods in optimization]. Moscow: Fizmatlit, 2009, 384 p.
19. *Karpenko A.P.* Sovremennyye algoritmy poiskovoy optimizatsii. Algoritmy, vdokhnovlennyye prirodoy: ucheb. posobie [Contemporary algorithms of search engine optimization. Bio-inspired algorithms: textbook]. Moscow: Izd-vo MGTU im. N.E. Baumana, 2014, 446 p.
20. *Safronenkova I.B., Mel'nik E.V., Kureychik V.M.* Svidetel'stvo o gosudarstvennoy registratsii programmy dlya EVM 2020616523 Rossiyskaya Federatsiya. Modul' ontologicheskogo analiza dlya raspredelelynoy sistemy avtomatizirovannogo proektirovaniya [State registration certificate 2020616523 Russian Federation. A module of ontological analysis for distributed computer-aided design system]; applicants and copyright holders Federal state budgetary institution of science "Federal research center southern scientific center of the Russian Academy of Sciences" (UNC RAS), Safronenkova Irina Borisovna, No 2020614892; declared 18.05.2020; published 18.06.2020, 1 p.

Статью рекомендовал к опубликованию д.т.н. С.Г. Капустян.

Курейчик Виктор Михайлович – Южный федеральный университет; e-mail: vmkureychik@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: +78634311487; кафедра САПР; д.т.н.; профессор.

Сафроненкова Ирина Борисовна – Федеральный исследовательский центр Южный научный центр Российской академии наук; e-mail: safronenkova050788@yandex.ru; г. Ростов-на-Дону, ул. Чехова, 41; тел.: +79604678753; м.н.с.

Kureichik Viktor Mikhailovich – Southern Federal University, e-mail: vmkureychik@sfedu.ru; 44, Nekrasovskiy, Taganrog, 347928, Russia; phone: +78634311487; the department of computer aided design; dr. of eng. sc.; professor.

Safronenkova Irina Borisovna – Federal Research Centre The Southern Scientific Centre of the Russian Academy of Sciences; e-mail: safronenkova050788@yandex.ru; 41, Chekhov street, Rostov-on-Don, 344006, Russia; phone: +79604678753; junior researcher.

УДК 004.896

DOI 10.18522/2311-3103-2020-4-82-93

Б.К. Лебедев, В.Б. Лебедев, О.Б. Лебедев**ПОПУЛЯЦИОННЫЙ АЛГОРИТМ ПОСТРОЕНИЯ ДЕРЕВА РЕШЕНИЙ
МЕТОДОМ КРИСТАЛЛИЗАЦИИ РОССЫПИ АЛЬТЕРНАТИВ***

В ряде случаев возникает необходимость установления соответствия между заявленным и фактическим значением категориальной переменной на основе совокупности признаков объекта. В этом случае возникает потребность в классификаторе с оптимальной последовательностью рассматриваемых атрибутов с заданным значением целевой функции. Значением целевой переменной может быть: да, нет, номер сорта, номер класса и т.д. В работе решается задача построения классификационной модели в виде оптимальной последовательности рассматриваемых атрибутов и их значений, входящих в состав маршрута от корневой вершины к концевой вершине с заданным значением целевой переменной. Если требуется классификатор, включающий возможность альтернативных ответов, то вначале строятся независимо друг от друга оптимальные маршруты для каждого значения целевой переменной, а затем эти маршруты объединяются («склеиваются») в единое бинарное дерево решений. В алгоритме построения классификатора на основе метода кристаллизации россыпи альтернатив, каждое решение Q_k интерпретируется в виде ориентированного маршрута M_k на бинарном дереве решений. Назовем порядковый номер элемента в ориентированном маршруте M_k позицией $s_i \in S = \{s_i | i=1, 2, \dots, n_A\}$. Элементом маршрута M_k является пара (x_i, u_i) , где x_i соответствует A_i , u_i в маршруте M_k является ребром, выходящим из x_i и соответствует выбранному вместе с A_i значению A_i . Вторым индексом элемента u_i определится после выбора A_i , помещенного в соседнюю с s_i позицию s_{j+1} . Работа алгоритма построения дерева решений базируется на использовании коллективной эволюционной памяти, под которой подразумевается информация, отражающая историю поиска решения. Алгоритм учитывает тенденции к использованию альтернатив из наилучших найденных решений. Особенности являются наличие непрямого обмена информацией – стигмержи. Совокупность данных об альтернативах и их оценках составляет россыпь альтернатив. Рассмотрены ключевые моменты анализа альтернатив в процессе эволюционной коллективной адаптации. Экспериментальные исследования показали, что разработанный алгоритм находит решения, не уступающие по качеству, а иногда и превосходящие своих аналогов в среднем на 3–4 %. Временная сложность алгоритма, полученная экспериментальным путем, лежит в пределах $O(n^2)$ – $O(n^3)$.

Классификация; дерево решений; оптимизация; популяционный алгоритм; адаптивное поведение; метод кристаллизации россыпи альтернатив.

В.К. Lebedev, V.B. Lebedev, O.B. Lebedev**POPULATION ALGORITHM FOR CONSTRUCTING A TREE
OF SOLUTIONS BY METHOD OF CRYSTALLIZATION OF ALTERNATIVES
FIELD**

In some cases, it becomes necessary to establish a correspondence between the declared and actual value of a categorical variable on the basis of a set of object characteristics. In this case, there is a need for a classifier with an optimal sequence of the considered attributes with a

* Работа выполнена при финансовой поддержке гранта РФФИ № 20–07–00260 А.

given value of the objective function. The target variable can be: yes, no, variety number, class number, etc. This paper solves the problem of constructing a classification model in the form of an optimal sequence of the considered attributes and their values included in the route from the root vertex to the terminal vertex with a given value of the target variable. If a classifier is required that includes the possibility of alternative answers, then first, independently from each other, optimal routes are built for each value of the target variable, and then these routes are combined ("glued") into a single binary decision tree. In the algorithm for constructing a classifier based on the method of crystallization of a placer of alternatives, each solution Q_k is interpreted as an oriented route M_k on a binary decision tree. Let us call the ordinal number of an element in the directed route M_k the position $s_i \in S = \{s_i | i=1, 2, \dots, n_A\}$. An element of the route M_k is the pair (x_i, u_i) , where x_i corresponds to A_i , u_i in the route M_k is an edge outgoing from x_i and corresponds to the value A_i chosen together with A_i . The second index of the element u_i is determined after the choice of A_i , placed in the position s_{j+1} adjacent to s_j . The work of the decision tree construction algorithm is based on the use of collective evolutionary memory, which is understood as information reflecting the history of the search for a solution. The algorithm takes into account the tendency to use alternatives from the best solutions found. The peculiarities are the presence of an indirect exchange of information – stigmerges. The totality of data on alternatives and their assessments constitutes a scattering of alternatives. The key points of the analysis of alternatives in the process of evolutionary collective adaptation are considered. Experimental studies have shown that the developed algorithm finds solutions that are not inferior in quality, and sometimes surpass their counterparts by an average of 3–4 %. The time complexity of the algorithm, obtained experimentally, lies within $O(n^2)$ - $O(n^3)$.

Classification; decision tree; optimization; population algorithm; adaptive behavior; method of crystallization of alternatives placer.

Введение. Наиболее распространенные методы решения задач классификации используют в качестве квалификационной модели дерево решений $D=(X,U)$, где $X=\{x_i | i=1, 2, \dots, n\}$ – множество вершин, $U=\{u_i | i=1, 2, \dots, m\}$ – множество ребер [1–3]. Множество X включает множество X_1 внутренних вершин и множество X_2 «концевых» вершин. Внутренние вершины дерева решений соответствуют признакам, характеризующим объект и подвергающимся разбиению, в зависимости от значений признаков. «Концевые» вершины соответствуют значениям категориальных переменных (конкретный класс, сорт и т.д.) [4,5]. Все ребра ориентированные. При этом ребро, выходящее из вершины x_i , соответствует значению признака x_i [4], другими словами все ребра дерева решений D помечены метками, соответствующими значениям признаков. Решение задачи классификации нового объекта, заключается в построении на дереве решений ориентированного маршрута M , начиная от корня до одной из концевых вершин. Порядок вершин в ориентированном маршруте определяет порядок учета признаков. Если целевая переменная принимает дискретные значения, то решается задача классификации. Самым распространенным, и наиболее простым случаем являются бинарные деревья решений (БДР) [3–4]. Эффективность БДР во многом зависит от правильного выбора последовательности и критерия ветвления внутренних вершин дерева решений.

Большинство из известных алгоритмов (CART, C4.5, NewId, ITrule, CHAID, CN2 и т.д. [1–4]) являются «жадными алгоритмами» последовательного типа. При использовании такого подхода построение БДР происходит сверху вниз. На каждом шаге жадного алгоритма разбиение множества объектов производится по признаку, обеспечивающему возникновение максимального различия между подмножествами. Последовательные алгоритмы отличаются наименьшей трудоемкостью, но дают наименьшее качество.

В работе решается задача построения классификационной модели в виде оптимальной последовательности рассматриваемых атрибутов и их значений, входящих в состав маршрута от корневой вершины к концевой вершине с заданным значением целевой переменной.

Если требуется классификатор, включающий возможность альтернативных ответов, то вначале строятся независимо друг от друга оптимальные маршруты для каждого значения целевой переменной, а затем эти маршруты объединяются («склеиваются») в единое БДР.

В ряде случаев возникает необходимость установления соответствия между заявленным и фактическим значением категориальной переменной на основе совокупности признаков объекта. В этом случае возникает потребность в классификаторе с оптимальной последовательностью рассматриваемых атрибутов с заданным значением целевой функции. Значением целевой переменной может быть: да, нет, номер сорта, номер класса и т.д.

Эффективным направлением повышения качества решений стало использование стохастических **популяционных** алгоритмов [5], которые, как правило, итерационные и работают в пространстве полных решений. Широкое распространение получили роевые и генетические алгоритмы. Исследования эффективности популяционных алгоритмов показали, что мощным средством повышения эффективности новых алгоритмов является их гибридизация [6, 7]. Рекомбинация метаэвристик популяционных алгоритмов обеспечивает более равномерный и обоснованный просмотр пространства поиска и более высокую эффективность интегрированных алгоритмов [8].

Разработка новых поисковых алгоритмов заключается в использовании и модификации метаэвристик, заложенных в природных механизмах принятия решений. В работе по аналогии с метаэвристикой, на которых построены роевые алгоритмы, используется метаэвристика [10], учитывающая тенденцию к использованию альтернатив (вариантов компонентов) из наилучших найденных решений. Особенности являются наличие непрямого обмена информацией – стигмержи [11]. Совокупность данных об альтернативах и их оценках составляет **россыпь альтернатив**. Рассмотрены ключевые моменты анализа альтернатив в процессе эволюционной коллективной адаптации [12, 13]. Алгоритм на основе кристаллизации россыпи альтернатив был успешно применен для решения задачи построения дерева решений. В процессе эволюционной коллективной адаптации методами дискриминантного анализа формируются оценки приспособленности альтернатив. Приспособленность альтернатив рассматривается как вероятность ее использования в формируемом решении. Совокупность данных об альтернативах и их оценках составляет **россыпь альтернатив**. Дискриминантный анализ и поиск эффективных альтернатив, в процессе эволюционной коллективной адаптации назван по аналогии с процессами вычленения объектов (формирования кристаллов) кристаллизацией. Другими словами, в процессе эволюционной коллективной адаптации производится вычленение из множества вариантов наиболее приспособленных альтернатив. Отсюда название метода оптимизации – метод кристаллизации россыпи альтернатив (КРА), (Crystallization of alternatives field (CAF)).

1. Постановка задачи построения бинарного дерева решений методами кристаллизации россыпи альтернатив. Имеется множество объектов $W = \{w_i | i = 1, 2, \dots, n_o\}$, каждый из которых характеризуется n_A признаками $A = \{A_i | i = 1, 2, \dots, n_A\}$. Задано некоторое обучающее множество примеров $P = \{P_i | i = 1, 2, \dots, n_p\}$ для объектов с описанием значений признаков, и указаний класса объекта. Каждый признак A_i имеет два отличных друг от друга значения Z_{i1}, Z_{i2} .

Необходимо разработать алгоритм построения решения бинарной классификационной модели в виде дерева решений, который позволил бы классифицировать новые поступающие извне данные. Целью построения дерева решения является определение значения категориальной зависимой переменной.

В общем случае маршрут M_k на дереве решений включает n_A вершин и n_A ребер. Вершины x_i и x_{i+1} маршрута M_k соответствуют признакам A_i, A_{i+1} . Каждое ребро u_{ij} соответствует одному из двух значений $Z_{it}(t=1,2)$ признака A_i , выходит из x_i и входит в x_{i+1} . Последнее ребро в маршруте выходит из последней вершины списка вершин S_k и входит в вершину L_c меткой $list$, значение которой соответствует номеру распознаваемого класса. Значения ребер, входящих в маршрут M_k задается списком W_k .

Пример. Для построения классификатора сорта пшеница задана обучающая выборка $P=\{P_i|i=1,2,\dots,n_p\}$, представленная в таблице 1. Каждый признак A_i имеет два значения Z_{i1}, Z_{i2} .

Таблица 1

Обучающая выборка

№	Признаки				
	A_1 . Влажность в%. Не более	A_2 . Сорная примесь в%. Не более	A_3 . Почерневшие Ядра в%. Не более	A_4 . Неочищенные ядра в%. Не более	Сорт
p_1	$Z_{11} \leq 10\%$	$Z_{21} \leq 0.2\%$	$Z_{31} \leq 2\%$	$Z_{41} = 0$	2
p_2	$Z_{11} \leq 10\%$	$Z_{21} \leq 0.2\%$	$Z_{31} \leq 2\%$	$Z_{42} \leq 2\%$	1
p_3	$Z_{11} \leq 10\%$	$Z_{21} \leq 0.2\%$	$Z_{32} = 0$	$Z_{42} \leq 2\%$	1
p_4	$Z_{12} \leq 18\%$	$Z_{21} \leq 0.2\%$	$Z_{32} = 0$	$Z_{42} \leq 2\%$	1
p_5	$Z_{12} \leq 18\%$	$Z_{22} \leq 0.3\%$	$Z_{32} = 0$	$Z_{42} \leq 2\%$	2
p_6	$Z_{12} \leq 18\%$	$Z_{21} \leq 30.2\%$	$Z_{32} = 0$	$Z_{41} = 0$	1
p_7	$Z_{11} \leq 10\%$	$Z_{22} \leq 0.3\%$	$Z_{31} \leq 2\%$	$Z_{41} = 0$	2

Рассмотрим построенный на базе обучающей выборки в дереве решений маршрут M_k конечная вершина которого соответствует заданному классу (1 сорту). ($M_k = \langle x_2, u_{24}; x_4, u_{41}; x_1, u_{13}; x_3, u_{3L}; L \rangle$. $u_{24} = Z_{21}, u_{41} = Z_{42}, u_{13} = Z_{11}, u_{3L} = Z_{32}$. $L = 1 \text{ сорт}$).

Маршруту M_k соответствует фрагмент классификатора, представленный на рис. 1. Все вершины БДР обладают памятью. В памяти внутренней вершины БДР хранится информации о признаке и его значениях, в соответствии с которыми вершина подвергалась разбиению. В памяти конечной вершины (листа) хранится информация о номере класса (сорта), в состав которого входят примеры данной вершины, и о последовательности признаков, входящих в состав маршрута M_k . В результате прохождения от корня дерева до конечной вершины решается задача о принадлежности объекта к номеру класса, хранящемуся в конечной вершине. Введем обозначения.

n_i – число примеров с признаком A_i .

n_{ij} – число примеров j -го сорта с признаком A_i .

Параметр $\Pi = (\pi_1, \pi_2)$ фиксирует отношение числа π_1 примеров первого сорта к числу π_2 примеров второго сорта.

В качестве оценки качества классификации выбрана величина $F_o = (n_o - n_o^*) / n_o$, где n_o – общее число объектов, n_o^* – количество правильно классифицированных объектов.

В работе на этапе построения модели формируется упорядоченная последовательность признаков, входящих в состав маршрута на дереве решений от корневой вершины к висячей вершине. Построение маршрута заканчивается, если достигнуто минимальное значение F_o (нулевое значение) или глубина поиска C (число признаков в последовательности) достигла предельного значения – C_{max} . При этом висячая вершина объявляется листом. Оценкой маршрута в первом случае является параметр C . Во втором случае оценкой маршрута является параметр F :

$$F = \alpha F_o + \beta C,$$

где α, β – коэффициенты пропорциональности.

Цель оптимизации – минимизация критерия F .

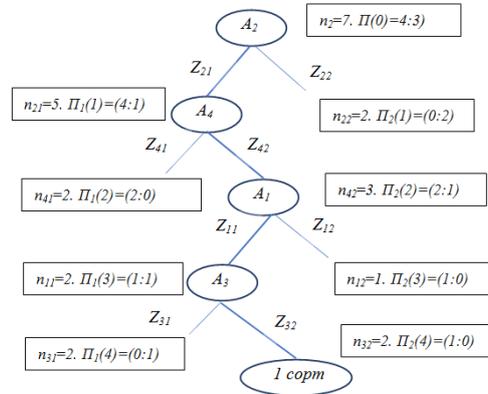


Рис. 1. Фрагмент классификатора сорта пшеницы, соответствующий маршруту M_k

2. Механизмы построения классификатора на основе метода кристаллизации россыпи альтернатив. В алгоритме построения классификатора на основе методе КРА [10-14] каждое решение Q_k интерпретируется в виде ориентированного маршрута M_k на БДР. Назовем порядковый номер элемента в ориентированном маршруте M_k позицией $s_i \in S = \{s_i | i=1, 2, \dots, n_A\}$. Элементом маршрута M_k является пара (x_i, u_i) , где x_i соответствует признаку A_i . Элемент u_i в маршруте M_k является ребром, выходящим из x_i и соответствует выбранному вместе с A_i значению A_i . Второй индекс элемента u_i определится после выбора A_i , помещенного в соседнюю с s_j позицию s_{j+1} .

Каждый агент A_i с выбранным значением может располагаться в одной из альтернативных позиций $s_i \in S = \{s_i | i=1, 2, \dots, n_A\}$, в соответствии с маршрутом M_k .

В каждой из позиций может располагаться только один агент. Каждый агент может иметь одно из двух значений. Решение Q_k определяется совокупностью альтернативных значений и позиций множества агентов A .

В работе под *россыпью альтернатив* (РА) решения Q_k называется структура данных, в виде матрицы $R_k = \|r_{kijl}\|_{m \times 2n}$, несущая информацию об занимаемых агентами альтернативных позициях (порядковых номерах), значениях агентов в ориентированном маршруте M_k и значениях их пригодностей, составляющих коллективную эволюционную память (КЭП) (рис.2).

Столбцы матрицы попарно соответствуют агентам, строки – альтернативным позициям, в которых может располагаться агент.

Номер строки (вектора $R_{kj} = \{r_{jl} | l=(1, 2), (3, 4), \dots, 2n_A\}$) матрицы R_k соответствует порядковому номеру j позиции в маршруте. Число строк равно числу позиций. Размер строки равен удвоенному числу агентов (признаков).

Каждому агенту A_i соответствует два столбца – два вектора $RI_{k\alpha} = \{r_{kja} | \alpha=(2i-1); i=1, 2, \dots, n_A\}$ и $R2_{k\beta} = \{r_{kjb} | \beta=2i; i=1, 2, \dots, n_A\}$. Элементы вектора $RI_{k\alpha}$ соответствуют паре (x_i, u_i) с первым значением признака. Элементы вектора $R2_{k\beta}$ соответствуют паре (x_i, u_i) со вторым значением признака. Размер векторов $RI_{k\alpha}$ и $R2_{k\beta}$ равен числу позиций агентов (признаков). Число столбцов матрицы R_k равно удвоенному числу агентов (признаков).

Для любого A_i в векторах $RI_{k\alpha}$ и $R2_{k\beta}$ только один из двух элементов r_{kja} или r_{kjb} соответствующих состоянию c_{ij} в котором находится агент имеет значение, отличное от нуля. Остальные элементы вектора R_{ki} имеют нулевые значения.

В каждой строке (векторе R_{kj}) только один элемент r_{kji} имеет значение, отличное от нуля. Для фиксации значения каждого агента A_i в векторе R_{kj} используются два соседних элемента r_{kja} и $r_{kj\beta}$, где $\alpha=(2i-1)$, $\beta=2i$: r_{kja} – для первого значения, $r_{kj\beta}$ – для второго значения.

Каждый, отличный от нуля элемент r_{kji} матрицы $R_k=||r_{kji}||_{m \times 2n}$, имеет значение, равное полезности δ_k решения Q_k , при котором агент a_i назначен в позицию j . $\delta_k=f(\zeta_k)$, где ζ_k – оценка решения, а δ_k – оценка полезности этого решения.

Пусть для формирования квалификационной модели построено множество решений $Q=\{Q_k/k=1,2,\dots,n_k\}$.

Рассмотрим процедуру отображения решения Q_k в матрице R_k . В качестве примера используем рассмотренный выше маршрут M_k конечная вершина которого соответствует заданному классу (1 сорту).

$(M_k=\langle x_2, u_{24}; x_4, u_{41}; x_1, u_{13}; x_3, u_{3L}; L \rangle. u_{24}=Z_{21}, u_{41}=Z_{42}, u_{13}=Z_{11}, u_{3L}=Z_{32}. L=1 \text{ сорт})$.

Решение Q_k формируется четырьмя агентами. Для решения Q_k рассчитаны оценки ζ_k и δ_k .

Для отображения решения Q_k в матрице R_k формируется набор ячеек Θ_k , соответствующих состояниям агентов в маршруте M_k .

Первый элемент в маршруте M_k – агент A_2 , следовательно, номер агента $i=2$, номер позиции агента $j=1$. Выбрано первое значение признака A_2 . Данной ситуации в матрице R_k соответствует элемент r_{kja} с индексами: $j=1, \alpha=2i-1=3. r_{k13}$.

Второй элемент в маршруте M_k – агент A_4 , следовательно, номер агента $i=4$, номер позиции агента $j=2$. Выбрано второе значение признака A_4 . Данной ситуации в матрице R_k соответствует элемент $r_{kj\beta}$ с индексами: $j=2, \beta=2i=8. r_{k28}$

Третий элемент в маршруте M_k – агент A_1 , следовательно, номер агента $i=1$, номер позиции агента $j=3$. Выбрано первое значение признака A_1 . Данной ситуации в матрице R_k соответствует элемент r_{kja} с индексами: $j=3, \alpha=2i-1=1. r_{k31}$

Четвертый элемент в маршруте M_k – агент A_3 , следовательно номер агента $i=3$, номер позиции агента $j=4$. Выбрано второе значение признака. Данной ситуации в матрице R_k соответствует элемент $r_{kj\beta}$ с индексами: $j=4, \beta=2i=6. r_{k46}$

Все определенные ячейки включаются в набор $\Theta_k=\{r_{k13}, r_{k28}, r_{k31}, r_{k46}\}$.

После формирования набора Θ_k и расчета оценки полезности δ_k этого решения все элементы набора Θ_k в матрице R_k увеличиваются на величину δ_k .

На рис. 2. всем элементам набора Θ_k в матрице R_k присвоено значение δ_k .

Позиция	Признаки							
	A_1		A_2		A_3		A_4	
	Z_{11}	Z_{12}	Z_{21}	Z_{22}	Z_{31}	Z_{32}	Z_{41}	Z_{42}
1			δ_k					
2								δ_k
3	δ_k							
4						δ_k		

Рис. 2. Россыпь альтернатив R_k

3. Построение классификатора (маршрута) в БДР с заданным значением конечной вершины (листа). Работа алгоритма построения квалификационной модели базируется на использовании коллективной эволюционной памяти (КЭП), под которой подразумевается любой вид информации, которая отражает прошлую историю развития и хранится независимо от индивидуумов. Алгоритм, связанный с эволюционной памятью, стремится к запоминанию и многократному использованию способов достижения лучших результатов. КЭП алгоритма построение

классификатора представляет собой набор статистических показателей, отражающих для каждого фрагмента решения число μ вхождений фрагмента в лучшие решения на предыдущих итерациях алгоритма и число δ показывающее полезность фрагмента при построении решений на предыдущих итерациях алгоритма. Рассматриваемый алгоритм относится к классу популяционных. Процесс поиска решений итерационный. В процессе поиска эволюционирует популяция решений. Дано: обучающая выборка; множество признаков; бинарные значения признаков; метка класса для концевой вершины. Каждая итерация l включает три этапа.

На первом этапе каждой итерации конструктивным алгоритмом формируется множество решений $Q = \{Q_k | k=1, 2, \dots, n_k\}$. Формирование каждого решения Q_k выполняется конструктивным алгоритмом *Маршрут*, путем последовательного выбора на каждом шаге признака (агента A_i), одного из двух значений признака $Z_{it}(t=1, 2)$ A_i , и позиции s_j , в которую назначается агент A_i . Работа конструктивного алгоритма базируется на базе показателей основной интегральной россыпи альтернатив – матрицы $R = \|r_{ij}\|_{m \times 2n}$, структура которой описана выше, и в которой хранятся интегральные показатели решений, полученных на предыдущих итерациях. Формирование основной интегральной россыпи альтернатив осуществляется в два этапа с использованием основной R и промежуточной R^* матриц россыпей альтернатив.

После построения очередного решения Q_k его показатели заносятся в промежуточную россыпь альтернатив – матрицу R^* . Для этого рассчитывается оценка ζ_k и оценка полезности δ_k решения Q_k (маршрута M_k). Далее формируется набор Θ_k ячеек для отображения решения Q_k в матрице R^* . После расчета оценки полезности δ_k и формирования Θ_k все элементы набора Θ_k в матрице R^* увеличиваются на величину δ_k . В работе используется циклический метод формирования решений. В этом случае наращивание оценок интегральной полезности δ_k в основной интегральной россыпи альтернатив R выполняется после полного формирования множества решений Q , в результате которого будет сформирована промежуточная матрица россыпей альтернатив R^* .

На втором этапе итерации производится наращивание оценок интегральной полезности δ_k в основной интегральной россыпи альтернатив – матрице R , путем добавления матрицы R^* к матрице R . На третьем этапе итерации осуществляется снижение интегральных оценок полезности δ_k в матрице R на величину δ^* .

Формирование конструктивным алгоритмом решения Q_k (маршрута M_k) производится последовательно, пошагово. Процесс выбора на каждом шаге признака (агента A_i), одного из двух значений признака $Z_{it}(t=1, 2)$, и позиции s_j , в которую назначается агент A_i , включает две стадии. На первой стадии выбирается агент и его значение (в маршруте M_k это выбор пары (x_i, u_i)), а на второй стадии – позиция s_j . При этом должно выполняться ограничение: каждому агенту множества A соответствует один единственный элемент множества S и наоборот.

На шаге t для каждого, еще не назначенного в позицию, агента $A_i \in A(t)$, путем просмотра двух соответствующих ему столбцов матрицы R (вектора $R1_{k\alpha} = \{r_{kj\alpha} | \alpha = (2i-1); i=1, 2, \dots, n_A\}$ и $R2_{k\beta} = \{r_{kj\beta} | \beta = 2i; i=1, 2, \dots, n_A\}$) среди свободных позиций $S(t)$ отыскивается позиция $s_j \in S(t)$ с максимальным значением полезности назначения в нее A_i , которое назовем стоимостью $\varepsilon_i(t)$ агента A_i на шаге t . Среди $A_i \in A(t)$ с вероятностью $P_{ik}(t) = \varepsilon_i(t) / \sum_i \varepsilon_i(t)$, пропорциональной стоимости $\varepsilon_i(t)$ агента A_i , выбирается агент A^*_i . В маршрут M_k в позицию s_j (порядковый номер) заносится пара элементов (x_i, u_i) . Элемент u_i в маршруте M_k является ребром, выходящим из x_i , и соответствует выбранному вместе с A_i значению A_i . Второй индекс элемента u_i определится после выбора A_i , помещенного в соседнюю с s_j позицию s_{j+1} .

Предварительно формируется структура основной R и промежуточной R^* матриц россыпей альтернатив. Введем обозначения:

$A(t)$ – множество агентов не размещенных в узлах элементов на шаге t ;

$S(t)$ – множество свободных позиций шаге t ;

k – номер решения;

l – номер итерации.

Задаются параметры:

N_k – объем популяции решений;

N_l – число итераций;

δ – начальное значение оценки полезности.

Алгоритм.

1. Выбор значений параметров N_k, N_l, δ .

2. Всем элементам основной матрицы россыпи альтернатив R присваивается начальное значение оценки полезности δ .

3. $l=1$.

4. Элементы промежуточной матрицы россыпи альтернатив R^* обнуляются.

5. $k=1$.

6. Формирование конструктивным алгоритмом решения Q_k (маршрута M_k) на базе матрицы R .

7. Расчет оценки полезности δ_k решения Q_k .

8. Формирование набора Θ_k ячеек для отображения решения Q_k в матрице $R^*(l)$.

9. Все элементы набора Θ_k в матрице $R^*(l)$ увеличиваются на величину δ_k .

10. Если $(k < n_k)$, то $k=k+1$ и переход к 6, иначе переход к 11.

11. Сложение матриц R и R^* . $R=R+R^*$.

12. Снижение всех интегральных оценок полезности r_{ij} интегральной россыпи альтернатив $R(l)$ на величину δ^* .

13. Если $(l < N_l)$, то $l=l+1$ и переход к 4, иначе переход к 14.

14. Конец работы алгоритма.

4. Экспериментальные исследования. Разработанный алгоритм построения квалификационной модели реализован в виде программы построения дерева решений **БДР-КРА**.

Тестирование программы **БДР-КРА** производилось на контрольных примерах с известным оптимумом $K_{\text{опт}}$ [14–21]. Уровень качества полученных решений оценивался по показателю $P=K_{\text{опт}}/K$, где K – значение критерия оптимизации, используемого в программе **БДР-КРА**. Число итераций, при котором алгоритм достигал максимального уровня качества не превышает 135.

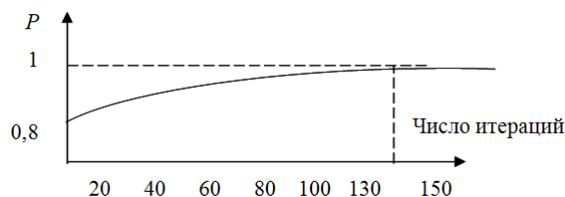


Рис. 3. Зависимость качества решений алгоритма БДР-КРА от числа итераций

Сравнительный анализ с разработанного алгоритма производился с генетическим алгоритмом и алгоритмом роя частиц. Результаты сравнения представлены в табл. 2.

Сравнение алгоритма БДР-КРА по показателю уровень качества с генетическим алгоритмом и алгоритмом роя частиц показало, что при сравнимых временных затратах у алгоритма БДР-КРА показатель P выше в среднем на 9–11%. В среднем уровень качества, достигнутый алгоритмом БДР-КРА на 135-ой итерации, отличается от максимального значения на 0,12 процента. Общая оценка временной сложности лежит в пределах $O(n^2)$ - $O(n^3)$, где n – число признаков.

Таблица 2

Сравнительная оценка работы алгоритмов

Алгоритм / Параметр		Тест 1	Тест 2	Тест 3	Тест 4	Тест 5
ГА	P	332	534	221	189	626
	t	5	12	5	72	27
РЧ	P	281	619	268	223	683
	t	4	12	7	78	34
БДР-КРА	P	268	502	209	169	611
	t	3	6	5	36	29

Заключение. В ряде случаев возникает необходимость установления соответствия между заявленным и фактическим значением категориальной переменной на основе совокупности признаков объекта. В этом случае возникает потребность в классификаторе с оптимальной последовательностью рассматриваемых атрибутов с заданным значением целевой функции. Значением целевой переменной может быть: да, нет, номер сорта, номер класса и т.д.

В работе решается задача построения классификационной модели в виде оптимальной последовательности рассматриваемых атрибутов и их значений, входящих в состав маршрута от корневой вершины к концевой вершине с заданным значением целевой переменной.

Если требуется классификатор, включающий возможность альтернативных ответов, то вначале строятся независимо друг от друга оптимальные маршруты для каждого значения целевой переменной, а затем эти маршруты объединяются («склеиваются») в единое бинарное дерево решений.

Эффективным направлением повышения качества решений стало использование стохастических **популяционных** алгоритмов, которые, как правило, итерационные и работают в пространстве полных решений. В работе по аналогии с метаэвристиками, на которых построены роевые алгоритмы, используется метаэвристика, учитывающая тенденцию к использованию альтернатив (вариантов компонентов) из наилучших найденных решений. Особенности являются наличие прямого обмена информацией – стигмержи. Совокупность данных об альтернативах и их оценках составляет **россыпь альтернатив**.

Разработан конструктивный алгоритм формирования решения (маршрута M_k). Процесс выбора на каждом шаге признака (агента A_i), одного из двух значений признака, и позиции, в которую назначается агент A_i , включает две стадии. На первой стадии выбирается агент и его значение, а на второй стадии – позиция. При этом должно выполняться ограничение: каждому агенту соответствует один единственный порядковый номер и наоборот.

Предложена эффективная методика расчета оценки полезности решения. Построение маршрута заканчивается, если достигнуто минимальное значение или глубина поиска достигла предельного значения.

В алгоритме построения классификатора на основе метода кристаллизации россыпи альтернатив, каждое решение Q_k интерпретируется в виде в ориентированного маршрута M_k на бинарном дереве решений. Работа алгоритма построения дерева решений базируется на использовании коллективной эволюционной памяти, под которой подразумевается информация, отражающая историю поиска решения. Рассмотрены ключевые моменты анализа альтернатив в процессе эволюционной коллективной адаптации. Экспериментальные исследования показали, что разработанный алгоритм находит решения, не уступающие по качеству, а иногда и превосходящие своих аналогов в среднем на 3–4 %.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Witten Ian H., Frank Eibe, Hall Mark A.* Data Mining: Practical Machine Learning Tools and Techniques. – 3rd Edition. – Morgan Kaufmann, 2011. – 178 p.
2. *Журавлев Ю.И., Рязанов В.В., Сенько О.В.* «Распознавание». Математические методы. Программная система. Практические применения. – М.: Фазис, 2006. – 122 с.
3. *Бериков В.С., Лбов Г.С.* Современные тенденции в кластерном анализе // Всероссийский конкурсный отбор обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы». – 2008. – 126 с.
4. *Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И.* Методы и модели анализа данных: OLAP и Data Mining. – СПб.: БХВ-Петербург, 2004. – 93 с.
5. *Карпенко А.П.* Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой: учеб. пособие. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2014. – 448 с.
6. *Wang X.* Hybrid nature-inspired computation method for optimization: Doctoral Dissertation. Helsinki University of Technology, TKK Dissertations, Espoo 2009. – 161 p.
7. *Лебедев Б.К., Лебедев О.Б., Лебедев В.Б.* Гибридизация роевого интеллекта и генетической эволюции на примере размещения // Программные продукты, системы и алгоритмы. – 2017. – №4.
8. *Лебедев Б.К., Лебедев О.Б.* Гибридный биоинспирированный алгоритм решения задачи символьной регрессии // Известия ЮФУ. Технические науки. – 2015. – № 6 (167). – С. 28-41.
9. *Курейчик В.М., Лебедев Б.К., Лебедев О.Б.* Поисковая адаптация: теория и практика. – М.: Физматлит, 2006. – 288 с.
10. *Лебедев Б.К., Лебедев В.Б.* Оптимизация методом кристаллизации россыпи альтернатив // Известия ЮФУ. Технические науки. – 2013. – № 7 (144). – С. 11-17.
11. *Лебедев Б.К., Лебедев О.Б., Лебедева Е.М.* Распределение ресурсов на основе гибридных моделей роевого интеллекта // Научно-технический вестник информационных технологий, механики и оптики. – 2017. – Т. 17, № 6. – С. 1063-1073.
12. *Курейчик В.В., Курейчик Вл.Вл.* Архитектура гибридного поиска при проектировании // Известия ЮФУ. Технические науки. – 2012. – № 7 (132). – С. 22-27.
13. *Лебедев Б.К., Лебедев О.Б., Лебедева Е.О.* Решение задачи символьной регрессии методами генетического поиска // Известия ЮФУ. Технические науки. – 2015. – № 2 (163). – С. 212-225.
14. *Kennedy J., Eberhart R.C.* Particle swarm optimization // In Proceedings of IEEE International Conference on Neural Networks. – 1995. – P. 1942-1948.
15. *Clerc M.* Particle Swarm Optimization. – ISTE, London, UK, 2006. – 133 p.
16. *Лебедев Б.К., Лебедев В.Б.* Эволюционная процедура обучения при распознавании образов // Известия ТРТУ. – 2004. – № 8 (43). – С. 83-88.
17. *Лебедев Б.К., Лебедев О.Б., Лебедева Е.О.* Разбиение на классы методом альтернативной коллективной адаптации // Известия ЮФУ. Технические науки. – 2016. – № 7 (180). – С. 89-101.
18. *Cong J., Romesis M., Xie M.* Optimality, Scalability and Stability Study of Partitioning and Placement Algorithms // Proc. of the International Symposium on Physical Design. – Monterey, CA, 2003. – P. 88-94.
19. *Захаров Д.О., Карпенко А.П.* Исследование эффективности популяционного алгоритма лиги чемпионов для задачи глобальной оптимизации // Математика и математическое моделирование. – 2020. – № 2. – С. 25-45.
20. *Карпенко А.П.* Методы повышения эффективности популяционных алгоритмов глобальной оптимизации // Матер. V межрегиональной научно-практической конференции. – Севастополь: Изд-во: Севастопольский государственный университет, 2019. – С. 87-88.
21. *Кучуганов В.Н., Кучуганов А.В., Каксимов Д.Р.* Алгоритм кластеризации множества деталей по чертежам // Программирование. – М.: Изд-во: Российская академия наук, 2020. – С. 29-38.

REFERENCES

1. *Witten Ian H., Frank Eibe, Hall Mark A.* Data Mining: Practical Machine Learning Tools and Techniques. 3rd Edition. Morgan Kaufmann, 2011, 178 p.
2. *Zhuravlev Yu.I., Ryzanov V.V., Sen'ko O.V.* «Raspознаvanie». Matematicheskie metody. Programmnyaya sistema. Prakticheskie primeneniya [Recognition". Mathematical methods. Software system. Practical applications]. Moscow: Fazis, 2006, 122 p.

3. *Berikov V.S., Lbov G.S.* Sovremennye tendentsii v klasternom analize [Modern trends in cluster analysis], *Vserossiyskiy konkursnyy otbor obzorno-analiticheskikh statey po prioritetnomu napravleniyu «Informatsionno-telekommunikatsionnye sistemy»* [All-Russian competitive selection of review and analytical articles in the priority direction "Information and telecommunication systems], 2008, 126 p.
4. *Barsegyan A.A., Kupriyanov M.S., Stepanenko V.V., Kholod I.I.* Metody i modeli analiza dannykh: OLAP i Data Mining [Methods and models for data analysis: OLAP and Data Mining]. Saint Petersburg: BKhV-Peterburg, 2004, 93 p.
5. *Karpenko A.P.* Sovremennye algoritmy poiskovoy optimizatsii. Algoritmy, vdokhnovlennyye prirodoy: ucheb. posobie [Modern search engine optimization algorithms. Algorithms inspired by nature: textbook]. Moscow: Izd-vo MGTU im. N.E. Baumana, 2014, 448 p.
6. *Wang X.* Hybrid nature-inspired computation method for optimization: Doctoral Dissertation. Helsinki University of Technology, TKK Dissertations, Espoo 2009, 161 p.
7. *Lebedev B.K., Lebedev O.B., Lebedev V.B.* Gibrizatsiya roevogo intellekta i geneticheskoy evolyutsii na primere razmeshcheniya [Hybridization of swarm intelligence and genetic evolution on the example of placement], *Programmnye produkty, sistemy i algoritmy* [Software products, systems and algorithms], 2017, No.4.
8. *Lebedev B.K., Lebedev O.B.* Gibridnyy bioinspirirovannyy algoritm resheniya zadachi simvol'noy regressii [Hybrid bioinspired algorithm for solving the problem of symbolic regression], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, No. 6 (167), pp. 28-41.
9. *Kureychik V.M., Lebedev B.K., Lebedev O.B.* Poiskovaya adaptatsiya: teoriya i praktika [Search engine adaptation: theory and practice]. Moscow: Fizmatlit, 2006, 288 p.
10. *Lebedev B.K., Lebedev V.B.* Optimizatsiya metodom kristallizatsii rossypi al'ternativ [Optimization by the method of crystallization of alternatives field], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2013, No. 7 (144), pp. 11-17.
11. *Lebedev B.K., Lebedev O.B., Lebedeva E.M.* Raspredelenie resursov na osnove gibridnykh modeley roevogo intellekta [Resource allocation based on hybrid models of swarm intelligence], *Nauchno-tekhnicheskiiy vestnik informatsionnykh tekhnologiy, mekhaniki i optiki* [Scientific and technical bulletin of information technologies, mechanics and optics], 2017, Vol. 17, No. 6, pp. 1063-1073.
12. *Kureychik V.V., Kureychik V.I.* Arkhitektura gibridnogo poiska pri proektirovanii [Architecture of hybrid search in design], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2012, No. 7 (132), pp. 22-27.
13. *Lebedev B.K., Lebedev O.B., Lebedeva E.O.* Reshenie zadachi simvol'noy regressii metodami geneticheskogo poiska [The solution of the problem of symbolic regression by methods of genetic search], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, No. 2 (163), pp. 212-225.
14. *Kennedy J., Eberhart R.C.* Particle swarm optimization, *In Proceedings of IEEE International Conference on Neural Networks*, 1995, pp. 1942-1948.
15. *Clerc M.* Particle Swarm Optimization. ISTE, London, UK, 2006, 133 p.
16. *Lebedev B.K., Lebedev V.B.* Evolyutsionnaya protsedura obucheniya pri raspoznavanii obrazov [Evolutionary training procedure for pattern recognition], *Izvestiya TRTU* [Izvestiya TSURE], 2004, No. 8 (43), pp. 83-88.
17. *Lebedev B.K., Lebedev O.B., Lebedeva E.O.* Razbienie na klassy metodom al'ternativnoy kollektivnoy adaptatsii [Division into classes by the method of alternative collective adaptation], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2016, No. 7 (180), pp. 89-101.
18. *Cong J., Romesis M., Xie M.* Optimality, Scalability and Stability Study of Partitioning and Placement Algorithms, *Proc. of the International Symposium on Physical Design*. Monterey, CA, 2003, pp. 88-94.
19. *Zakharov D.O., Karpenko A.P.* Issledovanie effektivnosti populyatsionnogo algoritma ligi chempionov dlya zadachi global'noy optimizatsii [Research of the efficiency of the population algorithm of the Champions League for the global optimization problem], *Matematika i matematicheskoe modelirovanie* [Mathematics and Mathematical Modeling], 2020, No. 2, pp. 25-45.

20. *Karpenko A.P.* Metody povysheniya effektivnosti populyatsionnykh algoritmov global'noy optimizatsii [Methods for increasing the efficiency of population algorithms for global optimization], *Mater. V mezhhregional'noy nauchno-prakticheskoy konferentsii* [Proceedings of the V interregional scientific and practical conference]. Sevastopol': Izd-vo: Sevastopol'skiy gosudarstvennyy universitet, 2019, pp. 87-88.
21. *Kuchuganov V.N., Kuchuganov A.V., Kaksimov D.R.* Algoritm klasterizatsii mnozhestva detaley po chertezham [Algorithm for clustering a set of parts according to drawings], *Programmirovaniye* [Programming]. Moscow: Izd-vo: Rossiyskaya akademiya nauk, 2020, pp. 29-38.

Статью рекомендовал к опубликованию д.т.н., профессор А.Г. Коробейников.

Лебедев Борис Константинович – Южный федеральный университет; e-mail: lebedev.b.k@gmail.com; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 89282897933; кафедра систем автоматизированного проектирования; профессор.

Лебедев Олег Борисович – e-mail: lebedev.ob@mail.ru; тел.: 89085135512; кафедра систем автоматизированного проектирования; доцент.

Лебедев Владимир Борисович – Московский государственный технический университет имени Н.Э. Баумана; e-mail: lebedev.vlad.bor@mail.ru; 105005, г. Москва, ул. Бауманская 2-я, д. 5, стр. 1; тел.: 89287775005; Научно-производственное объединение «Новые технологии»; с.н.с.

Lebedev Boris Konstantinovich – Southern Federal University; e-mail: lebedev.b.k@gmail.com; 44, Nekrasovsky, Taganrog, 347928, Russia; phone: +79282897933; the department of computer aided design; professor.

Lebedev Oleg Borisovich – e-mail: lebedev.ob@mail.ru; phone: +79085135512; the department of computer aided design; associate professor.

Lebedev Vladimir Borisovich – Moscow State Technical University named after N.E. Bauman; e-mail: lebedev.vlad.bor@mail.ru; 105005, Moscow, st. Baumanskaya 2-nd, 5, build. 1; phone: +79287775005; Research and Production Association "New Technologies"; senior researcher.

УДК 004.041

DOI 10.18522/2311-3103-2020-4-93-107

А.А. Сорокин, И.М. Бородинский, А.В. Дагаев

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ВОССТАНОВЛЕНИЯ ПРОПУЩЕННЫХ ДАННЫХ

В последние десятилетия качественно развиваются методы системного анализа, что связано с увеличением скорости технического развития, уплотнением временных процессов, быстрым ростом накапливаемой информации и новыми возможностями вычислительной техники. К этим методам относятся методы анализа большого объема данных, методы добычи данных, методы аналитического моделирования, методы параллельной обработки данных, нейросетевые методы, методы прогнозирования и другие. Представленные методы позволяют быстро и качественно обрабатывать разнородные кластеры информации, аккумулировать и синтезировать данные, обобщать и классифицировать информацию. К последним из представленных методов относятся методы интерполяции и экстраполяции потерянной, поврежденной или неполученной информации. Данные методы позволяют структурировать, восстанавливать и моделировать информацию на основе статистических данных, математических и алгоритмических методов. Таким образом в статье рассматривается проблема восстановления пропущенных данных в графических и сложных объектах. Приводятся литературные источники по рассматриваемым задачам. В них приводится обширная информация по рассматриваемой тематике: представлены генетические алгоритмы используемые для пространственной интерполяции; рассмотрено решение задач неоднородности интерполяции сейсмических данных; описано использо-

вание сплайн-аппроксимации для расчета характеристик нелинейных электронных компонентов; разобран метод построения модели трехмерных параметрических рациональных тел с помощью обобщенной интерполяции Безье, что позволяет моделировать форму тела и анизотропное пространство; описаны методы применяющие нечеткие линейные уравнения, которые широко распространены в компьютерном зрении; исследован метод адаптивной интерполяции на основе градиента учитывающий локальный градиент исходного изображения. В статье выполняется сравнение нескольких распространенных методов интерполяции и реставрации данных, таких как: билинейная интерполяция, поверхность Безье. Кратко описывается каждый метод и особенности его применения в рамках проведенного эксперимента. Приводится результат серии экспериментов с представленными методами с различным количеством испытаний. В заключении делаются выводы о рациональности выбора одного из предложенных методов без применения длительного натурного эксперимента в каждом случае.

Двумерные экспериментальные данные; интерполяция данных; реставрация данных; анализ методов интерполяции.

A.A. Sorokin, I.M. Borodyansky, A.V. Dagaev

COMPARATIVE ANALYSIS OF MISSING DATA RECOVERY METHODS

In recent decades, the methods of system analysis have been developing qualitatively. It is associated with an increase in the rate of technical development, the densification of time processes, the rapid growth of accumulated information and new capabilities of computer technology. These include methods for analyzing large amounts of data, methods of data mining, methods of analytical modeling, methods of parallel data processing, neural network methods, forecasting methods, and others. The presented methods make it possible to quickly and efficiently process heterogeneous clusters of information, accumulate and synthesize data, generalize and classify information. The last of the presented methods are methods of interpolation and extrapolation of lost, damaged or missing information. These methods allow to structure, restore and model information based on statistical data, mathematical and algorithmic methods. Thus, the article deals with the problem of recovering missing data in graphic and complex objects. Literary sources on the problems under consideration are given. They provide extensive information on the topic under consideration: present genetic algorithms used for spatial interpolation; the solution of problems of heterogeneity of interpolation of seismic data is considered; it is described the use of spline approximation to calculate the characteristics of nonlinear electronic components; the method of constructing a model of three-dimensional parametric rational bodies using generalized Bezier interpolation is analyzed, which allows modeling the shape of a body and anisotropic space; methods using fuzzy linear equations are described, which are widespread in computer vision; the method of adaptive interpolation based on the gradient and taking into account the local gradient of the original image is investigated. It is made comparing several common methods of interpolation and data restoration, in article, such as: bilinear interpolation, Bezier surface. Each method and features of its application within the framework of the experiment are briefly described. The result of a series of experiments with the presented methods with different numbers of tests is presented. In conclusion, summary is drawn about the rationality of choosing one of the proposed methods without the use of a long field experiment in each case.

Two-dimensional experimental data; data interpolation; data restoration; analysis of interpolation methods.

1. Введение. Сегодня анализ информации используется во всех областях знания человека. Моделирование и эксперименты создают базу для качественного развития науки. Анализ данных становится возможен только при наличии статистической информации, которая может быть получена физическим и математическим экспериментальным путём. Однако на практике ученые часто сталкиваются с проблемой нехватки информации или погрешности и неточности полученных данных. Кроме того, все чаще возникают задачи предсказания поведения выходной функции по имеющейся статистике. Таким образом, при сборе статистической

информации часто приходится дополнять недостающие данные модельными данными, представляющими характеристики состояния системы в требуемый промежуток времени. Задачи аппроксимации и интерполяции данных стали появляться с развитием компьютерной техники, а также алгоритмических и математических методов. Известны работы с шестидесятых годов прошлого столетия [22], за последние десятилетия появилось много методов данного направления исследования [1–15], но вопрос применимости методов при различном объеме и качестве исходных данных стоит и сейчас. В [16–18] представлены современные нелинейные методы фильтрации на основе PDE (Partial Differential Equations) и методы непрозрачного изображения, представлены основы теории интерполяции, приведена классификация методов, рассматриваются точные и приближенные, однопараметрические и многопараметрические методы. В источниках [19–21] дано описание теории интерполяции линейных операторов. Излагаются основные методы построения интерполяционных пространств, изучаются их свойства, рассматриваются различные методы многомерной аппроксимации и интерполяции в конечномерных вещественных и комплексных векторных пространствах. Как видно из литературных источников, теме интерполяции уделено большое внимание, однако вопросам применимости к различным наборам данных и оптимальности применения методов посвящено немного источников [17–19]. Таким образом основными аспектами при анализе статистических данных является проблема пропуска и противоречивости данных. В большинстве случаев эксперименты нужны, чтобы собрать достаточный для формирования математической модели набор данных. При таком подходе, изучаемый процесс имеет конечное математическое описание, которое на данный момент неизвестно и возникает необходимость заполнения пропущенных данных условно случайными значениями, влияние которых на разрабатываемую модель не должно быть существенным.

Целью текущего исследования было определение оптимального метода интерполяции или реставрации экспериментальных данных в двумерном пространстве, при известном допустимом пороге среднестатистического отклонения от исходных значений с учетом времени выполнения этой операции.

Следует выделить основные варианты решения задачи:

- ◆ повторные эксперименты с целью получить пропущенные данные;
- ◆ использование шаблонов для восстановления данных (на основе других экспериментов уже сформированы шаблоны возможных решений, производится выбор наиболее адекватного шаблона);
- ◆ использование некоторого метода интерполяции (на области вокруг пропущенных данных выбираются опорные точки и формируется некоторая поверхность, проходящая через них; полученные с этой поверхности значения в дальнейшем используются для заполнения пропусков).

В первом случае возможно систематическое повторение пропусков, связанное либо с отсутствием возможных значений, либо с проблемами сбора информации. Второй случай требует наличия повторяющихся или шаблонных областей значений в рамках исследованной области. Третий метод балансирует между возможным отклонением и сложностью и временем выполнения этой операции.

Следует отметить, что теория восстановления пропущенных данных постоянно развивается, появляются новые алгоритмы и модернизируются старые. Это связано с невозможностью разработки универсального алгоритма, выдающего наилучшие результаты во всех случаях. Чаще всего оптимальность отдельного метода решения приводится в источниках на конкретном примере, что априори делает алгоритм неуниверсальным [23].

В итоге, задачи исследования свелись к поиску адекватного метода для случайного эксперимента, который будет реже отклоняться от реальных значений (при условии, что все законы распределения результатов эксперимента заранее известен) в определенных временных рамках при определенном допуске объема пропущенных данных.

Общий набор методов (алгоритмов) заполнения пропусков можно представить в виде схемы, приведенной на рис. 1. Каждый конечный элемент на ней отражает группу похожих методик. Но отдельные методы внутри группы могут кардинально отличаться по результатам их применения на одних и тех же наборах данных.

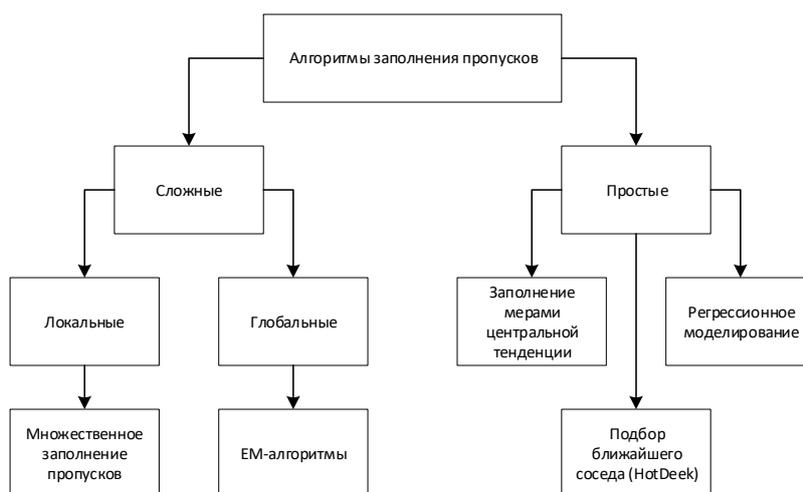


Рис. 1. Классификация алгоритмов

2. Интерполяция и реставрация двумерных данных. Одним из ключевых вопросов определения метода восстановления данных является понимание, допустимости применения существующих методов в различных ситуациях. Перед применением метода необходимо понять, что:

- ◆ при сборе исходных данных неизбежно возникла ошибка (точно измерительных приборов, наличие шумов и т.д.);
- ◆ данные были нанесены на двумерную сетку с неким шагом, т.е. фактически была произведена аппроксимация, а закон изменения между двумя соседними точками при выбранном шаге аппроксимации может кардинально различаться со случаем более мелкого шага (например, замеры с шагом 0,01 вместо 0,1);
- ◆ отсутствующие данные могут располагаться как внутри исследуемой области, так и с краю.

Последний случай связан с неполнотой классического понимания понятия интерполяция (методы нахождения точек в пределах заданного интервала), пропущенные значения расположенные на краю исследуемой области, относятся к ней, хотя, фактически, их получение с математической точки зрения должно опираться на понятие экстраполяция (методы нахождения точек за пределами заданного интервала). Тут и приходит на помощь понятие реставрация данных, по аналогии с классической реставрацией, например, старых фотографий, где, опираясь на известные фрагменты, можно попробовать частично, а иногда и полностью, воссоздать исходный вид.

Следует отметить, что оптимальный метод восстановления данных, расположенных на контурах исследуемой области, может кардинально отличаться от метода восстановления данных внутри нее. С учетом этого, в рамках данного исследования, вопрос реставрации пограничных значений не рассматривается.

3. Определение применяемых методов и условий эксперимента. Для построения адекватных моделей были определены несколько исходных неоднородных законов распределения данных в двумерном пространстве. В качестве основных отправных моментов использовались сложные тригонометрические функции с неравноценным влиянием на общую область исследования, например, существующие карты высот, карты распределения температуры и давления. Перед каждым экспериментом для соблюдения достоверности на данные в пределах области был наложен шум (случайные значения в пределах $\pm 5\%$ от разницы между минимальным и максимальным значением):

$$random\left(-\frac{|\max(A) - \min(A)|}{20}, +\frac{|\max(A) - \min(A)|}{20}\right),$$

где A – матрица значений. Это позволило в рамках исследования сделать закон распределения еще более неявным и смоделировать ситуации более близкие к реальным измерениям.

Для упрощения эксперимент был разбит на несколько итераций:

- 1) моделирование задачи для области 10×10 – выявление допустимого порога потери значений, после которого эффективность всех методов сводится на нет;
- 2) моделирование задачи для области 100×100 – общая цель та же самая, но предельное количество итераций заранее определено на основе предыдущего эксперимента;
- 3) моделирование конечной задачи с областью 1000×1000 – детализация двух предыдущих экспериментов.

В каждом случае удалялось 10% случайных значений (при условии сохранности двух крайних линий по обеим осям). Соответственно, допустимый порог доверительного интервала так же был задан 10% .

Для анализа были определены следующие методы:

- 1) заполнение нулевыми значениями;
- 2) заполнение минимальными значениями из известных;
- 3) заполнение максимальными значениями из известных;
- 4) заполнение средними значениями из известных [24];
- 5) заполнение минимальными значениями из окружающих;
- 6) заполнение максимальными значениями из окружающих;
- 7) заполнение средними значениями из окружающих;
- 8) билинейная интерполяция (4 опорные точки, образующие прямоугольную область);
- 9) квадратичная поверхность Безье (16 опорных точек, образующих прямоугольную область);
- 10) бикубическая интерполяция полиномом Лагранжа (16 опорных точек, образующих прямоугольную область);
- 11) бикубическая интерполяция натуральными сплайнами (16 опорных точек, образующих прямоугольную область);
- 12) метод итерационной взвешенной реставрации по соседним значениям [25].

Первый метод фактически не вносит корректировок (кроме случаев, когда известные значения находятся по разные стороны условной оси Oz). Следующие три метода строятся на простом принципе – заполняем одним из известных значе-

ний (в случае со средним – его необходимо сначала высчитать). Эти методы, после минимального набора действий, являются самыми быстрыми, т.к. требуют всего не более двух проходов, правда, по всей матрице.

Следующие три метода можно было бы назвать самыми быстрыми, если бы не случаи, когда окружающие, заполненные значениями ячейки матрицы, необходимо искать не для отдельной точки, а для целой области. В таких случаях число переборов возможных для решения значений может превысить разовый обход всей матрицы для первого метода. Кроме этого, необходимо определить общее минимальное количество известных значений для принятия конечного решения.

Билинейная интерполяция строится на выявление минимальной квадратной (прямоугольной) области вокруг пропущенного значения или области. Для этой операции нам потребуется четыре точки (рис. 2) – вершины нашего прямоугольника – $Q_{11}(x_1, y_1)$, $Q_{12}(x_2, y_1)$, $Q_{21}(x_1, y_2)$, $Q_{22}(x_2, y_2)$. Интерполяция производится пошагово. На первом шаге выявляются значения вспомогательных точек – R (здесь и далее будем считать, что сначала производится интерполяция в горизонтальном направлении, т.е. нахождение значений над и под пропущенными – x_r , а уже затем в вертикальном – y_r):

$$f(R_1) = \frac{x_r - x_1}{x_2 - x_1} \times (f(Q_{12}) - f(Q_{11}));$$

$$f(R_2) = \frac{x_r - x_1}{x_2 - x_1} \times (f(Q_{21}) - f(Q_{22})).$$

На втором – значение искомой точки – P :

$$f(P) = \frac{y_r - y_1}{y_2 - y_1} \times (f(R_2) - f(R_1)).$$

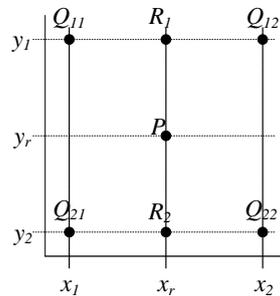


Рис. 2. Билинейная интерполяция

Интерполяция с помощью *поверхности Безье* (рис. 3) фактически можно считать идентичной билинейной линейной по сложности ее построения. Для ее построения будем брать 16 точек – четыре ряда по четыре точки (Q_{11} – Q_{44}). Вторая и третья точки будут крайними опорными (Bi_0 и Bi_2). Первая и вторая пара точек будут образовывать две прямых, на пересечении которых будет находиться промежуточная опорная точка (Bi_1) квадратичной кривой Безье:

$$\begin{cases} y_{\bar{i}} = \frac{(Q_{i1} - Q_{i2})x_{\bar{i}} + (x_1 Q_{i2} - x_2 Q_{i1})}{x_1 - x_2}; \\ y_{\bar{i}} = \frac{(Q_{i3} - Q_{i4})x_{\bar{i}} + (x_3 Q_{i4} - x_4 Q_{i3})}{x_3 - x_4}; \end{cases}$$

$$f(R_i) = \left(1 - \frac{x_r - x_2}{x_3 - x_2}\right)^2 B_{i0} + 2 \frac{x_r - x_2}{x_3 - x_2} \left(1 - \frac{x_r - x_2}{x_3 - x_2}\right) B_{i1} + \left(\frac{x_r - x_2}{x_3 - x_2}\right)^2 B_{i2}.$$

После определения ключевых значений каждой квадратичной кривой аналогичным образом определим значение конечной точки P :

$$\begin{cases} x_p = \frac{(f(R_1) - f(R_2))y_p + (y_1 f(R_2) - y_2 f(R_1))}{y_1 - y_2} \\ x_p = \frac{(f(R_3) - f(R_4))y_p + (y_3 f(R_4) - y_4 f(R_3))}{y_3 - y_4} \end{cases}$$

$$f(P) = \left(1 - \frac{y_r - y_2}{y_3 - y_2}\right)^2 f(R_2) + 2 \frac{y_r - y_2}{y_3 - y_2} \left(1 - \frac{y_r - y_2}{y_3 - y_2}\right) P + \left(\frac{y_r - y_2}{y_3 - y_2}\right)^2 f(R_3).$$

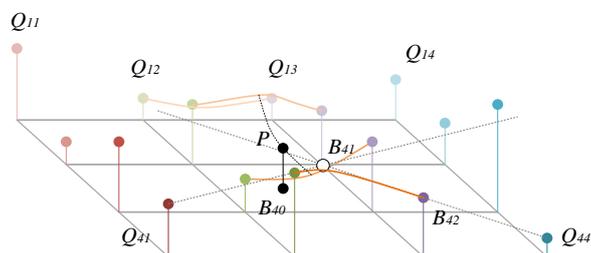


Рис. 3. Интерполяция с помощью поверхности Безье

При интерполяции с помощью *полинома Лагранжа* был выбран полином третьей степени, формирующий поверхность. Общая методика аналогична предыдущим – сначала строятся четыре опорных полинома и на каждом находится ключевая точка, потом по полученным точкам формируется конечный полином и определяется искомое значение (рис. 4):

$$L_i(x_r) = \sum_{j=1}^4 Q_{ij} l_{ij}(x_r);$$

$$l_{ij}(x_r) = \prod_{k=1, k \neq j}^4 \frac{x_r - x_k}{x_j - x_k};$$

$$L(y_p) = \sum_{i=1}^4 L_i l_i(y_p);$$

$$l_i(y_p) = \prod_{j=1, j \neq i}^4 \frac{y_p - y_j}{y_i - y_j}.$$

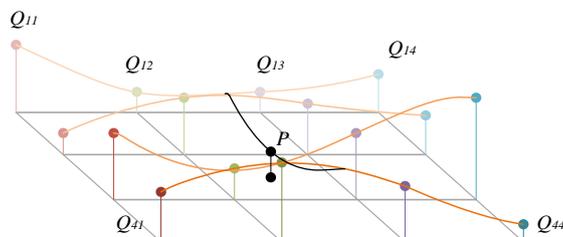


Рис. 4. Интерполяция с помощью полинома Лагранжа

Бикубическая интерполяция сплайнами один из самых популярных методов интерполяции двумерных массивов данных (например, в компьютерной графике). Существует множество методов построения сплайнов (отрезков функций, образующих общую поверхность). В нашем случае будем применять натуральные сплайны (рис. 5):

$$\begin{cases} f_{i1}(x_1) = Q_{i1} \\ f_{i1}(x_2) = Q_{i2} \\ f_{i2}(x_2) = Q_{i2} \\ f_{i2}(x_3) = Q_{i3} \\ f_{i3}(x_3) = Q_{i3} \\ f_{i3}(x_4) = Q_{i4} \\ f_{i1}(x_2) = f_{i2}(x_2); \\ f_{i2}(x_3) = f_{i3}(x_3) \\ f_{i1}(x_2) = f_{i2}(x_2) \\ f_{i2}(x_3) = f_{i3}(x_3) \\ f_{i1}(x_1) = 0 \\ f_{i3}(x_4) = 0 \end{cases}$$

$$R_i(x_r, f_{i2}(x_r)).$$

В общем виде методика сводится к получению все тех же четырех промежуточных значений (R_i) и нахождению искомого (P) на построенном по ним аналогичным образом конечном сплайне:

$$P(y_p, f_2(y_p)).$$

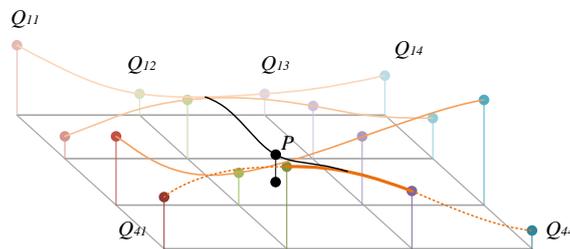


Рис. 5. Бикубическая интерполяция сплайнами

Последний метод, подробно описанный в [25], является разновидностью ZET-алгоритма [23, 26] или метода 7. Отличием от метода 7 является появление понятия вес значения, определяющийся на каждом шаге линейно относительно количества шагов, прошедших от появления значения в матрице данных. Стоит отметить, что сам метод работает лишь в комплексе с алгоритмом устранения выбросов и является частью комплекса методов поиска решения.

4. Анализ методов. В ходе эксперимента замерялось среднее отклонение конечного набора данных от исходного. Для эксперимента были взяты матрицы размерами 10×10 , 100×100 и 1000×1000 , каждая последующая размерность которых была детализацией предыдущей. На рис. 6, 7 отобразено поведение алгоритмов при первых 10 и 100 экспериментах, соответственно.

На основе полученных данных можно сделать выводы, что чаще всего наиболее адекватны исходным данным результаты методов 7, 8 и 12. Методы 7 (средние значения из известных окружающих) и 12 (итерационная взвешенная реставрация по соседним значениям) достаточно близки по реализации, метод 8 (билинейная интерполяция) использует малое значение (4 точки) вокруг искомого

го значения, что уменьшает искажение закона распределения данных по сравнению с поиском полностью описанной прямоугольной области, используемой в методах 9–11.

При этом стоит отметить, что при увеличении количества итераций и общего количества значений в исследуемом пространстве (матрице) положительно себя показывают методы, основанные на кубической интерполяции – менее 3 % отклонения. В то же время, метод 9 (квадратичная поверхность Безье), показавший в отдельных испытаниях как лучший, так и худший результат, показывает сложность построения адекватной поверхности для решения поставленной задачи, если она опирается на достаточно далеко разнесенные точки.

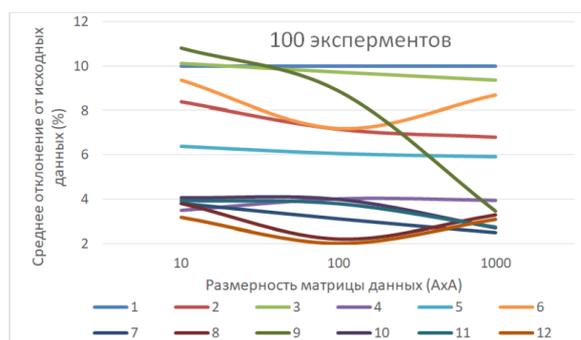


Рис. 6. Моделирование при 100 экспериментах

5. Формирование комплексного метода оценки оптимальных способов реставрации данных. Одной из целей проводимого исследования было исключение наименее точных методов интерполяции из перечня. Соответственно, результатом эксперимента стало ожидаемое полное исключение следующих методов:

- ◆ заполнение нулевыми значениями;
- ◆ заполнение минимальными значениями из известных;
- ◆ заполнение максимальными значениями из известных;
- ◆ заполнение максимальными значениями из окружающих.

Построение квадратичной поверхности Безье показало свою эффективность лишь на больших наборах значений. Остальные методы были достаточно эффективны.

Следующим этапом являлось построение метода выбора способа интерполяции для частных случаев и формирование обобщенного алгоритма. Комплексный подход заключался в анализе имеющихся фрагментов данных, поиске в них достаточно непрерывных участков, моделировании потери имеющихся данных и попытка их реставрации каждым из методов. За небольшое количество итераций (не более 10) набирается достаточно информации об отклонениях, чтобы определить, каким методом произвести полную реставрацию набора данных. В ходе экспериментов было определено, что: – для наборов данных, подчиняющихся, в целом, гладким законам (градиентные переходы, карты местности, тепловые карты и т.д.) наиболее успешными стали интерполяционные методы (8, 10, 11);

- ◆ для зашумленных и стохастических наборов данных успешнее работали методы, основанные на соседних значениях (4, 5, 7);
- ◆ для достаточно больших наборов данных ряде случаев с выгодной стороны себя показали методы 9 и 12;
- ◆ результат применения любого метода более гладкий, чем исходные данные.

Полученный в ходе эксперимента обобщенный алгоритм представлен на рис. 7. N и M – размеры матрицы значений; в случае, если обрабатывается не прямоугольная область, блок с ними можно модифицировать.

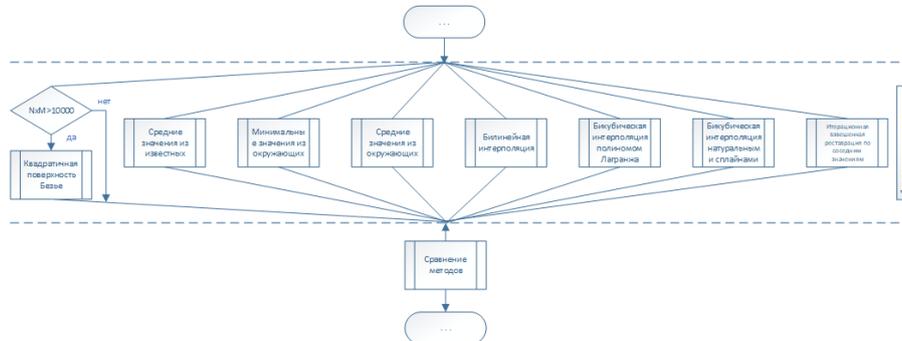


Рис. 7. Обобщенный алгоритм

Рассмотрим для примера вариант реставрации данных статистических данных. Имеется тепловая карта (рис. 9, слева), на которой присутствует дефект – в ходе замера температуры часть датчиков не работала (черная область). Попробуем провести реставрацию данных. Для этого:

- ◆ добавим контрольную область для проверки “качества” работы алгоритма (рис. 9., справа – второе черное пятно);
- ◆ выделим 3 достаточно большие прямоугольные области без поврежденных данных (рис. 9, справа);
- ◆ выполним поиск наиболее оптимального метода по обобщенному алгоритму (рис. 8).

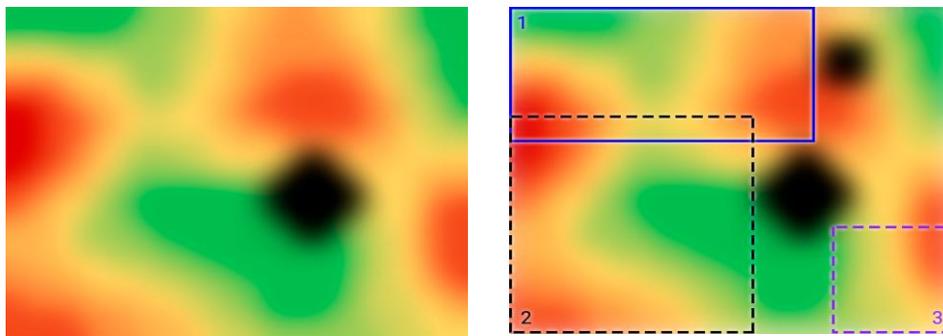


Рис. 9. Поврежденные данные: слева – исходные данные, черная область – отсутствующие значения, справа – добавлено контрольное повреждение данных и выделены 3 области для анализа

Общий объем данных слишком мал, чтобы применять метод 9, поэтому исключим его из списка анализируемых алгоритмов. Результат анализа приведен на рис. 10. Наиболее стабильно вели себя алгоритмы 7 и 4, но наименьшее среднее значение отклонения показал алгоритм 11 – бикубическая интерполяция натуральными сплайнами.



Рис. 10. Поиск оптимального алгоритма на областях для анализа

После реставрации с помощью выбранного алгоритма (подробнее было написано в пункте 3) получается полная температурная карта (рис. 11, слева). Достоверность работы алгоритма определим с помощью контрольного повреждения. Среднее отклонение значений в контрольной области получилось 2,79 % (рис. 11, справа – верхняя область). На рис. 11 отклонение показано с градациями серого. Для исходного повреждения разница считалась с 0.

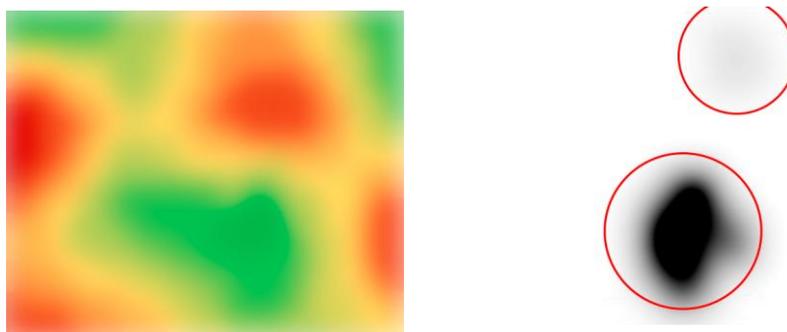


Рис. 11. Результат работы алгоритма: слева – восстановленные данные, справа – детализация разницы

Дальнейшие эксперименты с применением этого метода показали, что наиболее эффективным является выявление двух-трех наиболее оптимальных методов (чем больше исследуемых значений, тем больше методов стоит брать), реставрация с их помощью и усреднение конечных результатов. В таком случае показатель отклонения по отдельным случаям становится робастным, а среднее отклонение укладывается в пределах 2,25 %. В то же самое время попытка восстановления растровых изображений (JPEG со сжатием до 75 %) не дает желаемого результата с точки зрения визуального восприятия.

Заключение. Результатом поставленного эксперимента стало общее представление адекватности результатов некоторых из существующих методов восстановления потерянных экспериментальных данных исходным данным и получение комплексного метода оценки оптимальных способов реставрации данных. В качестве наиболее оптимальных методов для интерполяции (реставрации) потерянных данных, опираясь на результаты исследования, можно порекомендовать методы, основанные на определении среднего значения соседних с потерянным. В случае длительных экспериментов и обработки больших объемов данных, возможно, лучше себя покажут бикубические методы.

Независимо от результатов данного эксперимента, для каждого случая правильным будет провести базовый анализ, например, по приведенной выше методике. На основе его результатов можно будет определить один или несколько методов, с помощью которых можно провести восстановление данных и дальнейшую их обработку. Также в дальнейшем можно будет разработать кворум функцию, которая будет: оптимизировать существующие алгоритмы, создавать смешанные алгоритмы, выбирать оптимальный алгоритм в зависимости от начальных условий и объема статистических данных. Кроме того, в дальнейшем возможна модификация алгоритма с учетом градиентного и SVR методов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Wei Liu, Dongmei Zhang, Ao Wang*. Research of Spatial Data Interpolation Algorithm Based on SVR Optimization by GA // *Computational Intelligence and Intelligent Systems*. – 2010. – P. 137-145.
2. *Xihua Yang, Xiaojin Xie, De Li Liu, Fei Ji, Lin Wang*. Spatial Interpolation of Daily Rainfall Data for Local Climate Impact Assessment over Greater Sydney Region // *Advances in Meteorology*. – 2015.
3. *Zhihua Tian, Bin Xie, Yong Pan, Ningbo Wang, Hao Cheng*. Uermainaiti Wumaietjiang. Data Interpolation Method Based on Minimum Total Variation Method // *Proceedings of the 2019 International Conference on Artificial Intelligence and Computer Science*. – 2019. – P. 682-685.
4. *Зайцева Е.В.* Выбор метода интерполяции при решении задачи статистического восстановления непрерывных распределений геологических параметров при построении информационных систем экологического мониторинга // *Горный информационно-аналитический бюллетень*. – 2002. – № 2. – С. 197-198.
5. *Мехова Г.А., Руденко Н.В., Трофименко В.Н. Трофименко Е.Н.* Интерполяция экспериментальных данных при определении характеристик и параметров нелинейных электронных компонентов // *Тр. северо-кавказского филиала московского технического университета связи и информатики*. – 2014. – № 1. – С. 278-280.
6. *Аюшев Т.В., Булычев Р.Н.* Моделирование параметрических твердых тел с применением обобщенной линейной интерполяции // *26-я международная конференция "Graphicson 2016" АНО научного общества «Графикон» и Нижегородского государственного архитектурно-строительного университета*. – 2016. – С. 72-76.
7. *Жаринов И.О., Жаринов О.О.* Интерполяция линии спектральных цветностей цветового локуса на основе сплайнов Безье // *Известия высших учебных заведений. Приборостроение*. – 2015. – Т. 58 (12). – С. 985-992.
8. *Abd Wahab, Rozaimi Zakaria, Jamaludin Ali*. Fuzzy Interpolation Rational Bezier Curve // *Ninth International Conference on Computer Graphics, Imaging and Visualization*. – 2012. – P. 63-67.
9. *Аюшев Т.В., Булычев Р.Н.* Моделирование параметрических рациональных тел с использованием обобщенной интерполяции Безье // *Вестник Бурятского государственного университета. Математика, информатика*. – 2018. – № 1. – С. 83-94.
10. *Аюшев Т.В., Павлова С.В., Булычев Р.Н.* Геометрическое моделирование многослойных конструкций с применением кривых и поверхностей безье // *Динамика систем, механизмов и машин*. – 2016. – № 3. – С. 125-128.
11. *Федорук В.А.* Обработка экспериментальных данных на основе "Методики сдвига" при интерполяции кубическими сплайнами // *Вестник сибирской государственной автомобильно-дорожной академии*. – 2016. – № 2 (48). – С. 132-136.
12. *Ромаданова М.М.* Алгоритмы построения монотонного весового кубического сплайна // *Известия тульского государственного университета. Технические науки*. – 2018. – № 9. – С. 180-192.
13. *Yanlin Geng, Tong Lin, Zhouchen Lin, Pengwei Hao*. Refined Exponential Filter with Applications to Image Restoration and Interpolation. *Computer Vision - ACCV 2009*. – P. 33-42. – https://www.researchgate.net/publication/220745293_Refined_Exponential_Filter_with_Applications_to_Image_Restoration_and_Interpolation.
14. *Щерба Е.В.* Анализ применимости методов интерполяции и экстраполяции для решения задачи восстановления изображения // *Компьютерная оптика*. – 2009. – Т. 33 (3). – С. 336-339.

15. *Jinyu Chu, Ju Liu, Jianping Qiao, Xiaoling Wang, Yujun Li.* Gradient-based adaptive interpolation in super-resolution image restoration // International Conference on Signal Processing Proceedings. – 2009.
16. *Tudor Barbu.* Novel Diffusion-Based Models for Image Restoration and Interpolation. – 2019.
17. *Половко А.М., Бутусов П.Н.* Интерполяция. Методы и компьютерные технологии их реализации. – М.: БХВ-Петербург, 2016. – 320 с.
18. *Рассел Джесси.* Бикубическая интерполяция. – М.: Книга по Требованию, 2013. – 814 с.
19. *Асташкин С.В.* Интерполяция операторов и ее приложения. – М.: Книга по требованию, 2013. – 188 с.
20. *Крейн С.Г., Петунин Ю.И., Семенов Е.М.* Интерполяция линейных операторов. – М.: Главная редакция физико-математической литературы издательства "Наука", 2014. – 400 с.
21. *Михеев С.Е.* Многомерная аппроксимация и интерполяция. – СПб.: Изд-во С.-Петерб. ун-та, 2012. – 59 с.
22. *Уолли Дж.Л.* Интерполяция и аппроксимация рациональными функциями в комплексной области. – М.: Изд-во иностранной литературы, 1961. – 508 с.
23. *Мартышенко С.Н.* Методы восстановления пропусков в данных, представленных в различных измерительных шкалах // Территория новых возможностей. Вестник Владивостокского государственного университета экономики и сервиса. – 2013. – № 4 (22). – С. 242-255.
24. *Злоба Е., Яцкив И.* Статистические методы восстановления пропущенных данных // Computer Modelling & New Technologies. – 2002. – Vol. 6, No. 1. – P. 51-61.
25. *Сорокин А.А., Коваленко Р.А., Яковлева Е.А.* Прогнозирование стоимости аренды складских помещений на основе статистических данных // Евразийский союз ученых. – 2018. – № 12 (57). – С. 59-62.
26. *Рыженкова К.В.* Методы восстановления пропуска данных при проведении статистических исследований // Интеллект. Инновации. Инвестиции. – 2012. – № 3. – С. 127-133.

REFERENCES

1. *Wei Liu, Dongmei Zhang, Ao Wang.* Research of Spatial Data Interpolation Algorithm Based on SVR Optimization by GA, *Computational Intelligence and Intelligent Systems*, 2010, pp. 137-145.
2. *Xihua Yang, Xiaojin Xie, De Li Liu, Fei Ji, Lin Wang.* Spatial Interpolation of Daily Rainfall Data for Local Climate Impact Assessment over Greater Sydney Region, *Advances in Meteorology*, 2015.
3. *Zhihua Tian, Bin Xie, Yong Pan, Ningbo Wang, Hao Cheng.* Uermaimaiti Wumaierjiang. Data Interpolation Method Based on Minimum Total Variation Method, *Proceedings of the 2019 International Conference on Artificial Intelligence and Computer Science*, 2019, pp. 682-685.
4. *Zaytseva E.V.* Vybór metoda interpolýatsii pri reshenii zadachi statisticheskogo vosstanovleniya nepreryvnykh raspredeleniy geologicheskikh parametrov pri postroenii informatsionnykh sistem ekologicheskogo monitoringa [Selection of the interpolation method for solving the problem of statistical restoration of continuous distributions of geological parameters in the construction of information systems for environmental monitoring], *Gornyy informatsionno-analiticheskiy byulleten'* [Mining information and analytical Bulletin], 2002, No. 2, pp. 197-198.
5. *Mekhova G.A., Rudenko N.V., Trofimenko V.N., Trofimenko E.N.* Interpolyatsiya eksperimental'nykh dannykh pri opredelenii kharakteristik i parametrov nelineynykh elektronnykh komponentov [Interpolation of experimental data in determining the characteristics and parameters of nonlinear electronic components], *Tr. severo-kavkazskogo filiala moskovskogo tekhnicheskogo universiteta svyazi i informatiki* [Proceedings of the North Caucasus branch of the Moscow technical University of communications and Informatics], 2014, No. 1, pp. 278-280.
6. *Ayushev T.V., Bulychev R.N.* Modelirovanie parametricheskikh tverdykh tel s primeneniem obobshchennoy lineynoy interpolyatsii [Modeling of parametric solids using generalized linear interpolation], *26-ya mezhdunarodnaya konferentsiya "Graphicon 2016" ANO nauchnogo obshchestva «Grafikon» i Nizhegorodskogo gosudarstvennogo arkhitekturno-stroitel'nogo universiteta* [26th International Conference "Graphicon 2016" of the Scientific Society "Grafikon" and Nizhny Novgorod State Architectural and Construction University], 2016, pp. 72-76.

7. Zharinov I.O., Zharinov O.O. Interpolyatsiya linii spektral'nykh tsvetnostey tsvetovogo lokusa na osnove splaynov Bez'e [Interpolation of the spectral chromaticity line of a color locus based on Bezier splines], *Izvestiya vysshikh uchebnykh zavedeniy. Priborostroenie* [Izvestia of higher educational institutions. Equipment], 2015, Vol. 58 (12), pp. 985-992.
8. Abd Wahab, Rozaimi Zakaria, Jamaludin Ali. Fuzzy Interpolation Rational Bezier Curve, *Ninth International Conference on Computer Graphics, Imaging and Visualization*, 2012, pp. 63-67.
9. Ayushev T.V., Bulychev R.N. Modelirovanie parametricheskikh ratsional'nykh tel s ispol'zovaniem obobshchennoy interpolyatsii Bez'e [Modeling of parametric rational bodies using generalized interpolation Bezier], *Vestnik Buryatskogo gosudarstvennogo universiteta. Matematika, informatika* [Vestnik Buryat State University. Mathematics, informatics], 2018, No. 1, pp. 83-94.
10. Ayushev T.V., Pavlova S.V., Bulychev R.N. Geometricheskoe modelirovanie mnogoslounnykh konstruktivnykh s primeneniem krivyykh i poverkhnostey bez'e [Geometrical modeling of multi-layered constructions using curves and bezier surfaces], *Dinamika sistem, mekhanizmov i mashin* [Dynamics of systems, mechanisms and machines], 2016, No. 3, pp. 125-128.
11. Fedoruk V.A. Obrabotka eksperimental'nykh dannykh na osnove "Metodiki sdviga" pri interpolyatsii kubicheskimi splaynami [Processing of experimental data on the basis of the "Offset Methodology" for interpolation of cubic splines], *Vestnik sibirskoy gosudarstvennoy avtomobil'no-dorozhnoy akademii* [Bulletin of the Siberian State Road and Road Academy], 2016, No. 2 (48), pp. 132-136.
12. Romadanova M.M. Algoritmy postroeniya monotonnogo vesovogo kubicheskogo splayna [Algorithms of construction of monotonous weight cubic spline], *Izvestiya tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki* [Izvestia of Tula State University. Technical sciences], 2018, No. 9, pp. 180-192.
13. Yanlin Geng, Tong Lin, Zhouchen Lin, Pengwei Hao. Refined Exponential Filter with Applications to Image Restoration and Interpolation. *Computer Vision - ACCV 2009*, pp. 33-42. Available at: https://www.researchgate.net/publication/220745293_Refined_Exponential_Filter_with_Applications_to_Image_Restoration_and_Interpolation.
14. Shcherba E.V. Analiz primenimosti metodov interpolyatsii i ekstrapolyatsii dlya resheniya zadachi vosstanovleniya izobrazheniya [Analysis of the applicability of interpolation and extrapolation methods to solve the problem of image recovery], *Komp'yuternaya optika* [Computer optics], 2009, No. 33 (3), pp. 336-339.
15. Jinyu Chu, Ju Liu, Jianping Qiao, Xiaoling Wang, Yujun Li. Gradient-based adaptive interpolation in super-resolution image restoration, *International Conference on Signal Processing Proceedings*, 2009.
16. Tudor Barbu. Novel Diffusion-Based Models for Image Restoration and Interpolation, 2019.
17. Polovko A.M., Butusov P.N. Interpolyatsiya. Metody i komp'yuternye tekhnologii ikh realizatsii [Interpolation. Methods and computer technologies of their implementation]. Moscow: BKhV-Peterburg, 2016, 320 p.
18. Rassel Dzhessi. Bikubicheskaya interpolyatsiya [Bikubic Interpolation]. Moscow: Kniga po Trebovaniyu, 2013, 814 p.
19. Astashkin S.V. Interpolyatsiya operatorov i ee prilozheniya [Interpolation of operators and its applications]. Moscow: Kniga po trebovaniyu, 2013, 188 p.
20. Kreyin S.G., Petunin Yu.I., Semenov E.M. Interpolyatsiya lineynykh operatorov [Interpolation of line operators]. Moscow: Glavnaya redaktsiya fiziko-matematicheskoy literatury izdatel'stva "Nauka", 2014, 400 p.
21. Mikheev S.E. Mnogomernaya approksimatsiya i interpolyatsiya [Multidimensional approximation and interpolation]. Saint Petersburg: Izd-vo S.-Peterb. un-ta, 2012, 59 p.
22. Uolsh Dzh.L. Interpolyatsiya i approksimatsiya ratsional'nymi funktsiyami v kompleksnoy oblasti [Interpolation and approximation of rational functions in the complex area]. Moscow: Izd-vo inostrannoy literatury, 1961, 508 p.
23. Martysenko S.N. Metody vosstanovleniya propuskov v dannykh, predstavlenykh v razlichnykh izmeritel'nykh shkalakh [Methods for recovering gaps in data presented in various measuring scales], *Territoriya novykh vozmozhnostey. Vestnik Vladivostokskogo gosudarstvennogo universiteta ekonomiki i servisa* [Territory of new opportunities. Bulletin of the Vladivostok State University of Economics and Service], 2013, No. 4 (22), pp. 242-255.

24. Zloba E., Yatskiv I. Statisticheskie metody vosstanovleniya propushchennykh dannykh [Statistical methods for recovering missing data], *Computer Modelling & New Technologies*, 2002, Vol. 6, No. 1, pp. 51-61.
25. Sorokin A.A., Kovalenko R.A., Yakovleva E.A. Prognozirovaniye stoimosti arendy skladskikh pomeshcheniy na osnove statisticheskikh dannykh [Forecasting the cost of rental warehouse based on statistical data], *Evraziyskiy soyuz uchenykh* [Eurasian Union of Scientists], 2018, No. 12 (57), pp. 59-62.
26. Ryzhenkova K.V. Metody vosstanovleniya propuska dannykh pri provedenii statisticheskikh issledovaniy [Methods for recovering data gaps in statistical studies], *Intellekt. Innovatsii. Investitsii* [Intelligence. Innovation. Investments], 2012, No. 3, pp. 127-133.

Статью рекомендовал к опубликованию д.т.н. С.В. Колесниченко.

Сорокин Алексей Андреевич – Ивангородский гуманитарно-технический институт (филиал) Санкт-петербургского государственного университета аэрокосмического приборостроения; e-mail: A.A.Sorokin@ifguar.ru; 188491, Ивангород, ул. Котовского, 1; кафедра прикладной математики, информатики и информационных таможенных технологий; старший преподаватель.

Дагаев Александр Владимирович – e-mail: adagaev@list.ru; кафедра прикладной математики, информатики и информационных таможенных технологий; к.т.н.; доцент.

Бородинский Илья Михайлович – Южный федеральный университет; e-mail: iborodyanskiy@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: +79185027169; кафедра информационных измерительных технологий и систем; доцент.

Sorokin Aleksei Andreevich – St. Petersburg University of Aerospace Instrumentation; e-mail: A.A.Sorokin@ifguar.ru; 1, Kotovskogo street, Ivangorod, 188491, Russia; the department of applied mathematics, computer science and customs information technologies; senior lecturer.

Dagaev Aleksandr Vladimirovich – e-mail: adagaev@list.ru; the department of applied mathematics, computer science and customs information technologies; cand. of eng. sc.; associate professor.

Borodyansky Ilya Mikhailovich – Southern Federal University; e-mail: iborodyanskiy@sfedu.ru; 44, Nekrasovskiy lane, Taganrog, 347928, Russia; phone: +79185027169; the department of information measuring technologies and systems; associate professor.

УДК 519.27

DOI 10.18522/2311-3103-2020-4-107-117

Г.Ф. Филаретов, З. Бучаала

НЕПАРАМЕТРИЧЕСКИЙ МЕТОД ОБНАРУЖЕНИЯ РАЗЛАДКИ ВРЕМЕННОГО РЯДА С ИСПОЛЬЗОВАНИЕМ МЕХАНИЗМА СЛУЧАЙНЫХ БЛУЖДЕНИЙ

Рассмотрена задача оперативного обнаружения внезапного изменения вероятностных свойств временного ряда, обычно трактуемая как задача обнаружения разладки наблюдаемого стохастического процесса. Отмечается актуальность развития исследований по данной тематике, что обусловлено появлением всё новых прикладных задач, где методы и алгоритмы обнаружения разладки могут успешно использоваться – в частности, при создании мониторинговых систем в промышленности, экологии, медицине и др. Обсуждаются две основные разновидности методов обнаружения разладки: параметрические и непараметрические. Отмечено, что, хотя непараметрические методы при прочих равных условиях уступают параметрическим по эффективности (быстроте обнаружения разладки), но зато обладают рядом преимуществ, не требуя, в частности, контролируемого процесса. Это принципиально важно при построении мониторинговых систем, когда детальная информация об этих свойствах может либо полностью отсутствовать и тогда необходимо проводить достаточно трудоемкое его предварительное исследование, либо быть малодостоверной. Предложен оригинальный последовательный непараметрический алгоритм обнаружения разладки на основе

реализации механизма случайных блужданий или, более конкретно, с использованием теории серий «успехов». Объяснен принцип работы контролирующего алгоритма и дано его описание. Приведены результаты исследования основных статистических характеристик алгоритма, включая определение его эффективности, и результаты сопоставления с известными параметрическими методами. Выделена область возможного практического использования предложенного алгоритма, где его эффективность остается достаточно высокой. Отмечена перспективность применения предложенного алгоритма в составе программно-алгоритмического обеспечения систем мониторинга различного назначения.

Разладка временного ряда; обнаружение разладки; непараметрический алгоритм обнаружения; механизм случайных блужданий; теории серий «успехов»; вероятностные характеристики алгоритма; эффективность алгоритма; системы мониторинга.

G.F. Filaretov, Z. Bouchaala

NON-PARAMETRIC METHOD FOR DETECTING BREAKDOWN OF TIME SERIES USING THE RANDOM WALKS THEORY MECHANISM

The task of the on-line detection of a sudden change in the probability properties of a time series is considered, which is usually interpreted as the detecting task of change point the characteristics (breakdown) in the observed stochastic process. The actuality of the development of research on this topic is noted, which is due to the emergence of ever new applied problems where methods and algorithms for breakdown detecting can be successfully used - in particular, when creating monitoring systems in industry, ecology, medicine, etc. Two main varieties of methods for breakdown detecting are discussed: parametric and nonparametric. It is noted that, although nonparametric methods, ceteris paribus, are inferior to parametric methods in terms of efficiency (the speed of breakdown detecting), they also have a number of advantages, without requiring, in particular, for their application detailed information about the probabilistic properties of the controlled process. This is fundamentally important for building monitoring systems, when detailed information about these properties may either be completely absent and then it is necessary to conduct a rather laborious preliminary study of it, or to be unreliable. An original sequential nonparametric algorithm for detecting discord is proposed based on the implementation of the random walk mechanism or, more specifically, using the theory of success runs. The operating principle of the control algorithm is explained and its description is given. The results of the study of the basic statistical characteristics of the algorithm, including the determination of its effectiveness, and results of comparison with known parametric methods, are given. The area of possible practical use of the proposed algorithm is highlighted, where its effectiveness remains quite high. The prospects of using the proposed algorithm as part of the software and algorithmic support of monitoring systems for various purposes are noted.

Time series breakdown; breakdown detection; nonparametric detection algorithm; random walk mechanism; theory of success runs; probabilistic characteristics of the algorithm; algorithm efficiency; monitoring systems.

Введение. Рассматривается задача оперативного обнаружения внезапного изменения вероятностных свойств временного ряда. Данная задача обычно трактуется как задача обнаружения, так называемой, разладки наблюдаемого стохастического процесса [1] или как задача анализа его статистической стабильности с помощью контрольных карт [2]. Методы и алгоритмы ее решения относятся к категории последовательных, пригодных для работы в реальном времени и в любом варианте по сути остаются одними и теми же, разве что функционирование контрольных карт, обычно, сопровождается графическим дополнением.

Хотя первые работы по указанной тематике появились достаточно давно [3, 4], интерес к ней с годами не только не уменьшается, а скорее, наоборот, увеличивается, о чем говорят данные библиометрического анализа [5], зафиксировавшего экспоненциальный рост числа публикаций за последние годы. Это, по всей видимости, связано с появлением всё новых прикладных задач, где такого рода алгоритмы могут успешно использоваться – в частности, при создании мониторинговых систем в промышленности, экологии, медицине и др.

К настоящему времени разработано весьма большое число различных методов обнаружения разладки. Все их можно подразделить на две большие группы: методы параметрические и непараметрические.

Параметрические методы базируются на предположении, что наблюдаемый (контролируемый) процесс имеет априори известную функцию распределения вероятностей, входящую в некоторое параметрическое семейство распределений, причем разладка связана с изменением одного или нескольких параметров этого распределения. Чаще всего предполагается, что процесс гауссовский с некоррелированными отсчетами, а разладка связана со скачкообразным изменением математического ожидания или дисперсии в одномерном варианте или вектора математических ожиданий или ковариационной матрицы в многомерном.

Среди параметрических методов обнаружения разладки наиболее хорошо изученными и получившими наибольшее распространение являются различные варианты методов и алгоритмов, базирующихся на видоизменённом последовательном анализе. Чаще всего их называют алгоритмами кумулятивных сумм (АКС) или CUSUM-алгоритмами [1]. Доказано [6], что эти алгоритмы обладают определенными оптимальными свойствами в смысле максимизации показателя эффективности $E = \bar{T}_{лт} / \bar{\tau}_{зан}$, где $\bar{T}_{лт}$ – среднее время между ложными тревогами, когда контролирующий алгоритм подает сигнал о наличии разладки, когда на самом деле она отсутствует, $\bar{\tau}_{зан}$ – среднее время запаздывания в обнаружении номинальной разладки d , т.е. разладки, которую необходимо как можно быстрее обнаружить.

Непараметрические методы не требуют какой-либо существенной информации о виде функций распределения контролируемого процесса. Поэтому их часто именуют процедурами «свободными от распределения».

Их главная отличительная черта - они оперируют не с самими измеренными значениями, а с их знаковыми, ранговыми или порядковыми представлениями. Такие методы в настоящее время развиваются весьма интенсивно, поскольку у непараметрического подхода, как оказалось, имеется еще ряд преимуществ по сравнению с параметрическим [7–10]. Важнейшими из них являются: простота реализации контролирующего алгоритма, одинаковые характеристики эффективности для целого семейства процессов с непрерывными, возможно различными функциями распределения вероятностей, большая устойчивость результатов, отсутствие необходимости в оценке дисперсии для определения параметра разладки d .

Перечисленные преимущества имеют принципиально важный характер при построении мониторинговых систем, когда детальная информация о свойствах контролируемого процесса может либо полностью отсутствовать и тогда необходимо проводить достаточно трудоемкое его предварительное исследование, либо быть малодостоверной.

Как следует из рассмотрения и анализа имеющихся публикаций (см., например, [11–18]), подавляющее большинство известных на данный момент непараметрических методов обнаружения разладки базируется на модификации одного из стандартных непараметрических критериев проверки статистических гипотез. Дальнейшее изложение посвящено рассмотрению другого способа построения непараметрического алгоритма обнаружения разладки, а именно с использованием элементов теории случайных блужданий [19].

1. Постановка задачи. Впервые идея использования механизма одномерных случайных блужданий как средства решения задачи обнаружения разладки была высказана в работе [20], опубликованной в 1975 г. В ней рассматривалась классическая схема одномерных случайных блужданий, когда на каждом шаге в некото-

рый дискретный момент времени n текущая координата точки z_n , отображающей процесс блуждания, с вероятностью p получает приращение $+1$ или с вероятностью $q = 1 - p$ – приращение -1 в соответствии с соотношением:

$$z_n = \begin{cases} +1, & x_n \geq Me(X) \\ -1, & x_n < Me(X) \end{cases}, \quad n = 1, 2, \dots \quad (1)$$

где $Me(X)$ – значение медианы наблюдаемого временного ряда в состоянии «норма» (без разладки). Отметим, что в этом случае справедливо равенство $p = q = 1/2$; однако при наличии разладки эта симметрия нарушается. Можно, например, считать, что разладка приводит к сдвигу медианы в положительную сторону и, следовательно, тогда будем иметь вероятность $p > 1/2$. Исходная идея, к сожалению, не была затем доведена до уровня, позволяющего выработать четкие рекомендации по синтезу соответствующей контролирующей процедуры, оценить ее эффективность, а также возможности практического применения. Далее рассматривается одно из возможных направлений развития методов обнаружения разладки на базе применения механизма случайных блужданий, а именно с использованием теории серий «успехов» [19].

В соответствии с определением, приведенным в [19], под успехом понимается появление в последовательности испытаний Бернулли события, приводящего к движению отображающей точки случайного блуждания в положительную сторону, т.е. когда приращение равно $+1$. Тогда под серией успехов подразумевается появление нескольких следующих друг за другом значений $+1$, а под длиной серии – их количество k .

Целью данной работы является построение на этой основе непараметрического метода обнаружения разладки, исследование его основных статистических характеристик и анализ эффективности, в том числе при сопоставлении с известными параметрическими методами.

2. Непараметрический метод обнаружения разладки на основе теории серий успехов. Контролирующая процедура предлагаемого метода очень простая и сводится на каждом n -ом такте работы к вводу очередного значения контролируемого процесса x_n , его преобразованию в значение приращения z_n в соответствии с соотношением (1), подсчету k_n – текущей длины серии из следующих подряд значений $z_n = +1$: $k_n = k_{n-1} + 1$, если $z_n = +1$ и $k_n = 0$, если $z_n = -1$, фиксации номера такта n^* , на котором будет выполнено условие $k_{n^*} = k_n = k$, где k – заранее заданное пользователем значение, выбранное исходя из необходимости обеспечения желательного среднего времени между ложными тревогами $\bar{T}_{ЛТ}$. Фактически n^* фиксирует момент первого появления серии из $+1$ выбранной длины k .

Появление серии успехов длины k трактуется в [19] как рекуррентное событие. Для таких событий там приведены формулы, позволяющие найти математическое ожидание и дисперсию первого от начала осуществления испытаний появления серии длины k :

$$m_k = \frac{1 - p^k}{(1 - p)p^k}; \quad \sigma_k^2 = \frac{1}{[(1 - p)p^k]^2} - \frac{2k + 1}{(1 - p)p^k} - \frac{p}{(1 - p)^2} \quad (2)$$

При отсутствии разладки вероятность $p = 1/2$, а зафиксированное значение n^* отвечает моменту появления сигнала ложной тревоги. Полагая в соотношении для m_k вероятность $p = p_0 = 1/2$, можно определить средний интервал между ложными тревогами $\bar{T}_{ЛТ}$:

$$\bar{T}_{\text{лт}} = \frac{1 - (1/2)^k}{(1/2)^k} = 2(2^k - 1). \quad (3)$$

В табл. 1 приведены значения $\bar{T}_{\text{лт}}$, соответствующие разным величинам длин серий k , а также соответствующие значения σ_k .

Таблица 1

Зависимость $\bar{T}_{\text{лт}}$ и σ_k от длины серий успехов k

k	4	5	6	7	8	9	10	11	12
$\bar{T}_{\text{лт}}$	30	62	126	254	510	1022	2046	4094	8190
σ_k	27,09	58,22	121,3	248,3	503,4	1014,4	2037,5	4084,5	8179,5

Для оценки среднего времени запаздывания в обнаружении разладки $\bar{\tau}_{\text{зан}}$ также можно воспользоваться вышеприведенной формулой для m_k , если положить в ней значение $p = p_1 > 1/2$. При этом для получения общего представления о быстродействии алгоритма обнаружения выбор конкретных значений p может быть осуществлен пользователем достаточно произвольно, исходя из его собственных представлений о величине разладки, подлежащей обнаружению. Однако здесь целесообразнее выбирать p_1 с учетом их последующего использования для сопоставления предложенного метода с известными параметрическими методами обнаружения. В качестве таковых предлагается использовать наиболее изученные варианты CUSUM-алгоритмов, предназначенных для обнаружения разладки гауссовского случайного процесса по математическому ожиданию и дисперсии. Для таких алгоритмов известны расчетные величины $\bar{\tau}_{\text{зан}}$ для фиксированных $\bar{T}_{\text{лт}}$ и типовых наборов значений, характеризующих величину разладки d .

Для случая разладки по математическому ожиданию такой типовой набор обычно включает следующие значения d_m : 0,25; 0,5; 1,0; 1,5; 2,0; 2,5, где $d_m = |m_1 - m_0|/\sigma_x$, m_0 и m_1 – значения математических ожиданий процесса x_n до и после разладки соответственно.

Для случая разладки по дисперсии типовой набор содержит следующие значения $d_\sigma = \sigma_1^2/\sigma_0^2$; σ_0^2 и σ_1^2 – значения дисперсий процесса x_n до и после разладки: 1,5; 2,0; 2,5 – разладка в сторону увеличения дисперсии (d_σ^+) и 1/1,5; 1/2,0; 1/2,5 – разладка в сторону уменьшения дисперсии (d_σ^-).

Каждому из перечисленных типовых значений разладки могут быть сопоставлены значения вероятности p_1 . Соответствующая информация приведена в табл. 2 и табл. 3.

Таблица 2

Значения показателя разладки d_m и соответствующие им вероятности p_1

d_m	0,25	0,5	1,0	1,5	2,0	2,5
p_1	0,599	0,692	0,841	0,933	0,977	0,994

Таблица 3

Значения показателей разладки d_{σ}^{+} , d_{σ}^{-} и вероятности p_1^{+} , p_1^{-}

d_{σ}^{+}	1,25	1,5	2,0	2,5	3,0
p_1^{+}	0,546	0,582	0,634	0,669	0,697
d_{σ}^{-}	1/1,25	1/1,5	1/2,0	1/2,5	1/3,0
p_1^{-}	0,550	0,592	0,659	0,718	0,758

Далее при получении и анализе статистических характеристик рассматриваемого непараметрического алгоритма будет использованы именно значения p_1 из табл. 2 и табл. 3.

3. Статистические характеристики алгоритма обнаружения. Искомыми статистическими характеристиками алгоритма обнаружения разладки являются значения среднего времени запаздывания $\bar{\tau}_{\text{зап}}$ и показателя эффективности E , найденные для значений p_1 , p_1^{+} , p_1^{-} , соответствующих всем типовым вариантам параметров d_m , d_{σ}^{+} , d_{σ}^{-} и различным k . Полученные расчетные значения приведены в табл. 4 - 6 и отображены на рис. 1.

Таблица 4

Значения среднего времени запаздывания $\bar{\tau}_{\text{зап}}$ для различных p_1

d_m (p_1)	0,25 (0,599)	0,5 (0,692)	1,0 (0,841)	1,5 (0,933)	2,0 (0,977)	2,5 (0,994)
k						
4	16,88	10,94	6,28	4,77	4,24	4,06
5	29,84	17,27	8,66	6,19	5,36	5,09
6	51,49	26,42	11,49	7,70	6,51	6,13
7	87,64	39,66	14,85	9,33	7,69	7,17
8	147,97	58,81	18,84	11,07	8,90	8,22
9	248,70	86,51	23,59	12,94	10,13	9,28
10	416,87	126,57	29,24	14,94	11,39	10,34
11	697,61	184,51	35,96	17,08	12,68	11,42
12	1166,29	268,31	43,95	19,38	14,00	12,48

Таблица 5

Значения среднего времени запаздывания $\bar{\tau}_{\text{зап}}$ для различных p_1^{+} , p_1^{-}

k	d_{σ}^{+}			d_{σ}^{-}		
	1,5 ($p_1^{+} =$ 0,582)	2,0 ($p_1^{+} =$ 0,634)	2,5 ($p_1^{+} =$ 0,669)	1/1,5 ($p_1^{-} =$ 0,592)	1/2,0 ($p_1^{-} =$ 0,659)	1/2,5 ($p_1^{-} =$ 0,718)
8	176	101	72	277	80	47
9	310	162	107	268	122	66
10	532	254	163	466	185	94
11	907	406	244	781	284	134

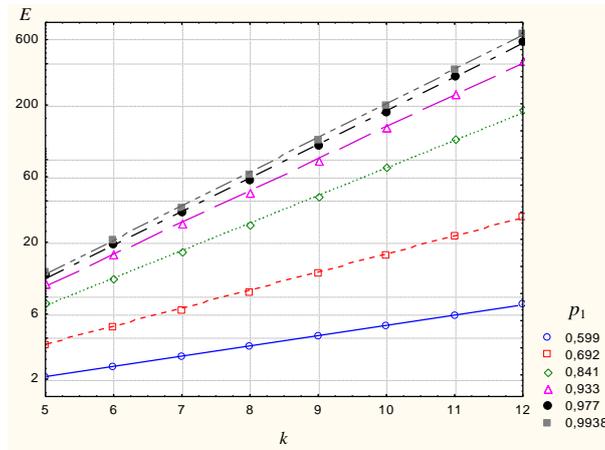


Рис. 1. Зависимости показателя эффективности E непараметрической контролирующей процедуры от k для различных вероятностей p_1

Таблица 6

Значения показателя эффективности контролирующей процедуры E_k для различных p_1^+ , p_1^-

k	d_σ^+			d_σ^-		
	1,5 ($p_1^+ = 0,582$)	2,0 ($p_1^+ = 0,634$)	2,5 ($p_1^+ = 0,669$)	1/1,5 ($p_1^- = 0,592$)	1/2,0 ($p_1^- = 0,659$)	1/2,5 ($p_1^- = 0,718$)
8	2,9	5,0	7,1	1,8	6,4	11,0
9	3,3	6,3	9,5	3,8	8,4	15,5
10	3,8	8,0	12,6	4,4	11,1	21,8
11	4,5	10,1	16,8	5,2	14,4	30,5

Ясно, что эффективность всех рассматриваемых процедур увеличивается с ростом вероятности p (т.е. с ростом величины разладки) и возрастанием значения k (т.е. с ростом значения $\bar{T}_{ЛТ}$).

4. Сопоставление с параметрическим cusum-алгоритмом. С точки зрения практического использования большой интерес представляет определение относительной эффективности данного непараметрического алгоритма по сравнению с известными параметрическими алгоритмами частности – с упомянутыми ранее вариантами CUSUM-алгоритма. Для проведения такого сопоставления используем коэффициент относительной эффективности ϵ , определяемый как отношение эффективности E непараметрического алгоритма к эффективности CUSUM-алгоритма E_{cusum} для фиксированных значений $\bar{T}_{ЛТ}$ и d :

$$\epsilon = \frac{E}{E_{\text{cusum}}} = \frac{\bar{\tau}_{\text{cusum}}}{\bar{\tau}_{\text{зап}}}, \quad (4)$$

где $\bar{\tau}_{\text{cusum}}$ – значение запаздывания в обнаружения разладки гауссовского процесса при использовании CUSUM-алгоритма.

Результаты сравнения эффективности непараметрического алгоритма с классическим CUSUM-алгоритмом при обнаружении разладки по математическому ожиданию представлены на диаграмме рис. 2.

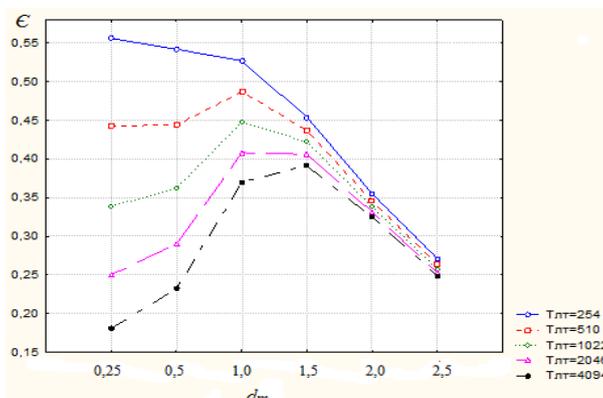


Рис. 2. Сравнение эффективности непараметрического алгоритма с CUSUM-алгоритмом (разладка по математическому ожиданию)

В итоге можно выделить некоторую область значений d_m и $\bar{T}_{лг}$, в которой рассматриваемый непараметрический алгоритм имеет достаточно высокую эффективность – не менее 35% по сравнению с классическим CUSUM-алгоритмом.

Аналогичный анализ был произведен для варианта разладки, связанной с изменением дисперсии. Его результаты приведены в табл. 7.

Таблица 7

Значения коэффициента относительной эффективности ϵ (разладка по дисперсии)

$\bar{T}_{лг}$	d_{σ}^{+}			d_{σ}^{-}		
	1,5 ($p_1^{+} = 0,582$)	2,0 ($p_1^{+} = 0,634$)	2,5 ($p_1^{+} = 0,669$)	1/1,5 ($p_1^{-} = 0,592$)	1/2,0 ($p_1^{-} = 0,659$)	1/2,5 ($p_1^{-} = 0,718$)
500	0,28	0,22	0,20	0,24	0,44	0,54
1000	0,20	0,16	0,16	0,31	0,35	0,45
2000	0,14	0,12	0,12	0,22	0,28	0,37
4000	0,10	0,09	0,09	0,15	0,21	0,29

Очевидно, то относительная эффективность данного контролирующего алгоритма:

- ♦ различна для разладок в сторону увеличения или уменьшения дисперсии, причем большую эффективность имеет процедура обнаружения разладки в сторону уменьшения дисперсии, что прямо противоположно свойствам классического CUSUM-алгоритма [1];

- ♦ выше для малых значений $\bar{T}_{лг}$ и увеличивается с ростом величины разладки;

- ♦ заметно меньше по сравнению с аналогичной процедурой для случая разладки по математическому ожиданию; действительно, если выделить 35 %-ый уровень относительной эффективности, то ему соответствует всего несколько значений из табл. 6 (все они выделены шрифтом).

В целом, как и ожидалось, предложенный непараметрический метод обнаружения разладки уступает по эффективности параметрическим методам. Тем не менее, в силу своей простоты он вполне может найти свое применение для целей оперативного контроля за текущим состоянием различного рода объектов, выявления нарушений установленных режимов их функционирования, а также в мониторинговых системах.

Заключение. В статье рассмотрена задача оперативного обнаружения спонтанного изменения (разладки) вероятностных характеристики временного ряда. Для ее решения в условиях дефицита априорной информации о свойствах контролируемого процесса предложен последовательный непараметрический алгоритм обнаружения разладки на основе реализации механизма случайных блужданий или, более конкретно, с использованием теории серий «успехов». Приведено описание контролирующего алгоритма, результаты исследования его основных статистических характеристик. Проведен анализ его эффективности, в том числе при сопоставлении с известными параметрическими методами. Выделена область возможного практического использования предложенного алгоритма, где его эффективность остается достаточно высокой. Отмечена перспективность применения предложенного алгоритма в составе программно-алгоритмического обеспечения систем мониторинга различного назначения.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Никифоров И.В.* Последовательное обнаружение изменения свойств временных рядов. – М.: Наука, 1983. – 200 с.
2. ГОСТ Р. ИСО 7870-2-2015 Национальный стандарт Российской Федерации. Статистические методы. Контрольные карты. – Стандартиформ, 2016.
3. *Page E.S.* Continuous inspection schemes // *Biometrika*. – 1954. – Vol. 41, No. 1. – P. 100-115.
4. *Lorden G.* Procedures for Reacting to a Change in Distribution // *Annals of Mathematical Statistics*. – 1971. – Vol. 42,6. – P. 1897-1908.
5. *Shafiq Ahmad.* Bibliometric Analysis of EWMA and CUSUM Control Chart Schemes // *IJTEE Journal*. – April 2018. – Vol. 7, Issue 2. – P. 1-11.
6. *Шуряев А.Н.* Задача скорейшего обнаружения нарушения стационарного режима // Доклады АН СССР. – 1961. – Т. 138, № 5. – С. 1039-1042.
7. *Chakraborti S., Laan van der P., Bakir S.T.* Nonparametric Statistical Process Control: An Overview and Some Results. (Memorandum COSOR; Vol.9908). Eindhoven: Technische Universiteit Eindhoven. – 1999.
8. *Chakraborti S., Van der Laan P., & Bakir S.T.* Nonparametric statistical process control: an overview and some results // *Journal of quality technology*. – 2001. – Vol. 33 (3). – P. 304-315.
9. *Saad Bakir.* Classification distribution-free quality control charts // *Proceedings of the Annual Meeting of the American Statistical Association*. – August 5-9, 2001. – 15 p.
10. *Chakraborti S., Graham M.A.* Nonparametric Statistical Process Control. – John Wiley & Sons Ltd, 2019. – 429 p.
11. *Bakir S.T. and Reynolds Jr. M.R.* A nonparametric procedure for process control based on within-group ranking // *Technometrics*. – 1979. – No. 21 (2). – P. 175-183.
12. *Amin R.W., Reynolds Jr. M.R., and Bakir S.T.* Nonparametric quality control charts based on the sign statistic // *Communications in Statistics: Theory and Methods*. – 1995. – No. 24 (6). – P. 1597-1623.
13. *Janacek G.J. and Meikle S.E.* Control charts based on medians // *The Statistician*. – 1997. – Vol. 46 (1). – P. 19-31.
14. *McDonald D.* A CUSUM procedure based on sequential ranks // *Naval Research logistics*. – 1999. – Vol. 37. – P. 627-646.
15. *Bakir S.T.* A distribution-free Shewhart quality control chart based on signed-ranks // *Quality Engineering*. – 2004. – No. 16 (4). – P. 613-623.
16. *Chakraborti S. and Van de Wiel M.A.* A nonparametric control chart based on the Mann-Whitney statistic. *IMS Collections // Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen*. – 2008. – No. 1. – P. 156-172.

17. Human S.W., Chakraborti S., and Smit C.F. Nonparametric Shewhart-type sign control charts based on runs // *Communications in Statistics: Theory and Methods*. – 2010b. – Vol. 39 (11). – P. 2046-2062.
18. Zhou Chunguang, Zou Changliang, Zhang Yujuan, Wang Zhaojun. Nonparametric control charts based on change-point model. *Statistical Papers*, 2009. – Vol. 50. – P. 13-28.
19. Феллер У. Введение в теорию вероятностей и ее приложения. Т. 1. – М.: Изд-во «Мир», 1970. – 499 с.
20. Mc. Gilchrist C.A., Woodyer K.D. Note on a distribution-free CUSUM technique // *Technometrics*. – 1975. – Vol. 17, No. 3. – P. 321-325.

REFERENCES

1. Nikiforov I.V. Posledovatel'noe obnaruzhenie izmeneniya svoystv vremennykh ryadov [Sequential detection of changes in the properties of time series]. Moscow: Nauka, 1983, 200 p.
2. GOST R. ISO 7870-2-2015 Natsional'nyy standart Rossiyskoy Federatsii. Statisticheskie metody. Kontrol'nye karty [GOST R ISO 7870-2-2015 national standard of the Russian Federation. Statistical technology. Control card]. Standartinform, 2016.
3. Page E.S. Continuous inspection schemes, *Biometrika*, 1954, Vol. 41, No. 1, pp. 100-115.
4. Lorden G. Procedures for Reacting to a Change in Distribution, *Annals of Mathematical Statistics*, 1971, Vol. 42,6, pp. 1897-1908.
5. Shafid Ahmad. Bibliometric Analysis of EWMA and CUSUM Control Chart Schemes, *ITEE Journal*, April 2018, Vol. 7, Issue 2, pp. 1-11.
6. Shiryaev A.N. Zadacha skoreyshego obnaruzheniya narusheniya statsionarnogo rezhima [The problem of early detection of violations of the stationary regime], *Doklady AN SSSR [Doklady Akademii Nauk USSR]*, 1961, Vol. 138, No. 5, pp. 1039-1042.
7. Chakraborti S., Laan van der P., Bakir S.T. Nonparametric Statistical Process Control: An Overview and Some Results. (Memorandum COSOR; Vol.9908). Eindhoven: Technische Universiteit Eindhoven, 1999.
8. Chakraborti S., Van der Laan P, & Bakir S.T. Nonparametric statistical process control: an overview and some results, *Journal of quality technology*, 2001, Vol. 33 (3), pp. 304-315.
9. Saad Bakir. Classification distribution-free quality control charts, *Proceedings of the Annual Meeting of the American Statistical Association*, August 5-9, 2001, 15 p.
10. Chakraborti S., Graham M.A. Nonparametric Statistical Process Control. John Wiley & Sons Ltd, 2019, 429 p.
11. Bakir S.T. and Reynolds Jr. M.R. A nonparametric procedure for process control based on within-group ranking, *Technometrics*, 1979, No. 21 (2), pp. 175-183.
12. Amin R.W., Reynolds Jr. M.R., and Bakir S.T. Nonparametric quality control charts based on the sign statistic, *Communications in Statistics: Theory and Methods*, 1995, No. 24 (6), pp. 1597-1623.
13. Janacek G.J. and Meikle S.E. Control charts based on medians, *The Statistician*, 1997, Vol. 46 (1), pp. 19-31.
14. McDonald D. A CUSUM procedure based on sequential ranks, *Naval Research logistics*, 1999, Vol. 37, pp. 627-646.
15. Bakir S.T. A distribution-free Shewhart quality control chart based on signed-ranks, *Quality Engineering*, 2004, No. 16 (4), pp. 613-623.
16. Chakraborti S. and Van de Wiel M.A. A nonparametric control chart based on the Mann-Whitney statistic. IMS Collections, *Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen*, 2008, No. 1, pp. 156-172.
17. Human S.W., Chakraborti S., and Smit C.F. Nonparametric Shewhart-type sign control charts based on runs, *Communications in Statistics: Theory and Methods*, 2010b, Vol. 39 (11), pp. 2046-2062.
18. Zhou Chunguang, Zou Changliang, Zhang Yujuan, Wang Zhaojun. Nonparametric control charts based on change-point model. *Statistical Papers*, 2009, Vol. 50, pp. 13-28.
19. Feller U. Vvedenie v teoriyu veroyatnostey i ee prilozheniya [Introduction to probability theory and its applications]. Vol. 1. Moscow: Izd-vo «Mir», 1970, 499 p.
20. Mc. Gilchrist C.A., Woodyer K.D. Note on a distribution-free CUSUM technique, *Technometrics*, 1975, Vol. 17, No. 3, pp. 321-325.

Статью рекомендовал к опубликованию д.т.н., профессор Э.К. Лецкий.

Филаретов Геннадий Федорович – Национальный исследовательский университет «МЭИ»; e-mail: gefefi@yandex.ru; 111141, Москва, Красноказарменная, 14; тел.: +79255176319; кафедра управления и интеллектуальных технологий; профессор.

Бучаала Зинеддин – e-mail: bouchaala.zinouzin@gmail.com; тел.: +79663039794; кафедра Управления и интеллектуальных технологий; аспирант.

Filaretov Gennady Fedorovich – National Research University “MPEI”; e-mail: gefefi@yandex.ru; 14, Krasnokazarmennaja, Moscow, 111141, Russia; phone: +79255176319; the department of control systems and intellectual technology, professor.

Bouchaala Zineddine – e-mail: bouchaala.zinouzin@gmail.com; phone: +79663039794; the department of control systems and intellectual technology; postgraduate student.

Раздел II. Автоматизация проектирования

УДК 621.3.049.771.14

DOI 10.18522/2311-3103-2020-4-118-125

Н.О. Васильев, М.А. Заплетина, Г.А. Иванова, А.Н. Щелоков

ЛОГИЧЕСКИЙ РЕСИНТЕЗ КОМБИНАЦИОННЫХ СХЕМ ДЛЯ ПОВЫШЕНИЯ СБОЕУСТОЙЧИВОСТИ*

При функционировании микроэлектронных устройств в условиях космоса необходимо учитывать внешние воздействия. Работа устройства в подобных условиях затрудняется негативным влиянием радиационного излучения на электронные компоненты схемы. Воздействие тяжелых заряженных частиц приводит к одиночным сбоям логических элементов, из-за чего логика работы устройства может быть нарушена. В связи с этим при проектировании электронных схем, которые будут использоваться в космических аппаратах, необходимо выполнение повышенных требований к устойчивости интегральных схем (ИС) к одиночным сбоям. По мере уменьшения технологических норм проектирования ИС проблема сбоеустойчивости становится актуальной и для изделий микроэлектроники гражданского применения. Решение данной задачи обычно осуществляется методами аппаратной защиты, к которым относятся методы помехоустойчивого кодирования, методы резервирования, а также методы логической защиты. В данной статье рассматриваются методы оценки устойчивости ИС к одиночным сбоям в логических элементах, а также основные методы защиты схем. В работе предлагается техника ресинтеза логических комбинационных схем, использующая логические ограничения, выводимые с помощью метода резолюций, для оценки устойчивости к одиночным сбоям. В ходе ресинтеза предлагается использовать методы логической защиты уязвимых участков схемы, что не влечет ощутимого роста занимаемой устройством площади, свойственного методам резервирования и помехоустойчивого кодирования.

Ресинтез; сбоеустойчивость; надежность; комбинационные схемы; логические корреляции; метод резолюций.

N.O. Vasilyev, M.A. Zapletina, G.A. Ivanona, A.N. Schelokov

LOGICAL RESYNTHESIS OF COMBINATIONAL CIRCUITS FOR RELIABILITY INCREASE

The external influences are necessary to take into account for microelectronic devices operating in space. In these conditions, the operation of the device is hampered by the negative effect of radiation on the electronic components of the circuit. Exposure of heavy charged particles leads to single faults of logic elements due to which the operation of a whole device can be violated. In this regard, the designed spacecraft electronic circuits must meet increased requirements for the fault tolerance of integrated circuits (ICs). The decrease of technological design standards for ICs makes the problem of fault tolerance to be relevant for civilian microelectronic products, also. The solution to this problem is usually carried out by methods of hardware protection, which include methods of error-correcting coding, methods of redundancy, as well as methods of logical protection. The paper considers the methods for assessing the IC tolerance to single faults in logic elements, as well as the main methods of circuits failure protection. The paper proposes a resynthesis technique for logical combinational circuits, using logical constraints derived from the resolution method to assess the IC resistance to single faults.

* Проект выполнен при поддержке РФФИ №17-19-01645.

During resynthesis, it is proposed to use the methods of logical protection of vulnerable parts of the circuit. This does not cause a perceptible increase in the area occupied by the device unlike in methods of redundancy and error-correcting coding.

Resynthesis; fault tolerance; reliability; combinational circuits; logic correlations; resolutions method.

Введение. При уменьшении технологических норм проектирования микроэлектронных устройств, помимо положительных факторов в виде увеличения плотности транзисторов на кристалле, улучшения энергоэффективности и повышения быстродействия, проявляются и отрицательные. Среди них выделяется возрастание роли динамического потребления тока, высокая сложность проверки правил топологического проектирования, учета требований к помехоустойчивости и радиационной стойкости устройства [1], а также рост количества дефектов при производстве [2].

В настоящее время логический синтез комбинационных схем с учетом повышенных требований к сбоеустойчивости является актуальной задачей маршрута проектирования микроэлектронных устройств. Методы оценки сбоеустойчивости традиционно делятся на две группы. Методы первой группы подразумевают использование наборов входных воздействий, содержащих ошибку [3]. Для каждого такого набора выполняется полная симуляция работы схемы, показывающая, приведет ли заданная ошибка к неправильной логике работы устройства. Подобные симуляции обычно требуют больших затрат вычислительных ресурсов, так как их вычислительная сложность имеет экспоненциальную зависимость от количества входов схемы [4]. Отсюда следует, что даже для схем с несколькими десятками первичных входов задача становится трудновыполнимой. Методы второй группы для решения задачи оценки помехоустойчивости используют методы статического анализа [5], при которых выполняется поиск нижней оценки сбоеустойчивости. Недостатком такого варианта является излишне пессимистическая оценка.

В данной работе предлагается метод ресинтеза логических комбинационных схем с целью повышения надежности критических участков с применением анализа логических корреляций. Использование информации о логических корреляциях в анализируемой схеме позволяет добиться большей точности в оценке сбоеустойчивости, чем при использовании статического анализа. В то же время подобный подход требует меньших вычислительных ресурсов, чем методы, подразумевающие полную симуляцию работы схемы [6].

1. Метод резолюций. Логические элементы имеет строго детерминированный характер работы. Каждая комбинация состояний входов логического элемента накладывает ограничение на возможное состояние выходов этого элемента. Например, подача логического нуля на один из входов элемента 2И ограничивает выходное состояние нулем. Нарушение логического ограничения возможно только в случае возникновения сбоя в элементе. Одним из способов представления логических ограничений является метод логических импликаций [7].

Логическая импликация представляет собой отношение между двумя узлами цепи следующего вида: $(a = n_1) \rightarrow (b = n_2)$, где n_1, n_2 – булевы состояния узлов. Также может быть использована альтернативная форма записи логической импликации с помощью литералов (литерал соответствует либо логической переменной, либо отрицанию этой переменной) [8]: $\bar{a} \rightarrow b$.

Для любого логического элемента существует множество отношений между входными и выходными узлами, которые могут быть представлены в виде логических импликаций. Например, для логического элемента 2И-НЕ они записываются следующим образом:

$$\begin{aligned} \bar{a} &\rightarrow y, \\ \bar{b} &\rightarrow y, \\ \bar{y} &\rightarrow a, \\ \bar{y} &\rightarrow b, \end{aligned}$$

где a и b – входы элемента, а y – выход. Таким образом, для элемента 2И-НЕ верно то, что логический ноль на любом из входных узлов определяет то, что выходной узел будет находиться в состоянии логической единицы, а состояние логического нуля на выходе схемы определяет то, что оба входных узла находятся в состоянии логической единицы (рис. 1).

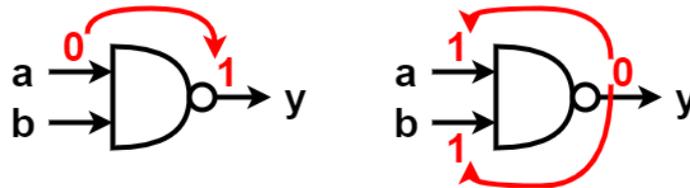


Рис. 1. Логические отношения между узлами элемента 2И-НЕ

Для вывода новых логических ограничений из уже существующих может быть использован модифицированный метод резолюций [9]. Новые логические ограничения представляются в форме суммы ограничений отношений между элементами. Например, для элемента 2И-НЕ ограничение выглядит следующим образом:

$$aby + \bar{a}\bar{y} + \bar{b}\bar{y} = 0.$$

Каждое слагаемое данной формулы наследуется от конкретной логической импликации и накладывает запрет на одну из комбинаций логических значений булевых переменных (например, $\bar{a}\bar{y}$ – на комбинацию $a = 0, y = 0$). При этом для вывода новых ограничений рекурсивно используется правило резолюций:

$$aB = 0, \bar{a}C = 0 \rightarrow BC = 0,$$

где B и C – произведения литералов.

1. Методы повышения сбоеустойчивости. При выполнении логического синтеза важно учитывать, что возникновение одиночного сбоя на одном из элементов комбинационной схемы не всегда приводит к возникновению ошибки на выходах схемы. Возникшая ошибка может быть маскирована и в итоге не окажет влияния на логику работы при некоторых входных воздействиях. Так, например, в логической схеме, изображенной на рис. 2, ошибка, возникшая на элементе $e1$, будет выявлена только в том случае, если на узел c будет подано входное воздействие, равное логическому нулю. При подаче логической единицы на узел c схема отработает корректно. Вероятность того, что ошибка не будет маскирована и приведет к сбою хотя бы на одном из выходов схемы называется наблюдаемостью.

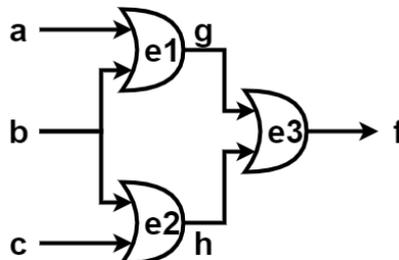


Рис. 2. Пример логической схемы

Классическим методом повышения надежности схем является использование методов резервирования [10–12]. Суть данных методов заключается в дублировании всей схемы либо её наиболее уязвимых участков с добавлением схемы-арбитра на выходе. Достоинство подобных методов заключается в том, что вероятность ошибочной работы защищенных участков схемы крайне мала, так как необходимо, чтобы сбой возник на каждой из копий [13, 14]. К недостаткам можно отнести то, что схема-арбитр также подвержена сбоям, следовательно, в случае возникновения сбоя в арбитра, логика работы схемы будет нарушена. Кроме того, резервирование существенно увеличивает необходимую для реализации схемы площадь.

Существуют также методы улучшения сбоеустойчивости с помощью использования самокорректирующихся схем. Подобные методы используют избыточное кодирование и помехоустойчивые коды [15, 16]. Надежность (сбоеустойчивость) схем, синтезированных с использованием данных методов, может превышать надежность схем, синтезированных с использованием резервирования критических участков схем [17]. Однако, такие методы в ещё большей степени увеличивают количество необходимых для реализации схемы логических элементов.

Компромиссом между необходимой для реализации микроэлектронного устройства площадью и надежностью его работы может служить ресинтез схемы с использованием целевой функции, зависящей от сбоеустойчивости. Для этого, прежде всего, необходимо определить методы и метрики, которые позволят дополнить оценку надежности схемы.

Одним из способов задания функции оценки сбоеустойчивости логических схем без привязки к технологии полупроводникового производства является коэффициент логической чувствительности, предложенный в работе [18]:

$$\alpha = \frac{1}{2^N} \sum_{\mathbf{X}, \mathbf{e}, |\mathbf{e}|=1} E(\mathbf{X}, \mathbf{e}),$$

где $E(\mathbf{X}, \mathbf{e})$ – характеристическая функция, зависящая от вектора входных сигналов \mathbf{X} и вектора ошибок \mathbf{e} . Данный коэффициент является суммой вероятностей возникновения нарушения в логике работы устройства при появлении одиночной ошибки на каждом из элементов при условии корректной работы остальных элементов [19].

Другим подходом к формированию функции надежности схемы, является использование рекурсивной функции потока ошибок (critical bit flow, CBF) [20]. Данная функция определяется следующим образом:

$$CBF(a) = CriticalBits(a) + \sum_{b \in fanins(a)} \frac{CBF(b)}{NumOfFanout(b)},$$

где $CriticalBits(a)$ – количество логических состояний элемента a , которые приведут к возникновению ошибки на выходах схемы в том случае, если в элементе схемы возник сбой; $fanins(a)$ – множество логических элементов, выходные узлы которых являются входными для элемента a ; $NumOfFanout(b)$ – количество входов в элементе b .

Функцию потока ошибок удобно использовать непосредственно во время выполнения ресинтеза, поскольку она не требует выполнения дополнительных симуляций при синтезе от первичных выходов к первичным входам. В то же время, коэффициент логической чувствительности дает более точную оценку и может быть использован для финальной оценки схемы на предмет надежности.

2. Техника логического ресинтеза комбинационных схем. Использование логических отношений между узлами схемы позволяет оценить влияние сбоя в одном из элементов на иные узлы и сделать вывод о критичности защиты опреде-

ленных участков схемы. Так, например, в схеме, изображенной на рис. 3, нахождение узла $n7$ в состоянии логического нуля определяет состояние пяти других узлов этой схемы, три из которых расположены в выходном конусе логики. Причем один из этих узлов является первичным выходом схемы. Следовательно, некорректное состояние данного узла способно вызвать ошибку функционирования всей схемы. При этом, если узел $n7$ будет находиться в состоянии логической единицы, то одной этой информации будет недостаточно для того, чтобы определить состояние других узлов схемы.

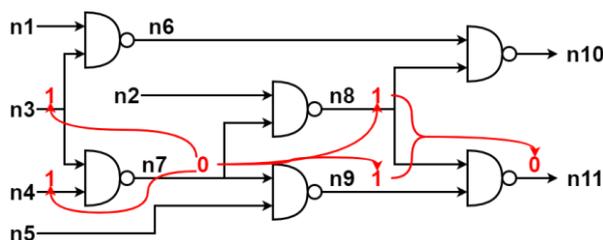


Рис. 3. Корреляции между узлами электрической схемы

Таким образом, методом резолюций построив логические импликации между элементом, в который вносится сбой, и выходными узлами схемы, можно оценить, в каких ситуациях сбой будет маскирован, а в каких окажет влияние на работу итогового устройства. При возникновении сбоя в элементе предполагается, что логика его работы инвертируется.

Применение логических ограничений потенциально может позволить добиться ускорения вычисления функции оценки сбоеустойчивости схемы. В связи с этим, для выполнения синтеза предлагается следующая последовательность действий:

- 1) вывести логические ограничения схемы с помощью метода резолюций;
- 2) с помощью векторов входных воздействий, вектора ошибок и полученных ограничений найти наиболее уязвимые участки схемы, нарушения работы которых наиболее часто приводят к сбоям на выходах;
- 3) с помощью функции потока ошибок произвести синтез схемы, заменив уязвимые участки более устойчивыми к одиночным сбоям аналогами;
- 4) оценить надежность полученной схемы с помощью коэффициента логической чувствительности.

Заключение. В представленной работе рассмотрены современные методы и подходы к оценке сбоеустойчивости логических схем, а также способы повышения их надежности. Авторами предложена методика логического синтеза комбинационных схем с применением метода резолюций, функции потока ошибок и коэффициента логической чувствительности. Предполагается, что предложенная авторами техника способна обеспечить более быстрое выполнение синтеза логических схем с целью повышения их помехоустойчивости к одиночным сбоям по сравнению с существующими подходами. Практическое исследование этой гипотезы представляет собой направление дальнейшей работы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Mahatme N.N. et al. Impact of technology scaling on the combinational logic soft error rate // 2014 IEEE international reliability physics symposium. – IEEE, 2014. – С. 5F. 2.1-5F. 2.6.
2. Heath J.R. et al. A defect-tolerant computer architecture: Opportunities for nanotechnology // Science. – 1998. – Vol. 280, No. 5370. – P. 1716-1721.

3. *Hu Y. et al.* Robust FPGA resynthesis based on fault-tolerant Boolean matching // 2008 IEEE/ACM International Conference on Computer-Aided Design. – IEEE, 2008. – P. 706-713.
4. *Luckenbill S. et al.* RALF: Reliability analysis for logic faults—An exact algorithm and its applications // 2010 Design, Automation & Test in Europe Conference & Exhibition (Date 2010). – IEEE, 2010. – P. 783-788.
5. *Han J. et al.* Reliability evaluation of logic circuits using probabilistic gate models // Microelectronics Reliability. – 2011. – Vol. 51, No. 2. – P. 468-476.
6. *Стемковский А.Л., Тельпухов Д.В., Соловьев Р.А., Тельпухова Н.В.* Исследование вероятностных методов оценки логической уязвимости комбинационных схем // Проблемы разработки перспективных микро- и наноэлектронных систем (МЭС) (IF 0,034). – 2016. – № 4. – С. 121-126.
7. *Glebov A. et al.* False-noise analysis using logic implications // IEEE/ACM International Conference on Computer Aided Design. ICCAD 2001. IEEE/ACM Digest of Technical Papers (Cat. No. 01CH37281). – IEEE, 2001. – P. 515-521.
8. *Гаврилов С.В.* Методы анализа логических корреляций для САПР цифровых КМОП СБИС: учеб. пособие. – 2011.
9. *Robinson J.A.* A machine-oriented logic based on the resolution principle // Journal of the ACM (JACM). – 1965. – Vol. 12, No. 1. – P. 23-41.
10. *Lyons R.E., Vanderkulk W.* The use of triple-modular redundancy to improve computer reliability // IBM journal of research and development. – 1962. – Vol. 6, No. 2. – P. 200-209.
11. *Kastensmidt F.L. et al.* On the optimal design of triple modular redundancy logic for SRAM-based FPGAs // Design, Automation and Test in Europe. – IEEE, 2005. – P. 1290-1295.
12. *Samudrala P.K., Ramos J., Katkooori S.* Selective triple modular redundancy (STMR) based single-event upset (SEU) tolerant synthesis for FPGAs // IEEE transactions on Nuclear Science. – 2004. – Vol. 51, No. 5. – P. 2957-2969.
13. *Hamamatsu M., Tsuchiya T., Kikuno T.* On the reliability of cascaded TMR systems // 2010 IEEE 16th Pacific Rim International Symposium on Dependable Computing. – IEEE, 2010. – P. 184-190.
14. *Gomes I.A. C. et al.* Exploring the use of approximate TMR to mask transient faults in logic with low area overhead // Microelectronics Reliability. – 2015. – Vol. 55, No. 9-10. – P. 2072-2076.
15. *Гуров С.И.* Спектральный R-код с проверками на чётность // Прикладная математика и информатика. – 2017. – С. 91-96.
16. *Стемковский А.Л., Тельпухов Д.В., Жукова Т.Д., Гуров С.И., Соловьев Р.А.* Методы синтеза сбоеустойчивых комбинационных КМОП схем, обеспечивающих автоматическое исправление ошибок // Известия ЮФУ. Технические науки. – 2017. – № 7 (192). – С. 197-210.
17. *Гаврилов С.В., Гуров С.И., Жукова Т.Д., Рыжова Д.И., Тельпухов Д.В.* Методы повышения сбоеустойчивости комбинационных ИМС методами избыточного кодирования // Прикладная математика и информатика: Тр. факультета Вычислительной математики и кибернетики. – М.: Изд-во факультета ВМК МГУ, 2016. – № 53. – С. 93-102.
18. *Стемковский А.Л., Тельпухов Д.В., Соловьев Р.А., Мячиков М.В., Тельпухова Н.В.* Разработка технологически независимых метрик для оценки маскирующих свойств логических схем // Вычислительные технологии. – 2016. – Т. 21, № 2.
19. *Тельпухов Д.В., Соловьев Р.А., Тельпухова Н.В., Щелоков А.Н.* Оценка параметра логической чувствительности комбинационной схемы к однократным ошибкам с помощью вероятностных методов // Известия ЮФУ. Технические науки. – 2016. – № 7 (180). – С. 149-158.
20. *Cong J., Minkovich K.* LUT-based FPGA technology mapping for reliability // Proceedings of the 47th Design Automation Conference. – 2010. – P. 517-522.

REFERENCES

1. *Mahatme N.N. et al.* Impact of technology scaling on the combinational logic soft error rate, 2014 IEEE international reliability physics symposium. IEEE, 2014, pp. 5F. 2.1-5F. 2.6.
2. *Heath J.R. et al.* A defect-tolerant computer architecture: Opportunities for nanotechnology, Science, 1998, Vol. 280, No. 5370, pp. 1716-1721.
3. *Hu Y. et al.* Robust FPGA resynthesis based on fault-tolerant Boolean matching, 2008 IEEE/ACM International Conference on Computer-Aided Design. IEEE, 2008, pp. 706-713.

4. Luckenbill S. et al. RALF: Reliability analysis for logic faults—An exact algorithm and its applications, *2010 Design, Automation & Test in Europe Conference & Exhibition (Date 2010)*. IEEE, 2010, pp. 783-788.
5. Han J. et al. Reliability evaluation of logic circuits using probabilistic gate models, *Microelectronics Reliability*, 2011, Vol. 51, No. 2, pp. 468-476.
6. Stempkovskiy A.L., Tel'pukhov D.V., Solov'ev R.A., Tel'pukhova N.V. Issledovanie veroyatnostnykh metodov otsenki logicheskoy uyazvimosti kombinatsionnykh skhem [Investigation of probabilistic methods for evaluating the logical vulnerability of combinational schemes], *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem (MES) (IF 0,034)* [Problems of development of advanced micro- and nanoelectronic systems (MES) (IF 0,034)], 2016, No. 4, pp. 121-126.
7. Glebov A. et al. False-noise analysis using logic implications, *IEEE/ACM International Conference on Computer Aided Design. ICCAD 2001. IEEE/ACM Digest of Technical Papers (Cat. No. 01CH37281)*. IEEE, 2001, pp. 515-521.
8. Gavrilov S.V. Metody analiza logicheskikh korrelyatsiy dlya SAPR tsifrovyykh KMOP SBIS: ucheb. posobie [Methods for analyzing logical correlations for CAD digital CMOS VLSI: textbook], 2011.
9. Robinson J.A. A machine-oriented logic based on the resolution principle, *Journal of the ACM (JACM)*, 1965, Vol. 12, No. 1, pp. 23-41.
10. Lyons R.E., Vanderkulk W. The use of triple-modular redundancy to improve computer reliability, *IBM journal of research and development*, 1962, Vol. 6, No. 2, pp. 200-209.
11. Kastensmidt F.L. et al. On the optimal design of triple modular redundancy logic for SRAM-based FPGAs, *Design, Automation and Test in Europe*. IEEE, 2005, pp. 1290-1295.
12. Samudrala P.K., Ramos J., Katkooori S. Selective triple modular redundancy (STMR) based single-event upset (SEU) tolerant synthesis for FPGAs, *IEEE transactions on Nuclear Science*, 2004, Vol. 51, No. 5, pp. 2957-2969.
13. Hamamatsu M., Tsuchiya T., Kikuno T. On the reliability of cascaded TMR systems, *2010 IEEE 16th Pacific Rim International Symposium on Dependable Computing*. IEEE, 2010, pp. 184-190.
14. Gomes I.A. C. et al. Exploring the use of approximate TMR to mask transient faults in logic with low area overhead, *Microelectronics Reliability*, 2015, Vol. 55, No. 9-10, pp. 2072-2076.
15. Gurov S.I. Spektral'nyy R-kod s proverkami na chetnost' [Spectral R-code with parity checks], *Prikladnaya matematika i informatika* [Applied mathematics and computer science], 2017, pp. 91-96.
16. Stempkovskiy A.L., Tel'pukhov D.V., Zhukova T.D., Gurov S.I., Solov'ev R.A. Metody sinteza sboeustoychivyykh kombinatsionnykh KMOP skhem, obespechivayushchikh avtomaticheskoe ispravlenie oshibok [Methods for synthesizing fault-tolerant combinational CMOS circuits that provide automatic error correction], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2017, No. 7 (192), pp. 197-210.
17. Gavrilov S.V., Gurov S.I., Zhukova T.D., Ryzhova D.I., Tel'pukhov D.V. Metody povysheniya sboeustoychivosti kombinatsionnykh IMS metodami izbytochnogo kodirovaniya [Methods for improving the fault tolerance of combinational ICS using redundant coding methods], *Prikladnaya matematika i informatika: Tr. fakul'teta Vychislitel'noy matematiki i kibernetiki* [Applied mathematics and computer science: Proceedings of the faculty of Computational mathematics and Cybernetics]. Moscow: Izd-vo fakul'teta VMK MGU, 2016, No. 53, pp. 93-102.
18. Stempkovskiy A.L., Tel'pukhov D.V., Solov'ev R.A., Myachikov M.V., Tel'pukhova N.V. Razrabotka tekhnologicheskikh nezavisimyykh metrik dlya otsenki maskiruyushchikh svoystv logicheskikh skhem [Development of technologically independent metrics for evaluating masking properties of logic circuits], *Vychislitel'nye tekhnologii* [Computing technologies], 2016, Vol. 21, No. 2.
19. Tel'pukhov D.V., Solov'ev R.A., Tel'pukhova N.V., ShchHelokov A.N. Otsenka parametra logicheskoy chuvstvitel'nosti kombinatsionnoy skhemy k odnokratnym oshibkam s pomoshch'yu veroyatnostnykh metodov [Estimation of the logical sensitivity parameter of a combinational scheme to single errors using probabilistic methods], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2016, No. 7 (180), pp. 149-158.
20. Cong J., Minkovich K. LUT-based FPGA technology mapping for reliability, *Proceedings of the 47th Design Automation Conference*, 2010, pp. 517-522.

Статью рекомендовал к опубликованию к.т.н. В.А. Иванов.

Васильев Николай Олегович – Институт проблем проектирования в микроэлектронике РАН (ИППМ РАН); e-mail: vasilyev_n@ippm.ru; 124365, Москва, Зеленоград, ул. Советская, 3; инженер-исследователь.

Заплетина Мария Андреевна – e-mail: zapletina_m@ippm.ru; м.н.с.

Иванова Галина Александровна – e-mail: ivanova_g@ippm.ru; с.н.с.; к.т.н.

Щелоков Альберт Николаевич – e-mail: schan@ippm.ru; зам. директора; к.ф.-м.н.

Vasilyev Nikolay Olegovich – The Institute for Design Problems in Microelectronics (IPPM RAS); e-mail: vasilyev_n@ippm.ru; 3, Sovetskaya street, Zelenograd, Moscow, 124365, Russia; research engineer.

Zapletina Mariya Andreevna – e-mail: zapletina_m@ippm.ru; junior researcher.

Ivanova Galina Aleksandrovna – e-mail: ivanova_g@ippm.ru; senior researcher; cand. of eng. sc.

Schelokov Albert Nikolaevich – e-mail: schan@ippm.ru; Deputy Director; cand. phys. and math. sc.

УДК 004.896

DOI 10.18522/2311-3103-2020-4-125-136

Б.К. Лебедев, В.Б. Лебедев, О.Б. Лебедев

ПОИСКОВЫЙ ПОПУЛЯЦИОННЫЙ АЛГОРИТМ РАЗМЕЩЕНИЯ ЭЛЕМЕНТОВ СБИС*

В работе рассматривается поисковый популяционный алгоритм размещения компонентов СБИС. По аналогии с процессом возникновения и формирования кристаллов из вещества, процесс порождения решения путем последовательного проявления и конкретизации решения на базе интегральной россыпи альтернатив назван методом кристаллизации россыпи альтернатив. Решение Q_k задачи размещения представляется в виде биективного отображения $F_k=A \rightarrow P$, каждому элементу множества A соответствует один единственный элемент множества P и наоборот. Лежащая в основе алгоритма метаэвристика кристаллизации россыпи альтернатив выполняет поиск решений с учетом коллективной эволюционной памяти, под которой подразумевается информация, отражающая историю поиска решения и памяти поисковой процедуры. Отличительной особенностью используемой метаэвристики является учет тенденции к использованию альтернатив из наилучших найденных решений. Предложены компактные структуры данных для хранения интерпретаций решений и памяти. Алгоритм, связанный с эволюционной памятью, стремится к запоминанию и многократному использованию способов достижения лучших результатов. Разработанный алгоритм относится к классу популяционных алгоритмов. Итерационный процесс поиска решений включает три этапа. На первом этапе каждой итерации конструктивным алгоритмом формируется n_a решений Q_k . Работа конструктивного алгоритма базируется на базе показателей основной интегральной россыпи альтернатив – матрицы R , в которой хранятся интегральные показатели решений, полученных на предыдущих итерациях. Процесс назначения элемента в позицию включает две стадии. На первой стадии выбирается элемент, а на второй стадии – позиция r_j . При этом должно выполняться ограничение: каждому элементу соответствует одна позиция r_j . Рассчитывается оценка ξ_k решения Q_k и оценка полезности δ_k множества позиций P_k выбранных агентами. В работе используется циклический метод формирования решений. В этом случае наращивание оценок интегральной полезности δ_k в основной интегральной россыпи альтернатив V выполняется после полного формирования множества решений Q . На втором этапе итерации производится наращивание оценок интегральной полезности δ_k в основной интегральной россыпи альтернатив – матрице R . На третьем этапе итерации осуществляется

* Работа выполнена при финансовой поддержке гранта РФФИ № 20-07-00260 А.

снижение оценок полезности δ_k интегральной россыпи альтернатив R на априори заданную величину δ^* . Работа алгоритма завершается после выполнения заданного числа итераций. Сравнительный анализ с другими алгоритмами решения производился на стандартных тестовых примерах (бенчмарках) корпорации IBM, при этом решения, синтезируемые алгоритмом CAF, превосходят по эффективности решения известных методов в среднем на 6%. Временная сложность алгоритма – $O(n^2)$ - $O(n^3)$.

Размещение; СБИС; коллективная эволюционная память; оценка полезности; оптимизация; популяционный алгоритм; метод кристаллизации россыпи альтернатив.

В.К. Lebedev, V.B. Lebedev, O.B. Lebedev

SEARCH POPULATION ALGORITHM FOR VLSI ELEMENTS PLACEMENT

The paper considers a population search algorithm for the placement of VLSI components. By analogy with the process of the emergence and formation of crystals from matter, the process of generating a solution by sequential manifestation and concretization of the solution based on an integral placer of alternatives is called the method of crystallization of a placer of alternatives. The solution Q_k of the placement problem is represented as a bijective mapping $F_k = A \rightarrow P$, each element of the set A corresponds to one single element of the set P and vice versa. The metaheuristic of crystallization of a placer of alternatives underlying the algorithm searches for solutions taking into account collective evolutionary memory, which means information reflecting the history of the search for a solution and the memory of the search procedure. A distinctive feature of the metaheuristic used is that it takes into account the tendency to use alternatives from the best found solutions. Compact data structures for storing solution interpretations and memory are proposed. An algorithm associated with evolutionary memory seeks to memorize and reuse ways to achieve better results. The developed algorithm belongs to the class of population. The iterative process of finding solutions includes three stages. At the first stage of each iteration, the constructive algorithm generates n_q solutions Q_k . The work of the constructive algorithm is based on the indicators of the main integral placer of alternatives – the matrix R , which stores the integral indicators of the solutions obtained at the previous iterations. The process of assigning an item to a position involves two stages. In the first stage, the element is selected, and in the second stage, the position p_j . In this case, the restriction must be fulfilled: each element corresponds to one position p_j . The estimate ξ_k of the solution Q_k and the estimate of the utility δ_k of the set of positions P_k selected by the agents are calculated. The work uses a cyclical method of forming decisions. In this case, the accumulation of estimates of the integral utility δ_k in the main integral placer of alternatives R is performed after the complete formation of the set of solutions Q . At the second stage of the iteration, the estimates of the integral utility δ_k are increased in the main integral placer of alternatives – the matrix R . At the third stage of the iteration, the estimates of the utility δ_k of the integral placer of alternatives R are reduced by a priori a given value δ^* . The algorithm ends after the specified number of iterations has been completed. Comparative analysis with other solution algorithms was carried out on standard test examples (benchmarks) of the IBM corporation, while the solutions synthesized by the CAF algorithm exceed the solution efficiency of the known methods by an average of 6%. The time complexity of the algorithm is $O(n^2)$ - $O(n^3)$.

Placement; VLSI; collective evolutionary memory; utility assessment; optimization; population algorithm; method of crystallization of alternatives placer.

Введение. Разработка поисковых алгоритмов базируется на использовании метаэвристик, заложенных в природных механизмах принятия решений [1–3]. Большинство известных алгоритмов [4–7] используют традиционные итерационные улучшающие структуры, основанные на случайном поиске. Основным недостатком, присущим этому подходу, является вхождение алгоритмов в локальный оптимум, часто далекий от глобального. В работе по аналогии с метаэвристиками на которых построены роевые алгоритмы [3], используется метаэвристика, учитывающая тенденцию к использованию альтернатив (вариантов) из наилучших найденных решений [8,9]. В процессе эволюционной коллективной адаптации методами дискриминантного анализа формируются оценки приспособленности альтер-

натив. Вероятность использования альтернативы в формируемом решении пропорциональна ее приспособленности. Совокупность данных об альтернативах и их оценках составляет *россыпь альтернатив* [8, 9]. По аналогии с процессом образования и роста кристаллов из вещества, процесс порождения решения путем последовательного проявления и конкретизации на базе интегральной россыпи альтернатив назван методом кристаллизации россыпи альтернатив. Другими словами, в процессе эволюционной коллективной адаптации решение «кристаллизуется» в состояние, определяемое совокупностью приспособленных альтернатив. Отсюда название метода оптимизации – *метод кристаллизации россыпи альтернатив* [8, 9].

Задача размещения является одним из наиболее важных этапов в процессе проектирования СБИС, поскольку качество полученного решения в значительной степени влияет на выполнение последующей трассировки соединений [10]. Задача размещения относится к классу NP -полных, поэтому не существует алгоритма, позволяющего получать ее решение за полиномиальное время. Проведенные исследования имеющихся подходов к ее решению показывают, что существующие алгоритмы не в полной мере отражают требования по эффективности [4], получаемые решения, с точки зрения их оптимальности неудовлетворительны. Необходимость повышения качества проектирования требует поиска новых путей и подходов к решению задач размещения. Изучение имеющихся методов и алгоритмов решения комбинаторных задач показало, что решения, чаще всего представляются списками. Основу таких алгоритмов составляет процедура формирования упорядоченного списка элементов, задающего последовательность формирования решения базовым конструктивным алгоритмом. Формально упорядоченный список – это код решения. Переход от кода (упорядоченный список) к фенотипу (решению) и вычисление значения функции пригодности выполняется процедурой декодирования. Обычно в качестве декодера используются конструктивные алгоритмы. Активно оперирование с решениями, представленными в закодированном виде, развиваются в алгоритмах поисковой оптимизации (биоинспирированными, роевыми, популяционными) [11–16]. Анализ таких алгоритмов показал, что их эффективность выше, чем у классических алгоритмов. В работе рассматривается поисковый популяционный алгоритм размещения элементов СБИС, основанный на методе кристаллизации россыпи альтернатив (КРА) [11–16].

В работе излагается алгоритм размещения элементов на основе метода кристаллизации россыпи альтернатив. С этой целью разработана единая структура данных. С учетом особенностей единой структуры данных разработан поисковый процесс, базирующийся на моделировании коллективной адаптации.

1. Постановка задачи размещения. Проблема размещения может быть сформулирована следующим образом [11–13]. Дано множество элементов (модулей) с расположенными на них терминалами (выводами). Задано множество цепей, связывающих терминалы модулей. Задано коммутационное поле (КП) с описанием позиций, в которых могут размещаться элементы. Необходимо разместить элементы на КП с оптимизацией некоторых критериев качества. Входная информация включает: описание модулей, в котором указываются форма, размеры, расположение терминалов на модулях, множество цепей, связывающих терминалы модулей, и описание КП [14]. Выходная информация включает координаты позиций на КП для размещения всех модулей.

Стандартная постановка задачи размещения элементов в позициях на коммутационном поле формулируется следующим образом [14]. Дано множество элементов $A = \{a_i | i = 1, 2, \dots, n\}$ и множество позиций $P = \{p_j | j = 1, 2, \dots, n_p\}$ на КП. В качестве модели схемы используется гиперграф $H = (X, E)$, где $X = \{x_i | i = 1, 2, \dots, n\}$ – множество вершин, моделирующих элементы, $E = \{e_j | e_j \subset X, j = 1, 2, \dots, m\}$ – множество гиперре-

бер, моделирующих цепи, связывающие элементы. Для размещения всех элементов необходимо выполнение условия $n_p \geq n$. Целевая функция ζ определяется выбранными критериями. Основными, известными критериями при размещении являются: минимальная суммарная длина связей, минимальная длина самой длинной связи, минимум числа возможных пересечений, минимум возможного числа изгибов соединений, минимальная площадь кристалла.

Каждая позиция p_j имеет координаты (x_j, y_j) . Между множеством порядковых номеров позиций и множеством их координат устанавливается взаимно однозначное соответствие. Пусть имеется множество элементов, размещенных в позициях. Произвольное размещение элементов в позициях представляет собой перестановку (упорядоченный список) $Q = \langle q_j | j=1, 2, \dots, n_p \rangle$, где q_j – номер элемента, размещенного в j -ой позиции. Задача размещения состоит в отыскании оптимального значения функции ζ на множестве перестановок Q [11-14]. Отметим, что в общем случае может использоваться любая модель соединений и функция их оценки. И это никоим образом не влияет на структуру кодирования и декодирования решения.

2. Механизмы размещения на основе метода кристаллизации россыпи альтернатив. В методе кристаллизации россыпи альтернатив [9,10] каждое решение Q_k формируется (представляется) множеством агентов $A = \{a_i | i=1, 2, \dots, n_a\}$, где n_a – число агентов. Каждому агенту a_i соответствует множество альтернативных состояний $S_i = \{s_{ij} | j=1, 2, \dots, n_s\}$, где n_s – число состояний агента a_i . Каждый агент a_i может находиться в одном из альтернативных состояний. Решение Q_k определяется совокупностью альтернативных состояний множества агентов [8, 9].

В рассмотренной задаче размещения агентами являются элементы, а альтернативами – множество позиций $P = \{p_j | j=1, 2, \dots, n_p\}$, где n_p – число позиций в которые помещаются элементы. Другими словами множество агентов $A = \{a_i | i=1, 2, \dots, n_a\}$, соответствует множеству всех элементов. Множество альтернативных состояний $S_i = \{s_{ij} | j=1, 2, \dots, n_p\}$ агента a_i соответствует множеству альтернативных вариантов размещения агента a_i – множеству позиций, в которые может быть помещен элемент.

Решение Q_k задачи размещения представляется в виде биективного отображения $F_k = A \rightarrow P$, каждому элементу множества A соответствует один единственный элемент множества P и наоборот. Номер позиции p_j в которую помещен элемент a_i определяется соотношением $p_j = F_k(a_i)$.

Под *россыпью альтернатив* (РА) решения в работе подразумевается структура данных, несущая информацию об альтернативах агентов в данном решении и об полезности решения, используемой в качестве коллективной эволюционной памяти (КЭП) [8].

Для хранения информации об занимаемых агентами решения Q_k альтернативных позициях (порядковых номерах), и значениях их пригодностей, составляющих коллективную эволюционную память (КЭП), используется матрица $R_k = \|r_{kij}\|_{m \times m}$, называемая *россыпью альтернатив* (РА) (рис. 1).

$$R_k =$$

P	p_1	p_2	p_3	p_4	p_5	p_6
a_1	0	6	0	0	0	0
a_2	0	0	0	0	6	0
a_3	6	0	0	0	0	0
a_4	0	0	6	0	0	0
a_5	0	0	0	0	0	6
a_6	0	0	0	6	0	0

Рис. 1. Россыпь альтернатив R_k

Представим структуру данных для отображения решения Q_k , формируемого множеством агентов A , в виде совокупности векторов $R_k = \{R_{ki} | i=1, 2, \dots, n_a\}$. Каждый вектор $R_{ki} = \{r_{kij} | j=1, 2, \dots, n_p\}$ соответствует агенту a_i и фактически является строкой матрицы R_k . Размерность вектора R_{ki} определяется числом возможных состояний агента a_i . В векторе R_{ki} только один элемент r_{kij} соответствующий состоянию s_{ij} , в котором находится агент a_i , имеет значение, отличное от нуля и это значение равно $\delta_k = f(\zeta_k)$, где ζ_k – оценка решения Q_k , δ_k – оценка полезности этого решения. Остальные элементы вектора R_{ki} имеют нулевые значения. Таким образом, в векторе R_{ki} хранится информация о состоянии, реализованном агентом a_i в решении Q_k , и об оценке полезности δ_k этого состояния.

Пример. Пусть шестью агентами сформировано решение $Q_k = \langle 2, 5, 1, 3, 6, 4 \rangle$. В решении Q_k агентами реализованы следующие состояния (альтернативы): a_1 в состоянии s_{12} , $a_2 - s_{25}$, $a_3 - s_{31}$, $a_4 - s_{43}$, $a_5 - s_{56}$, $a_6 - s_{64}$, которым соответствуют множество Θ_k ячеек r_{kij} матрицы R_k : $\Theta_k = \{r_{k12}, r_{k25}, r_{k31}, r_{k43}, r_{k56}, r_{k64}\}$. Элементам множества Θ_k присваивается значение оценки полезности $\delta_k = 6$. Тогда россыпь альтернатив R_k для решения Θ_k имеет вид, представленный на рис. 1.

Пусть для решения задачи размещения построено множество решений $Q = \{Q_k | k=1, 2, \dots, n_q\}$:

$Q_1 = \langle 2, 5, 1, 3, 6, 4 \rangle$; $Q_2 = \langle 2, 4, 1, 6, 3, 5 \rangle$; $Q_3 = \langle 2, 4, 1, 6, 3, 5 \rangle$; $Q_4 = \langle 5, 2, 4, 3, 6, 1 \rangle$; $Q_5 = \langle 1, 4, 2, 5, 3, 6 \rangle$; $Q_6 = \langle 5, 6, 3, 4, 2, 1 \rangle$. Для каждого решения Q_k рассчитаны оценки ζ_k и δ_k и сформирована индивидуальная россыпь альтернатив R_k . $\delta_1 = 6$, $\delta_2 = 8$, $\delta_3 = 4$, $\delta_4 = 2$, $\delta_5 = 9$, $\delta_6 = 5$.

Посредством объединения всех россыпей формируется интегральная россыпь альтернатив R альтернатив:

$$\begin{aligned} R &= \|r_{ij}\|_{m \times n} = \cup R_k | k=1, 2, \dots, n_q; \\ R_k &= \|r_{kij}\|_{m \times n} = \{R_{ki} | i=1, 2, \dots, n_a\}; \\ R_{ki} &= \{r_{kij} | j=1, 2, \dots, n_p\}; \\ r_{kij} &= \sum_k (r_{kij}), k = \{1, n_q\}. \end{aligned} \tag{1}$$

Фактически r_{kij} является суммарным значением полезностей решений, в которых агентом a_i была реализована альтернатива s_{ij} .

Произведем интеграцию решений и построим интегральную россыпь альтернатив множества решений $Q = \{Q_1, Q_2, Q_3, Q_4, Q_5, Q_6\}$. Агентами в каждом из решений реализованы следующие альтернативы (табл. 1):

Таблица 1
Реализация альтернатив

Q	Агент ы						δ
	a_1	a_2	a_3	a_4	a_5	a_6	
Q_1	2	5	1	3	6	4	$\delta_1=6$
Q_2	2	4	1	6	3	5	$\delta_2=8$
Q_3	2	4	1	6	3	4	$\delta_3=4$
Q_4	5	2	4	3	6	1	$\delta_4=6$
Q_5	2	4	2	5	3	6	$\delta_5=9$
Q_6	5	6	3	4	2	1	$\delta_6=5$

На рис. 2. представлены россыпи альтернатив $\{R_1, R_2, R_3, R_4, R_5, R_6\}$ множества решений Q .

Произведем расчет значений элементов r_{ij} интегральной россыпи альтернатив по формуле 1. Интегральная россыпь альтернатив представлена на рис. 3.

$R_1 =$	<table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr><th>P</th><th>p_1</th><th>p_2</th><th>p_3</th><th>p_4</th><th>p_5</th><th>p_6</th></tr> </thead> <tbody> <tr><td>a_1</td><td>0</td><td>6</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_2</td><td>0</td><td>0</td><td>0</td><td>0</td><td>6</td><td>0</td></tr> <tr><td>a_3</td><td>6</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_4</td><td>0</td><td>0</td><td>6</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_5</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>6</td></tr> <tr><td>a_6</td><td>0</td><td>0</td><td>0</td><td>6</td><td>0</td><td>0</td></tr> </tbody> </table>	P	p_1	p_2	p_3	p_4	p_5	p_6	a_1	0	6	0	0	0	0	a_2	0	0	0	0	6	0	a_3	6	0	0	0	0	0	a_4	0	0	6	0	0	0	a_5	0	0	0	0	0	6	a_6	0	0	0	6	0	0	$R_2 =$	<table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr><th>P</th><th>p_1</th><th>p_2</th><th>p_3</th><th>p_4</th><th>p_5</th><th>p_6</th></tr> </thead> <tbody> <tr><td>a_1</td><td>0</td><td>8</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_2</td><td>0</td><td>0</td><td>0</td><td>8</td><td>0</td><td>0</td></tr> <tr><td>a_3</td><td>8</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>8</td></tr> <tr><td>a_5</td><td>0</td><td>0</td><td>8</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_6</td><td>0</td><td>0</td><td>0</td><td>0</td><td>8</td><td>0</td></tr> </tbody> </table>	P	p_1	p_2	p_3	p_4	p_5	p_6	a_1	0	8	0	0	0	0	a_2	0	0	0	8	0	0	a_3	8	0	0	0	0	0	a_4	0	0	0	0	0	8	a_5	0	0	8	0	0	0	a_6	0	0	0	0	8	0
P	p_1	p_2	p_3	p_4	p_5	p_6																																																																																															
a_1	0	6	0	0	0	0																																																																																															
a_2	0	0	0	0	6	0																																																																																															
a_3	6	0	0	0	0	0																																																																																															
a_4	0	0	6	0	0	0																																																																																															
a_5	0	0	0	0	0	6																																																																																															
a_6	0	0	0	6	0	0																																																																																															
P	p_1	p_2	p_3	p_4	p_5	p_6																																																																																															
a_1	0	8	0	0	0	0																																																																																															
a_2	0	0	0	8	0	0																																																																																															
a_3	8	0	0	0	0	0																																																																																															
a_4	0	0	0	0	0	8																																																																																															
a_5	0	0	8	0	0	0																																																																																															
a_6	0	0	0	0	8	0																																																																																															
$R_3 =$	<table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr><th>P</th><th>p_1</th><th>p_2</th><th>p_3</th><th>p_4</th><th>p_5</th><th>p_6</th></tr> </thead> <tbody> <tr><td>a_1</td><td>0</td><td>4</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_2</td><td>0</td><td>0</td><td>0</td><td>4</td><td>0</td><td>0</td></tr> <tr><td>a_3</td><td>4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>4</td></tr> <tr><td>a_5</td><td>0</td><td>0</td><td>4</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_6</td><td>0</td><td>0</td><td>0</td><td>0</td><td>4</td><td>0</td></tr> </tbody> </table>	P	p_1	p_2	p_3	p_4	p_5	p_6	a_1	0	4	0	0	0	0	a_2	0	0	0	4	0	0	a_3	4	0	0	0	0	0	a_4	0	0	0	0	0	4	a_5	0	0	4	0	0	0	a_6	0	0	0	0	4	0	$R_4 =$	<table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr><th>P</th><th>p_1</th><th>p_2</th><th>p_3</th><th>p_4</th><th>p_5</th><th>p_6</th></tr> </thead> <tbody> <tr><td>a_1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>2</td><td>0</td></tr> <tr><td>a_2</td><td>0</td><td>2</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_3</td><td>0</td><td>0</td><td>0</td><td>2</td><td>0</td><td>0</td></tr> <tr><td>a_4</td><td>0</td><td>0</td><td>2</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_5</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>2</td></tr> <tr><td>a_6</td><td>2</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	P	p_1	p_2	p_3	p_4	p_5	p_6	a_1	0	0	0	0	2	0	a_2	0	2	0	0	0	0	a_3	0	0	0	2	0	0	a_4	0	0	2	0	0	0	a_5	0	0	0	0	0	2	a_6	2	0	0	0	0	0
P	p_1	p_2	p_3	p_4	p_5	p_6																																																																																															
a_1	0	4	0	0	0	0																																																																																															
a_2	0	0	0	4	0	0																																																																																															
a_3	4	0	0	0	0	0																																																																																															
a_4	0	0	0	0	0	4																																																																																															
a_5	0	0	4	0	0	0																																																																																															
a_6	0	0	0	0	4	0																																																																																															
P	p_1	p_2	p_3	p_4	p_5	p_6																																																																																															
a_1	0	0	0	0	2	0																																																																																															
a_2	0	2	0	0	0	0																																																																																															
a_3	0	0	0	2	0	0																																																																																															
a_4	0	0	2	0	0	0																																																																																															
a_5	0	0	0	0	0	2																																																																																															
a_6	2	0	0	0	0	0																																																																																															
$R_5 =$	<table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr><th>P</th><th>p_1</th><th>p_2</th><th>p_3</th><th>p_4</th><th>p_5</th><th>p_6</th></tr> </thead> <tbody> <tr><td>a_1</td><td>9</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_2</td><td>0</td><td>0</td><td>0</td><td>9</td><td>0</td><td>0</td></tr> <tr><td>a_3</td><td>0</td><td>9</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>9</td><td>0</td></tr> <tr><td>a_5</td><td>0</td><td>0</td><td>9</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_6</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>9</td></tr> </tbody> </table>	P	p_1	p_2	p_3	p_4	p_5	p_6	a_1	9	0	0	0	0	0	a_2	0	0	0	9	0	0	a_3	0	9	0	0	0	0	a_4	0	0	0	0	9	0	a_5	0	0	9	0	0	0	a_6	0	0	0	0	0	9	$R_6 =$	<table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr><th>P</th><th>p_1</th><th>p_2</th><th>p_3</th><th>p_4</th><th>p_5</th><th>p_6</th></tr> </thead> <tbody> <tr><td>a_1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>5</td><td>0</td></tr> <tr><td>a_2</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>5</td></tr> <tr><td>a_3</td><td>0</td><td>0</td><td>5</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_4</td><td>0</td><td>0</td><td>0</td><td>5</td><td>0</td><td>0</td></tr> <tr><td>a_5</td><td>0</td><td>5</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>a_6</td><td>5</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	P	p_1	p_2	p_3	p_4	p_5	p_6	a_1	0	0	0	0	5	0	a_2	0	0	0	0	0	5	a_3	0	0	5	0	0	0	a_4	0	0	0	5	0	0	a_5	0	5	0	0	0	0	a_6	5	0	0	0	0	0
P	p_1	p_2	p_3	p_4	p_5	p_6																																																																																															
a_1	9	0	0	0	0	0																																																																																															
a_2	0	0	0	9	0	0																																																																																															
a_3	0	9	0	0	0	0																																																																																															
a_4	0	0	0	0	9	0																																																																																															
a_5	0	0	9	0	0	0																																																																																															
a_6	0	0	0	0	0	9																																																																																															
P	p_1	p_2	p_3	p_4	p_5	p_6																																																																																															
a_1	0	0	0	0	5	0																																																																																															
a_2	0	0	0	0	0	5																																																																																															
a_3	0	0	5	0	0	0																																																																																															
a_4	0	0	0	5	0	0																																																																																															
a_5	0	5	0	0	0	0																																																																																															
a_6	5	0	0	0	0	0																																																																																															

Рис. 2. Россыпи альтернатив решений $\{R_1, R_2, R_3, R_4, R_5, R_6\}$

$R =$	<table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr><th>P</th><th>p_1</th><th>p_2</th><th>p_3</th><th>p_4</th><th>p_5</th><th>p_6</th></tr> </thead> <tbody> <tr><td>a_1</td><td>9</td><td>18</td><td>0</td><td>0</td><td>7</td><td>0</td></tr> <tr><td>a_2</td><td>0</td><td>2</td><td>0</td><td>21</td><td>6</td><td>5</td></tr> <tr><td>a_3</td><td>18</td><td>9</td><td>5</td><td>2</td><td>0</td><td>0</td></tr> <tr><td>a_4</td><td>0</td><td>0</td><td>8</td><td>5</td><td>9</td><td>12</td></tr> <tr><td>a_5</td><td>0</td><td>5</td><td>13</td><td>0</td><td>0</td><td>8</td></tr> <tr><td>a_6</td><td>7</td><td>0</td><td>0</td><td>6</td><td>12</td><td>14</td></tr> </tbody> </table>	P	p_1	p_2	p_3	p_4	p_5	p_6	a_1	9	18	0	0	7	0	a_2	0	2	0	21	6	5	a_3	18	9	5	2	0	0	a_4	0	0	8	5	9	12	a_5	0	5	13	0	0	8	a_6	7	0	0	6	12	14
P	p_1	p_2	p_3	p_4	p_5	p_6																																												
a_1	9	18	0	0	7	0																																												
a_2	0	2	0	21	6	5																																												
a_3	18	9	5	2	0	0																																												
a_4	0	0	8	5	9	12																																												
a_5	0	5	13	0	0	8																																												
a_6	7	0	0	6	12	14																																												

Рис. 3. Интегральная россыпь альтернатив R

Работа алгоритма размещения базируется на использовании коллективной эволюционной памяти (КЭП), под которой подразумевается любой вид информации, которая отражает прошлую историю развития и хранится независимо от индивидуумов. Алгоритм, связанный с эволюционной памятью, стремится к запоминанию и многократному использованию способов достижения лучших результатов. КЭП алгоритма размещения представляет собой набор статистических показателей, отражающих для каждого фрагмента решения число δ , показывающее полезность фрагмента при построении решений на предыдущих итерациях алгоритма. Основу многоагентной системы составляет самоорганизация, обеспечивающая достижения общих целей коллектива на основе низкоуровневого взаимодействия. Коллектив не имеет централизованного управления, и его особенностями являются наличие непрямого обмена информацией. Непрямой обмен – стигмержи (stigmergy), представляет собой разнесённое во времени взаимодействие, при котором один агент изменяет некоторые области КЭП, а другие агенты позже используют уже изменённую информацию в этих областях, когда в неё попадают [1, 2].

Разработанный алгоритм относится к классу популяционных алгоритмов. Всем элементам основной матрицы россыпи альтернатив R присваивается начальное значение оценки полезности δ . Процесс поиска решений итерационный. Каждая итерация l включает три этапа.

На первом этапе каждой итерации конструктивным алгоритмом (КА) формируется n_q решений Q_k . Работа конструктивного алгоритма базируется на базе показателей основной интегральной россыпи альтернатив – матрицы $R = \|r_{ij}\|_{m \times m}$, структура которой описана выше, и в которой хранятся интегральные показатели решений, полученных на предыдущих итерациях.

Формирование каждого решения Q_k выполняется посредством выбора каждым агентом a_i позиции p_j . Каждое решение Q_k является биективным отображением $F_k = A \rightarrow P$ и формируется путем последовательного назначения элементов в позиции [14]. Процесс назначения элемента в позицию включает две стадии. На первой стадии выбирается элемент, а на второй стадии – позиция p_j . При этом должно выполняться ограничение: каждому элементу множества A соответствует один единственный элемент множества P и наоборот. Рассчитывается оценка ζ_k решения Q_k и оценка полезности δ_k множества позиций P_k выбранных агентами в решении Q_k . Запоминание показателей построенных решений осуществляется в два этапа с использованием основной R и промежуточной R^* матриц россыпей альтернатив. Формируется набор $\Theta_k = \{r_{kij}\}$ ячеек матрицы R^* , соответствующих альтернативам, выбранным агентами, для отображения решения Q_k в матрице R^* . После построения очередного решения Q_k его результаты заносятся в промежуточную россыпь альтернатив – матрицу R^* : все элементы набора Θ_k в матрице R^* увеличиваются на величину δ_k . В работе используется циклический метод формирования решений. В этом случае наращивание оценок интегральной полезности δ_k в основной интегральной россыпи альтернатив R выполняется после полного формирования множества решений Q на итерации, и внесения показателей этих решений в промежуточную матрицу россыпи альтернатив R^* .

На втором этапе итерации производится наращивание оценок интегральной полезности δ_k в основной интегральной россыпи альтернатив – матрице R , путем добавления матрицы R^* к матрице R . На третьем этапе итерации осуществляется снижение оценок полезности δ_k интегральной россыпи альтернатив R на априори заданную величину δ^* . Работа алгоритма завершается после выполнения заданного числа итераций.

3. Базовый алгоритм размещения методом кристаллизации россыпи альтернатив. Предварительно формируется структура основной – R и промежуточной R^* матриц россыпей альтернатив.

Введем обозначения:

$A(t)$ – множество агентов еще не размещенных в списке Q_k на шаге t ;

$P(t)$ – множество свободных позиций в списке n_p на шаге t ;

N_l – число итераций;

n_k – число решений;

l – номер итерации;

k – номер решения;

ζ_{min} – минимальное значение оценки размещения Q_k ;

δ – априорное (начальное) значение оценки полезности альтернатив.

1. Задаются параметры $N_k, N_b, \delta, Q_{min} = \emptyset, \zeta_{min} = 1000$.

2. $l = 1$.

3. Всем элементам основной матрицы россыпи альтернатив $R(l)$ присваивается начальное значение оценки полезности – δ .

4. $k = 1$.

5. Элементы промежуточной матрицы россыпи альтернатив $R^*(l)$ обнуляются.

6. Формирование конструктивным алгоритмом решения Q_k на базе матрицы $R(l)$.
7. Список Q_k процедурой декодирования трансформируется в решение Q_k .
8. Расчет оценки ζ_k решения и оценки полезности δ_k решения Q_k .
9. Формирование набора Θ ячеек для отображения решения Q_k в матрице $R^*(l)$.
10. Если $\zeta_k < \zeta_{min}$, то $\zeta_{min} = \zeta_k$, $Q_{min} = Q_k$.
11. Все элементы набора Θ_k в матрице $R^*(l)$ увеличиваются на величину δ_k .
12. Если $(k < n_k)$, то $k = k + 1$ и переход к 5, иначе переход к 13.
13. Сложение матриц $R(l)$ и $R^*(l)$. $R(l+1) = R(l) + R^*(l)$.
14. Снижение всех интегральных оценок полезности r_{ij} интегральной россыпи альтернатив $R(l)$ на величину δ^* .
15. Если $(l < N_l)$, то $l = l + 1$ и переход к 4, иначе переход к 16.
16. Конец работы алгоритма. Фиксация и вывод лучшего решения.

4. Конструктивный алгоритм формирование решения Q_k на базе интегральной россыпи альтернатив $R(l)$.

1. Приводится в начальное состояние память:

$$A(t) = A;$$

$$P(t) = P;$$

$$Q_k(t) = \emptyset;$$

$$t = 1.$$

2. Выбор очередного агента. Для каждого не размещенного $a_i \in A(t)$ в соответствующем векторе $R_i(l)$ – строке матрицы $R(l)$ отыскивается элемент r_{ij}^* с максимальным значением σ_i^* суммарной оценки полезности назначения a_i в p_j .

3. Среди множества не размещенных агентов $A(t)$ с максимальными значениями σ_i^* с вероятностью:

$$\Psi_i = \sigma_i^* / \sum_i (\sigma_i^*), (i/a_i \in A(t)),$$

пропорциональной σ_i^* , выбирается агент $a_i^* \in A(t)$.

4. На второй стадии – выбор позиции. Пусть μ_j значение суммарной оценки полезности назначения выбранного на первой стадии агента $a_i^* \in A(t)$ в позицию $p_j \in P(t)$. Среди множества свободных позиций $P(t)$ с вероятностью:

$$\Psi_j = \mu_j / \sum_j (\mu_j), (j/p_j \in P(t))$$

пропорциональной значению оценки полезности μ_j выбирается позиция p_j^* .

5. Агент a_i^* помещается в позицию p_j^* списка $Q_k(t)$.

6. Если $t < n_a$, то $t = t + 1$ и переход к 6, иначе переход к 7.

7. $A(t) = A(t) \setminus a_i^*$; $P(t) = P(t) \setminus p_j^*$. Переход к 2.

8. Конец работы алгоритма.

5. Экспериментальные исследования. Для формирования тестов были использованы процедуры синтеза контрольных примеров с известным оптимумом F_{opt} по аналогии с известными методами РЕКУ (Placement Examples with Known Upper bounds on wirelength) [15–16].

Разработанный алгоритм САФ сравнивался с современными алгоритмами размещения KraftWerk [6], Capo 8.6 [21] и гибридным биоинспирированным трехуровневым алгоритмом ГТА [18].

Экспериментальные исследования проводились на наборе тестовых схем РЕКУ [23, 22]. Размещение оценивалось по суммарной длине соединений. Результаты экспериментов приведены в табл. 2.

Таблица 2

Результаты экспериментов

№ теста	Kraft Werk	Саро8.6	ГТА	CAF
	F, (е6)	F, (е6)	F, (е6)	F, (е6)
1	5,02	4,97	4,94	4,74
2	11,45	10,32	10,11	9,82
3	16,2	14,06	13,84	13,8
4	18,98	18,13	18,02	17,9
5	24,7	21,96	21,68	21,51
6	38,1	36,06	35,96	35,01
7	31,7	30,28	29,98	29,76
8	39,43	37,86	37,74	37,30
9	66,6	61,25	61,05	59,32
10	58,2	56,45	56,23	54,84

По представленным в таблице результатам можно заключить, что по качеству решения на тестовых примерах РЕКУ в среднем алгоритм CAF превосходит известные алгоритмы KraftWerk на 6 %, Саро8.6 на 4 % и ГТА) на 2 %.

Сравнительный анализ с другими алгоритмами размещения производился на стандартных тестовых примерах (бенчмарках) корпорации IBM [19, 20]. В этом случае сравнение разработанного алгоритма CAF с известными методами QPlace [22], mPG [23], A4 [7] и производилось на тестовых схемах корпорации IBM. Результаты экспериментов приведены в табл. 3. Алгоритм CAF в среднем превосходит программные средства Qplace, mPG и A4.

Таблица 3

Результаты сравнения

№ теста	QPlace	mPG	Ap1	CAF
	F, (е6)	F, (е6)	F, (е6)	F, (е6)
1	4,94	4,97	5,06	4,94
2	10,10	10,32	10,87	10,11
3	13,81	14,06	15,32	13,84
4	18,01	18,13	18,94	18,02
5	21,62	21,96	23,77	21,68
6	35,91	36,06	37,24	35,96
7	29,92	30,28	31,29	29,98
8	37,71	37,86	38,93	37,74
9	61,01	61,25	62,52	61,05
10	56,22	56,47	57,38	56,26

Еще одной важной задачей проведенного тестового исследования является определение сходимости разработанного алгоритма. Тестирование определило, что разработанный алгоритм размещения методом кристаллизации россыпи альтернатив (CAF) в среднем имеет сходимость на 125 итерации. Проведенное тестирование показало, что вероятность того, что будет найдено оптимальное решение, в среднем равняется – 0.9. В заключении отметим, что предложенный алгоритм CAF способствует формированию решений, не уступающими полученным известными методами, а иногда качество решений выше в среднем на 6%.

Заключение. Задача размещения – является одной из широко востребованных задач синтеза топологии СБИС. В работе рассматривается поисковый популяционный алгоритм размещения компонентов СБИС. Лежащая в основе алгоритма метаэвристика кристаллизации россыпи альтернатив выполняет поиск решений с учетом коллективной эволюционной памяти (КЭП), под которой подразумевается информация, отражающая историю поиска решения и памяти поисковой процеду-

ры. Отличительной особенностью используемой метаэвристики является учет тенденции к использованию альтернатив из наилучших найденных решений. Предложены компактные структуры данных для хранения интерпретаций решений и памяти. Разработанный конструктивный алгоритм с временной сложностью $O(n)$ для размещения интегрирован с итерационной структурой поиска. Алгоритм использует представление решения задачи размещения в виде упорядоченного списка. Переход от списка к решению задачи размещения производится с помощью декодера. Для перехода от упорядоченного списка к решению разработана процедура декодирования с временной сложностью $O(n)$. Рассмотрены главные принципы анализа и выбора альтернатив в процессе адаптации популяции. К преимуществам алгоритма относятся: невысокая трудоемкость, временная сложность алгоритма – $O(n^2)$; возможность диверсификации областей применения; алгоритм отличается быстрой сходимостью и высокой степенью эффективности.

Эффективность предложенного подхода подтверждена в процессе тестовых испытаний.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Карпенко А.П. Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой: учеб. пособие. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2014. – 448 с.
2. Лебедев О.Б. Модели адаптивного поведения муравьиной колонии в задачах проектирования. – Таганрог. Изд-во ЮФУ, 2013. – 199 с.
3. Лебедев Б.К., Лебедев О.Б., Лебединский А.Е. Гибридный биоинспирированный алгоритм размещения базовых стандартных библиотечных элементов при проектировании топологии полужаказной СБИС // Известия ЮФУ. Технические науки. – 2019. – № 3 (205). – С. 97-110.
4. Курейчик В.В., Курейчик В.В. Интегрированный алгоритм размещения фрагментов СБИС // Известия ЮФУ. Технические науки. – 2014. – № 7 (156). – С. 84-91.
5. Лебедев Б.К., Лебедев О.Б., Лебедев В.Б. Гибридизация роевого интеллекта и генетической эволюции на примере размещения // Программные продукты, системы и алгоритмы. – 2017. – № 4.
6. Caldwell A.E. Can Recursive Bisection Alone Produce Routable Placements // DAC. – 2000. – P. 477-482.
7. Mourelle M. Swarm intelligent systems. – Berlin: Heidelberg: Springer Verlag, 2006. – 217 p.
8. Лебедев Б.К., Лебедев В.Б. Оптимизация методом кристаллизации россыпи альтернатив // Известия ЮФУ. Технические науки. – 2013. – № 7 (144). – С. 11-17.
9. Лебедев Б.К., Лебедев В.Б. Метод кристаллизации россыпи альтернатив // Сб. научных трудов VII Международной научно-практической конференции “Интегрированные модели и мягкие вычисления в искусственном интеллекте”. Т. 2. – М.: Изд-во Физматлит, 2013. – С. 903-912.
10. Аюпов А.Б., Марченко А.М. Метод оптимизации трассируемости в аналитическом алгоритме размещения // Информационные технологии. – 2008. – № 2. – С. 12-17.
11. Норенков И.П. Основы автоматизированного проектирования: учебник. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2006. – 336 с.
12. Sherwani N.A. Algorithms for VLSI Physical Design Automation // Third Edition, Kluwer Academic Publisher. – USA: 2013. – 572 p.
13. Alpert C.J. Handbook of Algorithms for Physical design Automation // Auerbach Publications Taylor & Francis Group. – USA, 2009. – 349 p.
14. Лебедев Б.К., Лебедев В.Б., Лебедев О.Б. Методы, модели и алгоритмы размещения: монография. – Ростов-на-Дону: Изд-во ЮФУ, 2015. – 150 с.
15. Cong J., Romesis M., Xie M. UCLA Optimality Study Project. – <http://cadlab.cs.ucla.edu/~pubbench>. 2004.
16. Cong J., Romesis M., Xie M. Optimality, Scalability and Stability Study of Partitioning and Placement Algorithms // Proc. of the International Symposium on Physical Design. – Monterey, CA, 2003. – P. 88-94.

17. Zaporozhets D.Y., Zaruba D.V., Kureichik V.V. Hierarchical approach for VLSI components placement // *Advances in Intelligent Systems and Computing*. – 2015. – P. 79-87.
18. Kureichik V.I., Kureichik V., Leschanov D., Zaruba D. Hybrid approach for VLSI fragments placement // *Advances in Intelligent Systems and Computing*. – 2018. – P. 349-358.
19. IBM-PLACE 2.0 benchmark suits. – <http://er.cs.ucla.edu/benchmarks/-ibm-place2/bookshelf/ibm-place2-all-bookshelf-nopad.tar.gz>.
20. Adya S.N. ISPD02 IBM-MS Mixed-size Placement Benchmarks. – <http://vlsicad.eecs.umich.edu/BK/ISPD02bench/>.
21. Roy J.A., Papa D.A., Markov I.L. Capo: Congestion-Driven Placement for Standard-cell and RTL Netlists with Incremental Capability // In: *Modern Circuit Placement. Series on Integrated Circuits and Systems*. Springer, Boston, MA Springer Science + Business Media, LLC, 2007. – P. 123-146.
22. Adya S.N. Consistent placement of macro-blocks using floor planning and standard-cell placement // In *Proc. Intl. Symp. on Physical Design*. – 2002. – P. 12-17.
23. Wang M., Yang X., Sarrafzadeh M. Dragon 2000: Standard-cell Placement Tool for Large Industry Circuits. – *ICCAD*, 2000. – P. 260-263.

REFERENCES

1. Karpenko A.P. *Sovremennyye algoritmy poiskovoy optimizatsii. Algoritmy, vdokhnovlennyye prirodoy: ucheb. posobie [Modern search engine optimization algorithms. Algorithms inspired by nature: textbook]*. Moscow: Izd-vo MGTU im. N.E. Bauman, 2014, 448 p.
2. Lebedev O.B. *Modeli adaptivnogo povedeniya murav'inoy kolonii v zadachakh proektirovaniya [Models of adaptive behavior of an ant colony in design problems]*. Taganrog. Izd-vo YuFU, 2013, 199 p.
3. Lebedev B.K., Lebedev O.B., Lebedinskiy A.E. *Gibridnyy bioinspirirovannyy algoritm razmeshcheniya bazovykh standartnykh biblioteknykh elementov pri proektirovanii topologii poluzakaznoy SBIS [Hybrid bioinspired algorithm for placing basic standard library elements in the design of the topology of a semi-custom VLSI]*, *Izvestiya YuFU. Tekhnicheskie nauki [Izvestiya SFedU. Engineering Sciences]*, 2019, No. 3 (205), pp. 97-110.
4. Kureychik V.I., Kureychik V.V. *Integrirovannyy algoritm razmeshcheniya fragmentov CBIS [Integrated algorithm for placement of VLSI fragments]*, *Izvestiya YuFU. Tekhnicheskie nauki [Izvestiya SFedU. Engineering Sciences]*, 2014, No. 7 (156), pp. 84-91.
5. Lebedev B.K., Lebedev O.B., Lebedev V.B. *Gibridizatsiya roevogo intellekta i geneticheskoy evolyutsii na primere razmeshcheniya [Hybridization of swarm intelligence and genetic evolution on the example of placement]*, *Programmnyye produkty, sistemy i algoritmy [Software products, systems and algorithms]*, 2017, No. 4.
6. Caldwell A.E. *Can Recursive Bisection Alone Produce Routable Placements*, *DAC*, 2000, pp. 477-482.
7. Mourelle M. *Swarm intelligent systems*. Berlin: Heidelberg: Springer Verlag, 2006, 217 p.
8. Lebedev B.K., Lebedev V.B. *Optimizatsiya metodom kristallizatsii rossypi al'ternativ [Optimization by the method of crystallization of alternatives field]*, *Izvestiya YuFU. Tekhnicheskie nauki [Izvestiya SFedU. Engineering Sciences]*, 2013, No. 7 (144), pp. 11-17.
9. Lebedev B.K., Lebedev V.B. *Metod kristallizatsii rossypi al'ternativ [The method of crystallization of a placer of alternatives]*, *Sb. nauchnykh trudov VII Mezhdunarodnoy nauchno-prakticheskoy konferentsii "Integrirovannyye modeli i myagkie vychisleniya v iskusstvennom intellekte"* [Collection of scientific papers of the VII International scientific-practical conference "Integrated models and soft computing in artificial intelligence"]. Vol. 2. M.: Izd-vo Fizmatlit, 2013, pp. 903-912.
10. Ayupov A.B., Marchenko A.M. *Metod optimizatsii trassiruемости v analiticheskom algoritme razmeshcheniya [Traceability optimization method in the analytical placement algorithm]*, *Informatsionnyye tekhnologii [Information technologies]*, 2008, No. 2, pp. 12-17.
11. Norenkov I.P. *Osnovy avtomatizirovannogo proektirovaniya: uchebnik [Basics of computer-aided design: textbook]*. M.: Izd-vo MGTU im. N.E. Bauman, 2006, 336 p.
12. Sherwani N.A. *Algorithms for VLSI Physical Design Automation, Third Edition*, *Kluwer Academic Publisher*. USA: 2013, 572 p.

13. *Alpert C.J.* Handbook of Algorithms for Physical design Automation, *Auerbach Publications Taylor & Francis Group*. USA, 2009, 349 p.
14. *Lebedev B.K., Lebedev V.B., Lebedev O.B.* Metody, modeli i algoritmy razmeshcheniya: monografiya [Placement methods, models and algorithms: monograph]. Rostov-on-Don: Izd-vo YuFU, 2015, 150 p.
15. *Cong J., Romesis M., Xie M.* UCLA Optimality Study Project. Available at: <http://cadlab.cs.ucla.edu/~pubbench>. 2004.
16. *Cong J., Romesis M., Xie M.* Optimality, Scalability and Stability Study of Partitioning and Placement Algorithms, *Proc. of the International Symposium on Physical Design*. Monterey, CA, 2003, pp. 88-94.
17. *Zaporozhets D.Y., Zaruba D.V., Kureichik V.V.* Hierarchical approach for VLSI components placement, *Advances in Intelligent Systems and Computing*, 2015, pp. 79-87.
18. *Kureichik V.I., Kureichik V., Leschanov D., Zaruba D.* Hybrid approach for VLSI fragments placement, *Advances in Intelligent Systems and Computing*, 2018, pp. 349-358.
19. IBM-PLACE 2.0 benchmark suits. Available at: <http://er.cs.ucla.edu/benchmarks/-ibm-place2/bookshelf/ibm-place2-all-bookshelf-nopad.tar.gz>.
20. *Adya S.N.* ISPD02 IBM-MS Mixed-size Placement Benchmarks. Available at: <http://vlsicad.eecs.umich.edu/BK/ISPD02bench/>.
21. *Roy J.A., Papa D.A., Markov I.L.* Capo: Congestion-Driven Placement for Standard-cell and RTL Netlists with Incremental Capability, *In: Modern Circuit Placement. Series on Integrated Circuits and Systems*. Springer, Boston, MA Springer Science + Business Media, LLC, 2007, pp. 123-146.
22. *Adya S.N.* Consistent placement of macro-blocks using floor planning and standard-cell placement, *In Proc. Intl. Symp. on Physical Design*, 2002, pp. 12-17.
23. *Wang M., Yang X., Sarrafzadeh M.* Dragon 2000: Standard-cell Placement Tool for Large Industry Circuits. ICCAD, 2000, pp. 260-263.

Статью рекомендовал к опубликованию д.т.н., профессор А.Г. Коробейников.

Лебедев Борис Константинович – Южный федеральный университет; e-mail: lebedev.b.k@gmail.com; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 89282897933; кафедра систем автоматизированного проектирования; профессор.

Лебедев Олег Борисович – e-mail: lebedev.ob@mail.ru; ттел.: 89085135512; кафедра систем автоматизированного проектирования; доцент.

Лебедев Владимир Борисович – Московский государственный технический университет имени Н.Э. Баумана; e-mail: lebedev.vlad.bor@mail.ru; 105005, г. Москва, ул. Бауманская 2-я, д. 5, стр. 1; тел.: 89287775005; Научно-производственное объединение «Новые технологии»; с.н.с.

Lebedev Boris Konstantinovich – Southern Federal University; e-mail: lebedev.b.k@gmail.com; 44, Nekrasovsky, Taganrog, 347928, Russia; phone: +79282897933; the department of computer aided design; professor.

Lebedev Oleg Borisovich – e-mail: lebedev.ob@mail.ru; phone: +79085135512; the department of computer aided design; associate professor.

Lebedev Vladimir Borisovich – Moscow State Technical University named after N.E. Bauman; e-mail: lebedev.vlad.bor@mail.ru; 105005, Moscow, st. Baumanskaya 2-nd, 5, build. 1; phone: +79287775005; Research and Production Association "New Technologies"; senior researcher.

Н.О. Васильев, П.И. Фролова, Г.А. Иванова, А.Н. Щелоков

**МЕТОДЫ ЛОГИЧЕСКОГО РЕСИНТЕЗА ДЛЯ ТОПОЛОГИЧЕСКОГО
ПРОЕКТИРОВАНИЯ МИКРОЭЛЕКТРОННЫХ СХЕМ***

С уменьшением технологических норм возрастает число правил проектирования. Для сокращения временных затрат на проверку правил проектирования для технологий 22 нм и ниже переходят к использованию регулярных структур в нижних слоях топологии. При проектировании схем на основе регулярного шаблона становится возможным совмещение логического и топологического этапов проектирования. Данная задача также актуальна для проектирования схем на ПЛИС. В данной работе рассматривается метод структурной оптимизации логических схем на этапе топологического проектирования. Метод адаптирован для применения в маршруте проектирования схем с регулярными структурами в нижних слоях топологии, а также для ресинтеза технологических отображений на ПЛИС. Для схем с применением регулярных структур предлагается метод логического синтеза в базисе элементов, для которых построены компактные топологические шаблоны. Это позволяет упростить этап топологического проектирования, а также ведет к дополнительному снижению площади проектируемого устройства. Оптимизация логических схем для ПЛИС проводится при помощи алгоритма моделирования отжига, производящего логические операции над специальной графовой моделью, учитывающей особенности ПЛИС. Учет особенностей различных технологий в предлагаемом методе позволяет добиться хороших результатов по необходимым параметрам, в частности по занимаемой проектируемой схемой площади.

Ресинтез; транзисторы с вертикальным затвором (FinFET); КМОП технология; ПЛИС.

N.O. Vasilyev, P.I. Frolova, G.A. Ivanova, A.N. Schelokov

**LOGIC RESYNTHESIS METHODS FOR LAYOUT DESIGN
OF MICROELECTRONIC CIRCUITS**

As the size of electronic components decreases, the number of design rules increases. To reduce design rules checking runtime for 22 nm and below technologies, regular structures are used in the lower layers of the layout. When designing circuits based on a regular template, it becomes possible to combine the logical and layout design stages. This task is also relevant for designing circuits on FPGAs. This paper discusses a method for structural optimization of logic circuits at the stage of layout design. The method is adapted for use in the design route of circuits with regular structures in the lower layers of the layout, as well as for resynthesis of technology mappings on FPGAs. When working with circuits with regular structures, logical synthesis is used in the basis of elements for which compact layout templates are built. This approach simplifies the layout design stage, and also leads to an additional reduction in the area of the designed device. Optimization of logic circuits for FPGAs is carried out using a simulated annealing algorithm that performs logic operations on a special graph model that takes into account the features of the FPGA. Taking into account the features of various technologies in the proposed method allows achieving good results in terms of such parameters as the area occupied by the circuit.

Resynthesis; vertical gate transistors (FinFET); CMOS technology; FPGA.

Введение. Перспективными направлениями развития отечественной микроэлектронной промышленности в данный момент времени являются программируемые логические интегральные схемы (ПЛИС) и схемы с применением транзисторов с вертикальным затвором (FinFET).

* Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 20-07-00425, 18-07-00621.

Существует несколько причин, объясняющих возросший интерес к развитию таких технологий, как FinFET. Одна из основных причин приближение существующих КМОП-технологии к пределу своего масштабирования, что выражается в существенной деградации электрических параметров транзисторов при дальнейшем уменьшении технологических норм производства [1, 2]. Применение FinFET технологии позволяет снизить негативные эффекты, возникающие при уменьшении размеров транзисторов [3].

Популярности ПЛИС также способствовали несколько факторов. Одним из таких факторов является то, что уменьшение технологических норм ведет к увеличению стоимости производства интегральных схем. Из-за этого стоимость ошибки проектировщика многократно возросла за последнее время. В связи с этим среди компаний, имеющих отношение к микроэлектронной области, возник повышенный интерес к использованию ПЛИС для прототипирования разрабатываемых устройств [4, 5]. Другим фактором, объясняющим популярность ПЛИС, является то, что при производстве схем небольшими партиями стоимость единицы продукции получается довольно высокой. Компании, потребность которых в интегральных схемах ограничена небольшими объемами, предпочитают использовать ПЛИС.

Однако для работы с описанными технологиями необходимы специализированные методы проектирования. В частности, необходима разработка методов структурной оптимизации схем на этапах логического и топологического синтеза. В данной работе предлагаются методы оптимизации цифровых схем с применением FinFET, а также для схем, синтезируемых на ПЛИС.

1. Структурная оптимизация комбинационных схем с применением регулярных структур. Применение FinFET позволяет увеличить эффективную ширину затвора полевых транзисторов, которая будет определяться как $2 \cdot n \cdot h$, где n – число плавников, h – высота плавника, по сравнению со схемами, применяющими планарные структуры [6]. Таким образом, более широкие транзисторы с более высокими токами получаются при использовании нескольких плавников. Шаг плавника (p) – минимальное расстояние между соседними плавниками, которое определяется разрешающей способностью литографии на конкретном технологическом узле. Используя спейсерную литографию, p может уменьшиться до значения, равного половине разрешающей способности литографии [7]. Однако проектирование схем с применением данной технологии значительно усложняется. Для уменьшения количества проверок соблюдения правил проектирования на этапе топологического проектирования предъявляются требования к регулярности нижних слоев в транзисторных структурах. При этом, требования к регулярности возрастают с уменьшением технологических норм.

Для удовлетворения требований регулярности транзисторных структур в предлагаемом методе были разработаны компактные топологические шаблоны логических элементов (Рис. 1) [8]. Компактность достигается за счет выбора направления ориентации каждого из транзисторов. Под направлением ориентации транзистора понимается взаимное расположение стока и истока в топологии транзистора. Если в топологии исток находится левее стока, то ориентация считается нормальной, иначе – зеркальной. Выбор ориентации позволяет уменьшить необходимое количество изолирующих затворов путем объединения областей диффузии [9]. Для этого необходимо, чтобы контакты, расположенные слева и справа от изолирующего затвора, были подключены к одному узлу в цепи.

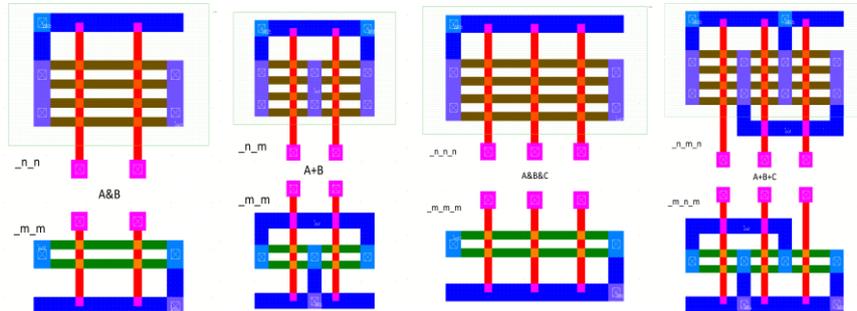


Рис. 1. Шаблоны функций $A \& B$, $A + B$, $A \& B \& C$ и $A + B + C$

Для представления логических функций используется графовая модель на основе графа вложенности последовательно-параллельных структур (SP-граф) [10]. SP-граф – это направленный ациклический граф $G = (V, E)$ [11], вершины которого из множества V соответствуют одной из следующих логических функций: конъюнкция, дизъюнкция, инверсия, конъюнкция с отрицанием, дизъюнкция с отрицанием; а ребра из множества E ссылаются на аргументы соответствующей функции (Рис. 2).

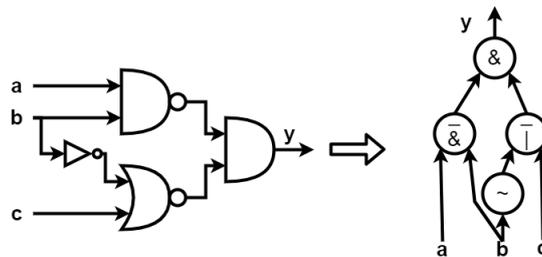


Рис. 2. SP-граф для некоторой схемы

На этапе ресинтеза производится оптимизация логической структуры схемы. Для этого в цикле оптимизации на каждой итерации выбирается подграф, над которым производятся логические преобразования с целью оптимизации схемы по необходимым параметрам. В качестве таких параметров обычно выступают занимаемая схемой площадь, быстродействие и потребляемая мощность.

При логической оптимизации синтез производится в базис элементов, для которых доступны топологические шаблоны регулярных структур. При этом, если контакты топологических шаблонов разных элементов расположены слева и справа от места соединения шаблонов подключены к одному и тому же сигналу, то данные шаблоны могут быть соединены без образования изолирующего затвора. Данный метод позволяет дополнительно снизить площадь итогового устройства.

Для сокращения количества изолирующих затворов при соединении шаблонов по аналогии с ориентацией транзистора используется понятие ориентации шаблона. Ориентация шаблонов позволяет выбрать наиболее удачный способ соединения с точки зрения занимаемой площади нескольких шаблонов.

Набор шаблонов для функций от 2 до 4 переменных позволяет эффективно отобразить в него все логические функции, встречающиеся в комбинационных схемах, с возможностью выбора наиболее оптимальной конфигурации. Использование в шаблонах большого числа переменных значительно усложняет задачу, а также приводит к формированию длинных цепочек последовательно соединенных транзисторов, которые обладают низким быстродействием [12].

2. Логическая оптимизация схем для ПЛИС. При проектировании на ПЛИС необходимо учитывать то, что базовый логический элемент для реализации задач комбинационной логики на ПЛИС выполнен в виде Look-Up Table (LUT) [13]. LUT представляет из себя логический элемент, логика работы которого может быть настроена с помощью задания конфигурационных битов. Фактически это означает, что LUT может реализовывать любую Булеву функцию, количество переменных которой не превышает количества входов в LUT в целевой архитектуре ПЛИС.

При логической оптимизации схемы для LUT обычно применяются следующие упрощения оптимизируемых параметров:

1. Для оценки площади схемы рассматривается количество необходимых для реализации данной схемы LUT [14];
2. Для быстрой оценки быстродействия используется наибольшее количество элементов на пути от первичных входов до первичных выходов [15, 16];
3. Оценка трассируемости схемы на данном этапе может производиться с помощью количества межсоединений в схеме [17].

При оптимизации схем на ПЛИС для представления логических элементов используется направленный ациклический граф $G = (V, E)$, вершины которого из множества V выполняют любую Булеву функцию и ссылаются на логические элементы схемы, а ребра множества E ссылаются на межсоединения между логическими элементами. Достоинством подобной графовой модели является то, что каждая вершина может быть отображена на один LUT.

Над этим графом определены следующие логические преобразования, которые изменяют составляющие структуру проектируемой схемы логические элементы:

1. Декомпозиция вершин разбивает одну вершину на несколько, в которых используется меньше входов, чем в исходной вершине (Рис. 3) [18];
2. Слияние вершин объединяет две вершины в одну (Рис. 3);
3. Переупорядочивание вершин меняет порядок двух вершин в графе (Рис. 4).

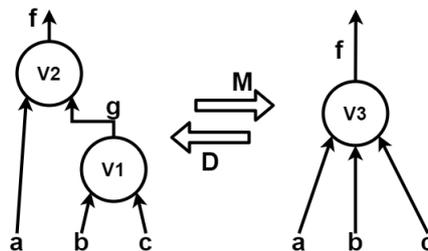


Рис. 3. Слияние (M) и декомпозиция (D) вершин графа

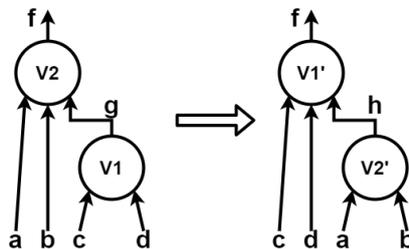


Рис. 4. Переупорядочивание вершин графа

Данные логические операции используются в цикле оптимизации, который производится так же, как и ранее описанный метод оптимизации комбинационных схем с применением регулярных структур, над окнами исходного графа [19]. Это

позволяет уменьшить пространство возможных логических состояний в каждом окне по сравнению с пространством логических состояний всего графа, что ведет к уменьшению вычислительной сложности задачи. Предложенный метод оптимизирует логическую схему по необходимым параметрам, выраженным с помощью целевой функции, в цикле алгоритма моделирования отжига [20]. Алгоритм моделирования отжига по сравнению с жадными алгоритмами стремится к достижению глобального минимума целевой функции путем принятия негативных решений в некоторых случаях.

Заключение. В данной работе были предложены методы оптимизации комбинационных схем для различных технологий микроэлектронной промышленности.

Применение логической оптимизации в сочетании с применением компактных топологических шаблонов для регулярных структур позволяет добиться дополнительного снижения занимаемой схемой площади.

В случае оптимизации логических схем для ПЛИС применение алгоритма моделирования отжига в цикле оптимизации ведет к дополнительной оптимизации схемы по необходимым параметрам, определенным в целевой функции, путем стремления к глобальному минимуму заданной функции.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Hu C.* Gate oxide scaling limits and projection // International Electron Devices Meeting. Technical Digest. – IEEE, 1996. – P. 319-322.
2. *Yeo Y.C., King T.J., Hu C.* MOSFET gate leakage modeling and selection guide for alternative gate dielectrics based on leakage considerations // IEEE Transactions on Electron Devices. – 2003. – Vol. 50, No. 4. – P. 1027-1035.
3. *Agostinelli M.* Leakage-delay tradeoff in FinFET logic circuits: A comparative analysis with bulk technology // IEEE Transactions. Very Large Scale Integrated (VLSI) Systems. – 2010. – P. 232-245.
4. *Najibi M. et al.* Prototyping globally asynchronous locally synchronous circuits on commercial synchronous FPGAs // 16th IEEE International Workshop on Rapid System Prototyping (RSP'05). – IEEE, 2005. – P. 63-69.
5. *Юрлин С.В., Бычков И.Н.* Прототипирование на основе ПЛИС для верификации многоядерных микропроцессоров // Проблемы разработки перспективных микро- и нанoeлектронных систем (МЭС). – 2014. – №. 4. – С. 45-50.
6. *Bhattacharya D., Jha N.K.* FinFETs: From Devices to Architectures // Advances in Electronics. – 2014. – 21 p.
7. *Манукян А.А.* Исследование и разработка методов логико-топологического синтеза библиотечных элементов и блоков для КМОП технологий с трехмерным затвором транзистора: дисс. ... канд. техн. наук. – 2015. – С. 31-79.
8. *Гаврилов С.В., Иванова Г.А., Манукян А.А.* Методы проектирования заказных сложнофункциональных блоков в базе элементов с регулярной топологической структурой в слоях поликремния и диффузии // Проблемы разработки перспективных микро- и нанoeлектронных систем (МЭС). – 2014. – №. 1. – С. 161-166.
9. *Рыжова Д.И., Васильев Н.О., Жукова Т.Д.* Алгоритм межвентильного ресинтеза на транзисторном уровне для автоматизированного проектирования микроэлектронных схем // Проблемы разработки перспективных микро- и нанoeлектронных систем (МЭС). – 2018. – №. 1. – С. 193-198.
10. *Гаврилов С.В., Карева Е.С., Рыжова Д.И.* Алгоритмы логико-топологического синтеза библиотечных элементов и блоков с регулярной структурой для технологических норм проектирования 32 нм // Известия высших учебных заведений. – 2017. – Т. 22, №. 4. – P. 42017369.
11. *Bryant R.E.* Graph-based algorithms for boolean function manipulation // Computers, IEEE Transactions on. – 1986. – Vol. 100, No. 8. – P. 677-691.
12. *Талалай М.С.* Метод логико-топологического синтеза нанометровых КМОП схем на основе транзисторных шаблонов: дисс. ... канд. техн. наук. – 2012.

13. Betz V., Rose J. How much logic should go in an FPGA logic block // IEEE Design & Test of Computers. – 1998. – Vol. 15, No. 1. – P. 10-15.
14. Jones D., Lewis D.M. A time-multiplexed FPGA architecture for logic emulation // Proceedings of the IEEE 1995 Custom Integrated Circuits Conference. – IEEE, 1995. – P. 495-498.
15. Ling A., Singh D.P., Brown S.D. FPGA technology mapping: a study of optimality // Proceedings. 42nd Design Automation Conference, 2005. – IEEE, 2005. – P. 427-432.
16. Cong J., Ding Y. On area/depth trade-off in LUT-based FPGA technology mapping // IEEE Transactions on Very Large Scale Integration (VLSI) Systems. – 1994. – Vol. 2, No. 2. – P. 137-148.
17. Jang S. et al. WireMap: FPGA technology mapping for improved routability // Proceedings of the 16th international ACM/SIGDA symposium on Field programmable gate arrays. – 2008. – P. 47-55.
18. Lehman E. et al. Logic decomposition during technology mapping // IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. – 1997. – Vol. 16, No. 8. – P. 813-834.
19. Vasilyev N.O., Tiunov I.V., Ryzhova D.I. Resynthesis for FPGA During Technology Mapping Stage // 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). – IEEE, 2019. – P. 1644-1647.
20. Vasilyev N.O., Tiunov I.V., Ryzhova D.I. The Simulated Annealing Based Logical Resynthesis Method for LUT-based FPGAs // 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). – IEEE, 2020. – P. 1892-1894.

REFERENCES

1. Hu C. Gate oxide scaling limits and projection, *International Electron Devices Meeting, Technical Digest*. IEEE, 1996, pp. 319-322.
2. Yeo Y.C., King T.J., Hu C. MOSFET gate leakage modeling and selection guide for alternative gate dielectrics based on leakage considerations, *IEEE Transactions on Electron Devices*, 2003, Vol. 50, No. 4, pp. 1027-1035.
3. Agostinelli M. Leakage-delay tradeoff in FinFET logic circuits: A comparative analysis with bulk technology, *IEEE Transactions. Very Large Scale Integrated (VLSI) Systems*, 2010, pp. 232-245.
4. Najibi M. et al. Prototyping globally asynchronous locally synchronous circuits on commercial synchronous FPGAs, *16th IEEE International Workshop on Rapid System Prototyping (RSP'05)*. IEEE, 2005, pp. 63-69.
5. Yurlin S.V., Bychkov I.N. Prototipirovanie na osnove PLIS dlya verifikatsii mnogoyadernykh mikroprotessorov [FPGA-based prototyping for multi-core microprocessor verification], *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem (MES)* [Problems of development of advanced micro-and nanoelectronic systems (MES)], 2014, No. 4, pp. 45-50.
6. Bhattacharya D., Jha N.K. FinFETs: From Devices to Architectures, *Advances in Electronics*, 2014, 21 p.
7. Manukyan A.A. Issledovanie i razrabotka metodov logiko-topologicheskogo sinteza biblioteknykh elementov i blokov dlya KMOP tekhnologiy s trekhmernym zatvorom tranzistora: diss. ... kand. tekhn. nauk [Research and development of methods for logical-topological synthesis of library elements and blocks for CMOS technologies with a three-dimensional transistor gate: cand. of eng. sc. diss.], 2015, pp. 31-79.
8. Gavrilov S.V., Ivanova G.A., Manukyan A.A. Metody proektirovaniya zakaznykh slozhno-funktional'nykh blokov v bazise elementov s regulyarnoy topologicheskoy strukturoy v sloyakh polikremniya i diffuzii [Methods for designing custom complex-functional blocks in the basis of elements with a regular topological structure in polysilicon and diffusion layers], *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem (MES)* [Problems of development of advanced micro- and nanoelectronic systems (MES)], 2014, No. 1, pp. 161-166.
9. Ryzhova D.I., Vasil'ev N.O., Zhukova T.D. Algoritm mezhventil'nogo resinteza na tranzistornom urovne dlya avtomatizirovannogo proektirovaniya mikroelektronnykh skhem [Transistor-level interventional resynthesis algorithm for computer-aided design of microelectronic circuits], *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem (MES)* [Problems of development of advanced micro-and nanoelectronic systems (MES)], 2018, No. 1, pp. 193-198.

10. Gavrilov S.V., Kareva E.S., Ryzhova D.I. Algoritmy logiko-topologicheskogo sinteza biblioteknykh elementov i blokov s regulyarnoy strukturoy dlya tekhnologicheskikh norm proektirovaniya 32 nm [Algorithms for logical-topological synthesis of library elements and blocks with a regular structure for technological design standards of 32 nm], *Izvestiya vysshikh uchebnykh zavedeniy* [Proceedings of higher educational institutions], 2017, Vol. 22, No. 4, pp. 42017369.
11. Bryant R.E. Graph-based algorithms for boolean function manipulation, *Computers, IEEE Transactions on*, 1986, Vol. 100, No. 8, pp. 677-691.
12. Talalay M.S. Metod logiko-topologicheskogo sinteza nanometrovykh KMOP skhem na osnove tranzistornykh shablonov: diss. ... kand. tekhn. nauk [Method of logical-topological synthesis of nanometer CMOS circuits based on transistor templates: cand. of eng. sc. diss.], 2012.
13. Betz V., Rose J. How much logic should go in an FPGA logic block, *IEEE Design & Test of Computers*, 1998, Vol. 15, No. 1, pp. 10-15.
14. Jones D., Lewis D.M. A time-multiplexed FPGA architecture for logic emulation, *Proceedings of the IEEE 1995 Custom Integrated Circuits Conference*. IEEE, 1995, pp. 495-498.
15. Ling A., Singh D.P., Brown S.D. FPGA technology mapping: a study of optimality, *Proceedings. 42nd Design Automation Conference, 2005*. IEEE, 2005, pp. 427-432.
16. Cong J., Ding Y. On area/depth trade-off in LUT-based FPGA technology mapping, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 1994, Vol. 2, No. 2, pp. 137-148.
17. Jang S. et al. WireMap: FPGA technology mapping for improved routability, *Proceedings of the 16th international ACM/SIGDA symposium on Field programmable gate arrays*, 2008, pp. 47-55.
18. Lehman E. et al. Logic decomposition during technology mapping, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 1997, Vol. 16, No. 8, pp. 813-834.
19. Vasilyev N.O., Tiunov I.V., Ryzhova D.I. Resynthesis for FPGA During Technology Mapping Stage, *2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. IEEE, 2019, pp. 1644-1647.
20. Vasilyev N.O., Tiunov I.V., Ryzhova D.I. The Simulated Annealing Based Logical Resynthesis Method for LUT-based FPGAs, *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. IEEE, 2020, pp. 1892-1894.

Статью рекомендовал к опубликованию к.т.н. В.А. Иванов.

Васильев Николай Олегович – Институт проблем проектирования в микроэлектронике РАН (ИППМ РАН); e-mail: vasilyev_n@ippm.ru; 124365, Москва, Зеленоград, ул. Советская, 3; инженер-исследователь.

Фролова Пелагея Ильинична – e-mail: frolova_p@ippm.ru; инженер-исследователь.

Иванова Галина Александровна – e-mail: ivanova_g@ippm.ru; с.н.с.; к.т.н.

Щелоков Альберт Николаевич – e-mail: schan@ippm.ru; зам. директора; к.ф.-м.н.

Vasilyev Nikolay Olegovich – The Institute for Design Problems in Microelectronics (IPPM RAS); e-mail: vasilyev_n@ippm.ru; 3, Sovetskaya street, Zelenograd, Moscow, 124365, Russia; research engineer.

Frolova Pelageya Ilyinichna – e-mail: frolova_p@ippm.ru; research engineer.

Ivanova Galina Aleksandrovna – e-mail: ivanova_g@ippm.ru; senior researcher; cand. of eng. sc.

Schelokov Albert Nikolaevich – e-mail: schan@ippm.ru; deputy director; cand. of phys. and math. sc.

Л.А. Гладков, Н.В. Гладкова, М.Д. Ясир

ГИБРИДНЫЙ ПОДХОД К СОВМЕСТНОМУ РЕШЕНИЮ ЗАДАЧ РАЗМЕЩЕНИЯ И ТРАССИРОВКИ*

В статье предложен интегрированный подход к решению задач размещения и трассировки элементов схем электронной вычислительной аппаратуры. Подход основан на совместном решении задач размещения и трассировки с использованием нечетких генетических методов. Приведено описание рассматриваемой проблемы и выполнен краткий анализ существующих подходов к ее решению. В статье рассматриваются интегрированные подходы к решению оптимизационных задач автоматизированного проектирования схем цифровой электронно-вычислительной аппаратуры. Подчеркнута актуальность и важность разработки новых эффективных методов решения подобных задач. Отмечено, что важным направлением развития методов оптимизации является разработка гибридных методов и подходов, сочетающих достоинства различных методов вычислительного интеллекта. В статье описаны следующие основные моменты: структура предлагаемого алгоритма и его основные этапы; модифицированные генетические операторы кроссовера; предложены модели формирования текущей популяции; модифицированные эвристики, операторы и стратегии поиска оптимальных решений. Приведены результаты вычислительных экспериментов. Проведенные эксперименты подтверждают эффективность предложенного подхода. В заключении приводится краткий анализ полученных результатов.

Системы автоматизированного проектирования; задачи проектирования; эволюционные вычисления; нечеткие генетические алгоритмы; контроллер нечеткой логики; биоинспирированные алгоритмы; гибридные методы.

L.A. Gladkov, N.V. Gladkova, M.J. Yasir

HYBRID APPROACH THE JOINT SOLUTION OF PLACEMENT AND TRACING PROBLEMS

The article proposes an integrated approach to solving the problems of placing and tracing elements of circuits of electronic computing equipment. The approach is based on the joint solution of placement and tracing problems using fuzzy genetic methods. A description of the problem under consideration is given and a brief analysis of existing approaches to its solution is performed. The article discusses integrated approaches to solving optimization problems of computer-aided design of digital electronic computing equipment circuits. The urgency and importance of developing new effective methods for solving such problems is emphasized. It is noted that an important direction in the development of optimization methods is the development of hybrid methods and approaches that combine the advantages of various methods of computational intelligence. The article describes the following main points: the structure of the proposed algorithm and its main stages; modified genetic crossover operators; models for the formation of the current population are proposed; modified heuristics, operators and strategies for finding optimal solutions. The results of computational experiments are presented. The experiments carried out confirm the effectiveness of the proposed approach. In conclusion, a brief analysis of the results obtained is given.

Computer-aided design systems; design tasks; evolutionary computing; fuzzy genetic algorithms; fuzzy logic controller; bioinspired algorithms; hybrid methods.

Введение. Современные тенденции научно-технического развития включают постоянное усложнение систем управления и технологических процессов. Это стимулирует разработчиков создавать и внедрять новые инструменты для повышения эффективности систем проектирования.

* Исследование выполнено при поддержке РФФИ в рамках научного проекта № 18-07-01054.

Проектирование является важным шагом в цикле компьютерного проектирования цифровых устройств. Среди основных задач этапа проектирования – задачи разбиения, размещения и трассировки [1, 2]. В современных средствах автоматизированного проектирования успешно применяются эволюционные и генетические методы и подходы. Эволюционные методы особенно эффективны при решении комбинаторных задач большой размерности, когда достаточно найти одно или несколько квазиоптимальных решений, удовлетворяющих заданным ограничениям [3–5].

Одним из современных направлений в науке является создание интегрированных и гибридных моделей и методов. Гибридную систему можно рассматривать как систему, которая содержит две или более интегрированные гетерогенные подсистемы (разных видов), которые объединены совместными действиями (хотя эти подсистемы могут иметь различную природу) [6, 7].

Интеграция и гибридизация различных методов проектирования позволяют нам решать сложные проблемы, которые не могут быть решены каким-либо конкретным методом и технологией [8, 9].

В настоящее время также активно развивается научное направление по разработке моделей и методов гибридных принципов нечеткого управления и поисковых способностей биоинспирированных методов. Можно выделить два основных подхода к использованию таких гибридных методов [10–12]:

1. Использовать эволюционные алгоритмы для решения задач оптимизации и поиска в условиях нечеткой, неоднозначной или недостаточной информации об объекте, параметрах критериев решаемых задач [10–12].

2. Использовать методы, основанные на нечеткой логике, для моделирования структуры и операторов эволюционных алгоритмов, для управления и адаптации параметров эволюционных алгоритмов [13, 14].

Рассмотрим второй подход. Для контроля и настройки параметров эволюционного алгоритма мы предлагаем включить нечеткий логический контроллер (НЛК), который использует опыт и знания экспертов в рассматриваемой области и может динамически изменять параметры поиска [13, 14]. Система правил производства на основе рассуждений и знаний выполняет логическое заключение, которое превращается в управляющие действия после дефаззификации. Изменение параметров алгоритма приводит к изменению процесса поиска и получению новых результатов, которые преобразуются в нечеткие множества в блоке нечеткости Система правил производства на основе рассуждений и знаний выполняет логическое заключение, которое превращается в управляющие действия после дефаззификации. Изменение параметров алгоритма приводит к изменению процесса поиска и получению новых результатов, которые преобразуются в нечеткие множества в блоке нечеткости [15, 16]

Также используются следующие инструменты: нечеткие операторы и нечеткие правила для создания генетических операторов с различными свойствами; нечеткая логическая система управления параметрами генетического алгоритма в соответствии с принятыми критериями; нечеткие критерии останова процесса генетического поиска. Математический аппарат теории нечетких систем используется в данном случае для кодирования, выбора оптимальных параметров генетических алгоритмов, значений вероятностей генетических операторов, функции пригодности и критерия останова, а также создания нечетких генетических операторов [15, 16].

Целью данной статьи является разработка гибридного алгоритма для эффективного решения сложных многокритериальных задач проектирования, таких как размещение, трассировка. [10–16].

Постановка задачи. Интеграция различных направлений научных исследований, отсутствие четких границ проблемной области, и, наконец, использование соответствующего инструментария, позволяет рассматривать научное направление «вычислительный интеллект» в качестве актуальной междисциплинарной научной области.

Гибридные системы состоят из различных элементов (компонентов), объединенных в интересах достижения поставленных целей. Интеграция различных методов и технологий позволяет решать сложные задачи, которые невозможно решить на основе каких-либо отдельных методов или технологий. В гибридной архитектуре, объединяющей несколько парадигм, эффективность одного подхода может компенсировать слабость другого. Комбинируя различные подходы, можно обойти недостатки, присущие каждому из них в отдельности. Этот эффект принято называть синергетическим [4, 6].

В задаче размещения определяется оптимальная конфигурация позиций с точки зрения взаимного расположения элементов и связей между ними. При этом должны быть соблюдены заданные конструктивные и технологические ограничения. Условно проблему размещения можно разделить на две части: определение взаимного расположения элементов конструкции на рабочем поле и оценка возможной длины соединений между ними.

При таком подходе проблема размещения сводится к поиску оптимального положения элементов и внешних контактов. В ряде алгоритмов размещение элементов выполняется без учета их связей с внешними выводами, поэтому элементы со ссылками на внешние выводы могут находиться на значительном расстоянии от них, что затрудняет отслеживание последующих соединений. Исходными данными для задачи размещения являются: список элементов схемы с учетом их соединений, параметры метрики и топологические свойства поля.

Основная сложность в постановке задачи определения местоположения заключается в выборе целевой функции. Это связано с тем, что одной из основных целей размещения является создание наилучших условий для дальнейшего решения задачи трассировки, что невозможно точно установить без выполнения самой трассировки. Таким образом, оптимальный результат может быть достигнут только с помощью одновременного решения задач размещения, выбора приоритетов прокладки соединений и оценочной трассировки, что практически невозможно из-за огромных затрат компьютерного времени.

Поэтому все алгоритмы, используемые в настоящее время с использованием временных критериев, лишь косвенно помогают решить основную проблему: получения оптимальной конфигурации соединений.

Для оценки качества решения задач размещения и трассировки используются следующие критерии [2]:

- ◆ общая длина взвешенных соединений;
- ◆ количество соединений превышающих заданную длину;
- ◆ количество пересечений проводников;
- ◆ количество соединений между элементами, которые находятся в соседних позициях или в позициях которые указывает разработчик.

Критерий, наиболее распространенный в алгоритмах размещения, является первым. Это объясняется следующими факторами: уменьшение длины связи может улучшить электрические характеристики цепи; упрощает процесс отслеживания; и это относительно просто реализовать.

Отслеживание соединений, как правило, является завершающей стадией конструирования электронно-вычислительного оборудования. Задача состоит в определении линий, соединяющих эквипотенциальные контактные элементы, то есть проектируемое устройство.

Исходными данными задачи являются: множество конструктивных элементов $E = \{e_i \mid i = 1, \dots, N\}$, множество соединяющих их цепей $C = \{c_k \mid k = 1, K\}$ и множество фиксированных положений для размещения элементов $S = \{s_j \mid j = 1, M\}$, where $M \geq N$. Нам нужно найти отображение множества E на множество T , при котором достигается экстремум целевой функции F .

Пусть задан набор ячеек $E = (e_1, \dots, e_n)$, при чем ячейка e_1 имеет длину в горизонтальном направлении a_1 . Оптимизация выполняется в два этапа: построение исходного размещения и его улучшение.

Первая ячейка размещается в строке $e_x \in E$, для которой число цепей, объединяющих ее с другими ячейками из E , является максимальным. Положительные ячейки e_x не фиксированы. На каждом из следующих $n-1$ шагов у нас есть набор размещенных ячеек E_h , набор неразмещенных ячеек E_k и набор возможных положений $S_h = (s_1, \dots, s_k)$ размещения ячеек в ряду $e_i \in E_h$ (рис. 1):

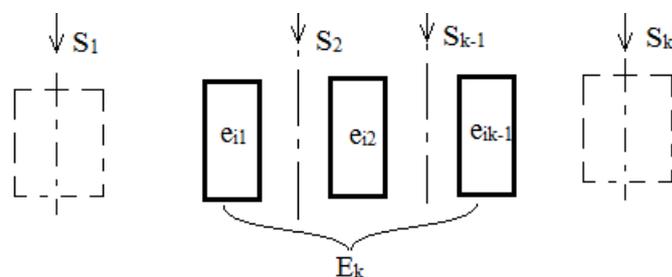


Рис. 1. Позиции для размещения элементов

Есть два положения s_1 и s_h , слева и справа от размещенных элементов и промежуточных положений s_2, \dots, s_{k-1} . Обозначим через V_{ik} множество цепей, соединяющих ячейку e_i и ячейку из E_h . На k -м шаге общая длина соединений для цепей V_{ik} рассчитывается при условии, что ячейка e_i находится в положении s_j ($j = 1, \dots, k$):

$$L_{ij} = \sum_{u_i \in V_{ik}} (x_{i_{max}} - x_{i_{min}})$$

где $x_{i_{max}}$ and $x_{i_{min}}$ координаты крайнего правого и левого выходов цепи v_i с линейным расположением ячеек $E_k \cup e_i$. Координаты рассчитываются с учетом размера ячеек и изменения положения ячеек в случае размещения ячеек в промежуточном положении.

Следующий шаг алгоритма состоит в выборе ячейки $e_i^* \in \overline{E_k}$ в позиции S_j^* , для некоторых:

$$L_{i^*j^*} = \max_{e_i \in B_k} \min_{s_j \in S_k} L_{ij}$$

После размещения всех ячеек множества E мы фиксируем их координаты в ряду.

Этот алгоритм используется совместно с алгоритмами разбиения ячеек на строки или блоки. Таким образом, создание строк или блоков и размещение их в ячейках будет происходить одновременно [8].

Для решения задачи размещения существует множество цепей, которые делят множество B на выходы элементов на непересекающиеся подмножества B_i , которые $B = \{B_i \mid i = 1, M\}$, $B_i = \{b_i, k = 1 / k_i\}$, где M – количество цепей; k_i – количество контактов, соединяемых i -й цепью. Пространство, представленное множеством $E = \{E_r \mid r = 1, R\}$.

В процессе улучшения исходного размещения мы будем использовать оценку качества полученной трассировки соединений. Эта оценка включает в себя следующие два критерия, такие как количество дорожек и общая неспособность связаться с контактами. На основе оценки качества трассировки разработанная программа может улучшить результат по этапу размещения или отобразить результат на экране.

Целевая функция Q является нормализованным аддитивным критерием:

$$Q = k_1q_1 + k_2q_2,$$

где k_1, k_2 – весовые коэффициенты используемых критериев; q_1 – число непроведенных соединений, q_2 – общая длина соединений.

В процессе проектирования, как правило, используются оценочные (приблизительные) значения данных. При этом их истинные значения этих данных становятся известны только после последующих процедур. Это обуславливает итеративный процесс проектирования с возвратами с последующих этапов предыдущих, конечно, значительно увеличивает стоимость проектирования. Поэтому поиск методов уменьшения количества итераций в цикле проектирования продолжается и задача разработки комплексных методов решения задач размещения и маршрутизации является актуальной. Такие методы должны позволять выполнять эти задачи за один цикл с учетом взаимного влияния ограничения и текущих результатов.

Для решения многокритериальной задачи размещения элементов с учетом последующей трассируемости предлагается использовать нечеткий генетический алгоритм [5, 9]. Для этого необходимо решить следующие задачи:

- ◆ разработать метод кодирования хромосомы с использованием действительных чисел;
- ◆ разрабатывать новые поколения поисковых стратегий на основе минимального разрыва между поколениями моделей и синтеза поколений;
- ◆ разработать новый оператор кроссовера: «восхождение на гору», основанный на методе градиентных методов, а также модифицированные арифметические операторы и линейный кроссовер;
- ◆ определить граничные условия и целевую функцию задачи;
- ◆ определить структуру генетического алгоритма с помощью реальных решений кодирования;
- ◆ разработать программное обеспечение для решения проблемы;
- ◆ проводить вычислительные эксперименты и анализировать полученные результаты.

Описание алгоритма. На основе анализа существующих методов и алгоритмов решения задач размещения и трассировки была предложена следующая процедура совместного решения этих задач.

1. Ввод исходных данных (количество элементов каждого типа, количество цепей, количество контактов и т.д.).
2. Настройка начального размещения элементов.
3. Построение начального набора решений (население).
4. Установка начальных значений параметров контроля генетического алгоритма.
5. Порядок повышения качества размещения с использованием нечетких генетических операторов:
 - 5.1. подбор решений для выступления операторов;

- 5.2. стратегия поиска компромиссов;
- 5.3. применение генетических операторов;
- 5.4. подбор лучших решений;
- 5.5. формирование нового набора решений.
6. Выбор и использование указанного алгоритма маршрутизации.
7. Расчет полученных значений целевой функции.
8. Проверка критерия останова.
9. Работа нечеткой логики контроллера, изменение параметров алгоритма управления.
10. Возврат к шагу 5, чтобы улучшить текущее размещение, или конец алгоритма.

Генетическая информация о процедурах размещения сохраняется в буфере после каждой итерации улучшения размещения элементов схемы. Затем проводится оценочная трассировка с использованием одного из известных алгоритмов (например, лучевого алгоритма), определяется оценка качества полученного решения и, при необходимости, возврат к этапу размещения.

Для повышения качества размещения используется нечеткий генетический алгоритм с решениями на основе вещественного кодирования [17–20].

Использование вещественного кодирования может повысить точность найденных решений и скорость нахождения экстремума.

Структура хромосомы представляет собой вектор действительных чисел, соответствующий координатам (x и y) позиций рабочего поля, где расположены элементы схемы, например.

$$x/0.63222/0.16464/0.9325\dots$$

$$y/0.13334/0.46561/0.7568\dots$$

Длина хромосомы равна числу позиций, и каждый ген соответствует одному элементу. Позиционирование для размещения следующего элемента выполняется таким образом, чтобы фактическая длина приращения была минимальной.

После определения нужной точки производится поиск ближайшей доступной позиции и размещение в эту позицию выбранного элемента. Предложенный метод не приводит к генерации «нелегальных решений».

Назначение элементов продолжается до тех пор, пока не будет выполнено заранее определенное количество итераций. При кодировании определяются сразу две хромосомы: первая определяет порядок размещения элементов, а вторая – ориентацию рабочего поля.

Для повышения эффективности поиска предлагаются две модифицированные модели формирования новой популяции: модель, основанная на сохранении минимального разрыва между поколениями и модель обобщения поколений [17, 20].

Схема процедуры создания новой популяции представлена на рис. 2.

Также было предложено использовать в разработанном алгоритме модифицированные генетические операторы, позволяющие оперировать с вещественными числами. В алгоритме используется модифицированный оператор поиска экстремума «Hill-Climbing» на основе градиента локальных методов поиска [19, 20]. Также были использованы два модифицированных оператора кроссинговера: арифметический и линейный операторы кроссинговера [18, 20].

В процессе использования оператора кроссинговера «Hill-Climbing» проводилось сравнение результатов его работы с результатами работы арифметического и линейного кроссинговеров. Результаты сравнения показаны на рис. 3–8.

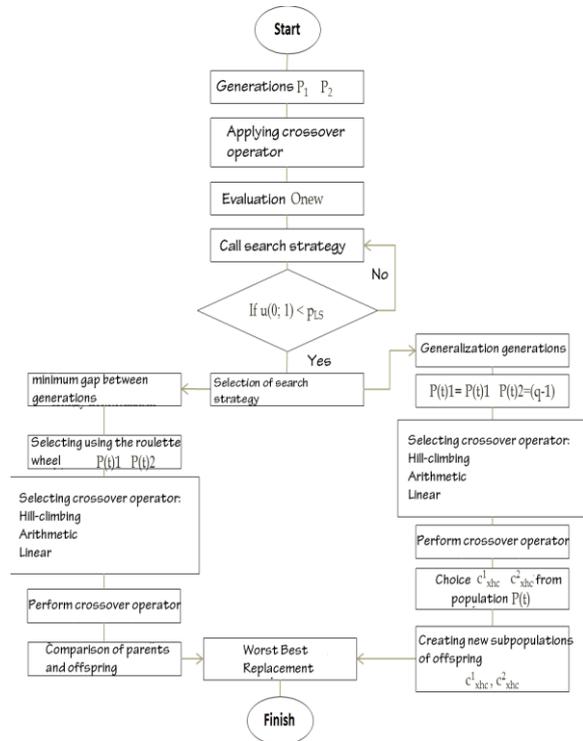


Рис. 2. Обобщенная схема работы процедуры создания новой популяции

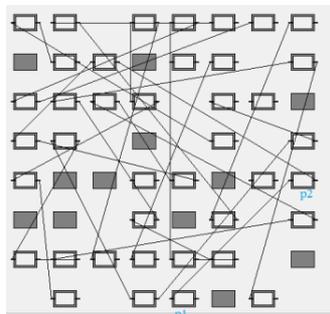


Рис. 3. Родительские хромосомы P1 и P2

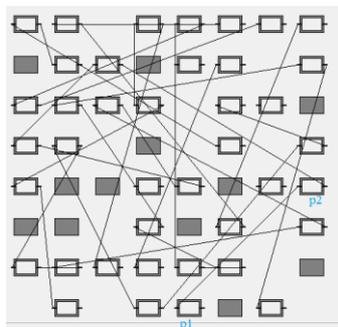


Рис. 4. Оценка потомков O1 и O2

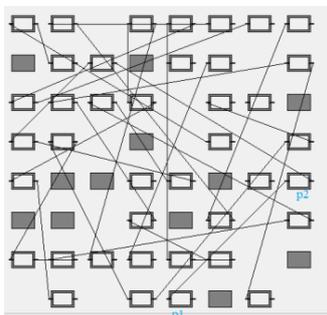


Рис. 5. Начальное размещение элементов

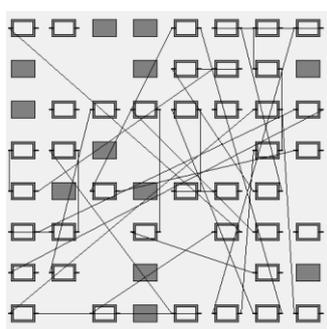


Рис. 6. Результат применения кроссингвера "hill-climbing"

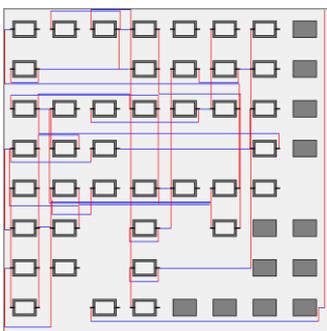


Рис. 7. Решение, полученное с применением арифметического оператора кроссингвера

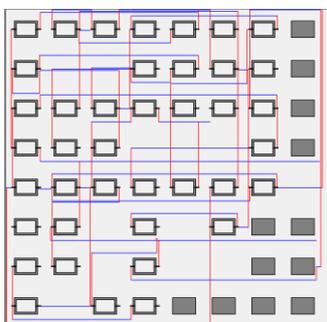


Рис. 8. Решение, полученное с применением линейного оператора кроссингвера

Отметим, что наилучшие результаты получены при использовании оператора кроссовера «восхождение на гору». Кроме того, преимуществом использования метода градиента является простота реализации и возможность начать процесс поиска и оптимизации с любыми действительными и не обязательно с эталонными решениями.

Модифицированный неоднородный оператор мутации также предоставляется. Неоднородный оператор мутации – это оператор, с помощью которого решения в алгоритме модифицируются таким образом, что на начальном этапе алгоритм поиска предоставляет униформу, а затем пытается улучшить результаты локально оптимальных решений.

Стратегия инцеста используется как механизм самоадаптивного мутационного оператора. Он заключается в том, что плотность мутаций (каждый ген вероятности мутации) определяется для каждого ребенка на основе его близких генетических родителей. Например, это может быть отношение количества совпадающих генов родительских хромосом к общему количеству генов. В результате инцеста, на начальных этапах алгоритм очень мал с большим разнообразием генофонда популяции с вероятностью мутации, то есть почти будет происходить кроссинговер. При уменьшении разнообразия, возникающем в случае попадания в локальный оптимальный алгоритм, вероятность будет увеличивать мутацию. Очевидно, что полный спуск населения будет стохастическим алгоритмом, поэтому вероятность выхода из местного населения увеличит оптимум.

Механизм реализации оператора можно представить следующим образом: когда значение гена y_i , изменяется, новое значение y'_i случайным образом генерируется в интервале $[min_i, max_i]$,

$$y'_i \begin{cases} y_i + (max_i - y_i)(1 - r(1 - \frac{1}{T})^b), & \text{если } q = 0, \\ y_i - (y_i - min_i)(1 - r(1 - \frac{b}{T})^b), & \text{если } q = 1. \end{cases}$$

где q – случайные значения в интервале от 0 до 1;
 r – случайные значения в интервале от 0 до 1;
 t – число итераций;
 T – максимальное число поколений;
 b – параметр, в зависимости от условий задачи;
 min_i и max_i – верхний и нижний пределы для значения y .

После применения генетических операторов рассчитывается значение целевой функции. Процесс повторяется до тех пор, пока не будут выполнены заданные критерии останова.

Для ускорения процесса поиска в алгоритме используется контроллер нечеткой логики. Используя набор хромосом и их целевые функции, мы можем установить определенный диапазон значений целевой функции, который нам нужен. С помощью лингвистических переменных мы можем изменять значения управляющих параметров алгоритма (рис. 9).

Результаты компьютерных экспериментов. На основе предложенных алгоритмов, методов и процедур была выполнена программная реализация для комплексного решения задач размещения и трассировки. Поиск оптимальных решений осуществляется с использованием двух вариантов стратегии поиска. Кроме того, предусмотрена возможность распараллеливания серии вычислительных экспериментов, и для каждого варианта можно выбрать собственные параметры [13–14]. Результаты проведенных серий экспериментов показаны на рис. 10.

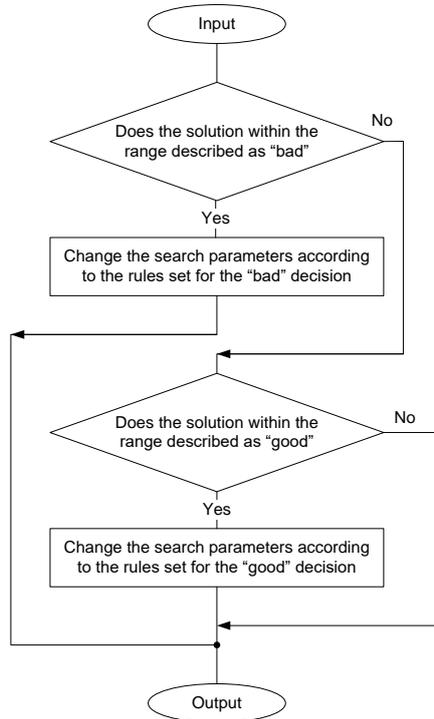


Рис. 9. Обобщенная схема алгоритма работы нечеткого логического контроллера

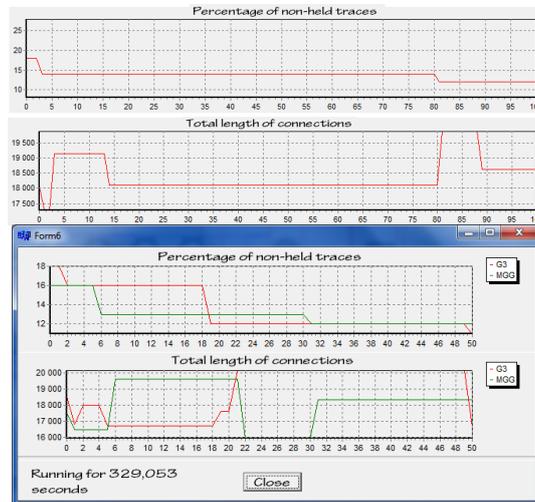


Рис. 10. Результаты экспериментальных исследований

При сравнении результатов программы с использованием контроллера нечеткой логики и без него очевидно, что в первом случае целевая функция сходится быстрее (рис. 11, 12).

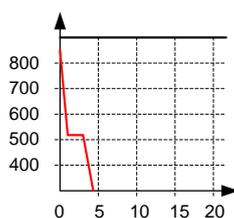


Рис. 11. Экспериментальные результаты с использованием НЛК

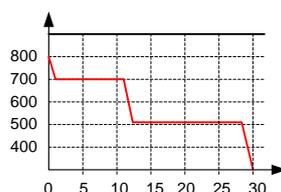


Рис. 12. Экспериментальные результаты без использования НЛК

Заключение. Разработанный оператор кроссингвера «hill- climbing» показал результаты в среднем на 2,5 % быстрее, чем классические операторы, такие как арифметические и линейные. Если сравнить результаты разных моделей формирования популяции, то можно сделать вывод, что модель, основанная на обобщении поколений с использованием оператора «hill- climbing» позволяет получить средний результат на 1,5 % быстрее. На основе анализа результатов работы моделей обобщений поколений и минимального разрыва между поколениями по числу не проведенных связей можно сделать вывод, что для схем с большим количеством элементов более эффективно использовать модель минимального разрыва между поколениями.

Исходя из полученных результатов, можно сделать вывод, что разработанный оператор кроссингвера «hill- climbing» позволяет сократить число не проведенных соединений в среднем на 1,5–2 %, при этом не уступая классическим операторам кроссингвера в плане производительности.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Charles J. Alpert, Dinesh P. Mehta, Sachin S. Sapatnekar. Handbook of algorithms for physical design automation. – CRC Press, New York, USA, 2009.
2. Shervani N. Algorithms for VLSI physical design automation. – USA, Kluwer Academy Publisher, 1995. – 538 p.
3. Cohoon J.P., Karro J., Lienig J. Evolutionary Algorithms for the Physical Design of VLSI Circuits. Advances in Evolutionary Computing: Theory and Applications, Ghosh, A., Tsutsui, S. (eds.). – Springer Verlag, London, 2003. – P. 683-712.
4. Гладков Л.А., Курейчик В.М., Курейчик В.В., Сороколетов П.В. Биоинспирированные методы в оптимизации. – М.: Физматлит, 2009.
5. Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы. – М.: Физматлит, 2010.
6. Ярушкина Н.Г. Основы теории нечетких и гибридных систем. – М.: Финансы и статистика, 2004.
7. Батыршин И.З., Недосекин А.О. и др. Нечеткие гибридные системы. Теория и практика / под ред. Н.Г. Ярушкиной. – М.: Физматлит, 2007.
8. Luger G.F. Artificial Intelligence. Structures and Strategies for Complex Problem Solving. – 6th ed. Addison Wesley, Boston MA, 2009.

9. *Russel S.J., Norvig P.* Artificial Intelligence. A modern Approach. – Prentice Hall, 2003.
10. *Michael A., Takagi H.* Dynamic control of genetic algorithms using fuzzy logic techniques // Proc. of the 5th International Conference on Genetic Algorithms. – Morgan Kaufmann, 1993. – P. 76-83.
11. *Lee M.A., Takagi H.* Integrating design stages of fuzzy systems using genetic algorithms // Proceedings of the 2nd IEEE International Conference on Fuzzy System. – 1993. – P. 612-617.
12. *Herrera F., Lozano M.* Fuzzy Adaptive Genetic Algorithms: design, taxonomy, and future directions // Soft Computing. – Springer-Verlag, 2003. – No. 7. – P. 545-562.
13. *King R.T.F.A., Radha B., Rughooputh H.C.S.* A fuzzy logic controlled genetic algorithm for optimal electrical distribution network reconfiguration // Proceedings of 2004 IEEE International Conference on Networking, Sensing and Control, Taipei, Taiwan. – 2004. – P. 577-582
14. *Praveen T., Arun Raj Kumar P.* Multi-Objective Memetic Algorithm for FPGA Placement Using Parallel Genetic Annealing // International Journal of Intelligent Systems and Applications (IJISA). – 2016. – Vol. 8, No. 4. – P. 60-66,
15. *Gladkov L.A., Gladkova N.V., Gromov S.A.* Hybrid Fuzzy Algorithm for Solving Operational Production Planning Problems // Advances in Intelligent Systems and Computing. – Springer International Publishing, Switzerland, 2017. – Vol. 573. – P. 444-456.
16. *Gladkov L.A., Gladkova N.V., Leiba S.N., Strakhov N.E.* Development and research of the hybrid approach to the solution of optimization design problems // Advances in Intelligent Systems and Computing. – Springer, Cham, 2019. – Vol. 875. – P. 246-257.
17. *Deb K., Joshi D., Anand A.* Real-coded evolutionary algorithms with parent-centric recombination // Proc. Evol. Comput. – 2002. – Vol. 1. – P. 61-66.
18. *Herrera F., Lozano M., Sánchez A.M.* A taxonomy for the crossover operator for real-coded genetic algorithms: an experimental study // Int. J. Intell. Syst. – 2003. – Vol. 18 (3). – P. 309-338.
19. *Lozano M., Herrera F., Krasnogor N., Molina D.* Real-coded memetic algorithms with crossover hill-climbing // Evol. Comput. – 2004. – No. 12 (3). – P. 273-302.
20. *Gladkov L.A., Gladkova N.V., Gusev N.Y., Semushina N.S.* Integrated approach to the solution of computer-aided design problems // Proceedings of the 4th International Scientific Conference “Intelligent Information Technologies for Industry” (IITI'19). Advances in Intelligent Systems and Computing. – Vol. 875. – Springer, Cham, 2020. – P. 246-257.
21. *Gladkov L.A., Gladkova N.V., Leiba S.N.* Manufacturing Scheduling Problem Based on Fuzzy Genetic Algorithm // Proc. of IEEE East-West Design & Test Symposium (EWDTS'2014). Kiev, Ukraine, September 26–29, 2014. –P. 209-213.
22. *Gladkov L.A., Gladkova N.V., Legebokov A.A.* Organization of Knowledge Management Based on Hybrid Intelligent Methods // Software Engineering in Intelligent Systems. Proceedings of the 4th Computer Science On-line Conference 2015 (CSOC 2015). – Vol 3: Software Engineering in Intelligent Systems. – Springer International Publishing, Switzerland 2015. – P. 107-113.
23. *Gladkov L.A., Gladkova N.V., Leiba S.N., Strakhov N.E.* Development and research of the hybrid approach to the solution of optimization design problems // Proceedings of the Third International Scientific Conference “Intelligent Information Technologies for Industry” (IITI'18). IITI'18 2018. Advances in Intelligent Systems and Computing. – Vol. 875. – Springer, Cham. – P. 246-257.

REFERENCES

1. *Charles J. Alpert, Dinesh P. Mehta, Sachin S. Sapatnekar.* Handbook of algorithms for physical design automation. CRC Press, New York, USA, 2009.
2. *Shervani N.* Algorithms for VLSI physical design automation. USA, Kluwer Academy Publisher, 1995, 538 p.
3. *Cohoon J.P., Karro J., Lienig J.* Evolutionary Algorithms for the Physical Design of VLSI Circuits. Advances in Evolutionary Computing: Theory and Applications, Ghosh, A., Tsutsui, S. (eds.). Springer Verlag, London, 2003, pp. 683-712.
4. *Gladkov L.A., Kureychik V.M., Kureychik V.V., Sorokoletov P.V.* Bioinspirirovannyye metody v optimizatsii [Bioinspired methods in optimization]. Moscow: Fizmatlit, 2009.
5. *Gladkov L.A., Kureychik V.V., Kureychik V.M.* Geneticheskie algoritmy [Genetic algorithms]. Moscow: Fizmatlit, 2010.

6. Yarushkina N.G. Osnovy teorii nechetkikh i gibridnykh system [Fundamentals of the theory of fuzzy and hybrid systems]. Moscow: Finansy i statistika, 2004.
7. Batyrshin I.Z., Nedosekin A.O. i dr. Nchetkie gibridnye sistemy. Teoriya i praktika [Fuzzy hybrid systems. Theory and practice], ed. by N.G. YArushkinoy. Moscow: Fizmatlit, 2007.
8. Luger G.F. Artificial Intelligence. Structures and Strategies for Complex Problem Solving. 6th ed. Addison Wesley, Boston MA, 2009.
9. Russel S.J., Norvig P. Artificial Intelligence. A modern Approach. Prentice Hall, 2003.
10. Michael A., Takagi H. Dynamic control of genetic algorithms using fuzzy logic techniques, *Proc. of the 5th International Conference on Genetic Algorithms. Morgan Kaufmann*, 1993, pp. 76-83.
11. Lee M.A., Takagi H. Integrating design stages of fuzzy systems using genetic algorithms, *Proceedings of the 2nd IEEE International Conference on Fuzzy System*, 1993, pp. 612-617
12. Herrera F., Lozano M. Fuzzy Adaptive Genetic Algorithms: design, taxonomy, and future directions, *Soft Computing*. Springer-Verlag, 2003, No. 7, pp. 545-562.
13. King R.T.F.A., Radha B., Rughooputh H.C.S. A fuzzy logic controlled genetic algorithm for optimal electrical distribution network reconfiguration, *Proceedings of 2004 IEEE International Conference on Networking, Sensing and Control, Taipei, Taiwan*, 2004, pp. 577-582
14. Praveen T., Arun Raj Kumar P. Multi-Objective Memetic Algorithm for FPGA Placement Using Parallel Genetic Annealing, *International Journal of Intelligent Systems and Applications (IJISA)*, 2016, Vol. 8, No. 4, pp. 60-66,
15. Gladkov L.A., Gladkova N.V., Gromov S.A. Hybrid Fuzzy Algorithm for Solving Operational Production Planning Problems, *Advances in Intelligent Systems and Computing*. Springer International Publishing, Switzerland, 2017, Vol. 573, pp. 444-456.
16. Gladkov L.A., Gladkova N.V., Leiba S.N., Strakhov N.E. Development and research of the hybrid approach to the solution of optimization design problems, *Advances in Intelligent Systems and Computing*. Springer, Cham, 2019, Vol. 875, pp. 246-257.
17. Deb K., Joshi D., Anand A. Real-coded evolutionary algorithms with parent-centric recombination, *Proc. Evol. Comput.*, 2002, Vol. 1, pp. 61-66.
18. Herrera F., Lozano M., Sánchez A.M.: A taxonomy for the crossover operator for real-coded genetic algorithms: an experimental study, *Int. J. Intell. Syst.*, 2003, Vol. 18 (3), pp. 309-338.
19. Lozano M., Herrera F., Krasnogor N., Molina D.: Real-coded memetic algorithms with crossover hill-climbing, *Evol. Comput.*, 2004, No. 12 (3), pp. 273-302.
20. Gladkov L.A., Gladkova N.V., Gusev N.Y., Semushina N.S. Integrated approach to the solution of computer-aided design problems, *Proceedings of the 4th International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'19). Advances in Intelligent Systems and Computing*, Vol. 875. Springer, Cham, 2020, pp. 246-257.
21. Gladkov L.A., Gladkova N.V., Leiba S.N. Manufacturing Scheduling Problem Based on Fuzzy Genetic Algorithm, *Proc. of IEEE East-West Design & Test Symposium (EWDTS'2014). Kiev, Ukraine, September 26–29, 2014*, pp. 209-213.
22. Gladkov L.A., Gladkova N.V., Legebokov A.A. Organization of Knowledge Management Based on Hybrid Intelligent Methods, *Software Engineering in Intelligent Systems. Proceedings of the 4th Computer Science On-line Conference 2015 (CSOC 2015)*, Vol 3: Software Engineering in Intelligent Systems. Springer International Publishing, Switzerland 2015, pp. 107-113.
23. Gladkov L.A., Gladkova N.V., Leiba S.N., Strakhov N.E. Development and research of the hybrid approach to the solution of optimization design problems, *Proceedings of the Third International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'18). IITI'18 2018. Advances in Intelligent Systems and Computing*, Vol. 875. Springer, Cham, pp. 246-257.

Статью рекомендовал к опубликованию д.т.н., профессор С.М. Ковалев.

Гладков Леонид Анатольевич – Южный федеральный университет; e-mail: leo_gladkov@mail.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371625; кафедра САПР; доцент.

Гладкова Надежда Викторовна – e-mail: nadyusha.gladkova77@mail.ru; кафедра САПР; старший преподаватель.

Ясир Муханад Джаббар Ясир – e-mail: yasir_82@mail.ru; кафедра САПР; аспирант.

Gladkov Leonid Anatol'evich – Southern Federal University; e-mail: leo_gladkov@mail.ru; 44, Nekrasovskiy lane, Taganrog, 347928, Russia; phone: +78634371625; CAD department; associate professor.

Gladkova Nadezhda Viktorovna – e-mail: nadyusha.gladkova77@mail.ru; CAD department; senior teacher.

Yasir Mukhanad Dzhabbar Yasir – e-mail: yasir_82@mail.ru; CAD department; postgraduate student.

УДК 658.512

DOI 10.18522/2311-3103-2020-4-157-164

В.И. Данильченко, Е.В. Данильченко, В.М. Курейчик

**АВТОМАТИЗИРОВАННЫЙ СТРУКТУРНО-ПАРАМЕТРИЧЕСКИЙ
СИНТЕЗ СТУПЕНЧАТОГО НАПРАВЛЕННОГО ОТВЕТВИТЕЛЯ
НА СВЯЗАННЫХ ЛИНИЯХ НА ОСНОВЕ ГЕНЕТИЧЕСКОГО
АЛГОРИТМА**

Описывается автоматизированный подход к структурно-параметрическому синтезу ступенчатого направленного ответвителя на связанных линиях на основе генетического алгоритма (ГА), позволяющий создать алгоритмическую среду в области генетического поиска для решения NP полных задач, в частности структурно-параметрический синтез ступенчатого направленного ответвителя на связанных линиях. Цель данной работы заключается в нахождении путей структурно-параметрического синтеза ступенчатого направленного ответвителя на связанных линиях на основе бионспирированной теории. Научная новизна заключается в разработке модифицированного генетического алгоритма для автоматизированного структурно-параметрического синтеза ступенчатого направленного ответвителя на связанных линиях. Постановка задачи в данной работе заключается в следующем: оптимизировать синтез схем пассивных и активных СВЧ цепей путем применения, модифицированного ГА. Принципиальное отличие от известных подходов в применении новых модифицированных генетических структур в автоматизированном структурно-параметрическом синтезе, кроме того в работе проведен новый метод расчёта ступенчатого направленного ответвителя на связанных линиях на основе модифицированного ГА. Таким образом, проблема создания методов, алгоритмов и программного обеспечения для автоматизированного структурного синтеза СВЧ модулей в настоящее время имеет особую актуальность. Ее решение позволит улучшить качественные характеристики проектируемых устройств, сократит сроки и затраты на проектирование, снизит требования к квалификации разработчика.

Генетические алгоритмы; графы и гиперграфы; эволюционные вычисления; автоматизация синтеза СВЧ модулей; САПР; принципиальная схема; топология.

V.I. Danilchenko, Y.V. Danilchenko, V.M. Kureichik

**AUTOMATED STRUCTURAL-PARAMETRIC SYNTHESIS OF A STEPPED
DIRECTIONAL RESPONDER ON CONNECTED LINES BASED
ON A GENETIC ALGORITHM**

An automated approach to the structural-parametric synthesis of a stepped directional coupler on connected lines based on a genetic algorithm (GA) is described, which makes it possible to create an algorithmic environment in the field of genetic search for solving NP complete problems, in particular, the structural-parametric synthesis of a stepped directional coupler on connected lines. The purpose of this work is to find ways of structural-parametric synthesis of a stepped directional coupler on coupled lines based on the bionspiration theory. The scientific novelty lies in the development of a modified genetic algorithm for automated structural-parametric synthesis of

a stepped directional coupler on connected lines. The problem statement in this work is as follows: to optimize the synthesis of passive and active microwave circuits by using a modified GA. A fundamental difference from the known approaches in the use of new modified genetic structures in automated structural-parametric synthesis, in addition, a new method for calculating a stepped directional coupler on connected lines based on a modified GA is righteous in the work. Thus, the problem of creating methods, algorithms and software for automated structural synthesis of microwave modules is currently of particular relevance. Its solution will improve the quality characteristics of the designed devices, reduce design time and costs, and reduce the requirements for developer qualifications.

Genetic algorithms, graphs and hypergraphs, evolutionary calculations, automation of microwave modules synthesis, CAD, schematic diagram, topology.

Введение. К проблемам при решении задач структурно-параметрического синтеза схем пассивных и активных СВЧ цепей на основе классического ГА можно отнести большое многообразие этих схем, что затрудняет их систематизацию и разработку универсального алгоритма кодирования-декодирования, а также приводит к получению избыточных или нереализуемых на практике решений. Согласно работе [1], задача структурного синтеза РЭУ относится к NP полным задачам [2], это означает, что в общем случае ее решение в принципе не может быть найдено за конечное время ни одним алгоритмом. Для устранения данной проблемы были разработаны модифицированные версии ГА, использующие предварительные знания о проектируемом классе устройств. Значительный вклад в развитие данного направления внесли следующие авторы: Л.С. Берштейн, Г.Г. Казеннов, В.П. Корячко, В.М. Курейчик, И.П. Норенков, Л.А. Растрин, Г.Г. Рябов, П.И. Соснин, А.Л. Стемповский, Д. Гольдберг, Д. Холланд и многие другие.

Теоретическая часть. Направленные ответители (НО) служат для отбора части мощности, падающей или отраженной волны из основного канала во вспомогательный (рис. 1). По схемному признаку они относятся к симметричным восьмиполосникам, у которых каждый из четырех входов связан с двумя другими и развязан с третьими. Согласно принципу взаимности, НО может быть и делителем мощности одного сигнала, и сумматором мощностей двух сигналов одинаковой частоты при соблюдении определенного фазового соотношения. При делении мощности поровну между двумя выходными каналами, т.е. при выполнении равенства $P_{3пад} = P_{4пад} = 0.5P_{1пад}$ (рис. 1), но называется гибридным соединением, которое, в свою очередь, является мостовым устройством при постоянном сдвиге фаз выходных сигналов в определенной полосе частот. S-матрица идеального НО имеет вид [3, 4]:

$$\begin{vmatrix} 0 & 0 & S_{13} & S_{14} \\ 0 & 0 & S_{14} & S_{13} \\ S_{13} & S_{14} & 0 & 0 \\ S_{14} & S_{13} & 0 & 0 \end{vmatrix}.$$

S-параметры позволяют определить характеристики НО, в том числе переходное затухание между каналами 1 и 3, 1 и 4: $b_{13} = -20 \lg |S_{13}|$, $b_{14} = -20 \lg |S_{14}|$ и коэффициент деления $K_d = 20 \lg |S_{13}|/|S_{14}|$. В радиопередатчиках, НО выполняют две основные функции: деления и суммирования мощностей из основного канала во вспомогательный для контроля параметров сигнала. В первом случае, НО является мостовым устройством, которое как типовое звено может входить в схему сумматора – делителя большого числа сигнала [4].

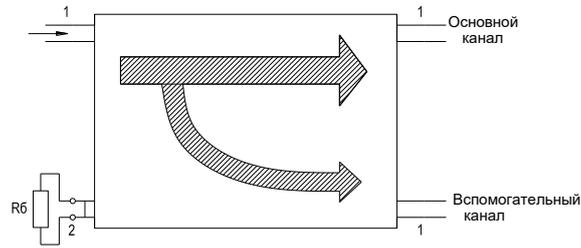


Рис. 1. Направленный ответвитель

Рассмотрим структурно-параметрический синтез одного из наиболее важных типов НО, применяемых в СВЧ диапазоне – на связанных линиях.

Синтез НО на связанных линиях осуществляется обычно восстановлением матрицы [Т] или [S] для реализации требуемого переходного затухания b_{13} и b_{14} в заданной полосе частот. Коэффициент деления НО, определяющий переходное затухание, является одним из элементов такой матрицы. Сама процедура синтеза, включающая вычисление корней функции рабочего затухания, отбор устойчивых корней, восстановление матриц, обрушение к четырёхколёсному аналогу и другие этапы, достаточно сложна. Кроме того, применение метода ограничено невозможностью учета влияния неоднородностей в местах стыка ступенек и расчетом параметров НО по таблицам только при определенных значениях коэффициента деления K_d . Поэтому определять параметры НО при промежуточном значении K_d можно только путем интерпретации. Обе причины снижают точность расчета НО в СВЧ диапазоне.

Предлагаемый далее генетический метод дихотомии параметрического синтеза НО основан на применении бионспирированных методов. Его сущность заключается в анализе большого множества вариантов, например, 1000, среди которых по определенному критерию выбирается наилучший. К достоинствам такого метода, практическая реализация которого возможна только с применением ЭВМ, относятся:

- ◆ повышение точности расчета, поскольку можно учесть все неоднородности, в том числе и в месте стыка ступенек;
- ◆ получение не только заданной частотной характеристики переходного затухания, но и других параметров, например, минимального КСВ по входу устройства;
- ◆ возможность синтеза не только при активной, но и при комплексной нагрузке;
- ◆ отсутствие обширных таблиц при дискретных значениях переходного затухания НО, поскольку расчет на ЭВМ можно производить при любом значении данного параметра;
- ◆ возможность выбора ступени НО на связанных линиях любой длины, а не только, как обычно принято, четвертьволновых.

Метод реализуется за счет механизма обычного перебора возможных точек разрыва. Он аналогичен методу деления отрезка пополам для нахождения точки, в которой значение целевой функции имеет локальный оптимум.

Например, на отрезке $A = [a, b]$, представленном в виде набора точек $A : a|b|c|d|f|g|i|h$, возможно максимально 8 точек разрыва. Однако, точность (эффективность) такого процесса определяется количеством выбранных точек n , причем $n=L-1$, где L – длина отрезка, т.е. количество выделенных дискретных точек. Дискретный интервал [1–8] обозначим L_0 . Реализованный интервал назовем интервалом неопределенности L_H .

Эффективность поиска будет составлять $2L_0/(n+1)$, что приблизительно равно двум. Такой метод деления отрезка пополам для нахождения точки, в которой целевая функция имеет локальный оптимум, называется методом дихотомии.

Приведем укрупненный алгоритм метода дихотомии:

- ◆ получить набор или одно альтернативное решение;
- ◆ ввести границы интервала, параметр ε ;
- ◆ разделить исходный отрезок пополам (при нечетном размере в любую часть берется ближайшее большее число);
- ◆ вблизи точки деления (по разные стороны с наименьшим интервалом) $\varepsilon = 0, 1, 2 \dots$ точки с экстремальным значением целевой функции;
- ◆ каждую половину отрезка снова разделить пополам и процесс расчета продолжать по исходной схеме до тех пор, пока не будет получено наилучшее значение целевой функции или ранее не будет закончено деление на основе заданной метки;
- ◆ конец работы алгоритма.

Целевая функция. Целевая функция программы синтеза НО направлена на решение двух задач: получение требуемого затухания $b_{13\tau} = b_{13k}$ или коэффициента передачи падающей волны со входа 1К в 3К НО: $K_{п.тр} = 10^{-0,05b_{12T}}$ в полосе частот $f_1 \dots f_2$ и минимума КСВ по входу 1К:

$$F_{ц} = V_1 \sum_{f_i=f_1}^{f_n} [S_{13K}(f_i) - K_{п.тр}]^2 + V_2 \sum_{f_i=f_1}^{f_n} [S_{11K}(f_i)]^2, \quad (1)$$

где V_1, V_2 – коэффициенты веса.

Варьируемыми параметрами при определении минимума целевой функции (1) являются коэффициенты связи $k_{св1}, \dots, k_{свk}$ отдельных ступеней. Следует найти такую комбинацию этих параметров при которой обеспечивается получение $\min F_{ц}$ ($k_{св1}, \dots, k_{свk}$).

Поиск глобального минимума $F_{ц}$ целевой функции осуществляется в алгоритме на основе метода перебора путем 3х вложенных друг в друга циклов ДО. предусмотрена возможность синтеза симметричного НО, состоящего из числа ступеней $K=7,5$ или 3 с коэффициентами связи $k_{св5} = k_{св3}, k_{св6} = k_{св2}, k_{св7} = k_{св1}$, каждый из которых меняется по закону:

$$k_{свk} = (C_{Hk} + \Delta C_{шк} MK) K_{п.тр}, \quad (2)$$

где C_{Hk} – начальное значение, $\Delta C_{шк}$ – шаг изменения параметра, $MK=1 \dots KN$ – переменное число, принимающее в цикле ДО значения от 1 до KN .

Все параметры, входящие в (2), устанавливаются пользователем программы, что позволяет проектировать НО с разными значениями переходного затухания в широкой полосе частот. Модули и фазы коэффициентов отражения нагрузок, подключенных к входам 2 и 4 (рис. 1), определяются с помощью полиномов:

$$\begin{aligned} |G_2| &= |G_{2н}| + A_1(f - f_0) + A_2(f - f_0)^2 + A_3(f - f_0)^3, \\ \phi_2 &= \phi_{2н} + B_1(f - f_0) + B_2(f - f_0)^2 \\ |G_4| &= |G_{4н}| + D_1(f - f_0) + D_2(f - f_0)^2 + D_3(f - f_0)^3, \\ \phi_4 &= \phi_{4н} + T_1(f - f_0) + T_2(f - f_0)^2 + T_3(f - f_0)^3, \end{aligned}$$

где постоянные коэффициенты задаются в исходных данных к программе.

Целевая функция и все остальные параметры, НО рассчитываются при тринадцати значениях частот, равномерно расположенных в заданной полосе. Число частот при необходимости можно увеличить. Длина одной ступени связанных линий устанавливается равной $l_{ст} = RL * \lambda_d/4$, где RL – произвольно выбранный коэффициент.

Эксперимент. Проведем расчет СВЧ ответвителя:

Таблица 1

Пример результата расчета ответвителя

	B3T=10.0	FT=1.5	DF=1.0	E=9.6	H=1.0		
	RL=0.5	BK1=0.8	BK2=0.2	K=5	SH=0.00008		
	CH1=0.1	CP2=0.3	CH3=0.8	CH4=1.2			
	VR2=0.0	A1=0.0	A2=0.0	A3=0.0			
	FH2=0.0	B1=0.0	B2=0.0	B3=0.0			
	VR4=0.0	D1=0.0	D2=0.0	D3=0.0			
	FH4=0.0	T1=0.0	T2=0.0	T3=0.0			
	CT	W	S	Y			
1	0.0474	0.9669	1.6262	9.7484			
2	0.1265	0.9663	0.9530	9.7487			
3	0.3795	0.7775	0.1723	9.8400			
4	0.5060	0.5764	0.0733	9.9572			
5	0.3795	0.7775	0.1723	9.8400			
6	0.1265	0.9663	0.9530	9.7487			
7	0.0474	0.9669	1.6262	9.7484			
	F	B3P	RPMT	RBMT	RXMT	RKCB	RKCX
1	1.0000	10.13622	0.31131	0.0	0.0	1.00	1.00
2	1.0833	9.99836	0.31629	0.0	0.0	1.00	1.00
3	1.1667	9.92232	0.31907	0.0	0.0	1.00	1.00
4	1.2500	9.89422	0.32010	0.0	0.0	1.00	1.00
5	1.3333	9.90085	0.31986	0.0	0.0	1.00	1.00
6	1.4167	9.92949	0.31880	0.0	0.0	1.00	1.00
7	1.500	9.96794	0.31740	0.0	0.0	1.00	1.00
8	1.5833	10.00472	0.31606	0.0	0.0	1.00	1.00
9	1.6667	10.02955	0.31515	0.0	0.0	1.00	1.00
10	1.7500	10.03389	0.31500	0.0	0.0	1.00	1.00
11	1.8333	10.01147	0.31581	0.0	0.0	1.00	1.00
12	1.9167	9.95877	0.31773	0.0	0.0	1.00	1.00
13	2.0000	9.87513	0.32081	0.0	0.0	1.00	1.00

FT и DF – центральная частота и полоса пропускания, в гигагерцах;

B3T и B3P – требуемое и полученное значение затухания НО, в децибелах;

CT(M) – вычислительный коэффициент связи (M=1...7);

W(M), Y(M), S(M) – ширина и длина микрополосковых линий и зазор между ними, в миллиметрах, в отдельных ступенях;

VR2, VR4 – модули коэффициентов отражения Γ_2 , Γ_4 нагрузок, подключенных к выходам 2, 4 (рис. 1) на центральной частоте FT;

FH 2, FH 4 – фазы этих коэффициентов, в градусах;

CH1, CH2, CH3, CH 4 – начальное значение параметров, определяющих пределы изменения коэффициентов связи ступеней при поиске.

Представлена модифицированная схема ГА и проведена ее оценка при расчете ступенчатого направленного осветителя на связанных линиях, которая показала эффективность схемы для решения проблем попадания в локальные ямы и преждевременной сходимости. Такая модель вычислений позволяет эффективно решать задачи на многоядерных процессорах.

Заклучение. В статье представлена методика автоматизированного синтеза ступенчатого направленного осветителя на основе ГА, которая дает возможность получить одновременно практически реализуемое схмотехническое решение с учетом особенностей технологии изготовления и первоначальный вариант топологии. Описана программная среда для реализации методики. Представленный пример расчета ступенчатого направленного осветителя подтверждает ее эффективность.

Поддержка. Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-37-90151.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Данильченко В.И., Курейчик В.М. Генетический алгоритм планирования размещения СБИС // Известия ЮФУ. Технические науки. – 2019. – № 2 (204). – С. 75-79.
2. Лебедев Б.К., Лебедев В.Б. Планирование на основе роевого интеллекта и генетической эволюции // Известия ЮФУ. Техническиенауки. – 2009. – № 4 (93). – С. 25-33.
3. Ahmet Karli, Vasfi Emre Omurlu, Utku Buyuksahin, RemziArtar, Ender Ortak. Self tuning fuzzy PD application on TI TMS320F 28335 for an experimental stationary quadrotor. – URL: <http://ieeexplore.ieee.org/document/6151404/> (дата обращения 23.04.2017).
4. Калентьев А.А., Гарайс Д.В., Добуш И.М., Бабак Л.И. Структурно-параметрический синтез СВЧ транзисторных усилителей на основе генетического алгоритма с использованием моделей монолитных элементов // Доклады ТУСУРа. – Декабрь 2012. – № 2 (26). – Ч. 2. – С. 104-112.
5. Tang, Maolin and Yao, Xin. A memetic algorithm for VLSI floorplanning // IEEE Transactions on Systems, Man, And Cybernetics–Part B: Cybernetics. – 2007. – No. 37 (1).
6. Горяинов А.Е., Добуш И.М., Бабак Л.И. Построение параметрических моделей пассивных компонентов СВЧ монолитных интегральных схем с использованием программы Extraction-P // Доклады ТУСУР. – С. 94-99.
7. Kokolov A.A., Salnikov A.S., Sheyerman F.I. and Babak L.I. Broadband Double-Balanced SiGe BiCMOS Mixer With Integrated Asymmetric MBaluns // Int. Conf. “Dynamics of Systems, Mechanisms and Machines” (Dynamics-2017), Omsk, Russia, 2017 (accepted for publication).
8. Wenyuan L. and Qian Z. A 0.7–1.9GHz Broadband Pseudo-differential Power Amplifier Using 0.13-um SiGe HBT Technology // 2012 Int. Conf. on Microwave and Millimeter Wave Technology (ICMMT), July 2012. – P. 1-4.
9. Коколов А.А., Добуш И.М., Шеерман Ф.И., Бабак Л.И. и др. Сложно-функциональные блоки широкополосных усилителей радиочастоты для однокристалльных приемников L- и S-диапазонов на основе технологии SiGe // 3-я Междунар. науч. конф. «ЭКБ и электронные модули» (Международный форум «Микроэлектроника-2017»), г. Алушта, октябрь 2017. – М.: Техносфера, 2017. – С. 395-401.
10. Bocklemann D.E. and Eisenstadt W.R. Combined Differential and Common-Mode Scattering Parameters: Theory and Simulation // IEEE Trans. on Microwave Theory and Techniques, Vol. MTT-43. – July 1995. – No. 7. – P. 520-523.
11. Kurokawa K. Power Waves and the Scattering Matrix // IEEE Trans. on Microwave Theory and Techniques. – 1965. – Vol. MTT-13, No. 2. – P. 194-202.
12. Zhabin D.A., Garays D.V., Kalentyev A.A., Dobush I.M. and Babak L.I. Automated Synthesis of Low Noise Amplifiers Using S-parameter Sets of Passive Elements, Asia-Pacific Microwave Conference (APMC 2017), Kuala Lumpur, Malaysia, 2017 (accepted for publication).
13. Kalentyev A.A., Garays D.V. and Babak L.I. Genetic-Algorithm-Based Synthesis of Low-Noise Amplifiers with Automatic Selection of Active Elements and DC Biases // European Microwave Week 2014, Rome, Italy, October 2014. – P. 520-523.
14. Babak L.I., Kokolov A.A. and Kalentyev A.A. A New Genetic-Algorithm-Based Technique for Low Noise Amplifier Synthesis, European Microwave Week 2012, Amsterdam, The Netherlands, November 2012. Genetic-Algorithm-Based Synthesis of Low-Noise Amplifiers with Automatic Selection of Active Elements and DC Biases, European Microwave Week 2014, Rome, Italy, October. – P. 520-523.
15. Mann G.K.I., Gosine R.G. Three-dimensional min–max-gravity based fuzzy PID inference analysis and tuning // Fuzzy Sets and Systems. – 2005. – Vol. 156. – P. 300-323.

16. Курейчик В.М. Гибридные генетические алгоритмы // Известия ЮФУ. Технические науки. – 2007. – № 2 (77). – С. 5-12.
17. Кошевой С.Е., Дорофеев С.Ю., Бабак Л.И. Структурный синтез СВЧ-устройств на основе генетического алгоритма в системе автоматизированного проектирования INDESYS // Всерос. науч.-техн. конф. с междунар. участ. «Современные проблемы радиоэлектроники». – Красноярск: СФУ, 2009. – С. 421-424.
18. Коколов А.А., Добуш И.М., Шеерман Ф.И., Бабак Л.И. и др. Сложно-функциональные блоки широкополосных усилителей радиочастоты для однокристалльных приемников L- и S-диапазонов на основе технологии SiGe // 3-я Междунар. науч. конф. «ЭКБ и электронные модули» (Международный форум «Микроэлектроника-2017»), г. Алушта, октябрь 2017. – М.: Техносфера, 2017. – С. 395-401.
19. Запорожец Д.Ю., Кравченко Ю.А., Лежебоков А.А. Способы интеллектуального анализа данных в сложных системах // Известия Кабардино-Балкарского научного центра РАН. – 2013. – № 3. – С. 52-54.
20. Zhiqiang Yang, Jimin Zhang, Zhongchao Chen, Baoan Zhang. Semi-active control of high-speed trains based on fuzzy PID control // Procedia Engineering. – 2011. – Vol. 15. – P. 521-525.

REFERENCES

1. Danil'chenko V.I., Kureychik V.M. Geneticheskiy algoritm planirovaniya razmeshcheniya SBIS [Genetic algorithm for VLSI placement planning], *Izvestiya YuFU. Tekhnicheskienauki* [Izvestiya SFedU. Engineering Sciences], 2019, No. 2 (204), pp. 75-79.
2. Lebedev B.K., Lebedev V.B. Planirovanie na osnove roevogo intellekta i geneticheskoy evolyutsii [Planning based on swarm intelligence and genetic evolution]. *Izvestiya YuFU. Tekhnicheskienauki* [Izvestiya SFedU. Engineering Sciences], 2009, No. 4 (93), pp. 25-33.
3. Ahmet Karli, Vasfi Emre Omurlu, Utku Buyuksahin, RemziArtar, Ender Ortak. Self tuning fuzzy PD application on TI TMS320F 28335 for an experimental stationary quadrotor. Available at: <http://ieeexplore.ieee.org/document/6151404/> (accessed 23 April 2017).
4. Kalent'ev A.A., Garays D.V., Dobush I.M., Babak L.I. Strukturno-parametricheskiiy sintez SVCh tranzistornykh usiliteley na osnove geneticheskogo algoritma s ispol'zovaniem modeley monolitnykh elementov [Structural-parametric synthesis of microwave transistor amplifiers based on a genetic algorithm using models of monolithic elements], *Doklady TUSUR* [Proceedings of TUSUR University]. Dekabr' 2012, No. 2 (26), Part 2, pp. 104-112.
5. Tang, Maolin and Yao, Xin. A memetic algorithm for VLSI floorplanning, *IEEE Transactions on Systems, Man, And Cybernetics–Part B: Cybernetics*, 2007, No. 37 (1).
6. Goryainov A.E., Dobush I.M., Babak L.I. Postroenie parametricheskikh modeley passivnykh komponentov SVCh monolitnykh integral'nykh skhem s ispol'zovaniem programmy Extraction-P [Construction of parametric models of passive components of monolithic microwave integrated circuits using the Extraction-P program], *Doklady TUSUR* [Proceedings of TUSUR University], pp. 94-99.
7. Kokolov A.A., Salnikov A.S., Sheyerman F.I. and Babak L.I. Broadband Double-Balanced SiGe BiCMOS Mixer With Integrated Asymmetric MBaluns, *Int. Conf. "Dynamics of Systems, Mechanisms and Machines" (Dynamics-2017), Omsk, Russia, 2017 (accepted for publication)*.
8. Wenyuan L. and Qian Z. A 0.7–1.9GHz Broadband Pseudo-differential Power Amplifier Using 0.13-um SiGe HBT Technology, *2012 Int. Conf. on Microwave and Millimeter Wave Technology (ICMMT)*, July 2012, pp. 1-4.
9. Kokolov A.A., Dobush I.M., Sheerman F.I., Babak L.I. i dr. Slozhno-funktsional'nye bloki shirokopolosnykh usiliteley radiochastoty dlya odnokristal'nykh priemnikov L- i S-diapazonov na osnove tekhnologii SiGe [Complex-functional blocks of broadband radio frequency amplifiers for single-chip l- and S-band receivers based on SiGe technology], *3-ya Mezhdunar. nauch. konf. «EKБ i elektronnyye moduli» (Mezhdunarodnyy forum «Mikroelektronika-2017»)*, g. Alushta, oktyabr' 2017 [3rd international. scientific conference " ECB and electronic modules "(international forum "Microelectronics-2017"), Alushta, October 2017]. Moscow: Tekhnosfera, 2017, pp. 395-401.
10. Bocklemann D.E. and Eisenstadt W.R. Combined Differential and Common-Mode Scattering Parameters: Theory and Simulation, *IEEE Trans. on Microwave Theory and Techniques*, Vol. MTT-43, July 1995, No. 7, pp. 520-523.

11. Kurokawa K. Power Waves and the Scattering Matrix, *IEEE Trans. on Microwave Theory and Techniques*, 1965, Vol. MTT-13, No. 2, pp. 194-202.
12. Zhabin D.A., Garays D.V., Kalentyev A.A., Dobush I.M. and Babak L.I. Automated Synthesis of Low Noise Amplifiers Using S-parameter Sets of Passive Elements, Asia-Pacific Microwave Conference (APMC 2017), Kuala Lumpur, Malaysia, 2017 (accepted for publication).
13. Kalentyev A.A., Garays D.V. and Babak L.I. Genetic-Algorithm-Based Synthesis of Low-Noise Amplifiers with Automatic Selection of Active Elements and DC Biases, *European Microwave Week 2014, Rome, Italy, October 2014*, pp. 520-523.
14. Babak L.I., Kokolov A.A. and Kalentyev A.A. A New Genetic-Algorithm-Based Technique for Low Noise Amplifier Synthesis, *European Microwave Week 2012, Amsterdam, The Netherlands, November 2012. Genetic-Algorithm-Based Synthesis of Low-Noise Amplifiers with Automatic Selection of Active Elements and DC Biases, European Microwave Week 2014, Rome, Italy, October*, pp. 520-523.
15. Mann G.K.I., Gosine R.G. Three-dimensional min-max-gravity based fuzzy PID inference analysis and tuning, *Fuzzy Sets and Systems*, 2005, Vol. 156, pp. 300-323.
16. Kureychik V.M. Gibridnye geneticheskie algoritmy [Hybrid genetic algorithms], *Izvestiya YuFU. Tekhnicheskienauki* [Izvestiya SFedU. Engineering Sciences], 2007, No. 2 (77), pp. 5-12.
17. Koshevoy S.E., Dorofeev C.Yu., Babak L.I. Strukturnyy sintez SVCh-ustroystv na osnove geneticheskogo algoritma v sisteme avtomatizirovannogo proektirovaniya INDESYS [Structural synthesis of microwave devices based on a genetic algorithm in the computer-aided design system INDESYS], *Vseros. nauch.-tekh. konf. s mezhdunar. uchast. «Sovremennye problemy radioelektorniki»* [All-Russian scientific and technical conference with international participation "Modern problems of Radioelectronics"]. Krasnoyarsk: SFU, 2009, pp. 421-424.
18. Kokolov A.A., Dobush I.M., Sheerman F.I., Babak L.I. i dr. Slozhno-funktsional'nye bloki shirokopolosnykh usiliteley radiochastoty dlya odnokristal'nykh priemnikov L- i S-diapazonov na osnove tekhnologii SiGe [Complex-functional blocks of broadband radio frequency amplifiers for single-chip L- and S-band receivers based on SiGe technology], *3-ya Mezhdunar. nauch. konf. «EKB i elektronnyye moduli» (Mezhdunarodnyy forum «Mikroelektronika-2017»), g. Alushta, oktyabr' 2017* [3rd international. scientific conference " ECB and electronic modules "(international forum "Microelectronics-2017"), Alushta, October 2017]. Moscow: Tekhnosfera, 2017, pp. 395-401.
19. Zaporozhets D.Yu., Kravchenko Yu.A., Lezhebokov A.A. Sposoby intellektual'nogo analiza dannykh v slozhnykh sistemakh [Methods of data mining in complex systems], *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN* [Izvestiya Kabardino-Balkar scientific center of the Russian Academy of Sciences], 2013, No. 3, pp. 52-54.
20. Zhiqiang Yang, Jimin Zhang, Zhongchao Chen, Baoan Zhang. Semi-active control of high-speed trains based on fuzzy PID control, *Procedia Engineering*, 2011, Vol. 15, pp. 521-525.

Статью рекомендовал к опубликованию к.т.н. Ю.П. Волощенко.

Данильченко Евгения Владимировна – Южный Федеральный университет; e-mail: lipkina@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: +79525691761; кафедра САПР.

Данильченко Владислав Иванович – e-mail: vdanielchenko@sfedu.ru; тел.: +79526088561; кафедра САПР.

Курейчик Виктор Михайлович – e-mail: vmkureychik@sfedu.ru; тел.: +79282132730; кафедра САПР; д.т.н.; профессор.

Danilchenko Yevgenia Vladimirovna – Southern Federal University; e-mail: lipkina@sfedu.ru; 44, Nekrasovsky lane, Taganrog, 347928, Russia; phone: +79525691761; the department of computer-aided design; postgraduate.

Danilchenko Vladislav Ivanovich – e-mail: vdanielchenko@sfedu.ru; phone: +79526088561; the department of computer-aided design; postgraduate.

Kureichik Viktor Mikhailovich – e-mail: vmkureychik@sfedu.ru; phone: +78634311487; the department of computer-aided design; dr. of eng. sc.; assistant professor.

Д.В. Тельпухов, Т.Д. Жукова, А.Н. Щелоков

АНАЛИЗ ХАРАКТЕРИСТИК СФК НА ОСНОВЕ МЕТОДОВ ИЗБЫТОЧНОГО КОДИРОВАНИЯ*

Обычно сбои, возникающие в электронной аппаратуре под действием различных дестабилизирующих факторов, таких как, например, высокая или низкая температура или ионизирующее излучение, находились под пристальным вниманием разработчиков элементов памяти. Но последние исследования в данной области показывают, что с развитием микроэлектронной промышленности число сбоев в комбинационных участках схемы растет и в скором времени их частота возникновения будет сопоставима с частотой в незащищенных элементах памяти. На сегодняшний день для решения проблемы проектирования комбинационных схем повышенной сбоеустойчивости в условиях экстремального применения особое внимание стали уделять методам синтеза схем функционального контроля (СФК). Данные методы, позволяя за счет внесения дополнительной структурной избыточности, наделять схему способностью автоматически выполнять обнаружение и/или исправление возникающих в ней ошибок. Однако, в результате применения различных методов синтеза СФК в зависимости от исходных параметров и внутреннего строения защищаемой схемы реализуются устройства, обладающие различной эффективностью и характеристиками надежности. Поэтому возникает необходимость в определении и разработке оценочных функций для выполнения анализа по нахождению наилучшего метода построения схемы контроля для конкретного устройства без проведения предварительного моделирования. Данная работа посвящена разработке спецификации оценочных функций структурной избыточности и характеристик надежности на примере разработанных методов синтеза схем функционального контроля на базе спектрального и низкоплотностного кода. Был проведен сравнительный и корреляционный анализ аналитических данных с экспериментальными значениями с целью оценки эффективности полученных в результате исследования функций. Полученные в рамках данной статьи оценочные функции продемонстрировали высокую точность в вычислении характеристик СФК.

Отказоустойчивость; схемы функционального контроля (СФК); характеристики надежности; комбинационные схемы; спектральный R-код; код с малой плотностью проверок на четность.

D.V. Telpukhov, A.N. Schelokov, T.D. Zhukova

ANALYSIS OF CHARACTERISTIC FOR CED CIRCUITS BASED ON REDUNDANT ENCODING METHODS

Typically, soft errors that occur in electronic equipment under influence of various destabilizing factors, were under the scrutiny of memory element developers. But recent research in this area shows that with development of microelectronics, the number of soft errors in combination circuits is increasing and soon their frequency of occurrence will be comparable to that in unprotected memory elements. Presently, to address this problem, special attention has been paid to methods based on control devices. These methods, by introducing additional structural redundancy, enable scheme to automatically detect and/or correct errors that occur in it. However, as a result of application of various methods of synthesis concurrent error detection (CED) circuits depending on initial parameters and internal structure of protected scheme devices possessing various efficiency and reliability characteristics are realized. However, as a result of application of various methods of synthesis CED circuits depending on initial parameters and internal structure protected circuit, the devices possessing different efficiency and reliability characteristics are realized. That is why there is a necessity to define and develop evaluation functions for analysis in

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-07-00626 А.

order to find the best method of synthesis CED circuit for certain device without any preliminary modeling. This work is devoted to development specification of structural redundancy and reliability characteristics evaluation functions on the example of CED circuits on basis of spectral and low-density parity-check code. The comparative and correlation analysis of analytical data with experimental values was carried out to evaluate efficiency of the functions obtained as a result of study.

Fault tolerance; concurrent error detection (CED) circuits; characteristic of reliability; combinational circuit; spectral R-code; low-density parity-check code.

Введение. При экстремальном применении на электронную аппаратуру зачастую воздействуют различные дестабилизирующие факторы (температура, ионизирующее излучение и т.д.), которые могут привести к возникновению кратковременных нарушений работоспособности устройства – сбоям. До настоящего времени исследования данной проблемы, а также разработанные методы борьбы с последствиями возникновения сбоев в основном касались последовательностных схем из-за наличия в них большого числа битов, которые могут быть подвержены искажению. В то время как комбинационная логика за счет присутствия внутренней защиты на основе механизмов маскирования [1–3] обладала меньшей чувствительностью к их возникновению.

Стремительное развитие микроэлектронной промышленности в ближайшем будущем грозит тем, что частота возникновения сбоев в логических схемах будет сопоставима с частотой в незащищенных элементах памяти [4–7]. Для решения данной проблемы в последнее время все чаще используют методы на основе средств контроля. Их применение позволяет решить проблему большой избыточности, к которой приводят традиционные методы повышения сбоеустойчивости, такие как тройное модульное резервирование (Triple Modular Redundancy, TMR).

В рамках комбинационных схем реализация схемы функционального контроля (СФК) осуществляется путем добавления некоторой подсхемы, которая позволяет выполнить обнаружение и/или исправление ошибки [8–12]. Под одиночным сбоем в данной работе понимается инверсия на логическом элементе схемы.

На рис. 1 представлена общая структурная схема СФК, позволяющая выполнить исправление и обнаружение ошибок. Как видно из рисунка, в состав схемы контроля входят следующие блоки:

- ◆ Основная схема (ОС) – комбинационная схема, для защиты которой было произведено формирование СФК.
- ◆ Кодер, в состав которого входит копия ОС и подсхема, отвечающая за вычисление битов четности.
- ◆ Декодер, выполняющий исправление/обнаружение возникших в защищаемой схеме ошибок.



Рис. 1. Общая структура схемы функционального контроля

В результате работы схемы контроля на выходе устройства могут возникнуть следующие ситуации:

- ◆ **Ошибка маскирована.** Возникновение ошибки не повлияло на результирующий выход схемы. При этом сигнал о возникновении в схеме ошибки отсутствует. Зачастую данная ситуация возникает благодаря механизму логического маскирования.

- ◆ **Ошибка исправлена.** Возникшая ошибка была обнаружена и исправлена декодером схемы контроля.

- ◆ **Ложная тревога.** СФК сигнализировала о наличии ошибки, но при этом реальная ошибка отсутствует.

- ◆ **Обнаружение ошибки.** Ошибка схемой контроля была обнаружена, но не исправлена.

- ◆ **Ошибка пропущена.** СФК не обнаружила ошибку и в следствии этого не выполнила ее исправление.

Применение для синтеза СФК разнообразных помехоустойчивых кодов, а также методов декодирования для каждого конкретного устройства приводит к построению схем, каждая из которых обладает своими структурными особенностями и характеристиками надежности. Это приводит к тому, что в определенных ситуациях каждый из методов реализации схем контроля будет обладать различной эффективностью. Из-за чего возникает необходимость в исследовании и разработке спецификации надежности СФК [13]. Полученные оценочные функции позволили бы определить наиболее эффективный метод построения схемы в зависимости от параметров защищаемой комбинационной схемы.

В данной работе в качестве основных характеристик СФК для оценки эффективности применения схем контроля рассматриваются структурная избыточность и характеристика надежности. В следующих разделах будет рассмотрен вывод оценочных функций данных характеристик для схем контроля на основе спектрального и низкоплотностного кодов.

1. Определение характеристик СФК на основе спектрального кода. Схема функционального контроля на основе спектрального R-кода способна выполнить исправление однократной и обнаружение двукратной ошибок [14–17]. Особенность данной схемы заключается в определении и защите специальными технологическими средствами подсхемы кодера, отвечающей за вычисление общего бита четности, необходимого для корректного функционирования СФК.

Стоит отметить, что в данной работе используется модель ошибок, при которой сбой с равной долей вероятности вносится во все части схемы, а именно в любой логический элемент схемы контроля. Поэтому для получения оценочной функции характеристики надежности [18] схемы был проведен ряд исследований по определению ненадежных блоков, ошибка на которых могла привести к возникновению неисправимой и необнаруживаемой ошибки на выходе устройства, т.е. к возникновению в СФК пропуска ошибки. В результате таких исследований для схемы контроля на основе спектрального кода были определены следующие ненадежные участки:

- ◆ Основная схема.
- ◆ Блок, отвечающий за вычисление общего синдрома.
- ◆ Блок, отвечающий за вычисление общего бита четности ОС.
- ◆ Блок, определяющий режим работы схемы контроля.

На рис. 2 представлена структурная схема СФК на основе спектрального R-кода на примере комбинационной схемы с 4 выходами. Красным отмечены ненадежные блоки, ошибка на которых может привести к возникновению пропуска.

Для получения итоговой функции расчета характеристики надежности необходимо было определить вероятность пропуска ошибки для каждого из выявленных ненадежных блоков:

$$P = P_{OC} + P_{Обч} + P_{S0} + P_{PP},$$

где P_{OC} – вероятность пропуска при возникновении ошибки в ОС, $P_{Обч}$ – вероятность пропуска при возникновении ошибки в блоке, выполняющем вычисление общего бита четности ОС, P_{S0} – вероятность пропуска в блоке вычисления общего синдрома, P_{PP} – вероятность пропуска в блоке определения режима работы.

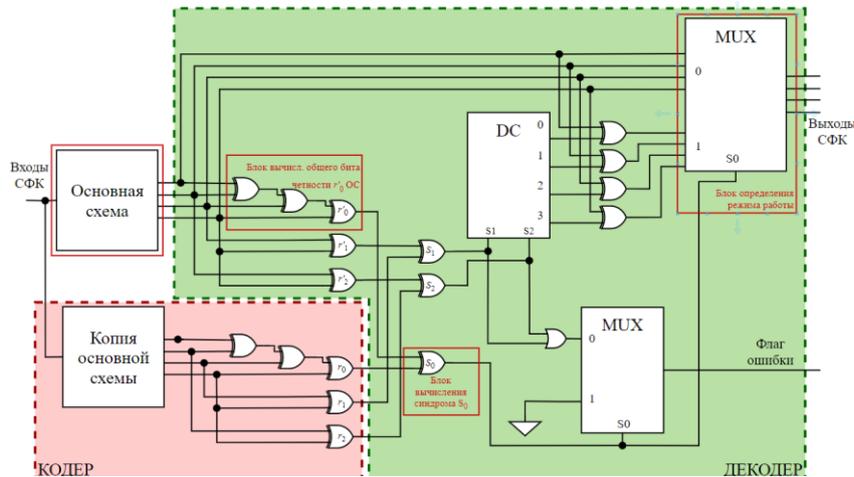


Рис. 2. Структурная схема СФК на основе спектрального R-кода на примере комбинационной схемы с 4 выходами с выделенными ненадежными блоками

В результате проведения ряда аналитических и практических исследований для каждой из ненадежных подсхем были получены оценочные функции вероятности пропуска:

$$P_{OC} = \sum_{i=3,5,\dots}^{rate} \frac{M_{OC}}{M_{СФК} - M_K} p_i + \sum_{i=3,5,\dots}^{rate} \frac{M_{OC}}{M_{СФК} - M_K} p_i f_i = \sum_{i=3,5,\dots}^{rate} \beta p_i + \sum_{i=3,5,\dots}^{rate} \beta p_i f_i,$$

$$P_{Обч} = \frac{k-1}{M_{СФК} - M_K}, \quad P_{S0} = \frac{1}{M_{СФК} - M_K},$$

$$P_{PP} = \left(1 - \frac{2^{2k+1} - 1}{2^{3k+1} - 1} \right) \left(\frac{3k+1}{M_{СФК} - M_K} \right),$$

где $rate$ – максимальная кратность ошибки, которая может возникнуть в ОС, p_i – вероятность возникновения в ОС ошибки i кратности, f_i – вероятность возникновения $flag = 0$ при возникновении четной ошибки, что приведет к появлению пропуска, β – вероятность того, что ошибка попадет в ОС (отношение размерности ОС к схеме контроля без учета защитного конуса, k – число информационных разрядов ОС.

С целью определения оценочной функции структурной избыточности для рассматриваемой СФК был проведен анализ блоков, из которой она состоит:

$$M_{СФК} = M_{OC} + M_{кодер} + M_{декодер},$$

где M_{OC} – число элементов основной комбинационной схемы, которую необходимо защитить, $M_{кодер}$ – число элементов кодера СФК, $M_{декодер}$ – число элементов декодера СФК.

Для определения итоговой функции необходимо было проанализировать структуру каждого из блоков СФК в отдельности.

Как было сказано ранее, в состав кодера входит копия основной комбинационной схемы и подсхема, вычисляющая проверочные разряды:

$$M_{кодер} = M_{OC} + M_{gx}.$$

Размер подсхемы M_{gx} , а именно число ЛЭ XOR2, используемых для вычисления проверочных разрядов копии ОС, определяется порождающей матрицей спектрального кода.

Вносимая структурная избыточность декодером вычисляется следующим образом:

$$M_{декодер} = M_{gx} + M_s + M_{dc} + M_{подсх. обнаруж.} + M_{подсх. исправ.},$$

где число M_{gx} для копии основной схемы вычисляется по тому же алгоритму, что и для кодера, а M_s совпадает с числом проверочных разрядов m .

Для вычисления размера дешифратора M_{dc} , выполняющего функцию определения номера ошибочного информационного разряда, необходимо рассмотреть его структурную схему. На рис. 3 представлен пример схемы дешифратора, используемого в СФК на основе спектрального R-кода для 3 выходов. Число выходов дешифратора зависит от числа его входов следующим соотношением:

$$n_{вых.dc} = 2^n_{вх.dc}.$$

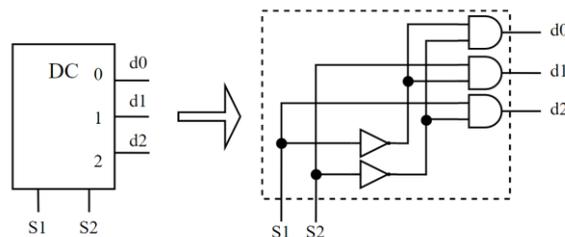


Рис. 3. Структурная схема дешифратора в СФК на основе спектрального R-кода для $k = 3$ выходов

В зависимости от числа выходов реализуемый в СФК дешифратор может быть неполным, т.е. обладать числом выходов $< 2^n_{вх.dc}$. Это возникает по причине того, что для СФК число выходов дешифратора зависит от числа выходов схемы и зачастую требуется дешифратор реализующий не все минтермы.

Функция, вычисляющая размер дешифратора для СФК будет иметь следующий вид:

$$M_{dc} = 2 + (m - 2) k.$$

На рис. 4 представлена структура подсхемы обнаружения декодера схемы контроля. Из нее видно, что размер данной подсхемы целиком зависит от числа синдромов схемы. Поэтому оценочная функция для вычисления структурной избыточности подсхемы обнаружения имеет следующий вид:

$$M_{подсх.обнаруж.} = m + 2.$$

Для вычисления размера подсхемы исправления (на рис. 5 представлен пример для схемы с 3 выходами) была получена следующая функция:

$$M_{\text{подсх.исправ.}} = k + 1 + 3k = 4k + 1.$$

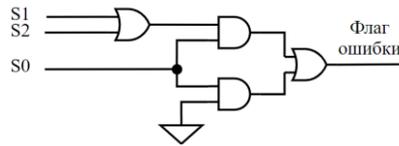


Рис. 4. Структура подсхемы обнаружения ошибок на примере схемы с $k = 3$

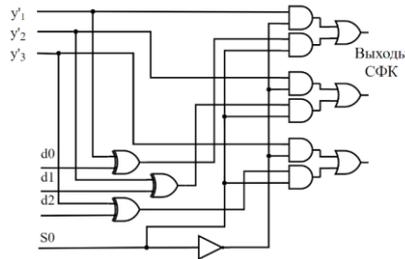


Рис. 5. Структура подсхемы декодера СФК, выполняющего функцию исправления однократных ошибок на примере схемы с $k = 3$

В результате проведенных исследований была получена оценочная функция структурной избыточности для рассматриваемой в качестве примера схемы функционального контроля:

$$M_{\text{СФК}} = 2M_{\text{ОС}} + 2M_{\text{гх}} + 3m + m \cdot k + 2k + 2.$$

Данная функция позволяет точно вычислить размер схемы контроля, реализованной на основе спектрального R-кода.

2. Определение характеристик СФК на основе низкоплотного кода.

Для определения оценочных функций структурной избыточности и характеристик надежности для СФК на основе низкоплотного кода [19–20] также были проведены исследования по оценке ненадежности блоков схемы. На рис. 6 представлена общая структура рассматриваемой схемы контроля.

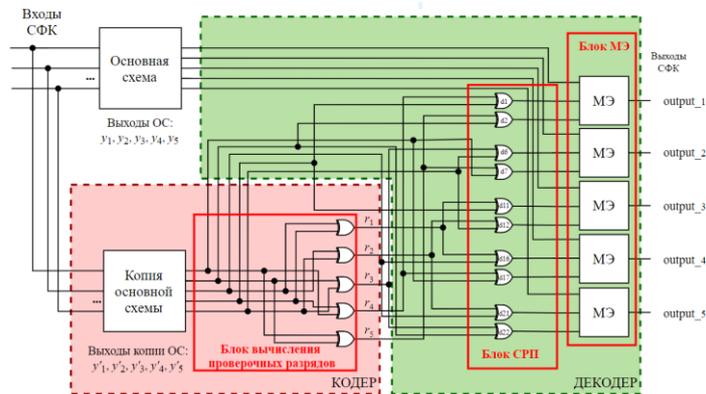


Рис. 6. Структурная схема СФК на основе низкоплотного кода на примере комбинационной схемы с 5 выходами с выделенными блоками, ошибка на которых приводит к пропуску на выходе

Проведенные исследования помогли определить в качестве ненадежных следующие подсхемы:

- ◆ Копия основной схемы, входящая в состав кодера СФК.
- ◆ Блок мажоритарных элементов (МЭ) для системы раздельных проверок.

В результате была получена следующая оценочная функция вероятности пропуска ошибки для СФК на основе низкоплотного кода:

$$P_{OC} = \sum_{i=1,2,\dots}^{rate} \frac{M_{OC}}{M_{СФК}} p_i + \frac{k \cdot M_{МЭ}}{M_{СФК}} p_{МЭ} = \sum_{i=1,2,\dots}^{rate} \beta p_i + \frac{k M_{МЭ}}{M_{СФК}},$$

где $M_{МЭ}$ – число элементов в блоке мажоритарного элемента, $p_{МЭ}$ – вероятность пропуска ошибки блоком мажоритарного элемента.

Общая функция для вычисления размера СФК на основе низкоплотного кода:

$$M_{СФК} = M_{OC} + M_{кодер} + M_{декодер}.$$

Стоит отметить, что в кодере число элементов необходимых для вычисления проверочных разрядов копии ОС зависит от разреженной порождающей матрицы специального вида. Поэтому функция для вычисления размера данной подсхемы имеет следующий вид:

$$M_{кодер} = M_{OC} + k.$$

Структурная схема кодера для комбинационной схемы с $k = 5$ представлена на рис. 7.

Декодер для рассматриваемой СФК реализован по методу мажоритарного декодирования, и состоит из систем раздельных проверок и мажоритарных элементов (МЭ), отдельно реализованных для каждого выхода основной схемы. Пример структуры данной подсхемы представлен на рис. 8.

Размерность блока декодера определяется с помощью следующего соотношения:

$$M_{декодер} = M_{МЭ} \cdot k + 2k = 6k.$$

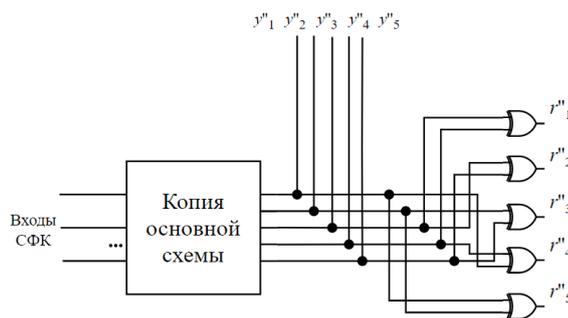


Рис. 7. Структурная схема кодера СФК на основе низкоплотного кода для комбинационной схемы с $k = 5$

Структурная схема мажоритарного элемента для системы раздельных проверок одного из выходов схемы изображена на рис. 9.

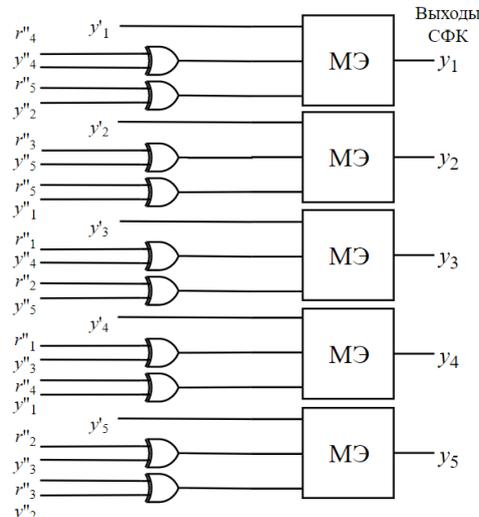


Рис. 8. Структурная схема декодера СФК на основе низкоплотного кода для комбинационной схемы с $k = 5$

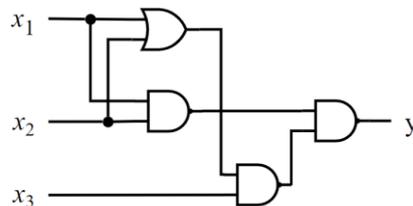


Рис. 9. Структурная схема мажоритарного элемента (МЭ) для системы отдельных проверок

В результате проведенного исследования оценочная функция для определения вносимой избыточности имеет следующий вид:

$$M_{\text{СФК}} = 2M_{\text{ОС}} + 7k.$$

Как видно из представленной функции, основное влияние на размер подсхемы, добавляемой при генерации СФК, оказывает значение параметра k - число выходов основной комбинационной схемы, для которой происходит синтез.

Как и в случае схемы на основе спектрального кода, выведенная оценочная функция позволяет точно вычислить размер результирующего устройства.

3. Результаты численных экспериментов. В рамках работы был проведен сравнительный анализ с целью сопоставления рассчитанных аналитически характеристик надежности СФК с практическими значениями, полученными в результате моделирования. В случае структурной избыточности такой анализ не проводился по причине того, что оценочные функции позволяют точно вычислять размер схем контроля.

Результаты проведенного анализа для тестовых схем из наборов ISCAS'85 и LGSynth89 [21] представлены в табл. 1 и 2, соответственно.

Таблица 1

Результаты сравнительного анализа аналитического значения вероятности пропуска ошибки P с экспериментальным $P_{\text{экс}}$ для схем контроля на основе спектрального кода

Benchmark	P	$P_{\text{экс}}$	$ P_{\text{экс}} - P $
ISCAS'85			
c1355	13,46	11,17	2,29
c1908	8,76	8,05	0,71
c432	6,63	7,05	0,42
c499	14,12	12,75	1,37
c880	9,28	8,75	0,53
LGSynth'89			
5xp1	17,36	14,85	2,51
apex7_synth	19,52	19,07	0,45
C5315_synth	14,01	13,45	0,56
C880_synth	15,10	13,11	1,99
cu_synth	24,81	26,6	1,79

Таблица 2

Результаты сравнительного анализа аналитического значения вероятности пропуска ошибки P с экспериментальным $P_{\text{экс}}$ для схем контроля на основе низкоплотного кода

Benchmark	P	$P_{\text{экс}}$	$ P_{\text{экс}} - P $
ISCAS'85			
c1355	20,66	18,19	2,48
c1908	18,47	18,18	0,29
c432	19,14	18,00	1,14
c499	24,71	22,89	1,82
c880	21,02	22,07	1,05
LGSynth'89			
5xp1	21,89	22,30	0,41
apex7_synth	26,04	31,08	5,04
b9_synth	24,98	25,72	0,74
C1908_synth	20,73	24,09	3,36
C5315_synth	22,22	24,15	1,94

Как видно из представленных в таблицах данных, рассчитанные значения вероятности пропуска оказались достаточно близки к экспериментальным. Разница между данными значениями для рассматриваемых комбинационных схем в среднем составила 1,3 % (для спектрального кода) и 4,7 % (для низкоплотного кода).

Для определения коэффициентов корреляций для полученных оценочных функций был проведен корреляционный анализ, результаты которого изображены на рис. 10.

Из графиков видно, что результаты, полученные с помощью оценочных функций, хорошо коррелируют с практическими значениями. Данное утверждение подтверждается вычисленными коэффициентами корреляции, которые составили 0,95 (для спектрального кода) и 0,82 (для низкоплотного кода).

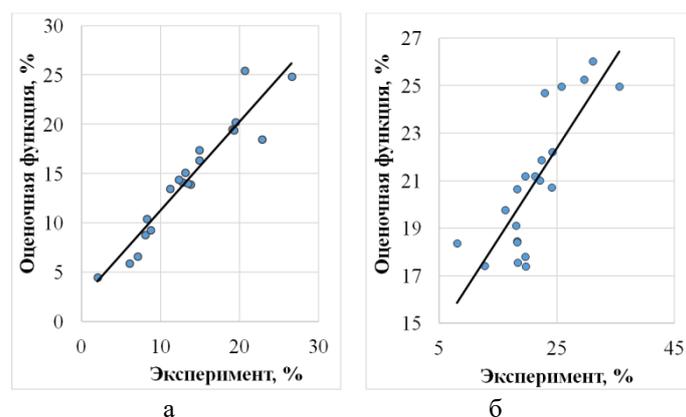


Рис. 10. Корреляционный анализ зависимости аналитических значений вероятности пропуска ошибки с экспериментальными данными, полученными в результате моделирования, для СФК на основе: а – спектрального и б – низкоплотностного кодов

Заключение. В работе предложена спецификация оценочных функций для схем функционального контроля на основе спектрального и низкоплотностного кодов. Полученные функции позволяют выполнить вычисление теоретических значений структурной избыточности и характеристик надежности СФК без проведения предварительного моделирования.

Для оценки их эффективности был проведен сравнительный анализ рассчитанных аналитических данных со значениями, полученными на практике, по результатам которого полученные функции продемонстрировали высокую точность в вычислении характеристик схем контроля. При этом разница между значениями составила 1,3 % (для спектрального кода) и 4,7 % (для низкоплотностного кода).

Кроме того, был проведен корреляционный анализ, который показал высокую корреляцию полученных данных с результатами моделирования.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Rao N.P. and Desai M.P. A detailed characterization of errors in logic circuits due to single-event transients // Proc. Euromicro Conf. Digit. Syst. Design. – 2015. – P. 714-721.
2. Saremi M., Privat A., Barnaby H.J., and Clark L.T. Physically based predictive model for single event transients in CMOS gates // IEEE Trans. Electron Devices. – 2016. – Vol. 63, No. 6. – P. 2248-2254.
3. Гаврилов С.В., Иванова Г.А., Рыжова Д.И., Соловьев А.Н., Степковский А.Л. Методы синтеза помехозащищенных комбинационных блоков // Информационные технологии. – 2015. – Т. 21, № 11. – С. 821-826.
4. Гаврилов С.В., Гуров С.И., Жукова Т.Д., Рыжова Д.И. Применение теории кодирования для повышения помехозащищенности комбинационных схем // Информационные технологии. – 2016. – Т. 22, № 12. – С. 931-937.
5. Гаврилов С.В., Жукова Т.Д., Рыжова Д.И. Методы оптимизации схем кодирования на основе диаграмм двоичных решений для синтеза отказоустойчивых микро- и нанoeлектронных схем // VII Всероссийская научно-техническая конференция «Проблемы разработки перспективных микро- и нанoeлектронных систем – 2016»: Сб. научн. тр. / под общ. ред. А.Л. Степковского. – Ч. IV. – С. 158-165.
6. Mahatme N.N. et al. Impact of technology scaling on the combinational logic soft error rate // 2014 IEEE international reliability physics symposium. – IEEE, 2014. – P. 5F. 2.1-5F. 2.6.
7. Yaran T.T.G., Tosun S. Improving combinational circuit resilience against soft errors via selective resource allocation // 2017 IEEE 20th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS). – IEEE, 2017. – P. 12-15.

8. *Mitra, S., McCluskey E.J.* Which Concurrent Error Detection Scheme To Choose? // Proc. International Test Conference. – 2000. – P. 985-994.
9. *Сапожников В.В., Сапожников Вл.В., Ефанов Д.В., Дмитриев В.В.* Новые структуры систем функционального контроля логических схем // Автоматика и телемеханика. – 2017. – №. 2. – С. 128-143.
10. *Сапожников В.В., Сапожников Вл.В., Ефанов Д.В., Пивоваров Д.В.* Синтез систем функционального контроля многовыходных комбинационных схем на основе метода логического дополнения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2017. – №. 41. – С. 69-80.
11. *Ефанов Д.В., Сапожников В.В., Сапожников Вл.В.* Коды с суммированием с фиксированными значениями кратностей обнаруживаемых монотонных и асимметричных ошибок для систем технического диагностирования // Автоматика и телемеханика. – 2019. – №. 6. – С. 121-141.
12. *Хетагуров Я.А., Руднев Ю.П.* Повышение надёжности цифровых устройств методами избыточного кодирования. – М.: Энергия, 1974. – 270 с.
13. *Жукова Т.Д.* Разработка системы автоматизированного проектирования СФК на основе методов избыточного кодирования // Проблемы разработки перспективных микро-и наноэлектронных систем (МЭС). – 2020. – №. 4. – С. 51-57.
14. *Гуров С.И.* Спектральный R-код с проверками на чётность // Прикладная математика и информатика: Тр. факультета Вычислительной математики и кибернетики МГУ им. М.В. Ломоносова. – 2017. – №. 55. – С. 91-96.
15. *Тельпухов Д.В., Жукова Т.Д., Деменева А.И., Гуров С.И.* Схема функционального контроля для комбинационных схем на основе R-кода // Проблемы разработки перспективных микро-и наноэлектронных систем (МЭС). – 2018. – №. 4. – С. 98-104.
16. *Stempkovskiy A.L., Telpukhov D.V., Gurov S.I., Zhukova T.D., Demeneva A.I.* R-code for concurrent error detection and correction in the logic circuits // Young Researchers in Electrical and Electronic Engineering (EIConRus), 2018 IEEE Conference of Russian. – IEEE, 2018. – P. 1430-1433.
17. *Стемковский А.Л., Тельпухов Д.В., Жукова Т.Д., Деменева А.И., Надоленко В.В., Гуров С.И.* Синтез схемы функционального контроля на основе спектрального R-кода с разбиением выходов на группы // Микроэлектроника. – 2019. – Т. 48, №. 4. – С. 284-294.
18. *Стемковский А.Л., Тельпухов Д.В., Деменева А.И., Жукова Т.Д.* Маршрут проектирования схем функционального контроля комбинационных устройств // Вестник РГРТУ. – 2018. – №. 65. – С. 92-98.
19. *Стемковский А.Л., Тельпухов Д.В., Гуров С.И., Жукова Т.Д., Щелоков А.Н., Новиков А.Д.* Синтез СФК на основе LDPC кода с использованием мажоритарного декодирования // Известия ЮФУ. Технические науки. – 2019. – №. 4 (206). – С. 195-206.
20. *Gallager R.* Low-density parity-check codes // IRE Transactions on information theory. – 1962. – Vol. 8, No. 1. – P. 21-28.
21. URL: <https://ddd.fit.cvut.cz/prj/Benchmarks/> (дата обращения: 31.08.2020).

REFERENCES

1. *Rao N.P. and Desai M.P.* A detailed characterization of errors in logic circuits due to single-event transients, *Proc. Euromicro Conf. Digit. Syst. Design.*, 2015, pp. 714-721.
2. *Saremi M., Privat A., Barnaby H.J., and Clark L.T.* Physically based predictive model for single event transients in CMOS gates, *IEEE Trans. Electron Devices*, 2016, Vol. 63, No. 6, pp. 2248-2254.
3. *Gavrilov S.V., Ivanova G.A., Ryzhova D.I., Solov'ev A.N., Stempkovskiy A.L.* Metody sinteza pomekhozashchishchennykh kombinatsionnykh blokov [Methods of noise immunity combinational blocks synthesis], *Informatsionnye tekhnologii* [Information technology], 2015, Vol. 21, No. 11, pp. 821-826.
4. *Gavrilov S.V., Gurov S.I., Zhukova T.D., Ryzhova D.I.* Primenenie teorii kodirovaniya dlya povysheniya pomekhozashchishchennosti kombinatsionnykh skhem [Application of coding theory to improve the noise immunity of combinational circuits], *Informatsionnye tekhnologii* [Information technology], 2016, Vol. 22, No. 12, pp. 931-937.

5. Gavrilov S.V., Zhukova T.D., Ryzhova D.I. Metody optimizatsii skhem kodirovaniya na osnove diagramm dvoichnykh resheniy dlya sinteza otkazoustoychivyykh mikro- i nanoelektronnykh skhem [Optimization methods of coding circuits based on the binary decision diagrams for synthesis of fault-tolerant microand nanoelectronic circuits], *VII Vserossiyskaya nauchno-tehnicheskaya konferentsiya «Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem – 2016»*: Sb. nauchn. tr. [VII all-Russian scientific and technical conference "Problems of development of advanced micro- and nanoelectronic systems-2016": Collection of scientific papers], ed. by A.L. Stempkovskogo. Part IV, pp. 158-165.
6. Mahatme N.N. et al. Impact of technology scaling on the combinational logic soft error rate, *2014 IEEE international reliability physics symposium*. IEEE, 2014, pp. 5F. 2.1-5F. 2.6.
7. Yaran T.T.G., Tosun S. Improving combinational circuit resilience against soft errors via selective resource allocation, *2017 IEEE 20th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS)*. IEEE, 2017, pp. 12-15.
8. Mitra, S., McCluskey E.J. Which Concurrent Error Detection Scheme To Choose?, *Proc. International Test Conference*, 2000, pp. 985-994.
9. Sapozhnikov V.V., Sapozhnikov V.I., Efanov D.V., Dmitriev V.V. Novye struktury sistem funktsional'nogo kontrolya logicheskikh skhem [New structures of the concurrent error detection systems for logic circuits], *Avtomatika i telemekhanika* [Automation and telemechanics], 2017, No. 2, pp. 128-143.
10. Sapozhnikov V.V., Sapozhnikov V.I., Efanov D.V., Pivovarov D.V. Sintez sistem funktsional'nogo kontrolya mnogovykhodnykh kombinatsionnykh skhem na osnove metoda logicheskogo dopolneniya [Synthesis of concurrent error detection systems of multioutput combinational circuits based on Boolean complement method], *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika* [Bulletin of Tomsk state University. Management, computer engineering and computer science], 2017, No. 41, pp. 69-80.
11. Efanov D.V., Sapozhnikov V.V., Sapozhnikov V.I. Kody s summirovaniem s fiksirovannymi znacheniyami kratnostey obnaruzhivaemykh monotonnykh i asimmetrichnykh oshibok dlya sistem tekhnicheskogo diagnostirovaniya [Sum codes with fixed values of multiplicities for detectable unidirectional and asymmetrical errors for technical diagnostics of discrete systems], *Avtomatika i telemekhanika* [Automation and telemechanics], 2019, No. 6, pp. 121-141.
12. Khetagurov YA.A., Rudnev Yu.P. Povyshenie nadezhnosti tsifrovyykh ustroystv metodami izbytochnogo kodirovaniya [Improving the reliability of digital devices methods of redundant coding]. Moscow: Energiya, 1974, 270 p.
13. Zhukova T.D. Razrabotka sistemy avtomatizirovannogo proektirovaniya SFK na osnove metodov izbytochnogo kodirovaniya [Functional control circuits cad system based on redundant coding methods], *Problemy razrabotki perspektivnykh mikro-i nanoelektronnykh sistem (MES)* [Problems of developing advanced micro-and nanoelectronic systems (MES)], 2020, No. 4, pp. 51-57.
14. Gurov S.I. Spektral'nyy R-kod s proverkami na chetnost' [Spectral R-code with parity checks], *Prikladnaya matematika i informatika: Tr. fakul'teta Vychislitel'noy matematiki i kibernetiki MGU imeni M.V. Lomonosova* [Applied mathematics and computer science: Proceedings of the faculty of Computational mathematics and Cybernetics of Lomonosov Moscow state University], 2017, No. 55, pp. 91-96.
15. Tel'pukhov D.V., Zhukova T.D., Demeneva A.I., Gurov S.I. Skhema funktsional'nogo kontrolya dlya kombinatsionnykh skhem na osnove R-koda [Circuit of functional control for combinational circuits based on R-code], *Problemy razrabotki perspektivnykh mikro-i nanoelektronnykh sistem (MES)* [Problems of developing advanced micro-and nanoelectronic systems (MES)], 2018, No. 4, pp. 98-104.
16. Stempkovskiy A.L., Telpukhov D.V., Gurov S.I., Zhukova T.D., Demeneva A.I. R-code for concurrent error detection and correction in the logic circuits, *Young Researchers in Electrical and Electronic Engineering (EConRus), 2018 IEEE Conference of Russian*. IEEE, 2018, pp. 1430-1433.
17. Stempkovskiy A.L., Tel'pukhov D.V., Zhukova T.D., Demeneva A.I., Nadolenko V.V., Gurov S.I. Sintez skhemy funktsional'nogo kontrolya na osnove spektral'nogo R-koda s razbieniem vykhodov na gruppy [Synthesis of a concurrent error detection circuit based on the spectral R-code with the partitioning of outputs into groups], *Mikroelektronika* [Microelectronics], 2019, Vol. 48, No. 4, pp. 284-294.

18. *Stemkovskiy A.L., Tel'pukhov D.V., Demeneva A.I., zhukova T.D.* Marshrut proektirovaniya skhem funktsional'nogo kontrolya kombinatsionnykh ustroystv [Design flow of concurrent error detection schemes for combinational circuits], *Vestnik RGRU* [Vestnik of RSREU], 2018, No. 65, pp. 92-98.
19. *Stemkovskiy A.L., Tel'pukhov D.V., Gurov S.I., Zhukova T.D., SHCHelokov A.N., Novikov A.D.* Sintez SFK na osnove LDPC koda s ispol'zovaniem mazhoritarnogo dekodirovaniya [Synthesis method of fault-tolerant combination circuits with CED based on LDPC code], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2019, No. 4 (206), pp. 195-206.
20. *Gallager R.* Low-density parity-check codes, *IRE Transactions on information theory*, 1962, Vol. 8, No. 1, pp. 21-28.
21. Available at: <https://ddd.fit.cvut.cz/prj/Benchmarks/> (accessed 31 August 2020).

Статью рекомендовал к опубликованию к.т.н. А.В. Коршунов.

Тельпухов Дмитрий Владимирович – Институт проблем проектирования в микроэлектронике Российской академии наук; e-mail: nofrost@inbox.ru; 124365, Москва, Зеленоград, ул. Советская, 3; д.т.н.; руководитель отдела методологии проектирования интегральных схем.

Жукова Татьяна Дмитриевна – e-mail: zhukova_t@ippm.ru; инженер-исследователь.

Щелоков Альберт Николаевич – e-mail: schan@ippm.ru; к.ф.-м.н.; с.н.с.

Telpukhov Dmitry Vladimirovich – Institute for Design Problems in Microelectronics of Russian Academy of Sciences, e-mail: nofrost@inbox.ru; 3, Sovetskaya street, Moscow, 124365, Russia; dr. of eng. sc.; head of department of integrated circuits design methodology.

Zhukova Tatyana Dmitrievna – e-mail: zhukova_t@ippm.ru; research engineer.

Schelokov Albert Nikolaevich – e-mail: schan@ippm.ru; cand. of phis.-math. sc.; senior researcher.

Раздел III. Системы управления и нелинейная динамика

УДК 517.977.56

DOI 10.18522/2311-3103-2020-4-178-191

А.Г. Клово, Г.В. Куповых, И.А. Ляпунова

МАТЕМАТИЧЕСКАЯ ЗАДАЧА ОБ ОПТИМАЛЬНОМ УПРАВЛЕНИИ СТРУНОЙ

Общепринято, что задачи оптимального управления или задачи проектирования системы определяют для заданного объекта или системы объектов управления закон или некоторую управляющую последовательность действий, которые обеспечивают максимум или минимум заданной совокупности критериев качества системы. При этом может рассматриваться задача быстрогодействия, т.е. задача о приведении системы в заданное состояние за наименьшее время. Также изучаются задачи минимизации заданного функционала при фиксированном времени управления системой. Оптимальное управление тесно связано с выбором наиболее рациональных режимов управления сложными объектами. Проблеме управления посвящено много работ, кроме того в настоящее время подобными исследованиями занимаются известные математические школы. В задачах с сосредоточенными параметрами исследуемые системы описываются обыкновенными дифференциальными уравнениями или их системами. В этом случае важную роль в таком исследовании играет принцип максимума Понтрягина. Для уравнений с частными производными говорят о системах с распределенными параметрами. В данной работе исследуется возможность синтеза оптимального управления одной системой с распределенными параметрами. Рассмотрена модель колебаний струны под воздействием управляющих функций в граничных условиях. Показана роль выбора минимизируемого функционала в создании возможностей синтеза оптимального управления. В этом случае осуществляется поиск управляющего воздействия в каждой точке временного промежутка, что приводит к возможности построения его в явном виде. Сформулированы условия, при которых существуют всюду оптимальные управления в соответствующих функциональных пространствах. В конкретной постановке задачи всюду оптимальное управление построено в явном виде.

Управление; колебания струны; синтез; оптимальное управление; управляющая функция; краевая задача.

A.G. Klovo, G.V. Kupovykh, I.A. Lyapunova

THE MATHEMATICAL PROBLEM OF OPTIMAL CONTROL OF THE STRING

It is generally accepted that optimal control problems or system design problems determine for a given object or system of control objects a law or a certain control sequence of actions that provide a maximum or minimum of a given set of system quality criteria. In this case, the speed problem can be considered, i.e. the problem of bringing the system to a given state in the shortest time. We also study the problems of minimizing a given functional for a fixed time of system management. Optimal control is closely related to the choice of the most rational modes for managing complex objects. A lot of works has been devoted to the problem of control, in addition, well-known mathematical schools are currently engaged in such research. In problems with concentrated parameters, the systems under study are described by ordinary differential equations or their systems. In this case, the Pontryagin maximum principle plays an important role in this study. For partial differential equations, we talk about systems with distributed parameters. In this

paper, we investigate the possibility of synthesizing optimal control of a single system with distributed parameters. A model of string oscillation under the influence of control functions under boundary conditions is considered. The role of the choice of the functional to be minimized in creating opportunities for the synthesis of optimal control. In this case, the control action is searched for at each point of the time interval, which leads to the possibility of constructing it explicitly. The conditions for the existence of optimal control everywhere in the corresponding functional spaces are formulated. In a specific statement of the problem, everywhere optimal control is explicitly constructed.

Control; string oscillation; synthesis; optimal control; control function; boundary value problem.

Введение. Задачи оптимального управления колебаниями струны рассматривались в ряде работ, например, в [1–7]. В работах В.А. Ильина, Е.И. Моисеева и их учеников рассмотрена серия задач управления струной. В частности, изучены различные типы граничных управлений струной на ее границе. При этом решения рассматривались в разнообразных функциональных пространствах. Рассмотрена серия задач, в которых оптимальное управление гиперболическими уравнениями может быть найдено в явном виде.

В работах [8–17] изучены условия, при которых существуют классические решения задач для гиперболических уравнений. При этом построены функциональные пространства, в которых существуют обобщенные решения и оптимальные управления.

Французские математики [18, 19] показали специфику задач управления колебаниями струны и минимальное время, за которое система может быть переведена в заданное состояние. Однако функционал, который минимизируется в этих задачах оказался, как показано в работах [20, 21], оказался не самым удачным для задач оптимального управления гиперболическими уравнениями.

В этой работе показано, что использование стандартного минимизируемого функционала не обеспечивает возможность построения оптимального управления на отдельных временных промежутках.

Целью исследования является установление возможности синтеза оптимального управления, т.е. возможности нахождения оптимального управления на отдельных временных участках. Найдены достаточные условия, при которых оптимальное управление может быть найдено на отдельных участках, вплоть до поточечного поиска. При таком подходе показана возможность построения оптимального управления в явном виде.

Начально краевая задача с управлением. Рассмотрим задачу, связанную с математическими моделями колебаний струны. Если управление струной происходит на ее границе, то можно рассмотреть вопрос о приведении струны в заданное состояние. В силу корректности обратных задач для гиперболического уравнения, в частности, для уравнения колебаний струны, задача о ее приведении в заданное состояние равносильна задаче о приведении в нулевое состояние, при котором равны нулю отклонения точек струны от состояния покоя и их скорости.

Пусть $u(t, x)$ – отклонение точки струны с координатой $x \in [0; l]$ в момент времени $t \in [0; T]$. Тогда функция $u(t, x)$ удовлетворяет уравнению

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \quad (1)$$

начальным

$$u|_{t=0} = \varphi(x), \quad \frac{\partial u}{\partial t}|_{t=0} = \psi(x) \quad (2)$$

и граничным условиям

$$u|_{x=0} = p(t), \quad u|_{x=l} = 0. \quad (3)$$

Задача заключается в том, чтобы подобрать управляющую функцию $u|_{x=0} = p(t)$ таким образом, чтобы в заданный момент времени $t = T$ значения отклонений точек струны от положения равновесия и скорости точек струны стали минимальными в соответствующей метрике.

В книге [2] рассматривается задача минимизации в этот момент времени функционала

$$J_1(p) = \int_0^l \left((u(T, x))^2 + \left(\frac{\partial u}{\partial t} \Big|_{t=T} \right)^2 \right) dx. \quad (4)$$

В этой работе показано, что существует минимальный промежуток времени, в данном случае равный $T = \frac{2}{l}$, за который система может быть приведено в любое

заданное состояние. Если временной промежуток меньше, то можно ставить задачу о минимизации функционала (4) или более подходящего. В этом случае надо формулировать теоремы существования и единственности решения поставленной задачи и оптимального управления.

Построение обобщенных решений поставленных задач. При решении задач, связанных с оптимальным управлением системами, как правило, приходится рассматривать последовательности решений и управляющих функций. Для существования предельных решений необходимо перейти к обобщенным решениям поставленных задач.

Следуя О.А. Ладыженской [1], рассмотрим вначале классические решения, которые гарантированно существуют, если выполнены условия $\varphi(x) \in C^2[0; l]$, $\psi(x) \in C^1[0; l]$, $p(t) \in C^2[0; T]$ гладкости и согласования: $\varphi(0) = p(0)$, $p'(0) = \psi(0)$, $\varphi(l) = 0$, $\psi'(l) = 0$.

Пространством $W_2^1(D)$ мы будем называть замыкание рассматриваемой совокупности функций по норме

$$\|u(t, x)\|_{W_2^1(D)} = \left(\iint_D \left(u^2 + \left(\frac{\partial u}{\partial t} \right)^2 + \left(\frac{\partial u}{\partial x} \right)^2 \right) \right)^{\frac{1}{2}}. \quad (5)$$

Под этим понимается следующее. Если функции $\{u_n(t, x)\}$, $n \in N$ в каждой точке области $D = [0; l] \times [0; T]$ сходятся к предельной функции $u(t, x)$, у которой существует конечная норма (5) и при этом выполнено $\lim_{n \rightarrow +\infty} \|u(t, x) - u_n(t, x)\|_{W_2^1(D)} = 0$, то мы будем говорить, что $u(t, x) \in W_2^1(D)$.

Теперь перейдем к определению обобщенного решения задачи (1)–(3). Для этого рассмотрим множество $\dot{C}_1(D)$ таких достаточно гладких функций, которые могут быть отличны от 0 только лишь в области $D_\delta = [\delta, l - \delta] \times [0, T]$. При этом положительное δ может быть различным для разных функций. Рассмотрим замыкание введенных функций $\dot{C}_1(D)$ в пространстве $W_2^1(D)$ и обозначим его через $\dot{C}_1(D)$.

Пусть $v(t, x) = p(t) \cdot (1 - x/l)$ и введем функцию

$$w(t, x) = u(t, x) - v(t, x). \quad (6)$$

Функция $w(t, x)$ равна 0 при $x = 0$ и при $x = l$ и является решением задачи

$$\frac{\partial^2 w}{\partial t^2} - \frac{\partial^2 w}{\partial x^2} = -p''(t)(1 - x/l), \quad (7)$$

$$w|_{t=0} = \varphi(x) - \varphi(0) \cdot (1 - x/l), \quad \frac{\partial w}{\partial t}|_{t=0} = \psi(x) - \psi(1) \cdot (1 - x/l), \quad (8)$$

$$w|_{x=0} = 0, \quad w|_{x=l} = 0. \quad (9)$$

Естественно, что для решения задачи (1)–(3) достаточно решить задачу (7)–(9).

Определение 1. Функция $w(t, x) \in \overset{\circ}{C}_1(D)$ такая, что ее предельные значения при $t \rightarrow 0$ равны $\varphi(x) - \varphi(0) \cdot (1 - x/l)$ и для каждой функции $\sigma(t, x) \in \overset{\circ}{C}_2(D)$ выполнено интегральное тождество

$$\iint_D \left(\frac{\partial w}{\partial t} \frac{\partial \sigma}{\partial t} - \frac{\partial w}{\partial x} \frac{\partial \sigma}{\partial x} + p'(t)(1 - x/l) \frac{\partial \sigma}{\partial t} \right) dx dt = \int_0^1 \psi(x) \sigma(0, x) dx \quad (10)$$

Называется обобщенным решением задачи (7)–(9).

Соотношение (10) получено умножением (7) на $\sigma(t, x)$ и интегрирования по области $D = [0; l] \times [0; T]$ с учетом дополнительных условий.

Невозможность синтеза оптимального управления для функционала (4). Рассмотрим конкретную задачу, когда оптимальное управление может быть построено в явном виде. Если взять $l = 2$, $\varphi(x) = 0$, $\psi(x) = \pi \sin \frac{\pi x}{2}$, то функция

$u_1(t, x) = 2 \sin \frac{\pi x}{2} \cdot \sin \frac{\pi t}{2}$ является, очевидно, решением задачи (1)–(3) с $p(t) = 0$.

Также несложно проверить, что функция

$$u_2(t; x) = \begin{cases} 0, & t - x \leq 0, \\ p(t - x), & t - x > 0 \end{cases}$$

является решением при $0 \leq t \leq 2$ задачи (1)–(3) с нулевыми начальными условиями (2) и граничными условиями (3). Поэтому на этом временном промежутке функция $u(t; x) = u_1(t; x) + u_2(t; x)$ является задачи (1)–(3) при указанных условиях.

Пусть $T = 1$, тогда для полученного решения выполнено:

$$u(1; x) = \begin{cases} 2 \sin \frac{\pi x}{2} + p(1 - x), & x \leq 1, \\ 2 \sin \frac{\pi x}{2}, & x > 1, \end{cases} \quad (10)$$

$$\frac{\partial u}{\partial t}|_{t=1} = \begin{cases} p'(1 - x), & x \leq 1, \\ 0, & x > 1. \end{cases} \quad (11)$$

Если эти значения в (4), то мы приходим к задаче минимизации функционала

$$\int_0^1 \left(\left(2 \sin \frac{\pi x}{2} + p(1 - x) \right)^2 + (p'(1 - x))^2 \right) dx + \int_1^2 \left(2 \sin \frac{\pi x}{2} \right)^2 dx.$$

Сделаем замену $1 - x = t$, после чего потребуется найти наименьшее значение интеграла

$$\int_0^1 \left(\left(2 \cos \frac{\pi t}{2} + p(t) \right)^2 + (p'(t))^2 \right) dt. \quad (12)$$

Для минимизации функционала вида $J = \int_a^b F(x, f(x), f'(x)) dx$ используется уравнение Лагранжа $\frac{\partial F}{\partial f} - \frac{d}{dx} \frac{\partial F}{\partial f'} = 0$, которое для (12) примет вид:

$$\int_0^1 \left((p(t))^2 + (p'(t) - \pi \sin \pi t)^2 \right) dt.$$

Тем самым для нахождения оптимального управления на временном промежутке $t \in [0; 1]$ мы приходим с учетом условий согласования к задаче Коши

$$p''(t) - p(t) = 2 \cos \frac{\pi t}{2}, \quad p(0) = p'(0) = 0,$$

решение которой задачи запишется в виде

$$p(t) = \frac{8}{\pi^2 + 4} \left(\cos \frac{\pi t}{2} - 1 \right).$$

Найдем теперь оптимальное управление системой на временном промежутке $t \in [0; 2]$. В этом случае при $T = 2$ выполнены условия:

$$u(2; x) = p(2 - x), \quad (13)$$

$$\left. \frac{\partial u}{\partial t} \right|_{t=1} = -\pi \sin \frac{\pi x}{2} + p'(2 - x). \quad (14)$$

Здесь мы приходим к нахождению наименьшего значения величины

$$\int_0^2 \left(p^2(2 - x) + \left(-\pi \sin \frac{\pi x}{2} + p'(2 - x) \right)^2 \right) dx.$$

И после замены $2 - x = t$ уравнение Лагранжа приводит нас к задаче Коши

$$p''(t) - p(t) = 2 \cos \frac{\pi t}{2}, \quad p(0) = p'(0) = 0,$$

имеющее решение

$$p(t) = \frac{2\pi^2}{\pi^2 - 4} \left(\cos \frac{\pi t}{2} - 1 \right).$$

Мы видим, что оптимальные управления на участках $t \in [0; 1]$ и $t \in [0; 2]$ не совпадают на общем временном участке. Это означает, что при минимизации функционала (4) синтез оптимального управления невозможен.

Пример синтеза оптимального управления для модифицированного функционала. Для задачи, рассмотренной в предыдущем пункте, рассмотрим оптимальное управление, минимизирующее функционал

$$J_2(p) = \int_0^1 \left(\left(\frac{\partial u}{\partial x} \Big|_{t=T} \right)^2 + \left(\frac{\partial u}{\partial t} \Big|_{t=T} \right)^2 \right) dx \quad (15)$$

в те же промежутки времени.

Решение при $T = 1$, указанные в (10), (11) при подстановке в функционал (15) дают интеграл

$$\int_0^1 \left(\left(\pi \cos \frac{\pi x}{2} - p'(1-x) \right)^2 + (p'(1-x))^2 \right) dx + \int_1^2 \left(\pi \cos \frac{\pi x}{2} \right)^2 dx.$$

После такой же замены $1 - x = t$ мы придем к минимизации функционала

$$\int_0^1 \left(\left(\pi \sin \frac{\pi t}{2} - p'(t) \right)^2 + (p'(t))^2 \right) dt. \quad (16)$$

С помощью уравнения Лагранжа мы приходим к задаче Коши

$$p''(t) = \frac{\pi^2}{4} \cos \frac{\pi t}{2}, \quad p(0) = p'(0) = 0,$$

решение которой задачи запишется в виде

$$p(t) = 1 - \cos \frac{\pi t}{2}.$$

При $T = 2$ подстановка полученного решения и (15) приведет нас к минимизации интеграла

$$\int_0^2 \left((-p'(2-x))^2 + \left(-\pi \sin \frac{\pi x}{2} + p'(2-x) \right)^2 \right) dx$$

и после замены $2 - x = t$ получим:

$$\int_0^2 \left((p'(t))^2 + \left(-\pi \sin \frac{\pi t}{2} + p'(t) \right)^2 \right) dt. \quad (17)$$

С помощью уравнения Лагранжа мы приходим к той же самой задаче Коши

$$p''(t) = \frac{\pi^2}{4} \cos \frac{\pi t}{2}, \quad p(0) = p'(0) = 0,$$

решение которой также запишется в виде

$$p(t) = 1 - \cos \frac{\pi t}{2}.$$

Следовательно, в данном случае оптимальные управления на общем участке совпадают.

Эквивалентное определение оптимального управления. При изучении возможности синтеза оптимального управления для минимизируемого функционала (15) формализуем постановку задачи. Начнем с двух вариантов определения оптимального управления.

Определение 2. Управление $p^*(t)$ будем называть оптимальным среди множества допустимых управлений P на отрезке $[0; T]$, если оно реализует минимум функционала (15).

Для двух допустимых на отрезке $[0; T]$ управлений $p^*(t)$, $p(t)$ и соответствующих им обобщенных решений $u^*(t, x)$ и $u(t, x)$ задачи (1)–(3) рассмотрим функцию

$$K_T(u^*, u) = \int_0^l \left(\frac{\partial u^*}{\partial t} \cdot \left(\frac{\partial u^*}{\partial t} - \frac{\partial u}{\partial t} \right) + \frac{\partial u^*}{\partial x} \cdot \left(\frac{\partial u^*}{\partial x} - \frac{\partial u}{\partial x} \right) \right) \Big|_{t=T} dx. \quad (18)$$

Определение 3. Управление $p^*(t)$ будем называть оптимальным среди множества допустимых управлений P , если для произвольного допустимого управления $p(t)$ и соответствующих решений $u^*(t, x)$ и $u(t, x)$ справедливо соотношение

$$K_T(u^*, u) \leq 0. \quad (19)$$

Введем обозначение

$$f_T(u^*, u) = \int_0^l \left(\frac{\partial u^*}{\partial t} \cdot \frac{\partial u}{\partial t} + \frac{\partial u^*}{\partial x} \cdot \frac{\partial u}{\partial x} \right) \Big|_{t=T} dx$$

и отметим, что $K_T(u^*, u) = f_T(u^*, u - u^*)$, $J_2(p) = f_T(u, u)$. В последнем соотношении $p(t)$ – произвольное управление, а $u(t, x)$ – соответствующее решение задачи (1)–(3). Как связаны между собой определения 2 и 3 показывает следующая теорема.

Теорема 1. Пусть множество допустимых управлений P является выпуклым, тогда определения 2 эквивалентно определению 3.

Доказательство. Рассмотрим управление $p^*(t)$, оптимальное по определению 2. Докажем, что оно оптимально по определению 3. В самом деле, если это не так, то найдется управление $p(t) \in P$ такое что справедливо неравенство $K_T(u^*, u) > 0$. Рассмотрим при $0 \leq \lambda \leq 1$ управление

$$\bar{p}(t) = p^*(t) + \lambda(p(t) - p^*(t)), \quad (20)$$

которое является допустимым в силу выпуклости P . Соответствующее решение задачи (1)–(3) имеет вид $\bar{u}(t) = u^*(t, x) + \lambda(u(t, x) - u^*(t, x))$ и для этого решения функционал (15) примет вид

$$J_2(\bar{p}(t)) = J(p^*(t)) - 2\lambda K_T(u^*, u) + \lambda^2 \int_0^l \left(\left(\frac{\partial u^*}{\partial t} - \frac{\partial u}{\partial t} \right)^2 + \left(\frac{\partial u^*}{\partial x} - \frac{\partial u}{\partial x} \right)^2 \right) \Big|_{t=T} dx. \quad (21)$$

При $\lambda = 0$ выражение (21) равно $J(p^*(t))$. В то же время само по себе выражение (21) является квадратным трехчленом относительно λ , вершина которого λ_0 положительна. Следовательно, на промежутке $(0; \lambda_0]$ квадратный трехчлен (21) принимает значения, меньшие, чем $J(p^*(t))$. Это противоречит предположению, что $p^*(t)$ оптимально по определению 2. Поэтому из оптимальности по определению 2 следует оптимальность по определению 3.

Докажем вторую часть теоремы. Пусть теперь $p^*(t)$ является оптимальным в смысле определения 3. Тогда при $\lambda = 1$ выполнено равенство $\bar{p}(t) = p^*(t) + 1 \cdot (p(t) - p^*(t)) = p(t)$ и $J(p^*(t)) < J(p(t))$ в силу (19) и (21). Теорема доказана.

Критерий возможности синтеза оптимального управления. Возможность синтеза оптимального управления, т.е. нахождение такого управления на различных промежутках времени тесно связана с понятием всюду оптимального управления.

Определение 4. Всяду оптимальным на промежутке $[0; T]$ называется управление $\bar{p}(t)$, оптимальное на каждом промежутке $[0; \tau]$ при $0 < \tau \leq T$.

Уточним, что понимается под синтезом управления. Пусть задано управление $p_{12}(t)$ на отрезке $[t_1; t_2]$ и управление $p_{23}(t)$ на отрезке $[t_2; t_3]$. Если при этом $p_{12}(t_2) = p_{23}(t_2)$, то синтезированным этими управлениями мы будем называть управление $p_{13}(t)$ на отрезке $[t_1; t_3]$, обладающее свойством

$$(p_{13}(t))' = \begin{cases} (p_{12}(t))', & t_1 < t < t_2, \\ (p_{23}(t))', & t_2 < t < t_3. \end{cases} \quad (22)$$

Пусть рассматривается задача оптимального управления системой на временном промежутке $[0; T]$. Рассмотрим следующие условия:

1. Множеством допустимых управлений P – выпуклое множество, содержащее нулевое управление.
2. Под оптимальным понимается управление, при котором достигается минимум заданного квадратичного, положительно определенного функционала.
3. Существует число $\delta > 0$, обладающее тем свойством, что на любом промежутке $[t_1; t_2]$, принадлежащем $[0; T]$, таком, что $t_2 - t_1 < \delta$, существует единственное оптимальное управление.
4. Если $p_{0,t_1}^*(t)$ является оптимальным управлением на промежутке $[0; t_1]$, $p_{t_1,t_2}^*(t)$ – оптимальным на $[t_1; t_2]$, то синтезированное управление $p_{0,t_2}^*(t)$ будет оптимальным на объединенном промежутке $[0; t_2]$.

Теорема 2. Пусть для задачи (1)–(3) и произвольного $T > 0$ выполнены условия 1.–4. Тогда для нее существует единственное всюду оптимальное управление $\bar{p}(t)$.

Доказательство. Точками $\{t_i\}$, $i = 0, 1, \dots, N$ разделим отрезок $[0; T]$ на N отрезков таким образом, что каждый из них по длине меньше $\delta > 0$ из пункта 3. Построим оптимальное управление на каждом из этих промежутков. На первом из них берем начальные условия (2), а затем на каждом последующем промежутке в качестве начальных берем конечные условия предыдущего промежутка. Докажем, что полученное после синтеза таких управление $\bar{p}(t)$ является искомым всюду оптимальным управлением.

Возьмем произвольные отрезки $[a;b]$ и $[b;c]$, где $0 \leq a < b \leq T$. Пусть $p_{ab}^*(t)$ является оптимальным управлением на $[a;b]$ при некоторых начальных условиях в точке a , а $p_{bc}^*(t)$ – оптимальным управлением на $[b;c]$ с согласованными начальными условиями в b . При этом соответствующими решениями задачи (1)–(3) являются, функции $u_{ab}^*(t, x)$ и $u_{bc}^*(t, x)$. При этом синтезированное на $[a;c]$ управление обозначим как $p^*(t)$ и соответствующее ему решение – $u^*(t, x)$.

Рассмотрим произвольное управление $p(t)$ на промежутке $[a;c]$ с теми же начальными условиями и соответствующее решение $u(t, x)$. Заметим, что с учетом введенных выше обозначений:

$$K_c(u^*, u) = f_c(u^*, u^* - u) = f_c(u^*, u_0^* - u_0) + f_c(u^*, u^* - u_0^* - (u - u_0)). \quad (23)$$

В качестве $u_0(t, x)$ возьмем решение, соответствующее управлению $p_0(t)$, определяемому соотношением

$$(p_0(t))' = \begin{cases} (p(t))', & a \leq t \leq b, \\ 0, & b < t \leq c, \end{cases}$$

а в качестве $u_0^*(t, x)$ возьмем решение, соответствующее управлению $p_0^*(t)$, определяемому соотношением

$$(p_0^*(t))' = \begin{cases} (p^*(t))', & a \leq t \leq b, \\ 0, & b < t \leq c. \end{cases}$$

Из оптимальности $p^*(t)$ следует неравенство

$$f_c(u^*, u^* - u_0^* - (u - u_0)) \leq 0$$

и нам осталось доказать, что первое слагаемое в (23) $f_c(u^*, u_0^* - u_0)$ не положительно. Пусть это не так и $f_c(u^*, u_0^* - u_0) > 0$. Тогда силу условия 1. допустимое управление $\tilde{p}_\lambda(t) = p_\lambda^*(t) + \lambda(p_\lambda(t) - p_\lambda^*(t))$ при $0 \leq \lambda \leq 1$ и соответствующее решение $\tilde{u}(t, x)$ порождают квадратный трехчлен $f_c(\tilde{u}, \tilde{u}) = f_c(u^*, u^*) - 2\lambda f_c(u^*, u_0^* - u_0) + \lambda^2 f_c(u_0^* - u_0, u_0^* - u_0)$, который при достаточно малых положительных λ удовлетворяет неравенству $f_c(\tilde{u}, \tilde{u}) < f_c(u^*, u^*)$, что противоречит условию 4.

В итоге $K_c(u^*, u) \leq 0$ и, с учетом условия 3 полученное управление будет всюду оптимальным. Теорема доказана.

Существование всюду оптимального управления в задаче о колебаниях струны. Несложно проверить, что для задачи (1)–(3) условия 1.-3. выполнены. Проверим выполнение свойства 4. при условии минимизации функционала (15).

Пусть функция $p_{0,t_1}(t)$ является в этом плане оптимальной в области $D_{0,t_1} = [0, l] \times [0, t_1]$, а функция $p_{t_1,t_2}(t)$ – оптимальна в $D_{t_1,t_2} = [0, l] \times [t_1, t_2]$. Нам надо убедиться в том, что синтезированное управление оптимально в объединенной области $D_{0,t_2} = [0, l] \times [0, t_2]$.

Умножим обе части уравнения (1) на $\frac{\partial u}{\partial t}$ и проинтегрируем по области D_{t_1,t_2} ,

$0 \leq t_1 < t_2 \leq T$. Заметим, что

$$\begin{aligned} \iint_{D_{t_1,t_2}} \frac{\partial^2 u}{\partial t^2} \cdot \frac{\partial u}{\partial t} dt dx &= \frac{1}{2} \iint_{D_{t_1,t_2}} \frac{\partial}{\partial t} \left(\left(\frac{\partial u}{\partial t} \right)^2 \right) dt dx = \frac{1}{2} \int_0^l \left(\frac{\partial u}{\partial t} \right)^2 \Big|_{t=t_1}^{t=t_2} dx, \\ \iint_{D_{t_1,t_2}} \frac{\partial^2 u}{\partial x^2} \cdot \frac{\partial u}{\partial t} dt dx &= \int_{t_1}^{t_2} \left(\frac{\partial u}{\partial x} \cdot \frac{\partial u}{\partial t} \right) \Big|_{x=0}^{x=l} dt - \iint_{D_{t_1,t_2}} \frac{\partial u}{\partial x} \cdot \frac{\partial^2 u}{\partial t \partial x} dt dx = \\ &= \int_{t_1}^{t_2} \left(\frac{\partial u}{\partial x} \cdot \frac{\partial u}{\partial t} \right) \Big|_{x=0}^{x=l} dt - \frac{1}{2} \iint_{D_{t_1,t_2}} \frac{\partial}{\partial t} \left(\left(\frac{\partial u}{\partial x} \right)^2 \right) dt dx = \int_{t_1}^{t_2} \left(\frac{\partial u}{\partial x} \cdot \frac{\partial u}{\partial t} \right) \Big|_{x=0}^{x=l} dt - \frac{1}{2} \int_0^l \left(\frac{\partial u}{\partial x} \right)^2 \Big|_{t=t_1}^{t=t_2} dx, \end{aligned}$$

поэтому

$$\int_0^l \left(\left(\frac{\partial u}{\partial t} \right)^2 + \left(\frac{\partial u}{\partial x} \right)^2 \right) \Big|_{t=t_1}^{t=t_2} dx = 2 \int_{t_1}^{t_2} \left(\frac{\partial u}{\partial x} \cdot \frac{\partial u}{\partial t} \right) \Big|_{x=0}^{x=l} dt.$$

Отсюда, с учетом нулевых граничных условий на правом конце струны и управляющей функции на левом, имеем

$$\int_0^l \left(\left(\frac{\partial u}{\partial t} \right)^2 + \left(\frac{\partial u}{\partial x} \right)^2 \right) \Big|_{t=t_1}^{t=t_2} dx = -2 \int_{t_1}^{t_2} p'(t) \left(\frac{\partial u}{\partial x} \right) \Big|_{x=0} dt. \quad (24)$$

Полученная формула демонстрирует закон сохранения энергии при отсутствии управляющей функции. Левая часть равна удвоенной разности энергии струны в моменты времени $t = t_2$ и $t = t_1$. В этой ситуации нам необходимо исследовать

величину $p'(t) \left(\frac{\partial u}{\partial x} \right) \Big|_{x=0}$ в формуле (24). Пусть в результате оптимального управления

$p_{0,t_1}(t)$ решение задачи (1)-(3) при $0 \leq x \leq l$ приняло значения

$$u \Big|_{t=t_1} = \varphi_{t_1}(x), \quad \frac{\partial u}{\partial t} \Big|_{t=t_1} = \psi_{t_1}(x).$$

Если величина $t_2 - t_1$ достаточно мала, искомые значения решения при $x = 0$ на промежутке $[t_1, t_2]$ зависят от значений $\varphi_{t_1}(x)$, $\psi_{t_1}(x)$ на промежутке $0 \leq x \leq t_2 - t_1$ и от управляющей функции на этом временном промежутке. Решение в области $[0, t_2 - t_1] \times [t_1, t_2]$ представим в виде суммы $u(t, x) = u_1(t, x) + u_2(t, x)$. Здесь

$$u_1(t, x) = \frac{\varphi_{t_1}(x)}{2} + \frac{1}{2} \cdot \int_0^{t-x} \psi_{t_1}(\xi) d\xi, \quad u_2(t, x) = p_{t_1, t_2}(t - x),$$

откуда на этом временном участке

$$\left(\frac{\partial u}{\partial x} \right) \Big|_{x=0} = \frac{(\varphi_{t_1}(t))'}{2} - \frac{\psi_{t_1}(t)}{2} - (p_{t_1, t_2}(t))', \quad \frac{\partial u}{\partial t} \Big|_{x=0} = (p_{t_1, t_2}(t))'.$$

Следовательно, исследуемая величина имеет вид:

$$p'(t) \left(\frac{\partial u}{\partial x} \right) \Big|_{x=0} = \left(\frac{(\varphi_{t_1}(t))'}{2} - \frac{\psi_{t_1}(t)}{2} - (p_{t_1, t_2}(t))' \right) \cdot (p_{t_1, t_2}(t))',$$

и ее наибольшее значение принимается при

$$(p_{t_1, t_2}(t))' = \frac{(\varphi_{t_1}(t))' - \psi_{t_1}(t)}{4}. \quad (25)$$

Так как возмущения при свободных колебаниях струны возмущения распространяются вдоль характеристик, то уменьшение функционала (15) на временном промежутке $[t_1, t_2]$ при достаточно малой величине $t_2 - t_1$ никак не связано с его минимизацией на промежутке $[0, t_1]$. Следовательно, условие 4 выполнено и всюду оптимальное управление задачи (1)–(3) существует.

Вид всюду оптимального управления в задаче о колебаниях струны. Перейдем к построению в явном виде оптимального управления струной в задаче (1)–(3).

Отметим, что в координатах $\xi = x - t$, $\eta = x + t$ уравнение (1) примет вид:

$$\frac{\partial^2 u}{\partial \xi \partial \eta} = 0. \quad (26)$$

Если в этих координатах обе части (26) проинтегрировать по прямоугольнику с вершинами $M(\xi_1; \eta_1)$, $N(\xi_1 + \xi_1; \eta_1)$, $P(\xi_1 + \xi_1; \eta_1 + \eta_2)$, $P(\xi_1; \eta_1 + \eta_2)$, то мы придем к соотношению

$$u(\xi_1; \eta_1) + u(\xi_1 + \xi_1; \eta_1 + \eta_2) = u(\xi_1 + \xi_1; \eta_1) + u(\xi_1; \eta_1 + \eta_2).$$

Отсюда следует, что производная по одной из характеристик не меняется при движении вдоль другой характеристики в той области, где выполнено (26).

Формула (25) показывает вид всюду оптимального управления на временном участке $[0, T]$ при $T = l$. Здесь надо учесть, что скорость распространения волны в данном случае равна 1. На следующем промежутке надо учитывать, что вследствие отражения от правой границы у вектора $\{\varphi'; \psi\}$ меняет знак вторая компонента. Поэтому на временном промежутке $[0, T]$ при $T = 2l$ всюду оптимальное управление определяется формулой

$$(p_{0, 2l}(t))' = \frac{(\varphi(l - |l - t|))' - \text{sign}(l - t)\psi(l - |l - t|)}{4}. \quad (27)$$

С помощью управляющей функции, производная от которой определяется формулой (27), произвольное начальное состояние струны за время $2l$ приводится в состояние покоя.

Заключение. Построена и исследована задача оптимального управления колебаниями струны. Исследованы функциональные пространства, в которых существуют обобщенные решения поставленных задач.

Доказан ряд теорем о существовании и единственности соответствующих оптимальных управлений. При этом сформулированы достаточные условия, при которых оптимальное управление можно автономно находить на различных временных промежутках. Это позволяет свести задачу к поиску такого управления в каждой точке временного промежутка. В итоге удалось получить всюду оптимальное управление в явном виде.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ильин В.А., Моисеев Е.И. Оптимизация граничных управлений колебаниями струны // Успехи математических наук. – 2005. – 60:6 (366). – С. 89-114.
2. Моисеев Е.И., Холмеева А.А., Фролов А.А. Граничное управление смещением процессом колебаний при граничном условии типа торможения за время, меньшее критического // Итоги науки и техники. Серия «Современная математика и ее приложения. Тематические обзоры». – 2019. – № 160. – С. 74-84.
3. Моисеев Е.И., Фролов А.А. Граничное управление процессом колебаний струны при условии сопротивления среды на правом конце за время, меньшее критического // Дифференциальные уравнения. – 2019. – Т. 55, № 4. – С. 555-566.
4. Моисеев Е.И., Моисеев Т.Е., Попиванов Н.И., Холмеева А.А. Разрешимость нелокальных краевых задач для уравнения смешанного типа с различными краевыми условиями // Дифференциальные уравнения. – 2018. – Т. 54, № 10. – С. 541-552.
5. Ильин В.А. Избранные труды: в 2 т. Т. 1. – М.: Изд-во ООО «Макс-пресс», 2008. – 727 с.
6. Моисеев Е.И., Холмеева А.А. Оптимальное граничное управление смещением на одном конце струны при заданной упругой силе на другом конце // Тр. ИММ УрО РАН. – 2011. – 17:2. – С. 151-158.
7. Попов А.Ю. Минимизация вариации производной граничного управления гашением колебаний струны с одним закрепленным концом // Доклады РАН. – 2006. – 409:1. – С. 22-25
8. Ладыженская О.А. Краевые задачи математической физики. – М.: Наука, 1973. – 407 с.
9. Ладыженская О.А. Смешанная задача для гиперболических уравнений. – М.: Гостехиздат, 1953. – 282 с.
10. Соболев С.Л. Некоторые применения функционального анализа в математической физике. – М.: Наука, 1988. – 336 с.
11. Михайлов В.П. Дифференциальные уравнения в частных производных. – М.: Наука, 1976. – 448 с.
12. Петровский И.Г. Лекции об уравнениях с частными производными. – М.; Л.: ГИТТЛ, 1953. – 360 с.
13. Стеклов В.А. Основные задачи математической физики. – М.: Наука, 1983. – 432 с.
14. Крылов А.Н. О некоторых дифференциальных уравнениях математической физики, имеющих приложения в технических вопросах. – М.; Л.: ГИТТЛ, 1950. – 368 с.
15. Ильин В.А. О разрешимости смешанных задач для гиперболического и параболического уравнений // Успехи математических наук. – 1960. – 15:2 (92). – С. 97-154.
16. Егоров Ю.В. Некоторые вопросы теории оптимального управления // Журнал вычислительной математики и математической физики. – 1963. – Т. 3, № 5. – С. 887-904.
17. Егоров Ю.В. Необходимые условия оптимальности управления в банаховых пространствах // Матем. сб. – 1964. – Т. 64 (106), № 1. – С. 79-101.
18. Латтес Р., Лионс Ж.-Л. Метод квазиобращения и его приложения. – М.: Мир, 1970. – 335 с.
19. Лионс Ж.-Л. Оптимальное управление системами, описываемыми уравнениями с частными производными. – М.: Мир, 1972. – 414 с.

20. Клово А.Г., Гончаров А.В. Условия всюду-оптимальности управления одной системой с распределенными параметрами // Матер. IV Международной научной конференции «Донецкие чтения 2019: образование, наука, инновации, культура и вызовы современности». Т. 1. Физико-математические и технические науки. Ч. 1. – Донецк: ДонНУ, 2019. – С. 27-30.
21. Клово А.Г., Куповых Г.В., Ляпунова И.А. О возможности синтеза оптимального управления колебаниями струны // Международная научная конференции по дифференциальным уравнениям и динамическим системам. Тезисы докладов. – Владимир: Изд-во ВлГУ, 2020. – С 72-73.

REFERENCES

1. Il'in V.A., Moiseev E.I. Optimizatsiya granichnykh upravleniy kolebaniyami struny [Optimization of boundary controls for string vibrations], *Uspekhi matematicheskikh nauk* [Advances in mathematical Sciences], 2005, 60:6 (366), pp. 89-114.
2. Moiseev E.I., Kholomeeva A.A., Frolov A.A. Granichnoe upravlenie smeshcheniem protsessom kolebaniy pri granichnom uslovii tipa tormozheniya za vremya, men'shee kriticheskogo [Boundary control of the displacement of the oscillation process under a boundary condition of the braking type for a time less than the critical one], *Itogi nauki i tekhniki. Seriya «Sovremennaya matematika i ee prilozheniya. Tematicheskie obzory»* [Results of science and technology. Series "Modern mathematics and its applications. Thematic review"], 2019, No. 160, pp. 74-84.
3. Moiseev E.I., Frolov A.A. Granichnoe upravlenie protsessom kolebaniy struny pri uslovii soprotivleniya srede na pravom kontse za vremya, men'shee kriticheskogo [Boundary control of the string oscillation process under the condition of the medium resistance at the right end for a time less than the critical one], *Differentsial'nye uravneniya* [Differential equations], 2019, Vol. 55, No. 4, pp. 555-566.
4. Moiseev E.I., Moiseev T.E., Popivanov N.I., Kholomeeva A.A. Razreshimost' nelokal'nykh kraevykh zadach dlya uravneniya smeshannogo tipa s razlichnymi kraevymi usloviyami [Solvability of non-local boundary value problems for a mixed-type equation with different boundary conditions], *Differentsial'nye uravneniya* [Differential equations], 2018, Vol. 54, No. 10, pp. 541-552.
5. Il'in V.A. Izbrannye trudy [Selected works]: in 2 vol. Vol. 1. Moscow: Izd-vo OOO «Makspress», 2008, 727 p.
6. Moiseev E.I., Kholomeeva A.A. Optimal'noe granichnoe upravlenie smeshcheniem na odnom kontse struny pri zadannoy uprugoy sile na drugom kontse [Optimal boundary control of displacement at one end of the string for a given elastic force at the other end], *Tr. IMM UrO RAN* [Trudy Instituta Matematiki i Mekhaniki UrO RAN], 2011, 17:2, pp. 151-158.
7. Popov A.Yu. Minimizatsiya variatsii proizvodnoy granichnogo upravleniya gasheniem kolebaniy struny s odnim zakreplennym kontsom [Minimization of the variation of the derivative of the boundary control for damping the vibrations of a string with one fixed end], *Doklady RAN* [Reports of the Russian Academy of Sciences], 2006, 409:1, pp. 22-25
8. Ladyzhenskaya O.A. Kraevye zadachi matematicheskoy fiziki [Boundary value problems of mathematical physics]. Moscow: Nauka, 1973, 407 p.
9. Ladyzhenskaya O.A. Smeshannaya zadacha dlya giperbolicheskikh uravneniy [Mixed problem for hyperbolic equations]. Moscow: Gostekhizdat, 1953, 282 p.
10. Sobolev S.L. Nekotorye primeneniya funktsional'nogo analiza v matematicheskoy fizike [Some applications of functional analysis in mathematical physics]. Moscow: Nauka, 1988, 336 p.
11. Mikhaylov V.P. Differentsial'nye uravneniya v chastnykh proizvodnykh [Differential equations in partial derivatives]. Moscow: Nauka, 1976, 448 p.
12. Petrovskiy I.G. Lektsii ob uravneniyakh s chastnymi proizvodnymi [Lectures on partial differential equations]. Moscow; Leninsrad: GITTL, 1953, 360 p.
13. Steklov V.A. Osnovnye zadachi matematicheskoy fiziki [Main problems of mathematical physics]. Moscow: Nauka, 1983, 432 p.
14. Krylov A.N. O nekotorykh differentsial'nykh uravneniyakh matematicheskoy fiziki, imeyushchikh prilozheniya v tekhnicheskikh voprosakh [On some differential equations of mathematical physics that have applications in technical issues]. Moscow; Leninsrad: GITTL, 1950, 368 p.

15. *Il'in V.A.* O razreshimosti smeshannykh zadach dlya giperbolicheskogo i parabolicheskogo uravneniy [On the solvability of mixed problems for hyperbolic and parabolic equations], *Uspekhi matematicheskikh nauk* [Advances in mathematical Sciences], 1960, 15:2 (92), pp. 97-154.
16. *Egorov Yu.V.* Nekotorye voprosy teorii optimal'nogo upravleniya [Some questions of optimal control theory], *Zhurnal vychislitel'noy matematiki i matematicheskoy fiziki* [Journal of computational mathematics and mathematical physics], 1963, Vol. 3, No. 5, pp. 887-904.
17. *Egorov Yu.V.* Neobkhodimye usloviya optimal'nosti upravleniya v banakhovykh prostranstvakh [Necessary conditions for optimal control in Banach spaces], *Matem. sb.* [Mathematical collection], 1964, Vol. 64 (106), No. 1, pp. 79-101.
18. *Lattes R., Lions Zh.-L.* Metod kvaziobrashcheniya i ego prilozheniya [The method of quasi-conversion and its applications]. Moscow: Mir, 1970, 335 p.
19. *Lions Zh.-L.* Optimal'noe upravlenie sistemami, opisyvaemymi uravneniyami s chastnymi proizvodnymi [Optimal control of systems described by partial differential equations]. Moscow: Mir, 1972, 414 p.
20. *Klovo A.G., Goncharov A.V.* Usloviya vsyudu-optimal'nosti upravleniya odnoy sistemoy s raspredelennymi parametrami [Conditions for everywhere-optimal control of a single system with distributed parameters], *Mater. IV Mezhdunarodnoy nauchnoy konferentsii «Donetskie chteniya 2019: obrazovanie, nauka, innovatsii, kul'tura i vyzovy sovremennosti». T. 1. Fiziko-matematicheskie i tekhnicheskije nauki* [Proceedings of the IV International scientific conference "Donetsk read 2019: education, science, culture, innovations and challenges". Vol. 1. Physico-mathematical and technical Sciences. Part 1]. Donetsk: DonNU, 2019, pp. 27-30.
21. *Klovo A.G., Kupovykh G.V., Lyapunova I.A.* O vozmozhnosti sinteza optimal'nogo upravleniya kolebaniyami struny [On the possibility of synthesizing optimal control of string vibrations], *Mezhdunarodnaya nauchnaya konferentsii po differentsial'nyim uravneniyam i dinamicheskim sistemam. Tezisy dokladov* [International scientific conference on differential equations and dynamical systems. Thesis of reports]. Vladimir: Izd-vo VIGU, 2020, pp. 72-73.

Статью рекомендовал к опубликованию д.ф.-м.н., профессор А.Х. Аджиев.

Куповых Геннадий Владимирович – Южный федеральный университет; e-mail: kupovykh@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 89289543642; д.ф.-м.н.; профессор.

Клово Александр Георгиевич – e-mail: klovo_ag@mail.ru; тел.: 89281221064; к.ф.-м.н.; доцент.

Ляпунова Ирина Артуровна – e-mail: ialyapunova@sfedu.ru; тел.: 89034310026; к.ф.-м.н.; доцент.

Kupovykh Gennady Vladimirovich – Southern Federal University; e-mail: kupovykh@sfedu.ru; 44, Nekrasovsky street, Taganrog, 347928, Russia; phone: +79289543642; dr. of phys. and math. sc.; professor.

Klovo Alexander Georgievich – e-mail: klovo_ag@mail.ru; phone: +79281221064; cand. of phys. and math. sc.; associate professor.

Lyapunova Irina Arturovna – e-mail: ialyapunova@sfedu.ru; phone: +79034310026; cand. of phys. and math. sc.; associate professor.

А.С. Черепанцев

СТАТИСТИЧЕСКИЕ МЕТОДЫ ОЦЕНКИ СВЯЗНОСТИ ДИНАМИЧЕСКИХ СИСТЕМ ПО ОТДЕЛЬНОЙ ВРЕМЕННОЙ ПРОЕКЦИИ

На основе подходов нелинейной динамики к оценке инвариантов динамической системы рассмотрена возможность определения степени связности различных динамических систем. Под динамической связностью исследуемых систем понимается число общих компонент в системах, определяющих временную эволюцию наблюдаемых проекций. Предложенный метод протестирован на модельных динамических системах и использован при анализе поведения сложных динамических систем наблюдаемых в геофизике – кажущегося электрического сопротивления по двум ортогональным направлениям и относительным вертикальным смещениям поверхности. Используемые в расчетах данные длительных режимных наблюдений в сейсмически активном регионе интересны имеющимися фактами чувствительности к напряженно-деформированному состоянию геофизической среды. Предполагая параметр состояния среды общей компонентой наблюдаемых динамических процессов различной природы, проведена оценка числа общих компонент систем на основе предложенной методики. В работе предложен статистический метод выделения отдельных отсчетов синхронного изменения вариаций динамических параметров наблюдаемого комплекса геофизических полей. Предполагая нестационарный характер формирования динамической системы при наличии большого числа воздействующих внешних факторов, актуальным является определение временных интервалов синхронизации свойств динамических систем при появлении доминирующего воздействия. Результатом применения разработанного метода является вывод о синхронизации вариаций корреляционной размерности объемной деформации на различных временных масштабах в фазе возникновения сильных сейсмических событий.

Динамическая система; размерность динамической системы; инварианты динамической системы.

A.S. Cherepantsev

STATISTICAL METHODS FOR EVALUATING THE CONNECTIVITY OF DYNAMIC SYSTEMS BASED ON A SEPARATE TIME PROJECTION

Based on the approaches of nonlinear dynamics to the estimation of invariants of a dynamical system, the possibility of determining the degree of coupling of various dynamical systems is considered. The dynamic coupling of the studied systems is understood as the number of common components in the systems that determine the time evolution of the observed projections. The proposed method has been tested on model dynamic systems and used to analyze the behavior of complex dynamic systems observed in geophysics – apparent electrical resistance in two orthogonal directions and relative vertical surface displacements. The data of long regime observations used in the calculations in the seismically active region are interesting by the available facts of sensitivity to the stress-strain state of the geophysical medium. Assuming a parameter of the state of the medium as a common component of the observed dynamic processes of various nature, the number of common components of the systems is estimated based on the proposed methodology. The paper proposes a statistical method for finding individual samples of synchronous changes in variations of the dynamic parameters of the observed number of geophysical fields. Assuming the unsteady nature of the formation of a dynamic system in the presence of a large number of acting external factors, it is relevant to determine the time intervals for synchronizing the properties of dynamic systems when a dominant effect appears. The result of the application of the developed method is the conclusion about synchronization of variations in the correlation dimension of volumetric deformation at different time scales in the phase of the occurrence of strong seismic events.

Dynamical system; dimension of a dynamical system; invariants of a dynamical system.

Введение. При анализе сложных динамических процессов, включающих большое число независимых факторов и соответственно степеней свободы, большое внимание уделяется выявлению связи исследуемого явления с наблюдаемыми процессами иной природы. Наличие подобной связи помогает как в построении адекватной модели явления и возможности прогноза, так и активного воздействия с целью управления ее поведением. Проводимые научные наблюдения геофизических полей различной природы являются характерным примером исследования сложных нелинейных динамических систем, включающих в себя как общие факторы воздействия, так и факторы, присущие исключительно наблюдаемому физическому явлению.

Исходя из представления о наблюдаемом геофизическом явлении как эволюции некоторой нелинейной динамической системы, в настоящее время получили активное развитие подходы нелинейной динамики [1–5]. Проведенные исследования позволили выявить закономерности и свойства нелинейных диссипативных структур таких сложных природных процессов, как атмосферных вихревых образований развитой турбулентности, сейсмоакустических процессов в земной коре и многих других [6–10].

Аппарат анализа сложных нелинейных динамических систем опирается на разработанные математические методы восстановления многомерной динамической системы по временной выборке лишь единственной компоненты и вычислительные методы оценки основных инвариантных параметров неизвестной динамической системы [11, 12]. Данные оценки основываются на анализе эволюции траекторий в фазовом пространстве. К основным оцениваемым инвариантным параметрам системы относятся корреляционная размерность d_2 и энтропия K_2 , размерность динамической системы m , показатели Ляпунова λ_i .

В настоящей работе рассмотрена возможность выделения в различных наблюдаемых геофизических процессах общих компонент формирующих их динамических систем. При этом под динамической связностью таких процессов подразумевается число независимых компонент систем, являющихся общими. Следует отметить, что подобная связность в общем случае не обязательно подразумевает влияние изменений одного процесса на изменение другого. Она лишь указывает на присутствие в обеих системах компонент, сформированных одинаковыми динамическими системами. При этом, они могут иметь различную природу и не быть связанными между собой.

Оценка числа общих компонент динамических систем. Параметр размерности вложения фазового пространства динамической системы m является одним из инвариантных параметров нелинейной динамической системы. Параметр m определяет число независимых переменных, входящих в неизвестную динамическую систему и формирующую наблюдаемую одномерную проекцию m -мерного фазового пространства. Он может быть оценен по наблюдаемой временной выборке лишь одной из компонент. Общий подход к оценке параметра m представляет собой достаточно общую процедуру. Она включает в себя оценку некоторого инвариантного параметра динамической системы при различных размерностях вложения фазового пространства $\tilde{m}=1,2,3,\dots$. Оценка \tilde{m}_{\min} , начиная с которой расчетный параметр остается неизменным с ростом \tilde{m} и есть оценка m : $m=\tilde{m}_{\min}$.

Традиционно в качестве расчетного инвариантного параметра динамической системы принимается величина корреляционной размерности d_2 [13–15]. В работе [9] предложена еще одна возможность оценки m с использованием в качестве инвариантного параметра скорости расхождения двух близких траекторий в фазовом \tilde{m} -мерном пространстве.

Методика расчета заключается в оценке средней величины увеличения расстояний между близкими точками $\rho(\vec{y}(t_i), \vec{y}(t_j))$ точек \tilde{m} -мерного фазового пространства за заданный интервал времени $\Delta t : \frac{\langle \rho(\tilde{m}) \rangle}{\varepsilon}$. В данном соотношении знаменатель определяет расстояние между точками фазовой траектории в начальный момент времени.

Выше рассмотрена оценка m по одномерной выборке неизвестной динамической системы. Тем более данный параметр может быть оценен при наличии временных вариаций s компонент ($1 \leq s \leq m$).

Так исходная система $\vec{y}(t) = \{y_1(t), y_2(t), \dots, y_m(t)\}$ и система $\vec{x}(t) = \left\{ \begin{array}{l} y_i(t), \dots, y_{i+s}(t), y_i(t + \Delta t), \dots, y_{i+s}(t + \Delta t), \dots \\ \dots, y_i(t + (m/s - 1)\Delta t), \dots, y_{i+s}(t + (m/s - 1)\Delta t) \end{array} \right\}$ ($i=1, \dots, m/s$), обладают инвариантным значением параметра m .

Рассмотрим временные вариации нескольких динамических систем и задачу оценки числа независимых переменных, присутствующие во всех исследуемых динамических системах. В геофизике актуальность такой задачи связана с выделением параметра общего управляющего деформационного процесса, являющегося фактором, участвующим в формировании вариаций физических полей различной природы - поля деформаций и напряжений, кажущегося электрического сопротивления, уровня воды в скважине и т.д.

Рассмотрим задачу совместного анализа нескольких динамических систем и оценки числа независимых переменных, описывающих наблюдаемые вариации отдельной проекции каждой системы, но являющихся при этом общим параметром. Актуальность такой задачи в геофизике определяется выделением общего управляющего тектонического процесса по наблюдениям физических полей различной природы - поля деформаций, электрического сопротивления, уровня воды в скважине и т.д.

Пусть $\mathbf{X}(t) = \{x_1(t), x_2(t), \dots, x_p(t)\}$, $\mathbf{Y}(t) = \{y_1(t), y_2(t), \dots, y_n(t)\}$ - временные вариации двух различных динамических систем с параметром $m=p$ и $m=n$ соответственно. Пусть m - размерность общего фазового пространства двух динамических систем $(x_k(t), y_l(t))$. А m_x, m_y - размерности отдельных систем $x_k(t)$, $y_l(t)$.

В случае отсутствия общих параметров и независимости систем:

$$m = m_x + m_y. \quad (1)$$

Если рассматриваемые вариации являются проекциями общей динамической системы, то:

$$m = \max(m_x, m_y). \quad (2)$$

При наличии общих компонент в каждой системе: $\max(m_x, m_y) < m < m_x + m_y$. Их число определяется как:

$$s = (m_x + m_y) - m. \quad (3)$$

С целью проверки предложенной методики рассмотрим расчет параметра m известных модельных систем: системы Лоренца:

$$\begin{aligned}\dot{x} &= -\sigma x + \sigma y, \\ \dot{y} &= rx - y - xz, \\ \dot{z} &= xy - bz,\end{aligned}\tag{4}$$

и системы затухающего осциллятора с детерминированным поведением:

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= -xz + x \\ \dot{z} &= xy.\end{aligned}\tag{5}$$

В качестве исходных данных для расчета взят вектор $\vec{w}(t) = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$, где u, v – x -компоненты систем (4),(5).

На рис. 1 представлены зависимости $\frac{\langle \rho(\tilde{m}) \rangle}{\varepsilon}$. Указанные доверительные интервалы выбраны на уровне 0.7 в предположении нормального распределения значений расстояния ρ в евклидовой метрике. Для системы Лоренца $m=3$. Это следует из того, что при $\tilde{m} \geq m$ величина $\frac{\langle \rho(\tilde{m}) \rangle}{\varepsilon}$ перестает значимо изменяться. При этом величина расхождения близких траекторий (1.3 ± 0.3) оказывается близкой к теоретическому значению, определяемых старшим показателем Ляпунова λ_1 :

$$\frac{\rho}{\varepsilon} = \exp(\lambda_1 \Delta t).$$

Зависимости $\frac{\langle \rho(\tilde{m}) \rangle}{\varepsilon}$ для системы (5) и совместной системы, полученной по временному ряду двумерного вектора даны на рис. 1.

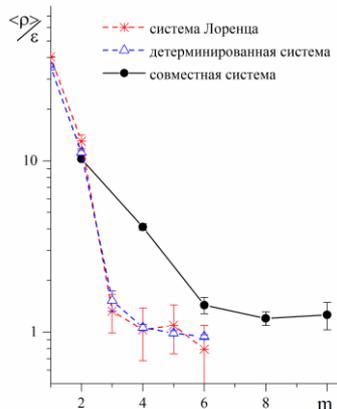


Рис. 1. Оценки размерности модельных динамических систем. $\varepsilon = 0.07, \Delta t = 0.3$

Исходя из независимости систем (4) и (5) оценка $m = 6$ для совместной системы согласуется с величиной размерности общего фазового пространства. При этом предельное значение $\frac{\langle \rho(\tilde{m}) \rangle}{\varepsilon} \approx 1.3 \pm 0.1$ определяется параметром скорости разбегания фазовых траекторий системы Лоренца [16].

Для анализа связности геофизических процессов различной природы рассмотрено поле кажущегося электрического сопротивления и вариаций относительного вертикального смещения поверхности в тектонически активной зоне геофизического полигона в КНР.

Для исследования динамической связности процессов проведен расчет параметра m двумерных векторов кажущихся электрических сопротивлений по двум направлениям N-S и E-W, а также пары процессов различной природы: кажущегося сопротивления N-S и относительного вертикального смещения. Результаты расчета показаны на рис. 2 и 3. соответственно.

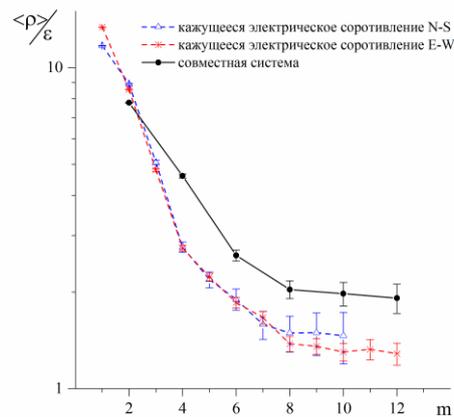


Рис. 2. Оценка связности динамических систем кажущегося электрического сопротивления в ортогональных направлениях

Как следует из полученных зависимостей на рис. 2, размерности фазового пространства динамических систем вариаций кажущегося электрического сопротивления оказываются близкими к значению $m=7-8$. Данные оценки близки к оценкам, полученным по расчету корреляционной размерности [17]. Используемый метод оценки m динамической системы вертикальных смещений дает $m=6$, что также согласуется с данными [17].

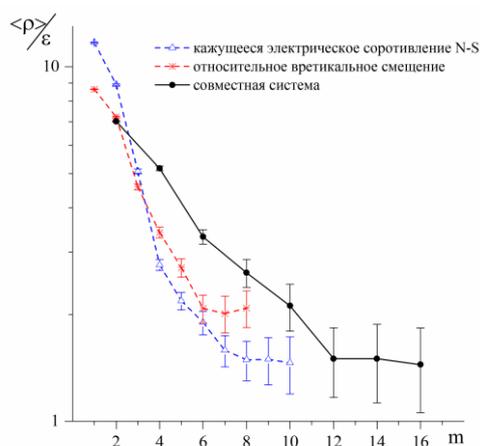


Рис. 3. Оценка связности динамических систем кажущегося электрического сопротивления и вертикального смещения

Для совместной динамической системы кажущихся электрических сопротивлений в двух ортогональных направлениях полученная оценка размерности $m_{N-S,E-W}=8$ показывает, что наблюдения кажущегося электрического сопротивления в перпендикулярных направлениях не являются независимыми процессами. Исходя из соотношения (2) их можно считать различными компонентами одной и той же динамической системы.

Значение параметра размерности $m_{N-S,D}=12$, для совместной системы кажущегося сопротивления N-S и относительного смещения указывает на частичную связность процессов. С учетом значений $m_{N-S}=8$ и $m_D=6$, в каждой динамической системе присутствуют и общие параметры. Оценка их числа в соответствии с (3) дает величину $s=2$. Полученный результат указывает на существование как общих параметров, имеющих значительные пространственные масштабы, так и индивидуальных параметров, характерных для каждого процесса.

Оценка корреляции временных вариаций параметров динамических систем. Еще одним важным инструментом при анализе связи различных динамических процессов во времени является выделение временных участков синхронного изменения амплитуд исследуемых динамических параметров и возможного присутствия общего управляющего фактора.

Рассмотрим задачу выделения участков с синхронизацией вариаций динамического параметра в различных точках наблюдения. Используем корреляционный анализ для выделения подобной линейной связи.

Оценим вклад связи отдельной пары отсчетов (x_i, y_i) с помощью расчета коэффициента корреляции $\tilde{r}_{x,y}$, вычисленной по выборке, в которой исключен этот отсчет. При положительной корреляции процессов в точке отсчета: $|\tilde{r}_{x,y}| < |r_{x,y}|$, где $r_{x,y}$ – коэффициент корреляции по всей выборке. Таким образом отклонение $\tilde{r}_{x,y}$ от $r_{x,y}$ может служить характеристикой вклада (x_i, y_i) в параметр связи процессов.

Важной характеристикой при использовании данного параметра является определение значимости оценки величины отклонения коэффициента корреляции неполной выборки от ее значения по полной выборке. Оценку значимости удобнее рассмотреть для выборочной ковариации двух случайных рядов [18,19]:

$$k(t_i) = \text{cov}_{XY}(t_1, \dots, t_N) - \text{cov}_{XY}(t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_N), \quad i = 1, \dots, N. \quad (6)$$

Оценка ковариации независимых случайных величин $x(t_i), y(t_i)$, имеющих стандартное нормальное распределение: $\text{cov}_{XY}(t_1, \dots, t_N) = \frac{1}{N} \sum_{j=1}^N x(t_j) y(t_j)$.

Рассмотрим распределение статистики (6). Соотношение (6) представим в виде:

$$k(t_i) = \frac{1}{N} \sum_j x(t_j) y(t_j) - \frac{1}{N-1} \sum_{j \neq i} x(t_j) y(t_j), \quad i, j = 1, \dots, N$$

или:

$$k(t_i) = \frac{1}{N} x(t_i) y(t_i) - \frac{1}{N(N-1)} \sum_{j \neq i} x(t_j) y(t_j) = k1(t_i) + k2(t_i).$$

Для $k1(t_i) = \frac{1}{N} x(t_i) y(t_i)$ функция распределения определяется соотношением [20]:

$$f(k1) = \frac{N}{\pi} K_0(N \cdot |k1|),$$

где $K_0(z) = \int_0^{\infty} \frac{\cos(z\xi)}{\sqrt{\xi^2 + 1}} d\xi$ – модифицированная функция Бесселя второго рода нулевого порядка.

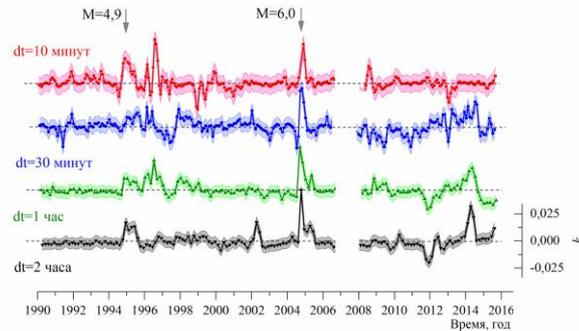


Рис. 4. Вариации параметра связи k корреляционных размерностей динамических систем объемной деформации при различных интервалах дискретизации dt

Для оценки вклада второго слагаемого учтем соотношение для среднеквадратичных отклонений:

$$\sigma(k2) = \frac{\sigma(k1)}{\sqrt{N-1}}. \quad (7)$$

Таким образом, вклад в дисперсию $k(t_i)$ второго слагаемого оказывается в N раз меньше вклада первого слагаемого. То есть для выборок $N > 100$ второе слагаемое слабо влияет на характер распределения и им можно пренебречь.

Практическая реализация данного метода рассмотрена на примере анализа временных участков изменений корреляционной размерности d_2 вариаций объемной деформации в тектонически активном регионе Северной Америки.

Особенностью рассмотренной методики является возможность сопоставления связи процессов на различных временных масштабах. При наличии временного ряда с постоянным временным интервалом дискретизации, стандартная оценка корреляции в скользящем временном окне возможна лишь на младших временных масштабах. При переходе к старшим масштабам временное окно содержит число отсчетов, недостаточное для получения надежной оценки.

На рис. 4 показана временная изменчивость синхронности вариаций параметра d_2 на различных временных масштабах, определяемых интервалом дискретизации dt . Выбранный уровень значимости $\alpha = 0.3$ дает возможность выделить участки перестройки динамических систем в режим синхронной чувствительности к напряженному состоянию, определяемому временными интервалами сильных землетрясений в регионе наблюдений. Магнитуды землетрясений и их время указаны в верхней части рис. 4. Как следует из полученных зависимостей, уменьшение размерности, сопровождающее сильное землетрясение, носит синхронный характер и длится в течении нескольких месяцев.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Анищенко В.С., Астахов В.В., Вадивасова Т.Е., Нейман А.Б., Стрелкова Г.И., Шиманский-Гайер Л. Нелинейные эффекты в хаотических и стохастических системах. – М.-Ижевск: Институт компьютерных исследований, 2003. – 544 с.
2. Badii R., Broggi G., Derighetti B. et al. Dimension Increase in Filtered Chaotic Signals // *Physical Review Letters*. – 1988. – Vol. 60 (11). – P. 979.
3. Кузнецов С.П. Динамический хаос. – М.: Физматлит, 2001. – 295 с.
4. Черепанцев А.С. Характеристики динамической системы геофизических полей на различных временных масштабах // *Физика Земли*. – 2018. – № 4, Приложение. – С. s3-s19.
5. Cherepantsev A.S. Dimension of the phase space of a dynamic system from geophysical data // *Izvestiya. Physics of the Solid Earth*. – 2007. – Vol. 43, No. 12. – P. 1047-1055.
6. Кейлис-Борок В.И., Кособоков В.Г., Мажкенов С.А. О подобии в пространственном распределении сейсмичности // *Вычислительная сейсмология*. – 1989. – № 22. – С. 28-40.
7. Черепанцев А.С. Выделение проекции динамической системы по данным наблюдений поля деформаций // *Физика Земли*. – 2008. – № 2. – С. 39-58.
8. Черепанцев А.С. Выделение динамической составляющей в вариациях геофизических полей на основе сходимости выборочного среднего // *Физика Земли*. – 2008а. – № 11. – С. 31-46.
9. Takens F. Detecting strange attractors in turbulence // *Lect. Notes in Math*. – 1981. – P. 898.
10. Li Q., Nyland E., Is the Dynamics of the Lithosphere Chaotic? // *Nonlinear Dynamics and Predictability of Geophysical Phenomena*. – 1994. – Geophysical Monograph 83, IUUG. Vol. 18. – P. 37-41.
11. Gibson J.F., Farmer J.D., Casdagli M., Eubank S. An analytic approach to practical state space reconstruction // *Physica D*. – 1992. – Vol. 57. – P. 1-30.
12. Parker T., Chua L. Practical numerical algorithms for chaotic systems. – New York: Springer-Verlag, 1989. – 348 p.
13. Grassberger P., Procaccia I., On the Characterization of Strange Attractors // *Phys.Rev. Lett*. – 1983. – Vol. 50. – P. 346-349.
14. Табор М. Хаос и интегрируемость в нелинейной динамике. – М.: Эдиториал УРСС, 2001. – 320 с.
15. Шустер Г. Детерминированный хаос. – М.: Мир, 1988. – 240 с.
16. Малинецкий Г.Г., Потанов А.Б. Современные проблемы нелинейной динамики. – М.: Эдиториал УРСС, 2002. – 360 с.
17. Смирнов В.Б., Пономарев А.В., Qian Jiadong, Черепанцев А.С. Ритмы и детерминированный хаос в геофизических временных рядах // *Физика Земли*. – 2005. – № 6. – С. 6-28.
18. Бриллинджер Д. Временные ряды. Обработка данных и теория. – М.: Мир, 1980. – 536 с.
19. Кендалл М., Стюарт А. Т. 2. Статистические выводы и связи. – М.: Наука, 1973. – 899 с.
20. Левин Б.Р. Теоретические основы статистической радиотехники. – М.: Радио и связь, 1989. – 656 с.

REFERENCES

1. Anishchenko V.S., Astakhov V.V., Vadivasova T.E., Neyman A.B., Strelkova G.I., SHimanskiy-Gayer L. Nelineynye efekty v khaoticheskikh i stokhasticheskikh sistemakh [Non-linear effects in chaotic and stochastic systems]. Moscow.-Izhevsk: Institut komp'yuternykh issledovaniy, 2003, 544 p.
2. Badii R., Broggi G., Derighetti B. et al. Dimension Increase in Filtered Chaotic Signals, *Physical Review Letters*, 1988, Vol. 60 (11), pp. 979.
3. Kuznetsov S.P. Dinamicheskiy khaos [Dynamic Chaos]. Moscow: Fizmatlit, 2001, 295 p.
4. Cherepantsev A.S. Kharakteristiki dinamicheskoy sistemy geofizicheskikh poley na razlichnykh vremennykh masshtabakh [Characteristics of a Dynamic System of Geophysical Fields on Different Time Scales], *Fizika Zemli* [Phisica Zemli], 2018, No. 4, Appendix, pp. s3-s19.
5. Cherepantsev A.S. Dimension of the phase space of a dynamic system from geophysical data, *Izvestiya. Physics of the Solid Earth*, 2007, Vol. 43, No. 12, pp. 1047-1055.
6. Keylis-Borok V.I., Kosobokov V.G., Mazhkenov S.A. O podobii v prostranstvennom raspredelenii seysmichnosti [On similarity in the spatial distribution of seismicity], *Vychislitel'naya seysmologiya* [Vichislitelnaya Seismologia], 1989, No. 22, pp. 28-40.

7. Cherepantsev A.S. Vydelenie proektsii dinamicheskoy sistemy po dannym nablyudeniya polya deformatsiy [Determination of the Phase Projection of a Dynamic System from Strain Field Observations], *Fizika Zemli* [Physica Zemli], 2008, No. 2, pp. 39-58.
8. Cherepantsev A.S. Vydelenie dinamicheskoy sostavlyayushchey v variatsiyakh geofizicheskikh poley na osnove skhodimosti vyborochnogo srednego [Extraction of a dynamic component from variations in geophysical fields using the convergence of a sample average], *Fizika Zemli* [Physica Zemli], 2008a, No. 11, pp. 31-46.
9. Takens F. Detecting strange attractors in turbulence, *Lect. Notes in Math.*, 1981, pp. 898.
10. Li Q., Nyland E., Is the Dynamics of the Lithosphere Chaotic?, *Nonlinear Dynamics and Predictability of Geophysical Phenomena*, 1994. Geophysical Monograph 83, IUUG Vol. 18, pp. 37-41.
11. Gibson J.F., Farmer J.D., Casdagli M., Eubank S. An analytic approach to practical state space reconstruction, *Physica D*, 1992, Vol. 57, pp. 1-30.
12. Parker T., Chua L. Practical numerical algorithms for chaotic systems. New York: Springer-Verlag, 1989, 348 p.
13. Grassberger P., Procaccia I., On the Characterization of Strange Attractors, *Phys.Rev. Lett.*, 1983, Vol. 50, pp. 346-349.
14. Tabor M. Khaos i integriruemost' v nelineynoy dinamike [Chaos and integrability in nonlinear dynamics]. Moscow: Editorial URSS, 2001, 320 p.
15. Shuster G. Determinirovanny khaos [Deterministic Chaos]. Moscow: Mir, 1988, 240 p.
16. Malinetskiy G.G., Potapov A.B. Sovremennyye problemy nelineynoy dinamiki [Modern problems of nonlinear dynamics]. Moscow: Editorial URSS, 2002, 360 p.
17. Smirnov V.B., Ponomarev A.V., Qian Jiadong, Cherepantsev A.S. Ritmy i determinirovanny khaos v geofizicheskikh vremennykh ryadakh [Rhythms and Determined Chaos in Geophysical Time Series], *Fizika Zemli* [Physica Zemli], 2005, No. 6, pp. 6-28.
18. Brillindzher D. Vremennyye ryady. Obrabotka dannykh i teoriya [Time series: data analysis and theory]. Moscow: Mir, 1980, 536 p.
19. Kendall M., Styuart A. T. 2. Statisticheskie vyvody i svyazi [Vol. 2. Statistical conclusions and links]. Moscow: Nauka, 1973, 899 p.
20. Levin B.R. Teoreticheskie osnovy statisticheskoy radiotekhniki [Theoretical Foundations of Statistical Radio Engineering]. Moscow: Radio i svyaz', 1989, 656 p.

Статью рекомендовал к опубликованию д.ф.-м.н., профессор Г.В. Куповых.

Черепанцев Александр Сергеевич – Южный федеральный университет; e-mail: acherpantsev@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371663; к.ф.-м.н.; доцент.

Cherepantsev Alexandr Sergeevich – Southern Federal University; e-mail: acherpantsev@sfedu.ru; 44, Nekrasovsky, Taganrog, 347928, Russia; phone: +78634371663; cand. of phis.-math. sc.; associate professor.

УДК 681.51

DOI 10.18522/2311-3103-2020-4-200-211

Е.Н. Обухова

ПРИМЕНЕНИЕ МЕТОДА ИНТЕГРАЛЬНОЙ АДАПТАЦИИ ДЛЯ СИНТЕЗА АДАПТИВНЫХ ЗАКОНОВ УПРАВЛЕНИЯ ПНЕВМОПРИВОДОМ В УСЛОВИЯХ ГАРМОНИЧЕСКОГО ВОЗМУЩЕНИЯ

Активное использование электропневматических систем в различных сферах промышленной автоматизации обусловлено такими достаточно высокими эксплуатационными показателями пневмопривода как надежность, быстрдействие, низкая стоимость, доступность использования в условиях высокой влажности, а так же во взрыво- и пожароопасных средах. В статье приведен краткий анализ отечественных и зарубежных научных работ, посвященных разработке различных методов управления пневматической сис-

темой, в которых ставится задача синтеза эффективных законов управления обладающих адапционными свойствами к внешним возмущениям. Целью данной работы является разработка адаптивного нелинейного синергетического закона управления для подавления возмущающего воздействия, которое было задано и аддитивно введено в математическую модель в виде гармонической функции. Синтез адаптивного закона управления проводился посредством метода интегральной адаптации, входящего в концепцию синергетической теории управления. Полученные результаты компьютерного моделирования подтверждают адапционные свойства полученных нелинейных синергетических законов управления и достижения поставленной технологической цели управления – перемещение штока в заданное положение в условиях гармонического возмущения. Одним из важных этапов анализа представленных в данной работе результатов является проведение экспериментальных исследований, которые позволяют проверить работоспособность полученных аналитическим путем синтезированных нелинейных синергетических законов управления на учебно-экспериментальном стенде пневматических приводов вертикального и горизонтального перемещения компании Camozzi. Для практической реализации полученных синергетических законов управления было произведено программирование контроллера в инструментальной среде разработки программ для промышленной автоматизации CoDeSys на графическом языке функциональных блок-диаграмм FBD (Function Block Diagram) МЭК 61131-3 программирования.

Электропневматическая система; инвариантные многообразия; фазовые переменные; гармоническое возмущение; интегральная адаптация.

E.N. Obukhova

APPLICATION OF THE METHOD OF INTEGRAL ADAPTATION FOR SYNTHESIS OF ADAPTIVE LAWS OF CONTROL IN A PNEUMATIC DRIVE UNDER HARMONIC DISTURBANCES

The active use of electro-pneumatic systems in various areas of industrial automation is due to such rather high performance indicators of the pneumatic drive as reliability, speed, low cost, availability of use in high humidity conditions, as well as in explosive and fire hazardous environments. The article provides a brief analysis of domestic and foreign scientific works devoted to the development of various methods of pneumatic system control, in which the problem of synthesis of effective control laws with adaptive properties to external disturbances is posed. The aim of this work is to develop an adaptive nonlinear synergetic control law to suppress the disturbing effect, which was specified and additively introduced into the mathematical model in the form of a harmonic function. The synthesis of the adaptive control law was carried out using the method of integral adaptation, which is part of the concept of synergetic control theory. The obtained results of computer modeling confirm the adaptive properties of the obtained nonlinear synergistic control laws and the achievement of the set technological control goal - moving the rod to a given position under conditions of harmonic disturbance. One of the important stages in the analysis of the results presented in this work is the conduct of experimental studies that make it possible to test the performance of synthesized nonlinear synergetic control laws obtained by analytical means on the educational and experimental stand of pneumatic drives of vertical and horizontal displacement of the Camozzi Company. For the practical implementation of the obtained synergetic control laws, the controller was programmed in the CoDeSys programming environment for industrial automation in the graphical language of functional block diagrams FBD (Function Block Diagram) IEC 61131-3 programming.

Electro-pneumatic system; invariant manifolds; phase variables; harmonic disturbance; integral adaptation.

Ведение. Сложная специфика математического описания существенно нелинейных термодинамических процессов, связанных со свойством высокой сжимаемости воздуха, приводит к тому, что в математической модели пневматической системы принимаются допущения о неизменяемости некоторых параметрических характеристик в процессе работы системы.

Совершенно очевидно, что в реальных условиях эксплуатации на систему управления также будут воздействовать некоторые внешние возмущающие факторы, действие которых может значительно отклонить желаемое изменение управляемой величины. При этом, принятые при формировании модели допущения могут привести к значительному расхождению результатов математического моделирования синтезированных законов управления со значениями изменения регулируемого параметра при реальном функционировании замкнутой системы управления.

Как известно, методы классического управления на основе линейных типовых регуляторов применительно к параметрически неопределенным моделям показывают не удовлетворительные результаты. Поэтому для исследования моделей с неопределенными параметрами в условиях внешних возмущений применяют специальные методы робастного и адаптивного управления.

В работе [1] робастные свойства пневматической системы обеспечиваются введением динамической обратной связи по положению штока привода, так же предлагается использование адаптивного нейросетевого динамического контроллера для оценки внешних возмущений. В работах [2, 3] представлен адаптивный робастный регулятор, основанный, на введении обратной связи, которая повышает способность отслеживания траектории выходного звена. Была доказана устойчивость разработанного адаптивного регулятора по методу функции Ляпунова.

В работах [4, 5] рассматривается построение регулятора на основе косвенного адаптивного робастного управления содержащего расширенный наблюдатель возмущений. При этом адаптивный закон управления позволяет оценить не только влияния возмущений, но и неопределенности параметров.

В работе [6] была разработана гибридная нечеткая адаптивная схема управления пневмоприводом с эталонной моделью. Управление основано на мультипликативной адаптации сигнала с линейным регулятором в прямом направлении и нечетким регулятором в адаптивном контуре обратной связи. Подобный принцип управления относится к методу прямой адаптации с эталонной моделью, который объединяет регулятор и алгоритм его адаптации одной целью – уменьшению сигнала ошибки по положению эталонной модели пневмопривода. Принцип прямой адаптации так же применялся для дискретной модели управления пневматической системой в работе [7], а так же в работе [8] для онлайн-оценки параметров эталонной модели пневмопривода.

В работах [9–11] в качестве неопределенного параметра рассматривается сила трения, как изменяющаяся во времени функция. Отдельно стоит выделить работы [12–15] в которых адаптивное управление строится на основе метода управления скользящим режимом. Это так же позволяет ослабить влияние ошибок оценки параметров, и внешних возмущений.

В основе рассмотренных выше работ лежит либо принцип косвенной адаптации либо принцип адаптации на основе эталонной модели или управления скользящим режимом. Принцип косвенной адаптации имеет существенный недостаток, заключающийся в том, что цель работы синтезируемого адаптивного регулятора расходится с целью построенного наблюдателя, вследствие чего управление по выходной координате оказывается разомкнутым, что естественным образом влияет на ошибку управления.

В адаптивном управлении на основе построения эталонной модели описанный выше недостаток устраняется. Это осуществляется за счет построения блока эталонной модели, вырабатывающей желаемое значение регулируемого параметра управления. Однако в данном принципе управления построение эталонной модели основывается на линейной модели исследуемой системы, что, на наш взгляд, является существенным недостатком. При этом введение в систему блока эталонной

модели усложняет процесс управления, возникают не контролируемые ошибки. Заметим, что в адаптивном управлении скользящим режимом также рассматриваются линейные модели, представленные в матрично-векторной форме.

В данной статье предлагается синтез адаптивных нелинейных синергетических законов управления, в основе которых положен метод аналитического конструирования агрегированных регуляторов синергетической теории управления (СТУ) [16], разработанный А.А. Колесниковым. При этом адаптивные свойства системы управления к рассматриваемому внешнему гармоническому воздействию достигаются за счет введения принципа интегральной адаптации СТУ [17–19].

Постановка задачи управления. Возмущения колебательного характера являются достаточно распространенными силами, которые могут возникать при различных режимах работы пневматической системы. Подобные возмущения, происходят под влиянием некоторых внешних переменных сил, действие которых не зависит от естественного движения системы в фазовом пространстве.

В качестве одного из примеров возникновения подобных возмущений в пневматической системе можно привести тренажер кабины карьерного экскаватора. При этом внешние возмущающие колебания могут иметь произвольный вид и для их точного функционального описания необходимо либо иметь доступ к экспериментальным данным о характере действующих возмущений [20], либо использовать функцию, которая наиболее полно описывает различные вариации колебаний. В качестве такой функции рассмотрим гармоническое воздействие: $z_f = A \cdot \sin(\omega t)$.

Компенсировать внешние воздействия позволяет использование принципа интегральной адаптации СТУ, базирующегося на фундаментальных свойствах астатизма подавлять возмущения до полного исчезновения ошибки регулирования. Причем применение принципа интегральной адаптации СТУ позволяет подавлять возмущения без их оценки и измерения.

Синтез законов управления с использованием принципа интегральной адаптации СТУ. Интегральная адаптация предполагает расширение фазового пространства путем введения в систему ДУ ЭПС дополнительных интеграторов, которые по сути формируют обратные связи системы, за счет которых и достигается результат подавления внешних возмущающих факторов.

Исходя из метода АКАР расширим фазовое пространство пневматической системы введя в математическую модель интегральные соотношения переменных x_5 и x_6 , обеспечивающие подавление функции возмущения z_f :

$$\begin{cases} \dot{x}_1(t) = x_2; \\ \dot{x}_2(t) = a_{21} \cdot x_3 - a_{22} \cdot x_4 - a_{23} \cdot x_2 - a_{24} + x_5 + z_f; \\ \dot{x}_3(t) = \left[a_{31} \cdot (x_1 + l_{01})^{-1} \cdot U_{1f} \cdot \sqrt{p_M^2 - x_3^2} \right] - \frac{kx_3x_2}{(x_1 + l_{01})}; \\ \dot{x}_4(t) = - \left[a_{41} \cdot (a_{42} - x_1)^{-1} \cdot U_{2f} \cdot \left(\frac{x_4}{p_A} \right)^{\frac{k-1}{2k}} \cdot \sqrt{x_4^2 - p_A^2} \right] + \frac{kx_4x_2}{(a_{42} - x_1)}; \\ \dot{x}_5(t) = x_6; \\ \dot{x}_6(t) = \beta(x_1 - x_1^*). \end{cases} \quad (1)$$

В математическую модель для упрощенного представления и дальнейшего анализа введены следующие коэффициенты: $a_{21} = \frac{S_1}{M}$; $a_{22} = \frac{S_2}{M}$; $a_{23} = \frac{k_{ВТР}}{M}$; $a_{24} = \frac{P_A(S_1 - S_2)}{M}$; $a_N = \frac{N}{M}$; $a_{31} = \frac{k\sqrt{RT_M}}{S_1\sqrt{\xi}}$; $a_{41} = \frac{k\sqrt{RT_M}}{S_2\sqrt{\xi}}$; $a_{42} = (L + l_{02})$.

В модели следующие параметры обозначены как: x_1 – координата перемещения поршня l (м); x_2 – скорость перемещения движущихся масс V (м/с); x_3 – давление в камере наполнения p_1 (Па); x_4 – давление в камере выхлопа p_2 (Па); M – масса подвижной части поршня и штока (0,5 кг); S_1 и S_2 – эффективные площади поршневой и штоковой полостей пневмоцилиндра ($8 \cdot 10^{-4} \text{ м}^2$ и $6 \cdot 10^{-4} \text{ м}^2$); $k_{\text{ВТР}}$ – коэффициент вязкого трения (100 Н · с/м); T_m – абсолютная температура газа (293 К); k – показатель адиабаты для воздуха (1,4); R – универсальная газовая постоянная (287 Дж/(кг · К)); l_{01} и l_{02} – начальная и конечная координаты положения поршня (0,002 м); ξ – суммарный коэффициент сопротивления входящих в линию дросселей (30); p_m – уровень давления на входе ли ($5 \cdot 10^5$ Па); p_a – атмосферное давление (10^5 Па).

Управлениями U_{1f} и U_{2f} являются площади поперечного сечения отверстий первого и второго пневмораспределителей которые находятся в правых частях математической модели.

Сформулируем задачу управления следующим образом: необходимо синтезировать управляющие воздействия U_{1f} и U_{2f} на ПР1 и ПР2 таким образом, чтобы положение штока пневмопривода x_1 достигало заданного значения с учетом равновесия давлений $x_3 = x_4$ камерах пневмоцилиндра, при условии воздействия внешних возмущений в виде гармонической функции времени.

Одним из первоначальных этапов синтеза синергетических законов управления по методу АКАР является выделение целей управления – инвариантов, которые задаются исходя из технологической задачи, с учетом физической сущности динамики процессов происходящих в исследуемой системе. Поскольку рассматриваемая пневматическая система описывается уравнениями динамики движения поршня и уравнениями термодинамического равновесия (1), то необходимо задать технологические и термодинамические инварианты которые будут являться конечной целью состояния системы.

Введем термодинамическое инвариантное многообразие ψ_1 , отражающее равновесие давлений в камерах пневмоцилиндра в конце хода поршня и технологическое инвариантное многообразие ψ_2 являющееся функцией равновесия давления в камере наполнения, порождающего движение поршня и внутреннего управления, как сложной функции скорости, перемещения и действия возмущающего фактора:

$$\begin{cases} \psi_1 = x_4 - x_3 = 0 \\ \psi_2 = x_3 - \varphi_1(x_1, x_2, x_5, x_6) = 0, \end{cases} \quad (2)$$

где $\varphi_1(x_1, x_2, x_5, x_6)$ – внутреннее управление, которым является желаемое изменение давления в камере наполнения x_3 и непосредственно влияющее на скорость штока и его перемещение, с учетом возмущения.

Для устойчивого асимптотического движения системы из условно допустимой области фазового пространства к пересечению инвариантных многообразий (2) согласно методу АКАР ψ_1 и ψ_2 должны удовлетворять следующим функциональным уравнениям:

$$\begin{cases} T_1 \dot{\psi}_1(t) + \psi_1 = 0, \\ T_2 \dot{\psi}_2(t) + \psi_2 = 0, \end{cases} \quad (3)$$

где $T_1 > 0$ и $T_2 > 0$ постоянные времени, характеризующие длительность движения системы к инвариантным многообразиям.

При попадании системы на пересечение инвариантных многообразий (2) исходная система (1) преобразуется в следующий декомпозированный вид:

$$\begin{cases} \dot{x}_1(t) = x_2, \\ \dot{x}_2(t) = (a_{21} - a_{22}) \cdot \varphi_1(x_1, x_2, x_5, x_6) - a_{23} \cdot x_2 - a_{24} + x_5, \\ \dot{x}_5(t) = x_6, \\ \dot{x}_6(t) = \beta(x_1 - x_1^*). \end{cases} \quad (4)$$

Внутреннее управление φ_1 является сложной функцией, отражающей зависимость скорости от желаемого перемещения с учетом возмущения, поэтому введем следующее многообразие:

$$\psi_3 = x_2 - k(x_1 - x_1^*) + b_1 x_5 + b_2 x_6 = 0, \quad (5)$$

где b_1, b_2 и k – параметры синергетического регулятора, отвечающие за устойчивость движения на многообразии ψ_3 к цели управления. Очевидно, что устойчивость заданного инвариантного многообразия (5), отражающего траекторию движения исследуемой системы в фазовом пространстве ее состояния к конечной цели управления $x_1 = x_1^*$, будет при следующих значениях коэффициентов: $k < 0, b_1, b_2$ и $\beta > 0$. Данные условия устойчивости будут учтены при моделировании системы.

Функциональное уравнение обеспечивающее асимптотическую устойчивость движения к инвариантному многообразию (5):

$$T_3 \dot{\psi}_3(t) + \psi_3 = 0, \quad (6)$$

где постоянная времени $T_1 > 0$.

Совместно решая (5) с (6) с учетом (4) выразим $\varphi_1(x_1, x_2, x_5, x_6)$:

$$\varphi_1 = \frac{T_3(x_5 - kx_2 + b_1 x_6 - a_{23} x_2 - a_{24}) + b_2 x_6 + x_2 - k(x_1 - x_1^*) + b_1 x_5}{T_3(a_{21} - a_{22})}. \quad (7)$$

Закон управления выраженный относительно площади поперечного сечения ПР1 и ПР2 находится из совместного решения (7) и (2) с учетом функциональных уравнений (3) и модели (1):

$$U_{1f} = \frac{T_2(\lambda_1 + \lambda_2 + \lambda_4 - \lambda_5) - x_3 - \lambda_6}{T_2 \lambda_3}; \quad U_{2f} = \frac{T_1(\lambda_3 U_1 + \lambda_4 + \lambda_8) - x_4 + x_3}{-T_1 \lambda_7}, \quad (8)$$

где

$$\begin{aligned} \lambda_1 &= \frac{kx_2}{T_3(a_{21} - a_{22})}; \quad \lambda_2 = \frac{(T_3 k + T_3 a_{23} - 1)(a_{21} x_3 - a_{22} x_4 - a_{23} x_2 - a_{24} - z_f)}{T_3(a_{21} - a_{22})}; \\ \lambda_3 &= \frac{a_{31} \sqrt{p_m^2 - x_3^2}}{x_1 + l_{01}}; \quad \lambda_4 = \frac{kx_3 x_2}{(x_1 + l_{01})}; \quad \lambda_5 = \frac{b_2 x_6}{T_3(a_{21} - a_{22})}; \\ \lambda_6 &= \frac{T_3(kx_2 + a_{23} x_2 + a_{24} + z_f + b_1 x_6) + x_2 - k(x_1 - x_1^*) + b_1 x_5 + b_2 x_6}{T_3(a_{21} - a_{22})}; \\ \lambda_7 &= \frac{a_{41}(x_4/P_m)^{\frac{k-1}{2k}} \cdot \sqrt{(x_4^2 - p_A^2)}}{a_{42} - x_1}; \quad \lambda_8 = \frac{kx_4 x_2}{a_{42} - x_1}. \end{aligned}$$

Моделирование. Результаты моделирования системы с учетом гармонического возмущения представлены на рис. 1–3. При моделировании были заданы следующие начальные условия системы: $x_1 = 0$ м; $x_2 = 0$ м/с; $x_3 = 10^5$ Па; $x_4 = 10^5$ Па; желаемые значения перемещения поршня $x_1^* = 0,1$ м и давления в выхлопной камере $x_4 = x_3 = 10^5$ Па; значения возмущающего воздействия: $A = 0,1; \omega = 0,5$. Значения параметров регулятора: $T_1 = T_2 = T_3 = 0,5$ с, $k = -4; b_1 = 0,01; b_2 = 0,1; \beta = 1$.

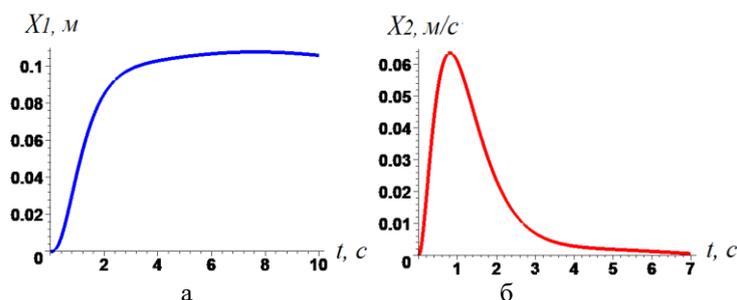


Рис. 1. Переходные процессы ЭПС при синусоидальном возмущении: а – перемещения поршня; б – скорости поршня

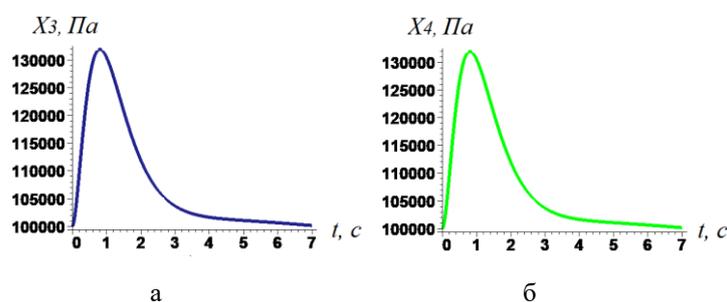


Рис. 2. Переходные процессы ЭПС при синусоидальном возмущении: а – давления во впускной камере; б – давления в выпускной камере

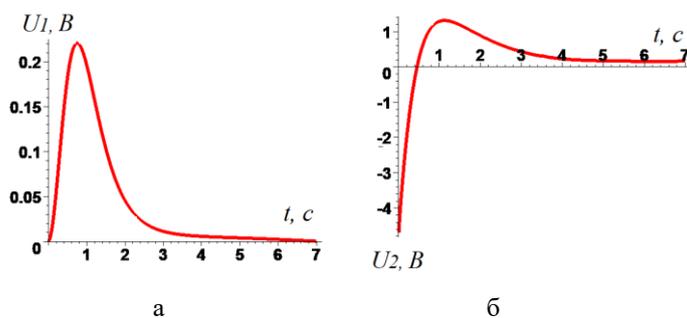


Рис. 3. Графики управляющих воздействий при синусоидальном возмущении: а – по напряжению ПР1; б – по напряжению ПР2

Экспериментальное исследование полученных законов управления на стенде пневматических приводов компании Camozzi. Полученные законы управления были опытным путем апробированы на учебно-экспериментальном стенде ЭПС, фото которого приведено на рис. 4. Ниже представлены основные функциональные элементы стенда ЭПС: пневматические приводы горизонтального и вертикального перемещений Camozzi серии QCT2A032A200 (1); пропорциональные пневматические распределители входящих и исходящих потоков сжатого воздуха Camozzi серии LRWA2-36-2A00 (2); программируемый логический контроллер ПЛК Овен 150; магнитный датчик положения поршня цилиндра серии MPS-128TSTR0.

Данный стенд ЭПС используется в учебных целях для исследования горизонтального и вертикального позиционирования пневматических приводов, с пропорциональными распределителями потоков сжатого воздуха, обеспечивающих торможение и остановку привода в заданной координате.

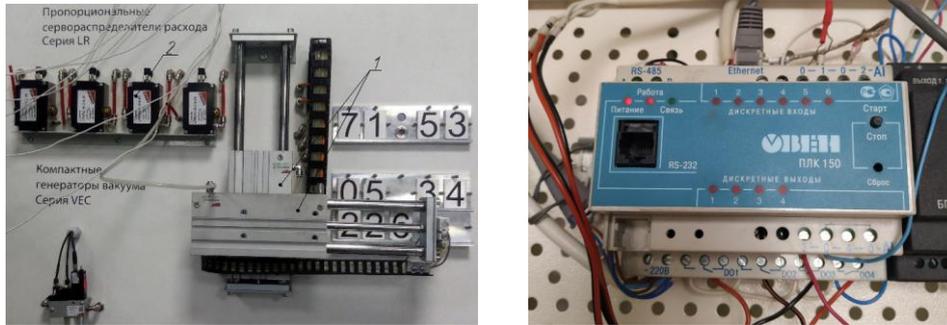


Рис. 4. Учебно-экспериментальный стенд горизонтального и вертикального позиционирования пневмоприводов и ПЛК Овен 150

Сбор и обработку информации о положении штока привода обеспечивает аналоговый магнитный датчик, который крепится на корпус пневмоцилиндра. Кабель датчика положения подключается к ПЛК Овен 150, который представлен на рис. 4. При этом, сигнал, который магнитный датчик положения получает при перемещении пневмопривода, передаётся в центральный процессорный модуль ПЛК, который в свою очередь связан с компьютером, через кабель Ethernet. Пропорциональные распределители в системе управления выполняют функцию точного распределения потоков сжатого воздуха, обеспечивая при этом нужную скорость и точность остановки штока привода в заданной точке позиционирования. Входы пропорциональных распределителей подключаются к ПЛК, которые в зависимости от значения управляющего сигнала, осуществляют открытие или закрытие проходных клапанов. Таким образом, в замкнутой системе ПЛК осуществляет организацию управления положением штока привода, изменения заданной координаты положения, а так же принятия, обработки и выдачи сигнала при помощи задаваемого программой алгоритма управления.

Проектирование, отладка и загрузка программы в ПЛК осуществляется в инструментальной среде разработки программ для промышленной автоматизации CoDeSys на одном из пяти языков МЭК 61131-3 программирования. В данном случае, в качестве языка программирования был выбран графический язык функциональных блочных диаграмм FBD (Function Block Diagram), который позволяет доступно и наглядно представить обработку и передачу управляющего сигнала от ПЛК к ПР, в виде цепей из определенных элементов-программных компонентов (POU).

На рис. 5,а представлен интерфейс среды CoDeSys и реализованная графическая программа для ПЛК на языке FBD. При разработке алгоритма программы использовались базовые функциональные блоки, имеющиеся в библиотеки программы CoDeSys, при этом связи между блоками отражают обмен данных в виде некоторых программно заданных переменных.

На рис. 5,б представлена одновременная цифровая трассировка двух управляющих сигналов – напряжение U1 на ПР1 (зеленый график) и напряжение U2 на ПР2 (красный график). По осям абсцисс представлены циклы работы контроллера, при этом в режиме эмуляции в среде программной отладки CoDeSys 1 цикл ПЛК составляет 55 миллисекунд. Таким образом 500 циклов ПЛК в переводе на время работы контроллера составляет 27,5 секунд. По осям ординат представлены значе-

ния выходных напряжений – единица измерения Вольт. Как видно из рисунка 5б контроллер начинает вырабатывать управляющий сигнал напряжения U_1 примерно на 150 цикле, в то время как сигнал напряжения U_2 примерно на 170 цикле, который соответствует времени равной 8,25 секунды. Эта естественная временная задержка на миллисекунды обусловлена тем, что вначале срабатывает ПР1 на выпуск сжатого воздуха в камеру наполнения, а затем открывается клапан на ПР2.

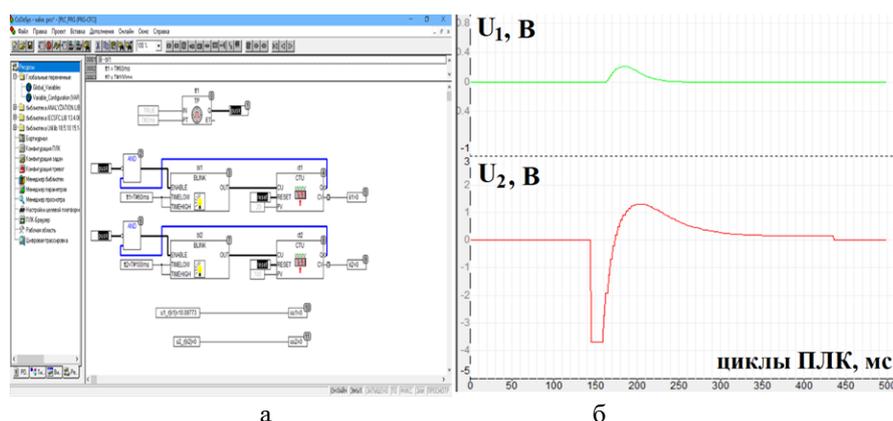


Рис. 5. Интерфейс среды CoDeSys: а – реализованная графическая программа для ПЛК на языке FBD; б – график управляющих напряжений U_1 на ПР1 и U_2 на ПР2

При этом наличие начальных циклов контроллера при которых отсутствуют значения управляющих сигналов обусловлены временем задержки, в которую входят: задержка операционной системы при подключении среды программирования к контроллеру, физические задержки распространения сигналов на входы ПЛК, а так же задержка на воспроизводство цифровой трассировки в режиме эмуляции контроллера.

Судя по графику напряжения U_2 , (рис. 5,б), время за которое клапан ПР2 полностью закроется составляет 11 секунд, при этом, по графику, полученному в ходе математического моделирования в пакете Maple, время полного закрытия ПР2 составляет 6-7 секунд (рис. 3б). Это связано с математической идеализацией получения аналитических данных, а так же отсутствие учета, при моделировании процессов, времени на поступление, обработку и выдачи управляющего сигнала контроллером.

Выводы. В целом, сравнивая графики управляющих воздействий – напряжений U_1 и U_2 (рис. 3), полученные с применением синергетического метода АКАР и графики сигналов дискретных выходов контроллера на рис.5б, которые являются управляющими воздействиями на клапаны ПР1 и ПР2, можно совершенно определенно констатировать тождественность данных графиков, что свидетельствует о высокой точности аналитически синтезированных нелинейных синергетических законов управления.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Liu G., Li G., Z. Peng, Pan H. Adaptive Neural Network Dynamic Surface Control Algorithm for Pneumatic Servo System // Proceedings of the 11th International Conference on Modelling, Identification and Control (ICMIC2019). – P. 821-829.
2. Zuo H., Tao G. Cross-coupling integral adaptive robust posture control of a pneumatic parallel platform // Mechanical Engineering, Control Science and Information Engineering, Journal of Central South University. – 2016. – Vol. 23. – P. 2036-2047.

3. *Ren H., Fan J.* Adaptive backstepping slide mode control of pneumatic position servo system // Chinese Journal of Mechanical Engineering. – 2016. – Vol. 29. – P. 1003-1009.
4. *Meng D., Tao G., Li A., Li W.* Motion synchronization of dual-cylinder pneumatic servo systems with integration of adaptive robust control and cross-coupling approach // Journal of Zhejiang University SCIENCE C. – 2014. – Vol. 15. – P. 651-663.
5. *Li A., Meng D., Lu B., Li Q.* Nonlinear cascade control of single-rod pneumatic actuator based on an extended disturbance observer // Journal of Central South University. – 2019. – Vol. 26. – P. 1637-1648.
6. *Tóthová M., Pitel J., Hošovský A.* Simulation of Hybrid Fuzzy Adaptive Control of Pneumatic Muscle Actuator // Proceedings of the 4th Computer Science On-line Conference (CSOC2015), Intelligent Systems in Cybernetics and Automation Theory. – 2015. – Vol. 2. – P. 239-246.
7. *Qian P., Tao G., Meng D., Liu H.* A modified direct adaptive robust motion trajectory tracking controller of a pneumatic system // Journal of Zhejiang University SCIENCE C. – 2014. – Vol. 15. – P. 878-891.
8. *Tao G., Shang C., Meng D., Zhou C.* Posture control of a 3-RPS pneumatic parallel platform with parameter initialization and an adaptive robust method // Frontiers of Information Technology & Electronic Engineering. – 2017. – Vol. 18. – P. 303-316.
9. *Meng D., Tao G., Zhu X.* Adaptive robust motion trajectory tracking control of pneumatic cylinders // Journal of Central South University. – 2013. – Vol. 20. – P. 3445-3460.
10. *Meng D., Tao G., Liu H., Zhu X.* Adaptive robust motion trajectory tracking control of pneumatic cylinders with LuGre model-based friction compensation // Chinese Journal of Mechanical Engineering. – 2014. – Vol. 27. – P. 802-815.
11. *Tran X., Nguyen V., Nguyen N.C., Pham D.T., Pha V.L.* Sliding mode control for a pneumatic servo system with friction compensation // International Conference on Engineering Research and Applications (ICERA 2019): Advances in Engineering Research and Application. – P. 648-656.
12. *Kien C.V., Anh H.P.H.* Enhanced adaptive fuzzy sliding mode control for nonlinear uncertain serial pneumatic artificial muscle robot system // International Conference on Advances in Computational Mechanics ACOME 2017: Proceedings of the International Conference on Advances in Computational Mechanics. – 2017. – P. 1033-1050.
13. *Chen, H.-Y., Liang J.-W.* Adaptive wavelet neural network controller for active suppression control of a diaphragm-type pneumatic vibration isolator // International Journal of Control, Automation and Systems. – 2017. – Vol. 15. – P. 1456-1465.
14. *Meng D., Tao G., Ban W., Qian P.* Adaptive robust output force tracking control of pneumatic cylinder while maximizing/minimizing its stiffness // Journal of Central South University. – 2013. – Vol. 20. – P. 1510-1518.
15. *Lee L., Li I.* Design and implementation of a robust FNN-based adaptive sliding-mode controller for pneumatic actuator systems // Journal of Mechanical Science and Technology. – 2016. – Vol. 30. – P. 381-396.
16. *Колесников А.А., Веселов Г.Е., Попов А.Н., Колесников Ал.А., Топчиев Б.В., Мушенко А.С., Кобзев В.А.* Синергетические методы управления сложными системами: механические и электромеханические системы. – 2-е изд.. – М.: Либроком, 2013. – 304 с.
17. *Кузьменко А.А.* Интегральная адаптация высокого порядка в задачах синтеза нелинейных систем управления // Информатика и системы управления. – 2018. – № 1 (55). – С. 142-153.
18. *Кузьменко А.А.* Синергетическое управление нелинейными техническими системами: интегральная адаптация высокого порядка // VIII Всероссийская научная конференция "Системный синтез и прикладная синергетика": Сб. научных трудов. Южный федеральный университет. – 2017. – С. 50-60.
19. *Колесников А.А., Колесников Ал.А., Кузьменко А.А.* Метод АКАР и теория адаптивного управления в задачах синтеза нелинейных систем управления // Мехатроника, автоматизация, управление. – 2017. – Т. 18, № 9. – С. 579-589.
20. *Веселов Г.Е., Синецын А.С.* Синтез системы управления адаптивной подвеской с учётом физических ограничений амортизатора // Известия ЮФУ. Технические науки. – 2015. – № 7 (168). – С. 170-184.

REFERENCES

1. Liu G., Li G., Z. Peng, Pan H. Adaptive Neural Network Dynamic Surface Control Algorithm for Pneumatic Servo System, *Proceedings of the 11th International Conference on Modelling, Identification and Control (ICMIC2019)*, pp. 821-829.
2. Zuo H., Tao G. Cross-coupling integral adaptive robust posture control of a pneumatic parallel platform, *Mechanical Engineering, Control Science and Information Engineering, Journal of Central South University*, 2016, Vol. 23, pp. 2036-2047.
3. Ren H., Fan J. Adaptive backstepping slide mode control of pneumatic position servo system, *Chinese Journal of Mechanical Engineering*, 2016, Vol. 29, pp. 1003-1009.
4. Meng D., Tao G., Li A., Li W. Motion synchronization of dual-cylinder pneumatic servo systems with integration of adaptive robust control and cross-coupling approach, *Journal of Zhejiang University SCIENCE C*, 2014, Vol. 15, pp. 651-663.
5. Li A., Meng D., Lu B., Li Q. Nonlinear cascade control of single-rod pneumatic actuator based on an extended disturbance observer, *Journal of Central South University*, 2019, Vol. 26, pp. 1637-1648.
6. Tóthová M., Pitel J., Hošovský A. Simulation of Hybrid Fuzzy Adaptive Control of Pneumatic Muscle Actuator, *Proceedings of the 4th Computer Science On-line Conference (CSOC2015), Intelligent Systems in Cybernetics and Automation Theory*, 2015, Vol. 2, pp. 239-246.
7. Qian P., Tao G., Meng D., Liu H. A modified direct adaptive robust motion trajectory tracking controller of a pneumatic system, *Journal of Zhejiang University SCIENCE C*, 2014, Vol. 15, pp. 878-891.
8. Tao G., Shang C., Meng D., Zhou C. Posture control of a 3-RPS pneumatic parallel platform with parameter initialization and an adaptive robust method, *Frontiers of Information Technology & Electronic Engineering*, 2017, Vol. 18, pp. 303-316.
9. Meng D., Tao G., Zhu X. Adaptive robust motion trajectory tracking control of pneumatic cylinders, *Journal of Central South University*, 2013, Vol. 20, pp. 3445-3460.
10. Meng D., Tao G., Liu H., Zhu X. Adaptive robust motion trajectory tracking control of pneumatic cylinders with LuGre model-based friction compensation, *Chinese Journal of Mechanical Engineering*, 2014, Vol. 27, pp. 802-815.
11. Tran X., Nguyen V., Nguyen N.C., Pham D.T., Pha V.L. Sliding mode control for a pneumatic servo system with friction compensation, *International Conference on Engineering Research and Applications (ICERA 2019): Advances in Engineering Research and Application*, pp. 648-656.
12. Kien C.V., Anh H.P.H. Enhanced adaptive fuzzy sliding mode control for nonlinear uncertain serial pneumatic artificial muscle robot system, *International Conference on Advances in Computational Mechanics ACOME 2017: Proceedings of the International Conference on Advances in Computational Mechanics*, 2017, pp. 1033-1050.
13. Chen H.-Y., Liang J.-W. Adaptive wavelet neural network controller for active suppression control of a diaphragm-type pneumatic vibration isolator, *International Journal of Control, Automation and Systems*, 2017, Vol. 15, pp. 1456-1465.
14. Meng D., Tao G., Ban W., Qian P. Adaptive robust output force tracking control of pneumatic cylinder while maximizing/minimizing its stiffness, *Journal of Central South University*, 2013, Vol. 20, pp. 1510-1518.
15. Lee L., Li I. Design and implementation of a robust FNN-based adaptive sliding-mode controller for pneumatic actuator systems, *Journal of Mechanical Science and Technology*, 2016, Vol. 30, pp. 381-396.
16. Kolesnikov A.A., Veselov G.E., Popov A.N., Kolesnikov A.A., Topchiev B.V., Mushenko A.S., Kobzev V.A. Sinergeticheskie metody upravleniya slozhnymi sistemami: mekhanicheskie i elektromekhanicheskie sistemy [Synergetic methods for managing complex systems: mechanical and electromechanical systems]. 2nd ed. Moscow: Librokom, 2013, 304 p.
17. Kuz'menko A.A. Integral'naya adaptatsiya vysokogo poriyadka v zadachakh sinteza nelineynykh sistem upravleniya [High-order integral adaptation in problems of synthesis of nonlinear control systems], *Informatika i sistemy upravleniya* [Informatics and control systems], 2018, No. 1 (55), pp. 142-153.
18. Kuz'menko A.A. Sinergeticheskoe upravlenie nelineynymi tekhnicheskimi sistemami: integral'naya adaptatsiya vysokogo poriyadka [Synergetic control of nonlinear technical systems: high-order integral adaptation], *VIII Vserossiyskaya nauchnaya konferentsiya "Sistemnyy sintez i prikladnaya sinergetika": Sb. nauchnykh trudov. Yuzhnyy federal'nyy universitet* [VIII all-Russian scientific conference "System synthesis and applied synergetics": Collection of scientific papers. Southern Federal University, 2017], 2017, pp. 50-60.

19. *Kolesnikov A.A., Kolesnikov A.A., Kuz'menko A.A.* Metod AKAR i teoriya adaptivnogo upravleniya v zadachakh sinteza nelineynykh sistem upravleniya [AKAR method and adaptive control theory in problems of synthesis of nonlinear control systems], *Mekhatronika, avtomatizatsiya, upravlenie* [Mechatronics, automation, control], 2017, Vol. 18, No. 9, pp. 579-589.
20. *Veselov G.E., Sinitsyn A.S.* Sintez sistemy upravleniya adaptivnoy podveskoj s uchedom fizicheskikh ograniczeniy amortizatora [Synthesis of adaptive suspension control system taking into account the physical limitations of the shock absorber], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, No. 7 (168), pp. 170-184.

Статью рекомендовал к опубликованию д.т.н., профессор А.Б. Чернышев.

Обухова Елена Николаевна – Донской государственный технический университет; e-mail: elena21@spark-mail.ru; 344016, г. Ростов-на-Дону, пл. Гагарина, 1; тел.: 89064146401; старший преподаватель.

Obukhova Elena Nikolayevna – Don State Technical University; e-mail: elena21@spark-mail.ru; 1, Gagarin square, Rostov-on-Don, 344016, Russia; phone: +79064146401; senior lecturer.

Раздел IV. Информационная безопасность

УДК 004.422

DOI 10.18522/2311-3103-2020-4-212-221

Л.К. Бабенко, И.Д. Русаловский**МЕТОД РЕАЛИЗАЦИИ ГОМОМОРФНОГО ДЕЛЕНИЯ***

Рассматриваются проблемы гомоморфной криптографии. Гомоморфная криптография – одно из молодых направлений криптографии. Его особенность заключается в том, что можно обрабатывать зашифрованные данные без их предварительной расшифровки таким образом, что результат операций над зашифрованными данными эквивалентен после расшифровки результату операции над открытыми данными. В статье приводится краткий обзор областей применения гомоморфного шифрования. Для решения различных прикладных задач требуется поддержка всех математических операций, в том числе и операции деления, а возможность выполнить эту операцию гомоморфно позволит расширить возможности применения гомоморфного шифрования. В работе предлагается метод гомоморфного деления, основанный на абстрактном представлении шифротекста в виде обыкновенной дроби. В работе подробно описывается предложенный метод. Кроме этого статья содержит пример практической реализации предложенного метода. Предлагается разделить уровни обработки данных на 2 уровня – криптографический и математический. На криптографическом уровне используется некоторый полностью гомоморфный алгоритм шифрования и выполняются базовые гомоморфные математические операции – сложение, умножение и разность. Математический уровень является надстройкой над криптографическим и расширяет его возможности. На математическом уровне шифротекст представляется в виде простой дроби и появляется возможность выполнения операции гомоморфного деления. Также в работе приводится практический пример применения метода гомоморфного деления на базе алгоритма Джендри для целых чисел. Приводятся выводы и возможные пути дальнейшего развития.

Гомоморфное шифрование; криптографическая защита; методы и алгоритмы; гомоморфное деление.

L.K. Babenko, I.D. Rusalovsky**METHOD OF IMPLEMENTING HOMOMORPHIC DIVISION**

The article deals with the problems of homomorphic cryptography. Homomorphic cryptography is one of the young directions of cryptography. Its peculiarity lies in the fact that it is possible to process encrypted data without preliminary decryption in such a way that the result of operations on encrypted data is equivalent, after decryption, to the result of operations on open data. The article provides a brief overview of the areas of application of homomorphic encryption. To solve various applied problems, support for all mathematical operations is required, including the division operation, and the ability to perform this operation homomorphically will expand the possibilities of using homomorphic encryption. The paper proposes a method of homomorphic division based on an abstract representation of the ciphertext in the form of an ordinary fraction. The paper describes in detail the proposed method. In addition, the article contains an example of the practical implementation of the proposed method. It is proposed to divide the levels of data processing into 2 levels – cryptographic and mathematical. At the cryptographic level, a complete-

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-37-90140.

ly homomorphic encryption algorithm is used and the basic homomorphic mathematical operations are performed – addition, multiplication and difference. The mathematical level is a superstructure on top of the cryptographic level and expands its capabilities. At the mathematical level, the ciphertext is represented as a simple fraction and it becomes possible to perform the homomorphic division operation. The paper also provides a practical example of applying the homomorphic division method based on the Gentry algorithm for integers. Conclusions and possible ways of further development are given.

Homomorphic encryption; cryptographic protection; methods and algorithms; homomorphic division.

Введение. Процессы информатизации и глобализации оказали значительное значение на все сферы жизни. Информационные технологии стали частью современной жизни. Человечество повсеместно пользуется благами информатизации: персональные компьютеры, удаленные сервисы, онлайн банки и покупки, интернет вещей и многое другое. Трудно представить жизнь современного человека без компьютера и глобальной сети интернет. В результате процесса информатизации вырос объем информации, возросли информационные потоки. В условиях бурного роста информационных технологий как никогда ранее стала актуальна проблема обеспечения информационной безопасности, в частности, проблема обеспечения криптографической защиты.

Криптография позволяет сохранять данные в секрете, обеспечивая их безопасную передачу. В настоящее время получило развитие новое направление – гомоморфная криптография. Его отличительной особенностью является то, что появилась возможность обрабатывать зашифрованные данные без их предварительной расшифровки таким образом, что результат операций над зашифрованными данными эквивалентен после расшифровки результату операции над открытыми данными. При этом решается одна из важных проблем криптографии, связанная с решением вопросов генерации, хранения и распространения (распределения) общих сеансовых ключей. В результате применения гомоморфного шифрования повышается уровень защищенности данных – сервер получает зашифрованные данные, обрабатывает их и возвращает зашифрованный результат, а открытые данные и ключи шифрования не покидают безопасный сегмент при сетевом взаимодействии.

Гомоморфное шифрование имеет существенные преимущества, однако и не лишено на данный момент недостатков. Основными из них являются:

- ◆ необходимость обеспечения целостности пересылаемых данных;
- ◆ высокая трудоемкость операций над зашифрованными данными;
- ◆ быстрый рост коэффициентов после выполнения операций над данными.

Первая и последняя из перечисленных проблем легко решаются. Для обеспечения целостности можно воспользоваться различными криптографическими протоколами, а чтобы избежать быстрого роста коэффициентов, можно использовать схемы с коррекцией размеров шифротекстов после выполнения операций над ними (технология бутстрэппинга, о которой будет подробно рассказано в другой статье).

На данный момент существует большое количество алгоритмов полностью гомоморфного шифрования, основанных на различных принципах. Для ряда из них выполнены практические реализации, которые находятся в общем доступе, есть и коммерческие продукты. Однако ни в одной из найденных во время анализа реализаций не было поддержки операции деления. Таким образом, актуальным является разработка метода, позволяющего расширить существующие гомоморфные алгоритмы функциональностью деления. Разработке такого метода посвящена данная статья.

Теоретические основы. Идея полностью гомоморфного шифрования была предложена изобретателями алгоритма шифрования RSA Рональдом Ривестом, Леонардом Адлеманом в соавторстве с Майклом Дертусосом в 1978 году [1]. Они выдвинули предположение, что возможно выполнение произвольных опера-

ций над зашифрованными данными без необходимости их предварительной расшифровки. В то время все попытки разработки полностью гомоморфной схемы не привели к успеху. Все криптосистемы, которые были разработаны в последующие годы, были лишь частично гомоморфными.

Лишь в начале 2000-ых годов стали появляться криптосистемы, которые выходили за рамки частичного гомоморфизма. Большой прорыв в исследовании гомоморфного шифрования сделан в 2009 году, когда Крейг Джентри, воспользовавшись криптографией на решетках, впервые представил вариант полностью гомоморфной схемы шифрования [2]. Следом за этим было опубликовано немало работ, предлагающих модификации криптосистемы Джентри, вносящих в нее улучшения и повышающие ее быстродействие. Вместе с этим активно идет работа над созданием альтернативных симметричных гомоморфных криптосистем на основе полиномов от одной, нескольких переменных и матричных полиномов.

Гомоморфное шифрование – это такая форма шифрования, которая позволяет проводить некоторые математические операции над зашифрованными данными и получать зашифрованный результат, который будет соответствовать результату операций, выполняемых над открытыми данными [3–6].

Различают два вида гомоморфного шифрования – частично и полностью гомоморфное. Криптосистема считается частично гомоморфной, если она является гомоморфной относительно умножения или сложения. Полностью гомоморфная схема гомоморфна одновременно относительно двух математических операций – сложения и умножения.

Гомоморфное шифрование – достаточно молодое направление. Его развитие и совершенствование актуально [7–18] и позволит использовать его в таких сферах как:

- ◆ Облачные вычисления.
- ◆ Облачная обработка фотографий.
- ◆ Электронные голосования (выборы).
- ◆ Защищенный поиск информации.

Анализ актуальности. Гомоморфное шифрование не так давно получило свое начальное развитие, но уже является довольно известным. Разрабатываются новые алгоритмы шифрования и совершенствуются старые, создаются библиотеки для гомоморфной обработки данных. На данный момент уже выполнено несколько реализаций библиотек для гомоморфного шифрования [19–20]. Наиболее серьезными реализациями, доступными для общего пользования, можно считать две:

- ◆ библиотека `HElib`, созданная Шаем Хавели и Виктором Шоуп, которая реализует криптосистему `BGV` с `GHS` оптимизацией;
- ◆ библиотека `FHEW` созданная Лео Дуглас и Даниэль Миккианакио, которая является реализацией комбинации криптосистемы обучения с ошибками Регева и техники создания гибкой схемы Алперин-Шериффа и Пейкерта.

Библиотеки имеют высокую скорость работы, хорошую оптимизацию. Но, проанализировав области применения гомоморфного шифрования, можно сделать вывод о необходимости выполнения операций над целыми числами, включая операцию деления. В результате поиска в открытых источниках не было найдено информации о возможной реализации метода или алгоритма гомоморфного деления. В данной статье будет рассмотрен вопрос гомоморфного деления, предложен метод и примеры его использования.

Метод гомоморфного деления. Существующие схемы и алгоритмы гомоморфного шифрования не позволяют использовать операцию деления над зашифрованными данными. Для решения данной проблемы предлагается использовать некую абстракцию, построенную над шифротекстом и расширяющую возможно-

сти по выполнению математических операций над ним. Для этого необходимо выделить два уровня представления данных – криптографический и математический. На рис. 1 приведен пример взаимодействия двух уровней и пользователя. Пунктиром выделена часть схемы, которая может иметь различные варианты реализации. Возможность замены реализации криптографического уровня может быть достигнута за счет единства интерфейса.

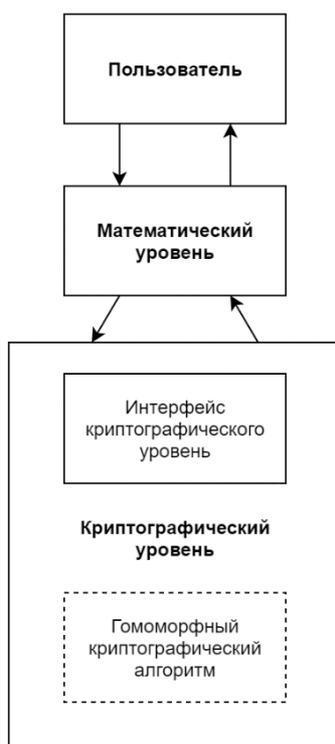


Рис. 1. Уровни представления данных и их взаимосвязь

Криптографическое ядро предложенной выше схемы может быть представлено абсолютно любым полностью гомоморфным алгоритмом шифрования. При этом алгоритм может быть и симметричным и ассиметричным. Конечный же пользователь данной схемы будет взаимодействовать только с математическим уровнем, интерфейс которого будет неизменен, какой бы гомоморфный алгоритм шифрования не был бы выбран.

Криптографический уровень представляет собой модуль, основным типом данных которого является зашифрованное гомоморфно число. На данном уровне должны быть реализованы возможности по созданию новых шифротекстов на базе открытых текстов и ключей шифрования (операция шифрования данных), получению открытых данных из шифротекста на основании ключа расшифрования (операция расшифрования данных), а также основные математические операции над шифротекстами – сложение, разность и умножение. Подобный функционал позволит реализовать любой полностью гомоморфный алгоритм, работающий с целыми числами. Криптографический уровень принимает запросы от математического, гомоморфно обрабатывает данные и возвращает результат обратно (на математический слой схемы). Пользователь не имеет доступа к криптографическому уровню напрямую, а взаимодействует только с математическим уровнем.

Математический уровень является надстройкой над криптографическим. Данный уровень необходим, чтобы обеспечить поддержку операции деления. Для этого на данном уровне данные будут содержаться в виде простой дроби, числителем и знаменателем которой являются экземпляры объектов криптографического уровня. Математический уровень является прослойкой между пользователем и криптографическим ядром. Все действия на математическом уровне выполняются как операции над простыми дробями, числитель и знаменатель которых поддерживают операции сложения, разности и умножения так, как представлены объектами криптографического уровня. К примеру, когда выполняется операция умножения двух объектов математического уровня, выполняются два запроса к криптографическому уровню для перемножения числителей (делимое – *dividend*) и знаменателей (делитель – *divider*) этих чисел соответственно, а в результате будет получена результирующая дробь:

$$A * B = \left\{ \begin{array}{l} C.dividend = A.dividend * B.dividend \\ C.divider = A.divider * B.divider \end{array} \right\} = C$$

Реализация операций математического уровня будет подробнее описана далее. Предложенный выше метод гомоморфного деления обладает явными плюсами – решение является гибким, так как позволяет реализовать гомоморфное деление на базе любого частично гомоморфного алгоритма шифрования. Вторым плюсом метода можно считать простоту реализации. Однако имеются и негативные стороны. Главным минусом метода является рост объема шифротекста в два раза, что является следствием использования простой дроби. При условии использования полностью гомоморфных алгоритмов, увеличение объема шифротекста практически не будет иметь негативного влияния на процесс шифрования и передачи данных.

Пример практической реализации. Приведем ниже пример реализации предложенного алгоритма в виде псевдокода. Для реализации необходимо будет создать два класса – базовый, который будет реализовывать всю криптографическую функциональность и надстройку над ним, реализующую математическую функциональность. Базовый класс будет иметь наименование EncryptedData, математического – EncryptedFraction. Ниже приведена более подробная информация по каждому из уровней.

Криптографический уровень. Класс, определяющий основной тип данных криптографического уровня, содержит в себе функционал, позволяющий генерировать ключи шифрования, выполнять операции шифрования и расшифрования а также основные математические операции над зашифрованными гомоморфно числами – сложение, разность и умножение. Данный уровень может быть реализован любым аддитивным и мультипликативным гомоморфным алгоритмом. Использование же полностью гомоморфного алгоритма позволит сдерживать уровень шума и, следовательно, ограничивать объем шифротекстов. Подробнее природа и рост шума в гомоморфных алгоритмах будут рассмотрены отдельно. Интерфейс криптографического уровня представлен на рис. 2.

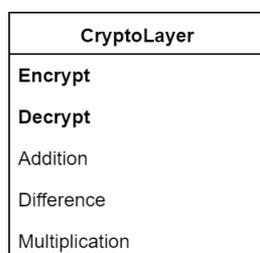


Рис. 2. Интерфейс класса криптографического уровня

Математический уровень. Данный уровень является надстройкой над криптографическим. В своей реализации он представляет собой обыкновенную дробь, в числителе и знаменателе которой объекты криптографического уровня – зашифрованные гомоморфно числа. Благодаря использованию простой дроби в математическом уровне стала возможна реализация деления с использованием только операции гомоморфного умножения. Выполнение всех математических операций на данном уровне – это выполнение операций над простыми дробями, используя гомоморфные операции сложения, разности и умножения криптографического уровня. Архитектура (интерфейс) класса представлена на рис. 3.

MathLayer
CryptoLayer divider
CryptoLayer dividend
Encrypt
Decrypt
Addition
Difference
Division
Multiplication

Рис. 3. Интерфейс класса математического уровня

Пример использования метода. Предложенный метод предполагает взаимодействие с математическим уровнем. Объект математического уровня содержит внутри себя делимое и делитель – объекты криптографического уровня. Во время выполнения операций над объектами математического уровня вызываются соответствующие операции над делимым и делителем – объектами криптографического уровня. Все операции выполняются как операции над обычными дробями – приведение к общему знаменателю и операции над числителями для сложения и разности, перемножение числителей и знаменателей между собой соответственно для умножения и умножение числителя первого на знаменатель второго и знаменателя первого на числитель второго для деления.

Пусть A, B – объекты математического уровня, тогда операции над ними будут реализованы следующим образом:

Сложение и разность:

$$A \pm B = \left\{ \begin{array}{l} C.dividend = A.dividend * B.divider \pm B.dividend * A.divider \\ C.divider = A.divider * B.divider \end{array} \right\} = C$$

Умножение:

$$A * B = \left\{ \begin{array}{l} C.dividend = A.dividend * B.dividend \\ C.divider = A.divider * B.divider \end{array} \right\} = C$$

Деление:

$$A \div B = \left\{ \begin{array}{l} C.dividend = A.dividend * B.divider \\ C.divider = A.divider * B.dividend \end{array} \right\} = C$$

Пример с численными значениями. В рамках данного примера в качестве алгоритма шифрования будет использован алгоритм гомоморфного шифрования целых чисел, предложенный Джентри [21–22]. Схема шифрования в предложенном алгоритме имеет вид:

$$C = P * r + s * e + m,$$

где P – секретный ключ, r – небольшая случайная константа, s – пространство открытого текста, e – небольшая случайная величина (ошибка).

Схема расшифрования может быть описана как:

$$m = (C \bmod P) \bmod s.$$

Пусть параметры схемы шифрования $P = 997$, $r = 5$, $s = 11$, $e = 2$. Тогда вычислим объекты A , B математического уровня для исходных текстов $a = 4$, $b = 2$:

$$C1 = P * r + s * e + a = 997 * 5 + 11 * 2 + 4 = 5011$$

$$C2 = P * r + s * e + a = 997 * 5 + 11 * 2 + 2 = 5009$$

$$A = \left\{ \begin{array}{l} A. \text{dividend} = 5011 \\ A. \text{divider} = 1 \end{array} \right\}$$

$$B = \left\{ \begin{array}{l} B. \text{dividend} = 5009 \\ B. \text{divider} = 1 \end{array} \right\}.$$

Сложение:

$$C = A + B = \left\{ \begin{array}{l} C. \text{dividend} = 5011 * 1 + 5009 * 1 \\ C. \text{divider} = 1 * 1 \end{array} \right\}$$

$$C = \left\{ \begin{array}{l} C. \text{dividend} = 10020 \\ C. \text{divider} = 1 \end{array} \right\}$$

$$m = (10020 \bmod 997) \bmod 11 / (1 \bmod 997) \bmod 11 = 6 / 1 = 6.$$

Разность:

$$C = A - B = \left\{ \begin{array}{l} C. \text{dividend} = 5011 * 1 - 5009 * 1 \\ C. \text{divider} = 1 * 1 \end{array} \right\}$$

$$C = \left\{ \begin{array}{l} C. \text{dividend} = 2 \\ C. \text{divider} = 1 \end{array} \right\}$$

$$m = (2 \bmod 997) \bmod 53 / (1 \bmod 997) \bmod 53 = 2 / 1 = 2$$

Умножение:

$$C = A * B = \left\{ \begin{array}{l} C. \text{dividend} = 5011 * 5009 \\ C. \text{divider} = 1 * 1 \end{array} \right\}$$

$$C = \left\{ \begin{array}{l} C. \text{dividend} = 25\,100\,099 \\ C. \text{divider} = 1 \end{array} \right\}$$

$$m = (25\,100\,099 \bmod 997) \bmod 53 / (1 \bmod 997) \bmod 53 = 8 / 1 = 8$$

Деление:

$$C = A \div B = \left\{ \begin{array}{l} C. \text{dividend} = 5011 * 1 \\ C. \text{divider} = 1 * 5009 \end{array} \right\}$$

$$C = \left\{ \begin{array}{l} C. \text{dividend} = 5011 \\ C. \text{divider} = 5009 \end{array} \right\}$$

$$m = (5011 \bmod 997) \bmod 53 / (5009 \bmod 997) \bmod 53 = 4 / 2 = 2.$$

Заключение. В рамках данной статьи рассмотрена проблема гомоморфного деления целых чисел, предложен и описан метод реализации гомоморфного деления и приведены практические примеры применения вышеописанного метода. Практическая ценность работы состоит в решении одной из проблем гомоморфного шифрования – реализация гомоморфного деления позволяет расширить область практического применения гомоморфного шифрования в таких областях, как об-
лачные вычисления, решение задач защиты информации, машинное обучение.

Предложенный в работе метод гомоморфного деления обладает рядом недостатков:

- ◆ Требуется использования в криптографическом ядре алгоритмов полностью гомоморфного шифрования, удовлетворяющих критериям Джендри.
- ◆ Независимо от используемого алгоритма шифрования объем шифротекстов увеличивается вдвое.

В перспективе планируется работать над улучшением предложенного метода и продолжать поиск лучшего решения для реализации гомоморфного деления. Также предполагается практическая реализация предложенного метода.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Rivest R.L., Adleman L., Dertouzos M.L. On data banks and privacy homomorphisms // Foundations of secure computation. – 1978. – Vol. 32, No. 4. – P. 169-178.
2. Gentry C. A fully homomorphic encryption scheme // A dissertation submitted to the department of computer science and the committee on graduate students of Stanford University. – 2009.
3. Diffie W. and Hellman M. New directions in cryptography // IEEE Transactions on Information Theory. – 1976. – Vol. IT-22. – P. 644-654.
4. Gentry C. Fully homomorphic encryption using ideal lattices // STOC. – 2009. – P. 169-178.
5. Rivest R., Adleman L., Dertouzos M. On data banks and privacy homomorphisms // Foundations of Secure Computation. Academic Press. – 1978. – P. 169-177.
6. Goldwasser S., Micali S. Probabilistic encryption // Journal of Computer and System Sciences. – 1984. – Vol. 28, No. 2. – P. 270-299.
7. Бабенко Л.К., Буртыка Ф.Б., Макаревич О.Б., Трепачева А.В. Методы полностью гомоморфного шифрования на основе матричных полиномов // Вопросы кибербезопасности, – 2015. – № 1. – С. 17-20.
8. Бабенко Л.К., Буртыка Ф.Б., Макаревич О.Б., Трепачева А.В. Полностью гомоморфное шифрование (обзор) // Вопросы защиты информации. – 2015. – № 3. – С. 3-26.
9. Бабенко Л.К., Буртыка Ф.Б., Макаревич О.Б., Трепачева А.В. Защищенные вычисления и гомоморфное шифрование // Программные системы: теория и приложения. – 2014. – 25 с.
10. Макаревич О.Б., Буртыка Ф.Б. Защищенная облачная база данных с применением гомоморфной криптографии // Тез. докл. 6-й Росс. мультиконференции «Информационные технологии в управлении» (ИТУ–2014). – СПб., 2014. – С. 567-572.
11. Буртыка Ф.Б. Пакетное симметричное полностью гомоморфное шифрование на основе матричных полиномов // Тр. Института системного программирования РАН. – 2014. – Т. 26, № 5. – С. 99-116.
12. Буртыка Ф.Б. Симметричное полностью гомоморфное шифрование с использованием неприводимых матричных полиномов // Известия ЮФУ. Технические науки. – 2014. – № 8 (157). – С. 107-122.
13. Трепачева А.В. Криптоанализ симметричных полностью гомоморфных линейных криптосистем на основе задачи факторизации чисел // Известия ЮФУ. Технические науки. – 2015. – № 5 (166). – С. 89-102.
14. Brakerski Vaikuntathan. Efficient fully homomorphic encryption from (standard) LWE // FOCS. – 2011.
15. Rao G.V., Kakulapati V., Purushoththaman M. Privacy homomorphism in mobile ad hoc networks // International Journal of Research & Reviews in Computer Science. – 2011.
16. Жиров А.О., Жирова А.О., Кренделев С.Ф. Безопасные облачные вычисления с помощью гомоморфной криптографии // Безопасность информационных технологий. – 2013. – № 1. – С. 6-12.
17. Regev O. New lattice-based cryptographic constructions // J. ACM. – 2004. – Vol. 51, No. 6. – P. 899-942.
18. Regev O. On lattices, learning with errors, random linear codes, and cryptography // STOC. – 2005. – P. 84-93.
19. Helib. – URL: <https://github.com/homenc/HElib> (дата обращения: 01.06.2020).
20. FHEW. – URL: <https://github.com/lducas/FHEW> (дата обращения: 01.06.2020).

21. Gentry C., Halevi S. Implementing gentry's fully-homomorphic encryption scheme // EUROCRYPT, ser. Lecture Notes in Computer Science, K. G. Paterson, Ed. – Vol. 6632. – Springer, 2011. – P. 129-148.
22. Dijk M., Gentry C., Halevi S., Vaikuntanathan V. Fully Homomorphic Encryption over the Integers // Eurocrypt. – 2010.

REFERENCES

1. Rivest R.L., Adleman L., Dertouzos M.L. On data banks and privacy homomorphisms, *Foundations of secure computation*, 1978, Vol. 32, No. 4, pp. 169-178.
2. Gentry C. A fully homomorphic encryption scheme, *A dissertation submitted to the department of computer science and the committee on graduate students of Stanford University*, 2009.
3. Diffie W. and Hellman M. New directions in cryptography, *IEEE Transactions on Information Theory*, 1976, Vol. IT-22, pp. 644-654.
4. Gentry C. Fully homomorphic encryption using ideal lattices, *STOC*, 2009, pp. 169-178.
5. Rivest R., Adleman L., Dertouzos M. On data banks and privacy homomorphisms, *Foundations of Secure Computation*. Academic Press, 1978, pp. 169-177.
6. Goldwasser S., Micali S. Probabilistic encryption, *Journal of Computer and System Sciences*, 1984, Vol. 28, No. 2, pp. 270-299.
7. Babenko L.K., Burtyka F.B., Makarevich O.B., Trepacheva A.V. Metody polnost'yu gomomorfnoy shifrovaniya na osnove matrichnykh polinomov [Methods of fully homomorphic encryption based on matrix polynomials], *Voprosy kiberbezopasnosti* [Cybersecurity issues], 2015, No. 1, pp. 17-20.
8. Babenko L.K., Burtyka F.B., Makarevich O.B., Trepacheva A.V. Polnost'yu gomomorfnoe shifrovanie (obzor) [Fully Homomorphic Encryption (Overview)], *Voprosy zashchity informatsii* [Information Security Issues], 2015, No. 3, pp. 3-26.
9. Babenko L.K., Burtyka F.B., Makarevich O.B., Trepacheva A.V. Zashchishchennyye vychisleniya i gomomorfnoe shifrovanie [Secure computing and homomorphic encryption], *Programmnyye sistemy: teoriya i prilozheniya* [Software systems: theory and applications], 2014, 25 p.
10. Makarevich O.B., Burtyka F.B. Zashchishchennaya oblachnaya baza dannykh s primeneniem gomomorfnoy kriptografii [Secure cloud database using homomorphic cryptography], *Tez. dokl. 6-y Ross. mul'tikonferentsii «Informatsionnyye tekhnologii v upravlenii» (ITU-2014)* [Proceedings of 6th Russian multiconference «Information Technologies in Control» (ITU-2014). Saint Petersburg, 2014, pp. 567-572.
11. Burtyka F.B. Paketnoe simmetrichnoe polnost'yu gomomorfnoe shifrovanie na osnove matrichnykh polinomov [Batch symmetric fully homomorphic encryption based on matrix polynomials], *Tr. Instituta sistemnogo programmirovaniya RAN* [Proceedings of the Institute for System Programming RAS], 2014, Vol. 26, No. 5, pp. 99-116.
12. Burtyka F.B. Simmetrichnoe polnost'yu gomomorfnoe shifrovanie s ispol'zovaniem neprivodimyykh matrichnykh polinomov [Symmetric fully homomorphic encryption using irreducible matrix polynomials], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2014, No. 8 (157), pp. 107-122.
13. Trepacheva A.V. Kriptoanaliz simmetrichnykh polnost'yu gomomorfnykh lineynykh kriptosistem na osnove zadachi faktorizatsii chisel [Cryptanalysis of symmetric fully homomorphic linear cryptosystems based on the number factorization problem], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, No. 5 (166), pp. 89-102.
14. Brakerski Vaikuntanathan. Efficient fully homomorphic encryption from (standard) LWE, *FOCS*, 2011.
15. Rao G.V., Kakulapati V., Purushoththaman M. Privacy homomorphism in mobile ad hoc networks, *International Journal of Research & Reviews in Computer Science*, 2011.
16. Zhirov A.O., Zhirova A.O., Krendelev S.F. Bezopasnyye oblachnyye vychisleniya s pomoshch'yu gomomorfnoy kriptografii [Secure Cloud Computing with Homomorphic Cryptography], *Bezopasnost' informatsionnykh tekhnologiy* [Information technology security], 2013, № 1, pp. 6-12.
17. Regev O. New lattice-based cryptographic constructions, *J. ACM*, 2004, Vol. 51, No. 6, pp. 899-942.

18. *Regev O.* On lattices, learning with errors, random linear codes, and cryptography, *STOC*, 2005, pp. 84-93.
19. *Helib*. Available at: <https://github.com/homenc/HElib> (accessed 01 June 2020).
20. *FHEW*. Available at: <https://github.com/lducas/FHEW> (accessed 01 June 2020).
21. *Gentry C., Halevi S.* Implementing gentry's fully-homomorphic encryption scheme, *EUROCRYPT*, ser. *Lecture Notes in Computer Science*, K.G. Paterson, Ed., Vol. 6632. Springer, 2011, pp. 129-148.
22. *Dijk M., Gentry C., Halevi S., Vaikuntanathan V.* Fully Homomorphic Encryption over the Integers, *Eurocrypt*, 2010.

Статью рекомендовал к опубликованию профессор И.А. Калмыков.

Бабенко Людмила Климентьевна – Южный федеральный университет; e-mail: blk@tsure.ru; 347928, г. Таганрог, ул. Чехова, 2, корпус "И"; тел.: +79054530191; кафедра безопасности информационных технологий; д.т.н.; профессор.

Русаловский Илья Дмитриевич – e-mail: ilya.rusalovskiy@mail.ru; тел.: +79885526701; кафедра безопасности информационных технологий; аспирант.

Babenko Lyudmila Kliment'evna – Southern Federal University; e-mail: blk@tsure.ru; Block "I", 2, Chekhov street, Taganrog, 347928, Russia; phone: +79054530191; the department of information technologies security; dr. of eng. sc.; professor.

Rusalovsky Ilya Dmitrievich – e-mail: ilya.rusalovskiy@mail.ru; phone: +79885526701; the department of information technologies security; postgraduate student.

УДК 621.396.96

DOI 10.18522/2311-3103-2020-4-221-229

Я.К. Миронов, П.Д. Миронова, К.Е. Румянцев

ВЕРОЯТНОСТНЫЕ ХАРАКТЕРИСТИКИ ПОРОГОВОГО АЛГОРИТМА ОБНАРУЖЕНИЯ СИНХРОИМПУЛЬСОВ В СИСТЕМЕ КВАНТОВОГО РАСПРЕДЕЛЕНИЯ КЛЮЧА НА ОСНОВЕ ИНФОРМАЦИИ СО СМЕЖНОЙ ПАРЫ ВРЕМЕННЫХ СЕГМЕНТОВ*

Системы квантового распределения ключа (КРК) обеспечивают повышенную защищённость передаваемой информации. Для стабильной работы системы КРК необходима точная синхронизация станций пользователей при минимальных временных затратах. Предложен алгоритм обнаружения синхросигнала с пороговым тестом. Предполагается, что синхроимпульс одновременно находится в двух соседних временных сегментах. Вероятность обнаружения пары временных сегментов, где присутствует синхроимпульс, определяется вероятностью превышения порогового уровня суммарным количеством сигнальных и шумовых импульсов, регистрируемых в двух соседних сегментах. Цель исследований направлена на сравнительный анализ порогового уровня и вероятностных характеристик аппаратуры синхронизации при пороговом тестировании каждой пары временных сегментов внутри временного кадра, полученных при ориентации на модели Гаусса и Пуассона для числа фотонов и импульсов темного тока (ИТТ), принимаемых за время анализа временного сегмента. Исследованы вероятностные характеристики алгоритма обнаружения синхросигналов в системе квантового распределения ключа на основе сравнения числа фотонов со смежной пары временных сегментов с пороговым уровнем. Анализируется применение аппроксимации статистических свойств процессов на выходе фотодетектора законом Пуассона и нормальным распределением. Оценивается влияния модели Пуассона и Гаусса на выбор порогового уровня и расчёт эффективности синхронизации при пороговом тестировании каждой пары временных сегментов внутри временного кадра,

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-37-90040.

полученных при ориентации на модели Гаусса и Пуассона для числа фотонов и ИТТ, принимаемых за время анализа временного сегмента. Установлено, что выбор порогового уровня, исходя из нормального распределения, даёт заниженное значение. Аппроксимация статистики фотонов и импульсов темнового тока нормальным законом обеспечивает пороговый уровень ниже требуемого. Причём различие растёт с ужесточением требований к вероятности ложного срабатывания. Полученные вероятностные свойства алгоритма обнаружения синхросигнала на основе анализа суммы отсчетов со смежной пары сегментов с пороговым уровнем позволяют сформулировать рекомендации по выбору аппроксимации статистики сигнала: для экспресс-расчётов вероятностных характеристик целесообразно использовать модель Гаусса, в случае необходимости более высокой точности анализа рекомендуется использовать модель Пуассона.

Квантовое распределение ключа; синхронизация; пороговый уровень; вероятностные характеристики; модель Гаусса; модель Пуассона.

Ya.K. Mironov, P.D. Mironova, K.E. Rumyantsev

PROBABILISTIC CHARACTERISTICS OF THE THRESHOLD ALGORITHM FOR DETECTING SYNCHRONIZING PULSES IN THE QUANTUM KEY DISTRIBUTION SYSTEM BASED ON INFORMATION FROM AN ADJACENT PAIR OF TIME SEGMENTS

Quantum key distribution systems (QKD) provide increased security of transmitted information. For the stable operation of the QKD system, accurate synchronization of user stations is required with minimal time costs. An algorithm for detecting a sync signal with a threshold test is proposed. It is assumed that the sync pulse is simultaneously in two adjacent time segments. The probability of detecting a pair of time segments where a sync pulse is present is determined by the probability of exceeding the threshold level by the total number of signal and noise pulses recorded in two adjacent segments. The purpose of the research is aimed at a comparative analysis of the threshold level and probabilistic characteristics of synchronization equipment during threshold testing of each pair of time segments within a time frame, obtained when orienting on the Gauss and Poisson model for the number of photons and dark current pulses (DCP) received during the time segment analysis. The probabilistic characteristics of the detection algorithm for sync signals are studied in a quantum key distribution system based on a comparison of the number of photons from an adjacent pair of time segments with a threshold level. The application of the approximation of the statistical properties of the processes at the output of the photodetector by the Poisson law and the normal distribution is analyzed. The influence of the Poisson and Gaussian models on the choice of the threshold level and the calculation of the synchronization efficiency during the threshold testing of each pair of time segments within the time frame are estimated, obtained by orientation on the Gauss and Poisson models for the number of photons and DCP received during the analysis of the time segment. It was established that the choice of the threshold level based on the normal distribution gives an underestimated value. The approximation of the statistics of photons and pulses of dark current by a normal law provides a threshold level lower than the required one. Moreover, the difference grows with stricter requirements for the probability of false positives. The obtained probabilistic properties of the sync signal detection algorithm based on the analysis of the sum of counts from an adjacent pair of segments with a threshold level allow us to formulate recommendations for choosing an approximation of the signal statistics: for express calculations of probabilistic characteristics, it is advisable to use the Gaussian model; if a higher analysis accuracy is required, it is recommended to use the Poisson model.

Quantum key distribution; synchronization; threshold level; probabilistic characteristics; Gaussian model; Poisson model.

Введение. Защищённые системы связи, использующие квантовую криптографию для безопасного распределения ключа (КРК) между удалёнными пользователями, рассматриваются как основа развития информационных сетей будущего [1–6]. При квантовом распределении ключа разнесённые приёмная и передающая станции должны работать синхронно для устранения влияния слу-

чайных задержек сигнала в информационном канале. Временная синхронизация обеспечивается подсистемами синхронизации, являющимися составными частями системы КРК [7, 8].

Вопросы синхронизации станций исследовались в работах отечественных и зарубежных учёных [9–16]. Причём период следования квантовых импульсов принимается за временной кадр, который, в свою очередь, разбивается на конечное число временных сегментов (окон).

В работах [17–19] положение временных сегментов, на которых может находиться синхроимпульс, определяется с помощью подсчёта однофотонных импульсов (ОФИ) и импульсов темнового тока (ИТТ) в каждом временном сегменте. Предполагается, что на первом этапе синхронизации время анализа временного сегмента значительно превышает длительность оптического синхроимпульса.

В отличие от этого в [20] предложено положение сигнального временного сегмента фиксировать посредством подсчёта фотонов в каждой паре временных сегментов. При использовании порогового теста для каждой пары сегментов, испытания продолжают до тех пор, пока пороговый уровень не будет превышен. При этом принимается решение об обнаружении синхроимпульса в этой паре временных сегментов. Аппаратура переходит ко второму этапу уточнения местоположения синхроимпульса.

Основанием для использования такого алгоритма служит тот факт, что временные сегменты должны анализироваться попарно, так как синхроимпульс может одновременно принадлежать сразу двум соседним временным сегментам. Причём вероятность такого события велика при соизмеримости длительности оптического синхроимпульса и времени анализа временного сегмента.

Здесь для выбора порогового уровня и расчёта вероятностных характеристик предлагается использовать нормальное распределение для числа фотонов и ИТТ, принимаемых за время анализа временного сегмента. Однако теоретическое обоснование отказа от использования дискретного закона Пуассона для выбора порогового уровня и расчётов вероятностных характеристик не обосновано.

Цель исследований направлена на сравнительный анализ порогового уровня и вероятностных характеристик аппаратуры синхронизации при пороговом тестировании каждой пары временных сегментов внутри временного кадра, полученных при ориентации на модели Гаусса и Пуассона для числа фотонов и ИТТ, принимаемых за время анализа временного сегмента.

Алгоритм обнаружения синхросигнала с пороговым тестом. Рассмотрим процесс обнаружения, где синхроимпульсы имеют длительность τ_s . Количество сигнальных ОФИ за длительность синхроимпульса равно n_s , а фоновых ОФИ и ИТТ – n_b . Предположим, что при непрерывном поиске синхроимпульса используется пороговый тест для каждой пары временных сегментов.

Пусть число зарегистрированных импульсов в i -м временном сегменте равно n_i , а в соседнем $(i+1)$ -м сегменте – n_{i+1} . Причём предполагаем, что синхроимпульс одновременно находится в i -м и $(i+1)$ -м временных сегментах.

Вероятность обнаружения $p_{th,s}$ пары i -го и $(i+1)$ -го временных сегментов, где присутствует синхроимпульс, определяется вероятностью превышения порогового уровня k_{th} суммарным количеством $k = n_i + n_{i+1}$ сигнальных и шумовых импульсов, регистрируемых в двух соседних сегментах.

Если распределение отсчётов подчиняется закону Пуассона, то вероятность обнаружения в сигнальной паре i -го и $(i+1)$ -го временных сегментов при среднем суммарном числе сигнальных ОФИ и шумовых импульсов $\overline{n_{sb}} = \overline{n_s} + 2 \cdot \overline{n_b}$ может быть рассчитана по формуле

$$p_{th.s} = \sum_{k=k_{th}}^{\infty} \frac{\bar{n}_{sb}^k}{k!} \cdot \exp(-\bar{n}_{sb}).$$

Утверждается, что при $\bar{n}_{sb} \geq 9$ (многофотонный или токовый режим синхронизации) распределение Пуассона стремится к нормальному распределению с математическим ожиданием \bar{n}_{sb} и дисперсией \bar{n}_{sb} :

$$p(k, \bar{n}_{sb}) = \frac{1}{\sqrt{2\pi\bar{n}_{sb}}} \cdot \exp\left(-\frac{(k - \bar{n}_{sb})^2}{2\bar{n}_{sb}}\right).$$

Используя функцию распределения стандартного нормального распределения, находим

$$p_{th.s} = 1 - \Phi\left(\frac{k_{th} - \bar{n}_{sb}}{\sqrt{\bar{n}_{sb}}}\right).$$

Из формулы при $k_{th} = \bar{n}_{sb}$ находим $p_{th.s} = 0,5$. Следовательно, пороговый уровень k_{th} должен превышать значение среднего суммарного числа сигнальных ОФИ и шумовых импульсов \bar{n}_{sb} для обеспечения вероятности обнаружения в сигнальной паре более 0,5.

При ориентации на использование функции ошибок расчёт выполняется по формуле

$$p_{th.s} = \frac{1}{2} \cdot \left[1 - \operatorname{erf}\left(\frac{k_{th} - \bar{n}_{sb}}{\sqrt{2\bar{n}_{sb}}}\right)\right] = \frac{1}{2} \cdot \operatorname{erfc}\left(\frac{k_{th} - \bar{n}_{sb}}{\sqrt{2\bar{n}_{sb}}}\right).$$

Значение порога выбирается из требования допустимой вероятности $p_{th.b}$ превышения порогового уровня k_{th} суммарным количеством $k = n_i + n_{i+1}$ шумовых импульсов, регистрируемых в паре шумовых временных сегментов.

Вероятность $p_{th.b}$ связана с пороговым уровнем k_{th} соотношением

$$p_{th.b} = \sum_{k=k_{th}}^{\infty} \operatorname{Pos}(k, 2 \cdot \bar{n}_b).$$

При ориентации на нормальное распределение вероятность $p_{th.b}$ рассчитывается по формуле

$$p_{th.b} = 1 - \Phi\left(\frac{k_{th} - 2 \cdot \bar{n}_b}{\sqrt{2 \cdot \bar{n}_b}}\right).$$

При $\bar{n}_{sb} \geq 9$ (многофотонный или токовый режим синхронизации) пороговый уровень k_{th} для заданной вероятности $p_{th.b}$ определяется с помощью обратных функций $\Phi^{-1}(x)$ к $\Phi(x)$ или $\operatorname{erf}^{-1}(x)$ к $\operatorname{erf}(x)$:

$$k_{th} = \begin{cases} 2 \cdot \bar{n}_b + \sqrt{2 \cdot \bar{n}_b} \cdot \Phi^{-1}(1 - p_{th.b}); \\ 2 \cdot \bar{n}_b + \sqrt{4 \cdot \bar{n}_b} \cdot \operatorname{erf}^{-1}(1 - 2 \cdot p_{th.b}). \end{cases}$$

На возможности использования и определении границ применимости нормального распределения для расчёта вероятности ложного срабатывания при наблюдении пары фоновых временных сегментов следует остановиться подробнее.

В табл. 1 приведен числовой материал, полученный в ходе анализа зависимостей вероятности ложного срабатывания от среднего числа шумовых импульсов за длительность синхроимпульса для различных пороговых уровней.

Видно, что вероятность ложного срабатывания при анализе пары сегментов, рассчитанная при ориентации на нормальное распределение, даёт заниженное значение. Причём различие превышает 6 порядков при среднем числе шумовых импульсов $\bar{n}_b=0,5$ и пороговом уровне $k_{th}=8$ (вероятность ложного срабатывания равна $1,3 \cdot 10^{-12}$). Различие уменьшается до двух порядков при среднем числе шумовых импульсов $\bar{n}_b=1$.

Отметим, что различие сокращается с уменьшением порогового уровня. Так, например, если при среднем числе шумовых импульсов $\bar{n}_b=1,5$ различие в вероятности ложного срабатывания составляет 6,12 раз при пороговом уровне $k_{th}=8$ (вероятность ложного срабатывания равна 0,002), то уже всего 25 % при $k_{th}=4$ (вероятность ложного срабатывания равна 0,28).

Приемлемое различие вероятностей ложного срабатывания (порядка 20...30 %) обеспечивается при условии $k_{th} \cdot p_{th.b} > 1,2$.

Таблица 1

Зависимости вероятности ложного срабатывания от среднего числа шумовых импульсов за длительность синхроимпульса для трёх пороговых уровней. Модели Пуассона и Гаусса

Порог	Условия. Параметр	Среднее число шумовых импульсов							
		0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0
2	Модель Пуассона	0,26	0,59	0,80	0,91	0,96	0,98	0,99	1,00
	Модель Гаусса	0,16	0,50	0,72	0,84	0,91	0,95	0,97	0,98
	Различие, раз	1,67	1,19	1,12	1,08	1,05	1,04	1,02	1,01
	Произведение порога на вероятность	0,53	1,19	1,60	1,82	1,92	1,97	1,99	1,99
4	Модель Пуассона	0,02	0,14	0,35	0,57	0,73	0,85	0,92	0,96
	Модель Гаусса	0,001	0,08	0,28	0,50	0,67	0,79	0,87	0,92
	Различие, раз	14,1	1,82	1,25	1,13	1,09	1,07	1,05	1,04
	Произведение порога на вероятность	0,08	0,57	1,41	2,27	1,94	3,40	3,67	3,83
8	Модель Пуассона	10^{-5}	10^{-3}	0,01	0,05	0,13	0,26	0,40	0,55
	Модель Гаусса	$1,3 \cdot 10^{-12}$	10^{-5}	0,002	0,02	0,09	0,21	0,35	0,50
	Различие, раз	$8 \cdot 10^6$	100	6,12	2,25	1,48	1,24	1,14	1,09
	Произведение порога на вероятность	0,00	0,01	0,10	0,41	1,07	2,05	3,21	4,38

На рис. 1 представлены зависимости вероятности ложного срабатывания от порогового уровня k_{th} при 4-х средних числах шумовых импульсов в сегменте \bar{n}_b : 2 (сплошная линия), 3 (штриховая линия), 4 (пунктирная линия) и 5 (штрихпунктирная линия). Ступенчатыми зависимостями представлены результаты расчётов при ориентации на распределение Пуассона, а непрерывными зависимостями – на нормальное распределение.

Из графиков видно, что выбор порогового уровня при ориентации на распределение Гаусса даёт при фиксированной вероятности ложного срабатывания заниженное значение. Например, при среднем числе шумовых импульсов в паре

шумовых временных сегментов 2 при ориентации на распределение Гаусса потребуются для обеспечения вероятности ложного срабатывания менее 10^{-6} выбор порогового уровня 14, а при ориентации на распределение Пуассона – 18 (различие в 1,3 раза). Для более высоких значений среднего числа шумовых импульсов в паре шумовых временных сегментов результаты аналогичны. При обеспечении вероятности ложного срабатывания в паре шумовых временных сегментов ниже 0,01 при среднем числе шумовых импульсов в сегменте $\bar{n}_b \geq 1$ пороговый уровень должен быть более 7, при $\bar{n}_b \geq 2 - k_{th} \geq 10$, а при $\bar{n}_b \geq 5 - k_{th} \geq 18$.

Аппроксимация статистики сигнала моделью Гаусса даёт значение порога, которое меньше на 3-4 единицы реально требуемого. Причём различие растёт с ужесточением требований к вероятности ложного срабатывания.

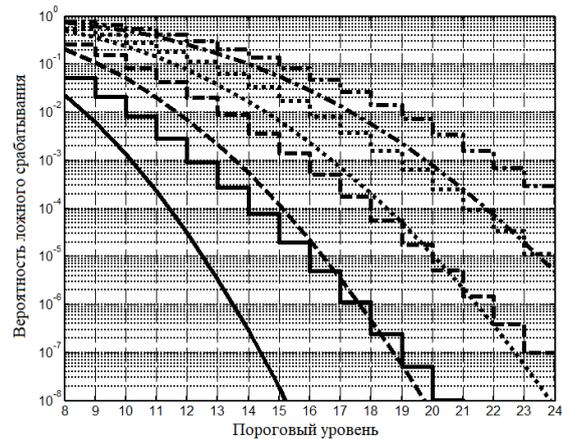


Рис. 1. Зависимости вероятности ложного срабатывания от порогового уровня

Полученные вероятностные свойства алгоритма обнаружения синхросигнала на основе анализа суммы отсчетов со смежной пары сегментов с пороговым уровнем позволяют сформулировать рекомендации по выбору аппроксимации статистики сигнала: для экспресс-расчётов вероятностных характеристик целесообразно использовать модель Гаусса, в случае необходимости более высокой точности анализа рекомендуется использовать модель Пуассона.

Выводы. Получены формулы для расчета порогового уровня, вероятностей ложного срабатывания и обнаружения в сигнальной паре временных сегментов, ориентированные на распределения Пуассона и Гаусса.

Проведен сравнительный анализ порогового уровня и вероятностных характеристик аппаратуры синхронизации при пороговом тестировании каждой пары временных сегментов внутри временного кадра, полученных при ориентации на модели Гаусса и Пуассона для числа фотонов и ИТТ, принимаемых за время анализа временного сегмента.

Выбор порогового уровня при ориентации на распределение Гаусса даёт при фиксированной вероятности ложного срабатывания меньшее значение, чем при использовании распределения Пуассона. Однако распределение Гаусса уместно использовать в случае необходимости проведения быстрого анализа. В связи с этим сформулированы рекомендации по выбору аппроксимации статистики сигнала: для экспресс-расчётов вероятностных характеристик целесообразно использовать модель Гаусса, в случае необходимости более высокой точности анализа рекомендуется использовать модель Пуассона.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ростелеком объявил о внедрении квантовой криптографии на своих сетях. – URL: <https://tass.ru/ekonomika/5685597> (дата обращения: 17.10.2018).
2. Румянцев К.Е. Системы квантового распределения ключа: монография. – Таганрог: Изд-во ТТИ ЮФУ, 2011. – 264 с.
3. Физика квантовой информации: Квантовая криптография. Квантовая телепортация. Квантовые вычисления / под ред. Д. Боумейстера, А. Экерта, А. Цайлингера. – М.: Постмаркет, 2002. – 376 с.
4. Gisin N., Ribordy G., Tittel W., Zbinden H. Quantum cryptography // Reviews of Modern Physics. – 2002. – Vol. 74, No. 1. – P. 145-195.
5. Bennett C., Brassard G. Quantum cryptography: Public key distribution and coin tossing // Proceedings of IEEE international conference on computers, systems and signal processing. Bangalore, India. – New York: Institute of Electrical and Electronics Engineers, 1984. – P. 175-179.
6. Shor P.W., Preskill J. Simple proof of security of the BB84 quantum key distribution protocol // Physical Review Letters. – 2000. – Vol. 85. – P. 441-444.
7. Румянцев К.Е. Синхронизация в системе квантового распределения ключа с автоматической компенсацией поляризационных искажений // Телекоммуникации. – 2017. – № 2. – С. 32-40.
8. Румянцев К.Е., Плёнкин А.П. Безопасность режима синхронизации системы квантового распределения ключей // Известия ЮФУ. Технические науки. – 2015. – № 5 (166). – С. 135-153.
9. Курочкин В.Л. и др. Экспериментальные исследования в области квантовой криптографии // Фотоника. – 2012. – Т. 5. – С. 54-66.
10. Mironov Y.K., Rumyantsev K.E. Single-Photon Algorithm for Synchronizing the System of Quantum Key Distribution with Polling Sections of a Fiber-Optic Line // Futuristic Trends in Networks and Computing Technologies. – 2020. – P. 87-97. DOI: https://doi.org/10.1007/978-981-15-4451-4_8.
11. Rumyantsev K.E., Linenko P.D., Shakir H.H.-Sh. Evaluation of the Influence of the Dispersion Properties of a Fiber-Optic Line on the Efficiency of an Algorithm for Single-Photon Synchronization of Quantum Key Distribution System // Conference Proceedings - 2019 Radiation and Scattering of Electromagnetic Waves, RSEMW 2019. – 2019. – P. 392-395. DOI: 10.1109/RSEMW.2019.8792769.
12. Румянцев К.Е., Рудинский Е.А. Двухэтапный временной алгоритм синхронизации в системе квантового распределения ключа с автоматической компенсацией поляризационных искажений // Известия ЮФУ. Технические науки. – 2017. – №5 (190). – С. 75-89.
13. Rumyantsev K., Rudinsky E. Parameters of the two-stage synchronization algorithm for the quantum key distribution system // Proceedings of the 10th International Conference on Security of Information and Networks (SIN'17). – 2017. – P. 140-147. DOI: 10.1145/3136825.3136888.
14. Lindsey W.C. Synchronization Systems in Communication and Control. Prentice-Hall, Englewood Cliffs, New Jersey, 1972.
15. Стиффлер Дж. Теория синхронной связи: пер. с англ. / под ред. Э.М. Габидулина. – М.: Связь, 1975.
16. Румянцев К.Е., Шакир Хайдер Хуссейн. Ограничения на дальность двухэтапной синхронизации в автокомпенсационной системе квантового распределения ключа // Телекоммуникации. – 2019. – № 12. – С. 2-10.
17. Румянцев К.Е., Плёнкин А.П. Синхронизация системы квантового распределения ключа при использовании фотонных импульсов для повышения защищённости // Известия ЮФУ. Технические науки. – 2014. – № 8 (157). – С. 81-96.
18. Rumyantsev K.E., Pljonkin A.P. Preliminary Stage Synchronization Algorithm of Autocompensation Quantum Key Distribution System with an Unauthorized Access Security // International Conference on Electronics, Information, and Communications (ICEIC). 2016. Vietnam, Danang. – P. 1-4. DOI: 10.1109/ELINFOCOM.2016.7562955. WOS:000389518100035. IDS: VG5KP.
19. Румянцев К.Е., Плёнкин А.П. Эффективность синхронизации системы квантового распределения ключа на однофотонных лавинных фотодиодах // Известия ЮФУ. Технические науки. – 2016. – № 9 (182). – С. 4-15.
20. Гальярди Р.М., Карп Ш. Оптическая связь: пер. с англ. / под ред. А.Г. Шереметьева. – М.: Связь, 1978. – 424 с.

REFERENCES

1. Rostelecom ob'yavil o vnedrenii kvantovoy kriptografii na svoikh setyakh [Rostelecom announced the introduction of quantum cryptography on its networks]. Available at: <https://tass.ru/ekonomika/5685597> (accessed 17 October 2018).
2. Rumyantsev K.E. Sistemy kvantovogo raspredeleniya klyucha: monografiya [Systems of quantum key distribution: monograph]. Taganrog: Izd-vo TTI YuFU, 2011, 264 s.
3. Fizika kvantovoy informatsii: Kvantovaya kriptografiya. Kvantovaya teleportatsiya. Kvantovye vychisleniya [Physics of Quantum Information: Quantum Cryptography. Quantum teleportation. Quantum computing], ed. by D. Boumeystera, A. Ekerta, A. T'Saylingera. Moscow: Postmarket, 2002, 376 p.
4. Gisin N., Ribordy G., Tittel W., Zbinden H. Quantum cryptography, *Reviews of Modern Physics*, 2002, Vol. 74, No. 1, pp. 145-195.
5. Bennett C., Brassard G. Quantum cryptography: Public key distribution and coin tossing, *Proceedings of IEEE international conference on computers, systems and signal processing. Bangalore. India*. New York: Institute of Electrical and Electronics Engineers, 1984, pp. 175-179.
6. Shor P.W., Preskill J. Simple proof of security of the BB84 quantum key distribution protocol, *Physical Review Letters*, 2000, Vol. 85, pp. 441-444.
7. Rumyantsev K.E. Sinkhronizatsiya v sisteme kvantovogo raspredeleniya klyucha s avtomaticheskoy kompensatsiyey polarizatsionnykh iskazheniy [Synchronization in a quantum key distribution system with automatic compensation of polarization distortions], *Telekommunikatsii* [Telecommunications], 2017, No. 2, pp. 32-40.
8. Rumyantsev K.E., Plenkin A.P. Bezopasnost' rezhima sinkhronizatsii sistemy kvantovogo raspredeleniya klyuchey [Security of the synchronization mode of a system of quantum key distribution], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, No. 5 (166), pp. 135-153.
9. Kurochkin V.L. i dr. Eksperimental'nye issledovaniya v oblasti kvantovoy kriptografii [Experimental research in the field of quantum cryptography], *Fotonika* [Photonics], 2012, Vol. 5, pp. 54-66.
10. Mironov Y.K., Rumyantsev K.E. Single-Photon Algorithm for Synchronizing the System of Quantum Key Distribution with Polling Sections of a Fiber-Optic Line, *Futuristic Trends in Networks and Computing Technologies*, 2020, pp. 87-97. DOI: https://doi.org/10.1007/978-981-15-4451-4_8.
11. Rumyantsev K.E., Linenko P.D., Shakir H.H.-Sh. Evaluation of the Influence of the Dispersion Properties of a Fiber-Optic Line on the Efficiency of an Algorithm for Single-Photon Synchronization of Quantum Key Distribution System, *Conference Proceedings - 2019 Radiation and Scattering of Electromagnetic Waves, RSEMW 2019*, 2019, pp. 392-395. DOI: 10.1109/RSEMW.2019.8792769.
12. Rumyantsev K.E., Rudinskiy E.A. Dvukhetapnyy vremennyy algoritm sinkhronizatsii v sisteme kvantovogo raspredeleniya klyucha s avtomaticheskoy kompensatsiyey polarizatsionnykh iskazheniy [Two-stage time synchronization algorithm in a quantum key distribution system with automatic compensation for polarization distortion], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2017, No. 5 (190), pp. 75-89.
13. Rumyantsev K., Rudinsky E. Parameters of the two-stage synchronization algorithm for the quantum key distribution system, *Proceedings of the 10th International Conference on Security of Information and Networks (SIN'17)*, 2017, pp. 140-147. DOI: 10.1145/3136825.3136888.
14. Lindsey W.C. Synchronization Systems in Communication and Control. Prentice-Hall, Englewood Cliffs, New Jersey, 1972.
15. Stiffler Dzh. Teoriya sinkhronnoy svyazi [Synchronous communication theory]: transl. from engl., ed. by E.M. Gabdulina. Moscow: Svyaz', 1975.
16. Rumyantsev K.E., Shakir Khayder khusseyn. Ogranicheniya na dal'nost' dvukhetapnoy sinkhronizatsii v avtokompensatsionnoy sisteme kvantovogo raspredeleniya klyucha [Restrictions on the range of two-stage synchronization in the autocompensation system of quantum key distribution], *Telekommunikatsii* [Telecommunications], 2019, No. 12, pp. 2-10.
17. Rumyantsev K.E., Plenkin A.P. Sinkhronizatsiya sistemy kvantovogo raspredeleniya klyucha pri ispol'zovanii fotonnykh impul'sov dlya povysheniya zashchishchennosti [Synchronization of the system of quantum key distribution when using photon pulses to increase security], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2014, No. 8 (157), pp. 81-96.

18. *Rumyantsev K.E., Pljonkin A.P.* Preliminary Stage Synchronization Algorithm of Autocompensation Quantum Key Distribution System with an Unauthorized Access Security, *International Conference on Electronics, Information, and Communications (ICEIC)*. 2016. Vietnam, Danang, pp. 1-4. DOI: 10.1109/ELINFOCOM.2016.7562955. WOS:000389518100035. IDS: BG5KP.
19. *Rumyantsev K.E., Plenkin A.P.* Effektivnost' sinkhronizatsii sistemy kvantovogo raspredeleniya klyucha na odnofotonnykh lavinnykh fotodiodakh [Efficiency of synchronization of a system of quantum key distribution based on single-photon avalanche photodiodes], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2016, No. 9 (182), pp. 4-15.
20. *Gal'yardi R.M., Karp Sh.* Opticheskaya svyaz': [Optical communication]. Moscow: Svyaz', 1978, 424 p.

Статью рекомендовала к опубликованию к.т.н. К.Б. Дахкильгова.

Румянцев Константин Евгеньевич – Южный федеральный университет; e-mail: rke2004@mail.ru; 347922, г. Таганрог, ул. Чехова, 2; тел.: 89281827209; кафедра информационной безопасности телекоммуникационных систем; зав. кафедрой; д.т.н.; профессор.

Миронов Яков Константинович – e-mail: tmiyap117@gmail.com; 347922, г. Таганрог, пер. Некрасовский, 19; тел.: 89285723456; аспирант.

Миронова Полина Демьяновна – e-mail: linenkopdem@gmail.com; тел.: 89081924053; аспирант.

Rumyantsev Konstantin Evgenievich – Southern Federal University; e-mail: rke2004@mail.ru; 2, Chekhov street, Taganrog, 347922, Russia; phone: +79281827209; the department of information security of telecommunication systems; head of department; dr .of eng. sc.; professor.

Mironov Yakov Konstantinovich – e-mail: tmiyap117@gmail.com; 19, Nekrasovsky lane, Taganrog, 347922, Russia; phone: +79285723456; graduate student.

Mironova Polina Demyanovna – e-mail: linenkopdem@gmail.com; phone: +79081924053; graduate student.

ПРАВИЛА ОФОРМЛЕНИЯ РУКОПИСЕЙ

1. Объем статьи должен быть не менее 12 и не более 18 страниц. Формат (А 4). Редактор **Word 7 for Windows**, шрифт Times New Roman, размер 14, интервал 1,5. Авторы представляют в редакцию 1 экз. статьи и идентичный электронный вариант.

2. Названию статьи предшествует индекс УДК, соответствующий заявленной теме.

3. Текст статьи начинается с названия статьи (на русском и английском языках), фамилии, имени и отчества автора (полностью) и снабжается аннотацией на русском и английском языках объемом **не менее 250-300 слов**. В тексте аннотации указывается цель, задачи исследования и краткие выводы. В аннотации **не следует** давать ссылку на номер публикации в списке литературы к статье. После аннотаций приводятся ключевые слова (словосочетания), несущие в тексте основную смысловую нагрузку (на русском и английском языках).

4. В тексте статьи следует использовать минимальное количество таблиц и иллюстраций. Рисунок должен иметь объяснения значений всех компонентов, порядковый номер, название, расположенное под рисунком. В тексте на рисунок дается ссылка. Таблица должна иметь порядковый номер, заголовок, расположенный над ней. Данные таблиц и рисунков не должны дублировать текст. Формулы должны быть набраны **в редакторе формул Word 7 for Windows**.

5. Цитаты тщательно сверяются с первоисточником и визируются автором на обратной стороне последней страницы: "Цитаты и фактический материал сверены". Подпись, дата.

6. Наличие пристатейного библиографического списка на русском и английском языках обязательно. **Ссылок должно быть не менее 20-ти**, из них на зарубежные источники – не менее 35 %. В тексте ссылки должны быть в квадратных скобках.

Примеры оформления литературы: а) для книг: фамилия, инициалы автора(ов), полное название книги, место, год издания, страницы; б) для статей: фамилия и инициалы автора(ов), полное название сборника, книги, газеты, журнала, где опубликована статья, место и год издания (сборника, книги), номер (для журнала), год и дата (для газеты), выпуск, часть (для сборника), страницы, на которых опубликована статья. Иностранная литература оформляется по тем же правилам.

Ссылки на неопубликованные работы не допускаются.

7. Рукопись должна быть тщательно вычитана. Редакционная коллегия оставляет за собой право при необходимости сокращать статьи, редактировать и отсылать авторам на доработку.

8. Статьи сопровождаются сведениями об авторе(ах) (фамилия, имя, отчество, ученое звание, должность, место работы, адрес, электронный адрес и номер телефона) на русском и английском языках.

9. Плата с аспирантов за публикацию рукописей не взимается.

Адрес журнала в Интернете: <http://izv-tn.tti.sfedu.ru/>.