



№2-2020

ISSN 1999-9429

ИЗВЕСТИЯ ЮФУ

ТЕХНИЧЕСКИЕ НАУКИ

- Интеллектуальные системы
- Вычислительные и информационно-управляющие системы
- Эволюционное моделирование
- Информационные системы и технологии
- Нанотехнологии

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

№ 2 (212). 2020 г.

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.; Стемпковский 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.

Ответственные за выпуск: Клевцов С.И., Самойлов А.Н.

Технический редактор Ярошевич Н.В.

Оригинал-макет выполнен Ярошевич Н.В.

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

Офсетная печать. Усл. печ. л. – 22,0. Уч.-изд. л. – 19,5.

Заказ № . 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. ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

В.В. Курейчик, А.Е. Глушенко МНОГОУРОВНЕВЫЙ ПОДХОД ДЛЯ РЕШЕНИЯ ЗАДАЧИ ТРЕХМЕРНОЙ УПАКОВКИ БОЛЬШОЙ РАЗМЕРНОСТИ.....	6
Д.Ю. Кравченко, Ю.А. Кравченко, В.В. Марков ГИБРИДНЫЙ БИОИНСПИРИРОВАННЫЙ АЛГОРИТМ ОТОБРАЖЕНИЯ ОНТОЛОГИЙ В ЗАДАЧАХ ИЗВЛЕЧЕНИЯ И УПРАВЛЕНИЯ ЗНАНИЯМИ.....	16
В.С. Корнилов, В.М. Глушань, А.Ю. Лозовой ОПРЕДЕЛЕНИЕ НАБОРА УСЛОВИЙ ДЛЯ АВТОМАТИЧЕСКОГО ПОИСКА ОПТИМАЛЬНОГО ВАРИАНТА ГИБРИДНОГО МАШИННОГО ПЕРЕВОДА ТЕКСТА НА ГРАФЕМНОМ УРОВНЕ	29
Н.К. Полуянович, Н.В. Азаров, А.В. Огреничев, М.Н. Дубяго МНОГОЭТАПНЫЙ МЕТОД КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ ТЕМПЕРАТУРНЫХ РЕЖИМОВ В СИЛОВОМ КАБЕЛЕ	38
Н.К. Полуянович, М.Н. Дубяго АНАЛИЗ И ВЫБОР МЕТОДИКИ В РЕШЕНИИ ЗАДАЧ ИНТЕЛЛЕКТУАЛИЗАЦИИ СИСТЕМ ПРОГНОЗИРОВАНИЯ ТЕРМОФЛУКТУАЦИОННЫХ ПРОЦЕССОВ В КАБЕЛЬНЫХ СЕТЯХ	52
В.В. Бова, Э.В. Кулиев, С.Н. Щеглов ОЦЕНКА ЭФФЕКТИВНОСТИ МЕТОДА ПОИСКА АССОЦИАТИВНЫХ ПРАВИЛ ДЛЯ ЗАДАЧ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ	66
Д.Ю. Кравченко, Ю.А. Кравченко, В.В. Курейчик, А.Э. Саак БИОИНСПИРИРОВАННЫЙ МЕТОД ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ ДЛЯ ДИСПЕТЧЕРИЗАЦИИ ПОТОКОВ ПАРАЛЛЕЛЬНЫХ ЗАЯВОК В GRID- СИСТЕМАХ.....	79

РАЗДЕЛ II. ВЫЧИСЛИТЕЛЬНЫЕ И ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

А.С. Синицын СИНТЕЗ СИНЕРГЕТИЧЕСКОЙ СИСТЕМЫ РАЗРЫВНОГО УПРАВЛЕНИЯ АКТИВНОЙ ПОДВЕСКОЙ АВТОМОБИЛЯ	89
Г.Е. Веселов, Алин Ингабире СИНЕРГЕТИЧЕСКИЙ СИНТЕЗ ЗАКОНА УПРАВЛЕНИЯ БПЛА В УСЛОВИЯХ ВЕТРОВЫХ ВОЗМУЩЕНИЙ С ВХОДНЫМИ ОГРАНИЧЕНИЯМИ	101
А.Н. Целых, В.С. Васильев, Л.А. Целых ПРОБЛЕМА ВЫБОРА ДЕМПИНГ-ФАКТОРА В МОДЕЛИ ЭФФЕКТИВНЫХ УПРАВЛЕНИЙ ДЛЯ НАПРАВЛЕННЫХ ВЗВЕШЕННЫХ ЗНАКОВЫХ ГРАФОВ	112
А.Л. Береснев, М.А. Береснев РАЗРАБОТКА МАТЕМАТИЧЕСКОЙ МОДЕЛИ ДЕТОНАЦИИ В СОВРЕМЕННОМ ДВС	123
В.А. Павский, К.В. Павский МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДЛЯ РАСЧЕТА ПОКАЗАТЕЛЕЙ НАДЕЖНОСТИ МАСШТАБИРУЕМЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ С УЧЕТОМ ВРЕМЕНИ ПЕРЕКЛЮЧЕНИЯ.....	134

РАЗДЕЛ III. ЭВОЛЮЦИОННОЕ МОДЕЛИРОВАНИЕ

Б.К. Лебедев, О.Б. Лебедев, Е.О. Лебедева ЭВОЛЮЦИОННЫЙ АЛГОРИТМ РАЗБИЕНИЯ МЕТОДОМ КРИСТАЛЛИЗАЦИИ РОССЫПИ АЛЬТЕРНАТИВ	146
А.А. Могилев, В.М. Курейчик МОДИФИЦИРОВАННЫЙ ГЕНЕТИЧЕСКИЙ АЛГОРИТМ ПЛАНИРОВАНИЯ ПРОЕКТОВ, РЕАЛИЗОВАННЫЙ С ИСПОЛЬЗОВАНИЕМ ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ.....	157

Э.В. Кулиев, А.А. Лежебоков, Семенова М.М., В.А. Семенов ПОДХОД К КОДИРОВАНИЮ РЕШЕНИЙ В ЭВОЛЮЦИОННЫХ МЕТОДАХ ДЛЯ СОЗДАНИЯ ИНСТРУМЕНТАЛЬНОЙ ПЛАТФОРМЫ ПРОЕКТИРОВАНИЯ	169
Е.В. Данильченко, В.И. Данильченко, В.М. Курейчик КЛАССИФИКАЦИЯ И АНАЛИЗ ЭВОЛЮЦИОННЫХ МЕТОДОВ КОМПОНОВКИ БЛОКОВ ЭВА	179
РАЗДЕЛ IV. ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ	
В.Н. Гридин, В.И. Анисимов, С.А. Васильев МЕТОДЫ ПОВЫШЕНИЯ ПРОИЗВОДИТЕЛЬНОСТИ СОВРЕМЕННЫХ ВЕБ-ПРИЛОЖЕНИЙ	193
Д.Е. Чикрин, А.А. Егорчев, Д.В. Ермаков МЕТОДОЛОГИЯ S.M.A.R.T.E.S.T. H-GQM ДЛЯ КОНТРОЛИРУЕМОЙ ЭВОЛЮЦИИ СИСТЕМ ADAS	200
А.И. Рыбаков, Р.Е. Кротов, С.А. Кокин АДАПТАЦИЯ ИНФОРМАЦИОННО-ТЕХНИЧЕСКИХ ХАРАКТЕРИСТИК К ПОСТОЯННО ИЗМЕНЯЮЩИМСЯ ПАРАМЕТРАМ ИОНОСФЕРНОГО РАСПРОСТРАНЕНИЯ	209
Л.К. Бабенко, И.Д. Русаловский БИБЛИОТЕКА ПОЛНОСТЬЮ ГОМОМОРФНОГО ШИФРОВАНИЯ ЦЕЛЫХ ЧИСЕЛ	218
РАЗДЕЛ V. НАНОТЕХНОЛОГИИ	
А.А. Гелдаш, Л.Э. Левенец, Е.Ю. Гусев, В.Н. Джуплин ИССЛЕДОВАНИЯ РЕЖИМОВ ФОРМИРОВАНИЯ НАНОКРИСТАЛЛИЧЕСКИХ ПЛЁНОК ZnO:Ga МЕТОДОМ МАГНЕТРОННОГО РАСПЫЛЕНИЯ	228
И.В. Писаренко, Е.А. Рындин ФОТОДЕТЕКТОР С УПРАВЛЯЕМОЙ ПЕРЕДИСЛОКАЦИЕЙ: ДИФфуЗИОННО-ДРЕЙФОВАЯ МОДЕЛЬ И ИСПОЛЬЗОВАНИЕ В ОПТИЧЕСКИХ МЕЖСОЕДИНЕНИЯХ	238

CONTENT

SECTION I. INTELLIGENT SYSTEM

V.V. Kureichik, A.E. Glushchenko MULTILEVEL APPROACH FOR HIGH DIMENSIONAL 3D PACKING PROBLEM.....	6
D.Yu. Kravchenko, Yu.A. Kravchenko, V.V. Markov HYBRID BIOINSPIRED ALGORITHM FOR ONTOLOGIES MAPPING IN THE TASKS OF EXTRACTION AND KNOWLEDGE MANAGEMENT.....	17
V.S. Kornilov, V.M. Glushan, A.Yu. Lozovoy DETERMINING A SET OF CONDITIONS FOR AUTOMATICALLY FINDING THE BEST OPTION FOR HYBRID MACHINE TRANSLATION OF TEXT AT THE LEVEL OF GRAPHEMES	29
N.K. Poluyanovich, N. V. Azarov, A.V. Ogrenichev, M. N. Dubyago MULTI-STAGE METHOD FOR SHORT-TERM FORECASTING OF TEMPERATURE CHANGES MODES IN THE POWER CABLE.....	38
N.K. Poluyanovich, M.N. Dubyago ANALYSIS AND SELECTION OF METHODOLOGIES IN THE SOLUTION OF THE PROBLEMS OF INTELLECTUALIZATION IN SYSTEMS FOR PROGNOSIS OF THERMOFLUCTUATION PROCESSES IN CABLE NETWORKS	52
V.V. Bova, E.V. Kuliev, S.N. Shcheglov ESTIMATING THE EFFECTIVENESS OF THE METHOD FOR SEARCHING THE ASSOCIATIVE RULES FOR THE TASKS OF PROCESSING BIG DATA	67
D.Yu. Kravchenko, Yu.A. Kravchenko, V.V. Kureichik, A.E. Saak BIOINSPIRED SIMULATION METHOD FOR SCHEDULING OF PARALLEL FLOWS APPLICATIONS IN GRID-SYSTEMS.....	79

SECTION II. COMPUTING AND INFORMATION AND CONTROL SYSTEMS

A.S. Sinitsyn SYNERGETIC SYNTHESIS OF SLIDING MODE CONTROL FOR VEHICLE'S ACTIVE SUSPENSION SYSTEM	89
G.E. Veselov, Alin Ingabire SYNERGETIC SYNTHESIS OF CONTROL LAW FOR UAV IN THE PRESENCE OF WIND DISTURBANCES WITH INPUT CONSTRAINTS.....	102
A.N. Tselykh, V.S. Vasilev, L.A. Tselykh THE PROBLEM OF CHOOSING A DUMPING FACTOR IN THE EFFECTIVE CONTROL MODEL FOR DIRECTED WEIGHTED SIGNED GRAPHS.....	113
A.L. Beresnev, M.A. Beresnev DEVELOPMENT OF KNOCK MATHEMATICAL MODEL FOR MODERN IC ENGINE.....	124
V.A. Pavsky, K.V. Pavsky MATHEMATICAL MODEL FOR CALCULATING RELIABILITY INDICES OF SCALABLE COMPUTER SYSTEMS WITH SWITCHING TIME.....	134

SECTION III. EVOLUTIONARY MODELING

B.K. Lebedev, O.B. Lebedev, E.O. Lebedeva EVOLUTION ALGORITHM FOR PARTITION BY METHOD OF CRYSTALLIZATION OF ALTERNATIVES FIELD	147
A.A. Mogilev, V.M. Kureichik MODIFIED GENETIC PROJECT PLANNING ALGORITHM IMPLEMENTED WITH THE USE OF CLOUD COMPUTING.....	158

E.V. Kuliev, A.A. Lezhebokov, M.M. Semenova, V.A. Semenov SOLUTIONS' ENCODING IN EVOLUTIONARY METHODS FOR INSTRUMENTAL DESIGN PLATFORM	170
E.V. Danilchenko, V.I. Danilchenko, V.M. Kureychik CLASSIFICATION AND ANALYSIS OF EVOLUTIONARY METHODS OF EVA BLOCK LAYOUT.....	180

SECTION IV. INFORMATION SYSTEMS AND TECHNOLOGIES

V.N. Gridin, V.I. Anisimov, S.A. Vasilev METHODS OF IMPROVING PERFORMANCE OF MODERN WEB-APPLICATIONS	193
D.E. Chickrin, A. A. Egorchev, D.V. Ermakov EVOLVABLE ADAS: H-GQM S.M.A.R.T.E.S.T. APPROACH.....	201
A.I. Rybakov, R.E. Krotov, S.A. Kokin ADAPTATION OF INFORMATION AND TECHNICAL CHARACTERISTICS TO THE CONSTANTLY CHANGING PARAMETERS OF IONOSPHERIC PROPAGATION	210
L.K. Babenko, I.D. Rusalovsky THE LIBRARY OF FULLY HOMOMORPHIC ENCRYPTION OVER THE INTEGERS	219

SECTION V. NANOTECHNOLOGIES

A.A. Geldash, L.E. Levenets, E.Y. Gusev, V.N. Dzhuplin STUDY OF THE MODES OF FORMING ZnO: Ga NANOCRYSTALLINE FILMS BY MAGNETRON SPUTTERING.....	228
I.V. Pisarenko, E.A. Ryndin PHOTODETECTOR WITH CONTROLLED RELOCATION: DRIFT-DIFFUSION MODEL AND APPLICATION IN OPTICAL INTERCONNECTIONS.....	238

Раздел I. Интеллектуальные системы

УДК 004.896

DOI 10.18522/2311-3103-2020-2-6-16

В.В. Курейчик, А.Е. Глущенко

МНОГОУРОВНЕВЫЙ ПОДХОД ДЛЯ РЕШЕНИЯ ЗАДАЧИ ТРЕХМЕРНОЙ УПАКОВКИ БОЛЬШОЙ РАЗМЕРНОСТИ*

Рассмотрена одна из важных комбинаторных задач оптимизации – задача трехмерной упаковки разноразмерных элементов в объеме. Она относится к классу NP-сложных и трудных оптимизационных задач. В работе приведена и описана постановка задачи трехмерной упаковки в объеме, введена комбинированная целевая функция учитывающая все ограничения. В связи со сложностью данной задачи предлагается многоуровневый подход заключающийся в разделении задачи трехмерной упаковки на 3-и подзадачи и решения каждой подзадачи в строгом порядке. При этом для каждой из подзадач определен уникальный набор объектов, не повторяющихся в остальных подзадачах. Для реализации многоуровневого подхода авторами разработан комбинированный биоинспирированный алгоритм, основанный на эволюционном и генетическом поиске. Такой подход позволяет значительно сократить время получения результата, частично решить проблему предварительной сходимости алгоритмов и получить наборы квазиоптимальных решений за полиномиальное время. Разработан программный комплекс и реализованы на ЭВМ алгоритмы автоматизированной трехмерной упаковки на основе комбинированного биоинспирированного поиска. Проведен вычислительный эксперимент на тестовых примерах (бенчмарках). Качество упаковки, полученное, на основе разработанного комбинированного биоинспирированного алгоритма, в среднем на 5 % превосходит результаты упаковки, полученные с использованием известных алгоритмов, а время решения меньше от 5 % до 20 %, что говорит об эффективности предложенного подхода. Проведенные серии тестов и экспериментов позволили уточнить теоретические оценки временной сложности алгоритмов упаковки. В лучшем случае временная сложность алгоритмов $O(n^2)$, в худшем случае – $O(n^3)$.

Трехмерная упаковка; упаковка в контейнеры; многоуровневый подход; комбинированный биоинспирированный алгоритм; генетический алгоритм; эволюционный алгоритм.

V.V. Kureichik, A.E. Glushchenko

MULTILEVEL APPROACH FOR HIGH DIMENSIONAL 3D PACKING PROBLEM

The article considers one of the important combinatorial optimization problems, the problem of 3D packing of different elements in a fixed volume. It belongs to the class of NP-complex and difficult optimization problems. The paper presents and describes the formulation of the 3D packing problem, introduces a combined objective function that takes into account all the restrictions. Due to the complexity of this task, a multilevel approach is proposed. It is consisting in dividing the 3D packing problem into 3 subtasks and solving each subtask in a strict order. Moreover, for each of the subtasks a unique set of objects is defined that are not repeated in the remaining subtasks. To implement a multi-level approach, the authors developed a combined bio-inspired algorithm based on

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-01-00148.

evolutionary and genetic search. This approach can significantly reduce the time to obtain the result, partially solve the problem of preliminary convergence of the algorithms and obtain sets of quasi-optimal solutions in polynomial time. A software package was developed and computer-based algorithms for automated 3D packaging based on a combined bio-inspired search were implemented. A computational experiment was conducted on test examples (benchmarks). The packaging quality obtained on the basis of the developed combined bio-inspired algorithm is on average 5 % higher than the packaging results obtained using known algorithms, and the solution time is less than 5 % to 20 %, which indicates the effectiveness of the proposed approach. The series of tests and experiments carried out made it possible to refine the theoretical estimates of the time complexity of the packaging algorithms. In the best case the time complexity of the $O(n^2)$ algorithms; in the worst, $O(n^3)$.

Three-dimensional packaging; packing in containers; multi-level approach; combined bio-inspired algorithm; genetic algorithm; evolutionary algorithm.

Введение. Массовое увеличение товарооборота оказывает влияние и на транспортировку продуктов в контейнерах, которая ежегодно составляет более 69 % от всех объемов перевозок. Число контейнерных перевозок в период с 2009 по 2019 г.г. выросла более чем на 34 %. Рост связан с постоянным наращиванием производственных мощностей, способствующих увеличению производства и разнообразия товаров [1–3], в связи с чем возникает необходимость в увеличении объемов грузооборота и уменьшению сокращением времени доставки продуктов конечному покупателю. Путей решения данной проблемы существует несколько: увеличение количества транспортных средств, способных перевозить товары, оптимизация транспортной логистики. Оптимальным путем и, зачастую единственным, является оптимизация (уплотнение) упаковки товаров с помощью новых методов и подходов к решению задачи трехмерной упаковки. Плотность (эффективность) упаковки оказывает большое влияние на стоимость и скорость доставки товаров. Для решения данной задачи, многие компании внедряют системы автоматизированного проектирования (САПР). Снижение материальных затрат, увеличение качества проектирования, снижение сроков проектирования, являются основными целями автоматизации.

Основными критериями грузовых перевозок являются – время обработки заказа, скорость доставки грузов до конечного потребителя, а также возможность перемещения наибольшего количества грузов на одном транспортном средстве.

Трехмерная упаковка является NP-трудной и NP-сложной задачей, не имеющей точных алгоритмов ее решения за полиномиальное время [1–14]. В настоящее время разработанные алгоритмы трехмерной упаковки позволяют получать приемлемые результаты за разумное время только для задач небольшой размерности, в остальных случаях решение задачи может занимать непозволительно много времени.

В связи с вышеизложенным, создание новых подходов, методов и алгоритмов решения задачи трехмерной упаковки большой размерности является актуальной и важной задачей.

Постановка задачи трехмерной упаковки. Исходными данными задачи упаковки являются:

- ♦ область трехмерного пространства;
- ♦ список блоков;
- ♦ набор ограничений.

Дадим формальное описание задачи упаковки грузов в ограниченном трехмерном пространстве. Дана область трехмерного пространства шириной W , длиной L и высотой H . Также дано множество блоков $A=\{a_i\}$ количеством N . Все множество блоков разбито на множество типов $T=\{t_j\}$, количество каждого типа указывается пользователем. Каждый тип определяет заданные параметры каждого блока, такие как, длина, ширина, высота и вес, ориентация по умолчанию [5–12].

Критерий – это отношение суммарного объема упакованных блоков к объему трехмерной области размещения. Таким образом, снижается расход на транспортировку грузов из-за их более компактного распределения. Результаты решения задачи в виде плана погрузки позволяют заказчику за один рейс перевезти большее количество товара. Количество совершаемых рейсов уменьшается и следовательно уменьшается стоимость грузоперевозок и эффективность процесса транспортировки [4].

Каждый тип груза характеризуется кортежем длины 4 $\langle l_i, w_i, h_i, m_i \rangle$, где l_i, w_i, h_i – габаритные размеры элемента, m_i – площадь блока. В ходе решения задачи положение элементов, обозначенное заданным типом, в пространстве задается множеством $S = \{s_i = (\langle x_{1i}, y_{1i}, z_{1i} \rangle, \langle x_{2i}, y_{2i}, z_{2i} \rangle) \mid i=1, 2, \dots, n\}$, где $\langle x_{1i}, y_{1i}, z_{1i} \rangle, \langle x_{2i}, y_{2i}, z_{2i} \rangle$ – координаты углов элемента – самого близкого к началу осей координат и самого удаленного соответственно [5–12].

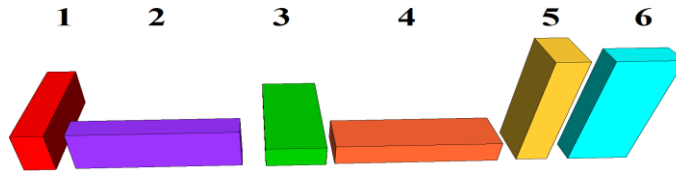


Рис. 1. Пример возможного расположения блоков в пространстве

Так же имеется параллелепипед $M = \{L, W, H\}$, где L, W, H – габаритные размеры области упаковки.

Используя критерий V – это отношение объема упакованных блоков к объему параллелепипеда $V_{\text{кон}}$, требующийся для упаковки всех элементов, введем следующие ограничения [12]:

1) ни один элемент не может выходить за границы заданного объема, т.е. удовлетворять системе неравенств (формула 1):

$$\begin{cases} x_{1i} \geq 0; \\ y_{1i} \geq 0; \\ z_{1i} \geq 0; \\ x_{2i} \leq L_x; \\ y_{2i} \leq W_y; \\ z_{2i} \leq H_z. \end{cases} \quad (1)$$

2) суммарный объем элементов не должен превышать объема области упаковки (формула 2):

$$\sum_{i=1}^n (l_i \cdot w_i \cdot h_i) \leq L_x \cdot W_y \cdot H_z. \quad (2)$$

3) отсутствие пересечений [7] (т.е. в одну область пространства нельзя поместить два объекта) (формула 3):

$$\begin{aligned} & (x_{2i} \cdot x_{1j} \& y_{2i} \geq y_{1j} \& z_{2i} \geq z_{1j}) \vee (x_{2i} \cdot x_{1j} \& y_{2i} \geq y_{1j} \& z_{2i} \geq z_{1j}) \vee \\ & (x_{2i} \cdot x_{1j} \& y_{2i} \geq y_{1j} \& z_{2i} \geq z_{1j}) \vee (x_{2i} \cdot x_{1j} \& y_{2i} \geq y_{1j} \& z_{2i} \geq z_{1j}) \vee \\ & (x_{2i} \cdot x_{1j} \& y_{2i} \geq y_{1j} \& z_{2i} \geq z_{1j}) \vee (x_{2i} \cdot x_{1j} \& y_{2i} \geq y_{1j} \& z_{2i} \geq z_{1j}) \vee \\ & (x_{2i} \cdot x_{1j} \& y_{2i} \geq y_{1j} \& z_{2i} \geq z_{1j}) = 1 \forall i \leq n, j \leq (i \neq j). \end{aligned} \quad (3)$$

4) суммарный вес всех блоков не может превышать заданный лимит (формула 4):

$$\sum G_{\text{бл}} \leq G_{\text{лим}}, \quad (4)$$

где $G_{\text{бл}}$ – вес блока; $G_{\text{лим}}$ – максимальный вес всех блоков.

- 5) все элементы должны опираться своим основанием на поверхность;
- 6) элементы должны лежать перпендикулярно заданной области.

Критерием оптимизации является объем, занимаемый блоками. Целевая функция имеет вид (формула 5):

$$F = \frac{\sum V_{\text{бл}}}{V_{\text{кон}}} \rightarrow 1, \quad (5)$$

где $V_{\text{бл}}$ – объем каждого блока, $V_{\text{кон}}$ – объем контейнера.

Это означает, что необходимо стремиться к уменьшению пустот в области упаковки. Максимально эффективный результат решения задачи, когда целевая функция (ЦФ) равна 1, область упаковки заполнена на 100 %.

Данная ЦФ не полностью учитывает всех критериев задачи трехмерной упаковки контейнеров, поэтому в работе авторы предлагают использовать комбинированную целевую функцию:

$$\text{ЦФ} = \frac{\sum_{i=1}^n V_i}{V} \times Q^1 + P^2 \times Q^2 + P^3 \times Q^3 + P^4 \times Q^4, \quad (6)$$

где P^2 – отношение числа упакованных блоков в контейнер к числу всех блоков, подлежащих упаковке; P^3 – число блоков одного размера, упакованных подряд так, что их стороны соприкасаются; P^4 – отношение горизонтально ориентированных блоков к числу всех блоков, подлежащих упаковке; Q^1 – коэффициент важности заполненности контейнера; Q^2 – коэффициент важности числа упакованных блоков в контейнер; Q^3 – коэффициент важности упаковки блоков одного размера с одинаковыми сторонами друг к другу; Q^4 – коэффициент важности упаковки блоков горизонтально ориентированно. Коэффициенты Q^1, Q^2, Q^3, Q^4 устанавливаются лицом принимающим решение на основе опыта эксперта. P^2, P^3 и P^4 вычисляются при расчёте ЦФ. Задача – максимизация целевой функции: $\text{ЦФ} \rightarrow \max$.

Многоуровневый подход для решения задачи трехмерной упаковки большой размерности. В настоящее время выделяют следующие подходы к решению NP-сложных задач. Первый подход – это упрощение алгоритмов, т.е. снижение их вычислительной сложности за счет применения эвристических процедур [5, 6, 8–14]. Второй подход – это упрощение решаемых задач за счет уменьшения их размерности или их декомпозиции [15, 16].

Используя гибридизацию этих двух подходов, в работе используется метод, который позволяет реализовывать различные виды алгоритмов на разных уровнях поиска.

Многоуровневый подход заключается в разделении задачи трехмерной упаковки на 3-и подзадачи и решения каждой подзадачи в строгом порядке [17]. При этом для каждой из подзадач определен уникальный набор объектов, не повторяющихся в остальных подзадачах. Такой подход позволяет значительно сократить время работы алгоритмов, а также позволит сохранить лучший результат решения 1-й подзадачи, который остается статичным. Для реализации многоуровневого подхода авторы предлагают комбинированный биоинспирированный алгоритм рис. 2 [12, 18].

Рассмотрим более подробно этапы работы комбинированного биоинспирированного алгоритма трехмерной упаковки:

- ♦ на первом этапе лицо принимающее решение (ЛПР) вводит основные параметры: размеры области упаковки, размеры и количество объектов упаковки, параметры сортировки объектов в группы, вероятности применения оператора кроссинговера и операторов мутации, а также критерий остановки;

- ♦ на втором этапе происходит генерация начальной популяции. Сгенерированные блоки изменяют свою ориентацию в пространстве, если блок повернуть нельзя, то происходит регенерация с последующей проверкой;
- ♦ на третьем этапе оцениваются габариты каждого блока (длина, ширина, высота, объем);

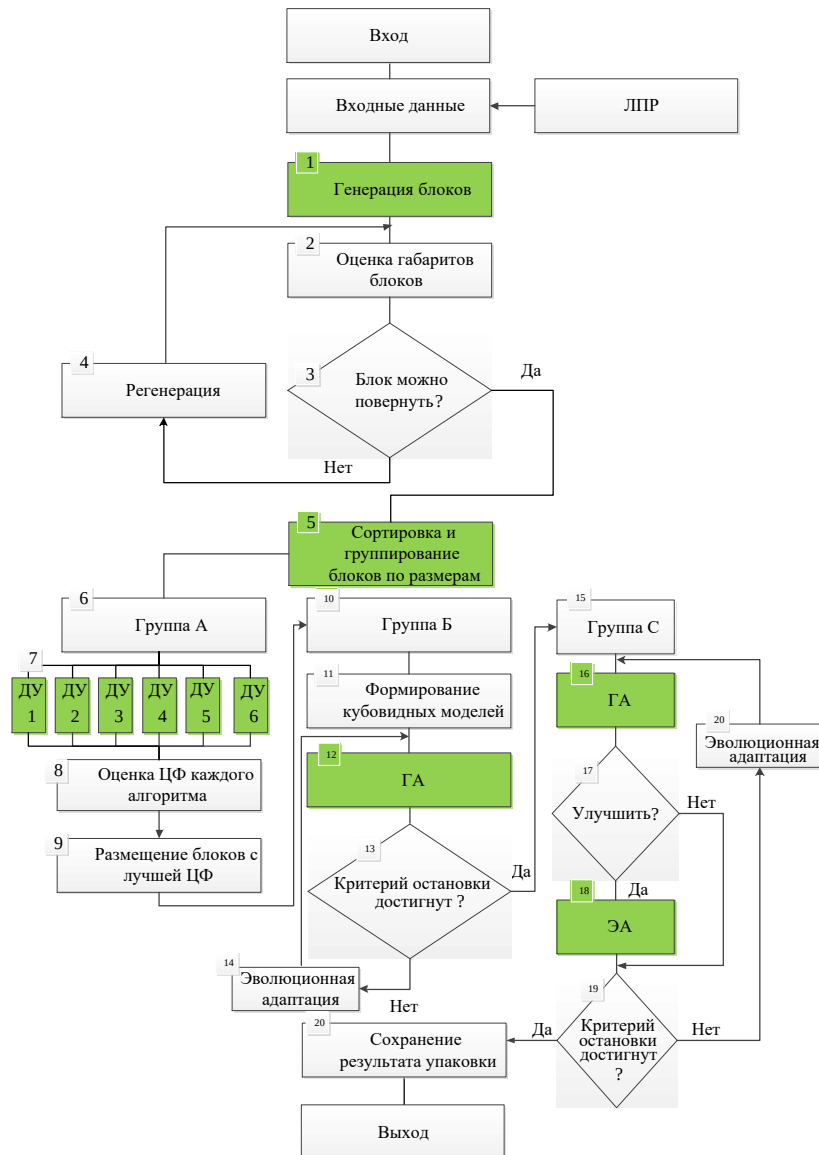


Рис. 2. Укрупненная схема комбинированного биоинспирированного алгоритма

- ♦ на четвертом этапе проводится сортировка блоков. Рассмотрим сортировку блоков более подробно:

1) Блоки одинаковых размеров объединяются в отдельные подгруппы ($A_1, A_2, A_3, \dots, A_i$), которые помещаются в группу А. Блоки без пары помещаются в группу С.

2) Подгруппы группы А сортируются по объему. Подгруппы, которые содержат четное количество блоков и объем которых меньше $Y\%$ (устанавливается ЛПР) помещаются в группу В. Подгруппы, которые содержат нечетное количество блоков и объем которых меньше $Y\%$ (устанавливается ЛПР) помещаются в группу С.

♦ на пятом этапе проводится размещение блоков группы А. Размещение проводится вдоль левой стенки области упаковки. На данном этапе решается 6 отдельных задач двумерной упаковки по завершению работы которых, выбирается лучшее значение ЦФ. Данный этап повторяется для всех подгрупп группы А. На данном этапе происходит быстрое и эффективное размещение большого количества блоков за минимальное время. На рис. 3 представлен пример размещения подгрупп группы А.

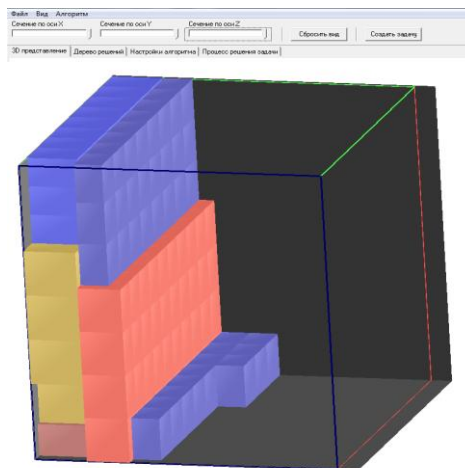


Рис. 3. Пример размещения подгрупп группы А

♦ на шестом этапе подгруппы группы Б сортируются по объему. Далее блоки подгруппы Б объединяются в кубовидные модели. Кубовидная модель содержит 2 блока, которые присоединяются друг к другу по самой длинной стороне. Данный подход позволяет уменьшить количество блоков, которое необходимо разместить, в 2 раза;

♦ на седьмом этапе сформированные кубовидные модели размещаются при помощи модифицированного генетического алгоритма. Процесс повторяется до тех пор, пока все блоки не будут размещены или критерий остановки не будет достигнут;

♦ на восьмом этапе все не размещенные блоки перемещаются в группу С, которая становится единственной. Далее блоки группы С сортируются по объему. Далее отсортированное количество блоков i группируется на k групп по объему, где i пропорционально k ;

♦ на девятом и десятом этапах отсортированные подгруппы группы С размещаются при помощи модифицированных генетического и эволюционного алгоритмов (ГА и ЭА) [19]. Размещение происходит с подгруппы блоков имеющих наибольший объем. В случае невозможности разместить блок осуществляется переход к блоку из этой же группы с шагом n (n – количество блоков, которое пропускается в выбранной группе при переходе к следующему блоку, при невозможности размещения предыдущего. Значение n определяет ЛПР), в случае если ни один из блоков группы невозможно разместить в области, осуществляется переход к следующей группе блоков. Процесс повторяется до тех пор, пока все блоки не будут размещены или критерий остановки не будет достигнут.

Применение предложенной схемы поиска имеет следующие преимущества: позволяет быстрее и эффективнее находить оптимальные решения, работать не с одним решением, а с множеством альтернативных решений, присутствует возможность рассматривать более перспективные и удалять заранее неперспективные решения.

Вычислительный эксперимент. Разработана программная среда для решения задачи трехмерной упаковки. При построении комплекса программ использовались пакеты Visual C++, Borland C++, Builder. Отладка и тестирование разработанных алгоритмов выполнялось на компьютере типа IBM PC с процессором ryzen 5 3600x с ОЗУ-16Гб. Проведен вычислительный эксперимент. Для определения эффективности разработанного комбинированного алгоритма были проведены исследования времени и качества решения для разного набора тестовых примеров (бенчмарок), различающихся количеством блоков [12, 20, 21]. Результаты проведенных исследований представлены в табл. 1 и 2, а также на рис. 4 и 5.

Таблица 1

Сравнение качества решений разработанного алгоритма с аналогами

Количество блоков	ЭА	ГА	Разработанный алгоритм	Тестовые примеры esicup № 1	Тестовые примеры esicup № 2
500	0,63	0,65	0,55	0,62	0,6
1000	0,66	0,71	0,79	0,81	0,8
1500	0,73	0,77	0,86	0,83	0,81
2000	0,74	0,78	0,9	0,88	0,86
2500	0,76	0,81	0,92	0,9	0,88

Таблица 2

Сравнение времени работы разработанного алгоритма с аналогами

Количество блоков шт.	ЭА	ГА	Разработанный алгоритм	Тестовые примеры esicup № 1	Тестовые примеры esicup № 2
500	15	22	28	23	25
1000	35	46	69	60	63
1250	64	81	105	120	125
1500	84	97	144	149	155
1750	176	218	261	280	310
2000	224	251	320	370	401

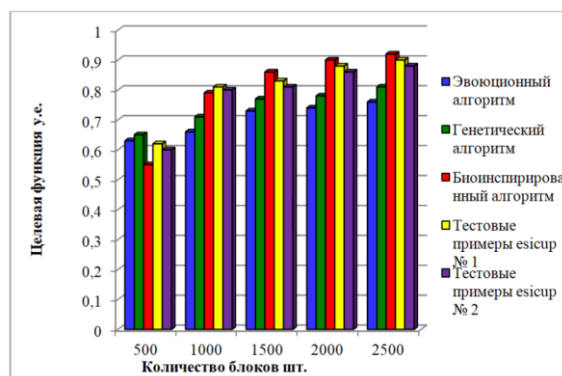


Рис. 4. Гистограмма сравнения значений целевой функции алгоритмов упаковки от числа блоков

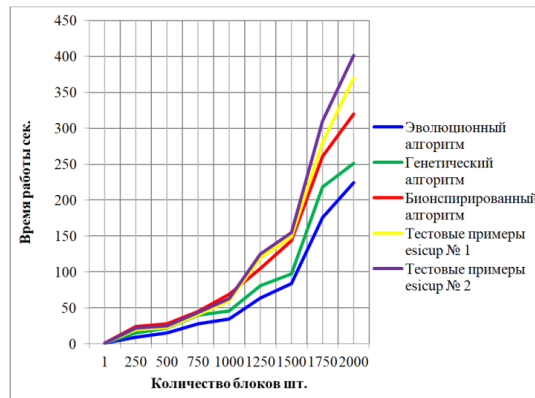


Рис. 5. Гистограмма сравнения времени работы алгоритмов упаковки от числа блоков

В результате анализа представленных таблиц и графиков зависимостей можно сделать вывод, что разработанный комбинированный биоинспирированный алгоритм при упаковке небольшого числа блоков получает менее качественные решения, используя большее количество времени, чем известные алгоритмы, но при упаковке большого числа блоков 1500 и более, разработанный комбинированный биоинспирированный алгоритм позволяет лучше производить упаковку блоков в среднем на 5 %, чем известные алгоритмы, затрачивая при этом меньше времени от 5 % до 20 % в зависимости от числа блоков, что говорит об эффективности предложенного подхода.

Заключение. Предложен многоуровневый подход для решения задачи трехмерной упаковки большой размерности. Отличительной особенностью данного подхода является разбиение задачи на несколько подзадач, с последовательным решением каждой задачи, при этом каждая подзадача имеет уникальный набор блоков для размещения. Для реализации этого подхода разработан комбинированный биоинспирированный алгоритм, позволяющий получать наборы квазиоптимальных решений за полиномиальное время и частично решать проблему предварительной сходимости алгоритмов. Разработана программная среда на языке C++. Проведен вычислительный эксперимент. Проведенные экспериментальные исследования, показали преимущество использования разработанного многоуровневого подхода для решения задач трехмерной упаковки большой размерности (свыше 1500 блоков), по сравнению с известными методами.

Качество упаковки, полученное, на основе разработанного комбинированного биоинспирированного алгоритма, в среднем на 5 % превосходит результаты упаковки, полученные с использованием известных алгоритмов, а время решения меньше от 5 % до 20 %, что говорит об эффективности предложенного подхода.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Koide S., Suzuki S., Degawa S. A Palletize-Planning System for Multiple Kinds of Loads using GA Search and Traditional Search // Intelligent Robots and Systems 95. 'Human Robot Interaction and Cooperative Robots', Proceedings. IEEE. – 1995. – Vol. 3. – P. 510-515.
2. Мухачева Э.А., Мухачева А.С. Технология блочных структур локального поиска оптимума в задачах прямоугольной упаковки // Информационные технологии. – 2004. – № 5. Приложение. – С. 19-31.
3. Юдаков П.В. Задача о трехмерной упаковке и методы ее решения. Обзор // Инженерный вестник. – 2015. – № 06. – С. 552-581.

4. Луцан М.В., Нужнов Е.В. Решение задачи трехмерной упаковки с палетированием контейнеров // Известия ЮФУ. Технические науки. – 2014. – № 7 (156). – С. 196-204.
5. Нужнов Е.В., Барлит А.В. Трехмерная упаковка несвязных элементов на основе эвристических процедур. – Таганрог: Изд-во ТРТУ, 2002. – С. 23-28.
6. Псиола В.В. О приближенном решении 3-х мерной задачи об упаковке на основе эвристики // Интеллектуальные системы. – 2007. – № 11. – Вып. 1-4. – С. 83-101.
7. Bortfeldt A., Wascher G. Constraints in container loading: a state-of-the-art review // European Journal of Operational Research. – 2013. – Vol. 229, Is. 1. – P. 1-20.
8. Гладков Л.А., Гладкова Н.В., Скубриева Е.С. Решение задачи трехмерной упаковки разногабаритных объектов с использованием бионических методов // Известия ЮФУ. Технические науки. – 2012. – № 7 (144). – С. 35-41.
9. Тимофеева О.П., Соколова Э.С., Милов К.В. Генетический алгоритм в оптимизации упаковки контейнеров // Тр. НГТУ им. Р.Е. Алексеева. – 2012. – № 4 (101). – С. 167-172.
10. Gehring H., Bortfeldt A. A genetic algorithm for solving the container loading problem // International Transactions in Operational Research. – 1997. – Vol. 4, Iss. 5-6. – P. 401-418.
11. Курейчик В.В., Заруба Д.В., Запорожец Д.Ю. Применение генетического алгоритма решения задачи трехмерной упаковки // Известия ЮФУ. Технические науки. – 2012. – № 7 (132). – С. 8-14.
12. Курейчик В.В., Глуценко А.Е., Орлов А.Н. Гибридный подход для решения задачи 3-х мерной упаковки // Известия ЮФУ Технические науки. – 2016. – № 6 (179). – С. 45-53.
13. Курейчик В.М., Курейчик Л.В. Комплексный метод упаковки блоков // Информатика, вычислительная техника и инженерное образование. – 2015. – № 1 (21). – С. 17-26.
14. Chauny F.A. Bloc Heuristic for the Container Loading Problem // Grouped'études et de recherche en analyse des decisions, Montréal. – 2005. – P. 1-18.
15. Базилевич Р.П. Декомпозиционные и топологические методы автоматизированного конструирования электронных устройств: монография. – Львов: Вища школа, 1981. – 81 с.
16. Базилевич Р.П. Метод оптимального свертывания схемы – эффективный подход для качественного решения неполиномиальных комбинаторных задач большой и сверхбольшой размерности в автоматизированном конструировании РЭА // Сб. научных трудов «Проблемы разработки перспективных микроэлектронных систем». – М.: ИПИМ РАН, 2005. – С. 94-100.
17. Chan F.T.S., Kumar N., Wong T.C. Three-Dimensional Air-Cargo Loading Problem: An Evolutionary Algorithm Based Approach // Proceedings of the 7th Asia Pacific Industrial Engineering and Management Systems Conference. – 2006. – P. 758-765.
18. Курейчик В.В., Глуценко А.Е. Комбинированный подход для решения задачи 3-х мерной упаковки разногабаритных элементов // XII Всероссийская научная конференция молодых ученых, аспирантов и студентов, информационные технологии, системный анализ и управление (ИТСА и У-2015). – Ростов-на-Дону: Изд-во ЮФУ, 2015. – С. 75-79.
19. Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы. – М.: Физматлит, 2010.
20. Курейчик В.В., Глуценко А.Е., Курейчик Л.В. Программный комплекс комбинированного поиска для решения задачи трехмерной упаковки // Тр. II Всероссийской научно-технической конференции Фундаментальные и прикладные аспекты компьютерных технологий и информационной безопасности. – Таганрог: Изд-во ЮФУ, 2016. – С. 216-220.
21. Жуков Л.А., Корчевская О.В. Метод плоскостей: численный эксперимент для задач двух и трехмерной ортогональной упаковки // Информационные технологии. – 2008. – № 11. – С. 41-45.

REFERENCES

1. Koide S., Suzuki S., Degawa S. A Palletize-Planning System for Multiple Kinds of Loads using GA Search and Traditional Search, *Intelligent Robots and Systems 95. 'Human Robot Interaction and Cooperative Robots', Proceedings. IEEE*, 1995, Vol. 3, pp. 510-515.
2. Mukhacheva E.A., Mukhacheva A.S. Tekhnologiya blochnykh struktur lokal'nogo poiska optimuma v zadachakh pryamougol'noy upakovki [Technology of block structures for local optimal search in rectangular packing problems], *Informatsionnye tekhnologii* [Information technologies], 2004, No. 5. Appendix, pp. 19-31.

3. Yudakov P.V. Zadacha o trekhmernoy upakovke i metody ee resheniya. Obzor [The problem of three-dimensional packaging and methods of its solution. Review], *Inzhenernyy vestnik* [Engineering Bulletin], 2015, No. 06, pp. 552-581.
4. Lutsan M.V., Nuzhnov E.V. Reshenie zadachi trekhmernoy upakovki s paletirovaniem konteynerov [Solving the problem of three-dimensional packaging with container palletization], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2014, No. 7 (156), pp. 196-204.
5. Nuzhnov E.V., Barlit A.V. Trekhmernaya upakovka nesvyaznykh elementov na osnove evristicheskikh protsedur [Three-dimensional packaging of disconnected elements based on heuristic procedures]. Taganrog: Izd-vo TRTU, 2002, pp. 23-28.
6. Psiola B.V. O priblizhennoy reshenii 3-kh mernoy zadachi ob upakovke na osnove evristik [Approximate solution of a 3-dimensional packaging problem based on heuristics], *Intellektual'nye sistemy* [Intelligent system], 2007, No. 11, Issue 1-4, pp. 83-101.
7. Bortfeldt A., Wascher G. Constraints in container loading: a state-of-the-art review, *European Journal of Operational Research*, 2013, Vol. 229, Issue 1, pp. 1-20.
8. Gladkov L.A., Gladkova N.V., Skubrieva E.S. Reshenie zadachi trekhmernoy upakovki raznogabaritnykh ob'ektov s ispol'zovaniem bionicheskikh metodov [Solving the problem of three-dimensional packaging of multi-dimensional objects using bionic methods], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2012, No. 7 (144), pp. 35-41.
9. Timofeeva O.P., Sokolova E.S., Milov K.V. Geneticheskiy algoritm v optimizatsii upakovki konteynerov [Genetic algorithm for optimizing container packaging], *Tr. NGTU im. R.E. Alekseeva* [Proceedings of the NSTU named after R. E. Alekseev], 2012, No. 4 (101), pp. 167-172.
10. Gehring H., Bortfeldt A. A genetic algorithm for solving the container loading problem, *International Transactions in Operational Research*, 1997, Vol. 4, Issue 5-6, pp. 401-418.
11. Kureychik V.V., Zaruba D.V., Zaporozhets D.Yu. Primenenie geneticheskogo algoritma resheniya zadachi trekhmernoy upakovki [Application of a genetic algorithm for solving the problem of three-dimensional packaging], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2012, No. 7 (132), pp. 8-14.
12. Kureychik V.V., Glushchenko A.E., Orlov A.N. Gibriddnyy podkhod dlya resheniya zadachi 3-kh mernoy upakovki [Hybrid approach for solving the problem of 3-dimensional packaging], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2016, No. 6 (179), pp. 45-53.
13. Kureychik V.M., Kureychik L.V. Kompleksnyy metod upakovki blokov [Complex method of packing blocks], *Informatika, vychislitel'naya tekhnika i inzhenernoe obrazovanie* [Computer science, computer engineering and engineering education], 2015, No. 1 (21), pp. 17-26.
14. Chauny F.A. Bloc Heuristic for the Container Loading Problem, *Grouped'études et de recherche en analyse des decisions, Montréal*, 2005, pp. 1-18.
15. Bazilevich R.P. Dekompozitsionnye i topologicheskie metody avtomatizirovannogo konstruirovaniya elektronnykh ustroystv: monografiya [Decomposition and topological methods of automated design of electronic devices: monograph]. L'vov: Vishchashkola, 1981, 81 p.
16. Bazilevich R.P. Metod optimal'nogo svertyvaniya skhemy – effektivnyy podkhod dlya kachestvennogo resheniya nepolinomial'nykh kombinatornykh zadach bol'shoy i sverkhbol'shoy razmernosti v avtomatizirovannom konstruirovanii REA [The method of optimal circuit folding is an effective approach for the qualitative solution of non-polynomial combinatorial problems of large and super-large dimensions in the automated design of REA], *Sb. nauchnykh trudov «Problemy razrabotki perspektivnykh mikroelektronnykh sistem»* [Collection of scientific papers "Problems of development of perspective microelectronic systems"]. Moscow: IPPM RAN, 2005, pp. 94-100.
17. Chan F.T.S., Kumar N., Wong T.C. Three-Dimensional Air-Cargo Loading Problem: An Evolutionary Algorithm Based Approach, *Proceedings of the 7th Asia Pacific Industrial Engineering and Management Systems Conference*, 2006, pp. 758-765.
18. Kureychik V.V., Glushchenko A.E. Kombinirovanny podkhod dlya resheniya zadachi 3-kh mernoy upakovki raznogabaritnykh elementov [Combined approach for solving the problem of 3-dimensional packaging of different-sized elements], *XII Vserossiyskaya nauchnaya konferentsiya molodykh uchenykh, aspirantov i studentov, informatsionnye tekhnologii, sistemnyy analiz i upravlenie (ITSA i U-2015)* [XII all-Russian scientific conference of young scientists, postgraduates and students, information technologies, system analysis and management (ITSA and U-2015)]. Rostov-on-Don: Izd-vo YuFU, 2015, pp. 75-79.

19. Gladkov L.A., Kureychik V.V., Kureychik V.M. Geneticheskie algoritmy [Genetic algorithm]. Moscow: Fizmatlit, 2010.
20. Kureychik V.V., Glushchenko A.E., Kureychik L.V. Programmyy kompleks kombinirovannogo poiska dlya resheniya zadachi trekhmernoy upakovki [Combined search software package for solving the problem of three-dimensional packaging] *Tr. II Vserossiyskoy nauchno-tekhnicheskoy konferentsii Fundamental'nye i prikladnye aspekty komp'yuternykh tekhnologiy i informatsionnoy bezopasnosti* [Proceedings of the II all-Russian scientific and technical conference Fundamental and applied aspects of computer technologies and information security]. Taganrog: Izd-vo YuFU, 2016, pp. 216-220.
21. Zhukov L.A., Korchevskaya O.V. Metod ploskostey: chislenny eksperiment dlya zadach dvukh i trekhmernoy ortogonal'noy upakovki [Method of planes: numerical experiment for two-and three-dimensional orthogonal packing problems], *Informatsionnye tekhnologii* [Information technologies], 2008, No. 11, pp. 41-45.

Статью рекомендовала к опубликованию д.т.н., профессор Л.С. Лисицына.

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 САПР; д.т.н.; профессор.

Глушенко Александр Евгеньевич – Таганрогская таможня; e-mail: alex-14-93@mail.ru; 347939, г. Таганрог, ул. Чучева, д. 46-2, кв. 29; тел.: 88634371651; Уполномоченный по особо важным делам отдела административных расследований.

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

Glushchenko Alexander Evgenyevich – Taganrog customs; e-mail: alex-14-93@mail.ru; 46-2, Chucheva st., apt. 29th, Taganrog, 347939, Russia; phone: +78634371651; Commissioner for Special Cases, Division of Administrative Investigations.

УДК 004.89

DOI 10.18522/2311-3103-2020-2-16-28

Д.Ю. Кравченко, Ю.А. Кравченко, В.В. Марков

ГИБРИДНЫЙ БИОИНСПИРИРОВАННЫЙ АЛГОРИТМ ОТОБРАЖЕНИЯ ОНТОЛОГИЙ В ЗАДАЧАХ ИЗВЛЕЧЕНИЯ И УПРАВЛЕНИЯ ЗНАНИЯМИ*

Статья посвящена решению задачи отображения онтологических моделей в процессах извлечения и управления знаниями. Актуальность и значимость данной задачи обусловлены необходимостью сохранения достоверности и исключения избыточности знаний при интеграции (объединении) структурированных информационных источников различного происхождения. Близость и непротиворечивость понятийной семантики объединенного ресурса при проводимом отображении является основным критерием эффективности предложенных решений. В статье рассмотрены проблемы выбора соответствующих задаче подходов решения, сохраняющих семантику при отображении концептов. Обоснована стратегия выбора биоинспирированного моделирования. Проанализированы аспекты эффективности применения различных децентрализованных биоинспирированных методов. Определены причины необходимости проведения гибридизации. Предложено решать задачу отображения онтологических моделей с применением биоинспирированного алгоритма, построенного на основе гибридизации оптимизационных механизмов алгоритмов бактериального и кукушкиного поиска. Проведенная гибридизация данных алгоритмов позволила объединить их основные преимущества: последовательный бактериальный поиск, обеспе-

* Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов №№ 19-07-00099 и 18-07-00050.

чивающий детальное исследование локальных областей, и значительное число глобальных перемещений агента-кукушки при реализации полетов Леви. Для оценки эффективности предложенного гибридного биоинспирированного алгоритма разработан программный продукт и проведены эксперименты по отображению онтологий разного размера. Каждый концепт любой онтологии имеет определенный набор атрибутов, являющийся семантическим вектором признаков. Степень сходства семантических векторов сравниваемых концептов отображаемых онтологий является критерием их интеграции. Для повышения качества процесса отображения введена новая кодировка решений. Полученные количественные оценки демонстрируют экономию времени при решении задач относительно большой размерности (от 500000 вершин онтографа) не менее 13 %. Временная сложность разработанного гибридного алгоритма составляет $O(n^2)$. Описанные исследования имеют высокий уровень теоретической и практической значимости и напрямую связаны с решением классических задач искусственного интеллекта, направленных на поиск скрытых зависимостей и закономерностей на множестве элементов знаний.

Онтология; гибридные алгоритмы; роевые методы; семантический вектор; биоинспирированный поиск; оптимизация; интеллектуальные агенты.

D.Yu. Kravchenko, Yu.A. Kravchenko, V.V. Markov

HYBRID BIOINSPIRED ALGORITHM FOR ONTOLOGIES MAPPING IN THE TASKS OF EXTRACTION AND KNOWLEDGE MANAGEMENT

The article is devoted to solving the problem of mapping ontological models in the processes of extracting and knowledge management. The relevance and significance of this task are due to the need to maintain reliability and eliminate redundancy of knowledge during the integration (unification) of various origins structured information sources. The proximity and consistency of the conceptual semantics of the combined resource during the mapping is the main criterion for the effectiveness of the proposed solutions. The article considers the problems of choosing appropriate solution approaches that preserve semantics when displaying concepts. The strategy of choosing bio-inspired modeling is substantiated. The aspects of the effectiveness of various decentralized bio-inspired methods are analyzed. The reasons for the need for hybridization are identified. The paper proposes to solve the problem of mapping ontological models using a bio-inspired algorithm based on hybridization of bacterial and cuckoo search algorithms optimization mechanisms. The hybridization of these algorithms allowed us to combine their main advantages: a consistent bacterial search that provides a detailed study of local areas, and a significant number of the cuckoo agent during the implementation global movements of Levy flights. To evaluate the effectiveness of the proposed hybrid bio-inspired algorithm, a software product was developed and experiments were performed on the mapping of different sizes ontologies. Each concept of any ontology has a certain set of attributes, which is a semantic vector of attributes. The degree of the semantic vectors similarity for the compared concepts of displayed ontologies is a criterion for their integration. To improve the quality of the display process, a new encoding of solutions has been introduced. The quantitative estimates obtained demonstrate time savings in solving problems of relatively large dimension (from 500,000 ontograph vertices) of at least 13 %. The time complexity of the developed hybrid algorithm is $O(n^2)$. The described studies have a high level of theoretical and practical significance and are directly related to the solution of classical problems of artificial intelligence aimed at finding hidden dependencies and patterns on a multitude of knowledge elements.

Ontology; hybrid algorithms; swarm methods; semantic vector; bio-inspired search; optimization; intelligent agents.

Введение. Проблема эффективного отображения онтологических структур имеет прямое отношение к решению следующих классических задач искусственного интеллекта: классификации, кластеризации, структурирования, интеграции и семантического поиска информации и знаний. Под отображением будем понимать процесс установления соответствий между понятиями (концептами) нескольких онтологических структур. Данная проблема является оптимизационной. Основным

критерием отображения концептов онтологий установим семантическую близость подразумеваемой понятийной семантики [1–5]. При соблюдении требований использования хорошо структурированных онтологий, точно и однозначно интерпретирующих семантику контекста ресурса, требуется создание формализованных интеллектуальных методов и алгоритмов отображения, позволяющих при сохранении высокого уровня обоснованности имеющихся системно значимых отношений обнаружить обоснованные скрытые зависимости и закономерности на множестве элементов информации и знаний. Подобные интеллектуальные методы и алгоритмы обладают способностью нахождения определенной эквивалентности в понятийной семантике, независимо от сходств и различий в имеющейся структуре отображаемых онтологий.

По мнению авторов, в наибольшей мере такими качествами обладают роевые популяционные биоинспирированные алгоритмы с децентрализованной схемой управления, что создает предпосылки для беспрецедентного параллелизма в проводимых вычислительных процедурах. В качестве основной сложности при решении проблемы отображения онтологий следует отметить наличие в большинстве рассматриваемых случаев области поиска большого объема, что требует одновременно с интенсификацией поиска в локальных фрагментах информационного пространства, также обеспечивать достаточное количество глобальных перемещений интеллектуальных агентов для обеспечения выхода из локальных оптимумов. Большинство известных биоинспирированных алгоритмов обладают только одной хорошо развитой особенностью из двух перечисленных [6–10]. Устранением данного недостатка является решение задачи гибридизации механизмов поиска различных роевых алгоритмов. В представленной работе для гибридизации выбраны алгоритмы бактериальной оптимизации (Bacterial Foraging Optimization, BFO) [11], предложенного Пассино (Passino) в 2002 году и кукушкиного поиска (Cuckoo Search, CS) [12], разработанного Янгом (Yang) и Дебом (Deb) в 2009 году. Для повышения эффективности получаемых решений оба алгоритма были модифицированы в соответствии с требованиями решаемой проблемы. Модифицированный бактериальный алгоритм предлагает реализацию наиболее последовательного поиска с детальным исследованием локальных областей, а модифицированный алгоритм, инспирированный поведением кукушек в процессе гнездового паразитизма, имеет достаточную вероятность глобальных перемещений.

1. Методы отображения онтологий. Рассмотрим основные методы отображения онтологических структур. Выбор среди данных методов производится не из соображений эффективного поиска совпадений в именах или структурных схем понятий, основными критериями является семантическая близость как мера сходства и непротиворечивость с точки зрения контекста.

Одним из таких методов является *отношение уточнения спецификаций сущностей онтологий*. Метод использует формальный критерий корректности отображения спецификаций онтологических концептов, соответствующих абстрактным типам данных. Установленное на основе данного критерия отношение уточнения спецификаций указывает на то, что уточняющая спецификация без ограничений применяется вместо уточняемой без потери качества имеющихся связей. Доказательство уточнения спецификаций поддерживают автоматический и интерактивный режимы в зависимости от сложности информационной модели.

Отметим, что одним из наиболее эффективных вариантов реализации процедуры уточнения спецификаций служит отношение поглощения, которое устанавливается на экстенционалах концептов [13, 14]. Это означает, что все экземпляры класса поглощаемого концепта являются одновременно экземплярами поглощающего. Учитывая тенденции развития онтологических моделей в настоящее время, в

большинстве выстраиваемых рассуждений достаточно применения процедуры поглощения. Отношение уточнения формируется между концептами как при построении одной онтологии, так и при интеграции (отображении) разных. Данная процедура обеспечивает корректность отображения концептов одной онтологии в другую, так как спецификации онтологических концептов отражают и сохраняют понятийную семантику и однозначность контекста.

Помимо формальных методов отображения онтологий значительной популярностью обладают методы, определяющие эквивалентность и семантические различия концептов на основании метаинформации. Реализацией такого метода является применение *метаонтологии*, связанной с обеими интегрируемыми онтологиями [1–5, 13, 14]. В случае, когда отображаемые онтологии построены на основе разных онтологических структур, применение обобщенной метамодели является наиболее эффективным. Метаонтология более абстрактна и содержит в себе большинство разновидностей понятий, составляющих исследуемую предметную область. Метаонтология используется в качестве подложки для отображаемых онтологий, что позволяет унифицировать результирующую структуру. Принцип создания данной подложки следующий. Каждый концепт отображаемых онтологий вместе с предикатами и отношениями становится экземпляром некоторого концепта метаонтологии. При отсутствии семантически близких концептов создаются так называемые служебные концепты как подпонятия метаонтологии, задающие необходимую семантику в терминах концептов метаонтологии. В итоге элементы спецификаций отображаемых онтологий будут распределены по классам, определяемым концептами метаонтологии с учетом семантики и контекста ресурса.

Предлагаемый принцип создания подложки на метаонтологии позволяет распределить спецификации по разным уровням иерархии классов, что делает их независимыми друг от друга. Сформированные в терминах метаонтологии спецификации используются:

- ♦ для проверки корректности отображения концептов;
- ♦ при семантическом поиске релевантных концептов.

При использовании данного метода поглощающий и поглощаемый концепты должны быть в одном классе, соответствующем концепту метаонтологии или служебному концепту. Также возможно в качестве подложки использовать общую онтологию верхнего уровня, что является более сложной задачей, так как требуется интеграция каждой онтологии в общую, что приводит к изменению начальных версий онтологических моделей при совместимости их концептов с онтологией верхнего уровня.

Оценка фундаментальных метасвойств лежит в основе следующего метода отображения онтологий. Любая онтология включает в себя множество свойств и отношений между концептами. Набор фундаментальных метасвойств позволяет оценить концепты или отношения по следующим критериям:

- 1) *существенность* – неотъемлемость свойства сущности;
- 2) *строгость* – принадлежность существенного свойства сущности в любом контексте или предметной области;
- 3) *идентификация* – является ли утверждение об объекте идентифицирующим его свойством;
- 4) *собственная идентификация* – несет ли эту идентификацию само утверждение, либо оно наследует ее из других свойств или сущностей;
- 5) *неизменность* – может ли свойство меняться во времени;
- 6) *зависимость* – может ли сущность существовать без других;
- 7) *постоянство* – как долго сущность остается таковой;
- 8) *объединение* – существование экземпляров контекста как целых сущностей в отношении часть-целое.

Обратим внимание на то, что формальное определение перечисленных метасвойств требует введения некоторых ограничений. Одни метасвойства совместимы, другие обладают признаком ортогональности и исключают друг друга. Для свойства W , поглощающего свойство V , верны следующие ограничения:

- ◆ если W строгое для любых сущностей, то V также строгое для любых сущностей;
- ◆ если W несет критерий идентификации сущностей, то и V также;
- ◆ если W несет критерий объединения, то и V также;
- ◆ если W не несет объединение, то и V также;
- ◆ всякая сущность должна быть значением наиболее общего свойства, несущего его идентификацию.

Поясним, что при построении последовательности поглощений между концептами существуют случаи наследования метасвойств более специфичных концептов, поэтому наличие конфликтных ситуаций в данной иерархии приведет к некорректному отображению. Описанные ограничения также выполняются для отношений, обнаруженных между концептами различных онтологий, что позволяет выявлять семантические конфликты, полученные при отображении онтологий. На присутствие конфликта указывает противоречие метасвойств связанных концептов, в этом случае отображение принято считать некорректным. Исследование метасвойств дает возможность строить более сложные рассуждения, которые необходимы для повышения эффективности решения задачи отображения онтологий.

В иерархии свойств концептов учитывают систему классов, в которой родовые концепты идентифицируют объекты реального мира, создавая при этом типы, видовые – описывают категории объектов, ролевые – распределяют роли. Причем, ролевые концепты входят в подкласс родовых, а видовые могут быть подклассами как ролевых, так и родовых концептов [1–5, 13, 14]. В метасвойствах отношений любая часть является неотъемлемой, что позволяет однозначно идентифицировать целое по части, а в случае инвариантности целого, части идентифицируется по нему, все эти особенности применяются для корректного отображения онтологий.

Описанный метод имеет связь с предыдущим, использующим общую онтологию верхнего уровня для проведения процедуры отображения. Фундаментальные концепты общей онтологии имеют вполне определенные наборы значений метасвойств. Совместное применение указанных методов дает дополнительные возможности для повышения эффективности решения поставленной задачи.

В заключении проведенного в данном пункте обзора рассмотрим метод отображения онтологий на основе экземпляров экстенционалов понятий. В данном методе релевантность определяется на основе принадлежности экземпляров одним и тем же концептам. В качестве экземпляров рассматриваются: объекты-сущности реального мира; модели реального мира; классифицированные в онтологии данные. Проверка корректности обнаруженных отношений проводится с применением экстенциональной составляющей связанных концептов онтологий. Если существует хотя бы один пример экземпляра, в котором сущности не принадлежат одновременно классам эквивалентным концептам отображаемых онтологий, тогда возникает конфликт, ставящий под сомнение корректность обнаруженной связи между концептами. Описанный экстенциональный метод проверки отображения онтологий «по образцу» реализуется формальным образом.

Реализация всех вышеперечисленных методов возможна только при наличии эффективных автоматических средств оценки семантической близости концептов, как меры сходства, позволяющей утверждать об эквивалентности рассматриваемых понятий и отношений. Создание таких средств на основе алгоритмов биоинспирированного поиска с децентрализованной схемой управления является основной целью данной работы.

2. Постановка задачи. Используем следующую модель онтологической структуры $O = \langle P, V, R, C \rangle$, где P – множество концептов; V – множество экземпляров концептов; R – множество предикатов; C – множество значимых отношений. Для проведения процедуры отображения построим две онтологии O_1 и O_2 . $O_1 = \langle P^1, V^1, R^1, C^1 \rangle$ и $O_2 = \langle P^2, V^2, R^2, C^2 \rangle$. Примем равным n количество концептов в онтологии O_1 , и равным m – в онтологии O_2 . Для оценки релевантности концептов применим эквивалентный тип семантической близости [1–5], в которой $mapping(P^1) = P^2$, if $sim(P^1, P^2) \geq d$, где d – пороговое значение сходства предикатов, достижения которого достаточно для отображения концепта P^1 в онтологию O_2 . Совпадение не ниже порогового значения d предикатов понятия P_i^1 онтологии O_1 с предикатами понятия P_j^2 онтологии O_2 указывает на эквивалентность множеств R^1, R^2 .

Итоговая онтологическая структура, полученная после отображения онтологии O_1 в онтологию O_2 , примет следующий вид:

$$Z = \langle O, H^O, S, U \rangle, \quad (1)$$

где $O = \langle P, V, R, C \rangle$ – обобщенная онтология; H^O – интегрированная интеллектуальная информационная система с онтологией O ; S – множество значений эквивалентной семантической близости; U – изменение информационной системы в виде отображения.

Целевой функцией поставленной задачи является следующее выражение:

$$f(R^1, R^2) \rightarrow \max, \quad (2)$$

где R^1, R^2 – множества предикатов концептов онтологий O_1 и O_2 . Каждое множество предикатов состоит из подмножеств конкретных атрибутов, принадлежащих каждому из понятий двух онтологий.

$$R^1 = \{R_1^1; R_2^1; \dots; R_i^1; \dots; R_n^1\}, \quad (3)$$

где $i = \overline{1, n}$, n – количество концептов в O_1 .

$$R^2 = \{R_1^2; R_2^2; \dots; R_j^2; \dots; R_m^2\}, \quad (4)$$

где $j = \overline{1, m}$, m – количество концептов в O_2 .

Решением задачи оценки эквивалентной семантической близости является следующее выражение:

$$f(R^1, R^2) = \sum_{i=1}^n \sum_{j=1}^m R_i^1 \cap R_j^2 \geq d, \quad (5)$$

где $d > 0$, d – пороговое значение, равенство или превышение которого означает присутствие эквивалентной семантической близости между подмножествами предикатов, принадлежащих концептам P_i^1 и P_j^2 . Чем выше показатель d , тем более значимой является эквивалентная семантическая близость между рассматриваемыми концептами.

Таким образом, при решении задачи отображения большего числа онтологий требуется найти такую пару онтологий O_a и O_b , для которых функция f примет наибольшее значение.

2. Разработка гибридного биоинспирированного алгоритма отображения онтологий. Оценка эквивалентной семантической близости является трансвычислительной задачей, так как в общем случае потребуются полный перебор для проверки всех оценок попарных сравнений предикатов концептов отображаемых онтологий. Вычислительная сложность задачи требует применения методов искусственного интеллекта.

В работе задача отображения онтологий решена с применением нового биоинспирированного алгоритма, созданного на основе гибридизации оптимизационных механизмов алгоритмов бактериального и кукушкиного поиска. Первый алго-

ритм позволяет интенсифицировать поиск в локальных областях, а второй дает возможность реализовывать глобальные перемещения интеллектуальных агентов для выхода из локальных оптимумов.

Отметим, что поставленная задача решается одновременно в двух информационных пространствах, первой и второй онтологий, что использовано авторами для дифференцирования процедур локальных и глобальных перемещений. Для этого каждая онтология получает определенный статус. Для верификации поиска предусмотрена возможность смены статусов онтологий. При первом запуске процедуры отображения условно будем считать онтологию с номером 1 – информационным пространством для реализации глобальных перемещений агентов на основе реализации полетов кукушки по закону Леви, а онтологию под номером 2 – информационным пространством для осуществления последовательного поиска на основе передвижений агентов-бактерий. Правило смены статуса онтологий описан ниже.

Популяционное децентрализованное поведение бактерий стало основой канонического бактериального алгоритма поисковой оптимизации (Bacterial Foraging Optimization, BFO), предложенного Пассино (Passino) в 2002 году [11, 15–17]. В представленной работе модификация данного метода коснулась структуры используемой области поиска, разделенной на две части с изменяемыми статусами в соответствии с онтографами двух отображаемых онтологий, а также привела к реализации нового гибридного асимметричного механизма перемещений популяции агентов, основанного на чередовании локальных последовательных передвижений агентов-бактерий с глобальными перемещениями агентов-кукушек. Алгоритм поиска кукушки (Cuckoo Search, CS) был разработан Янгом (Yang) и Дебом (Deb) в 2009 году [12, 15–17]. Основное достоинство алгоритма – незначительное количество свободных параметров, что делает метод более универсальным и эффективным по сравнению с другими роевыми методами оптимизации. Отличием использованного в работе модифицированного метода поиска кукушки от канонического является применение одного онтографа в качестве статичного пространства популяции подкладываемых яиц кукушки, а второго – динамичного пространства перемещений популяции агентов кукушек по чужим гнездам в поиске сходства подкладываемого яйца с оригиналом. Данная особенность также реализована с помощью периодических смен статуса онтологий во время поиска.

Связи между вершинами онтографов зададим множествами отношений C^1 и C^2 онтологий O_1 и O_2 соответственно. Вершины онтографов содержат значения предикатов концептов, представленных в виде подмножеств $R_1^1; R_2^1; \dots; R_i^1; \dots; R_N^1$ для онтологии O_1 , и подмножества $R_1^2; R_2^2; \dots; R_j^2; \dots; R_M^2$ для онтологии O_2 . Превышение порогового значения d показателем семантической близости между предикатами концептов P_i^1 и P_j^2 указывает на их эквивалентность [18–20]. Пара сравниваемых концептов с номерами a и b , где $a \in [1:N]$, $b \in [1:M]$ для каждого интеллектуального агента выбирается случайным образом на каждой итерации работы алгоритма.

На начальном этапе работы алгоритма подмножество предикатов концептов 1-ой онтологии будет статичным, а сама 1-ая онтология будет иметь статус константы, напротив, подмножество предикатов концептов 2-ой онтологии станет динамичной составляющей поиска, а сама 2-ая онтология станет координатным пространством для передвижений агентов (рис. 1). Динамика поиска во 2-ой онтологии реализована на основе бактериальной оптимизации, которая продолжается пока удастся получить прирост фитнес-функции. В случае отсутствия прогресса в бактериальном поиске происходит смена статусов онтологий. Текущие вершины-координаты бактерий во 2-ой онтологии назначаются константами, а поиск экви-

валентных им концептов возобновляется в пространстве 1-ой онтологии на основе перелетов агентов-кукушек. Перелеты продолжаются до увеличения значения фитнес-функции, либо до достижения их максимального количества, после чего поиск в 1-ой онтологии продолжается на основе механизма бактериальной оптимизации, который применяется до следующей смены статуса онтологий по уже описанному выше сценарию.

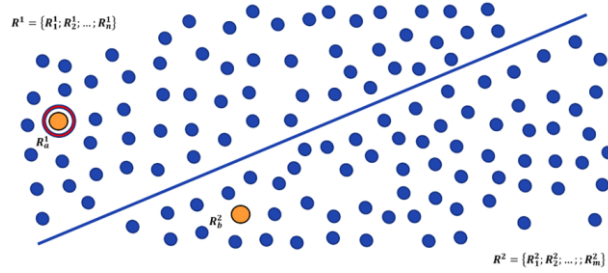


Рис. 1. Модель разделенного на подмножества пространства решений

Свободными параметрами алгоритма являются T, T^r, T^l – общие числа шагов хемотаксиса (перелетов), репродукции и ликвидации соответственно, а $|S|$ – четное число агентов в популяции. Координаты при локальном поиске $X_{i,r,l} = X_{i,r,l}(t)$ бактерии $s_i \in S$ на t -м шаге хемотаксиса, r -м шаге репродукции и l -м шаге ликвидации определяются направлением вектора «плавания» $V_i(\epsilon)$, где ϵ – мощность окрестности целевой вершины. Этот параметр отличает данный механизм поиска при бактериальной оптимизации от канонического и позволяет повысить вероятность обнаружения траектории передвижения, дающей наибольший прирост фитнес-функции, благодаря выбору по ходу движения бактерии вершины с наибольшей мощностью окрестности (рис. 2).

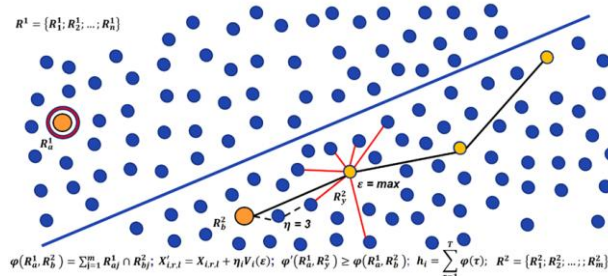


Рис. 2. Передвижение бактерии с учетом мощности окрестности целевой вершины

Выражение фитнес-функции имеет следующий канонический вид:

$$\varphi_{i,r,l} = \varphi_{i,r,l}(T). \quad (6)$$

Новые координаты $X'_{i,r,l}$ агента-бактерии вычисляются по следующему выражению:

$$X'_{i,r,l} = X_{i,r,l} + \eta_i V_i(\epsilon), \quad (7)$$

где $\eta \geq 1$, задает величину шага передвижения бактерии (свободный параметр алгоритма). При увеличении значений фитнес-функции направление вектора «плавания» остается неизменным $V'_i(\epsilon) = V_i(\epsilon)$. При уменьшении значений фитнес-

функции агент-бактерия совершает «кувырок», возвращается в предыдущую вершину и продолжает поиск, кардинальным образом меняя направление вектора своего движения (рис. 3).

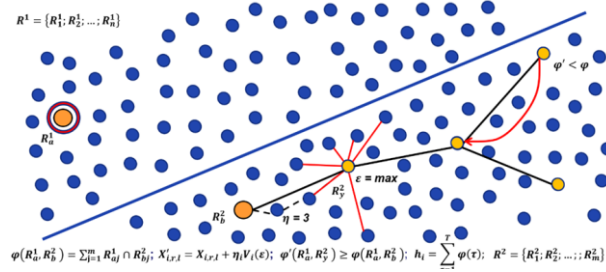


Рис. 3. «Кувырок» бактерии в условиях снижения значений фитнес-функции

Интенсификация процесса локального поиска достигается применением процедуры репродукции, которая позволяет на заданном шаге r воспроизвести копии наиболее удачных агентов бактерий, составляющих половину всей популяции, уничтожив при этом вторую половину безуспешных. Решение о репродукции основано на высоком показателе здоровья бактерии, вычисляемом по следующему выражению:

$$h_i = \sum_{\tau=1}^T \varphi_{i,r,l}(\tau), i \in [1: |S|]. \quad (8)$$

Например, если успешная бактерия $s_j, j \in [1: |S|]$ имеет координаты $X_{j,r,l}$, тогда после процедуры репродукции возникнет бактерия s_k , причем, $k = \frac{|S|}{2} + j$, $X_{j,r+1,l} = X_{j,r,l}$, $X_{k,r+1,l} = X_{j,r,l}$. При этом, общее количество бактерий в колонии останется неизменным.

Процедура репродукции совершается многократно. Максимальное количество репродукций задается свободным параметром алгоритма Z . Достижение данного значения запускает процедуру ликвидации, при которой случайное число агентов $w < |S|$ (свободный параметр алгоритма) уничтожается, а вместо них в случайных координатах возникает такое же количество новых агентов, что позволяет алгоритму выйти из локальных оптимумов.

В случае отсутствия прогресса в поиске какого-либо агента при достижении или превышении им показателя половины шагов до репродукции $r \geq T'/2$, статус онтологий меняется (рис. 4).

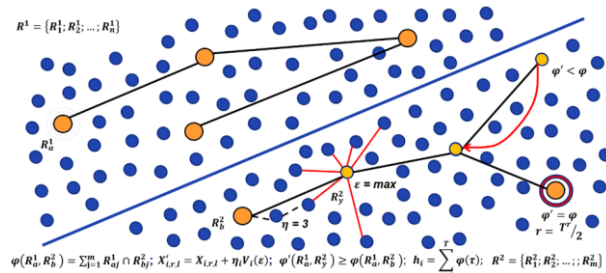


Рис. 4. Смена статусов онтологий, организация «перелетов»

Текущая координата бактерии замораживается и становится константой, процедура поиска переводится в информационное пространство другой онтологии, где в последующем производятся процедуры «перелетов» кукушки (рис. 5), что

обеспечивает глобальные перемещения. При обнаружении вершин дающих прирост фитнес-функции, «перелеты» прекращаются, агент-кукушка заменяется на агента-бактерию, и поиск по заданным выше правилам возобновляется в пространстве онтологии, получившей новый статус динамичной. Данный процесс продолжается до очередной смены статуса.

Обобщенная схема гибридного биоинспирированного алгоритма отображения онтологий, разработанного на основе алгоритмов бактериального и кукушкиного поиска, представлена на рис. 5. Основные модифицированные процедуры созданного гибридного алгоритма выделены на рис. 5 фоном.

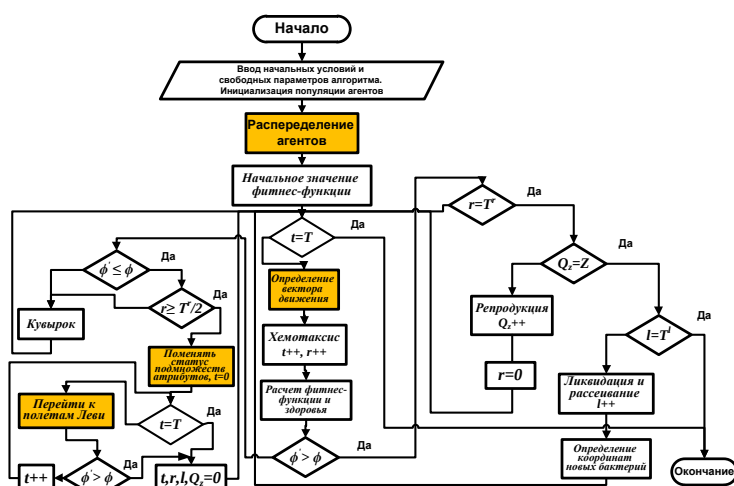


Рис. 5. Обобщенная схема гибридного биоинспирированного алгоритма отображения онтологий

Алгоритм, инспирированный поведением кукушек в процессе гнездового паразитизма, отличается от известных аналогов более эффективным способом определения траектории передвижения агента в пространстве поиска. Кукушка свои передвижения осуществляет в виде полетов Леви (Levy flights), реализуемых по следующей канонической формуле (9) с тензорным произведением вектора размера шагов и случайного вектора независимых вещественных чисел.

$$X' = X + U \otimes L_{|X|}(\lambda), \quad (9)$$

где $U = (u_a, a \in [1: |X|])$ – вектор размера шагов по соответствующим компонентам вектора координат местонахождения агента X ; $L_{|X|}(\lambda) = (|X| \times 1)$ – вектор независимых вещественных случайных чисел, распределенных по закону Леви [12, 21–25]. Данный способ обеспечивает высокую вероятность малых и достаточную вероятность больших перелетов агента-кукушки, что позволяет интенсифицировать поиск в локальных областях, с одной стороны, и реализовать выход из локальных оптимумов, с другой.

3. Экспериментальные исследования. Для подтверждения эффективности разработанного гибридного биоинспирированного алгоритма отображения онтологий был создан программный модуль, позволяющий сравнить результаты решения поставленной задачи, полученные с применением предложенного и канонических роевых алгоритмов. В качестве значений предикатов применены усредненные абстрактные показатели, полученные случайным образом с заданной вероятностью.

В результате проведенных комплексных исследований получена временная сложность (ВСА) предложенного гибридного алгоритма, которая составила $O(n^2)$, где n – количество анализируемых алгоритмом входных данных, которым в данном случае является число вершин онтографов (рис. 6). В представленном биоинспирированном алгоритме, построенном на основе гибридизации оптимизационных механизмов алгоритмов бактериального и кукушкиного поиска, предусмотрено два вложенных цикла прохождения по множествам предикатов концептов: первый, для реализации локального поиска, а второй, для осуществления глобальных перемещений. Таким образом, максимально возможное количество операций зависит от числа вершин онтографов N как $n * n$.

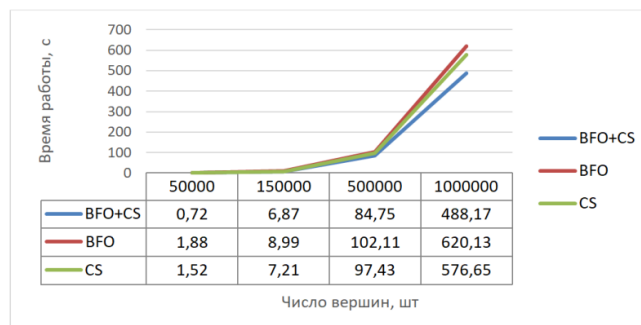


Рис. 6. Результаты сравнения скорости работы канонических алгоритмов бактериальной оптимизации (BFO) и кукушкиного поиска (CS) с предложенным гибридным алгоритмом (BFO + CS)

Экспериментальные исследования разработанного метода проводились на вычислительной системе с производительностью порядка 10^9 (гигафлопс) операций в секунду. В этом случае, например, при $n = 150000$ вершин, время выполнения алгоритма не превысило 7-ми секунд, при $n = 500000$ вершин – 90 секунд, при $n = 1000000$ вершин – 490 секунд, что наглядно иллюстрирует представленная на рисунке 6 полученная зависимость времени выполнения предложенного метода от числа входных вершин. Исследование скорости работы предложенного гибридного алгоритма биоинспирированной оптимизации для решения задачи отображения онтологий проводилось в сравнении с каноническими бактериальным и кукушкиным алгоритмами.

Полученные количественные оценки сравнения скорости работы канонических алгоритмов бактериальной оптимизации (BFO) и кукушкиного поиска (CS) с предложенным гибридным алгоритмом (BFO + CS) показывают, что разработанный гибридный алгоритм дает лучшие результаты по сравнению с каноническими. Отметим, что при относительно небольших размерностях задачи (от 50000 до 150000 вершин) время работы оцениваемых алгоритмов соизмеримо: от 0,72–6,87 секунды в предлагаемом гибридном алгоритме до 1,88–8,99 и 1,52–7,21 секунды соответственно в бактериальном и кукушкином алгоритмах по отдельности. При количестве вершин от 500000 до 1000000 разница в показателях времени работы становится более значительной. В целом, на указанных относительно больших размерностях задачи, алгоритм кукушкиного поиска быстрее бактериального алгоритма на 6–7 %, а разработанный гибридный алгоритм быстрее кукушкиного на 13–15 %. Эффективность предложенного гибридного алгоритма повышена за счет улучшения способа определения траектории передвижения агента в пространстве поиска, позволяющего обеспечить максимальное количество итераций локального поиска при достаточном количестве – глобального.

Заключение. В данной работе представлена разработка гибридного биоинспирированного алгоритма отображения онтологий. Предложено решать задачу отображения онтологических моделей с применением биоинспирированного алгоритма, построенного на основе гибридизации оптимизационных механизмов алгоритмов бактериального и кукушкиного поиска. Проведенная гибридизация данных алгоритмов позволила объединить их основные преимущества: последовательный бактериальный поиск, обеспечивающий детальное исследование локальных областей, и значительное число глобальных перемещений агента-кукушки, позволяющих алгоритму выходить из «локальных ям».

Для оценки эффективности предложенного метода разработан программный продукт, включающий в себя функции построения онтологий, генерации набора предикатов концептов и проведения оценки эквивалентной семантической близости. Проведенные экспериментальные исследования подтвердили эффективность разработки в сравнении с каноническими биоинспирированными алгоритмами бактериальной оптимизации и кукушкиного поиска. Полученные количественные оценки подтверждают снижение времени решения задачи отображения онтологий при помощи разработанного гибридного алгоритма (BFO + CS) не менее чем на 13 % по отношению в результатам работы канонических алгоритмов бактериальной оптимизации (BFO) и кукушкиного поиска (CS). При относительно небольших размерностях задачи (от 50000 до 150000 вершин в отображаемых онтологиях) время работы всех алгоритмов соизмеримо, наибольший прирост скорости решения достигается в задачах с размерностью не менее 500000 вершин. Временная сложность разработанного гибридного алгоритма составляет $O(n^2)$, где n – количество анализируемых методом входных данных, которым в данном случае является число вершин онтографов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Andreassen T., Knappe R., Bulskov H. Domain specific similarity and retrieval, *11th Int. Fuzzy Systems Association World Congress*, 2016, Vol. 1, pp. 496-502.
2. Castano S., Ferrara A., Montanelli S., Racca G. Semantic information interoperability in open networked systems, *Proc. of the Int. Conf. SNW*. Paris, 2004, pp. 215-230.
3. Kravchenko Yu.A., Kursitys I.O., Markov V.V. Bioinspired Algorithm for Acquiring New Knowledge based on Information Resource Classification, *2019 International Russian Automation Conference (RusAutoCon)*.
4. Haase P., Siebes R., Harmelen F. Peer selection in peerto-peer networks with semantic topologies, *Proc. of Int. Conf. on Semantics in a Networked World*. Paris, 2004, pp. 108125.
5. Maedche A., Zacharias V. Clustering ontology-based metadata in the Semantic Web, *Proc. 6th European PKDD Conf. LNCS*. Berlin: Springer, 2002, Vol. 2431, pp. 348-360.
6. Abraham A., Grosan G., Ramos V. *Swarm Intelligence in Data Mining*. Berlin. Heidelberg: SpringerVerlag, 2006, 267 p.
7. Sousa T., Silva A., Neves A. Particle Swarm based Data Mining Algorithms for classification tasks, *Parallel Computing*, 2004, Vol. 30, Issue 5–6, pp. 767-783.
8. Parsopoulos K.E., Vrahatis M.N. Recent Approaches to Global Optimization Problems Through Particle Swarm Optimization, *Natural Computing*, 2002, No. 1 (2–3), pp. 235-306.
9. Kravchenko Yu.A., Kravchenko D.Y., Kursitys I.O. Architecture and method of integrating information and knowledge on the basis of the ontological structure, *Advances in Intelligent Systems and Computing. 1st International Conference of Artificial Intelligence, Medical Engineering, and Education, AIMEE 2017*. Moscow, 2018, Vol. 658, pp. 93-103.
10. Kravchenko Yu.A., Kureichik V.V. Knowledge management based on multi-agent simulation in informational systems, *Conference proceedings. 8th IEEE International Conference "Application of Information and Communication Technologies – AICT 2014"*. – 15-17 October 2014, Astana, Kazakhstan, pp. 264-267.
11. Arijit S.D., Sambarta B., Abraham D.A. Bacterial Foraging Optimization Algorithm: Theoretical Foundations, Analysis, and Applications, *Foundations of Computational Intelligence*. Springer, 2009, Vol. 3, pp. 23-55.

12. Payne R.B., Sorenson M.D., and Klitz K. The Cuckoos. Oxford University Press, 2005.
13. Kravchenko Yu.A., Kuliev E.V., Kulieva N.V., Kureichik V.V. Problem-oriented knowledge processing on the basis of hybrid approach, *Information technologies in science. Management, Social Sphere and Medicine (ITSMSM 2016)*, pp. 510-513.
14. Martens D., De Backer M., Haesen R., Vanthienen J., Snoeck M. and Baesens B. Classification With Ant Colony Optimization, *IEEE Transactions on Evolutionary Computation*, 2007, Vol. 11, No. 5, pp. 651-665.
15. Falco I.D., Cioppa A.D., and Tarantino E. Evaluation of particle swarm optimization effectiveness in classification, *LNAI3849*, 2006, pp. 164-171.
16. Soliman O.S. and Adly A. Bio-inspired algorithm for classification association rules, *8th International Conference on Informatics and Systems (INFOS)*, Cairo, 2012, pp. 154-160.
17. Bova V., Zaporozhets D., and Kureichik V. Integration and processing of problem-oriented knowledge based on evolutionary procedures, *Advances in Intelligent Systems and Computing*, 2016, Vol. 450, pp. 239-249.
18. Semenova A.V. and Kureichik V.M. Ensemble of classifiers for ontology enrichment, *Journal of Physics: Conference Series*, 2018, Vol. 1015, Issue 3, article id. 032123.
19. Kureychik V.M. Overview and problem state of ontology models development, *9th International Conference on Application of Information and Communication Technologies, AICT 2015 - Proceedings 9*, 2015, pp. 558-564.
20. Semenova A.V. and Kureychik V.M. Application of swarm intelligence for domain ontology alignment, *Proceedings of the First International Scientific Conference "Intelligent Information Technologies for Industry" (ITI'16)*, 2016, Vol. 1, pp. 261-270.
21. Bova V., Kureichik V. and Zaruba D. Heuristic approach to model of corporate knowledge construction in information and analytical systems, *2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT)*, Baku, 2016, pp. 1-5.
22. Kureichik V., Zaporozhets D., and Zaruba D. Generation of bioinspired search procedures for optimization problems, *Application of Information and Communication Technologies, AICT 2016 - Conference Proceedings*, 2016, Vol. 10.
23. Kar A.K. Bio inspired computing – A review of algorithms and scope of applications, *Expert Systems with Applications*, 2016, Vol. 59, pp. 20-32.
24. Zaporozhets D., Zaruba D., and Kulieva N. Parallel approach for bioinspired algorithms, *Journal of Physics: Conference Series Ser. "International Conference Information Technologies in Business and Industry 2018 - Enterprise Information Systems"*, 2018.
25. 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.

Статью рекомендовал к опубликованию к.т.н. С.Г. Буланов.

Кравченко Даниил Юрьевич – Южный федеральный университет; e-mail: kravchenkodanil122@gmail.com; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371651; кафедра систем автоматизированного проектирования; студент.

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

Марков Владимир Васильевич – e-mail: vvmarkov@sfedu.ru; кафедра систем автоматизированного проектирования; доцент.

Kravchenko Daniil Yurievich – Southern Federal University; e-mail: kravchenkodanil122@gmail.com; 44, Nekrasovskiy lane, Taganrog, 347928, Russia; phone: +78634371651; the department of computer aided design; student.

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

Markov Vladimir Vasilievich – e-mail: vvmarkov@sfedu.ru; the department of computer aided design; associate professor.

В.С. Корнилов, В.М. Глушань, А.Ю. Лозовой**ОПРЕДЕЛЕНИЕ НАБОРА УСЛОВИЙ ДЛЯ АВТОМАТИЧЕСКОГО
ПОИСКА ОПТИМАЛЬНОГО ВАРИАНТА ГИБРИДНОГО МАШИННОГО
ПЕРЕВОДА ТЕКСТА НА ГРАФЕМНОМ УРОВНЕ***

Статья посвящена алгоритмическому поиску оптимальных решений при оценке и повышении качества гибридного машинного перевода текста. Объектом исследования являются тексты на любых алфавитных языках с различной базой (алфавитом), а также их переводы на другие алфавитные языки. В настоящее время существующие методы и средства гибридного машинного перевода отличаются большим многообразием алгоритмов оценки качества, однако, недостатком данных методов является отсутствие четких критериев, ограничений и схемы оценивания, в итоге результат перевода в большинстве случаев не соответствует уровню публикации. Целью работы является определение набора условий для автоматического поиска оптимального варианта гибридного машинного перевода текста на графемном уровне. Основными решаемыми задачами в ходе исследования являются поиск качественных и количественных условий, в том числе максимальных, минимальных и средних значений длин переводов, обратных переводов и редакционных расстояний между парами текстов, имеющими одинаковый смысл. Научная новизна заключается в использовании графического представления модели алфавитных языков на графемном уровне в виде декартовой системы координат с размерностью, равной единичному редакционному расстоянию (Левенштейна). При решении использованы следствия теоремы де Гуа, действующие Правила Стандартизации ПР 50.1.027–2014 «Правила оказания переводческих и особых видов лингвистических услуг», метод деканонизации и модель «оригинальный текст – перевод – обратный перевод». В результате получены действительные и практически применимые решения для рассматриваемых задач. В связи с этим данная работа может быть интересна для широкого круга специалистов, занимающихся проблемами машинного перевода и переводоведением.

Гибридный машинный перевод; качество перевода; обратный перевод; редакционные расстояния; графемный уровень; алфавитные языки.

V.S. Kornilov, V.M. Glushan, A.Yu. Lozovoy**DETERMINING A SET OF CONDITIONS FOR AUTOMATICALLY FINDING
THE BEST OPTION FOR HYBRID MACHINE TRANSLATION OF TEXT
AT THE LEVEL OF GRAPHEMES**

The article is devoted to the Algorithmic Search for Optimal Solutions for evaluating and improving the Quality of Hybrid Machine Translation of Text. The Object of the Research is Texts on any Alphabetical Languages with different Bases (Alphabets), as well as their Translations into other Alphabetical Languages. Currently, existing Methods and Means of Hybrid Machine Translation are characterized by a wide variety of Quality Assessment Algorithms, but the Disadvantage of these Methods is that most of them do not have Clear Criteria, Limitations and Schemes of Assessments, eventually, the Result of the Translation in most cases does not correspond to the Level of Publication. The Aim of the Work is to determine a Set of Conditions for automatic search for the Best Option of Hybrid Machine Translation of Text at the Level of Graphemes. The Main Tasks to be solved during the Research are the Search for Qualitative and Quantitative Conditions, including the maximum, minimum and average values of the Lengths of Translations, Reverse Translations and Editorial Distances between Pairs of Texts that have the Same Meaning. The Scientific Novelty lies in use the Graphical Representation of the Model of Alphabetic Languages at the Level of Graphemes in the Form of a Cartesian Coordinate System with a Dimension equal to a Unit Editorial Distance (by Levenstein). When solving the de Goui's Theorem, the current Rules of

* Работа выполнена при поддержке РФФИ (проект № 18-01-00041).

Standardization PR 50.1.027–2014 "Rules for the Provision of Translation and Special Types of Linguistic Services", the Method of Decanonization and the Model "Original Text – Translation – Reverse Translation" were used. As a Result, Actual and Practically Applicable Solutions for the Problems under consideration are obtained. In this regard; this Work may be interest to a wide range of Specialists engaged in Machine Translation and Translation Studies.

Hybrid machine translation; translation quality; reverse translation; editorial distances; level of graphemes; alphabetical languages.

Введение. Целью современных систем машинного перевода является получение текстов высокого качества. При этом нередко решаются задачи выбора оптимального перевода из нескольких вариантов. Для отдельных фраз и словосочетаний существует единственный вариант перевода, а для предложений и абзацев, как правило, множество вариантов или нечетких дубликатов текста, имеющих смысловое сходство. В настоящее время существует ряд методов оценки машинного перевода и методов сравнения нечетких дубликатов текста [1]. Одним из таких методов является вычисление редакционного расстояния Левенштейна между двумя текстами [2].

Преимуществом данного метода является его универсальность, выражающаяся в возможности вычислять длину и редакционное расстояние для любых текстов на разных алфавитных языках (не иероглифических), а также высокая математическая точность. К недостаткам данного метода относится невозможность отслеживания ошибок в переводе при сравнении текстов на разных языках. В данной статье предлагается один из возможных подходов к повышению универсальности метода редакционных расстояний Левенштейна для определения оптимального варианта машинного перевода текста на графемном уровне и отслеживания ошибок в переводе в автоматическом режиме.

Постановка задачи. Целью статьи является установление набора условий для автоматического поиска оптимального варианта гибридного машинного перевода текста на графемном уровне передаваемой информации. Объектом исследования являются тексты на любых символьных языках с различной базой (алфавитом). Решаемыми задачами являются определение минимальных, максимальных и средних значений длин текстов при переводе, а также допустимые пределы редакционных расстояний между текстами. При переводе текстов с языка на язык возникает необходимость определения соотношения длин оригинального текста и перевода как одного из возможных критериев оценки качества. Длины текстов оригинала и перевода является одним из критериев при оценке эквивалентности текстов. Поскольку длинный текст невозможно перевести короткой фразой, данная задача является актуальной. В рамках данной статьи не рассматривается сравнение оригинального текста и перевода на логико-семантическом и морфосинтаксическом уровнях, которое проводится параллельно в соответствии с Правилами Стандартизации ПР 50.1.027–2014 «Правила оказания переводческих и особых видов лингвистических услуг» (далее – Правила) [4] и методом деканонизации, который более подробно изложен в работе [5].

Обзор основных работ по исследуемой тематике. Рекомендация по применению обратного перевода при оценке самого перевода впервые приведена в статье [6]. Следует отметить, что обратные методы имеют широкое применение в математике [7], измерительной технике и метрологии [8]. При многократном применении прямых и обратных операций с данными методы носят итерационный характер [9, 10].

В работе [8] исследуются методы взаимных обратных преобразований для определения чувствительности обратимых преобразователей в измерительных системах. «Все методы взаимности основываются на том, что один из преобразо-

вателей является взаимным и для него отношение чувствительности в режиме приема к чувствительности в режиме передачи равно постоянной величине, называемой параметром взаимности, и этот параметр зависит от свойств среды, граничных условий, и не зависит от типа преобразователя» [8]. Учитывая наличие среднестатистического коэффициента пересчета длин текстов при переводе с произвольного языка на язык k , можно утверждать, что существует вероятность аналогичного поведения системы обратимого «черного ящика» и для машинного перевода.

Применение расстояния Левенштейна при оценке качества машинного перевода для средств стандартизации терминологии изложено в статье [11]. В данной статье использован подход, заключающийся «в прямом и обратном переводе русского термина на иностранный язык (английский) и последующем обратном переводе на русский. В случае совпадения оригинала и результата двойного перевода термина последний предполагается согласованным с международно-принятой терминологией. Расстояние Левенштейна использовалось для определения совпадения/расхождения между исходным термином и термином, полученным в результате прямого и обратного переводов». Указанный подход является частным случаем задачи определения набора условий для автоматического поиска оптимального варианта гибридного машинного перевода текста на графемном уровне, рассматриваемой в рамках данной статьи.

В работе [12] приведена классификация параметров алгоритмов, применяемых для сравнения строковых данных. Основными параметрами являются «соотношение длин источника (SourceSize – англ.) и шаблона (PatternSize – англ.), точность совпадения строк, применяемые вычислительные ресурсы на выполнение поиска и предобработку данных, возможность изменения шаблона, поведение алгоритма при изменении допустимого числа отличий» [12]. В цитируемой статье также отмечается использование преобразующего автомата «на основе суффиксного дерева предложения (алгоритм Укконена) и редакционных расстояний Дамерау-Левенштейна».

В пункте 6.1.2 Правил [4] устанавливаются пересчетные коэффициенты. «Из-за изменения объема текста в процессе перевода в зависимости от языковой пары при первичном расчете объема работ и услуг по письменному переводу используются условные коэффициенты» [4]. «Остальным языкам пересчетные коэффициенты определяются опытным путем, либо на базе накопленной статистики переводов в данной языковой паре, либо по выборочному расчету типовой для данного документа страницы путем деления числа знаков переведенной на русский язык стандартной страницы оригинала на число знаков в оригинальном тексте» [4].

Исходные данные. В связи с тем, что количество параметров, определяющих качество машинного перевода довольно велико, модель «оригинальный текст – перевод – обратный перевод» будем рассматривать как обратимый «черный ящик», имеющий определенную суммарную погрешность преобразования.

Учитывая наборы ограничений, изложенных выше (в работах [1, 4, 8, 11, 12]), зададим исходные данные к поставленной задаче. Величиной k_{IJ} будем называть значение пересчетного коэффициента [4] длин текстов, при переводе с одного языка на другой язык. Индексы I, J означают случайные алфавитные языки оригинального и целевого текста соответственно. Известно, что $k_{IJ} > 0$ и $k_{IJ} \neq 1$.

Даны три алфавитных языка M, N, P , имеющие в своем составе, m, n, p неповторяющихся символов соответственно. Пересчетный коэффициент k_{MN} при переводе с языка N на язык M имеет минимальное значение в диапазоне от 0 до 1, среди известных значений k_{IJ} . Аналогично, пересчетный коэффициент k_{MP} при переводе с языка P на язык M имеет максимальное значение среди известных значений $k_{IJ} > 1$.

Даны три текста A_M, B_N, C_P на языках M, N, P , с длинами a, b, c , соответственно, имеющие одинаковый смысл. При переводе текстов B_N и C_P на целевой язык M получены тексты B_{NM} и C_{PM} длинами $b \cdot k_{NM}$ и $c \cdot k_{PM}$, соответственно. При переводе отсутствуют смысловые ошибки, определяемые в соответствии с требованиями Правил [4].

Определение максимальных и минимальных редакционных расстояний на одном алфавитном языке. Длину любого текста можно представить в виде редакционного расстояния на одном языке между текстом нулевой длины и его собственной длиной, поэтому размеры любых текстов могут быть выражены при помощи редакционных расстояний. «Фактически редакционное расстояние корректно и адекватно вводит понятие расстояния на элементах свободной полугруппы (моноида), порожденной множеством букв» [13]. Редакционные расстояния (Левенштейна, Дамерау-Левенштейна) имеют дискретный характер [14] и целочисленные значения [15], в связи с этим все вычисления округляются до целого. Поскольку алфавитные языки не имеют общих символов (множества M, N, P не пересекаются) и имеют размер, который определяется через редакционные расстояния, используем следствия теоремы де Гуа (частного случая теоремы Пифагора для трехмерного пространства) [16–18]. Для удобства восприятия расположим тексты A_M, B_N, C_P в декартовой системе координат $Oxyz$, с началом в точке O и единичными отрезками, равными единичной операции редактирования. Языкам M, N, P будут соответствовать оси Ox, Oy, Oz , как показано на рис. 1,а. Длины текстов a, b, c будут соответственно отрезки OA, OB и OC .

Под *минимальными различиями на графемном уровне* Δ_{minM} на алфавитном языке M будем понимать минимальные редакционные расстояния между текстами, имеющими одинаковый смысл, на данном языке. Очевидно, что минимальные редакционные расстояния между текстами на одном языке будут:

$$\Delta_{minM} = 0.$$

Под *максимальными различиями на графемном уровне* Δ_{maxM} на алфавитном языке M будем понимать максимальные редакционные расстояния между текстами, имеющими одинаковый смысл, на данном языке. В общем случае Δ_{maxM} на языке M для произвольного набора текстов $a_1, a_2, \dots, a_n$ выражаются как:

$$\Delta_{maxM} \rightarrow \infty, \text{ при } a_1, a_2, \dots, a_n \rightarrow \infty. \quad (1)$$

Очевидно, что выражение (1) не имеет практического применения. При сравнении текстов на одном языке, в том числе полученных в результате перевода необходимо учитывать, что существует статистическая зависимость, ограничивающая диапазон допустимых значений Δ_{maxM} , между исходными и переведенными текстами:

$$i/j \approx k_{IJ}.$$

Практически применимое значение Δ_{maxM} , удовлетворяющее правилам [4], для алфавитного языка M , будет равно:

$$\Delta_{maxM} = |\Delta_{maxNM} + \Delta_{maxPM} + z_{NM} + z_{PM}|, \quad (2)$$

где Δ_{maxIJ} – максимальные различия на графемном уровне при переводе с алфавитного языка I на алфавитный язык J , между текстами на целевом языке J , z_{IJ} – ошибки в переводе.

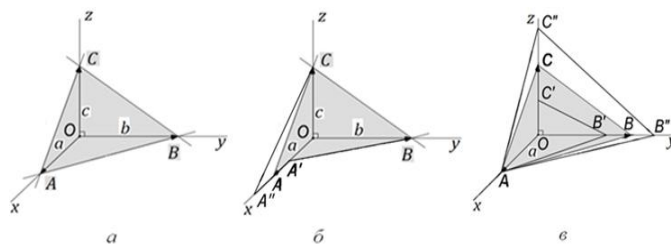


Рис. 1. Представление длин текстов в декартовой системе координат

На основании (2), для того, чтобы сумма ошибок $z_{NM} + z_{PM}$ соответствовала Правилам [4], возьмем предельный случай, когда единственная ошибка делает текст непригодным к публикации, учитывая, что редакционные расстояния может быть только целым числом, и численно выражается как: $z_{NM} + z_{PM} = 0$, тогда формула (2) принимает вид:

$$\Delta_{\max M} = |\Delta_{\max NM} + \Delta_{\max PM}|.$$

Предположим, что существуют отрезки AA' , AA'' отложенные на оси Ox , как показано на рис. 1,б, и являющиеся редакционными расстояниями между текстами A_M , B_{NM} и C_{PM} соответственно. Поскольку k_{NM} имеет минимальное значение в диапазоне от 0 до 1, среди k_{IJ} , а k_{MP} – максимальное среди $k_{IJ} > 1$, получим выражение для, максимального редакционного расстояния между текстами на одном языке (с учетом дискретности редакционных расстояний):

$$\Delta_{\max M} = |c \cdot k_{PM} - b \cdot k_{NM}|, \quad (3)$$

Определение максимальных и минимальных редакционных расстояний между текстами на двух алфавитных языках. Минимальные редакционные расстояния для текстов на двух алфавитных языках будут ограничены наименьшей из длин текстов, при $z=0$, независимо от k_{IJ} (меньший катет треугольника AOB на рис. 1,а), поскольку множества M , N , P непересекающиеся и в процессе перевода должны быть переведены все символы наименьшего текста.

Максимальные редакционные расстояния для текстов на двух алфавитных языков ограничены наибольшей из длин текстов, при $z=0$, независимо от k_{IJ} (большой катет треугольника AOB на рис. 1,а), поскольку редакционное расстояние не может быть больше чем длина переведенного текста.

Определение максимальных и минимальных редакционных расстояний между текстами на одном и двух алфавитных языках при использовании модели «оригинальный текст – перевод – обратный перевод». При использовании модели «оригинальный текст – перевод – обратный перевод» рассматриваемая система будет соответствовать рис. 1,в. Длина текста A_M на языке M будет равна a (отрезок AO). При переводе на другие алфавитные языки при отсутствии ошибок ($z=0$), необходимо учитывать величину k_{IJ} ; соответственно на целевом языке N :

- ♦ среднее значение длины перевода A_{MN} текста A_M будет $\approx a \cdot k_{MN}$;
- ♦ максимальное значение длины перевода $A_{MN\max}$ (отрезок OB'' на рис. 1,в) текста A_M будет равно:

$$A_{MN\max} = a;$$

- ♦ минимальное значение длины перевода $A_{MN\min}$ (отрезок OB' на рис. 1,в), текста A_M , учитывая выражение (3), будет равно:

$$A_{MN\min} = 2ak_{MN} - a,$$

при условии, что $A_{MN\max}$ и $A_{MN\min}$ равноудалены от среднего значения длины перевода A_{MN} текста A_M .

Аналогично на целевом языке P :

- ♦ среднее значение длины перевода A_{MP} текста A_M будет $\approx a \cdot k_{MP}$;
- ♦ максимальное значение длины перевода A_{MPmax} (отрезок OB'' на рис. 1,в) текста A_M , учитывая выражение (3), будет равно:

$$A_{MPmax} = 2ak_{MP} - a,$$

при условии, что A_{MPmax} и A_{MPmin} равноудалены от среднего значения длины перевода A_{MP} текста A_M .

- ♦ минимальное значение длины перевода A_{MPmin} (отрезок OB' на рис. 1,в), текста A_M будет равно:

$$A_{MPmin} = a.$$

При обратном переводе всегда справедливо соотношение $k_{IJ} = 1/k_{JI}$, тогда среднее значение длины обратного переводов A_{NM} и A_{PM} текста A_M будет приближенно равным a .

Для целевого языка N :

- ♦ максимальное значение длины обратного перевода A_{NMmax} (отрезок OC'' на рис. 1,в) текста A_M будет равно:

$$A_{NMmax} = ak_{NM} = a/k_{MN};$$

- ♦ минимальное значение длины обратного перевода A_{NMmin} (отрезок OC' на рис. 1,в), текста A_M будет, учитывая выражение (3), равно:

$$A_{NMmin} = (2ak_{MN} - a)k_{NM} = a(2k_{MN} - 1)/k_{MN},$$

при условии, что A_{NMmax} и A_{NMmin} равноудалены от среднего значения длины перевода A_{NM} текста A_M .

Аналогично для целевого языка P :

- ♦ максимальное значение длины обратного перевода A_{PMmax} (отрезок OC'' на рис. 1,в) текста A_M будет равно:

$$A_{PMmax} = ak_{PM} = a/k_{MP};$$

- ♦ минимальное значение длины обратного перевода A_{PMmin} (отрезок OC' на рис. 1,в), текста A_M будет, учитывая выражение (3), равно:

$$A_{PMmin} = (2ak_{MP} - a)k_{PM} = a(2k_{MP} - 1)/k_{MP},$$

при условии, что A_{PMmax} и A_{PMmin} равноудалены от среднего значения длины перевода A_{PM} текста A_M .

Результаты вычислительного эксперимента. Количественные сравнения производились при помощи стандартной метрики BLEU [3], анализирующей эталонные переводы человеком и машиной. В результате установлено, что точное (коэффициент схожести равен 1,0) совпадение составляет 43 % от общего числа аннотаций. Близкое совпадение (коэффициент схожести более 0,95) установлено в 32 % случаев. Частичное отклонение (коэффициент схожести более 0,8) зафиксировано в 22 % случаев. Существенное отклонение (коэффициент схожести менее 0,8) имеют 3 % результатов. Минимальный коэффициент схожести составляет 0,67.

Выводы. В общем виде решение задачи обработки текстовой информации в процессе подготовки к публикации результатов работы гибридного машинного переводчика сводится к решению следующих задач: вычисления погрешности входных данных; преобразования данных в соответствии с определенным алгоритмом; вычисления погрешности методом сравнения входных данных и повторно преобразованных в обратном направлении [19, 20].

Преимущество при использовании модели «оригинальный текст – перевод – обратный перевод» заключается в возможности точного установления границ множества вариантов перевода на графемном уровне при помощи редакционных расстояний Левенштейна по сравнению с однонаправленным переводом.

Заключение. Работа посвящена проблемам оценки и улучшения качества машинного перевода. Авторами предложен подход к реализации задачи определения набора условий для автоматического поиска оптимального варианта гибридного машинного перевода текста на графемном уровне на основе графического представление языковой модели алфавитных языков в виде декартовой системы координат с размерностью, равной единичному редакционному расстоянию (Левенштейна). При определении граничных условий использованы следствия теоремы де Гуа, действующие Правила Стандартизации ПР 50.1.027–2014 «Правила оказания переводческих и особых видов лингвистических услуг», метод деканонизации и модель «оригинальный текст – перевод – обратный перевод». Предложенное решение является обобщением известных аналогов и применимо для текстов на любом не иероглифическом алфавитном языке, не содержащих в своем составе слов из других языков. В ходе решения получены практически применимые результаты для рассматриваемых задач. Поэтому данная работа может представлять интерес для широкого круга специалистов, занимающихся проблемами машинного перевода в частности и переводоведением в целом.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Kornilov V., Glushan V., Lozovoy A. Numerical Assessment of Machine Translation Quality by Method of Near Duplicates Analysis // Proceedings of the 11th IEEE International Conference on Application of Information and Communication Technologies, AICT2017, Moscow, Russia. – Vol. 2. – P. 363-367. – DOI: 10.1109/ICAICT.2017.8686861.
2. Левенштейн В.И. Двоичные коды с исправлением выпадений, вставок и замещений символов // Доклады АН СССР. – 1965. – Т. 163, № 4. – С. 845-848.
3. Bhattacharyya P. Machine Translation // CRC Press Taylor & Francis Group. – 2015. – P. 5-6. – ISBN 9781439897188.
4. Правила Стандартизации ПР 50.1.027–2014, Правила оказания переводческих и особых видов лингвистических услуг. – Введ. 2014-04-01. – М.: Изд-во ФГУП «Стандартинформ», 2014. – 20 с.
5. Глушан В.М., Корнилов В.С., Лозовой А.Ю. Экспериментальное определение пригодности пристатейных аннотаций к русско-английскому гибриднему машинному переводу до и после автоматического предредактирования // Вестник РГРТУ. – 2019. – № 69. – С. 102-109. – DOI: 10.21667/1995-4565-2019-69-102-109.
6. Цвиллинг М.Я., Туровер Г.Я. О критериях оценки перевода // Тетради переводчика. – Вып. 15. – М.: Междунар. Отношения, 1973. – С. 5. – <https://wt-blog.net/perevodchiku/o-kriterijah-ocenki-perevoda-cvilling-turover.html> (дата обращения: 06.11.19).
7. Маслов С.Ю. Обратный метод установления выводимости для логических исчислений. Логические и логико-математические исчисления // Тр. МИАН СССР. – 1968. – № 98. – С. 26-87.
8. Bobber R.J. Underwater Electroacoustic Measurements // Naval Research Laboratory Underwater Sound Reference Division Orlando. – Florida, 1970. – P. 31-33.
9. Zeng J., Lau T., Lin Sh.-B., Yao Yu. Global Convergence of Block Coordinate Descent in Deep Learning // Proceedings of the 36th International Conference on Machine Learning. – Long Beach, California. – PMLR 97. – 2019. – <https://arxiv.org/pdf/1803.00225.pdf> (дата обращения: 06.11.19).
10. Nesterov Yu., Stich S. Efficiency of Accelerated Coordinate Descent Method on Structured Optimization Problems // CORE Discussion Paper. – 2016. – No. 03. – DOI: 10.1137/16M1060182.
11. Потемкин С.Б. Машинный перевод как средство стандартизации терминологии // Вестник Московского государственного областного университета. Серия: Лингвистика. – 2017. – № 5. – С. 77-84. – DOI: 10.18384/2310-712X-2017-5-77-84.
12. Филиппович Ю.Н., Сиренко А.В. Программный комплекс исследований психолингвистической модели вербального сознания на основе когнитивного и ассоциативного экспериментов // Вопросы психолингвистики. – 2011. – № 13. – С. 126-139. – <https://cyberleninka.ru/article/n/programmnyy-kompleks-issledovaniy-psiholingvisticheskoy-modeli-verbalnogo-soznaniya-na-osnove-kognitivnogo-i-assotsiativnogo> (дата обращения: 28.11.2019).

13. Вахрамеев М.А. Редакционное расстояние в свободных леворегулярных связках // Вестник Омского университета. – 2018. – Т. 23, № 3. – С. 15-19. – DOI: 10.25513/1812-3996.2018.23(3).15-19.
14. Пономарева Н.С., Реброва Г.Н., Колина Е.А. Применение расстояний редактирования при биоинформационном анализе геномов для задач оценки состояния репродуктивной системы // Фундаментальные исследования. – 2015. – № 7-4. – С. 774-777. – <https://www.fundamental-research.ru/pdf/2015/7-4/38819.pdf> (дата обращения: 24.11.2019).
15. Черненький В.М., Гапанюк Ю.Е. Методика идентификации пассажира по установочным данным // Инженерный журнал: наука и инновации. – 2012. – Вып. 3-89. – С. 30-39. – DOI: 10.18698/2308-6033-2012-3-89.
16. Теорема де Гуа. – [http://poivs.tsput.ru/Downloads/Article/5988/Теорема де Гуа.pdf](http://poivs.tsput.ru/Downloads/Article/5988/Теорема%20де%20Гуа.pdf) (дата обращения: 24.11.2019).
17. Amir-Moéz, A.R.; Byerly, R.E. Pythagorean theorem in unitary spaces // Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. – 1996. – No. 7. – P. 85-89. – <http://pefmath2.etf.rs/files/116/844.pdf> (дата обращения: 24.11.2019).
18. Kheifits A. The Theorem of Cosines for Pyramids // The College Mathematics Journal. – 2004. – Vol. 35, No. 5. – P. 385-388. – DOI: 10.2307/4146849.
19. Beck A. The 2-Coordinate Descent Method for Solving Double-Sided Simplex Constrained Minimization Problems // JOptim. Theory Appl. – 2014. – Vol. 162. – P. 892-919. – DOI: 10.1007/s10957-013-0491-5.
20. Николаев С.В. Алгебраические модели систем ЦОС: виды подобия и критерии близости // Известия ТРТУ. – 2002. – № 1 (24). – С. 108-109.

REFERENCES

1. Kornilov V., Glushan V., Lozovoy A. Numerical Assessment of Machine Translation Quality by Method of Near Duplicates Analysis, *Proceedings of the 11th IEEE International Conference on Application of Information and Communication Technologies, AICT2017, Moscow, Russia*, Vol. 2, pp. 363-367. DOI: 10.1109/ICAICT.2017.8686861.
2. Levenshteyn V.I. Dvoichnye kody s ispravleniem vypadeniy, vstavok i zameshcheniy simvolov [Binary Codes with Correction for Deletions and Insertions of the Symbols], *Doklady AN SSSR* [Reports of the USSR Academy of Sciences], 1965, Vol. 163, No. 4, pp. 845-848.
3. Bhattacharyya P. Machine Translation, *CRC Press Taylor & Francis Group*, 2015, pp. 5-6. ISBN 9781439897188.
4. Pravila Standartizatsii PR 50.1.027–2014, Pravila okazaniya perevodcheskikh i osobykh vidov lingvisticheskikh uslug [Standardization Rules PR 50.1.027–2014 Rules for the Provision of Translation and special types of Linguistic Services]. Moscow: Izd-vo FGUP «Standartinform», 2014, 20 p.
5. Glushan' V.M., Kornilov V.S., Lozovoy A.Yu. Eksperimental'noe opredelenie prigodnosti pristateynykh annotatsiy k russko-angliyskomu gibridnomu mashinnomu perevodu do i posle avtomaticheskogo predredaktirovaniya [Experimental Definition of Suitability of the Abstracts of Papers to the Russian-English Hybrid Machine Translation before and after the Automatic Pre-editing], *Vestnik RGRU* [Bulletin of RSTU], 2019, No. 69, pp. 102-109. DOI: 10.21667/1995-4565-2019-69-102-109.
6. Tsvilling M.Ya., Turover G.Ya. O kriteriyakh otsenki perevoda [About the criteria for assessing the translation], *Tetradi perevodchika* [Interpreters' notebooks]. Issue 15. Moscow: Mezhdunar. Otnosheniya, 1973, pp. 5. Available at: <https://wt-blog.net/perevodchiku/o-kriterijah-ocenki-perevoda-cvilling-turover.html> (accessed 06 November 2019).
7. Maslov S.Yu. Obratnyy metod ustanovleniya vyvodimosti dlya logicheskikh ischisleniy. Logicheskie i logiko-matematicheskie ischisleniya [An Inverse Method of Establishing Deducibility for Logical Calculi. Logical and Logical-Mathematical Calculus], *Tr. MIAN SSSR* [Proceedings of the Steklov Institute of Mathematics], 1968, No. 98, pp. 26-87.
8. Bobber R.J. Underwater Electroacoustic Measurements, *Naval Research Laboratory Underwater Sound Reference Division Orlando*. Florida, 1970, pp. 31-33.
9. Zeng J., Lau T., Lin Sh.-B., Yao Yu. Global Convergence of Block Coordinate Descent in Deep Learning, *Proceedings of the 36th International Conference on Machine Learning*. Long Beach, California. PMLR 97, 2019. Available at: <https://arxiv.org/pdf/1803.00225.pdf> (accessed 06 November 19).

10. Yu. Nesterov, S. Stich Efficiency of Accelerated Coordinate Descent Method on Structured Optimization Problems, *CORE Discussion Paper*, 2016, No. 03. DOI: 10.1137/16M1060182.
11. Potemkin S.B. Mashinnyy perevod kak sredstvo standartizatsii terminologii [Terminology Database and Machine Translation], *Vestnik Moskovskogo gosudarstvennogo oblastnogo universiteta. Seriya: Lingvistika* [Bulletin of the Moscow state regional University. Linguistics Series], 2017, No. 5, pp. 77-84. DOI: 10.18384/2310-712X-2017-5-77-84.
12. Filippovich Yu.N., Sirenko A.V. Programmnyy kompleks issledovaniy psikholingvisticheskoy modeli verbal'nogo soznaniya na osnove kognitivnogo i assotsiativnogo eksperimentov [Software to Research Psycholinguistic Model of Verbal Consciousness based on Cognitive and Associative Experiments], *Voprosy psikholingvistiki* [Questions of Psycholinguistics], 2011, No. 13, pp. 126-139. Available at: <https://cyberleninka.ru/article/n/programmnyy-kompleks-issledovaniy-psiholingvisticheskoy-modeli-verbalnogo-soznaniya-na-osnove-kognitivnogo-i-assotsiativnogo> (accessed 28 November 19).
13. Vakhrameev M.A. Redaktsionnoe rasstoyanie v svobodnykh levoregulyarnykh svyazkakh [Edit Distance in Free Left Regular Bands], *Vestnik Omskogo universiteta* [Bulletin of Omsk University], 2018, Vol. 23, No. 3, pp. 15-19. DOI: 10.25513/1812-3996.2018.23(3).15-19.
14. Ponomareva N.S., Rebrova G.N., Kolina E.A. Primenenie rasstoyaniy redaktirovaniya pri bioinformatsionnom analize genomov dlya zadach otsenki sostoyaniya reproduktivnoy sistemy [Application of editing distances in bioinformatic analysis of genomes for problems of evaluating the state of the reproductive system], *Fundamental'nye issledovaniya* [Fundamental study], 2015, No. 7-4, pp. 774-777. Available at: <https://www.fundamental-research.ru/pdf/2015/7-4/38819.pdf> (accessed 24 November 2019).
15. Chernen'kiy V.M., Gapanyuk Yu.E. Metodika identifikatsii passazhira po ustanovochnym dannym [Method of Passenger Identification according to Installation Data], *Inzhenernyy zhurnal: nauka i innovatsii* [Engineering journal: science and innovation], 2012, Issue 3-89, pp. 30-39. DOI: 10.18698/2308-6033-2012-3-89.
16. Teorema de Gua [de Gouï's Teorem]. Available at: [http://poivs.tsput.ru/Downloads/Article/5988/Teorema de Gua.pdf](http://poivs.tsput.ru/Downloads/Article/5988/Teorema%20de%20Gua.pdf) (accessed 24 November 2019).
17. Amir-Moéz, A.R.; Byerly, R.E. Pythagorean theorem in unitary spaces, *Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat.*, 1996, No. 7, pp. 85-89. Available at: <http://pefm2.etf.rs/files/116/844.pdf> (accessed 24 November 2019).
18. Kheifits A. The Theorem of Cosines for Pyramids, *The College Mathematics Journal*, 2004, Vol. 35, No. 5, pp. 385-388. DOI: 10.2307/4146849.
19. Beck A. The 2-Coordinate Descent Method for Solving Double-Sided Simplex Constrained Minimization Problems, *J Optim. Theory Appl.*, 2014, Vol. 162, pp. 892-919. DOI: 10.1007/s10957-013-0491-5.
20. Nikolaev S.V. Algebraicheskie modeli sistem TSOS: vidy podobiya i kriterii blizosti [Algebraic Models of DSP Systems: Types of Similarity and Proximity Criteria], *Izvestiya TRTU [Izvestiya TSURE]*, 2002, No. 1 (24), pp. 108-109.

Статью рекомендовал к опубликованию д.т.н., профессор Я.Е. Ромм.

Корнилов Василий Сергеевич – Южный федеральный университет; e-mail: koresh-jr@yandex.ru; 347928, г. Таганрог, пер. Некрасовский, 44; кафедра САПР; аспирант.

Глушань Валентин Михайлович – e-mail: gluval07@rambler.ru; кафедра САПР; профессор.

Лозовой Алексей Юрьевич – e-mail: lozovoy@sfedu.ru; кафедра ИЯ; доцент.

Kornilov Vasiliy Sergeevich – Southern Federal University; e-mail: koresh-jr@yandex.ru; 44, Nekrasovskiy, Taganrog, 347928, Russia; the department of computer aided design; postgraduate student.

Glushan Valentin Mihailovich – e-mail: gluval07@rambler.ru; the department of computer aided design; professor.

Lozovoy Aleksey Yurievich – e-mail: lozovoy@sfedu.ru; the department of foreign languages; associate professor.

Н.К. Полуянович, Н.В. Азаров, А.В. Огреничев, М.Н. Дубяго

**МНОГОЭТАПНЫЙ МЕТОД КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ
ТЕМПЕРАТУРНЫХ РЕЖИМОВ В СИЛОВОМ КАБЕЛЕ**

Статья посвящена исследованиям вопросам создания диагностики и прогнозирования термофлуктуационных процессов изоляционных материалов силовых кабельных линий (СКЛ) электроэнергетических систем на основе таких методов искусственного интеллекта, как нейронные сети и нечеткая логика. Показана необходимость разработки более совершенной методики анализа тепловых режимов в СКЛ. Обоснована актуальность задачи создания нейросетей (НС) для оценки пропускной способности, расчёта и прогнозирования температуры жил СКЛ в режиме реального времени на основе данных системы температурного мониторинга, с учетом изменения токовой нагрузки линии и внешних условий теплоотвода. По основным критериям проведено сравнение традиционных и нейросетевых алгоритмов для прогнозирования, показало преимущество НС методов. Проведена классификация НС методов и моделей прогнозирования температурных режимов КЛ. Предложенный нейросетевой алгоритм прогнозирования характеристик электрической изоляции был апробирован на контрольной выборке экспериментальных данных, по которым обучение искусственной нейронной сети не проводилось. Результаты прогноза показали эффективность выбранной модели. Для решения задачи прогнозирования ресурса СКЛ была выбрана сеть с прямым распространением данных и обратного распространения ошибки, т.к. сети такого типа в совокупности с активационной функцией в виде гиперболического тангенса являются в некоторой степени универсальной структурой для многих задач аппроксимации, приближения и прогнозирования. Разработана нейросеть для определения температурного режима токоведущей жилы силового кабеля. Проведен сравнительный анализ экспериментальных и расчетных характеристик распределений температуры, при этом исследовались различные нагрузочные режимы работы и функции изменения тока кабеля. При анализе данных было определено, что максимальное отклонение данных, полученных от нейросети от данных обучающей выборки, составило менее 2,2 % что является вполне приемлемым результатом. Модель может быть использована в устройствах и системах непрерывного диагностирования силовых кабелей по температурным режимам.

Искусственный интеллект; нейронные сети; термофлуктуационные процессы; изоляционные материалы; прогнозирование; надежность систем энергоснабжения.

N.K. Poluyanovich, N. V. Azarov, A.V. Ogrenichev, M. N. Dubyago

**MULTI-STAGE METHOD FOR SHORT-TERM FORECASTING OF
TEMPERATURE CHANGES MODES IN THE POWER CABLE**

The article is devoted to research on the creation of diagnostics and prediction of thermofluctuation processes of insulating materials of power cable lines (PCL) of electric power systems based on such methods of artificial intelligence as neural networks and fuzzy logic. The necessity of developing a better methodology for the analysis of thermal conditions in PCL is shown. The urgency of the task of creating neural networks (NS) for assessing the throughput, calculating and predicting the temperature of PCL conductors in real time based on the data of the temperature monitoring system, taking into account changes in the current load of the line and the external conditions of the heat sink, is substantiated. Based on the main criteria, traditional and neural network algorithms for forecasting are compared, and the advantage of NS methods is shown. The classification of NS methods and models for predicting the temperature conditions of cosmic rays has been carried out. The proposed neural network algorithm for predicting the characteristics of electrical isolation was tested on a control sample of experimental data on which training of an artificial neural network was not carried out. The forecast results showed the effectiveness of the selected model. To solve the problem of PCL resource prediction, a network was selected with direct data distribution and back propagation of the error, because Networks of this

type, together with the activation function in the form of a hyperbolic tangent, are to some extent a universal structure for many problems of approximation, approximation, and forecasting. A neural network was developed to determine the temperature regime of a current-carrying core of a power cable. A comparative analysis of the experimental and calculated characteristics of the temperature distributions was carried out, while various load modes and the functions of changing the cable current were investigated. When analyzing the data, it was determined that the maximum deviation of the data received from the neural network from the data of the training sample was less than 2.2 %, which is an acceptable result. The model can be used in devices and systems for continuous diagnosis of power cables by temperature conditions.

Artificial intelligence; neural networks; thermal fluctuation processes; insulation materials; forecasting; reliability of power supply systems.

Введение и постановка задачи. Устойчивое функционирование электросетевого комплекса невозможно без надежной и качественной работы распределительных электрических сетей 6–10/0,4 кВ, которые являются завершающим звеном в системе обеспечения потребителей электрической энергией и находятся в непосредственном взаимодействии с конкретным потребителем [1]. К настоящему моменту значительная часть работающих силовых кабельных линий (СКЛ) выработала свой нормативный ресурс, однако возможности по его замене весьма ограничены. Результаты эксплуатации СКЛ свидетельствуют о том, что основной причиной его отказов является снижение электрической прочности изоляции. В процессе эксплуатации изоляция СКЛ подвергаются тепловым, электрическим, химическим, механическим и другим видам воздействий, в результате чего происходит изменение её электрических свойств и, как следствие, некоторые изменения претерпевают и технические характеристики изоляционных материалов [2]. В этих условиях важно разработать долговременную программу по обновлению СКЛ. Реализация такой программы требует разработки новых методов оценки и поддержания технического состояния действующего оборудования кабельных систем на заданном уровне и продления его работоспособности на максимально возможный срок.

Важной характеристикой СКЛ систем электроснабжения является нагрузочная способность, определяемая тепловым режимом, по которой рассчитывается площадь поперечного сечения и их пропускная способность [3]. Расчёт пропускной способности необходим и осуществляется при проектировании СКЛ. Однако использование даже самых точных методик не позволяет учесть на практике все нюансы, которые весьма существенно влияют на температурный режим работы СКЛ. В связи с этим, при проектировании СКЛ стремятся выбирать коэффициенты с определённым запасом, поэтому в ряде случаев кабели оказываются недогруженными, а иногда работают на пределе температурного режима [3]. Одним из решений указанной проблемы является мониторинг температуры кабеля во время работы.

Текущий анализ диагностических систем подтверждает, что имеется объективная научно-техническая проблема создания комплексных систем диагностирования, построенных на разносторонних принципах, дающих высокий уровень достоверности постановки диагноза и прогнозирования технического состояния СКЛ. Перспективным направлением является создание инфраструктуры диагностирования неисправностей СКЛ, основанной на использовании искусственных нейронных сетей. Искусственные нейронные сети (ИНС) в задачах диагностирования и прогнозирования остаточного ресурса изоляционных материалов могут служить в роли подсистемы выборки и принятия решений, а затем отдавать полученные данные другим подсистемам управления.

Применение интеллектуальной системы управления приведет к повышению общей надежности энергосети и повышению энергоэффективности функционирования энергетического сектора, уменьшению влияния человеческого фактора и снижению аварийных ситуаций.

- ♦ задачей исследования являются разработка интеллектуальной системы прогнозирования температуры жилы СКЛ для планирования режимов работы электросети с целью повышения надежности и энергоэффективности их взаимодействия с объединенной энергосистемой;

- ♦ целью исследования является повышение надежности и энергоэффективности распределенных энергосистем за счёт применения новых неразрушающих методов и алгоритмов мониторинга и диагностики ресурса СКЛ на базе искусственных нейронных сетей для краткосрочного и среднесрочного прогнозирования температурных режимов токоведущей жилы силовых кабельных линий.

Задачи диагностики и прогнозирования термофлуктуационных процессов СКЛ. Различаются методики по расчету стационарных и нестационарных тепловых процессов в кабеле, методики без возможности анализа тепловых полей и с возможностью анализа тепловых полей, методики на основе эмпирических зависимостей и на основе дифференциальных уравнений.

Нормативным документом, устанавливающим расчет тепловых процессов в силовых кабелях, является МЭК 60287 и его российский аналог ГОСТ [5–7]. Методика расчета номинальной токовой нагрузки СКЛ, исходя из допустимой температуры жилы рассматривает в ГОСТ 60287-1-1 [5]. Методика корректировки соотношений из ГОСТ 60287-2-1 с учетом влияния соседних кабелей приводится в ГОСТ 60287-2-2 [7]. Тепловые процессы в одножильном одиночном кабеле, проложенном в воздухе, и соответствующие краевые условия описываются в [8] системой квазилинейных дифференциальных уравнений параболического типа. Методика расчета тепловых полей СКЛ на основе решения связанной задачи электродинамики и теплопередачи для СПЭ кабелей с помощью уравнения плоскопараллельного квазистационарного поля, рассматривается в [9]. Даны расчетные картины тепловых полей с учетом токов в экране кабеля. Расчет стационарных тепловых процессов в кабеле подробно рассмотрен [10], а приведенные соотношения показывают, что рассмотренные выше стандарты МЭК созданы на их основе. Методика расчета тепловых сопротивлений для кабелей, проложенных под землей, рассматривается в [11], а приведенные соотношения аналогичны с рекомендуемыми МЭК. Стационарные тепловые процессы в СКЛ, рассматриваемые МЭК и ГОСТ, больше освещены в литературе чем нестационарные режимы, что связано со сложностью получения достаточно адекватных соотношений и алгоритмов на их основе. Многофакторность задачи обусловила тот факт, что большинство традиционных методик расчета базируется на эмпирических зависимостях и коэффициентах, что ограничивает их применимость диапазоном воздействующих факторов и условиями окружающей среды. Даже в случае сложной системы дифференциальных уравнений уйти от эмпирических коэффициентов очень сложно. С точки зрения теплотехники силовой кабель не является сложной системой, однако факторы окружающей среды и электромагнитные процессы усложняют задачу анализа тепловых режимов на несколько порядков. Способ прокладки кабеля, изменение параметров почвы, способы заземления экранов, взаимоиндукция между жилами, краевой и поверхностный эффекты, электрические параметры изоляции – это и многое другое влияет на тепловые процессы в кабеле. Стремление существующих методик расчета учесть все эти факторы приводит к значительному усложнению алгоритмов и снижению гарантии их адекватности при изменении внешних условий.

Таким образом, существует необходимость разработки более совершенной методики анализа тепловых режимов СКЛ, максимально инвариантной к внешним условиям. Поставим задачу разработать математическую модель тепловых процессов силового кабеля, не использующую эмпирические коэффициенты и зависимости, но позволяющую вести расчет температуры наиболее нагретой точки изоляции кабеля с достаточной точностью.

1. Прогнозирование термофлуктуационных процессов. Анализ полученных результатов по прогнозированию изменения технического состояния изоляции СКЛ в условиях эксплуатации с использованием результатов эксплуатационного мониторинга [3] показал, что такие методы позволяют повысить эффективность эксплуатации кабельных систем.

Прогнозирование проводится в различных временных диапазонах. В работе [12], например, временные диапазоны прогнозирования классифицируются следующим образом: в пределах текущих суток – оперативный прогноз; вперед на сутки-неделю-месяц – краткосрочный; на месяц-квартал-год – долгосрочный.

Традиционные статистические модели могут быть условно разделены на регрессионные модели и модели на основе временных рядов. Подробное обсуждение этих моделей приведено в [13]. С развитием теории новых информационных технологий в последнее десятилетие было предложено решение задачи прогнозирования нагрузки нетрадиционными методами, а именно, с использованием моделей на основе экспертных систем и искусственных нейронных сетей [14–17]. Предпочтение таких моделей традиционным обусловлено тем, что не требуется построения модели объекта, не теряется работоспособность при неполной входной информации. Они обладают устойчивостью к помехам, имеют высокое быстродействие.

В настоящее время существует много методов, позволяющих производить прогнозирование в электроэнергетической системе [18].

Качество прогноза во многом зависит от выбранной математической модели [19]. Основной недостаток существующих методов прогноза нагрузки состоит в необходимости построения модели нагрузки, т.е. зависимости нагрузки от влияющих факторов, и постоянного ее уточнения. Другой недостаток состоит в неточном установлении соотношения между входными и выходными переменными, так как зависимости между ними нелинейные. Входные переменные представляют собой название дня недели, ретроспективные данные о нагрузке, информацию о температуре окружающего воздуха, влажности, и др. Выходные переменные – это предстоящая нагрузка на каждый час суток.

Искусственная нейронная сеть (ИНС) [20] способна установить сколь угодно сложные соотношения между различными оперативно вводимыми, непрерывно изменяющимися исходными данными.

Обычно прогнозирование какого-либо параметра выполняется на величину шага скользящего окна по обучающей выборке [21]. В случае необходимости долгосрочного прогноза применяется последовательный итерационный прогноз на величину шага, кратного шагу обучающей выборки. При этом качество долгосрочного прогноза обычно ухудшается с ростом числа итераций прогноза (например, при увеличении временного периода прогноза) [22].

Для решения проблемы ухудшения долгосрочного прогноза было предложено использовать в качестве оценки качества обучения не отклонение в величине прогноза на один шаг скользящего окна по обучающей выборке, а суммарное отклонение прогноза по всем значениям обучающей выборки, причем с получением новых значений на основе прогноза, полученного на предыдущем шаге. Полученные значения поступали в скользящее окно вместо экспериментальных данных, и на этих значениях строился новый прогноз, также по пяти параметрам.

2. Метод скользящего окна в многоэтапном прогнозировании. Прогнозирование временных рядов может быть реализовано с помощью нейронной сети в виде многослойного персептрона (MLP). В работе используется метод обучения с учителем. Прогнозируемой величиной является температура поверхности силового кабеля (θз.о). Обучение с учителем – метод, в котором существует некоторое количество входов (X) и выходов (y) и используется алгоритм обучения отображения от входа к выходу:

$$Y = f(X).$$

Метод основан на аппроксимации входных параметров таким образом, чтобы при подаче на вход новых данных (X), можно было бы спрогнозировать выходные данные (Y).

Цель исследования. Для использования метода обучения с учителем, необходимо преобразовать временной ряд ($\theta_{з.о}$) в определенный вид, приемлемый для метода подконтрольного обучения. В работе временной ряд представляет собой температуры силового кабеля, снятые послойно (рис. 1), в нашем случае используется температура на поверхности силового кабеля ($\theta_{з.о}$).

Пример обучения с учителем может быть проиллюстрирован табл. 1, где один из столбцов является входными данными (X), а второй – выходными данными (y).

Таблица 1

Пример данных обучения с учителем.

Время	$\theta_{з.о}$	$I_{ж}$	N_k
1	23	460	1
2	23.06	460	1
3	23.12	460	1
4	23.19	460	1
5	23.25	460	1

Алгоритм итеративно делает прогноз на основании данных обучающей выборки ($X(n)$) и корректируется, обновляя весовые коэффициенты. Обучение прекращается после того, как достигается приемлемый уровень эффективности (минимальная ошибка прогноза E).

2.1. Применение метода скользящего окна для преобразования временных рядов. Временной ряд в виде температуры поверхности СКЛ ($\theta_{з.о}$) необходимо преобразовать для последующего контролируемого обучения НС (обучения с учителем). Суть преобразования данного временного ряда заключается в формировании временного ряда ($\theta_{з.о}$) в определенном порядке, при котором первое значение временного ряда X_1 устанавливается как входной параметр ($\theta_{з.о_1}$), а последующее значение X_1 устанавливается как выходной параметр (Y_1). Пример представлен в табл. 2.

Таблица 2

Пример данных временного ряда.

Время	Входные переменные X_n		
	$\theta_{з.о}$	$I_{ж}$	N_k
1	23	460	1
2	23.06	460	1
3	23.12	460	1
4	23.19	460	1
5	23.25	460	1

После преобразования данные будут сформированы следующим образом (табл. 3), где X – это входной параметр ($\theta_{з.о_1}$), а Y – выходной ($\theta_{з.о_2}$):

Таблица 3

Преобразованные данные

Входные переменные X_n			Выходные переменные Y_n
$\theta_{з.о}(n)$	$I_{ж}$	N_k	$\theta_{з.о}(n+1)$
?	?	?	23
23	460	1	23.06
23.06	460	1	23.12
23.12	460	1	23.19
23.19	460	1	?

При сравнении второй и третьей таблиц видно:

- ♦ порядок данных (X_n) строго сохранен;
- ♦ так как первое значение столбца выходных данных (Y_n) не имеет предыдущего результата, то данную строку необходимо удалить;
- ♦ как и в первом случае, в последней строке входных данных (X_n) так же не хватает одного значения, поэтому ее необходимо удалить.

Ранее установлено, что использование предыдущего шага по времени, чтобы спрогнозировать следующий шаг называется скользящим окном. В статистике и анализе временных рядов это называется задержкой или методом задержки. Метод скользящего окна позволяет преобразовать любой набор временного ряда для подконтрольного обучения НС (обучения с учителем).

2.2. Метод скользящего окна в многомерном временном ряду. По количеству значений временные ряды делятся на:

- ♦ одномерные временные ряды – набор данных, где дано единственное значение в единицу времени, например показатель температуры каждый час;
- ♦ многомерные временные ряды – набор данных, где два или более значения даны за единицу времени, например показатель температуры каждый час, а также сила действующего тока и температура окружающей среды.

В работе временной ряд является многомерным, т.к. кроме температуры на поверхности кабеля ($\theta_{з.о}$) также учитывается сила действующего тока ($I_{ж}$) и номер кабеля (N_k).

Пример многомерного временного ряда приведен в табл. 4.

Таблица 4

Многомерные временной ряд

Время	Входные переменные X_n		
	$\theta_{з.о}$	$I_{ж}$	N_k
1	23	460	1
2	23.06	460	1
3	23.12	460	1
4	23.19	460	1
5	23.25	460	1

Для того чтобы спрогнозировать следующее значение временного ряда вышеуказанным методом, необходимо преобразовать временной ряд для обучения с учителем. Для этого необходимо воспользоваться методом скользящего окна. Таким образом, получили три входа ($X_n = \theta_{з.о}(n), I_{ж}, N_k$) и один выход $Y_n = \theta_{з.о}(n+1)$ для каждой записи (табл. 5):

Таблица 5

Сформированный многомерный временной ряд для подконтрольного обучения

Входные переменные X_n			Выходные переменные Y_n
$\theta_{з.о}(n)$	$I_{ж}$	N_k	$\theta_{з.о}(n+1)$
?	?	?	23
23	460	1	23.06
23.06	460	1	23.12
23.12	460	1	23.19
23.19	460	1	?

Для прогнозирования более чем на один шаг вперед необходимо воспользоваться методом многоэтапного прогнозирования.

2.3. Скользящее окно в многоэтапном прогнозировании. Прогнозирование бывает двух видов:

- ♦ одноэтапный прогноз – прогнозирование на один шаг вперед ($t + 1$);
- ♦ многоэтапный прогноз – прогнозирование на два и более шагов ($t + n$).

Вышеприведённые примеры являются одноэтапными прогнозами временных рядов. Для многоэтапного прогнозирования необходимо воспользоваться методом скользящего окна.

Для построения многоэтапного прогноза многомерного ряда воспользуемся данными, приведенными в табл. 5.

Входные ($X_n = \theta_{з.о}(n), I_{ж}, N_k$) и выходные ($Y_n = \theta_{з.о}(n+1), \theta_{з.о}(n+2)$) переменные формируются таким образом, чтобы ширина окна была равна одному шагу, а прогноз был двухэтапным, получаем преобразованный многомерный временной ряд (табл. 6).

Таблица 6

Многомерный временной ряд с двухэтапным прогнозом

Входные переменные X_n			Выходные переменные Y_n	
$\theta_{з.о}(n)$	$I_{ж}$	N_k	$\theta_{з.о}(n+1)$	$\theta_{з.о}(n+2)$
?	?	?	23	23.06
23	460	1	23.06	23.12
23.06	460	1	23.12	23.19
23.12	460	1	23.19	23.25
23.19	460	1	23.25	?
23.25	460	1	?	?

Из табл. 6 видно, что первая запись входных параметров (X_n) и две последних записи выходных параметров (Y_n) должны быть удалены, т.к. они бесполезны. Для достижения необходимой точности и производительности модели нейросети для прогнозирования температуры поверхности силового кабеля ($\theta_{з.о}$) необходимы аккуратность и множество экспериментов для подбора необходимой ширины скользящего окна.

2.4. Математическое описание алгоритма метода скользящего окна. Для надежного, бесперебойного электроснабжения потребителей необходима оценки длительно допустимых значений нагрева жилы кабеля напряжением 6–10 кВ с

изоляцией из сшитого полиэтилена при различных значениях тока жилы кабеля. В режиме динамически изменяющейся нагрузки силового кабеля необходим прогноз его температурного режима не на одну единицу наблюдения, а на несколько. Для этого в качестве целевых выбирается соответствующее число значений, которое называется горизонтом прогноза. Количество наблюдений ряда, которые берутся в качестве входных значений называется глубиной прогноза.

На первом этапе необходимо было решить задачи выбора и анализа модели НС, идентификация ее параметров и наконец проверка модели на адекватность. Основная задача заключается в построении обоснованных вычислительных алгоритмов прогноза и их нейросетевой реализации.

Рассмотрим задачу построения прогноза не на одну единицу наблюдения, а на несколько. Для этого в качестве целевых выбирается соответствующее число значений, которое называется горизонтом прогноза.

Количество наблюдений ряда, которые берутся в качестве входных значений называется глубиной прогноза.

Алгоритм трансформации, позволяющий из членов временного ряда ($X(n)$) входных параметров СКЛ:

$\theta_{з.о}(1...n)$ – температура защитной оболочки, $\theta_{о.с}(1...n)$ – температура окружающей среды, $I_{ж}(1...n)$ – ток в жиле СКЛ сформировать “окно”, являющееся обучающим множеством для построения прогноза.

Окно представляет собой временной интервал, содержащий набор значений ($\theta_{з.о}(1...n)$, $I_{ж}(1...n)$, $\theta_{о.с}(1...n)$), которые используются для формирования обучающего примера.

В процессе работы алгоритма окно смещается по временной последовательности на единицу наблюдения, и каждое положение окна образует один пример. Алгоритм трансформации представлен на рис. 1,а.

При краткосрочном прогнозировании чаще используются искусственные нейронные сети с нелинейными функциями активации. Для построения многоэтапного прогноза (рис. 1,б) многомерного ряда входные ($X_n = \theta_{з.о}(n)$ $I_{ж}$ N_k) и выходные ($Y_n = \theta_{з.о}(n+1)$, $\theta_{з.о}(n+2)$) переменные формируются таким образом, чтобы ширина окна была равна одному шагу, а прогноз был двухэтапным, получаем преобразованный многомерный временной ряд. Алгоритм итеративно делает прогноз на основании данных обучающей выборки $X(n)$ и корректируется, обновляя весовые коэффициенты. Обучение прекращается после того, как достигается приемлемый уровень эффективности (минимальная ошибка прогноза E).

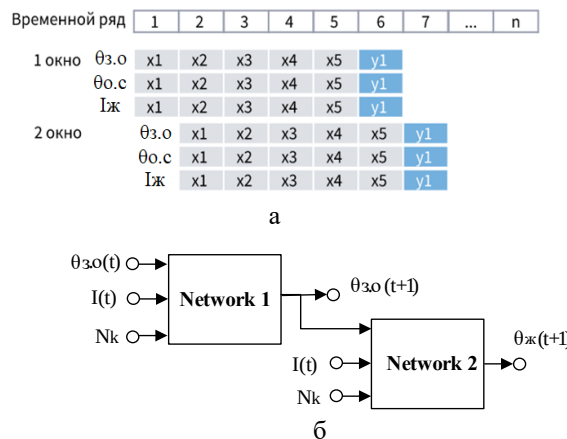


Рис. 1. Метод скользящего окна: а – алгоритм трансформации, б – архитектура последовательно распространяемой структуры

3. Экспериментальные исследования по обучающей выборки. Для составления модели прогнозирования на базе ИНС была выбрана система Matlab. При прогнозировании температуры жилы кабеля создание, обучение и моделирование нейронной сети осуществлялось с помощью Neural Network Toolbox в системе Matlab. Исследованы термофлуктуационные процессы в силовом кабеле (без повреждений) АПв Пу г-1х240/25-10 при реальной диаграмме действующего значения тока жилы кабеля и разработанной модели прогноза с использованием искусственных нейронных сетей (ИНС) на глубоких [23, 25] ретроспективных данных температур исследуемого кабеля за 2015–2019 годы.

База данных (рис. 2) содержит данные обучающей выборки, для исходных образцов кабеля (нового) и состаренных образцов – тестирующая выборка. Содержатся значения по температуре $\theta_{з.о.}$, $\theta_{о.с.}$ – защитной оболочки, окружающей среды, получены в ходе работы с реальными образцами силового кабеля АПв Пу г-1х240/25-10.



Рис. 2. Структура базы данных

С помощью многоэтапного метода краткосрочного прогнозирования были исследованы четыре образца СКЛ. Диаграммы прогнозируемых и экспериментальных температур кабеля № 5, 6, 7, 8 приведены на рис. 3. Величины ошибки многоэтапного прогноза температуры жил различных образцов кабеля, приведены в табл. 7.

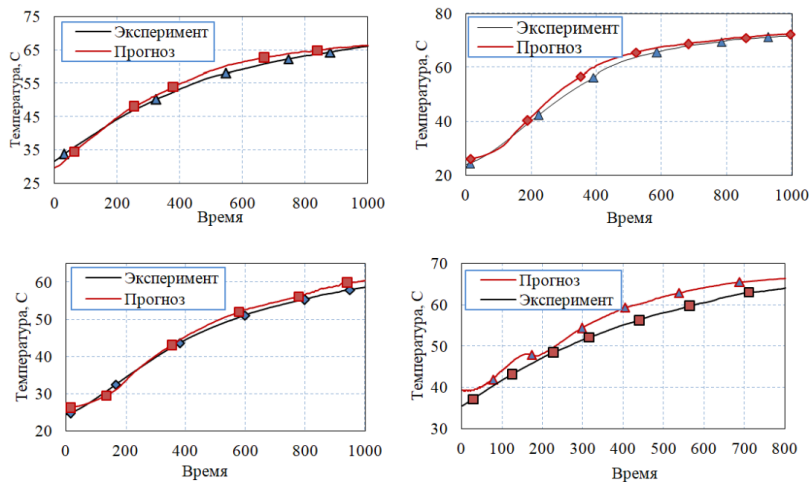


Рис. 3. График прогнозируемых и экспериментальных температур кабеля № 5, 6, 7, 8

Таблица 7

Результаты многоэтапного прогнозирования температуры кабеля

Тип нейронной сети	Номер образца кабеля	Средняя ошибка прогноза	
		ε , °C,	ε , %
Многослойный перцептрон (MLP) / Multilayered perceptron (MLP)	Кабель № 5	2.5	3.7 %
	Кабель № 6	3.8	5.3 %
	Кабель № 7	2.3	3.3 %
	Кабель № 8	4.6	5.8 %

Вывод. При анализе данных было определено, что максимальное отклонение данных нейросети (НС) от данных обучающей выборки, составило не более – 5.8 %, что является вполне приемлемым результатом.

Для подтверждения работоспособности модели проведено исследование. Для каждого значения тока жилы экспериментально сняты значения температур ($\theta_{з.о}$, $\theta_{о.с}$) по 1000 точек, построен график температурной зависимости $\theta_{о.с}$, рис. 3,а (эксперимент). Используя метод многоэтапного прогнозирования и отснятую в тысячу точек выборку построена спрогнозированная зависимость $\theta_{о.с}$ на 2000 точек, рис. 4,а (прогноз). Затем были до сняты экспериментальные данные до 2000 значений ($\theta_{з.о}$, $\theta_{о.с}$), рис. 3,б (эксперимент) и проведено сравнение этих диаграмм. Результаты прогноза подтвердились. Средняя ошибка прогноза для тестовой выборки составила 2,2 процента, табл. 8.

Таблица 8

Результаты погрешности прогноза

Номер образца кабеля	Средняя ошибка прогноза	
	ε , °C,	ε , %
Кабель № 7	1,13	2,2 %

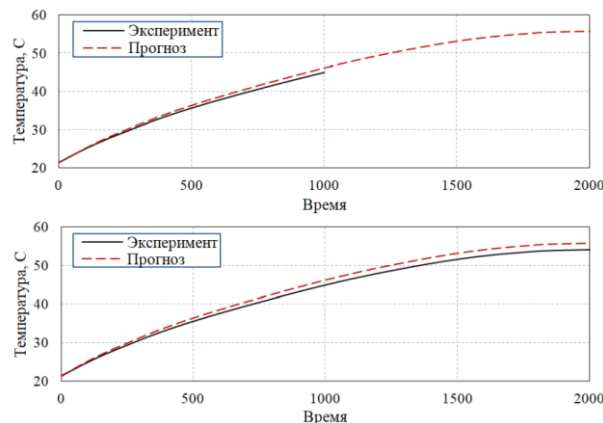


Рис. 4. Графики экспериментальных и прогнозируемых температур кабеля № 7 (состаренного)

Вывод. Предложенный нейросетевой алгоритм прогнозирования характеристик электрической изоляции был апробирован на контрольной выборке экспериментальных данных, по которым обучение искусственной нейронной сети не проводилось. Результаты прогноза показали эффективность выбранной модели. Многоэтапный метод прогнозирования температурного режима токоведущей жилы силового кабеля позволяет осуществлять прогноз для тестовых данных с погрешностью в 2,2 %

Заключение. Предложена методика и разработана нейросеть для определения температурного режима токоведущей жилы силового кабеля. Предложенный нейросетевой алгоритм прогнозирования характеристик электрической изоляции был апробирован на контрольной выборке экспериментальных данных, по которым обучение искусственной нейронной сети не проводилось. Результаты прогноза показали эффективность выбранной модели. Разработан многоэтапный метод прогнозирования ресурса позволяет осуществлять прогноз для тестовых данных с погрешностью в 2,2 %. Для прогнозирования более чем на один шаг вперед необходимо воспользоваться методом многоэтапного прогнозирования. Основная область применения разработанной нейросети для определения температурного режима токоведущей жилы заключается в диагностике и прогнозировании ресурса ЭИ силового кабеля.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Положение ОАО «Россети» о единой технической политике в электросетевом комплексе. – М.: ОАО «Россети», 2013. – 196 с.
2. Анищенко В.А., Колосова И.В. Основы надежности систем электроснабжения. – Мн.: БНТУ, 2007.
3. Дубяго М.Н., Полуянович Н.К. Совершенствование методов диагностики и прогнозирования электроизоляционных материалов систем энергоснабжения. Южный федеральный университет. – Ростов-на-Дону; Таганрог: Изд-во ЮФУ, 2019. – 192 с.
4. Каталог оборудования компании LIOS Technology GmbH – 2010.
5. ГОСТ Р МЭК 60287-1-1. Кабели электрические. Расчет номинальной токовой нагрузки. Ч. 1-1. Уравнения для расчета номинальной токовой нагрузки и расчет потерь. Общие положения. – М.: Стандартинформ, 2009. – 28 с.
6. ГОСТ Р МЭК 60287-1-2. Кабели электрические. Расчет номинальной токовой нагрузки. Ч. 1-2. Уравнения для расчета номинальной токовой нагрузки и расчет потерь. Коэффициенты потерь, обусловленных вихревыми токами в оболочке, для двух цепей, расположенных в одной плоскости. – М.: Стандартинформ, 2009. – 23 с.

7. ГОСТ Р МЭК 60287-2-2. Кабели электрические. Расчет номинальной токовой нагрузки. Ч. 2-2. Тепловое сопротивление. Метод расчета коэффициентов снижения максимально допустимой токовой нагрузки для групп кабелей, проложенных на воздухе и защищенных от прямого солнечного излучения. – М.: Стандартинформ, 2009. – 12 с.
8. Ларина Э.Т. [и др.]. Расчет переходных тепловых режимов одножильного силового кабеля с пластмассовой изоляцией, проложенного в воздухе // Электротехника. – 1991. – № 10. – С. 39-42.
9. Титков, В.В. К оценке теплового режима трехфазной линии из СПЭ-кабеля // Кабель-News. – 2009. – № 10. – С. 47-51.
10. Neher J.M., McGrath M.H. The calculation of the temperature rise and load capability of cable systems // Philadelphia electric company. – October 1957. – P. 752-763.
11. Leon F. Calculation of underground cable ampacity // CYME international T&D Inc. – 2005.
12. Меламед М.А. Современные методы анализа и прогнозирования режимов электропотребления в электроэнергетических системах // Итоги науки и техники. Серия «Энергетические системы и их автоматизация». – 1988. – Т. 4. – С. 4-111.
13. Gross G. and Galiana F.D. Short term load forecasting // Proc. IEEE. – 1987. – Vol. 75, No. 12. – P. 1558-1573.
14. Chen S.T., David C.Y., Moghaddamjo A.R. Weather sensitive short-term load forecasting using non fully connected artificial neural network // IEEE Trans. on Power Systems. – 1992. – Vol. 7, No. 3. – P. 1098-1105.
15. Hsy Y., Ho K. Fuzzy expert systems: An application to short term load forecasting // IEE Proceedings – C. – 1992. – Vol. 139, No. 6. – P. 471-477.
16. Lee K.Y., Park J.H. Short-term load forecasting using an artificial neural network // IEEE Trans. on Power Systems. – 1992. – Vol. 7, No. 1. – P. 124-130.
17. Meldorf M., Kilter J., Pajo R. Comprehensive Modelling of Load // CIGRE Regional Meeting, June 18-20, 2007, Tallinn, Estonia. – P. 145-150.
18. Шуилова Г.П., Готман Н.Э., Старцева Т.Б. Прогнозирование электрических нагрузок ЭЭС с использованием методов искусственного интеллекта // Сб. трудов Российского национального симпозиума по энергетике. Казань 10–14 сент. 2001 г. – М.: КГЭУ, 2001. – С. 103-106.
19. Хайкин С. Нейронные сети: полный курс. – 2-е изд. – М.: Изд. дом «Вильямс», 2006. – 1104 с.
20. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы: пер. с польск. И.Д. Рудинского. – М.: Горячая линия – Телеком, 2008. – 452 с.
21. Кричевский М.Л. Интеллектуальный анализ в менеджменте. – СПб.: СПбГУАП, 2005. – 48 с.
22. Сидоров С.Г. и др. Многопроцессорная реализация нейросетевого алгоритма прогнозирования изменения характеристик электрической изоляции // Вестник ИГЭУ. – 2011. – Вып. 1.
23. Полуянович Н.К., Дубяго М.Н. Прогнозирование ресурса кабельных линий с использованием метода искусственных нейронных сетей // Известия ЮФУ. Технические науки. – 2019. – № 3 (205). – С. 51-62.
24. Galushka V.V., Fathi V.A. Formation of the training sample by using artificial neural networks in search problems database errors // Inženernyj vestnik Dona (Rus). – 2013. – No. 2. – URL: ivdon.ru/ru/magazine/archive/n2y2013/1597/.
25. Дубяго М.Н., Полуянович Н.К. Метод оценки и прогнозирования остаточного ресурса изоляции кабельных линий // Известия ЮФУ. Технические науки. – 2019. – № 3. – С. 132-143.

REFERENCE

1. Polozhenie OAO «Rosseti» o edinoj tekhnicheskoy politike v elektrosетevom komplekse [Regulation of JSC "Rosseti" on the unified technical policy in the electric grid complex]. Moscow: OAO «Rosseti», 2013, 196 p.
2. Anishchenko V.A., Kolosova I.V. Osnovy nadezhnosti sistem elektrosnabzheniya [Fundamentals of reliability of power supply systems]. Minsk: BNTU, 2007.

3. *Dubyago M.N., Poluyanovich N.K.* Sovershenstvovanie metodov diagnostiki i prognozirovaniya elektroizolyatsionnykh materialov sistem energosnabzheniya [Improving methods of diagnostics and forecasting of electrical insulation materials of power supply systems]. Rostov-on-Don; Taganrog; Izd-vo YuFU, 2019, 192 p.
4. Katalog oborudovaniya kompanii LIOS Technology GmbH – 2010 [Directory of equipment of the company lios Technology GmbH – 2010].
5. GOST R MEK 60287-1-1. Kabeli elektricheskie. Raschet nominal'noy tokovoy nagruzki. Ch. 1-1. Uravneniya dlya rascheta nominal'noy tokovoy nagruzki i raschet poter'. Obshchie polozheniya [GOST R IEC 60287-1-1. Electrical cables. Calculation of the rated current load. Part 1-1. Equations for calculating the rated current load and calculating losses. Generalities]. Moscow: Standartinform, 2009, 28 p.
6. GOST R MEK 60287-1-2. Kabeli elektricheskie. Raschet nominal'noy tokovoy nagruzki. Ch. 1-2. Uravneniya dlya rascheta nominal'noy tokovoy nagruzki i raschet poter'. Koeffitsienty poter', obuslovlennyy vikhrevymi tokami v obolochke, dlya dvukh tsepey, raspolozhennykh v odnoy ploskosti [GOST R IEC 60287-1-2. Electrical cables. Calculation of the rated current load. Part 1-2. Equations for calculating the rated current load and calculating losses. Coefficients of losses due to eddy currents in the shell for two circuits located in the same plane]. Moscow: Standartinform, 2009, 23 p.
7. GOST R MEK 60287-2-2. Kabeli elektricheskie. Raschet nominal'noy tokovoy nagruzki. Ch. 2-2. Teplovoe soprotivlenie. Metod rascheta koeffitsientov snizheniya maksimal'no dopustimoy tokovoy nagruzki dlya grupp kabeley, prolozhennykh na vozdukh i zashchishchennykh ot pryamogo solnechnogo izlucheniya [GOST R IEC 60287-2-2. Electrical cables. Calculation of the rated current load. Part 2-2. Thermal resistance. Method for calculating the reduction coefficients of the maximum permissible current load for groups of cables laid in the air and protected from direct solar radiation]. Moscow: Standartinform, 2009, 12 p.
8. *Larina E.T. [i dr.]*. Raschet perekhodnykh teplovykh rezhimov odnozhil'nogo silovogo kabelya s plastmassovoy izolyatsiey, prolozhennogo v vozdukh [Calculation of transient thermal modes of a single-core power cable with plastic insulation laid in the air], *Elektrotehnika* [Electrical Engineering], 1991, No. 10, pp. 39-42.
9. *Titkov, V.V.* K otsenke teplovogo rezhima trekhfaznoy linii iz SPE-kabelya [On the estimation of the thermal regime of a three-phase line from a SPE cable], *Kabel'news* [Cable-news], 2009, No. 10, pp. 47-51.
10. *Neher J.M., McGrath M.H.* The calculation of the temperature rise and load capability of cable systems, *Philadelphia electric company*, October 1957, pp. 752-763.
11. *Leon F.* Calculation of underground cable ampacity, *CYME international T&D Inc.*, 2005.
12. *Melamed M.A.* Sovremennyye metody analiza i prognozirovaniya rezhimov elektropotrebleniya v elektroenergeticheskikh sistemakh [Modern methods of analysis and forecasting of power consumption in electric power systems], *Itogi nauki i tekhniki. Seriya «Energeticheskie sistemy i ikh avtomatizatsiya»* [Itogi Nauki i Tekhniki. Series Energy systems and their automation], 1988, Vol. 4, pp. 4-111.
13. *Gross G. and Galiana F.D.* Short term load forecasting, *Proc. IEEE*, 1987, Vol. 75, No. 12, pp. 1558-1573.
14. *Chen S.T., David C.Y., Moghaddamjo A.R.* Weather sensitive short-term load forecasting using non fully connected artificial neural network, *IEEE Trans. on Power Systems*, 1992, Vol. 7, No. 3, pp. 1098-1105.
15. *Hsy Y., Ho K.* Fuzzy expert systems: An application to short term load forecasting, *IEE Proceedings – C*, 1992, Vol. 139, No. 6, pp. 471-477.
16. *Lee K.Y., Park J.H.* Short-term load forecasting using an artificial neural network, *IEEE Trans. on Power Systems*, 1992, Vol. 7, No. 1, pp. 124-130.
17. *Meldorf M., Kilter J., Pajo R.* Comprehensive Modelling of Load, *CIGRE Regonal Meeting, June 18-20, 2007, Tallinn, Estonia*, pp. 145-150.
18. *Shumilova G.P., Gotman N.E., Startseva T.B.* Prognozirovaniye elektricheskikh nagruzok EES s ispol'zovaniem metodov iskusstvennogo intellekta [Prediction of electric loads of EPS using artificial intelligence methods], *Sb. trudov Rossiyskogo natsional'nogo simpoziuma po energetike. Kazan' 10–14 sent. 2001 g.* [Proceedings of the Russian National Symposium on Energy, Kazan, September 10-14. 2001]. Moscow: KGEU, 2001, pp. 103-106.

19. *Khaykin S.* Neyronnye seti: polnyy kurs [Neural networks: full course]. 2nd ed. Moscow: Izd. dom «Vil'yams», 2006, 1104 p.
20. *Rutkovskaya D., Pilin'skiy M., Rutkovskiy L.* Neyronnye seti, geneticheskie algoritmy i nechetkie sistemy [Neural networks, genetic algorithms and fuzzy systems]: trans. from polish I.D. Rudinskogo. Moscow: Goryachaya liniya – Telekom, 2008, 452 p.
21. *Krichevskiy M.L.* Intellectual'nyy analiz v menedzhmente [Intelligent analysis in management]. Saint Petersburg: SPbGUAP, 2005, 48 p.
22. *Sidorov S.G. i dr.* Mnogoprotsessornaya realizatsiya neyrosetevogo algoritma prognozirovaniya izmeneniya kharakteristik elektricheskoy izolyatsii [Multiprocessor implementation of a neural network algorithm for predicting changes in the characteristics of electrical insulation], *Vestnik IGEU* [Bulletin of the ISEU], 2011, Issue 1.
23. *Poluyanovich N.K., Dubyago M.N.* Prognozirovaniye resursa kabel'nykh liniy s ispol'zovaniem metoda iskusstvennykh neyronnykh setey [Prediction of the resource of cable lines using the method of artificial neural networks], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2019, No. 3 (205), pp. 51-62.
24. *Galushka V.V., Fathi V.A.* Formation of the training sample by using artificial neural networks in search problems database errors, *Inzhenernyy vestnik Dona (Rus.)*, 2013, No. 2. Available at: ivdon.ru/ru/magazine/archive/n2y2013/1597/.
25. *Dubyago M.N., Poluyanovich N.K.* Metod otsenki i prognozirovaniya ostatochnogo resursa izolyatsii kabel'nykh liniy [A method for assessing and predicting the residual life of insulation of cable lines], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2019, No. 3, pp. 132-143.

Статью рекомендовал к опубликованию д.т.н., профессор В.И. Финаев.

Полюянович Николай Константинович – Южный федеральный университет; e-mail: nik1-58@mail.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 89185693365; кафедра электротехники и мехатроники.

Азаров Николай Васильевич – e-mail: s.t.a.l.k.e.r999@yandex.ru; тел.: 89281758225; кафедра электротехники и мехатроники; аспирант.

Огреничев Александр Васильевич – e-mail: ogrenicheval@mail.ru; тел.: 89281758225; кафедра электротехники и мехатроники; аспирант.

Дубяго Марина Николаевна – e-mail: w_m88@mail.ru; тел.: 89281758225; кафедра электротехники и мехатроники; аспирант.

Poluyanovich Nikolay Konstantinovich – Southern Federal University; e-mail: nik1-58@mail.ru; 44, Nekrasovsky, Taganrog, 347928, Russia; phone: +79185693365; the department of electric technics and mechatronics.

Azarov Nikolay Vasilyevich – e-mail: s.t.a.l.k.e.r999@yandex.ru; phone: +79281758225; the department of electrical engineering and mechatronics; graduate student.

Ogrenichev Alexander Vasilyevich – e-mail: ogrenicheval@mail.ru; phone: +79281758225; the department of electrical engineering and mechatronics; graduate student.

Dubyago Marina Nikolaevna – e-mail: w_m88@mail.ru; phone: +79281758225; the department of electrical engineering and mechatronics; graduate student.

Н.К. Полуянович, М.Н. Дубяго

**АНАЛИЗ И ВЫБОР МЕТОДИКИ В РЕШЕНИИ ЗАДАЧ
ИНТЕЛЛЕКТУАЛИЗАЦИИ СИСТЕМ ПРОГНОЗИРОВАНИЯ
ТЕРМОФЛУКТУАЦИОННЫХ ПРОЦЕССОВ В КАБЕЛЬНЫХ СЕТЯХ**

Статья посвящена исследованиям вопросам создания диагностики и прогнозирования термофлуктуационных процессов изоляционных материалов силовых кабельных линий (СКЛ) электроэнергетических систем на основе таких методов искусственного интеллекта, как нейронные сети и нечеткая логика. Показана необходимость разработки более совершенной методики анализа тепловых режимов в СКЛ. Обоснована актуальность задачи создания нейросетей (НС) для оценки пропускной способности, расчёта и прогнозирования температуры жил СКЛ в режиме реального времени на основе данных системы температурного мониторинга, с учетом изменения токовой нагрузки линии и внешних условий теплоотвода. По основным критериям проведено сравнение традиционных и нейросетевых алгоритмов для прогнозирования, показало преимущество НС методов. Проведена классификация НС методов и моделей прогнозирования температурных режимов КЛ. Для решения задачи прогнозирования ресурса СКЛ была выбрана сеть с прямым распространением данных и обратного распространения ошибки, т.к. сети такого типа в совокупности с активационной функцией в виде гиперболического тангенса являются в некоторой степени универсальной структурой для многих задач аппроксимации, приближения и прогнозирования. Разработана нейросеть для определения температурного режима токоведущей жилы силового кабеля. Проведен сравнительный анализ экспериментальных и расчетных характеристик распределений температуры, при этом исследовались различные нагрузочные режимы работы и функции изменения тока кабеля. При анализе данных было определено, что максимальное отклонение данных, полученных от нейросети от данных обучающей выборки, составило менее 2,5 % что является вполне приемлемым результатом. Для повышения точности необходимо использовалось большое количество входных и выходных данных при обучении сети, а также некоторая доработка ее структуры. Модель позволяет оценивать текущее состояние изоляции и прогнозировать остаточный ресурс СКЛ. Модель может быть использована в устройствах и системах непрерывного диагностирования силовых кабелей по температурным режимам.

Искусственный интеллект; нейронные сети; термофлуктуационные процессы; изоляционные материалы; прогнозирование; надежность систем энергоснабжения.

N.K. Poluyanovich, M.N. Dubyago

**ANALYSIS AND SELECTION OF METHODOLOGIES IN THE SOLUTION
OF THE PROBLEMS OF INTELLECTUALIZATION IN SYSTEMS
FOR PROGNOSIS OF THERMOFLUCTUATION PROCESSES IN CABLE
NETWORKS**

The article is devoted to research on the creation of diagnostics and prediction of thermofluctuation processes of insulating materials of power cable lines (PCL) of electric power systems based on such methods of artificial intelligence as neural networks and fuzzy logic. The necessity of developing a better methodology for the analysis of thermal conditions in PCL is shown. The urgency of the task of creating neural networks (NS) for assessing the throughput, calculating and predicting the temperature of PCL conductors in real time based on the data of the temperature monitoring system, taking into account changes in the current load of the line and the external conditions of the heat sink, is substantiated. Based on the main criteria, traditional and neural network algorithms for forecasting are compared, and the advantage of NS methods is shown. The classification of NS methods and models for predicting the temperature conditions of cosmic rays has been carried out. To solve the problem of forecasting the PCL resource, a network was selected with direct data distribution and back propagation of the error, because net-

works of this type, together with an activation function in the form of a hyperbolic tangent, are to some extent a universal structure for many problems of approximation, approximation, and forecasting. A neural network has been developed to determine the temperature regime of a current-carrying core of a power cable. A comparative analysis of the experimental and calculated characteristics of the temperature distributions was carried out, while various load modes and the functions of changing the cable current were investigated. When analyzing the data, it was determined that the maximum deviation of the data received from the neural network from the data of the training sample was less than 2.5 %, which is an acceptable result. To increase the accuracy, a large amount of input and output data was used when training the network, as well as some refinement of its structure. The model allows you to evaluate the current state of isolation and predict the residual life of PCL. The model can be used in devices and systems for continuous diagnosis of power cables by temperature conditions.

Artificial intelligence; neural networks; thermal fluctuation processes; insulation materials; forecasting; reliability of power supply systems.

Введение и постановка задачи. Устойчивое функционирование электросетевого комплекса невозможно без надежной и качественной работы распределительных электрических сетей 6–10/0,4 кВ, которые являются завершающим звеном в системе обеспечения потребителей электрической энергией и находятся в непосредственном взаимодействии с конкретным потребителем [1]. К настоящему моменту значительная часть работающих силовых кабельных линий (СКЛ) выработала свой нормативный ресурс, однако возможности по его замене весьма ограничены. Результаты эксплуатации СКЛ свидетельствуют о том, что основной причиной его отказов является снижение электрической прочности изоляции. В процессе эксплуатации изоляция СКЛ подвергаются тепловым, электрическим, химическим, механическим и другим видам воздействий, в результате чего происходит изменение её электрических свойств и, как следствие, некоторые изменения претерпевают и технические характеристики изоляционных материалов [2]. В этих условиях важно разработать долговременную программу по обновлению СКЛ. Реализация такой программы требует разработки новых методов оценки и поддержания технического состояния действующего оборудования кабельных систем на заданном уровне и продления его работоспособности на максимально возможный срок.

Важной характеристикой СКЛ систем электроснабжения является нагрузочная способность, определяемая тепловым режимом, по которой рассчитывается площадь поперечного сечения и их пропускная способность [3]. Расчёт пропускной способности необходим и осуществляется при проектировании СКЛ. Однако использование даже самых точных методик не позволяет учесть на практике все нюансы, которые весьма существенно влияют на температурный режим работы СКЛ. В связи с этим, при проектировании СКЛ стремятся выбирать коэффициенты с определённым запасом, поэтому в ряде случаев кабели оказываются недогруженными, а иногда работают на пределе температурного режима [3]. Одним из решений указанной проблемы является мониторинг температуры кабеля во время работы.

Текущий анализ диагностических систем подтверждает, что имеется объективная научно-техническая проблема создания комплексных систем диагностирования, построенных на разносторонних принципах, дающих высокий уровень достоверности постановки диагноза и прогнозирования технического состояния СКЛ. Перспективным направлением является создание инфраструктуры диагностирования неисправностей СКЛ, основанной на использовании искусственных нейронных сетей. Искусственные нейронные сети (ИНС) в задачах диагностирования и прогнозирования остаточного ресурса изоляционных материалов могут служить в роли подсистемы выборки и принятия решений, а затем отдавать полученные данные другим подсистемам управления.

Применение интеллектуальной системы управления приведет к повышению общей надежности энергосети и повышению энергоэффективности функционирования энергетического сектора, уменьшению влияния человеческого фактора и снижению аварийных ситуаций.

Задачей исследования являются разработка интеллектуальной системы прогнозирования температуры жилы СКЛ для планирования режимов работы электросети с целью повышения надежности и энергоэффективности их взаимодействия с объединенной энергосистемой;

Целью исследования является повышение надежности и энергоэффективности распределенных энергосистем за счёт применения новых неразрушающих методов и алгоритмов мониторинга и диагностики ресурса СКЛ на базе искусственных нейронных сетей для краткосрочного и среднесрочного прогнозирования температурных режимов токоведущей жилы силовых кабельных линий.

Преимущества СКЛ в системах электроснабжения. Несмотря на то, что кабельные линии широко используются уже на протяжении половины века, только сейчас современные технологии проектирования и производства позволяют стать им эффективной альтернативой воздушных линий электропередач:

1. Подземные кабели обладают уникальными свойствами по передаче энергии – они не излучают электрических полей, имеют улучшенные характеристики по потери мощности, высокую стойкость при аварийных нагрузках. СКЛ можно использовать в местах плотной застройки, реках и сложных геологических условиях, местах, где требуется сохранение окружающей среды, ландшафтов, значимых строений и т.п.

2. Энергосбережение. Себестоимость кабельных сетей приближается к стоимости воздушных линий электропередач. Кабельные сети менее восприимчивы к тяжелым погодным условиям (штормы, землетрясения, и т.д), что сокращает стоимость обслуживания по сравнению с воздушными линиями. СКЛ содержат большое количество меди, наиболее токопроводящего металла, в результате чего на 30% снижаются потери при высоких нагрузках по сравнению с ВЭЛ, а следовательно повышается рентабельность всей энергосистемы. СКЛ используют в качестве проводника более эффективные медные сплавы, которые работают при более низких температурах.

3. Повышенная надежность. Современные кабельные сети используют поперечно сшитый полиэтилен (XLPE) в качестве основного изоляционного материала, который уже 20 лет подтверждает свою высокую надежность.

4. Мониторинг состояния кабеля (мониторинг температуры фаз кабеля по всей длине) обеспечивает надежную работу кабельной системы. Система помогает эксплуатационным организациям в решении задач: снижения количества перебоев в электроснабжении или системных аварий; оперативного реагирования на возникающие перегрузки; увеличение нагрузки без превышения допустимой температуры; прогнозирования срока эксплуатации и др.

Задачи диагностики и прогнозирования термофлуктуационных процессов СКЛ. Различаются методики по расчету стационарных и нестационарных тепловых процессов в кабеле, методики без возможности анализа тепловых полей и с возможностью анализа тепловых полей, методики на основе эмпирических зависимостей и на основе дифференциальных уравнений.

Нормативным документом, устанавливающим расчет тепловых процессов в силовых кабелях, является МЭК 60287 и его российский аналог ГОСТ [5–7]. Методика расчета номинальной токовой нагрузки СКЛ, исходя из допустимой температуры жилы рассматривает в ГОСТ 60287-1-1 [5]. Методика корректировки соотношений из ГОСТ 60287-2-1 с учетом влияния соседних кабелей приводится в

ГОСТ 60287-2-2 [7]. Тепловые процессы в одножильном одиночном кабеле, проложенном в воздухе, и соответствующие краевые условия описываются в [8] системой квазилинейных дифференциальных уравнений параболического типа. Методика расчета тепловых полей СКЛ на основе решения связанной задачи электродинамики и теплопередачи для СПЭ кабелей с помощью уравнения плоскопараллельного квазистационарного поля, рассматривается в [9]. Даны расчетные картины тепловых полей с учетом токов в экране кабеля. Расчет стационарных тепловых процессов в кабеле подробно рассмотрен [10], а приведенные соотношения показывают, что рассмотренные выше стандарты МЭК созданы на их основе. Методика расчета тепловых сопротивлений для кабелей, проложенных под землей рассматривается в [11], а приведенные соотношения аналогичны с рекомендуемыми МЭК. Стационарные тепловые процессы в СКЛ, рассматриваемые МЭК и ГОСТ, больше освещены в литературе чем нестационарные режимы, что связано со сложностью получения достаточно адекватных соотношений и алгоритмов на их основе. Многофакторность задачи обусловила тот факт, что большинство традиционных методик расчета базируется на эмпирических зависимостях и коэффициентах, что ограничивает их применимость диапазоном воздействующих факторов и условиями окружающей среды. Даже в случае сложной системы дифференциальных уравнений уйти от эмпирических коэффициентов очень сложно. С точки зрения теплотехники силовой кабель не является сложной системой, однако факторы окружающей среды и электромагнитные процессы усложняют задачу анализа тепловых режимов на несколько порядков. Способ прокладки кабеля, изменение параметров почвы, способы заземления экранов, взаимоиנדукция между жилами, краевой и поверхностный эффекты, электрические параметры изоляции – это и многое другое влияет на тепловые процессы в кабеле. Стремление существующих методик расчета учесть все эти факторы приводит к значительному усложнению алгоритмов и снижению гарантии их адекватности при изменении внешних условий.

Таким образом, существует необходимость разработки более совершенной методики анализа тепловых режимов СКЛ, максимально инвариантной к внешним условиям. Поставим задачу разработать математическую модель тепловых процессов силового кабеля, не использующую эмпирические коэффициенты и зависимости, но позволяющую вести расчет температуры наиболее нагретой точки изоляции кабеля с достаточной точностью.

1. Разработка тепловой схемы замещения СКЛ. Тепловые процессы. В общем случае допустимый длительный ток кабеля зависит от ряда факторов [18].

$$I_{\text{доп}} = f(n, \theta_{\text{ж}}, \theta_{\text{с}}, \tau, R, P_{\text{и}}), \quad (1)$$

где n – количество жил в кабеле, шт; $\theta_{\text{ж}}$ – допустимая температура нагрева жилы кабеля при продолжительной нагрузке, °C, $\theta_{\text{с}}$ – температура окружающей среды, °C, S – суммарное тепловое сопротивление кабеля и окружающей среды, град см/Вт; R – электрическое сопротивление токопроводящей жилы кабеля, Ом; $P_{\text{и}}$ – мощность, рассеиваемая в изоляции кабеля за счет диэлектрических потерь, Вт.

Уменьшение n , $\theta_{\text{с}}$, R , S , R и $P_{\text{и}}$, а также увеличение $\theta_{\text{ж}}$ способствует повышению $I_{\text{доп}}$. При одинаковом числе жил из одного и того же проводникового материала величина $I_{\text{доп}}$ главным образом зависит от значений $\theta_{\text{ж}}$, S и R . Следовательно, кабели напряжением 6–10 кВ с изоляцией из СПЭ, имеющие $\theta_{\text{ж}} = 90$ °C, при идентичных условиях прокладки допускают большие токи нагрузки по сравнению с кабелями с бумажной пропитанной изоляцией, у которых $\theta_{\text{ж}} = 65$ °C при напряжении 6 кВ и $\theta_{\text{ж}} = 60$ °C при напряжении 10 кВ. При прокладке в воздухе величина S складывается из тепловых сопротивлений

кабеля и его поверхности, а при прокладке в земле – из тепловых сопротивлений кабеля, защитных покровов и почвы, окружающей кабель. Тепловое сопротивление кабеля зависит в основном от его геометрических размеров и конструктивных особенностей, а также от удельного теплового сопротивления материалов, из которых изготовлен кабель. Удельное тепловое сопротивление полиэтилена небольшое (300–400 град см/Вт), что обеспечивает хорошую теплопроводность полимерной изоляции и оболочки кабеля и благоприятно сказывается на его пропускной способности. Отметим, что с ростом температуры нагрева жилы имеет место некоторое снижение удельного теплового сопротивления изоляционного материала. Это способствует увеличению допустимого тока кабеля.

Длительно допустимые токовые нагрузки проводников могут определяться на основе теплового расчета или экспериментально. Для изолированных проводников расчетные формулы получаются сложными, требующими большого количества исходных данных, которые часто носят неопределенный характер. Поэтому в справочной информации по кабельной продукции приводятся допустимые токи кабелей, полученные на основе расчета и уточненные экспериментально с учетом применяемой изоляции, площади сечения жил, условий прокладки и допустимых температур нагрева жил и окружающей среды. Особенностью одножильных кабелей с изоляцией из СПЭ напряжением 6–10 кВ является то, что они имеют большую пропускную способность при прокладке в воздухе, а не в земле.

Ресурс изоляционных материалов зависит от таких параметров как температура, ток, влажность [3], поэтому необходимо исследования температурного поля в сечении СКЛ [12, 13]. Низкая теплопроводность изоляции кабеля рис. 1, приводит к высокому температурному градиенту, поэтому температура наиболее нагретого участка в сечении кабеля (вблизи жилы) значительно отличается от измеренной температуры, а в переходных режимах эта разница может увеличиваться в несколько раз. Поэтому актуальна задача создания нейросетей для оценки пропускной способности, расчёта и прогнозирования температуры жил СКЛ в режиме реального времени на основе данных системы температурного мониторинга, с учетом изменения токовой нагрузки линии и внешних условий теплоотвода.

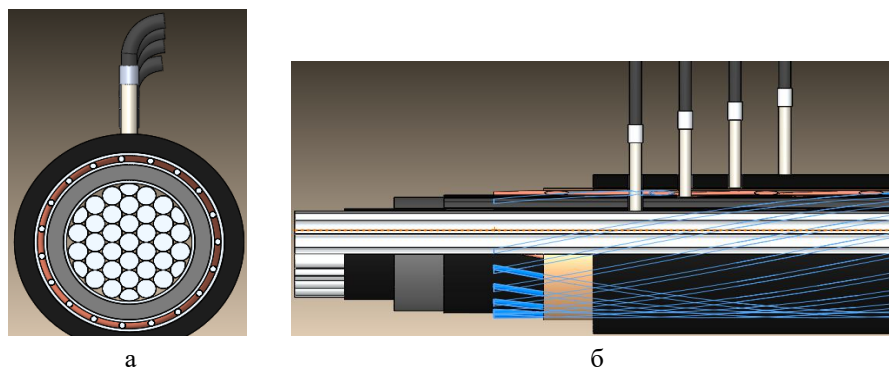


Рис. 1. Сечение кабеля и его эквивалентная тепловая схема замещения

Помимо оценки пропускной способности к задачам мониторинга теплового режима относят оценку температуры жилы в режиме реального времени и своевременное предотвращение токовых перегрузок СКЛ, требуется прогнозирование

температур жил кабелей в режиме реального времени. Последнее означает, что необходимо постоянно прогнозировать нагрев кабелей на некоторое время вперёд, причем такое, чтобы была возможность принять меры по снижению нагрузки КЛ. Нейронные сети не программируются в привычном смысле этого слова, они обучаются [12]. Возможность обучения – одно из главных преимуществ нейронных сетей перед традиционными алгоритмами.

Поэтому в работе рассматривается задача с помощью искусственных нейронных сетей (ИНС) рассчитывать процесс передачи тепла через изоляционные слои (рис.1), при сложном изменении температуры жилы. Причина выбора обусловлена необходимостью прогнозирования нелинейных изменений температуры.

2. Интеллектуальная система мониторинга силовых кабельных сетей.

Пропускная способность кабельных линий ограничивается допустимой температурой нагрева жилы. На рабочие показатели кабельной системы оказывают такие факторы как условия прокладки (тепловое сопротивление грунта, температура грунта, расстояние до соседних кабелей, источников тепла и т.д.). Ключевым элементом современных систем распределенного измерения температуры кабеля стало оптическое волокно.

Интеллектуальная система мониторинга кабельных сетей основана на современной технологии распределённого измерения температуры оптоволоконна (Distributed Fiber-Optic Sensing) [15]. Данная технология позволяет проводить высокоточные измерения температуры по всей длине высоковольтного кабеля [16] в реальном времени более чем в 40000 точках с разрешающей способностью 1 метр с помощью многомодового оптического волокна вмонтированного в кабель [17]. Система позволяет: получать данные о температурном профиле высоковольтного кабеля в реальном времени для постоянного мониторинга состояния кабеля и раннего выявления пожаров и аварийных ситуаций; использовать температурные данные для динамического управления нагрузкой высоковольтной сети; точно (до 50 см) определять места локальных перегревов и выдавать аварийные сообщения на пульт диспетчерской с указанием расположения аварийного участка. Краткое описание распределенных методов волоконно-оптического мониторинга приведено в табл. 1.

Мониторинг температуры КЛ посредством оптоволоконна является на сегодняшний день единственным способом диагностики КЛ на всей ее протяженности. На разных участках трассы при протекании одного и того же тока, температура токопроводящей жилы может отличаться. Ключевым элементом современных систем распределенного измерения температуры кабеля стало оптическое волокно. Информация о состоянии КЛ собранная за определенные периоды времени (сезоны) позволяет более правильно и экономично эксплуатировать КЛ, а также минимизировать возможные аварии в распределительных сетях. Пропускная способность кабеля ограничивается допустимой температурой нагрева жилы. Существует два способа применения оптоволоконна в качестве датчика измерения распределенной температуры КЛ. Первый – металлический модуль с оптоволоконнами находится в экране кабеля (интегрируется вместо одной из проволок экрана в процессе производства), второй – оптоволоконный кабель, с наименьшей толщиной изолирующего слоя, прокладывается вдоль кабельной трассы максимально близко к оболочке силового кабеля. Согласно СТО 56947007-29.060.20.170-2014, внутренние дефекты элементов КЛ (участков КЛ, концевых и соединительных муфт) выявляются по превышению измеренной температуры контролируемого элемента (зоны аномального нагрева поверхности) над температурой аналогичных элементов других фаз, находящихся в одинаковых условиях на 1 °C и более.

Таблица 1

Методы волоконно-оптического мониторинга

Метод мониторинга	Описание метода
DAS (Distributed Acoustic Sensing) – распределенный акустический мониторинг	Основывается на рассеянии Рэлея при использовании одномодового оптоволокна. Акустические колебания вызывают небольшие изменения показателя преломления, улавливаемые с помощью этого рассеяния.
DTS (Distributed Temperature Sensing) – распределенный температурный мониторинг	Основывается на использовании двух компонент рассеяния Рамана, смещенных по частоте относительно исходного сигнала. Амплитуда антистоксовой компоненты сильно зависит от температуры волокна, а стоксовая нет. Две эти частоты выделяются из отраженного сигнала, после из отношения их спектральных интенсивностей, определяется температура в точке рассеяния.
DSS (Distributed Strain Sensing) – распределенный мониторинг напряжений (механических)	Реализуется только на одномодовом волокне при использовании рассеяния Менделъштама – Бриллюэна. При изменении механического напряжения волокна происходит сдвиг частоты Бриллюэна, который может быть зафиксирован.

Таким образом, для эксплуатации кабельной линии необходимо знать температуры жилы кабеля, в то время, как измерение температуры производится в металлическом экране кабеля или на поверхности оболочки кабеля. Зная условия прокладки для различных участков кабельной линии можно рассчитать температуру проводника в зависимости от распределенной температуры экрана кабеля [3]. Сбор и анализ данных по температуре по всей длине кабельной линии дает полную картину процессов, происходящих в линии, что позволяет более рационально использовать кабельные электросети при разных условиях и режимах работы. Для мониторинга тепловых процессов в СКЛ оптимальным является метод, основанный на рассеянии Рамана.

2.1. Методы, основанные на моделях нейронных сетей. Многие нелинейные задачи не поддаются строгой формализации традиционными математическими методами [14]. В таких условиях, когда решение задачи аналитически в общем виде невозможно, оправдан нейросетевой подход, позволяющий обеспечить достаточно высокое качество выполнения задачи. В ряде работ касающихся аппроксимации нелинейных функций заложен математический базис нейросетевой теории, определяющий универсальные аппроксимирующие свойства нейронных сетей. В большинстве случаев аналитические модели диагностирования это нелинейные соотношения, затрудняющие формирование модели объекта диагностирования по моделям составляющих компонент. Формально высказывание об универсальных аппроксимационных свойствах нелинейности представляется в виде:

$$F(x_1, x_2, \dots, x_n) = \sum_q^{2n+1} h_p \left[\sum_p^n \Phi_q^p(x_p) \right].$$

Оно утверждает, что с помощью линейных операций и каскадного соединения можно из произвольных нелинейных элементов получить любой требуемый результат с заранее заданной точностью. Здесь h_p – непрерывная функция; $\Phi_q^p(x_p)$ – функция, зависящая от x .

Для решения задач аппроксимации нелинейностей важны методики разрешающие проблемы принятия решений в условиях неполных данных с учетом постоянно изменяющихся условий окружающей среды. Этим методикам отвечают возможности нейро - технологий. Нейронные сети не требуют традиционного программирования: информация обучения НС накапливается в весах, а не в программах. Это делает их устойчивыми к флуктуациям входных воздействий и обеспечивает устойчивость работоспособности сети при выходе из строя отдельных ее компонент. На возникший дефект сеть реагирует только изменением качества функционирования при сохранении общей работоспособности (это роднит ее с высокоорганизованными биологическими системами).

Изначальными процедурами построения однонаправленной сети являются задание топологии и задание правил обучения. Топология выбирается исходя из требуемой точности идентификации, содержания задачи, количества параметров процесса, размерности вектора входных данных. Настройка сети это многоходовой итерационный процесс, при котором периодически анализируются результаты и регулируются параметры: количество ассоциативных слоев, количество нейронов в слое, выбор функции активации.

Увеличение количества слоев позволяет выявить более тонкие статистические закономерности. Но размерность сети должна соответствовать размерности данных обучающей выборки. В противном случае, способность сети к обучению будет снижаться, или наоборот, будет утрачена способность сети определять основные параметры отображения.

Обзор НС методов прогнозирования. В 1991 году, Park D.C. и др. разработали нейронную сеть для моделирования и прогнозирования нагрузки [19–21]. Нейронные сети, по существу это нелинейные функции, которые имеют способность выполнения аппроксимации нелинейных зависимостей. Выходы искусственной нейронной сети являются линейными или нелинейными математическими функциями ее входов. Преимуществом технологии нейронных сетей является интеллектуальная обработка. Недостатком является то, что процесс обучения является относительно медленным, и это не гарантирует сходимости между фактическими и прогнозными данными. Кроме этого, определение оптимального набора входных переменных и размеры скрытых слоев следует исследовать на практике.

При применении нейронной сети для целей прогнозирования, необходимо решить следующие задачи: выбрать архитектуру нейронной сети; определить количество и соединение слоев и элементов; принять решение по использованию двунаправленной или однонаправленной связи и цифрового формата [22]. Самой популярной архитектурой искусственной нейронной сети для прогнозирования электрической нагрузки является архитектура обратного распространения. Эта сеть использует непрерывно оцениваемые функции и обучение с учителем. Фактические числовые данные весов, присваиваемые входам элементов определяются путем сопоставления архивных данных (например, время и погода) с заданными выходами (например, прогнозные данные нагрузки) во время обучения. Искусственные нейронные сети с неконтролируемым обучением не требуют предварительной оперативной подготовки.

Paralekhoroulos и др. в [23] также разработали и внедрили многоуровневую прямолинейную трехслойную ИНС для краткосрочного прогнозирования нагрузки. В этой модели используются три типа переменных в качестве входных данных нейронной сети: входы связанные с временем года, входы связанные с погодой и с

историческими нагрузками. Khotanzad и др. [24] описали модель ИНС для прогнозирования нагрузки, которая основана на нескольких стратегиях, которые отражают различные тенденции данных. В основе прогнозной модели используется многослойный персептрон, обучаемый по алгоритму обратного распространения ошибки. Эта прогнозная модель может учитывать влияние температуры и относительной влажности на нагрузку, а также содержит функцию, которая может прогнозировать почасовую температуру и относительную влажность, необходимые для работы системы. Chen и др. в [25] также разработали трехслойную прямолинейную модель нейронной сети, в которой при обучении также использовался алгоритм обратного распространения. Их ИНС также рассматривает цену на электроэнергию в качестве одной из основных характеристик загрузки энергосистемы. Многие опубликованные исследования используют искусственные нейронные сети в сочетании с другими методами прогнозирования, такими как временные ряды [26] и нечеткая логика. Искусственная нейронная сеть не нуждается в человеческом опыте и направлена на создание связей между входным набором данных и наблюдаемыми выходами. Это является достоинством при работе с нелинейными зависимостями между нагрузкой и влияющими на неё факторами.

Таблица 2

Сравнение алгоритмов прогнозирования

Требования к специалисту	
Традиционный алгоритм	Нейронная сеть
– Для применения алгоритмов работающему специалисту требуется большее количество знаний в данной области. – Несовершенство традиционных современных методик анализа электроэнергии (ЭЭ) и потерь.	НС обучается на подготовленных специалистом примерах. От специалиста требуются знания о том, какие данные необходимы, в каком виде и количестве, как их подготовить для передачи в сети и как в итоге интерпретировать полученные результаты. – Альтернативный подход к анализу потерь ЭЭ.
Обучаемость	
Традиционный алгоритм	Нейронная сеть
Появление новых данных вызывает необходимость в их пересчете. Реальные процессы (I , $t^{\circ}C$, P_n и т.д.) не могут быть адекватно описаны традиционными стохастическими моделями, т.к. являются существенно нелинейными.	Благодаря системе обучения они адаптируются к различным изменениям, каждый раз применяя новые данные для поиска решения поставленной задачи.
Универсальность	
Традиционный алгоритм	Нейронная сеть
Многие традиционные алгоритмы предназначены для решения конкретной задачи	– нахождение решения при неопределенности, отсутствии некоторых закономерностей или входных данных; – не обязательно устанавливать взаимоотношения между величинами заранее (сеть изучает взаимосвязи на уже готовых моделях). – могут работать с данными различных типов (дискретные данные, непрерывные или номинальные).

Сравнение традиционных и нейросетевых алгоритмов для прогнозирования, по нескольким основным критериям [27], представленных в табл. 2, показало преимущество НС методов.

Можно выделить следующие методы прогнозирования: – экстраполяция (распространение тенденций и закономерностей, действующих в прошлом, на будущее); – моделирование (разработка математических моделей, которые на основе анализа ретроспективных данных могут оценивать наиболее вероятные значения в будущем); – экспертные оценки (суждения о будущем выносят эксперты предметной области на основе личных знаний и опыта. Классификация НС методов прогнозирования представлена на рис. 2



Рис. 2. Классификация НС методов прогнозирования

3. Экспериментальные исследования по обучающей выборке. Для составления модели прогнозирования на базе ИНС была выбрана система Mathlab. При прогнозирования температуры жилы кабеля создание, обучение и моделирование нейронной сети осуществлялось с помощью Neural Network Toolbox в системе Mathlab. Исследованы термофлуктуационные процессы в силовом кабеле (без повреждений) АПв Пу г-1х240/25-10 при изменении действующего значения тока жилы кабеля и разработанной модели прогноза с использованием искусственных нейронных сетей (ИНС) на глубоких ретроспективных данных температур исследуемого кабеля за 2015-2019 годы. Данные по температуре $\theta_{з.о.}$, $\theta_{о.с.}$ – защитной оболочки, окружающей среды получены в ходе работы с реальными образцами силового кабеля АПв Пу г-1х30/25-10, рис. 1. Для каждого значения тока жилы бралось по 400 точек температуры $\theta_{з.о.}$, $\theta_{о.с.}$.

В статье приведены результаты прогноза температуры жилы кабеля ($\theta_{ж}$), с помощью ИНС, для четырех исследуемых образцов кабеля. Построены графики экспериментальной температуры исследуемых образцов и графики построенные на основе данных обучающей выборки и нейросети (рис. 3).

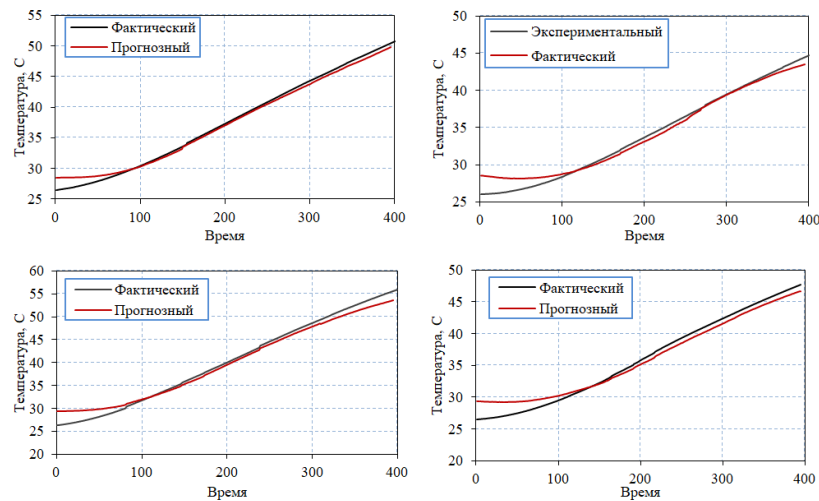


Рис. 3. График экспериментальной температуры образцов кабеля №1, 2, 3, 4 (фактический) и графики построенные на основе данных обучающей выборки нейросетью (прогнозный)

Результаты прогноза температуры жилы кабеля (θж) полученные на основе данных обучающей выборки нейросетью и графики экспериментально снятой температуры, одного из исследуемых образцов (№1), а также результаты анализа погрешностей приведены в табл. 3.

Таблица 3

Экспериментальные и спрогнозированные значения

Ток жилы, А	Средняя температура токоведущей жилы		Средняя величина ошибки прогноза	
	Экспериментальная, t°C	Спрогнозированная, t°C	t°, C	t°, %
440	34,22	34,47195	1,045001	3,031455
450	38,75	42,567	3,817004	8,967048
455	38	41,10783	3,107835	7,560201
460	39,155	42,80692	3,879702	9,063258
470	38,28	38,79928	2,442176	6,294385
475	35,03	36,72509	1,695086	4,615608
480	37,47	36,26801	1,201989	3,314185
485	36,125	36,16032	0,609145	1,684567
490	36,435	38,53838	2,103384	5,457894
495	32,785	37,63504	4,850041	12,88704
500	36,185	36,09096	1,1344	3,143169
510	42,59582	39,47206	3,155809	7,995045
515	37,095	37,93254	0,837542	2,207977
525	39,655	37,49795	3,234812	8,626637
530	37,375	37,97731	0,724587	1,907947
540	47,065	43,94306	3,121937	7,104504
580	45,845	48,68854	2,843538	5,840262
600	48,47	48,55757	1,442092	2,969861

Средняя ошибка прогноза температуры жилы кабеля $\theta_{ж}$ при различных токах жилы (табл. 1) не превышает 13 %, что свидетельствует о возможности применения метода искусственных нейронных сетей для целей прогнозирования температуры жилы кабеля по температуре на поверхности $\theta_{з.о.}$, рис. 1.

Заключение. Показана необходимость разработки более совершенной методики анализа тепловых режимов в СКЛ. По основным критериям проведено сравнение традиционных и нейросетевых алгоритмов для прогнозирования, которое показало преимущество НС методов. Разработана нейросеть для определения температурного режима токоведущей жилы силового кабеля. Исследованы термофлуктуационные процессы в силовом кабеле (без повреждений) АПв Пу г-1х240/25-10 при реальной диаграмме действующего значения тока жилы кабеля и разработанной модели прогноза с использованием искусственных нейронных сетей (ИНС). Приведены результаты экспериментальных исследований прогноза температуры жилы кабеля ($\theta_{ж}$), с помощью ИНС. При анализе данных было определено, что максимальное отклонение данных, полученных от нейросети от данных обучающей выборки, составило менее 2,5 % что является вполне приемлемым результатом. Модель позволяет оценивать текущее состояние изоляции и прогнозировать остаточный ресурс СКЛ. Основная область применения разработанной нейросети для определения температурного режима токоведущей жилы заключается в диагностике и прогнозировании ресурса ЭИ силового кабеля.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Положение ОАО «Россети» о единой технической политике в электросетевом комплексе. – М.: ОАО «Россети», 2013. – 196 с.
2. Анищенко В.А., Колосова И.В. Основы надежности систем электроснабжения. – Мн.: БНТУ, 2007.
3. Дубяго М.Н., Полуянович Н.К. Совершенствование методов диагностики и прогнозирования электроизоляционных материалов систем энергоснабжения. – Ростов-на-Дону; Таганрог: Изд-во ЮФУ, 2019. – 192 с.
4. Каталог оборудования компании LIOS Technology GmbH – 2010.
5. ГОСТ Р МЭК 60287-1-1. Кабели электрические. Расчет номинальной токовой нагрузки. Ч. 1-1. Уравнения для расчета номинальной токовой нагрузки и расчет потерь. Общие положения. – М.: Стандартинформ, 2009. – 28 с.
6. ГОСТ Р МЭК 60287-1-2. Кабели электрические. Расчет номинальной токовой нагрузки. Ч. 1-2. Уравнения для расчета номинальной токовой нагрузки и расчет потерь. Коэффициенты потерь, обусловленных вихревыми токами в оболочке, для двух цепей, расположенных в одной плоскости. – М.: Стандартинформ, 2009. – 23 с.
7. ГОСТ Р МЭК 60287-2-2. Кабели электрические. Расчет номинальной токовой нагрузки. Ч. 2-2. Тепловое сопротивление. Метод расчета коэффициентов снижения максимально допустимой токовой нагрузки для групп кабелей, проложенных на воздухе и защищенных от прямого солнечного излучения. – М.: Стандартинформ, 2009. – 12 с.
8. Ларина Э.Т. [и др.]. Расчет переходных тепловых режимов одножильного силового кабеля с пластмассовой изоляцией, проложенного в воздухе // Электротехника. – 1991. – № 10. – С. 39-42.
9. Титков В.В. К оценке теплового режима трехфазной линии из СПЭ-кабеля // Кабель-News. – 2009. – № 10. – С. 47-51.
10. Neher J.M., McGrath M.H. The calculation of the temperature rise and load capability of cable systems // Philadelphia electric company. – October 1957. – P. 752-763.
11. Leon F. Calculation of underground cable ampacity // CYME international T&D Inc., 2005.
12. Полуянович Н.К., Дубяго М.Н. Прогнозирование ресурса кабельных линий с использованием метода искусственных нейронных сетей // Известия ЮФУ. Технические науки. – 2019. – № 3. – С. 51-62.
13. Дубяго М.Н., Полуянович Н.К. Метод оценки и прогнозирования остаточного ресурса изоляции кабельных линий // Известия ЮФУ. Технические науки. – 2019. – № 3. – С. 132-143.

14. Горева Т.И., Порнягин Н.Н., Пюкке Г.А. Нейросетевые модели диагностики технических систем // Вестник КРАУНЦ. Физ.-мат. науки. – 2012. – № 1 (4). – С. 31-43.
15. Якунин А.В. Мониторинг теплового режима эксплуатации кабельных линий 110-500 кВ // Сб. докладов конференции «Линии электропередачи 2010: Проектирование, строительство, опыт эксплуатации и НТП». – 2010. – С. 306-310.
16. Ukil A., Braendle H., Krippner P. Distributed temperature sensing: review of technology and applications // Sensors Journal, IEEE. – 2012. – Vol. 12, No. 5. – P. 885-892.
17. Deng J., Xiao H., Huo W. and other. Optical fiber sensor-based detection of partial discharges in power transformers // Optics and Laser Technology. – 2001. – Vol. 33, No. 5. – P. 305-311.
18. Бачелис Д.С., Белорусов Н.И., Саакян А.Е. Электрические кабели, провода и шнуры: Справ. – М.: Энергия, 1971. – 704 с.
19. El-Sharkawi, Park D.C. Electric load forecasting using an artificial neural network // IEEE Transaction on Power Systems. – 1991. – No. 6 (2). – P. 442-449.
20. Peng M., Hubele N.F., Karady G.G. Advancement in the application of neural networks for short-term load forecasting // IEEE Transactions on Power Systems. – 1992. – Vol. 7. – P. 250-257.
21. Desouky A.A., Elkateb M.M. Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA // IEE Proceedings of Generation, Transmission and Distribution. – 2000. – Vol. 4. – P. 213-217.
22. Peng M., Hubele N.F., Karady G.G. Advancement in the application of neural networks for short-term load forecasting // IEEE Transactions on Power Systems. – 1992. – Vol. 7. – P. 250-257.
23. Papalexopoulos A.D., Hao S., Peng T.M. An implementation of a neural network based load forecasting model for the EMS // IEEE Transactions on Power Systems. – 1994. – Vol. 9. – P. 1956-1962.
24. Khotanzad A. ANNSTLF - A neural-network-based electric load forecasting system // IEEE Transactions on Neural Networks. – 1997. – Vol. 8. – P. 835-846.
25. Chen C.S., Tzeng Y.M., Hwang J.C. The application of artificial neural networks to substation load forecasting // Electric Power Systems Research. – 1996. – Vol. 38, Issue 2. – P. 153-160.
26. Chow T.W., Leung C.T. Nonlinear autoregressive integrated neural network model for short-term load forecasting // IEE Proceedings, Generation, Transmission and Distribution. – 1996. – Vol. 143. – P. 500-506.
27. Курбацкий В.Г., Томин Н.В. Практика использования новых информационных технологий для прогнозирования и анализа отдельных характеристик сетевых энергопредприятий // Проблемы энергетики. – 2006. – № 3-4.

REFERENCE

1. Polozhenie OAO «Rosseti» o edinoy tekhnicheskoy politike v elektrosetevom komplekse [Regulation of JSC "rosseti" on the unified technical policy in the electric grid complex]. Moscow: OAO «Rosseti», 2013, 196 p.
2. Anishchenko V.A., Kolosova I.V. Osnovy nadezhnosti sistem elektrosnabzheniya [Fundamentals of reliability of power supply systems]. Minsk.: BNTU, 2007.
3. Dubyago M.N., Poluyanovich N.K. Sovershenstvovanie metodov diagnostiki i prognozirovaniya elektroizolyatsionnykh materialov sistem energosnabzheniya [Improving methods of diagnostics and forecasting of electrical insulation materials of power supply systems]. Rostov-on-Don; Taganrog: Izd-vo YuFU, 2019, 192 p.
4. Katalog oborudovaniya kompanii LIOS Technology GmbH – 2010 [Directory of equipment of the company lios Technology GmbH – 2010].
5. GOST R MEK 60287-1-1. Kabeli elektricheskie. Raschet nominal'noy tokovoy nagruzki. Ch. 1-1. Uravneniya dlya rascheta nominal'noy tokovoy nagruzki i raschet poter'. Obshchie polozheniya [GOST R IEC 60287-1-1. Electrical cables. Calculation of the rated current load. Part 1-1. Equations for calculating the rated current load and calculating losses. Generalities]. Moscow: Standartinform, 2009, 28 p.
6. GOST R MEK 60287-1-2. Kabeli elektricheskie. Raschet nominal'noy tokovoy nagruzki. Ch. 1-2. Uravneniya dlya rascheta nominal'noy tokovoy nagruzki i raschet poter'. Koeffitsienty poter', obuslovlennyykh vikhrevymi tokami v obolochke, dlya dvukh tsepey, raspolozhennykh v odnoy ploskosti [GOST R IEC 60287-1-2. Electrical cables. Calculation of the rated current load. Part 1-2. Equations for calculating the rated current load and calculating losses. Coefficients of losses due to eddy currents in the shell for two circuits located in the same plane]. Moscow: Standartinform, 2009, 23 p.

7. GOST R MEK 60287-2-2. Kabeli elektricheskie. Raschet nominal'noy tokovoy nagruzki. Ch. 2-2. Teplovoe soprotivlenie. Metod rascheta koeffitsientov snizheniya maksimal'no dopustimoy tokovoy nagruzki dlya grupp kabeley, prolozhennykh na vozdukh i zashchishchennykh ot pryamogo solnechnogo izlucheniya [GOST R IEC 60287-2-2. Electrical cables. Calculation of the rated current load. Part 2-2. Thermal resistance. Method for calculating the reduction coefficients of the maximum permissible current load for groups of cables laid in the air and protected from direct solar radiation]. Moscow: Standartinform, 2009, 12 p.
8. Larina E.T. [i dr.]. Raschet perekhodnykh teplovykh rezhimov odnozhil'nogo silovogo kabelya s plastmassovoy izolyatsiei, prolozhennogo v vozdukh [Calculation of transient thermal modes of a single-core power cable with plastic insulation laid in the air], *Elektrotehnika* [Electrical Engineering], 1991, No. 10, pp. 39-42.
9. Titkov V.V. K otsenke teplovogo rezhima trekhfaznoy linii iz SPE-kabelya [on the estimation of the thermal regime of a three-phase line from a SPE cable], *Kabel'news* [Cable-news], 2009, No. 10, pp. 47-51.
10. Neher J.M., McGrath M.H. The calculation of the temperature rise and load capability of cable systems, *Philadelphia electric company*, October 1957, pp. 752-763.
11. Leon F. Calculation of underground cable ampacity, *CYME international T&D Inc.*, 2005.
12. Poluyanovich N.K., Dubyago M.N. Prognozirovanie resursa kabel'nykh liniy s ispol'zovaniem metoda iskusstvennykh neyronnykh setey [Forecasting the resource of cable lines using the method of artificial neural networks], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2019, No. 3, pp. 51-62.
13. Dubyago M.N., Poluyanovich N.K. Metod otsenki i prognozirovaniya ostatochnogo resursa izolyatsii kabel'nykh liniy [Method for estimating and predicting the residual insulation life of cable lines] *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2019, No. 3, pp. 132-143.
14. Goreva T.I., Pornyagin N.N., Pyukke G.A. Neyrosetevye modeli diagnostiki tekhnicheskikh sistem [Neural network model for diagnosis of technical systems], *Vestnik KRAUNTS. Fiz.-mat. nauki* [Vestnik KRAUNTS. Phys. -Mat. sciences], 2012, No. 1 (4), pp. 31-43.
15. Yakunin A.V. Monitoring teplovogo rezhima ekspluatatsii kabel'nykh liniy 110-500 kV [Monitoring of the thermal mode of operation of 110-500 kV cable lines], *Sb. dokladov konferentsii «Linii elektropredachi 2010: Proektirovanie, stroitel'stvo, opyt ekspluatatsii i NTP»* [Collection of reports of the conference "power transmission Lines 2010: Design, construction, operational experience and NTP"], 2010, pp. 306-310.
16. Ukil A., Braendle H., Krippner P. Distributed temperature sensing: review of technology and applications, *Sensors Journal, IEEE*, 2012, Vol. 12, No. 5, pp. 885-892.
17. Deng J., Xiao H., Huo W. and other. Optical fiber sensor-based detection of partial discharges in power transformers, *Optics and Laser Technology*, 2001, Vol. 33, No. 5, pp. 305-311.
18. Bachelis D.S., Belorusev N.I., Saakyan A.E. Elektricheskie kabeli, provody i shnury: Sprav. [Electric cables, wires and cords: Reference]. Moscow: Energiya, 1971, 704 p.
19. El-Sharkawi, Park D.C. Electric the load forecasting using an artificial neural network, *IEEE Transaction on Power Systems*, 1991, No. 6 (2), pp. 442-449.
20. Peng M., Hubele N.F., Karady G.G. Advancement in the application of neural networks for short-term load forecasting, *IEEE Transactions on Power Systems*, 1992, Vol. 7, pp. 250-257.
21. Desouky A.A., Elkateb M.M. Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA, *IEE Proceedings of Generation, Transmission and Distribution*, 2000, Vol. 4, pp. 213-217.
22. Peng M., Hubele N.F., Karady G.G. Advancement in the application of neural networks for short-term load forecasting, *IEEE Transactions on Power Systems*, 1992, Vol. 7, pp. 250-257.
23. Papalexopoulos A.D., Hao S., Peng T.M. An implementation of a neural network based load forecasting model for the EMS, *IEEE Transactions on Power Systems*, 1994, Vol. 9, pp. 1956-1962.
24. Khotanzad A. ANNSTLF - A neural-network-based electric load forecasting system, *IEEE Transactions on Neural Networks*, 1997, Vol. 8, pp. 835-846.
25. Chen C.S., Tzeng Y.M., Hwang J.C. The application of artificial neural networks to substation load forecasting Electric, *Power Systems Research*, 1996, Vol. 38, Issue 2, pp. 153-160.
26. Chow T.W., Leung C.T. Nonlinear autoregressive integrated neural network model for short-term load forecasting, *IEE Proceedings, Generation, Transmission and Distribution*, 1996, Vol. 143, pp. 500-506.

27. Kurbatskiy V.G., Tomin N.V. Praktika ispol'zovaniya novykh informatsionnykh tekhnologiy dlya prognozirovaniya i analiza otдел'nykh kharakteristik setevykh energopredpriyatiy [The Practice of using new information technologies for forecasting and analysis of individual characteristics of the network utilities], *Problemy energetiki* [Problems of power engineering], 2006, No. 3-4.

Статью рекомендовал к опубликованию д.т.н., профессор В.И. Финаев.

Полуянович Николай Константинович – Южный федеральный университет; e-mail: nik1-58@mail.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 89185693365; кафедра электротехники и мехатроники.

Дубяго Марина Николаевна – e-mail: w_m88@mail.ru; тел.: 89281758225; кафедра электротехники и мехатроники; аспирант.

Poluyanovich Nikolay Konstantinovich – Southern Federal University; e-mail: nik1-58@mail.ru; 44, Nekrasovsky, Taganrog, 347928, Russia; phone: +79185693365; the department of electric technics and mechatronics.

Dubyago Marina Nikolaevna – e-mail: w_m88@mail.ru; phone: +79281758225; the department of electrical engineering and mechatronics; graduate student.

УДК 004.853

DOI 10.18522/2311-3103-2020-2-66-78

В.В. Бова, Э.В. Кулиев, С.Н. Щеглов

ОЦЕНКА ЭФФЕКТИВНОСТИ МЕТОДА ПОИСКА АССОЦИАТИВНЫХ ПРАВИЛ ДЛЯ ЗАДАЧ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ*

Объемы современных баз данных имеют значительные объемы и содержат большие массивы информации. Одним из популярных методов обнаружения знаний для задач обработки и анализа больших данных стали алгоритмы поиска ассоциативных правил. В статье решается задача построения баз ассоциативных правил для анализа представленной большими массивами неструктурированных данных на основе поиска в них различных закономерностей с учетом значимости их признаков. Предложен метод синтеза баз ассоциативных правил, в котором выполняется построение транзакционной базы данных на основе вычисления пороговых значений поддержки и применением критериев оценивания косвенных ассоциаций, что позволяет извлекать как частые, так и неясные наборы ассоциативных правил. С целью повышения вычислительной эффективности извлечения ассоциативных правил, применяется генетический алгоритм оптимизации входных параметров признакового пространства поиска. Метод позволяет улучшить время извлечения правил, сократить число сгенерированных обобщенных правил, избежать затратной процедуры предобработки синтезированной базы правил. Разработан программно-алгоритмический модуль, с помощью которого проведены экспериментальные исследования метода синтеза ассоциативных правил на основе фильтрации входных параметров модели поиска для решения задач обработки неструктурированных данных. Проведенные серии экспериментов на тестовых транзакционных базах данных позволили уточнить теоретические оценки временной сложности метода, в котором для вычисления взвешенной поддержки наборов правил с учетом оценки априорной информативности признаков, входящих в данный набор применяется генетический алгоритм. Временная сложность разработанного метода составляет $\approx O(I^2)$. Сравнительный анализ проводился на тестовых данных корпуса Retail Data с алгоритмами Apriori и Frequent Pattern-Growth. Результаты исследований подтвердили эффективность метода поиска на больших наборах транзакций, позволяющего более чем на 40 % уменьшить мощность избыточного множества извлеченных ассоциативных правил по сравнению с известными алгоритмами и показали перспективность его применения для задачи обнаружения знаний при обработке данных большого объема.

Извлечение ассоциативных правил; неструктурированные данные; генетический алгоритм; база ассоциативных правил; большие данные.

* Работа выполнена при поддержке грантов РФФИ № 19-07-00099 и № 18-07-00055.

V.V. Bova, E.V. Kuliev, S.N. Shcheglov

ESTIMATING THE EFFECTIVENESS OF THE METHOD FOR SEARCHING THE ASSOCIATIVE RULES FOR THE TASKS OF PROCESSING BIG DATA

The modern databases have significant volume and consist of large masses of information. One of the popular methods of knowledge identification in terms of tasks of analysis and processing of large data volumes is composed of the algorithms for searching the associative rules. The paper solves the problem of building the bases of associative rules for the analysis of the unstructured large data volumes on the basis of searching different regularities considering the importance of their characteristics. The authors propose the method for synthesizing the bases and building the transaction database to calculate the threshold values of support and application of criteria of estimating implicit associations. This allows us to extract repeated and implicit associative rules. To improve the computational effectiveness of extracting the associative rules, the paper applies the genetic algorithm for optimization of input parameters of the characteristic searching space. The developed method shortens the time of rules extraction, reduces the number of generated common rules, and avoid the resource-consuming procedure of pre-processing the synthesized rule base. The authors developed the program and algorithmic module to carry out the experimental research of the proposed method for synthesizing the associative rules on the basis of filtering the input parameters of the search model for solving the tasks of processing the unstructured data. The experiments conducted on the test transaction bases allow us to clarify the theoretical estimations of time complexity of the proposed method that used the genetic algorithm to calculate the weighed support of the set of rules considering the assessment of a priori informative content of the characteristics included in the dataset. The time complexity of the developed method is estimated as $\approx O(I^2)$. The comparative analysis is performed using the test data of the Retail Data with the algorithms Apriori and Frequent Pattern-Growth. The results have proven the effectiveness of the search method on big sets of transactions. The method allows us to reduce the cardinal of an irredundant set of extracted associative rules in more than 40 % in comparison with the popular algorithms. The experiments have shown that the method can be effective for the tasks of knowledge discovery in terms of processing large volumes of data.

Associative rule extraction; unstructured data; genetic algorithm; associative rule base; big data.

Введение. В настоящее время для поиска закономерностей при решении задач, связанных с необходимостью извлечения новых знаний при обработке больших массивов неструктурированных данных широкое распространение получили ассоциативные правила [1–5]. Задача поиска закономерностей в больших данных связана с необходимостью анализа не только качественной, но и количественной информации [1–3, 5]. Исследуемые объекты данных описываются числовыми признаками, поэтому для поиска ассоциативных зависимостей и построения на их основе баз правил необходимо осуществлять предварительную обработку таких атрибутов с учетом значимости признаков [6–8].

Предобработка массивов данных производится на основе разбиения часто встречающихся наборов элементов (признаков) на непересекающиеся множества, каждый из которых представляется новым атрибутом признакового пространства транзакций. В теории ассоциативных правил понятие транзакции трактуется как набор событий или объектов, появляющихся одновременно в некотором наблюдении [6–10]. При решении задачи извлечения новых знаний об исследуемых зависимостях в данных возникают проблемы нахождения всех наборов элементов в том числе и неявных, позволяющих выявлять косвенные ассоциации. Это приводит к значительному возрастанию размерности решаемой задачи и повышает требования к вычислительным ресурсам [8].

Другой проблемой при поиске ассоциативных зависимостей, часто возникающей при анализе больших данных остается значительное число извлеченных правил, большинство из которых являются избыточными. Для устранения избыточности в представлении ассоциативных правил разрабатываются модификации известных алгоритмов поиска, основанные на свойствах частых замкнутых множеств [7–9].

Для оценки значимости выявленных ассоциативных правил производится фильтрация с варьированием параметров принадлежности поддержки (support) и достоверности (confidence) для различных наборов элементов. Анализу подвергаются только те правила, для которых значения мер близости могут быть больше порога минимальной поддержки [7]. Фильтрация по принадлежности элементов транзакций к той или иной группе позволяет снизить количество вычисленных ассоциативных правил, но полностью не решает проблему поиска в признаковом пространстве большой размерности [1–4].

Поэтому актуальной задачей является разработка и исследование эффективности метода синтеза баз ассоциативных правил, свободного от указанных недостатков и позволяющего извлекать не только часто встречающиеся наборы, но и неявные ассоциативные зависимости. *Предлагаемый метод решает задачу* отбора информативных признаков с использованием генетического алгоритма, что *позволяет* эффективно работать с категорийной входной информацией в процессе извлечения ассоциативных правил.

Метод построения базы обобщенных ассоциативных правил. Методы поиска обобщенных правил при вычислении используют информацию о группировке элементов (таксономии), что позволяет значительно расширить круг задач, решаемых алгоритмами поиска ассоциативных правил [6, 10]. При поиске обобщенных правил важной характеристикой, используемой в процессе их извлечения, является поддержка наборов элементов, а также ее пороговое значение, задаваемое экспертом в качестве параметра метода [8, 10–12]. Помимо априорных параметров для извлечения обобщенных правил: поддержки и достоверности представим такой параметр, как уровень интереса, позволяющий учитывать информацию об иерархической группировке элементов для отсека «неинтересных» правил [6–8]. Сформулируем постановку задачи генерации баз обобщенных ассоциативных правил.

Пусть задан набор (база) транзакций $D = \{T_1, T_2, \dots, T_{N_D}\}$, в котором каждый элемент T_j , $j = 1, 2, \dots, N_D$ содержит информацию о последовательности взаимосвязанных событий; $N_D = |D|$ – количество элементов в D ; $T_j = \{t_{1j}, t_{2j}, \dots, t_{N_{itemj}j}\} \subseteq I$ – j -я транзакция базы D , определяющая последовательность элементов t_{ij} с конкретным значением числовых атрибутов; t_{ij} – i -й элемент j -й транзакции T_j , $i = 1, 2, \dots, N_{|T_j|}$; $N_{|T_j|}$ – количество элементов в j -й транзакции T_j ; $I = \{\tau_1, \tau_2, \dots, \tau_{N_I}\}$ – множество возможных переменных (признаков), которые могут входить в набор элементов каждой транзакции T_j , $j = 1, 2, \dots, N_D$ набора данных D ; а τ – a -й элемент множества I , $a = 1, 2, \dots, N_I$; $N_I = |I|$ – количество элементов множества I . Представим задачу синтеза базы обобщенных ассоциативных правил (БП) на основе D как генерацию обобщенных ассоциативных правил в виде $\langle X, v(X) \rangle \rightarrow \langle Y, v(Y) \rangle$. Обобщенным ассоциативным правилом называется импликация $\langle X, v(X) \rangle \rightarrow \langle Y, v(Y) \rangle: X \subset I, Y \subset I, X \cap Y = \emptyset$, где $v(X)$ и $v(Y)$ – множества значений признаков, принадлежащих непересекающимся множествам X и Y [7–9].

Для всех транзакций T_j параметр поддержки рассчитывается как пересечение функций принадлежности наборов элементов (признаков) входящих в T_j :

$$\text{supp}(T_j) = \bigcap_{\tau_a \in T_j} \mu_a(T_j), \quad (1)$$

где $\mu_a(T_j)$ – значение функции принадлежности a -го признака, вычисленное для транзакции T_j . Согласно формуле (1) порог поддержки набора X вычисляется суммированием поддержек всех транзакций, содержащих это множество признаков:

$$\text{supp}(X) = \sum_{X \subseteq T_j} \text{supp}(T_j) = \sum_{X \subseteq T_j} \bigcap_{\tau_a \in T_j} \mu_a(T_j). \quad (2)$$

Для вычисления взвешенной поддержки набора X предлагается использовать генетический алгоритм поиска наилучшей оценки индивидуальной информативности признаков, предложенный авторами в работах [13–15]. Согласно формулам (1) и (2) оператор вычисления поддержки, определим следующим образом:

$$\text{wsupp}(X) = \text{supp}(X) \sum_{\tau_a \in X} w_a, \quad (3)$$

где величина $\sum_{\tau_a \in X} w_a$ определяет оценку информативности набора признаков X .

Взвешенная поддержка обобщенного правила $X \rightarrow Y$ определяется по формуле:

$$\text{wsupp}(X \rightarrow Y) = \text{supp}(X \cup Y) \sum_{\tau_a \in X \cup Y} w_a. \quad (4)$$

Рассмотрим случай, когда X – часто встречающийся взвешенный набор признаков, для которого справедливо условие: $\text{wsupp}(X) \geq \text{wminsupport}$, где wminsupport – минимальный порог взвешенной поддержки [7]. Для извлечения частых наборов X установим значение минимальной поддержки $\text{minsupport} = \alpha \cdot \text{wminsupport}$ с учетом α – коэффициента оценки информативности самой длинной иерархической последовательности признаков в БД D для транзакции T_j ($|T_j| = \max_{T_j \in D} |T_j|$) в D :

$$\alpha = \frac{1}{\sum_{T_j(|T_j| = \max_{T_j \in D} |T_j|)} w_a}. \quad (5)$$

Особым случаем для извлечения новых знаний об исследуемых объектах в массивах данных большого объема являются косвенные (непрямые) наборы признаков, позволяющие выявлять интересные ассоциативные зависимости [9, 14]. Для их поиска рассмотрим зависимость элементов множеств X и Y косвенно связанных с третьим набором Z , $Z: X \xrightarrow{Z} Y$. Наличие косвенной ассоциативной связи подтверждается, если подтверждается истинность условий [7, 15]:

1) значение взвешенной поддержки набора $X \cup Y$ меньше: $\text{wsupp}(X \cup Y) < \beta_{\text{wsupp}(X \cup Y)}$, где $\beta_{\text{wsupp}(X \cup Y)}$ – минимальный порог значения взвешенной поддержки для нечасто встречающихся множеств X и Y , т.е. $\beta_{\text{wsupp}(X \cup Y)} = \text{wminsupport}$;

2) для непустой набора Z ($\exists Z \neq \emptyset$) справедливы условия:

$$\begin{cases} \text{wsupp}(X \cup Z) \geq \beta_{\text{wsupp}(Z)}; \\ \text{wsupp}(Y \cup Z) \geq \beta_{\text{wsupp}(Z)}; \end{cases} \text{ и } \begin{cases} w(X, Z) \geq w_{\min}; \\ w(Y, Z) \geq w_{\min}, \end{cases} \quad (6)$$

где $\beta_{\text{wsupp}(Z)}$ – пороговое значение взвешенной поддержки между некоторым набором X и Y и набором Z . Параметр $\beta_{\text{wsupp}(Z)}$ задается по правилу: $\beta_{\text{wsupp}(Z)} \geq \beta_{\text{wsupp}(X \cup Y)}$; $w(X, Z)$ и $w(Y, Z)$ – меры близости взаимосвязи между множествами

X , Y и Z соответственно; w_{min} – минимально допустимое значение показателя значимости взаимосвязи между множествами элементов базы D . Мере оценки близости $w(X, Z)$ рассчитаем по следующей формуле [15]:

$$w(X, Z) = \frac{p(X \cap Z)}{\sqrt{p(X)p(Z)}}, \quad (7)$$

где $p(X)$, $p(Z)$, $p(X \cap Z)$ – вероятность появления наборов X , Z и $X \cap Z$ в базе данных D . Согласно формуле (7), использование предложенных выше критериев и их пороговых значений позволит извлекать не только часто встречающиеся, но и неявные наборы, являющиеся интересными и позволяющие выявлять новые знания об исследуемых транзакционных и категориальных данных.

После выполнения механизма фильтрации, связанной с предобработкой транзакционной БД и нахождением пороговых значений $wminsupport$ и $wminconfidence$, выполняется извлечение ассоциативных правил и построение на их основе БП [9].

Для построения базы ассоциативных правил задается транзакционная база данных D , содержащая транзакции T_j с числовыми значениями атрибутов τ_a , определяется набор функций принадлежности μ , используемых для разбиения диапазонов Δ_{ak} численных признаков на множества, определяется минимальная взвешенная поддержка $wminsupport$ и взвешенная достоверность $wminconfidence$, задаются пороговые значения $(\beta_{wsupp(X \cup Y)}, \beta_{wsupp(Z)}, wmin)$, необходимые для работы алгоритма генетического поиска [15].

На следующем этапе производится вычисление мощности k -го диапазона разбиения a -го признака: $C_{\Delta_{ak}} = \sum_{j=1}^{N_D} \mu_{ak}(\tau_a \in T_j)$ и находится его максимальная величина: $\max C_{\Delta_a} = \max_{k=1,2,\dots,N_{\text{инт.}a}} C_{\Delta_{ak}}$, $a = 1, 2, \dots, |I|$. Для всех наборов разбиения $\max \Delta_a$, $a = 1, 2, \dots, |I|$ вычисляется взвешенная поддержка $wsupp(\max \Delta_a)$ по формулам (3)-(5). В массив часто встречающихся наборов: $FI_1 = \{\max \Delta_a | wsupp(\max \Delta_a) \geq wminsupport\}$ заносятся все интервалы $\max \Delta_a$, с допустимым порогом поддержки $wminsupport$. А в массив $RI_1: \{\max \Delta_a | wsupp(\max \Delta_a) \geq wminsupport\}$ – нечасто встречающиеся последовательности элементов, выявленные на множествах с низкими значениями взвешенных поддержек $wsupp(\max \Delta_a)$.

На следующем шаге производится формирование множества $C_{d+1}(d+1)$ -элементных кандидатов в часто встречающиеся наборы на основе текущего множества FI_d . На этапе генерации множества кандидатов C_{d+1} отсекаются (не извлекаются и не сохраняются в C_{d+1}) те наборы, которые не относятся к часто встречающимся, на основе значений $wsupp$, рассчитанными по формуле (3) и хранящимися во множестве RI . По формуле (6), при вычислении нового множества C_{d+1} кандидатов, кандидат X , содержащий подмножество $Y \subset X$, отброшенный на предыдущих этапах как нечасто встречающееся правило ($Y \in RI$), не отбирается в следующее множество C_{d+1} кандидатов в частые наборы.

После формирования множества C_{d+1} для всех $X = \{\tau_1, \tau_2, \dots, \tau_{d+1}\} \in C_{d+1}$ ($|X| = d+1$) вычисляется его степень принадлежности T_j :

$$\mu_X(T_j) = \bigcap_{a: \tau_a \in X} \mu_a(\tau_a \in T_j), \quad (8)$$

далее с учетом формулы (8) определим взвешенную поддержку набора X как:

$$wsupp(X) = \sum_{X \in T_j, T_j \in D} \mu_X(T_j) \sum_{\tau_a \in X} w_a. \quad (9)$$

Если значение $wsupp$, найденное по формуле (9) не менее минимально допустимого порога $wminsupport$, множество X заносится в массив FI_{d+1} , в противном случае – в массив редко встречающихся наборов RI_{d+1} . В случае, если $FI_{d+1} \neq \emptyset$, выполняются действия, аналогичные описанным выше. Обобщенные правила генерируются исходя из формул (3)-(5) с учетом достоверности $wminconfidence$:

$$wconf(X \rightarrow Y) = \frac{wsupp(X \rightarrow Y)}{wsupp(X)} \geq wminconfidence, X \cap Y = \emptyset.$$

Рассмотренный метод построения базы обобщенных ассоциативных правил предполагает использование критериев для оценивания косвенных ассоциаций, что уменьшает вероятность извлечения правил, некорректно описывающих исследуемые объекты неструктурированной информации, а также позволяет находить не только часто встречающиеся наборы, но и интересные ассоциативные правила.

2. Оценка эффективности метода поиска и синтеза ассоциативных правил.

Для исследования эффективности предложенного метода оценим его вычислительную сложность O . Извлечение обобщенных правил связано с построением множества часто встречающихся наборов FI , что в свою очередь требует вычисления ГА пороговых значений поддержек каждого из кандидатов [15], максимальное количество которых не превышает $|I|^2$. Сложность этого процесса составляет $O_{FI}(|I|)$ операций. Процесс извлечения правил из множества FI предполагает обработку каждого подмножества $X \in FI$, на что потребуются $O_{извл}(|I|^2)$ операций. Поэтому вычислительная сложность предложенного метода составит $O = O_{FI}(|I|^2) + O_{извл}(|I|^2) = O(|I|^2)$. Поскольку зависимость элементарных операций от размера входных данных является полиномиальной (квадратичной), можно сделать вывод о том, что предложенный метод является вычислительно эффективным.

Для исследования свойств и характеристик предложенного метода построения баз ассоциативных правил была выполнена его программная реализация на языке программирования C# [16]. Программно-алгоритмический модуль позволяет настроить параметры алгоритма поиска обобщенных правил: предельно допустимые значения коэффициента поддержки, минимальную и максимальную достоверность процедуры поиска правил и максимальную мощность множества (рис. 1).

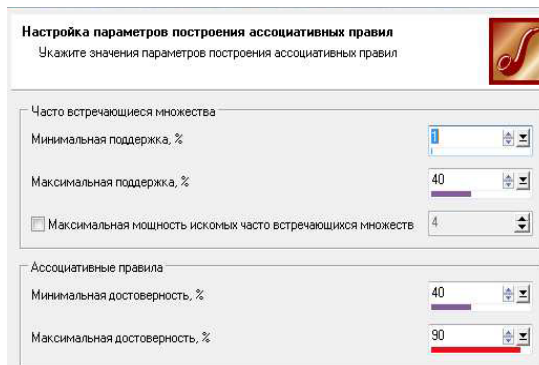


Рис. 1. Интерфейс настройки параметров

После запуска алгоритма поиска вычисления правил, на экране отображается информация о количестве множеств, количестве найденных правил, а также диаграмма распределения найденных часто встречающихся множеств по мощности (рис. 2).

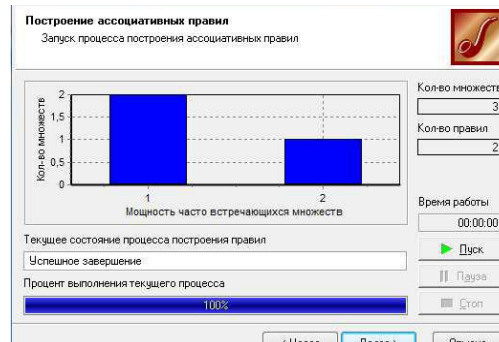


Рис. 2. Процесс выявления ассоциативных правил

Экспериментальное исследование разработанного метода выполнялось с помощью тестовых транзакционных баз данных. Результаты исследований приведены в табл. 1, в которой используются такие обозначения: N_D – количество транзакций T_j в базе D ; I – количество элементов (признаков), из которых могли формироваться транзакции, $|T_j|_{cp}$ – среднее количество признаков в транзакциях базы D ; $|БП|$ – количество извлеченных обобщенных ассоциативных правил в синтезированной базе правил $БП$; t – время работы метода.

В результате анализа данных, представленных в табл. 1, можно сделать вывод о том, что время работы предложенного метода существенно зависит от количества I элементов в базе D , что подтверждает оценку вычислительной сложности $O(I^2)$ метода.

Расчетные данные результатов экспериментов показали, что количество сгенерированных правил $|БП|$ незначительно увеличивается с ростом параметров N_D , I и T_j базы транзакций D .

Таблица 1

Результаты экспериментальных исследований

№	Характеристики базы транзакций D			Результаты синтеза баз ассоциативных правил	
	N_D	$ I $	$ T_j $	$ БП $	t
1	10000	100	10	212	0,22
2	10000	500	20	257	0,41
3	10000	1000	30	290	0,92
4	50000	100	10	491	0,53
5	50000	500	20	525	2,27
6	50000	1000	30	546	5,82
7	100000	1000	10	612	11,13
8	100000	5000	20	560	14,15
9	100000	10000	30	486	17,68

Так при увеличении среднего числа признаков от 100 до 10000 для постоянных значений элементов транзакций размерности $T_j = \{10, 20, 30\}$ количество сформированных взаимосвязей (правил) находится в диапазоне от 20 % до 6 %. Снижение данного показателя обусловлено большим количеством различных элементов в множестве часто встречающихся наборов FI , что позволило сгенерировать оптимальное множество ассоциативных правил. На сравнительных графиках зависимости найден-

ных обобщенных правил и времени работы метода от $|T_j|$ при различных значениях числа транзакций N_D , представленных на рис. 3 и 4 можно оценить, что наилучшие результаты работы метода достигаются при больших значениях $N=100000$ и увеличении усредненного показателя размерности признакового пространства, при которых количество извлеченных правил уменьшается в среднем на 14 %.

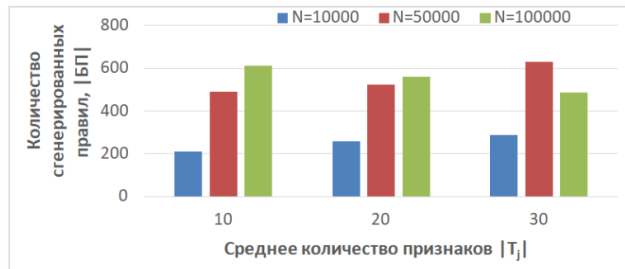


Рис. 3. Гистограмма зависимости количества сформированных ассоциативных правил от числа признаков наборов

Сравнительные исследования показали, что с увеличением числа транзакций и их признакового пространства в БД временные затраты на поиск частых предметных наборов растут пропорционально (рис. 4).



Рис. 4. График зависимости времени генерации ассоциативных правил от числа транзакций в БД

Выявленные факторы и зависимости позволяют сформулировать вывод о том, что метод является эффективным для задач извлечения ассоциативных зависимостей при обработке данных большого объема.

3. Экспериментальные исследования. С целью исследования эффективности метода для задачи поиска ассоциативных правил используется тестовый набор данных, на котором строится методика сравнительной оценки качества с алгоритмами Apriori и Frequent Pattern-Growth (FRG) [17–22]. Алгоритм Apriori предназначен для поиска всех частых множеств признаков. Он является поуровневым, использует стратегию поиска в ширину и осуществляет его снизу-вверх [17]. Apriori использует свойство анти-монотонности нахождения многоэлементных наборов. Благодаря этому свойству перебор не является «жадным» и позволяет снижать размерность пространства поиска и обрабатывать большие массивы информации за приемлемое время [20]. В основе метода FPG лежит предобработка базы транзакций, в процессе которой эта БД преобразуется в компактную структуру, называемую Frequent-Pattern Tree – дерево популярных предметных наборов [19]. Алгоритм FPG позволяет избежать затратной процедуры генерации кандидатов и сократить число проходов БД.

Алгоритмы Apriori, FPG и разработанный метод сравнивались по числу сгенерированных обобщенных правил и времени работы. В экспериментах оценивались максимальные наборы, для которых вычисляется поддержка $\text{supp}(\{t_1, t_2, \dots, t_n\}) \geq \text{minsupp}$ и для $\forall b$ осуществляется фильтрация элементов по правилу $\text{supp}(\{t_1, t_2, \dots, t_n, b\}) < \text{minsupp}$.

Для определения эффективности разработанного метода были проведены исследования времени и качества извлечения ассоциативных правил для разного набора тестовых примеров, различающихся количеством элементов транзакций T_j корпусов данных. Для тестирования эффективности предложенной методики фильтрации входных данных использовался входной набор корпуса данных Retail Data [23], содержащий 89238 транзакцию, а число различных элементов в ней равно 17562. Для рассматриваемых алгоритмов тестовая серия производилась на разных значениях коэффициента уровня поддержки от 0,2 до 0,001. В ходе вычислительного эксперимента было определено лучшее его значение 0,012. Это объясняется тем, что по мере уменьшения уровня поддержки общее число связей увеличивается и количество несвязанных элементов в транзакциях и сгенерированных наборов уменьшается [24, 25], а алгоритм FPG не может быть завершен за приемлемое время. Численные результаты экспериментов для алгоритмов по критериям сравнения приведены в табл. 2, а также на рис. 5 и 6.

Таблица 2

Результаты сравнения методов для задачи поиска ассоциативных правил

Характеристики транзакций корпуса данных		Методы					
		Apriori		FPG		Предложенный метод	
$ T_j $	$ I $	t	БП	t	БП	t	БП
1000	500	16,34	66	1,55	26	1,23	15
5000	500	17,54	72	2,62	35	2,76	17
10000	500	23,79	85	2,96	50	3,54	26
15000	500	25,90	97	3,88	62	4,71	30
20000	500	26,87	105	4,94	70	5,77	32
25000	500	27,12	127	4,98	84	6,85	33

В экспериментальном исследовании оценивались элементы наборов тестовых данных транзакции размерностью от 1000 до 25000 записей со средним количеством признаков $|I|$ равным 500. В результате анализа табл. 2 и графиков зависимостей можно сделать, что алгоритмы FPG и разработанный метод демонстрируют высокую работоспособность в сравнении с алгоритмом Apriori по числу извлеченных строгих ассоциативных правил $|БП|$ и по времени работы t .

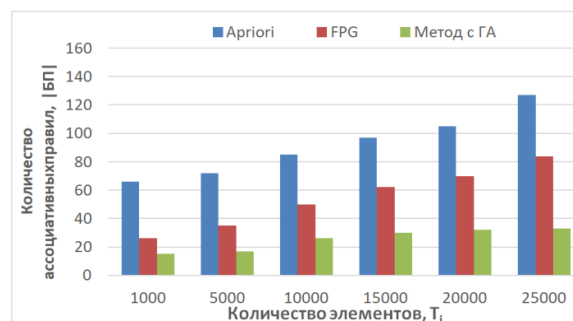


Рис. 5. Гисторграмма сравнения числа извлеченных ассоциативных правил от числа транзакций в БД

Методика предобработки данных позволяет добиться лучших результатов извлечения частых наборов алгоритмом FPG в среднем на 30 %, а с увеличением размерности пространства поиска до 25000 элементов предложенный метод дает улучшение качества сформированных правил более 40 %.

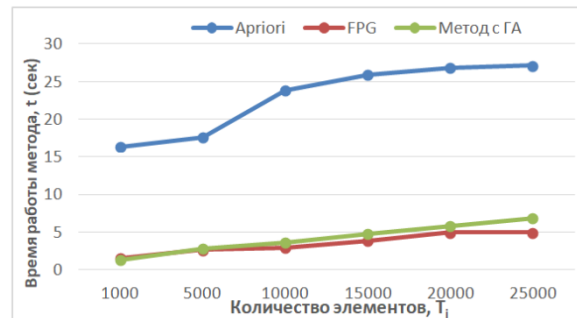


Рис. 6. График зависимости времени извлечения ассоциативных правил от числа элементов транзакций в БД

Предложенный метод с фильтрацией на основе ГА по времени работы сопоставим с алгоритмом FPG и позволяет более чем в два раза уменьшить мощность избыточного множества обобщенных ассоциативных правил, формируемого алгоритмом FPG.

Заключение. Предложен метод извлечения ассоциативных правил, основными этапами которого являются: фильтрация случайной выборки исходной базы данных транзакций, на основе ГА вычисления индивидуальной значимости признаков, вычисление пороговых значений поддержки и построение базы обобщенных ассоциативных правил. Отличительной особенностью метода является предобработка базы транзакций, в процессе которой БД преобразуется в компактную структуру, что обеспечивает эффективное и полное извлечение частых предметных наборов, позволяет избежать затратной процедуры генерации кандидатов, по сравнению с известными методами Apriori и FPG.

Разработан программно-алгоритмический модуль, позволяющий выполнять извлечение обобщенных ассоциативных правил. Теоретическая оценка эффективности метода показала, что для увеличения стабильности и точности работы метода размер выборки, полученной в результате работы ГА, позволит сократить необходимое число сканирований базы данных транзакций, обеспечивая приемлемые вычислительные затраты, сопоставимые с известным алгоритмом FPG и превосходящие по времени поиска ассоциативных правил алгоритм Apriori более чем в два раза. Временная сложность разработанного метода составляет $\approx O(I^2)$. Результаты проведенных экспериментов выявили, что предложенный метод более чем в два раза уменьшает мощность избыточного множества извлеченных ассоциативных правил, формируемого известными алгоритмами и при значительных показателях параметров транзакционной БД, показывает качество поиска на 4 % выше чем алгоритм FPG. Это подтверждает эффективность метода для решения задачи извлечения ассоциативных правил без потери информации о связях при обработке данных большого объема.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Газиев Г.З., Курдюкова Г.Н., Курдюков В.В. Кластеризация Big Data для их анализа и обработки // Сб. конференции «Направления и механизмы развития науки нового времени: от теории до внедрения результатов». – 2017. – С. 150-162.

2. Бова В.В., Щеглов С.Н., Лещанов Д.В. Модифицированный алгоритм ЕМ-кластеризации для задач интегрированной обработки больших данных // Известия ЮФУ. Технические науки. – 2018. – № 4 (165). – С. 197-211.
3. Bova V.V., Kureichik V.V., Scheglov S.N., Kureichik L.V. Multi-level ontological model of big data processing // *Advances in Intelligent Systems and Computing*. – 2019. – Vol. 874. – P. 171-181.
4. Wu X., Zhu X., Wu G., Ding W. Data mining with big data // *IEEE Transaction on Knowledge and Data Engineering*. – 2014. – Vol. 26. – P. 97-107.
5. Kravchenko Y.A., Kuliev E.V., Kursitys I.O. Information's semantic search, classification, structuring and integration objectives in the knowledge management context problems // 10th IEEE International Conference on «Application of Information and Communication Technologies, AICT 2016. – P. 136-141.
6. 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.
7. 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.
8. Ibrahim S., Chandran K.R. Compact Weighted Class Association Rule Mining using Information Gain // *International Journal of Data Mining and Knowledge Management Process*. – 2011. – Vol. 1. – P. 1-13.
9. Muyebe M., Khan M. S., Coenen F. Fuzzy weighted association rule mining with weighted support and confidence framework // *New Frontiers in Applied Data Mining Lecture Notes in Computer Science*. – 2009. – Vol. 5433. – P. 312-320.
10. Зайко Т.А., Олейник А.А., Субботин С.А. Ассоциативные правила в интеллектуальном анализе данных // *Вестник Национального технического университета Харьковский политехнический институт. Серия: Информатика и моделирование*. – 2013. – № 39 (1012). – С. 82-96.
11. Кравченко Ю.А. Модель фильтра знаний для задач семантической идентификации // Известия ЮФУ. Технические науки. – 2018. – № 4 (165). – С. 197-211.
12. Субботин С.А., Олейник А.А., Гофман Е.А. Интеллектуальные информационные технологии проектирования автоматизированных систем диагностирования и распознавания образов: монография / под ред. С.А. Субботина. – Харьков: ООО «Компания Смит», 2012. – 317 с.
13. Бова В.В., Щеглов С.Н., Лещанов Д.В. Применение методов генетического поиска для задач обработки ассоциативных правил // *XXI Международная конференция по мягким вычислениям и измерениям (SCM-2018)*. – СПб.: СПбГЭТУ «ЛЭТИ», 2018. – Т. 1. – С. 761-769.
14. Щеглов С.Н. Модифицированный алгоритм обработки и анализа неструктурированной информации на основе поиска ассоциативных правил // *Тр. Конгресса по интеллектуальным системам и информационным технологиям – «IS&IT'18»*. – Таганрог: Изд-во ЮФУ, 2018. – Т. 2. – С. 183-191.
15. Бова В.В., Щеглов С.Н., Лещанов Д.В. Modified Approach to Problems of Associative Rules Processing based on Genetic Search // 2019 International Russian Automation Conference (RusAutoCon). 10.1109/RUSAUTOCON. – 2019. – № 8867675.
16. 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.
17. Guo Z., Chi D., Wu J., Zhang W. A new wind speed forecasting strategy based on the chaotic time series modelling technique and the Apriori algorithm // *Energy Conversion and Management*. – 2014. – No. 84. – P. 140-151.
18. Kumar B.S. Rukmani K.V. Implementation of web usage mining using Apriori and FP Growth algorithms // *International Journal of Advanced Networking and Applications*. – 2010. – Vol. 400. – P. 400-404.
19. Пальмов С.В., Франузова Е.Н. Алгоритм поиска ассоциативных правил FP-GROWTH // *Национальная ассоциация ученых*. – М.: Изд-во: ООО «Евразийское Научное Содружество», 2016. – № 10-1 (26). – С. 27-32.

20. Qureshi Z., Bansal S. Improving Apriori Algorithm to get better performance with Cloud Computing // *International Journal of Software and Hardware Research in Engineerin.* – 2014. – Vol. 2. – P. 33-37.
21. Singh J., Ram H. Improving Efficiency of Apriori Algorithm Using // *International Journal of Scientific and Research Publications.* – 2013. – Vol. 3. – P. 1-4.
22. Yahya O., Hegazy O., Ezat E. An efficient implementation of Apriori algorithm based on Hadoop-Mapreduce model // *International Journal of Reviews in Computing.* – 2012. – Vol. 12. – P. 59-67.
23. Frequent Itemset Mining Implementations Repository. Retail. – URL: <http://fimi.ua.ac.be/data/retail.dat/>.
24. Zhao Y., Zhang C., Cao L. Post-mining of association rules: techniques for effective knowledge extraction. – New York: Information Science Reference. 2009. – 372 p.
25. Gkoulalas-Divanis A., Verykios V.S. Association Rule Hiding for Data Mining. – New York: Springer-Verlag. 2010. – 150 p.

REFERENCES

1. Gaziev G.Z., Kurdyukova G.N., Kurdyukov V.V. Klasterizatsiya Big Data dlya ikh analiza i obrabotki [Clusterization of Big Data for their analysis and processing], Sb. konferentsii «Napravleniya i mekhanizmy razvitiya nauki novogo vremeni: ot teorii do vnedreniya rezul'tatov» [Collection of the conference "Directions and mechanisms of modern science development: from theory to implementation of results"], 2017, pp. 150-162.
2. Bova V.V., Shcheglov S.N., Leshchanov D.V. Modifitsirovannyi algoritm EM-klasterizatsii dlya zadach integrirovannoy obrabotki bol'shikh dannykh [Modified EM clustering algorithm for integrated big data processing], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2018, No. 4 (165), pp. 197-211.
3. Bova V.V., Kureichik V.V., Scheglov S.N., Kureichik L.V. Multi-level ontological model of big data processing, *Advances in Intelligent Systems and Computing*, 2019, Vol. 874, pp. 171-181.
4. Wu X., Zhu X., Wu G., Ding W. Data mining with big data, *IEEE Transaction on Knowledge and Data Engineering*, 2014, Vol. 26, pp. 97-107.
5. Kravchenko Y.A., Kuliev E.V., Kursitys I.O. Information's semantic search, classification, structuring and integration objectives in the knowledge management context problems, *10th IEEE International Conference on «Application of Information and Communication Technologies, AICT 2016*, pp. 136-141.
6. 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.
7. Zayko T.A., Oleynik A.A., Subbotin S.A. Izvlechenie chislennykh assotsiativnykh 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.
8. Ibrahim S., Chandran K.R. Compact Weighted Class Association Rule Mining using Information Gain, *International Journal of Data Mining and Knowledge Management Process*, 2011, Vol. 1, pp. 1-13.
9. Mueyba M., Khan M. S., Coenen F. Fuzzy weighted association rule mining with weighted support and confidence framework, *New Frontiers in Applied Data Mining Lecture Notes in Computer Science*, 2009, Vol. 5433, pp. 312-320.
10. Zayko T.A., Oleynik A.A., Subbotin S.A. Assotsiativnye pravila v intellektual'nom analize dannykh [Associative rules in data mining], *Vestnik Natsional'nogo tekhnicheskogo universiteta Khar'kovskiy politekhnicheskiiy institut. Seriya: Informatika i modelirovanie* [Bulletin of the national technical University Kharkiv Polytechnic Institute. Series: computer Science and modeling], 2013, No. 39 (1012), pp. 82-96.
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. Subbotin S.A., Oleynik A.A., Gofman E.A. Intellektual'nye informatsionnye tekhnologii proektirovaniya avtomatizirovannykh sistem diagnostirovaniya i raspoznavaniya obrazov: monografiya [Intelligent information technologies of automated diagnosis and pattern recognition: monograph], ed. by S.A. Subbotina. Khar'kov: OOO «Kompaniya Smit», 2012, 317 p.

13. Bova V.V., Shcheglov S.N., Leshchanov D.V. Primenenie metodov geneticheskogo poiska dlya zadach obrabotki assotsiativnykh pravil [Application of genetic search methods for processing associative rules], *XXI Mezhdunarodnaya konferentsiya po myagkim vychisleniyam i izmereniyam (SCM-2018)* [XXI international conference on soft computing and measurement (SCM-2018)]. Saint Petersburg: SPbGETU «LETI», 2018, Vol. 1, pp. 761-769.
14. Shcheglov S.N. Modifitsirovannyi algoritm obrabotki i analiza nestrukturirovannoy informatsii na osnove poiska assotsiativnykh pravil [Modified algorithm for processing and analysis of unstructured information based on search for associative rules], *Tr. Kongressa po intellektual'nykh sistemam i informatsionnym tekhnologiyam – «IS&IT'18»* [Proceedings of the Congress on intelligent systems and information technologies – "IS&IT'18"]. Taganrog: Izd-vo YuFU, 2018, Vol. 2, pp. 183-191.
15. Bova V.V., Scheglov S.N., Lemanov D.V. Modified Approach to Problems of Associative Rules Processing based on Genetic Search, *2019 International Russian Automation Conference (RusAutoCon)*. 10.1109/RUSAUTOCON, 2019, No. 8867675.
16. 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 tsentra RAN* [Izvestiya Kabardino-Balkar scientific center of the Russian Academy of Sciences], 2019, No. 4 (90), pp. 14-23.
17. Guo Z., Chi D., Wu J., Zhang W. A new wind speed forecasting strategy based on the chaotic time series modelling technique and the Apriori algorithm, *Energy Conversion and Management*, 2014, No. 84, pp. 140-151.
18. Kumar B.S. Rukmani K.V. Implementation of web usage mining using Apriori and FP Growth algorithms, *International Journal of Advanced Networking and Applications*, 2010, Vol. 400, pp. 400-404.
19. Pal'mov S.V., Franuzova E.N. Algoritm poiska assotsiativnykh pravil FP-GROWTH [Search algorithm for associative rules FP-GROWTH], *Natsional'naya assotsiatsiya uchenykh* [National Association of scientists]. Moscow. Izd-vo: OOO «Evraziyskoe Nauchnoe Sodruzhestvo», 2016, No. 10-1 (26), pp. 27-32.
20. Qureshi Z. Bansal S. Improving Apriori Algorithm to get better performance with Cloud Computing, *International Journal of Software and Hardware Research in Engineerin*, 2014, Vol. 2, pp. 33-37.
21. Singh J., Ram H. Improving Efficiency of Apriori Algorithm Using, *International Journal of Scientific and Research Publications*, 2013, Vol. 3, pp. 1-4.
22. Yahya O., Hegazy O., Ezat E. An efficient implementation of Apriori algorithm based on Hadoop-Mapreduce model, *International Journal of Reviews in Computing*, 2012, Vol. 12, pp. 59-67.
23. Frequent Itemset Mining Implementations Repository. Retail. Available at: <http://fimi.ua.ac.be/data/retail.dat/>.
24. Zhao Y., Zhang C., Cao L. Post-mining of association rules: techniques for effective knowledge extraction. New York: Information Science Reference. 2009, 372 p.
25. Gkoulalas-Divanis A., Verykios V.S. Association Rule Hiding for Data Mining. New York: Springer-Verlag. 2010, 150 p.

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

Бова Виктория Викторовна – Южный федеральный университет; e-mail: vvbova@sfedu.ru, 347928, г. Таганрог, Некрасовский, 44; тел.: 88634371651; доцент.

Кулиев Эльмар Валерьевич – e-mail: elmar_2005@mail.ru; кафедра систем автоматизированного проектирования; доцент.

Щеглов Сергей Николаевич – e-mail: srg_sch@mail.ru; доцент.

Bova Victoria Victorovna – Southern Federal University; e-mail: vvbova@sfedu.ru; 44, Nekrasovskiy, Taganrog, 347928, Russia; phone: +78634371651; associate professor.

Kuliev Elmar Valerevich – e-mail: elmar_2005@mail.ru; the department of computer aided design; associate professor.

Scheglov Sergey Nikolaevich – e-mail: srg_sch@mail.ru; associate professor.

Д.Ю. Кравченко, Ю.А. Кравченко, В.В. Курейчик, А.Э. Саак

**БИОИНСПИРИРОВАННЫЙ МЕТОД ИМИТАЦИОННОГО
ДИСПЕТЧИРОВАНИЯ ДЛЯ ДИСПЕТЧЕРИЗАЦИИ ПОТОКОВ
ПАРАЛЛЕЛЬНЫХ ЗАЯВОК В GRID- СИСТЕМАХ***

Статья посвящена решению задачи диспетчеризации потоков параллельных заявок в пространственно-распределённых вычислительных системах. Актуальность задачи обоснована значительным ростом востребованности парадигмы распределённых вычислений в условиях информационного переполнения и неопределённости. В статье рассмотрены проблемы диспетчирования заявок пользователей, требующих для своего обслуживания несколько процессоров одновременно, что выходит за рамки классической теории расписаний. Проанализированы аспекты эффективности применения эвристических алгоритмов диспетчирования планарными ресурсами. Определены причины их недостаточности как по результативности, так и по эмпиричности подходов. Предложено решать задачу диспетчеризации параллельных заявок на основе комплексного применения коалиции интеллектуальных агентов и событийной имитационной модели. Классификацию поступающих на вход заявок предлагается проводить на основе применения модифицированного биоинспирированного метода оптимизации поиском кукушки. Совместное использование коалиции интеллектуальных агентов и биоинспирированного метода позволит обеспечить беспрецедентный параллелизм вычислений, а последующее определение путей обработки классифицированных заявок на основе имитационной модели сформирует наборы альтернативных решений, позволяющих ускорить решение задач и оптимизировать распределение имеющихся вычислительных ресурсов в зависимости от наборов поступающих заявок. Для оценки эффективности предложенного подхода разработан программный продукт и проведены эксперименты с разным количеством поступающих на вход заявок. Каждая поступающая на вход заявка имеет определенный набор атрибутов, являющийся вектором признаков заявки. Степень сходства вектора признаков заявки и эталонного вектора признаков вершины в распределяющей имитационной модели является критерием классификации заявки. Для повышения качества процесса диспетчеризации введены новые процедуры дублирования неклассифицированных заявок, позволяющие интенсифицировать поиск совпадений в векторах признаков. Также предусмотрены резервные траектории диспетчеризации необходимые для обработки прецедентов появления на входах заявок с абсолютным приоритетом. Полученные количественные оценки демонстрируют экономию времени при решении задач относительно большой размерности (от 500000 вершин) не менее 10 %. Временная сложность в рассмотренных примерах составила $O(n^2)$. Описанные исследования имеют высокий уровень теоретической и практической значимости и напрямую связаны с решением классических задач искусственного интеллекта, направленных на поиск скрытых зависимостей и закономерностей на множестве больших данных.

Диспетчеризация; параллельные вычисления; grid- системы; потоки заявок; биоинспирированный поиск; имитационное моделирование; интеллектуальные агенты.

D.Yu. Kravchenko, Yu.A. Kravchenko, V.V. Kureichik, A.E. Saak

**BIOINSPIRED SIMULATION METHOD FOR SCHEDULING OF PARALLEL
FLOWS APPLICATIONS IN GRID-SYSTEMS**

The article is devoted to solving the problem of parallel requests scheduling flows in spatially distributed computing systems. The relevance of the task is justified by a significant increase in the demand for the distributed computing paradigm in the conditions of information overflow and uncertainty. The article discusses the problems of scheduling user requests that require several

* Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов №№ 20-01-00148 и 18-01-00658.

processors at the same time, which goes beyond the classical theory of schedules. The aspects of the efficiency of using heuristic algorithms for scheduling planar resources are analyzed. The reasons for their insufficiency are determined both in terms of effectiveness and empirical approaches. The paper proposes to solve the problem of scheduling parallel applications based on the integrated application of intelligent agents coalition and an event simulation model. It is proposed to classify incoming applications on the basis of using a modified bio-inspired optimization method for cuckoo search. The joint use of a coalition of intelligent agents and a bio-inspired method will allow for unprecedented parallelism of calculations, and the subsequent determination of the processing classified applications ways on the basis of a simulation model will allow us to form sets of alternative solutions to speed up problem solving and optimize the distribution of available computing resources depending on the sets of incoming applications. To evaluate the effectiveness of the proposed approach, a software product was developed and experiments were conducted with a different number of incoming applications. Each incoming application has a certain set of attributes, which is a vector of the application characteristics. The degree of the application similarity feature vector and the vertex reference feature vector in the distributing simulation model is a classification criterion for the application. To improve the quality of the dispatch process, new procedures for duplicating unclassified applications have been introduced, which allow intensifying the search for matches in feature vectors. It also provides backup dispatching trajectories necessary for processing precedents for the appearance of applications with absolute priority at the inputs. The quantitative estimates obtained demonstrate time savings in solving problems of relatively large dimension (from 500,000 vertices) of at least 10%. The time complexity in the considered examples was $O(n^2)$. The described studies have a high level of theoretical and practical significance and are directly related to the solution of classical problems of artificial intelligence aimed at finding hidden dependencies and patterns on a large set of big data.

Scheduling; parallel computing; grid-systems; application flows; bio-inspired search; simulation; intelligent agents.

Введение. В настоящее время проблема эффективного диспетчирования параллельных информационных потоков с оптимизацией распределения используемых ресурсов до конца не решена. Не предложен формальный аппарат среды диспетчирования, который позволит выявить закономерности множества заявок, провести их классификацию и предложить эвристические алгоритмы распределения данных заявок по информационным процессам, обеспечивающим параллелизм выполнения и оптимальное использование вычислительных ресурсов. Разработка такого формального аппарата является актуальной задачей.

Рассмотрим базовые архитектуры для диспетчирования задач (job scheduling) и управления ресурсами Grid (Grid resource management systems), предложенные в работах [1–3]. В основе существующих классификаций данными архитектурами являются: централизованная (centralized), иерархическая (hierarchical) и распределённая (decentralized, distributed). В работах [4–9] эти архитектуры систем диспетчирования были названы парадигмами. Централизованная, иерархическая и распределённая структуры различаются способами принятия решения об управлении ресурсами и задачами. Так, в централизованной структуре решение принимается центральным диспетчером, обладающим полной информацией о вычислительных ресурсах и задачах. В иерархической структуре решение принимается иерархическим способом: центральный диспетчер распределяет задачи по локальным диспетчерам, которые производят их распределение в информационных процессах. В распределённой структуре решение принимается локальными диспетчерами, которые обмениваются между собой информацией, принимая решение параллельно.

В данной работе использована комбинированная распределённо-централизованная схема поиска решений при диспетчеризации. Авторами использован многоагентный подход, обеспечивающий необходимую связность субъектов принятия решений при достаточной автономности и целеустремлённости каждого интеллектуального агента. Задача классификации поступающих на вход заявок

решается на основе применения коалиции агентов обучения классификатора, функционирующих на основе модифицированного биоинспирированного метода оптимизации поиском кукушки, позволяющего обеспечить максимальное количество итераций локального поиска при достаточном количестве – глобального. Предложена имитационная модель, повышающая эффективность решения задачи построения возможных альтернатив распределения вычислительных ресурсов. После чего централизованное распределение данных альтернативных решений по вычислительным ресурсам осуществляет агент-диспетчер комбинирования классификатора.

1. Постановка задачи. Представим распределенную вычислительную систему в виде P вычислительных подсистем, которые состоят из R элементарных автоматов. Тогда r_k элементарных автоматов входит в состав подсистемы $k \in S = \{1, 2, \dots, P\}$ [5–8]. В отличие от известных постановок подобных задач авторы предлагают закрепить за каждой подсистемой вектор ее атрибутов-признаков $I^S = \{i_1^S, i_2^S, \dots, i_m^S\}$, ($m = 1, \dots, M$), где i – атрибут-признак подсистемы, m – номер атрибута-признака подсистемы; M – общее количество атрибутов-признаков в подсистеме. Данный вектор несет информацию о классе или классах обрабатываемых данной подсистемой заявок, о предельных возможностях параллельной обработки, о пропускной способности каналов подсистемы и т.д. Применение вектора атрибутов-признаков позволяет упростить процесс диспетчеризации заявок заменив диспетчеров каждой подсистемы одним агентом-диспетчером комбинирования классификатора.

Входом агента-диспетчера служат выходы обслуживающих аппаратов-переходов t_w имитационной модели $Q = \langle C, T, H, O \rangle$, где C – множество вершин возможных G классов входных заявок $C = \{c_1, c_2, \dots, c_g\}$, ($g = 1, \dots, G$), T – множество W обслуживающих аппаратов-переходов $T = \{t_1, t_2, \dots, t_w\}$, ($w = 1, \dots, W$), $W \leq 2^G$, задающих множество альтернативных комбинаций параллельной обработки входных заявок для последующего выбора вычислительной подсистемы агентом-диспетчером комбинирования классификатора, H – входные для аппаратов-переходов ребра связей с вершинами классов заявок, O – ребро-связь аппарата-перехода и агента-диспетчера. При этом, каждый класс имеет определенный вектор-набор признаков $K^g = \{k_1^g, k_2^g, \dots, k_z^g\}$, ($z = 1, \dots, Z$), где k – атрибут-признак класса, g – номер класса, z – номер признака, Z – общее количество атрибутов-признаков.

Зададим множество входных заявок $V = \{v_1, v_2, \dots, v_n\}$, ($n = 1, \dots, N$). Каждая заявка v_n из N поступивших на вход системы имеет свой вектор атрибутов-признаков $J^n = \{j_1^n, j_2^n, \dots, j_e^n\}$, ($e = 1, \dots, E$), где j – атрибут-признак заявки, n – целочисленный номер заявки, e – номер признака, E – общее количество атрибутов-признаков заявки. Данный вектор используется для решения задачи классификации заявок с целью определения класса каждой заявки. Параллельную классификацию заявок осуществляет коалиция интеллектуальных агентов обучения классификатора. Кодировка решений и организация механизма поиска реализованы на основе применения модифицированного биоинспирированного метода оптимизации.

Тогда при задании порогового значения D_g для вхождения заявки в класс g , целевая зависимость классификации входных заявок примет вид следующего выражения:

$$\forall v_n, c_g: \sum_{e=1}^E \sum_{z=1}^Z j_e^n \cap k_z^g \geq D_g. \quad (1)$$

Прошедшие процедуру классификации заявки заполняют входные вершины имитационной модели, активизируя при этом определенные альтернативные комбинации, распределение которых по вычислительным подсистемам выполняет агент-диспетчер комбинирования классификатора.

Предложенная постановка задачи отличается от существующих применением векторов признаков для классификации входных заявок и биоинспирированной оптимизации, а также распределенно-централизованной схемы поиска, что позволяет с одной стороны, распараллелить процесс классификации, а с другой, уменьшить размерность задачи принятия решений по диспетчеризации на основе применения одного агента-диспетчера комбинирования классификатора.

2. Разработка метода диспетчеризации потоков параллельных заявок в grid-системах. Эффективность применения агентных моделей для решения сложных вычислительных задач, отличительной особенностью которых является необходимость построения параллельных информационных процессов, обусловлена наличием необходимых качеств автономности и целенаправленности интеллектуальных агентов одновременно с их высокой связностью.

Интеллектуальный агент является сущностью, получающей информацию о состоянии управляемых ею процессов через систему сенсоров и осуществляющей влияние на них через систему актуаторов [10], при этом поведение агентов рационально, т.к. процессы управляемые ими содействуют достижению определённой цели. Наиболее близким аналогом поведению агентов в живой природе является инстинктивное поведение насекомых. Термин «интеллектуальный» не содержит прямого смысла, а подчеркивает более высокий уровень функций управления у агента по сравнению с примитивными триггерными системами автоматического управления.

Наиболее известными типами интеллектуальных агентов являются следующие:

1. Агенты с простым поведением, которые действуют только на основе текущих знаний и происходящих событий. Их агентская функция поддерживает схему условие-действие.
2. Целенаправленные агенты имеют возможность хранить информацию о прецедентах, которые для них желательны. Это позволяет агенту осуществлять выбор альтернативных траекторий поведения для достижения нужной цели. Целенаправленные агенты стремятся к достижению состояния достигнутой цели.
3. Практичные агенты также способны различать, насколько желанно для них текущее состояние. Такая оценка достигается на основе применения «функции полезности», которая проецирует множество состояний на множество критериев полезности состояний.
4. Обучающиеся агенты независимы и способны к обучению и адаптации к изменяющимся условиям.

В данной работе использован целенаправленный тип интеллектуальных агентов. Коалиция агентов моделирует процесс рассуждений референтной группы в процессе решения задачи классификации потоков параллельных заявок при диспетчеризации в grid-системах. В исследовании решена задача создания классификаторов в виде агентов с заданными условными атрибутами.

Зададим два типа агентов: для организации распределенной классификации входных потоков заявок будут использованы агенты обучения классификатора (АОК); для централизованного распределения вычислительных ресурсов между альтернативными наборами заявок – агент-диспетчер комбинирования классификатора (АДКК). Предлагаемый метод диспетчеризации потоков параллельных заявок включает в себя три основных этапа (рис. 1).

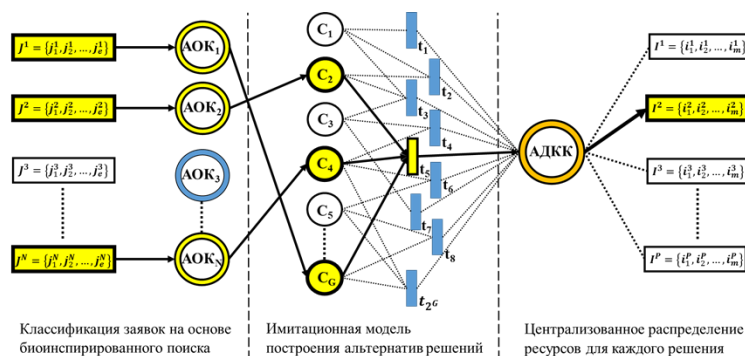


Рис. 1. Структура метода диспетчеризации потоков параллельных заявок

На этапе классификации потоков входных заявок реализацию модифицированного биоинспирированного метода оптимизации поиском кукушки осуществляет коалиция агентов обучения классификатора. Опишем разработку алгоритма классификации потоков входных заявок. Данный алгоритм составляет 1-ый этап реализации метода диспетчеризации потоков параллельных заявок.

В первую очередь формируется коалиция интеллектуальных АОК, размер которой соответствует максимально возможному числу параллельных заявок на входе системы. Также создается агент-диспетчер комбинирования классификатора. Агент обучения классификатора проверяет построенную модель классификации на основе встроенного биоинспирированного метода и определяет наличие элементов, принадлежащих разным классам. Агент-диспетчер комбинирования классификатора объединяет результаты проверок, проведенных агентами обучения классификатора, вносит коррективы в распределение вычислительных ресурсов на основе полученных имитационной моделью альтернатив решений.

Каждый АОК в рамках назначенной ему части пространства решений создает независимую модель классификации с учетом заданных системно значимых признаков, чем предсказывает наличие проблемной области. На этом заканчивается этап предсказания классификации, после чего АДКК производит разбиение обозначенных областей, которые предсказаны агентами обучения классификатора. АОК анализируют характеристики и присваивают определенную метку экземплярам, несоответствующим предварительной классификации, а также отправляют результаты полученной оценки с меткой класса АДКК. АДКК меняет статус класса и проводит необходимые разбиения проблемных областей.

Атрибуты каждой поступившей на вход заявки v_n присваиваются одному из коалиции АОК. В дальнейшем происходит запуск модифицированного биоинспирированного метода оптимизации поиском кукушки, позволяющего проверить атрибуты поступающих заявок на эквивалентность атрибутам имеющихся классов. По результатам работы биоинспирированного метода полученные заявки распределяются по классам.

Метод, инспирированный поведением кукушек в процессе гнездового паразитизма, отличается от известных аналогов более эффективным способом определения траектории передвижения агента в пространстве поиска. Кукушка свои передвижения осуществляет в виде полетов Леви (Levy flights), реализуемых по следующей канонической формуле (2) с тензорным произведением вектора размера шагов и случайного вектора независимых вещественных чисел.

$$X' = X + U \otimes L_{|X|}(\lambda), \quad (2)$$

где $U = (u_a, a \in [1: |X|])$ – вектор размера шагов по соответствующим компонентам вектора координат местонахождения агента X ; $L_{|X|}(\lambda) = (|X| \times \mathbf{1})$ – вектор независимых вещественных случайных чисел, распределенных по закону Леви [11]. Данный способ обеспечивает высокую вероятность малых и достаточную вероятность больших перелетов агента-кукушки (рис. 2), что позволяет интенсифицировать поиск в локальных областях, с одной стороны, и реализовать выход из локальных оптимумов, с другой.

В отличие от канонического метода в данном модифицированном авторы предлагают смоделировать ситуацию, когда в каждом гнезде, куда кукушка прилетает помимо оригинального яйца хозяйки гнезда (анализируемого класса), уже есть чужое подложенное (случайным образом присвоенная входная заявка), которое имеет определенную степень сходства с оригинальным. Задача кукушки – принести такое «яйцо» (не прошедшую процедуру классификации заявку) в случайно выбранную вершину множества классов, степень сходства которого (атрибутов заявки) с оригиналом (атрибутами класса) будет выше, чем у ранее подброшенного.

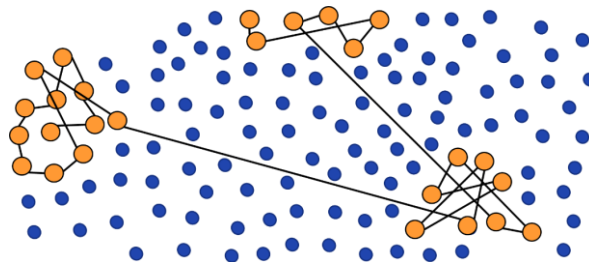


Рис. 2. Иллюстрация способа определения траектории передвижения агента в пространстве поиска

Таким образом, в процессе классификации происходит сравнение значений выражения (1) для случайных и сформированных в процессе поиска пар атрибутов заявка-класс. В процессе решения с заданной вероятностью (свободный параметр метода) периодически удаляется определенное число наихудших полученных решений. При исследовании большого пространства поиска, по мнению авторов, подобные (рис. 2) случайные блуждания (2) являются наиболее эффективными по сравнению с другими более упорядоченными передвижениями агентов, как например, в методе бактериальной оптимизации [12–24].

Полученные результаты классификации на 2-ом этапе выполнения метода диспетчеризации потоков параллельных заявок поступают на вход имитационной модели, построенной на основе применения аппарата сетей Петри (рис. 3). Данная модель на основе комбинирования своих входов G создает полный спектр 2^G возможных альтернативных решений распределения потоков параллельных заявок между имеющимися ресурсами – вычислительными подсистемами. За соблюдение условий выбора той или иной альтернативы отвечает множество переходов T имитационной модели. Заполнение необходимой для открытия перехода t_k комбинации классифицированных заявок происходит по результатам работы модифицированного биоинспирированного метода на 1-ом этапе решения задачи диспетчеризации.

Для повышения качества процесса диспетчеризации предусмотрены процедуры дублирования неклассифицированных заявок, позволяющие интенсифицировать поиск совпадений в векторах атрибутов-признаков. Также предусмотрены резервные траектории диспетчеризации необходимые для обработки прецедентов появления на входах заявок с абсолютным приоритетом. Описанные две дополни-

тельные возможности метода реализованы на основе применения отдельных агентов-диспетчеров комбинирования классификатора $АДКК_2$ и $АДКК_3$ (рис. 3). Каждое очередное срабатывание одного из агентов-диспетчеров через определенный для него переход активизирует работу модуля поддержки принятия решений (МПП).

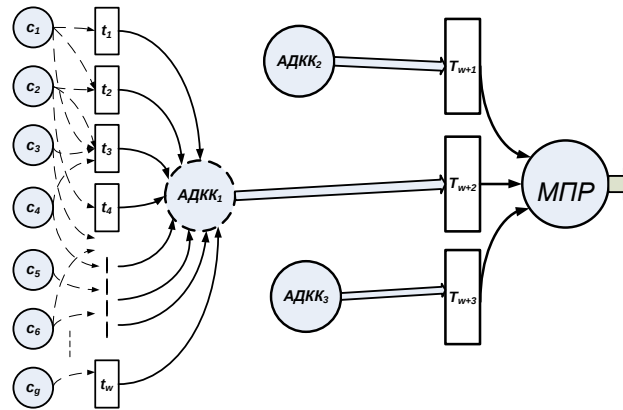


Рис. 3. Имитационная модель построения альтернативных решений

На 3-ем этапе разработанного метода диспетчеризации происходит централизованное распределение $АДКК_1$ прошедших имитационную модель альтернативных комбинаций решений между вычислительными подсистемами P соответствующими их атрибутам.

3. Экспериментальные исследования. Для подтверждения эффективности разработанного метода диспетчеризации потоков параллельных заявок был создан программный модуль, включающий в себя функции поэтапного решения задач классификации входных заявок на основе биоинспирированного поиска, имитации построения альтернативных вариантов решений и централизованного распределения вычислительных ресурсов. В качестве значений атрибутов применены усредненные абстрактные показатели, полученные случайным образом с заданной вероятностью.

Необходимо отметить, что наибольшее влияние на качество и скорость работы представленного метода оказывает процедура классификации потоков входных заявок. В результате проведенных комплексных исследований получена временная сложность предложенного метода (рис. 4). Временная сложность в рассмотренном примере составила $O(n^2)$, где n – количество анализируемых методом входных данных, которым в данном случае является число входных заявок. В представленной модификации метода биоинспирированной оптимизации поиском кукушки предусмотрено два вложенных цикла прохождения по множествам атрибутов-признаков входных заявок J^n : первый, для предварительного распределения заявок по классам, а второй, для увеличения значений фитнес-функций эквивалентности и перераспределения заявок. Таким образом, максимально возможное количество операций зависит от числа вершин входных заявок N как $n * n$.

Экспериментальные исследования разработанного метода проводились на вычислительной системе с производительностью порядка 10^9 (гигафлопс) операций в секунду. В этом случае, например, при $n = 50000$ вершин, время выполнения алгоритма не превысит 2,5 секунды, при $n = 200000$ вершин – 40 секунд, при $n = 500000$ вершин – 250 секунд, что наглядно иллюстрирует представленная на рисунке 4 полученная зависимость времени выполнения предложенного метода от числа входных вершин.

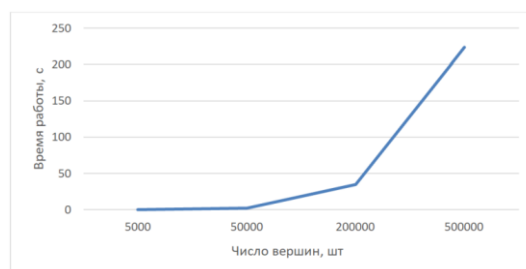


Рис. 4. Временная сложность в зависимости от числа входных вершин

Исследование скорости работы метода биоинспирированной оптимизации поиском кукушки для решения задачи классификации входных заявок в процессе диспетчеризации проводилось в сравнении с каноническим бактериальным методом и методом роя частиц (табл. 1).

Таблица 1

Сравнение времени работы алгоритмов

Размерность задачи (количество вершин, шт)	Модифицированный метод поиска кукушки (время работы, с)	Канонический бактериальный метод (время работы, с)	Метод роя частиц (время работы, с)
50000	1.91	2.28	2.37
200000	38.54	47.05	44.87
500000	229.18	313.20	259.33

Из анализа приведенных графика временной сложности и таблицы сравнения времени работы методов видно, что предложенный модифицированный метод показал лучший результат по сравнению с каноническим бактериальным методом и методом роя частиц. Отметим, что при относительно небольших размерностях задачи (от 50000 до 200000 вершин) время работы оцениваемых методов соизмеримо: от 1,91–38,54 секунды в предлагаемой модификации метода кукушкиного поиска до 2,28–47,05 и 2,37–44,87 секунды соответственно в бактериальном методе и методе роя частиц. Наибольший прирост скорости предложенные процедуры интенсификации локального поиска при достаточном количестве глобальных перемещений для выхода из локальных оптимумов позволили достичь при решении задач с размерностью не менее 500000 вершин. Из табл. 1 видно, что в этом случае, если принять за 100 % время работы наиболее медленного последовательного бактериального метода, тогда метод роя частиц затратил на решение этой же задачи 83 % времени от максимального, а предложенный модифицированный – 73 %. Эффективность предложенного модифицированного метода повышена за счет улучшения способа определения траектории передвижения агента в пространстве поиска, позволяющего обеспечить максимальное количество итераций локального поиска при достаточном количестве – глобального.

Заключение. В данной работе представлена разработка метода диспетчеризации потоков параллельных заявок в grid-системах. Задача диспетчеризации решена на основе комплексного применения коалиции интеллектуальных агентов и событийной имитационной модели. Использована комбинированная распределенно-централизованная схема поиска решений при диспетчеризации, что позволило уменьшить размерность задачи. Классификация поступающих на вход заявок проводится на основе применения модифицированного биоинспирированного метода

оптимизации поиском кукушки. Совместное использование коалиции интеллектуальных агентов и биоинспирированного метода обеспечивает беспрецедентный параллелизм вычислений, а последующее определение путей обработки классифицированных заявок на основе имитационной модели позволяет формировать наборы альтернативных решений.

Для оценки эффективности предложенного метода разработан программный продукт и проведены эксперименты с разным количеством поступающих на вход заявок. Проведенные экспериментальные исследования подтвердили эффективность разработки в сравнении с каноническим бактериальным методом и методом роя частиц. Полученные количественные оценки подтверждают снижение времени решения задачи диспетчеризации потоков параллельных заявок в grid-системах при помощи разработанного модифицированного метода не менее чем на 10 % по отношению в результатам работы канонических методов роя частиц и бактериальной оптимизации. При относительно небольших размерностях задачи (от 50000 до 200000 вершин) время работы методов соизмеримо, наибольший прирост скорости решения достигается в задачах с размерностью не менее 500000 вершин. Временная сложность в рассмотренных примерах составила $O(n^2)$, где n – количество анализируемых методом входных данных, которым в данном случае является число входных заявок.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Magoulès F., Nguyen T., Yu L. Grid resource management: toward virtual and services compliant grid computing, *Numerical analysis and scientific computing*. CRC Press, UK, 2009.
2. Magoulès F. (ed.). Fundamentals of grid computing: theory, algorithms and technologies, *Numerical analysis and scientific computing*. CRC Press, UK, 2010.
3. Patel S. Survey Report of Job Scheduler on Grids, *International Journal of Emerging Research in Management & Technology*, 2013, No. 2 (4), pp. 115-125.
4. Li M., Baker M. The grid: core technologies. John Wiley & Sons Ltd, England, 2005, 452 p.
5. Saak A.E., Kureichik V.V., Kravchenko Y.A. Scheduling quality of precise form sets which consist of tasks of circular type in GRID systems, *Journal of Physics: Conference Series*, 2018, 1015 (4).
6. Saak A.E., Kureichik V.V., Lezhebokov A.A. Scheduling of parabolic-type tasks arrays in GRID systems, *Advances in Intelligent Systems and Computing*, 2017, pp. 292-298.
7. Saak A., Kureichik V., Kravchenko Y. To scheduling quality of sets of precise form which consist of tasks of circular and hyperbolic type in grid systems, *Advances in Intelligent Systems and Computing*, 2016, pp. 157-166.
8. Saak A.E., Kureichik V.V., Kuliev E.V. Ring algorithms for scheduling in grid systems, *Advances in Intelligent Systems and Computing*, 2015, pp. 201-209.
9. Kravchenko Y.A., Kravchenko D.Y., Kursitys I.O. Architecture and method of integrating information and knowledge on the basis of the ontological structure, *Advances in Intelligent Systems and Computing. 1st International Conference of Artificial Intelligence, Medical Engineering, and Education, AIMEE 2017. Moscow: 2018, Vol. 658*, pp. 93-103.
10. Kravchenko Y.A., Kureichik V.V. Knowledge management based on multi-agent simulation in informational systems, *Conference proceedings. 8th IEEE International Conference "Application of Information and Communication Technologies – AICT 2014". – 15-17 October 2014, Astana, Kazakhstan*, pp. 264-267.
11. Payne R.B., Sorenson M.D., and Klitz K. The Cuckoos, Oxford University Press, 2005.
12. Shukran M.A.M., Chung Y.Y., Yeh W.C., Wahid N., and Zaidi A.M.A. Artificial Bee Colony based Data Mining Algorithms for Classification Tasks // *Mod. Appl. Sci.*, 2011, Vol. 5, pp. 217-231.
13. Martens D., De Backer M., Haesen R., Vanthienen J., Snoeck M. and Baesens B. Classification With Ant Colony Optimization // *IEEE Transactions on Evolutionary Computation*, 2007, Vol. 11, No. 5, pp. 651-665.
14. Falco I.D., Cioppa A.D., and Tarantino E. Evaluation of particle swarm optimization effectiveness in classification // *LNAI3849*, 2006, pp. 164-171.
15. Soliman O.S. and Adly A. Bio-inspired algorithm for classification association rules, *8th International Conference on Informatics and Systems (INFOS)*, Cairo, 2012, pp. 154-160.

16. Bova V., Zaporozhets D., and Kureichik V. Integration and processing of problem-oriented knowledge based on evolutionary procedures, *Advances in Intelligent Systems and Computing*, 2016, Vol. 450, pp. 239-249.
17. Semenova A.V. and Kureichik V.M. Ensemble of classifiers for ontology enrichment, *Journal of Physics: Conference Series*, 2018, Vol. 1015, Issue 3, article id. 032123.
18. Kureychik V.M. Overview and problem state of ontology models development, *9th International Conference on Application of Information and Communication Technologies, AICT 2015 - Proceedings 9*, 2015, pp. 558-564.
19. Semenova A.V. and Kureychik V.M. Application of swarm intelligence for domain ontology alignment, *Proceedings of the First International Scientific Conference "Intelligent Information Technologies for Industry" (ITI'16)*, 2016, Vol. 1, pp. 261-270.
20. Bova V., Kureichik V. and Zaruba D. Heuristic approach to model of corporate knowledge construction in information and analytical systems, *2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT), Baku, 2016*, pp. 1-5.
21. Kureichik V., Zaporozhets D., and Zaruba D. Generation of bioinspired search procedures for optimization problems, *Application of Information and Communication Technologies, AICT 2016 - Conference Proceedings*, 2016, Vol. 10.
22. Kar A.K. Bio inspired computing - A review of algorithms and scope of applications, *Expert Systems with Applications*, 2016, Vol. 59, pp. 20-32.
23. Zaporozhets D., Zaruba D., and Kulieva N. Parallel approach for bioinspired algorithms, *Journal of Physics: Conference Series Ser. "International Conference Information Technologies in Business and Industry 2018 - Enterprise Information Systems"*, 2018.
24. 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.

Статью рекомендовал к опубликованию к.т.н. С.Г. Буланов.

Кравченко Даниил Юрьевич – Южный федеральный университет; e-mail: kravchenkodanil122@gmail.com; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371651; кафедра систем автоматизированного проектирования; студент.

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

Курейчик Владимир Викторович – e-mail: vkur@sfedu.ru; кафедра систем автоматизированного проектирования; зав. кафедрой.

Саак Андрей Эрнестович – e-mail: aesaak@sfedu.ru; 347922, г. Таганрог, ул. Чехова, 22; тел.: +78634393373; кафедра государственного и муниципального управления зав. кафедрой.

Kravchenko Daniil Yurievich – Southern Federal University; e-mail: kravchenkodanil122@gmail.com; 44, Nekrasovskiy, Taganrog, 347928, Russia; phone: +78634371651; the department of computer aided design; student.

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

Kureichik Vladimir Viktorovich – e-mail: vkur@sfedu.ru; the department of computer aided design; head of department.

Saak Andrey Ernestovich – e-mail: aesaak@sfedu.ru; 22, Chehov street, Taganrog, 347922, Russia; phone: +78634393373; the department of state and municipal management; head of department.

Раздел II. Вычислительные и информационно-управляющие системы

УДК 681.51

DOI 10.18522/2311-3103-2020-2-89-101

А.С. Синицын

СИНТЕЗ СИНЕРГЕТИЧЕСКОЙ СИСТЕМЫ РАЗРЫВНОГО УПРАВЛЕНИЯ АКТИВНОЙ ПОДВЕСКОЙ АВТОМОБИЛЯ*

Рассматривается проблема проектирования активных систем подвески транспортных средств, в которых исполнительный механизм (ИМ) не является идеальным и подвержен влиянию гистерезиса и мертвой зоны. Целью данной работы является синтез системы управления, позволяющей уменьшить влияние гистерезиса и мертвой зоны на эффективность работы системы адаптивной подвески. Параметры системы подобные гистерезису требуют значительных усилий для их идентификации и более того могут в широких пределах изменяться на протяжении жизненного цикла системы. Таким образом учесть гистерезис при синтезе системы управления является весьма трудной задачей, равно как и построение наблюдателей. Решением данной проблемы может быть применение систем разрывного управления, которые в известной мере являются робастными к параметрическим и структурным изменениям в объекте управления. Существующие подходы к синтезу систем разрывного управления опираются на линейную или линеаризованную модель объекта управления. Таким образом эффективность подобного рода систем может существенно отличаться при работе регулятора в составе реального, нелинейного объекта управления. Предлагаемая система разрывного управления позволяет снизить чувствительность системы к возмущениям, обусловленным неидеальностью ИМ, а так-же учитывает нелинейную структуру объекта управления. Эффективность замкнутой системы исследована на динамической модели построенной при помощи пакета Simulink. Предлагаемая система разрывного управления сравнивается с адаптивным синергетическим регулятором. В качестве возмущающего воздействия выбрано дорожное покрытие класса С по классификации ISO 8608. Для оценки эффективности предлагаемой системы оцениваются следующие параметры: взвешенное ускорение поддрессоренной массы; относительное перемещение амортизатора и сила реакции шины. Для каждого параметра определяется среднеквадратическое и максимальное значение. Результаты численного моделирования позволяют сделать вывод что применение систем разрывного управления позволяет улучшить следующие показатели эффективности адаптивной подвески: более чем в два раза уменьшить максимальное значение взвешенного ускорения поддрессоренной массы и более чем на 20 % уменьшить максимальную нагрузку на шину.

Управляемая подвеска; нелинейное управление; адаптация; синергетическая теория управления; разрывное управление.

A.S. Sinitsyn

SYNERGETIC SYNTHESIS OF SLIDING MODE CONTROL FOR VEHICLE'S ACTIVE SUSPENSION SYSTEM

This article discusses the problem of designing vehicle's active suspension systems in which the actuator is not ideal and is a subject to the influence of hysteresis and dead zone. The main goal of this work is to synthesize a control system that reduces the influence of hysteresis and dead

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-08-00366.

zone on the efficiency of the adaptive suspension system. System parameters like hysteresis require significant efforts to identify them and, moreover, can vary widely over the life cycle of the system. Thus, it is very difficult to take into account hysteresis in the synthesis of the control system, as well as the construction of observers. The solution to this problem is use of sliding control systems, which to a certain extent are robust to parametric and structural changes in the control object. Existing approaches to the synthesis of sliding control systems are based on a linear or linearized model of the control object. Thus, the effectiveness of such systems can vary significantly when the regulator operates as part of a real, non-linear control object. The proposed sliding mode control allows to reducing sensitivity of the system to disturbances due to imperfect actuator, and also takes into account the nonlinear structure of the control object. The efficiency of a closed system is investigated on a dynamic model built in Simulink package. The proposed controller is compared with an adaptive synergetic regulator. Road of class C (according to ISO 8608 classification) was selected as a disturbance. To investigate the effectiveness of the proposed control system, the following parameters are evaluated: weighted acceleration of the sprung mass; relative motion of suspension and tire reaction force. The RMS and maximum values are calculated for each parameter. The results of numerical simulations allow to conclude that the use of sliding control systems can improve the following adaptive suspension performance indicators: reduce the maximum value of the weighted acceleration of the sprung mass by more than two times and reduce the maximum load on the tire by more than 20 %.

Active suspension; nonlinear control; adaptation; synergetic control theory; sliding mode control.

Введение. В настоящее время наблюдается большой интерес к исследованиям возможностей повышения комфортабельности движения легковых автомобилей за счет применения системы активных подвесок. Разработка систем управления активной подвеской является сложной задачей, поскольку имеют место значительные неопределенности параметров элементов подвески, поддрессоренной массы возмущающих воздействий со стороны дорожного полотна. При этом следует отметить, что модели пружин и демпферов являются нелинейными, что существенно усложняет проектирование эффективных систем управления такими системами с применением традиционных подходов теории управления. Различные подходы к управлению, такие как оптимальное управление [1], обход интегратора [2], адаптивное управление [3], нейронные сети [4], безмодельное управление [5], нелинейный компенсатор возмущений [6] – это лишь некоторые подходы, применение которых часто встречается в публикациях. В большинстве таких систем управления привод считается идеальным, что не соответствует реальности, так как практика показывает, что характеристики приводов обладают мертвой зоной, гистерезисом, насыщением, а также в таких системах присутствуют временные задержки. В отсутствие корректирующих мер эффективность замкнутой системы управления существенно снижается. Данная работа посвящена исследованию влияния мертвой зоны и гистерезиса в ИМ систем активной подвески.

Решению данной задачи посвящены множество результатов исследования с применением теории адаптивного управления [7, 8] и методов обратного управления [9], а также их комбинации. В работах [10, 11] было предложено снизить влияние мертвой зоны за счет использования её обратной модели. В работах [12, 13] предложен адаптивный регулятор с эталонной моделью, однако мертвая зона также компенсируется за счет обратной модели. Обратная модель мертвой зоны строится путем представления неизвестной мертвой зоны ограниченным возмущением, изменяющимся во времени. Обратные модели работают лучше всего, когда модель нелинейности привода полностью известна и идентифицирована.

Наличие же в системе гистерезиса может дестабилизировать систему [11], следовательно, наличие корректирующих звеньев в системе управления становится обязательным условием. Влияние гистерезиса на механические системы рассмотрено в работах [14, 15]. Хотя проблема неидеального привода в современной

литературе в основном решается с использованием нескольких методов линейного или нелинейного управления, существует также подход с использованием наблюдателей возмущения. Несмотря на то, что применение наблюдателя возмущения может с удивительной простотой компенсировать неидеальность привода, для корректной работы наблюдателя необходима модель возмущения и точные значения её параметров, получение которых связано со значительными трудностями.

Перспективным направлением решения указанной проблемы является применение систем управления с переменной структурой (sliding mode control). Наличие в системе поверхности переключения при определенных условиях обеспечивает высокую робастность системы к параметрическим и внешним возмущениям. Как было определено ранее объект управления является существенно нелинейным, следовательно, при синтезе законов управления целесообразно воспользоваться синергетической теорией управления (СТУ) [16].

Отличительными особенностями предлагаемой системы управления являются:

- ♦ индивидуальная оценка неопределенных параметров не требуется;
- ♦ не требуется знание ширины или наклона мертвой зоны, а также формы или амплитуды гистерезиса;
- ♦ знание профиля дороги не требуется, благодаря чему устраняется необходимость в датчиках предварительного просмотра дороги;
- ♦ контроллер не содержит динамических звеньев, требующих вычислений в реальном времени, что значительно снижает сложность контроллера.

Модель объекта управления. Модель одной опоры активной подвески автомобиля представлена на рис. 1. Данная модель активно использовалась в работах [7–11]. На рис. 1 обозначениям соответствуют: m_s – подрессоренная (изолируемая) масса, представляет собой массу кузова с полезной нагрузкой; m_u – неподрессоренная масса, представляет собой массу колеса в сборе с амортизатором; F_d и F_s обозначают силы, производимые амортизатором и пружиной соответственно, F_t , F_b – силы упругости и демпфирования шины.

Перемещение подрессоренной массы обозначено как z_s , неподрессоренной – z_u , а «перемещение» дорожного покрытия – z_r . Управляющее воздействие обозначено как u .

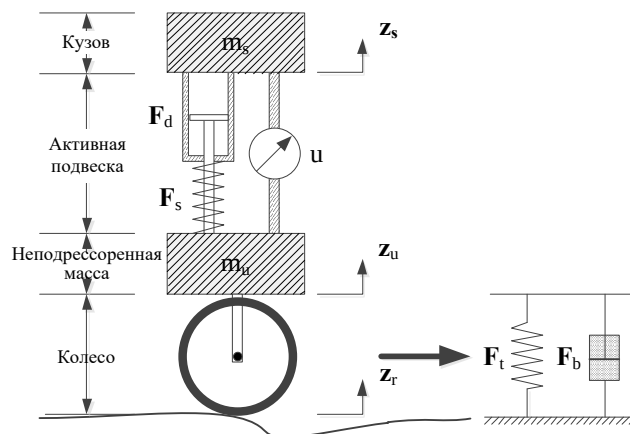


Рис. 1. Модель одной опоры активной подвески автомобиля

Динамика перемещения кузова и неподрессоренной массы описывается следующей системой дифференциальных уравнений:

$$\begin{aligned} m_s \ddot{z}_s + F_d(\dot{z}_s, \dot{z}_u, t) + F_s(z_s, z_u, t) &= N(u), \\ m_u \ddot{z}_u - F_d(\dot{z}_s, \dot{z}_u, t) - F_s(z_s, z_u, t) + F_t(z_u, z_r, t) + F_b(\dot{z}_u, \dot{z}_r, t) &= -N(u), \end{aligned} \quad (1)$$

где $N(u)$ – сила, производимая исполнительным механизмом. Силы, производимые нелинейно упругой пружиной, кусочно-линейным амортизатором и шиной заданы в следующем виде:

$$\begin{aligned} F_s(z_s, z_u, t) &= k_s(z_s - z_u) + k_{sn}(z_s - z_u)^3, \\ F_d(\dot{z}_s, \dot{z}_u, t) &= \begin{cases} b_e(\dot{z}_s - \dot{z}_u) \\ b_c(\dot{z}_s - \dot{z}_u) \end{cases}, \\ F_t(z_u, z_r, t) &= k_f(z_u - z_r), \\ F_b(\dot{z}_u, \dot{z}_r, t) &= b_f(\dot{z}_u - \dot{z}_r), \end{aligned} \quad (2)$$

где k_s и k_{sn} – линейный и нелинейный коэффициенты упругости; b_e и b_c – коэффициенты демпфирования амортизатора на ходе расширения и сжатия соответственно, k_f и b_f – коэффициенты упругости и демпфирования шины.

Определим переменные состояния:

$$x_1 = z_s, x_2 = \dot{z}_s, x_3 = z_u, x_4 = \dot{z}_u, \quad (3)$$

таким образом, дифференциальные уравнения (1) могут быть трансформированы в модель, записанную в пространстве состояний:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= \frac{1}{m_s} (-F_d(\dot{z}_s, \dot{z}_u, t) - F_s(z_s, z_u, t) + N(u)), \\ \dot{x}_3 &= x_4, \\ \dot{x}_4 &= \frac{1}{m_u} (F_d(\dot{z}_s, \dot{z}_u, t) + F_s(z_s, z_u, t) - F_t(z_u, z_r, t) - F_b(\dot{z}_u, \dot{z}_r, t) - N(u)). \end{aligned} \quad (4)$$

Для краткости записи далее будут использованы сокращенные обозначения сил F_s , F_d , F_t , F_b без явного указания аргументов.

Неидеальный привод. Привод, обозначенный как $N(u)$ в уравнениях (4), является неидеальным приводом. Предполагается, что на него влияет мертвая зона или гистерезис. Сила, производимая ИМ, обозначается как $N(u)$. Входное воздействие обозначено как u . В работах [10, 17] выход ИМ при наличии мертвой зоны описывается уравнением:

$$N(u) = p_s u + d_d, \quad (5)$$

где

$$d_d = \begin{cases} -mb_r, & \text{если } u \geq b_r \\ -mu, & \text{если } b_l < u < b_r \\ -mb_l, & \text{если } u \leq b_l \end{cases}, \quad (6)$$

где p_s – псевдо-наклон мертвой зоны, $b_r > 0$, $b_l < 0$ – границы мертвой зоны. При наличии гистерезиса выход привода может быть задан уравнением [14, 15, 18]:

$$N(u) = \mu_1 u + \mu_2 d_h, \quad (7)$$

где $\mu_1 > 0$, $\mu_2 > 0$ – коэффициенты жесткости, d_h – вспомогательная переменная, описываемая дифференциальным уравнением:

$$\dot{d}_h = \dot{u} - \gamma |\dot{u}| |d_h|^{n-1} - \alpha \dot{u} |d_h|^n, d_h(0) = 0, \quad (8)$$

где γ , n , α – положительные коэффициенты, определяющие форму и ширину петли гистерезиса. Коэффициенты, фигурирующие в уравнениях (5)–(8) могут быть оценены лишь приблизительно и более того, могут меняться в процессе функционирования системы, что существенно осложняет синтез точного наблюдателя.

Критерии эффективности. Пространство между подпрессоренной и неподпрессоренной массой называется рабочим. Относительное перемещение амортизатора подвески, определяется выражением:

$$\xi = \frac{x_1 - x_3}{x_r}, \quad (9)$$

где x_r – максимально допустимое перемещение.

Относительная реакция шины, обозначаемая как ψ_t , определяется как отношение динамического усилия шины к статическому:

$$\psi_t = \frac{F_t(z_u, z_r, t) - F_b(\dot{z}_u, \dot{z}_r, t)}{(m_s + m_{us})g}. \quad (10)$$

Относительное перемещение амортизатора и относительная реакция шины выступают в роли критериев безопасности. Обе величины должны быть по модулю меньше единицы, чтобы гарантировать удержание дороги.

В ISO 2631 [19] установлено, что организм человека более чувствителен к ускорению в диапазоне 4–8 Гц. Наиболее простой способ найти средневзвешенное среднеквадратичное ускорение – фильтрация сигнала ускорения подпрессоренной массы через полосовые фильтры. Фильтр, соответствующий стандарту ISO 2631 приведён в работах [20, 21] и имеет вид:

$$W(s) = \frac{87,72s^4 + 1138s^3 + 11336s^2 + 5453s + 5509}{s^5 + 92,6854s^4 + 2649,83s^3 + 25969s^2 + 81057s + 79783}. \quad (11)$$

Цели управления. Цель системы управления состоит в том, чтобы минимизировать $\dot{x}_2$ в условиях неопределенностей в элементах подвески и внешних возмущений. При этом желательно удерживать кузов в заданном положении при выполнении условий $|\psi_t| < 1$, $|\xi| < 1$.

Синтез адаптивного закона управления. Приступая к синтезу системы управления методом аналитического конструирования агрегированных регуляторов (АКАР), в первую очередь необходимо определить форму инварианта, погружаемого в систему. Инвариантное многообразие должно отражать требования к закону управления и при этом не должно содержать переменных состояния, которые недоступны прямому измерению, в данном случае такой переменной является z_r – профиль дороги. Измерение данной переменной не представляется возможным, а построение точного наблюдателя осложнено непредсказуемостью данного воздействия. Исходя из требований к системе управления, сформируем целевое многообразие в следующем виде:

$$x_2 + k_1 x_1 = 0, \quad (12)$$

где k_1 – ненулевая положительная константа.

Определившись с формой инвариантного многообразия, перейдем к синтезу закона управления, гарантирующего его погружение в управляемый объект. Нам известно, что в исходной системе имеется параметрическое возмущение, которое выражается в наличии гистерезиса и мертвой зоны в ИМ. Для обеспечения адаптации регулятора к возмущениям введем в регулятор одно интегрирующее звено вида:

$$\dot{v} = c_1(x_2 + k_1 x_1), \quad (13)$$

где c_1 – положительная константа.

Для синтеза закона управления запишем макропеременную следующего вида:

$$\psi = x_2 + k_1 x_1 + v. \quad (14)$$

Составим дифференциальное уравнение, описывающее движение изображающей точки системы (4) из исходного пространства состояний к инвариантному многообразию (12) [22, 23]:

$$T\dot{\psi}(t) + \psi = 0, \quad (15)$$

где T – положительная константа. Следуя процедуре метода АКАР, продифференцируем уравнение (14) по времени с учетом системы дифференциальных уравнений (4):

$$\dot{\psi}(t) = \frac{-F_d - F_s + N(u)}{m_s} + c_1(x_2 + k_1 x_1) + k_1 x_2. \quad (16)$$

Искомый закон управления является решением (15) относительно $N(u)$:

$$N(u) = -\frac{m_s(((1+T)k_1 + 1 + c_1 T)x_2 + c_1 T k_1 + v)}{T} + F_d + F_s. \quad (17)$$

Характерной особенностью найденного закона управления (13), (17) является адаптивность к любым полиномиальным возмущениям нулевого порядка. Эта особенность гарантирует, что инвариантное отношение (12) будет выполняться даже в условиях неточного определения параметров, страдать при этом будет лишь качество переходного процесса системы из исходного состояния к аттрактору.

Синтез системы разрывного управления. Согласно процедуре, изложенной в [22, 23] формируем поверхность переключения, отражающую технологическую задачу системы управления:

$$s = x_2 + k_1 x_1. \quad (18)$$

Введем агрегированную макропеременную следующего вида:

$$\psi = x_1^2 \beta_1 + \beta_2 |s|, \quad (19)$$

где β_1, β_2 – постоянные коэффициенты. Продифференцировав (19) по времени

$$\dot{\psi} = 2\beta_1 x_1 \dot{x}_1 + \beta_2 \text{sign}(s) \dot{s} \quad (20)$$

и воспользовавшись функциональным уравнением (15) можно получить управляющее воздействие, гарантирующее попадание изображающей точки системы (4) на поверхность переключения $s = 0$:

$$N(u) = \frac{T\beta_2(-m_s k_1 x_2 + F_s + F_d) \text{sign}(s) - m_s(\beta_2 |x_2 + k_1 x_1| + \beta_1 x_1(x_1 + 0,5Tx_2))}{T\beta_2 \text{sign}(s)}. \quad (21)$$

Условием возникновения скользящего режима на поверхности $s = 0$ является [22]:

$$s \cdot \dot{s} < 0. \quad (22)$$

Разрешив правую часть уравнения (20) относительно $\dot{s}$ и подставив результат в (22) запишем условия возникновения скользящего режима:

$$\frac{-\beta_1 - \beta_2 + 2\beta_1 kT}{T\beta_2} < 0. \quad (23)$$

Исследование синтезированной системы управления. Проведем исследование эффективности полученных результатов средствами численного моделирования замкнутой системы. Для наглядной демонстрации отличительных особенностей синтезированного разрывного управления (21) проводится его сравнение с нелинейным адаптивным законом (17). Параметры модели одной амортизационной опоры системы активной подвески автомобиля приведены в табл. 1, а параметры исследуемых законов управления приведены в табл. 2.

Таблица 1

Значения параметров модели активного амортизатора

Параметр	Значение	Параметр	Значение
m_s	600 кг	k_f	200 000 Н/м
m_u	60 кг	b_f	1 000 Н·с/м
k_s	18 000 Н/м	b_e	2 500 Н·с/м
k_{sn}	1 000 Н/м	b_c	2 200 Н·с/м
b_l	-3	b_r	2
γ	1	α	1
n	1	μ_2	1

Таблица 2

Значения параметров исследуемых систем управления

Параметр	Значение	Параметр	Значение
T	0,1	c_1	0,1
k_1	50	β_1	1000
β_2	$1e^{-6}$		

При моделировании будем полагать, что возмущающее воздействие представляет собой дорожное покрытие класса С по классификации ISO 8608 [24]:

$$\dot{z}_r = -\delta V z_r + \omega, \quad (24)$$

где $V = 30$ м/с – скорость автомобиля, ω – случайное возмущающее воздействие со спектральной плотностью мощности Φ_0 . Для дорожного покрытия С класса стандартом ISO 8608 определены следующие параметры: $\Phi_0 = 16 \times 10^{-6}$ м³, $\delta = 0,127$. Результаты моделирования замкнутой системы с различными системами управления при наличии в системе неучтенной зоны нечувствительности представлены на рис. 2–5.

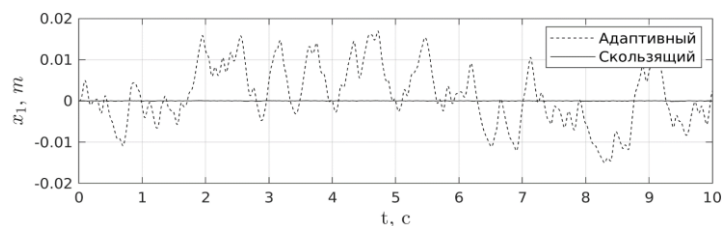


Рис. 2. Перемещение поддресоренной массы в случае наличия мертвой зоны

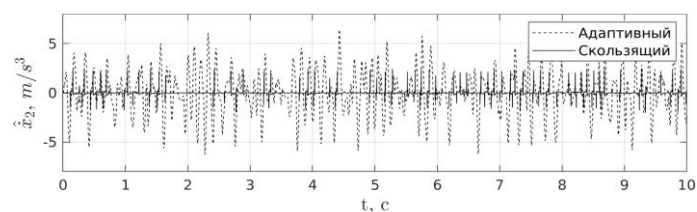


Рис. 3. Взвешенная оценка ускорения поддресоренной массы в случае наличия мертвой зоны

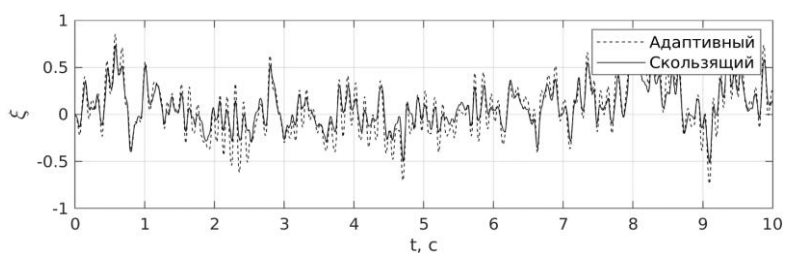


Рис. 4. Относительное перемещение амортизатора в случае наличия мертвой зоны

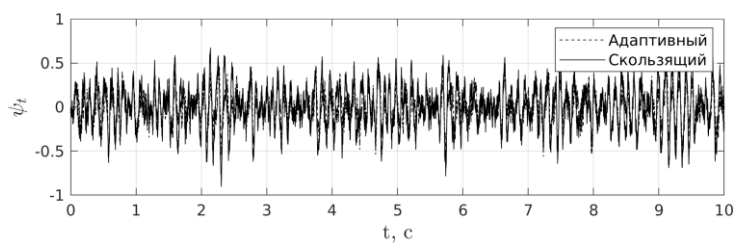


Рис. 5. Относительная реакция шины в случае наличия мертвой зоны

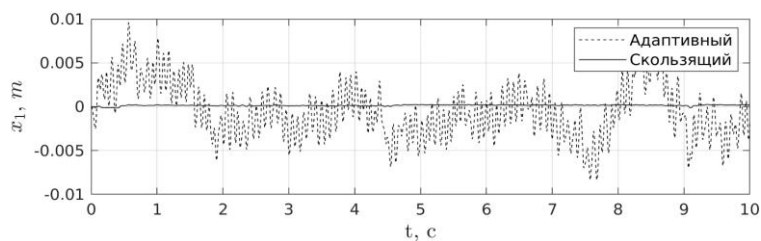


Рис. 6. Перемещение поддресоренной массы в случае наличия гистерезиса

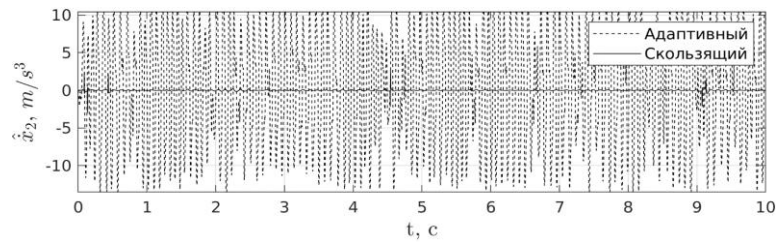


Рис. 7. Взвешенная оценка ускорения поддрессоренной массы в случае наличия гистерезиса

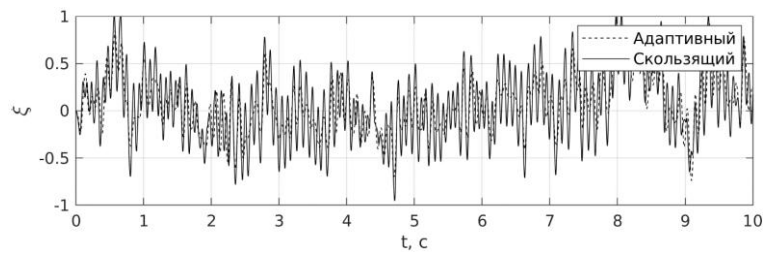


Рис. 8. Относительное перемещение амортизатора в случае наличия гистерезиса

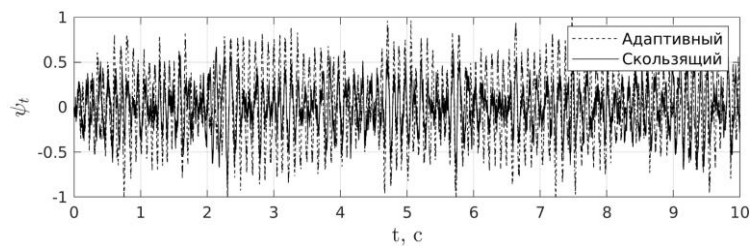


Рис. 9. Относительная реакция шины в случае наличия гистерезиса

Результаты моделирования замкнутой системы с различными системами управления при наличии в системе гистерезиса представлены на рис. 6–9.

Результаты оценки показателей эффективности системы представлены в табл. 3.

Таблица 3

Показатели эффективности

Тип звена	Регулятор	Взвешенное ускорение		Относительное перемещение		Относительная реакция шины	
		$\max \hat{x}_2 $	СКО	$\max \xi $	СКО	$\max \psi_t $	СКО
Мертвая зона	Адаптивный	6,414	2,247	0,828	0,214	0,903	0,239
	Скользящий	2,560	0,227	0,91	0,283	0,585	0,171
Гистерезис	Адаптивный	15,105	7,697	0,922	0,283	1,048	0,391
	Скользящий	3,004	0,103	0,998	0,350	0,892	0,238

Для более наглядной демонстрации отличительных особенностей исследуемых законов управления среднеквадратические значения оценок табл. 3 проиллюстрированы на рис. 10, 11.

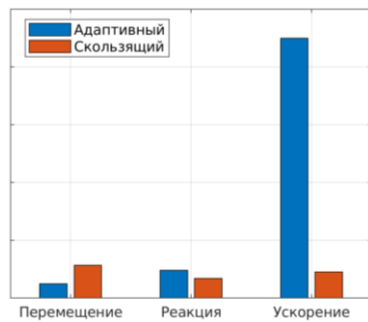


Рис. 10. Результаты оценки показателей эффективности при наличии мертвой зоны

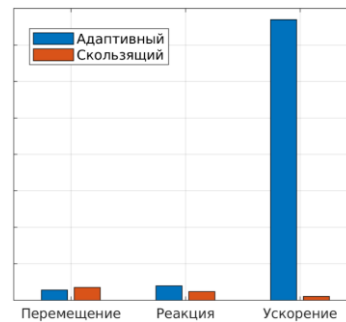


Рис. 11. Результаты оценки показателей эффективности при наличии гистерезиса

Анализируя полученные результаты, можно сделать вывод, что скользкий регулятор значительно превосходит адаптивный закон управления в части изоляции кузова автомобиля от ударов и вибраций, вызванных неровностями дорожного покрытия. Скользящий регулятор также обеспечивает большую безопасность движения транспортного средства, что следует из меньшего значения относительной реакции шины. Следует отметить, что при использовании скользкого регулятора увеличивается относительное перемещение амортизатора, что обусловлено противоречием между требованием к максимальной неподвижности кузова автомобиля и минимизацией перемещения амортизатора.

Заключение. В работе предложен скользкий синергетический регулятор для системы адаптивной подвески, который базируется на современной концепции синергетического управления. Синтезированный закон управления позволяет обеспечить стабилизацию изолируемой массы в нулевом положении из любых, физически приемлемых начальных условий. Введение в систему поверхности переключения позволяет добиться высокого уровня робастности системы и значительно увеличить показатели эффективности системы в условиях наличия неопределенностей в исполнительном механизме, а использование метода АКАР – учесть всю доступную информацию о динамике объекта управления. Сравнение полученного регулятора и адаптивной системы управления, синтезированной на основе метода АКАР, показывает, что предложенная система разрывного управления позволяет повысить следующие характеристики системы адаптивной подвески:

- 1) более чем в 2 раза уменьшить амплитуду взвешенного ускорения поддрессоренной массы и более чем на порядок уменьшить СКО при наличии мертвой зоны в исполнительном механизме подвески;
- 2) более чем на 30 % уменьшить максимальную силу реакции шины и СКО при наличии мертвой зоны в исполнительном механизме подвески;
- 3) более чем в 5 раза уменьшить амплитуду взвешенного ускорения поддрессоренной массы и более чем на порядок уменьшить СКО при наличии гистерезиса в исполнительном механизме подвески;
- 4) на 15 % уменьшить максимальную и на 39 % СКО силы реакции шины при наличии гистерезиса в исполнительном механизме подвески.

Уменьшение амплитуды и ускорения поддрессоренной массы приводит к повышению комфорта, а уменьшение нормальной реакции шины уменьшает ее износ и повышает управляемость автомобиля.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Hrovat D.* Survey of advanced suspension developments and related optimal control applications // *Automatica*. – 1997. – Vol. 33, No. 10. – P. 1781-1817.
2. *Yagiz N., Hacıoglu Y.* Backstepping control of a vehicle with active suspensions // *Contr. Eng. Pract.* – 2008. – Vol. 16, No. 12. – P. 1457-1467.
3. *Koch G., Spirk S., Pellegrini E., Pletschen N., Lohmann B.* Experimental validation of a new adaptive control approach for a hybrid suspension system // *American Control Conference 2011, June 29-July 1, San Francisco, California, USA*. – San Francisco, 2011. – P. 4580-4585.
4. *Zhao F., Ge S.S., Tu F., Qin Y., Dong M.* Adaptive neural network control for active suspension system with actuator saturation // *IET Control Theory & Appl.* – 2016. – Vol. 10, No. 14. – P. 1696-1705.
5. *Jue W., Jing Z.* Model-free tracking control for vehicle active suspension systems // *34th Chinese Control Conference (CCC), July 28-30, Hangzhou, China*. – Hangzhou, 2015. – Vol. 10, No. 14. – P. 1696-1705.
6. *Deshpande V.S., Shendge P.D., Phadke S.B.* Dual objective active suspension system based on a novel nonlinear disturbance compensator // *Veh. Syst. Dyn.* – 2016. – Vol. 54, No. 9. – P. 1269-1290.
7. *Hu C., Yao B., Wang Q.* Adaptive robust precision motion control of systems with unknown input dead-zones: a case study with comparative experiments // *IEEE Trans. Ind. Electron.* – 2016. – Vol. 58, No. 6. – P. 2454-2464.
8. *Ahmad N.J., Ebraheem H.K., Alnaser M.J., Alostath J.M.* Adaptive control of a dc motor with uncertain deadzone nonlinearity at the input // *Control and Decision Conference (CCDC), Dec 12-15, Orlando, FL, USA*. – Orlando, 2011. – P. 4295-4299.
9. *Tao G., Taware A.* Control by Compensation of Nonlinearities // *Control Systems, Robotics and Automation*. – 2009. – Vol. III. – P. 165-171.
10. *Recker D., Kokotovic P., Rhode D., Winkelmann J.* Adaptive nonlinear control of systems containing a deadzone // *Proceedings of the 30th IEEE Conference on Decision and Control, Dec. 11-13, Brighton, England*. – Brighton, 1991. – P. 2111-2115.
11. *Dong J., Mo B.* The adaptive PID controller design for motor control system with backlash // *Fourth International Conference on Intelligent Control and Information Processing (ICICIP), June 12-13, Beijing, China*. – Beijing, 2013. – P. 59-63.
12. *Tao G., Kokotovic P.V.* Adaptive control of plants with unknown dead-zones // *IEEE Trans. Automat. Contr.* – 1994. – Vol. 39, No. 1. – P. 59-68.
13. *Chang T., Yuan D., Hanek H.* Matched feedforward/model reference control of a high precision robot with dead-zone // *IEEE Trans. Contr. Syst. Technol.* – 2008. – Vol. 16, No. 1. – P. 94-102.
14. *Danilin N.A., Shalashilin D.A.* Hysteresis Modelling of Mechanical Systems at Nonstationary Vibrations // *Mathematical Problems in Engineering*. – 2018. – Vol. 2018. – P. 1-15.
15. *Solomon O.* Some typical shapes of hysteretic loops using the bouc-wen model, *Journal of Information Systems & Operations Management*. – 2013. – Vol. 7, No. 1. – P. 1-9.
16. *Колесников А.А.* Синергетическая теория управления. – М.: Энергоатомиздат, 1994.
17. *Tao G., Kokotovic P.V.* Adaptive Control of Systems with Actuator and Sensor Nonlinearities. – NY.: John Wiley & Sons, Inc., 1996.
18. *Pan H., Sun W., Jing X., Gao H., Yao J.* Adaptive tracking control for active suspension systems with non-ideal actuators // *J. Sound Vib.* – 2017. Vol. 399, No. 1. – P. 2-20.
19. ISO-2631, Mechanical Vibration and Shock: Evaluation of Human Exposure to Whole-body Vibration Part 1, General Requirements, ISO, Geneva.
20. *Zuo L., Nayfeh S.A.* Low order continuous-time filters for approximation of the ISO 2631-1 human vibration sensitivity weightings // *J. Sound Vib.* – 2003. – Vol. 265, No. 1. – P. 459-465.
21. *Deshpande V.S., Shendge P.D., Phadke S.B.* Nonlinear control for dual objective active suspension systems // *IEEE Trans. Intell. Transport. Syst.* – 2017. – Vol. 18, No. 3. – P. 656-665.
22. *Колесников А.А., Веселов Г.Е., Попов А.Н. и др.* Синергетические методы управления сложными системами: Механические и электромеханические системы / под общей ред. А.А. Колесникова. – Изд. стереотип. – М.: Книжный дом «ЛИБРОКОМ», 2013. – 304 с.
23. *Колесников А.А., Веселов Г.Е.* Синергетический принцип иерархизации и аналитический синтез регуляторов взаимосвязанных электромеханических систем // *Известия ЮФУ. Технические науки*. – 2001. – № 5 (23). – С. 80-99.
24. ISO-8608, Mechanical Vibration-road Surface Profiles-reporting of Measured Data, ISO, Geneva.

REFERENCES

1. Hrovat D. Survey of advanced suspension developments and related optimal control applications, *Automatica*, 1997, Vol. 33, No. 10, pp. 1781-1817.
2. Yagiz N., Hacıoglu Y. Backstepping control of a vehicle with active suspensions, *Contr. Eng. Pract.*, 2008, Vol. 16, No. 12, pp. 1457-1467.
3. Koch G., Spirk S., Pellegrini E., Pletschen N., Lohmann B. Experimental validation of a new adaptive control approach for a hybrid suspension system, *American Control Conference 2011, June 29-July 1, San Francisco, California, USA*. San Francisco, 2011, pp. 4580-4585.
4. Zhao F., Ge S.S., Tu F., Qin Y., Dong M. Adaptive neural network control for active suspension system with actuator saturation, *IET Control Theory & Appl.*, 2016, Vol. 10, No. 14, pp. 1696-1705.
5. Jue W., Jing Z. Model-free tracking control for vehicle active suspension systems, *34th Chinese Control Conference (CCC), July 28-30, Hangzhou, China*. Hangzhou, 2015, Vol. 10, No. 14, pp. 1696-1705.
6. Deshpande V.S., Shendge P.D., Phadke S.B., Dual objective active suspension system based on a novel nonlinear disturbance compensator, *Veh. Syst. Dyn.*, 2016, Vol. 54, No. 9, pp. 1269-1290.
7. Hu C., Yao B., Wang Q. Adaptive robust precision motion control of systems with unknown input dead-zones: a case study with comparative experiments, *IEEE Trans. Ind. Electron.*, 2016, Vol. 58, No. 6, pp. 2454-2464.
8. Ahmad N.J., Ebraheem H.K., Alnaser M.J., Alostath J.M. Adaptive control of a dc motor with uncertain deadzone nonlinearity at the input, *Control and Decision Conference (CCDC), Dec 12-15, Orlando, FL, USA*. Orlando, 2011, pp. 4295-4299.
9. Tao G., Taware A. Control by Compensation of Nonlinearities // *Control Systems, Robotics and Automation*, 2009, Vol. III, pp. 165-171.
10. Recker D., Kokotovic P., Rhode D., Winkelmann J. Adaptive nonlinear control of systems containing a deadzone, *Proceedings of the 30th IEEE Conference on Decision and Control, Dec. 11-13, Brighton, England*. Brighton, 1991, pp. 2111-2115.
11. Dong J., Mo B. The adaptive PID controller design for motor control system with backlash, *Fourth International Conference on Intelligent Control and Information Processing (ICICIP), June 12-13, Beijing, China*. Beijing, 2013, pp. 59-63.
12. Tao G., Kokotovic P.V. Adaptive control of plants with unknown dead-zones, *IEEE Trans. Automat. Contr.*, 1994, Vol. 39, No. 1, pp. 59-68.
13. Chang T., Yuan D., Hanek H. Matched feedforward/model reference control of a high precision robot with dead-zone, *IEEE Trans. Contr. Syst. Technol.*, 2008, Vol. 16, No. 1, pp. 94-102.
14. Danilin N.A., Shalashilin D.A. Hysteresis Modelling of Mechanical Systems at Nonstationary Vibrations, *Mathematical Problems in Engineering*, 2018, Vol. 2018, pp. 1-15.
15. Solomon O. Some typical shapes of hysteretic loops using the bouc-wen model, *Journal of Information Systems & Operations Management*, 2013, Vol. 7, No. 1, pp. 1-9.
16. Kolesnikov A.A. *Sinergeticheskaya teoriya upravleniya [Synergetic control theory]*. Moscow: Energoatomizdat, 1994.
17. Tao G., Kokotovic P.V. *Adaptive Control of Systems with Actuator and Sensor Nonlinearities*. NY.: John Wiley & Sons, Inc., 1996.
18. Pan H., Sun W., Jing X., Gao H., Yao J. Adaptive tracking control for active suspension systems with non-ideal actuators, *J. Sound Vib.*, 2017, Vol. 399, No. 1, pp. 2-20.
19. ISO-2631, *Mechanical Vibration and Shock: Evaluation of Human Exposure to Whole-body Vibration Part 1, General Requirements*, ISO, Geneva.
20. Zuo L., Nayfeh S.A. Low order continuous-time filters for approximation of the ISO 2631-1 human vibration sensitivity weightings, *J. Sound Vib.*, 2003, Vol. 265, No. 1, pp. 459-465.
21. Deshpande V.S., Shendge P.D., Phadke S.B. Nonlinear control for dual objective active suspension systems, *IEEE Trans. Intell. Transport. Syst.*, 2017, Vol. 18, No. 3, pp. 656-665.
22. Kolesnikov A.A., Veselov G.E., Popov A.N. i dr. *Sinergeticheskie metody upravleniya slozhnymi sistemami: Mekhanicheskie i elektromekhanicheskie sistemy [Synergetic control methods of complex systems: Mechanical and Electromechanical systems]*, under the General ed. of A.A. Kolesnikova. Moscow: Knizhnyy dom «LIBROKOM», 2013, 304 p.

23. Kolesnikov A.A., Veselov G.E. Sinergeticheskiy printsip ierarkhizatsii i analiticheskiy sintez regulyatorov vzaimosvyazannykh elektromekhanicheskikh sistem [Synergetic principle of hierarchy and analytical synthesis of regulators of interconnected Electromechanical systems], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2001, No. 5 (23), pp. 80-99.
24. ISO-8608, Mechanical Vibration-road Surface Profiles-reporting of Measured Data, ISO, Geneva.

Статью рекомендовал к опубликованию д.т.н. А.Б. Чернышов.

Синицын Александр Сергеевич – АО «Научно-конструкторское бюро вычислительных систем» (АО НКБ ВС); e-mail: axeom18@gmail.com; 347936, г. Таганрог, ул. 1-я Линия, 144-а; тел.: +78634682560, доб. 4760; Отдел алгоритмического обеспечения встроенных систем управления; научный сотрудник.

Sinityn Alexander Sergeevich – Joint Stock Company "Scientific Design Bureau of Computing Systems" (JSC SDB CS); e-mail: axeom18@gmail.com; 1 Liniya St., 144-a, Taganrog, 347936, Russia; phone: +78634682560, ad. 4760; the department of algorithmic support for embedded control systems; researcher.

УДК 681.51

DOI 10.18522/2311-3103-2020-2-101-112

Г.Е. Веселов, Алин Ингабире

СИНЕРГЕТИЧЕСКИЙ СИНТЕЗ ЗАКОНА УПРАВЛЕНИЯ БПЛА В УСЛОВИЯХ ВЕТРОВЫХ ВОЗМУЩЕНИЙ С ВХОДНЫМИ ОГРАНИЧЕНИЯМИ*

Рассматривается применение методов синергетической теории управления (СТУ) к проблеме синтеза системы управления беспилотным летательным аппаратом (БПЛА) в условиях действия ветровых возмущений. Основной задачей исследования является разработка синергетического метода синтеза нелинейных систем управления БПЛА с жёстким крылом, гарантирующих асимптотическую устойчивость замкнутых систем при движении по заданной траектории, устойчивость и адаптивность при значительной нелинейности математических моделей управления БПЛА с жёстким крылом в условиях влияния ветровых возмущений. Кроме того, важной задачей при синтезе систем управления различными объектами, в том числе и БПЛА, является учёт ограничений на переменные состояния объекта управления, которые могут обуславливаться как требованиями к энергоэффективности и безопасности систем, так и другими ограничениями и требованиями, накладываемые на эти координаты. В статье предлагается процедура синтеза нелинейных векторных систем управления БПЛА с жёстким крылом методами СТУ, обеспечивающих инвариантность к действию внешних неизмеряемых возмущающих воздействий, выполнение заданных технологических целей управления, асимптотическую устойчивость замкнутой системы, а также учёт введенных ограничений на внутренние координаты БПЛА. Предлагаемая в статье процедура синергетического синтеза векторных нелинейных систем управления БПЛА с жёстким крылом гарантирует эффективное использование такого типа БПЛА в решении различного рода задач, в том числе и при функционировании таких БПЛА в качестве элементов группы автономных объектов, решающих заданную групповую технологическую задачу. Эффективность предлагаемого подхода к синергетическому синтезу стратегий управления подтверждается результатами компьютерного моделирования синтезированной нелинейной векторной системы управления БПЛА с жёстким крылом. Предлагаемый метод синергетического синтеза системы управления БПЛА с жёстким крылом, может быть применён для разработки перспективных имитационных пилотажно-навигационных комплексов, моделирующих поведение БПЛА в условиях присутствия ветровых возмущений и послужить основой для улучшения лётно-технических характеристик БПЛА.

БПЛА с жёстким крылом; синергетическая теория управления; инвариантное многообразие; ветровые возмущения; инвариантность к внешним неизмеряемым возмущениям; ограничения на переменные пространства состояния объекта управления.

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-07-00645.

G.E. Veselov, Alin Ingabire

SYNERGETIC SYNTHESIS OF CONTROL LAW FOR UAV IN THE PRESENCE OF WIND DISTURBANCES WITH INPUT CONSTRAINTS

This paper discusses the application of synergetic control theory (SCT) methods to the problem of control system synthesis for fixed-wing unmanned aerial vehicle (UAV) in the presence of wind disturbances. The main purpose of this study is to develop a synergetic method for the synthesis of nonlinear control systems for fixed-wing UAVs, which guarantee the asymptotic stability of the closed-loop systems when moving along a given trajectory, stability and adaptability with significant nonlinearity of mathematical models for controlling fixed-wing UAVs in the presence of wind disturbances. Furthermore, an important task in the synthesis of control systems for various objects, including UAV, is to take into account constraints on the state variables of the control object, which can be determined by both the energy efficiency requirements and safety systems, as well as other constraints and requirements imposed on these coordinates. This article proposes a procedure for the synthesis of nonlinear vector control systems for fixed-wing UAV by applying SCT approaches that provide invariance to external unmeasured disturbances, fulfillment of specified technological control objectives, asymptotic stability of the closed-loop system, and also take into account the introduced constraints on the UAV internal coordinates. The procedure suggested in this article for the synergetic synthesis of nonlinear vector control systems of fixed-wing UAV ensures the effective use of this type of UAV in solving various tasks, including the operation of such UAV as elements of a group of autonomous objects that solve a given group technological task. The effectiveness of the proposed approach to the synergetic synthesis of control strategies is confirmed by the results of computer modeling of the synthesized nonlinear vector control system of fixed-wing UAV. The proposed synergetic method of control system synthesis for fixed-wing UAV can be applied for the development of advanced flight simulation and navigation complexes that simulate the UAV behavior in the presence of wind disturbances and serve as a basis for improving the flight performance of the fixed-wing UAV.

Fixed-wing UAV; synergetic control theory; invariant manifold; wind disturbances; invariance to external unmeasured disturbances; constraints on the state variables of the control object.

Введение. Беспилотные летательные аппараты (БПЛА) с жестким крылом являются наиболее распространённым типом БПЛА и нашли широкое применение в агропромышленных комплексах при опрыскивании посевов, при мониторинге пожаров, освещении спортивных и развлекательных мероприятий, в аварийно-спасательных службах и научных исследованиях [1–3].

БПЛА с жестким крылом отличается от БПЛА других видов тем, что время их автономного полёта больше, они могут летать на большей высоте и скорость их полёта больше чем у других типов БПЛА. Однако у БПЛА с жестким крылом есть главный недостаток: они чувствительны к внешним возмущениям из-за малых моментов инерции, размера и веса. Одним из главных ограничивающих факторов БПЛА является сильный ветер, поскольку он может оказать такое влияние на БПЛА, которое приведет к существенным отклонениям от желаемых траекторий движения, что, в свою очередь, может привести к нештатным ситуациям и авариям [4, 5]. Следует отметить, что отклонение БПЛА от требуемой траектории движения на несколько метров может быть допустимым только в малонаселенной местности. Таким образом, для того чтобы иметь возможность осуществлять выполнение технологической задачи в ветреную погоду, необходимо учитывать при синтезе законов управления БПЛА влияние ветровых возмущений [6].

В виду большой популярности БПЛА, имеется большое количество работ, посвященных разработке методов синтеза систем управления БПЛА. В работе [7] был предложен метод нелинейного управления, основанный на использовании так называемого «упреждающего вектора», который описывает желаемое направление скорости для БПЛА. «Упреждающий вектор» представляется в виде функции состояния системы, записанной совместно с влиянием ветровых возмущений в соответствии с целевой задачей объекта управления. Применение данного метода позволяет учитывать изменения ветровых возмущений, что в свою очередь, позволяет улучшить нави-

гацию БПЛА. В работе [8] приводится процедура расчёта пределов скорости поворота БПЛА, и представлен новый подход к вычислению безопасных расстояний вдоль сегментов трохонидального поворота для учета заранее не определенной скорости ветра. Это позволяет планировать траектории полета БПЛА, которые оптимизированы для ожидаемого состояния ветра и безопасны в случае, если самолет подвергается дополнительным ветровым возмущениям. В работе [9] изучается алгоритм траекторного управления БПЛА с учетом ветровых возмущений. В работе представлен синтез наблюдателя возмущений, который предназначен для оценки заранее неизвестных порывов ветра, объединённый с синтезом нелинейного закона управления методом Backstepping. В работе [10] разрабатывается подход к моделированию и идентификации динамики управления для БПЛА. Впоследствии синтезирован «прогнозирующий» нелинейный регулятор для одновременной стабилизации воздушной скорости, следования по заданной траектории с учетом ограничений.

Применение теории и методов синергетической теории управления [11–14] позволяет существенно улучшить процедуру синтеза за счет учета нелинейности поведения БПЛА и действия ветровых возмущений. Однако при применении синергетического подхода к синтезу системы управления БПЛА не учитывались ограничения, которые накладываются на рулевые механизмы БПЛА и существуют физически, а также ограничения на предельно-допустимое изменение некоторых переменных состояния БПЛА. В рамках данной статьи предлагается подход к синтезу системы пространственного БПЛА с использованием метода СТУ, обеспечивающего учёт ограничений на координаты объекта управления [15–17].

Математическая модель. Рассмотрим нелинейную математическую модель, описывающую пространственное поведение БПЛА [18]:

$$\begin{aligned}
 \dot{x}(t) &= V_x \cos \psi \cos \vartheta + V_y (\sin \gamma \sin \psi - \cos \gamma \cos \psi \sin \vartheta) + \\
 &+ V_z (\cos \gamma \sin \psi + \sin \gamma \cos \psi \sin \vartheta); \\
 \dot{y}(t) &= V_x \sin \vartheta + V_y \cos \gamma \cos \vartheta - V_z \sin \gamma \cos \vartheta; \\
 \dot{z}(t) &= -V_x \sin \psi \cos \vartheta + V_y (\sin \gamma \cos \psi + \cos \gamma \sin \psi \sin \vartheta) + \\
 &+ V_z (\cos \gamma \cos \psi - \sin \gamma \sin \psi \sin \vartheta); \\
 \dot{V}_x(t) &= V_y \omega_z - V_z \omega_y + \frac{1}{m} (c_x q S + P \cos \varphi - mg \sin \vartheta); \\
 \dot{V}_y(t) &= V_z \omega_x - V_x \omega_z + \frac{1}{m} (c_y q S + P \sin \varphi - mg \cos \vartheta \cos \gamma); \\
 \dot{V}_z(t) &= V_x \omega_y - V_y \omega_x + \frac{1}{m} (c_z q S + mg \cos \vartheta \sin \gamma); \\
 \dot{\omega}_x(t) &= \frac{I_y - I_z}{I_x} \omega_y \omega_z + \frac{1}{I_x} m_x S q l; \\
 \dot{\omega}_y(t) &= \frac{I_z - I_x}{I_y} \omega_x \omega_z + \frac{1}{I_y} m_y S q l; \\
 \dot{\omega}_z(t) &= \frac{I_x - I_y}{I_z} \omega_x \omega_y + \frac{1}{I_z} m_z S q b_A; \\
 \dot{\vartheta}(t) &= \omega_y \sin \gamma + \omega_z \cos \gamma; \\
 \dot{\gamma}(t) &= \omega_x - \tan \vartheta (\omega_y \cos \gamma - \omega_z \sin \gamma); \\
 \dot{\psi}(t) &= \frac{1}{\cos \vartheta} (\omega_y \cos \gamma - \omega_z \sin \gamma).
 \end{aligned} \tag{1}$$

Здесь x, y, z – координаты центра масс БПЛА в земной системе координат; V_x, V_y, V_z – проекции вектора линейной скорости БПЛА на оси связанной системы координат; $\omega_x, \omega_y, \omega_z$ – проекции вектора угловой скорости БПЛА на оси связанной системы координат; ϑ, γ, ψ – углы тангажа, крена и рыскания БПЛА, соответственно; P – сила тяги двигателя БПЛА; c_x, c_y, c_z – аэродинамические силы; m_x, m_y, m_z – аэродинамические моменты; g – ускорение свободного падения; m – масса БПЛА; I_x, I_y, I_z – моменты инерции БПЛА относительно соответствующих осей; S – площадь крыла самолета; q – скоростной напор определяемый формулой $q = \frac{\rho |V_a|^2}{2}$, ρ – плотность воздуха, V_a – величина воздушная скорость; l – размах крыла; b_A – средняя аэродинамическая хорда крыла БПЛА.

На динамику БПЛА оказывают существенное влияние ветровые возмущения (рис. 1) [19]. Поэтому необходимо, чтобы замкнутая система управления БПЛА была инвариантна к действию этих возмущений. В работах [20, 21] предложена процедура интегральной адаптации для сложных нелинейных систем управления.

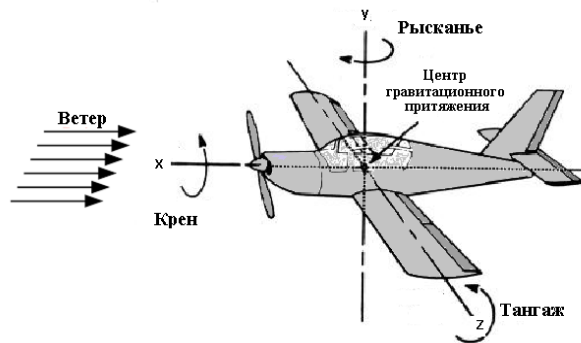


Рис. 1. БПЛА с жестким крылом, летящий в направлении ветра

В соответствии с процедурой интегральной адаптации необходимо ввести набор целей управления. В СТУ критерии управления выражаются в виде соответствующей системы инвариантов, которые играют значительную роль при проектировании системы управления. В качестве таких инвариантов выберем стабилизацию линейной скорости V_x , высоты y , бокового смещения z , а так же углов ϑ, γ, ψ . Запишем эти инварианты БПЛА в следующем виде:

$$\begin{aligned} V_x &= V_x^*; & \vartheta &= \vartheta_0; \\ y &= y_0; & \gamma &= \gamma_0; \\ z &= z_0; & \psi &= \psi_0; \end{aligned}$$

которые будут входить в уравнения оценок возмущений $z_i, i=1, \dots, 6$, дополняющих базовую модель динамики БПЛА. Таким образом, воспользовавшись этой процедурой и задав цели управления можно записать расширенную математическую модель синтеза:

$$\begin{aligned}
 \dot{x}(t) &= V_x \cos \psi \cos \vartheta + V_y (\sin \gamma \sin \psi - \cos \gamma \cos \psi \sin \vartheta) + \\
 &+ V_z (\cos \gamma \sin \psi + \sin \gamma \cos \psi \sin \vartheta); \\
 \dot{y}(t) &= V_x \sin \vartheta + V_y \cos \gamma \cos \vartheta - V_z \sin \gamma \cos \vartheta; \\
 \dot{z}(t) &= -V_x \sin \psi \cos \vartheta + V_y (\sin \gamma \cos \psi + \cos \gamma \sin \psi \sin \vartheta) + \\
 &+ V_z (\cos \gamma \cos \psi - \sin \gamma \sin \psi \sin \vartheta); \\
 \dot{V}_x(t) &= V_y \omega_z - V_z \omega_y + \frac{1}{m} (c_x q S + P \cos \varphi - mg \sin \vartheta) + z_1; \\
 \dot{V}_y(t) &= V_z \omega_x - V_x \omega_z + \frac{1}{m} (c_y q S + P \sin \varphi - mg \cos \vartheta \cos \gamma) + z_2; \\
 \dot{V}_z(t) &= V_x \omega_y - V_y \omega_x + \frac{1}{m} (c_z q S + mg \cos \vartheta \sin \gamma) + z_3; \\
 \dot{\omega}_x(t) &= \frac{I_y - I_z}{I_x} \omega_y \omega_z + \frac{1}{I_x} m_x S q l + z_4; \\
 \dot{\omega}_y(t) &= \frac{I_z - I_x}{I_y} \omega_x \omega_z + \frac{1}{I_y} m_y S q l + z_5; \\
 \dot{\omega}_z(t) &= \frac{I_x - I_y}{I_z} \omega_x \omega_y + \frac{1}{I_z} m_z S q b_A + z_6; \\
 \dot{\vartheta}(t) &= \omega_y \sin \gamma + \omega_z \cos \gamma; \\
 \dot{\gamma}(t) &= \omega_x - \tan \vartheta (\omega_y \cos \gamma - \omega_z \sin \gamma); \\
 \dot{\psi}(t) &= \frac{1}{\cos \vartheta} (\omega_y \cos \gamma - \omega_z \sin \gamma); \\
 \dot{z}_1(t) &= \eta_1 (V_x - V_x^*); \\
 \dot{z}_2(t) &= \eta_2 (y - y_0); \\
 \dot{z}_3(t) &= \eta_3 (z - z_0); \\
 \dot{z}_4(t) &= \eta_4 (\vartheta - \vartheta_0); \\
 \dot{z}_5(t) &= \eta_5 (\gamma - \gamma_0); \\
 \dot{z}_6(t) &= \eta_6 (\psi - \psi_0);
 \end{aligned} \tag{2}$$

где z_i – оценки внешних неизмеряемых возмущающих действий, η_i , $i = 1, \dots, 6$, – постоянные коэффициенты, V_x^* – желательное значение скорости, y_0 – требуемая высота полёта, z_0 – необходимое боковое смещение, ϑ_0 – требуемое значение угла тангажа, γ_0 – требуемое значение угла крена, ψ_0 – требуемое значение угла рыскания.

Процедура синтеза системы управления. На первом этапе синергетического синтеза вводится первый набор макропеременных:

$$\begin{aligned}
 \Psi_1 &= V_x - V_x^* + z_1; & \Psi_2 &= V_y - L_2 \tanh(\varphi_1); \\
 \Psi_3 &= V_z - L_3 \tanh(\varphi_2); & \Psi_4 &= \omega_x - \varphi_3; \\
 \Psi_5 &= \omega_y - \varphi_4; & \Psi_6 &= \omega_z - \varphi_5;
 \end{aligned} \tag{3}$$

удовлетворяющих решению системы основных функциональных уравнений

$$T_i \dot{\Psi}_i(t) + \Psi_i = 0, \quad n=1,2,\dots,6, \quad (4)$$

где T_i – константы времени; φ_i , $i=1,\dots,5$ – некоторые функции связи, которые будут определены далее в процессе проектирования; L_2 , L_3 – максимальные значения проекций линейной скорости на координаты x и y . В соответствии с СТУ переменные Ψ_2 и Ψ_3 сформированы таким образом, что при условии, что $\Psi_2 = 0$ и $\Psi_3 = 0$ значение проекций линейной скорости на оси y и z не будут превышать заданных максимальных значений, таким образом будет реализовано ограничение на внутренние координаты системы.

Таким образом, при попадании изображающей точки замкнутой системы в силу уравнений (3) и (4) на пересечение инвариантных многообразий $\Psi_i = 0$, $i = 1,\dots,6$ в системе будут гарантированно выполняться следующие соотношения

$$\begin{aligned} V_x &= V_x^* + z_1; & V_y &= L_2 \tanh(\varphi_1); \\ V_z &= L_3 \tanh(\varphi_2); & \omega_x &= \varphi_3; \\ \omega_y &= \varphi_4; & \omega_z &= \varphi_5. \end{aligned} \quad (5)$$

Соответственно, из совместного решения уравнений (3) и (4), с учётом расширенной модели объекта управления (2), определяются аэродинамические силы C_x, C_y, C_z и аэродинамические моменты m_x, m_y, m_z , необходимые для выполнения заданной технологической задачи управления БПЛА. При попадании изображающей точки системы на пересечение совокупности инвариантных многообразий (5) в замкнутой системе осуществляется динамического сжатия фазового пространства, в результате чего поведение системы на пересечении многообразий (5) будет описываться декомпозированной системой дифференциальных уравнений:

$$\begin{aligned} \dot{x}(t) &= V_x^* \cos \psi \cos \vartheta + \varphi_1 (\sin \gamma \sin \psi - \cos \gamma \cos \psi \sin \vartheta) + \\ &+ \varphi_2 (\cos \gamma \sin \psi + \sin \gamma \cos \psi \sin \vartheta); \\ \dot{y}(t) &= V_x^* \sin \vartheta + \varphi_1 \cos \gamma \cos \vartheta - \varphi_2 \sin \gamma \cos \vartheta; \\ \dot{z}(t) &= -V_x^* \sin \psi \cos \vartheta + \varphi_1 (\sin \gamma \cos \psi + \cos \gamma \sin \psi \sin \vartheta) + \\ &+ \varphi_2 (\cos \gamma \cos \psi - \sin \gamma \sin \psi \sin \vartheta); \\ \dot{\vartheta}(t) &= \varphi_4 \sin \gamma + \varphi_5 \cos \gamma; \\ \dot{\gamma}(t) &= \varphi_3 - \operatorname{tg} \vartheta (\varphi_4 \cos \gamma - \varphi_5 \sin \gamma); \\ \dot{\psi}(t) &= \frac{1}{\cos \vartheta} (\varphi_4 \cos \gamma - \varphi_5 \sin \gamma); \\ \dot{z}_1(t) &= \eta_1 (V_x - V_x^*); \\ \dot{z}_2(t) &= \eta_2 (y - y_0); \\ \dot{z}_3(t) &= \eta_3 (z - z_0); \\ \dot{z}_4(t) &= \eta_4 (\vartheta - \vartheta_0); \\ \dot{z}_5(t) &= \eta_5 (\gamma - \gamma_0); \\ \dot{z}_6(t) &= \eta_6 (\psi - \psi_0). \end{aligned} \quad (6)$$

Для декомпозированной системы (6) введем вторую совокупность макропеременных системы

$$\begin{aligned}
 \Psi_7 &= y - y_0 + z_2; \\
 \Psi_8 &= z - z_0 + z_3; \\
 \Psi_9 &= \vartheta - \vartheta_0 + z_4; \\
 \Psi_{10} &= \gamma - \gamma_0 + z_5; \\
 \Psi_{11} &= \psi - \psi_0 + z_6;
 \end{aligned} \tag{7}$$

которая также должна удовлетворять решению системы функциональных уравнений

$$T_i \dot{\Psi}_i(t) + \Psi_i = 0, \quad i = 7, 8, \dots, 11. \tag{8}$$

Из совместного решения систем уравнений (7) и (8), с учетом декомпозированной модели (6), определяются выражения для внутренних управлений:

$$\begin{aligned}
 \varphi_1 &= \frac{1}{\cos \psi T_7 (\sin^2 \gamma + \cos^2 \gamma) \cos \mathcal{G} T_8} (\sin \psi \sin^2 \mathcal{G} \sin \gamma T_7 T_8 V_x^* + \\
 &+ \sin \psi \sin \mathcal{G} \sin \gamma T_7 T_8 \eta_2 y - \sin \psi \sin \mathcal{G} \sin \gamma T_7 T_8 \eta_2 y_0 + \sin \psi \sin \gamma \cos^2 \mathcal{G} T_7 T_8 V_x^* - \\
 &- \cos \psi \sin \mathcal{G} \cos \gamma T_7 T_8 V_x^* - \cos \psi \cos \gamma T_7 T_8 \eta_2 y + \cos \psi \cos \gamma T_7 T_8 \eta_2 y_0 - \\
 &- \sin \gamma \cos \mathcal{G} T_7 T_8 \eta_3 z + \sin \gamma \cos \mathcal{G} T_7 T_8 \eta_3 z_0 + \sin \psi \sin \mathcal{G} \sin \gamma T_7 y - \\
 &- \sin \psi \sin \mathcal{G} \sin \gamma T_7 y_0 + \sin \psi \sin \mathcal{G} \sin \gamma T_7 z_2 - \cos \psi \cos \gamma T_7 y + \cos \psi \cos \gamma T_7 y_0 - \\
 &- \cos \psi \cos \gamma T_7 z_2 - \sin \gamma \cos \mathcal{G} T_8 z + \sin \gamma \cos \mathcal{G} T_8 z_0 - \sin \gamma \cos \mathcal{G} T_8 z_3); \\
 \varphi_2 &= \frac{1}{\cos \psi T_7 (\sin^2 \gamma + \cos^2 \gamma) \cos \mathcal{G} T_8} (\sin \psi \sin^2 \mathcal{G} \cos \gamma T_7 T_8 V_x^* + \\
 &+ \sin \psi \sin \mathcal{G} \cos \gamma T_7 T_8 \eta_2 y - \sin \psi \sin \mathcal{G} \cos \gamma T_7 T_8 \eta_2 y_0 + \sin \psi \cos^2 \mathcal{G} T_7 V_x^* \cos \gamma T_8 + \\
 &+ \cos \psi \sin \mathcal{G} \sin \gamma T_7 T_8 V_x^* + \cos \psi \sin \gamma T_7 T_8 \eta_2 y - \cos \psi \sin \gamma T_7 T_8 \eta_2 y_0 - \\
 &- T_7 \eta_3 z \cos \mathcal{G} \cos \gamma T_8 + T_7 \eta_3 z_0 \cos \mathcal{G} \cos \gamma T_8 + \sin \psi \sin \mathcal{G} \cos \gamma T_7 y - \\
 &- \sin \psi \sin \mathcal{G} \cos \gamma T_7 y_0 + \sin \psi \sin \mathcal{G} \cos \gamma T_7 z_2 + \cos \psi \sin \gamma T_7 y - \cos \psi \sin \gamma T_7 y_0 + \\
 &+ \cos \psi \sin \gamma T_7 z_2 - z \cos \mathcal{G} \cos \gamma T_8 + z_0 \cos \mathcal{G} \cos \gamma T_8 - z_3 \cos \mathcal{G} \cos \gamma T_8); \\
 \varphi_3 &= \frac{1}{T_{10} T_{11}} (tg \mathcal{G} \cos \mathcal{G} T_{10} T_{11} \eta_6 \psi - tg \mathcal{G} \cos \mathcal{G} T_{10} T_{11} \eta_6 \psi_0 + tg \mathcal{G} \cos \mathcal{G} T_{10} \psi - \\
 &- tg \mathcal{G} \cos \mathcal{G} T_{10} \psi_0 + tg \mathcal{G} \cos \mathcal{G} T_{10} z_6 + T_{10} T_{11} \eta_5 \gamma - T_{10} T_{11} \eta_5 \gamma_0 + T_{11} \gamma - T_{11} \gamma_0 + T_{11} z_5); \\
 \varphi_4 &= \frac{1}{T_{11} T_9 (\sin^2 \gamma + \cos^2 \gamma)} (\cos \mathcal{G} \cos \gamma T_{11} T_9 \eta_6 \psi - \cos \mathcal{G} \cos \gamma T_{11} T_9 \eta_6 \psi_0 + \\
 &+ \sin \gamma T_{11} T_9 \eta_4 \mathcal{G} - \sin \gamma T_{11} T_9 \eta_4 \mathcal{G}_0 + \cos \mathcal{G} \cos \gamma T_9 \psi - \cos \mathcal{G} \cos \gamma T_9 \psi_0 + \\
 &+ \cos \mathcal{G} \cos \gamma T_9 z_6 + \sin \gamma T_{11} \mathcal{G} - \sin \gamma T_{11} \mathcal{G}_0 + \sin \gamma T_{11} z_4); \\
 \varphi_5 &= \frac{1}{T_{11} T_9 (\sin^2 \gamma + \cos^2 \gamma)} (\cos \mathcal{G} T_{11} \eta_6 \psi \sin \gamma T_9 - \cos \mathcal{G} T_{11} \eta_6 \psi_0 \sin \gamma T_9 - \\
 &- \cos \gamma T_{11} T_9 \eta_4 \mathcal{G} + \cos \gamma T_{11} T_9 \eta_4 \mathcal{G}_0 + \psi \cos \mathcal{G} \sin \gamma T_9 - \psi_0 \cos \mathcal{G} \sin \gamma T_9 + \\
 &+ z_6 \cos \mathcal{G} \sin \gamma T_9 - \cos \gamma T_{11} \mathcal{G} + \cos \gamma T_{11} \mathcal{G}_0 - \cos \gamma T_{11} z_4).
 \end{aligned} \tag{9}$$

Нелинейные законы векторного управления БПЛА с учётом ограничений на координаты объекта в статье не приводятся ввиду своей громоздкости. Однако они могут быть получены из совместного решения уравнений (3) и (4), с учётом расширенной модели объекта управления (2) и полученных выражений для внутренних управляющих воздействий (9).

Компьютерное моделирование. Выполним компьютерное моделирование синтезированной замкнутой системы управления БПЛА в условиях ветровых возмущений с входными ограничениями. Параметры БПЛА с жестким крылом приведены в табл. 1.

Таблица 1

Параметры БПЛА с жестким крылом

Параметр	Обозначение	Значение	Единица измерения
Масса	m	22	кг
Момент инерции, x	I_x	5.32	кгм ²
Момент инерции, y	I_y	12	кгм ²
Момент инерции, z	I_z	8.7	кгм ²
Средняя аэродинамическая хорда крыла	b_A	4.3	м
Площадь крыла	S	8.3	м ²
Ускорение свободного падения	g	9.81	м / с ²
Сила тяги двигателя	P	24.5	Н
Плотность воздуха	ρ	1.225	кг / м ³
Размах крыла	l	4.1	м
Воздушная скорость	V_a	17	м / с

Зададим желаемые значения переменных состояния БПЛА с жестким крылом:

$$V_x^* = 80 \text{ м/с}, y_0 = 90 \text{ м}, z_0 = 50 \text{ м}, \vartheta_0 = 0, \gamma_0 = 0, \psi_0 = 0.$$

В табл. 2 представлены параметры ограничений на координаты состояния БПЛА с жестким крылом.

Таблица 2

Параметры ограничений БПЛА

Ограничения	Значение	Единица измерения
L_2	95	м / с
L_3	65	м / с

Результаты моделирования синтезированной замкнутой системы управления БПЛА с жестким крылом и каналов управления представлены на рис. 2–7. Представленные результаты моделирования демонстрируют выполнение поставленных задач управления, БПЛА выходит на заданную высоту, соблюдая требуемый курс движения и осуществляет движение с заданной скоростью, при этом проекции линейной скорости на оси y и z не превышают предельно допустимых значений (см. табл. 2).

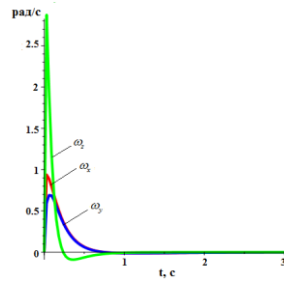


Рис. 2. Переходной процесс угловой скорости $\omega_x, \omega_y, \omega_z$

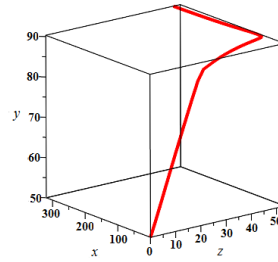


Рис. 3. Линейные координаты x, y, z

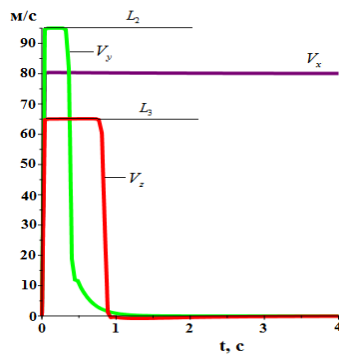


Рис. 4. Переходной процесс линейной скорости V_x, V_y, V_z

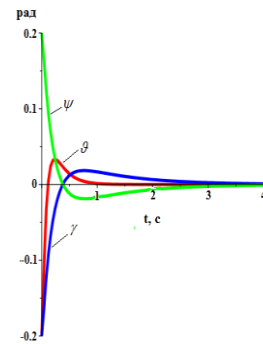


Рис. 5. Переходной процесс угловой координаты ϑ, γ, ψ

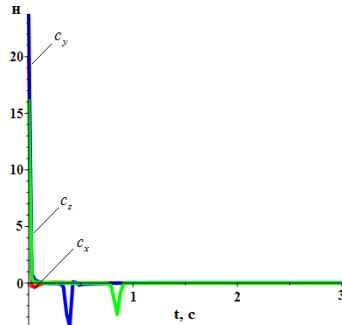


Рис. 6. Переходный процесс управления c_x, c_y, c_z

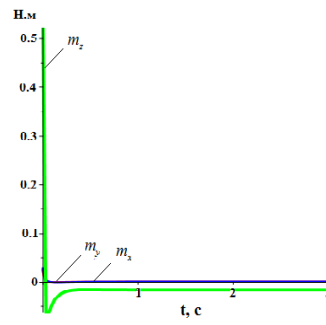


Рис. 7. Переходный процесс управления m_x, m_y, m_z

Заключение. Таким образом в данной работе была рассмотрена нелинейная модель пространственного движения БПЛА с жестким крылом в условиях ветровых возмущений. Представленная процедура синергетического синтеза векторной нелинейной системы управления БПЛА с жестким крылом позволяет учитывать требуемые ограничения на внутренние координаты состояния БПЛА, а также обеспечивать инвариантность к действию внешних возмущающих воздействий.

Гарантированное асимптотически устойчивое поведение БПЛА и отработка требуемых инвариантов управления обеспечивает возможность использования БПЛА с синергетическими законами векторного нелинейного управления как элементов более сложных систем, решающих сложные технические задачи, в том числе для группы автономных управляемых объектов. Эффективность предложенного подхода демонстрируется результатами компьютерного моделирования БПЛА с синтезированным векторным нелинейным регулятором. Предлагаемый метод синергетического синтеза системы управления БПЛА с жёстким крылом, может быть применён для разработки перспективных имитационных пилотажно-навигационных комплексов, моделирующих поведение БПЛА в условиях присутствия ветровых возмущений и послужить основой для улучшения лётно-технических характеристик БПЛА.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Honrado J., Solpico D.B., Favila C., Tongson E., Tangonan G.L., Libatique N.J.* UAV imaging with low-cost multispectral imaging system for precision agriculture applications // IEEE Global Humanitarian Technology Conference (GHTC). – 2017. – P. 1-7.
2. *Kaleem Z., Rehmani M.H., Ahmed E., Jamalipour A., Rodrigues J.J., Moustafa H., Guibene W.* Amateur drone surveillance: Applications, architectures, enabling technologies, and public safety issues: Part 1 // IEEE Communications Magazine. – 2018. – Vol. 56, No. 1. – P. 14-15.
3. *Scott J., Scott C.* Drone delivery models for healthcare // Proc. of the 50th Hawaii Int. Conf. on system sciences. – 2017. – P. 3297-3304.
4. *Liu C., Chen W.H.* Disturbance rejection flight control for small fixed-wing unmanned aerial vehicles // Journal of Guidance, Control, and Dynamics. – 2016. – P. 2810-2819.
5. *Ingabire A., Sklyarov A.A.* Fixed-wing UAVs navigation in the presence of wind: a survey // Инженерный вестник Дона. – 2019. – № 3 (54). – С. 1-10.
6. *Liu C., McAree O., Chen W.H.* Path following for small UAVs in the presence of wind disturbance // IEEE Proceedings of UKACC International Conference on Control. – 2012. – P. 613-618.
7. *Furieri L., Stastny T., Marconi L., Siegwart R., Gilitschenski I.* Gone with the wind: Nonlinear guidance for small fixed-wing aircraft in arbitrarily strong wind fields // American Control Conference (ACC). IEEE. – 2017. – P. 4254-4261.
8. *Schopferer S., Lorenz J.S., Keipour A., Scherer S.* Path planning for unmanned fixed-wing aircraft in uncertain wind conditions using trochoids // International Conference on Unmanned Aircraft Systems (ICUAS). IEEE. – 2018. – P. 503-512.
9. *Wu K., Fan B., Zhang X.* Trajectory following control of UAVs with wind disturbance // 36th Chinese Control Conference (CCC). IEEE. – 2017. – P. 4993-4997.
10. *Stastny T., Siegwart R.* Nonlinear model predictive guidance for fixed-wing UAVs using identified control augmented dynamics // International Conference on Unmanned Aircraft Systems (ICUAS). IEEE. – 2018. – P. 432-442.
11. *Колесников А.А., Кобзев В.А., Нгуен Ф.* Синергетический синтез систем управления движением самолетов-амфибий, функционирующих в экстремальных условиях // Известия ЮФУ. Технические науки. – 2010. – № 5 (106). – С. 150-155.
12. *Мушенко А.С.* Нелинейный синергетический регулятор системы автоматического управления беспилотным летательным аппаратом // Механика твердого тела. – 2002. – С. 165-171.
13. *Мотиенко Т.А.* Синергетический синтез астатических законов управления движением летательных аппаратов // Известия ЮФУ. Технические науки. – 2011. – № 5 (118). – С. 124-128.
14. *Фоменко А.А.* Синергетический синтез законов векторного управления пространственным движением беспилотного летательного аппарата // Известия ЮФУ. Технические науки. – 2011. – № 6 (119). – С. 162-170.
15. *Колесников А.А.* Синергетическая теория управления. – М.: Энергоатомиздат, 1994. – 344 с.
16. *Колесников А.А.* Синергетическая концепция системного синтеза: единство процессов самоорганизации и управления // Известия ТРТУ. – 2006. – № 6 (61). – С. 10-38.
17. *Kolesnikov A.A.* Introduction of synergetic control // Proceedings of the American Control Conference. – 2014. – P. 3013-3016.

18. Колесников А.А. Новые нелинейные методы управления полетом. – М.: Физматлит, 2013. – 196 с.
19. Gonzalez H.A. Robust tracking of dynamic targets with aerial vehicles using quaternion-based techniques: Doctoral dissertation. – 2019. – 18 p.
20. 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.
21. Kuz'menko A.A., Kolesnikov A.A., Kolesnitchenko D.A. Novel robust control of hydrogenerator: The synergetic approach // IFAC-PapersOnLine. – 2015. – Vol. 48 (11). – P. 451-456.

REFERENCES

1. Honrado J., Solpico D.B., Favila C., Tongson E., Tangonan G.L., Libatique N.J. UAV imaging with low-cost multispectral imaging system for precision agriculture applications, *IEEE Global Humanitarian Technology Conference (GHTC)*, 2017, pp. 1-7.
2. Kaleem Z., Rehmani M.H., Ahmed E., Jamalipour A., Rodrigues J.J., Moustafa H., Guibene W. Amateur drone surveillance: Applications, architectures, enabling technologies, and public safety issues: Part 1, *IEEE Communications Magazine*, 2018, Vol. 56, No. 1, pp. 14-15.
3. Scott J., Scott C. Drone delivery models for healthcare, *Proc. of the 50th Hawaii Int. Conf. on system sciences*, 2017, pp. 3297-3304.
4. Liu C., Chen W.H. Disturbance rejection flight control for small fixed-wing unmanned aerial vehicles, *Journal of Guidance, Control, and Dynamics*, 2016, pp. 2810-2819.
5. Ingabire A., Sklyarov A.A. Fixed-wing UAVs navigation in the presence of wind: a survey, *Inzhenernyy vestnik Dona [Engineering Journal of Don]*, 2019, No. 3 (54), pp. 1-10.
6. Liu C., McAree O., Chen W.H. Path following for small UAVs in the presence of wind disturbance, *IEEE Proceedings of UKACC International Conference on Control*, 2012, pp. 613-618.
7. Furieri L., Stastny T., Marconi L., Siegwart R., Gilitschenski I. Gone with the wind: Nonlinear guidance for small fixed-wing aircraft in arbitrarily strong wind fields, *American Control Conference (ACC). IEEE*, 2017, pp. 4254-4261.
8. Schopferer S., Lorenz J.S., Keipour A., Scherer S. Path planning for unmanned fixed-wing aircraft in uncertain wind conditions using trochoids, *International Conference on Unmanned Aircraft Systems (ICUAS). IEEE*, 2018, pp. 503-512.
9. Wu K., Fan B., Zhang X. Trajectory following control of UAVs with wind disturbance, *36th Chinese Control Conference (CCC). IEEE*, 2017, pp. 4993-4997.
10. Stastny T., Siegwart R. Nonlinear model predictive guidance for fixed-wing UAVs using identified control augmented dynamics, *International Conference on Unmanned Aircraft Systems (ICUAS). IEEE*, 2018, pp. 432-442.
11. Kolesnikov A.A., Kobzev V.A., Nguen F. Sinergeticheskii sintez sistem upravleniya dvizheniem samoletov-amfibi, funktsioniruyushchikh v ekstremal'nykh usloviyakh [Synergetics synthesis of amphibian aircraft motion control system operated under extreme conditions], *Izvestiya YuFU. Tekhnicheskie nauki [Izvestiya SFedU. Engineering Sciences]*, 2010, No. 5 (106), pp. 150-155.
12. Mushenko A.S. Nelineyny sinergeticheskii regulyator sistemy avtomaticheskogo upravleniya bespilotnym letatel'nyim apparatom [Nonlinear synergistic regulator automated control systems of Unmanned Aerial Vehicle], *Mekhanika tverdogo tela [Mechanics of Solids]*, 2002, pp. 165-171.
13. Motienko T.A. Sinergeticheskii sintez astaticheskikh zakonov upravleniya dvizheniem letatel'nykh apparatov [Synergetic synthesis of control laws for the movement of the aircraft], *Izvestiya YuFU. Tekhnicheskie nauki [Izvestiya SFedU. Engineering Sciences]*, 2011, No. 5 (118), pp. 124-128.
14. Fomenko A.A. Sinergeticheskii sintez zakonov vektornogo upravleniya prostranstvennym dvizheniem bespilotnogo letatel'nogo apparata [Synergistic synthesis of the vector control laws of spatial movement of the unmanned aerial vehicle], *Izvestiya YuFU. Tekhnicheskie nauki [Izvestiya SFedU. Engineering Sciences]*, 2011, No. 6 (119), pp. 162-170.
15. Kolesnikov A.A. Sinergeticheskaya teoriya upravleniya [Synergetic control theory]. Moscow: Energoatomizdat, 1994, 344 p.
16. Kolesnikov A.A. Sinergeticheskaya kontseptsiya sistemnogo sinteza: edinstvo protsessov samoorganizatsii i upravleniya [Synergetic conception of system synthesis: the unity of processes of self-organization and control], *Izvestiya TRTU [Izvestiya TSURE]*, 2006, No. 6 (61), pp. 10-38.

17. Kolesnikov A.A. Introduction of synergetic control, *Proceedings of the American Control Conference*, 2014, p. 3013-3016.
18. Kolesnikov A.A. Novye nelineynye metody upravleniya poletom [New nonlinear methods of flight control]. Moscow: Fizmatlit, 2013, 196 p.
19. Gonzalez H.A. Robust tracking of dynamic targets with aerial vehicles using quaternion-based techniques: Doctoral dissertation. 2019. 18 p.
20. Metod AKAR i teoriya adaptivnogo upravleniya v zadachakh sinteza nelineynykh sistem upravleniya [ADAR Method and Theory of Adaptive Control in the Tasks of Synthesis of the Nonlinear Control Systems], *Mekhatronika, avtomatizatsiya, upravlenie* [Mechatronics, automation, management], 2017, Vol. 18, No. 9, pp. 579-589.
21. Kuz'menko A.A., Kolesnikov A.A., Kolesnitchenko D.A. Novel robust control of hydrogenerator: The synergetic approach, *IFAC-PapersOnLine*, 2015, Vol. 48 (11), pp. 451-456.

Статью рекомендовал к опубликованию д.т.н., профессор И.М. Першин.

Веселов Геннадий Евгеньевич – Южный федеральный университет; e-mail: gev@sfedu.ru; 347900, г. Таганрог, ул. Чехова, 2; тел.: +78634360450; д.т.н.; директор Института компьютерных технологий и информационной безопасности.

Ингабире Алин – e-mail: ingabire@sfedu.ru; тел.: +79381069281; кафедра синергетики и процессов управления; аспирант.

Veselov Gennady Evgen'evich – Southern Federal University; e-mail: gev@sfedu.ru, 2, Chekhov street, Taganrog, 347900, Russia; phone: +78634360450; dr. of eng. sc.; Director of Institute of Computer Technology and Informational Security.

Ingabire Aline – e-mail: ingabire@sfedu.ru; phone: +79381069281; the department of synergetics and control processes; postgraduate student.

УДК 004.891.2

DOI 10.18522/2311-3103-2020-2-112-123

А.Н. Целых, В.С. Васильев, Л.А. Целых

ПРОБЛЕМА ВЫБОРА ДЕМПИНГ-ФАКТОРА В МОДЕЛИ ЭФФЕКТИВНЫХ УПРАВЛЕНИЙ ДЛЯ НАПРАВЛЕННЫХ ВЗВЕШЕННЫХ ЗНАКОВЫХ ГРАФОВ*

Рассматривается проблема выбора демпинг-фактора в модели эффективных управлений на основе максимизации передачи влияний для нечетких когнитивных моделей, представленных направленными взвешенными знаковыми графами. Для передачи влияния используется модель управления, реализующая развитие системы. Алгоритм эффективных управлений основан на решении оптимизационной задачи отыскания вектора внешних воздействий, максимизирующего накопленный рост приращений показателей вершин. Оптимальным управляющим воздействием признаётся управление, доставляющее максимум отношению квадрата нормы вектора отклика системы к квадрату нормы вектора управления. Демпинг-фактор такой модели управляет сравнительным масштабом прямого и косвенного влияния всех внутрифакторных связей системы в целом. Целью исследования является определение таких областей допустимых значений для получаемых решений, при которых (i) соблюдается условие непротиворечивости результата; (ii) изменение рангов вершин носит медленный характер. Под непротиворечивостью результата мы понимаем удовлетворение правилами работы системы в целом. Эти правила могут выражаться в наложении ограничений на статус вершин, на знак воздействий и откликов. В работе устанавливается значение демпинг-фактора, называемое резонансным, при котором происходит резонансный всплеск значения целевой функции задачи максимизации влияния, когда

* Работа выполнена при поддержке гранта Российского Фонда фундаментальных исследований № 19-01-00109.

сонаправленность между резонансным откликом и вызывающим его воздействием отсутствует. Выбор демпинг-фактора влияет на значение целевой функции задачи максимизации влияния и на вектор эффективного управления, на котором это решение достигается. Значение резонансного демпинг-фактора можно интерпретировать как предел возможной управляемости системы, т.е. предел потенциальной возможности воздействия на систему без причинения ей вреда. Оценка предложенного решения производится по степени устойчивости ранга узлов модели в зависимости от влияния изменений демпинг-фактора, алгоритмизации определения области его допустимых значений и форме проявления резонанса в границах значений демпинг-фактора.

Эффективное управление; демпинг-фактор; влияние; направленный взвешенный знаковый граф; нечеткая когнитивная карта.

A.N. Tselykh, V.S. Vasilev, L.A. Tselykh

THE PROBLEM OF CHOOSING A DUMPING FACTOR IN THE EFFECTIVE CONTROL MODEL FOR DIRECTED WEIGHTED SIGNED GRAPHS

The paper deals with the problem of choosing a dumping factor in the effective control model based on maximizing the transfer of impacts for fuzzy cognitive models represented by directed weighted signed graphs. To transfer influence, a management model is used that implements the development of the system. An effective control algorithm is based on solving the optimization problem of finding a vector of external influences that maximizes the accumulated increase in increments of vertex indicators. The optimal control effect is considered to be a control that provides the maximum ratio of the square of the norm of the response vector of the system to the square of the norm of the control vector. The dumping factor of this model controls the comparative scale of the direct and indirect influence of all intrafactor relationships of the system as a whole. The purpose of the study is to determine such areas of acceptable values for the obtained solutions, in which (i) the condition of consistency of the result is met; (ii) the change of vertex ranks is slow. By the sequence of results, we mean satisfaction with the rules of the system as a whole. These rules can be expressed in imposing restrictions on the status of vertices, on the sign of impacts and responses. Set the damping factor, called the resonance, where resonance occurs a surge in the value of the objective function the problem of maximization of the impact when there is no alignment between the resonant response and caused its effect. The choice of the dumping factor affects the value of the target function of the impact maximization problem and the vector of effective management on which this solution is achieved. The value of the resonant damping coefficient can be interpreted as the limit of the possible controllability of the system, i.e. limit the potential impact on the system without harming it. The proposed solution is evaluated based on the degree of stability of the rank of model nodes depending on the influence of changes in the damping coefficient, the algorithmization of determining the range of its acceptable values and the shape of the resonance within the values of the damping coefficient.

Effective control; dumping factor; influence; direct weighted sign graph; fuzzy cognitive map.

Введение. В данном исследовании рассматривается проблема выбора демпинг-фактора (ДФ) в модели эффективных управлений на основе максимизации передачи влияний на направленных взвешенных знаковых графах. Задача распространения влияния в сетях широко исследуется в последние годы. Сети могут быть представлены ненаправленными взвешенными знаковыми графами и направленными взвешенными знаковыми графами. Ненаправленные взвешенные знаковые графы представляют социальные сети, где веса на ребрах означают частоту контактов между вершинами. Самым известным алгоритмом, ранжирующим узлы в сетях, является алгоритм PageRank [1]. Вопросы выбора ДФ рассмотрены для ненаправленных взвешенных знаковых графов, матрица смежности которых является вещественной и симметричной [2].

Нечеткая когнитивная карта (НKK) является направленным взвешенным знаковым графом, где веса ребер имеют обязывающий (причинно-следственный) характер взаимоотношений между вершинами. Матрица смежности НKK является вещественной и несимметричной. Классическая модель Коско [3] предполагает

устойчивости в терминах технических систем и нахождение состояния равновесия в фиксированной точке. Ранжирование вершин происходит на основе достигаемых значений состояния вершин в этой неподвижной точке [4]. В некоторых работах упоминается ДФ, предназначенный для ускорения конвергенции [5]. Однако обоснование выбора ДФ не рассматривалось.

Задачи ранжирования вершин в моделях передачи влияний в направленных взвешенных знаковых графах с причинно-следственными отношениями на ребрах в литературе, как нам известно, не встречались. Таким образом, вопрос о том, как могут влиять вариации ДФ на рейтинг влиятельных вершин, индуцированный алгоритмом эффективных управлений [6, 7], в случае вещественной и несимметричной матрицы смежности требует соответствующих исследований.

1. Обзор публикаций по теме исследования. Анализ публикаций, использующих модели с ДФ, показал, что использование в модели ДФ может носить аддитивную [8, 9] или мультипликативную формы [10–14]. Аддитивная форма выражается в корректировке параметров модели за счет добавления или вычитания некоторой поправки. Мультипликативная форма выражается путем помещения множителя в модель. Характерным примером аддитивной формы ДФ является подход в статье [8]. За счет помещения в модель трех аддитивных демпинг-факторов здесь моделируется рост и задержка роста сети. Помещение в модель аддитивного ДФ проявляется в индивидуальной корректировке каждого показателя модели. В случае использования НКК такая корректировка отдельных параметров модели не допускается, но допустимо управление сравнительным масштабом прямого и косвенного влияния. Мультипликативность ДФ состоит во взвешивании прямых и косвенных влияний в целом для системы. Характерными представителями аддитивной формы можно назвать исследования α -centrality в [13, 14] и алгоритмов PageRank [10–12]. Здесь взвешивание прямых и косвенных влияний выражается помещением в модель множителя у матрицы смежности в целом. Вопрос выбора значения ДФ непосредственно рассматривается в статье [12], в котором отмечается чувствительность ранжирования узлов от изменения значения ДФ. Авторы в [12] обращают внимание на то, что малое изменение ДФ может приводить к существенным изменениям рангов вершин. Bressan и Peserico приводят пример графа, когда изменение ДФ даже на 0.001 “может вызвать полный разворот ранга”. Это указывает на важность и необходимость выбора значения ДФ в модели передачи влияний. Однако механизм применения ДФ в этих исследованиях применим для сетей с вероятностной матрицей и предполагаемым равномерным распределением переходов по исходящим дугам графа.

В статье [15] вводится параметр релаксации и оптимизируется для ускорения сходимости. В данном случае этот параметр относится не к модели, а к вычислительному методу нахождения решения. Можно провести аналогию между этим параметром релаксации и параметром регуляризации по Тихонову [18].

2. Описание метода.

2.1. Коэффициент демпфирования в алгоритме эффективного управления.

Рассматривается алгоритм эффективных управлений, предложенный в статьях [6, 7]. Метод, используемый в алгоритме, основан на решении оптимизационной задачи отыскания вектора (векторов) внешних воздействий, максимизирующего накопленный рост приращений показателей вершин для системы, представленной ориентированным знаковым взвешенным графом с причинно-следственными связями на ребрах. В качестве графа используется нечеткая когнитивная карта (НКК) исследуемой системы. Для передачи влияния используется модель управления, названная моделью роста:

$$x_j = u_j + \delta \sum_{i=1}^n a_{ij} x_i,$$

где x_j – значение приращения показателя вершины v_j ; u_j – управляющее воздействие на вершину v_j ; δ – ДФ; a_{ij} – вес дуги соединяющей вершину v_i с v_j , $-1 \leq a_{ij} \leq 1$; x_i – значение приращения показателя вершины v_i . Оптимальным управляющим воздействием признаётся управление $\mathbf{u}$, доставляющее максимум отношению квадрата нормы вектора показателей $\mathbf{x}$ к квадрату нормы вектора управления $\mathbf{u}$:

$$(\mathbf{x}, \mathbf{x})/(\mathbf{u}, \mathbf{u}) \rightarrow \max.$$

Результатом работы алгоритма являются два набора ранжированных в порядке важности узлов графа. Важность узла определяется значением (по модулю) компонент векторов воздействия и влияния. Один набор узлов графа представляет собой значения компонент вектора управляющих воздействий, другой – значения компонент вектора влияния для соответствующих вершин графа. Оба набора вершин упорядочены по возрастанию значений. Отбор наиболее влиятельных и влияющих узлов (эффективных управлений) производится на основе кумулятивного эффекта (компонент векторов), выражающего потенциал системы в целом. Метод основан на спектральных свойствах матрицы смежности графа и определении ее резонансных свойств, а именно нахождении собственных чисел и собственных векторов матрицы смежности.

Как показано в разделе 1 во многих моделях ДФ использовался для достижения сходимости решения. При этом предполагается, что наличие любого решения системы уравнений при достижении сходимости обеспечивает правильное решение. Тем не менее, в работе [5, 15–17] отмечается, что наблюдались случаи получения противоречивых результатов. В этом случае структура модели анализировалась и изменялась через корректировку весов и факторов для получения более разумных результатов. Такие случаи встречаются в большинстве НКК. Однако анализ показывает, что это не всегда объясняется неправильной структурой модели. Тогда изменение весов, связей и факторов, по сути, может являться простой “подгонкой” решения под ожидаемый результат.

Проиллюстрируем пример такого решения, где наблюдается эффект “перевернутого стола”. Рассмотрим НКК для определения клинического риска при администрировании назначений медицинских препаратов из статьи [19]. НКК содержит 30 вершин-факторов, 119 ребер с общим весом 84.2. Проведем анализ качества получаемых результатов:

- 1) для рангов, полученных вершинами НКК;
- 2) для соответствия воздействий и откликов, которые получили вершины, при разных значениях ДФ.

Анализ рангов, полученных вершинами НКК, показал, что для этой НКК изменение ДФ существенным образом влияет на получаемый результат. Рис. 1 показывает изменение ранга влиятельных вершин (рис. 1,а) и вершин, получивших отклик системы (рис. 1,б) при изменениях ДФ. Так, например, вершина 5 переместилась с 9-го места на 1-е, а вершина с номером 1 переместилась с 1-го места на 20-е место. То есть, если система работает в полную мощность так, как задал факторы и значения связей эксперт, тогда вершиной, получившей максимальный отклик, является вершина с номером 5 (под $\delta=1.0$), а вовсе не вершина с номером 1 (под $\delta=0.1$). И, соответственно, самой влиятельной вершиной, занявшей 1-е место под $\delta=1.0$, является вершина с номером 23, а вершина с номером 30 переместилась с 1-го места под $\delta=0.1$ на 4-е место под $\delta=1.0$. 24-я вершина переместилась с 11-го места на 3-е место, соответственно.

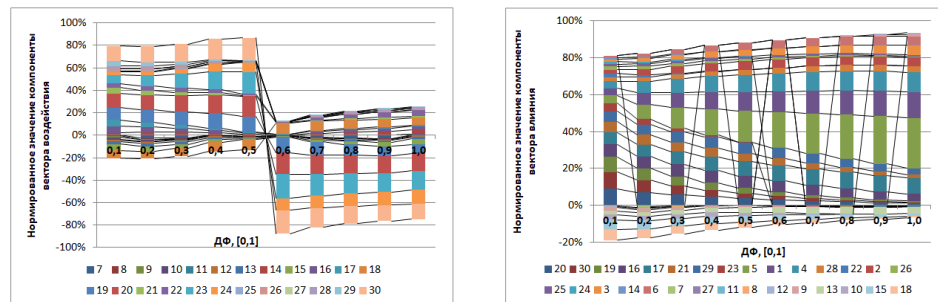


Рис. 1. Изменение нормированных вкладов (рангов) компонент (а) в векторе воздействия (влияющих вершин) и (б) в векторе влияния (вершин, получивших отклик системы) при изменении $\Delta\Phi$, $0 < \delta \leq 1.0$

Анализ соответствия воздействий и откликов, которые получили вершины при разных значениях $\Delta\Phi$, проведем с точки зрения смыслового содержания факторов системы. Например, увеличение факторов усталости, спешки, стресса, аварийности и рабочей нагрузки (соответственно, номера вершин 1-6), очевидно, увеличивают ошибки, допускаемые персоналом (соответственно, номера вершин 19, 20, 23, 24, 30). Такие вершины имеют “связанные знаки” изменений характеристик вершин. В данном случае знак у всех значений компонент векторов должен быть положительным. Однако полученное решение в интервале $0.6 \leq \delta \leq 1$, не удовлетворяет требованиям непротиворечивости результата (табл. 1), т.е. знак у компонент векторов вершин второй группы изменился на противоположный.

Таблица 1

Значения компонент откликов и воздействий для связанных знаков изменений характеристик вершин

Номер вершины	Значения компонент вектора при соответствующих значениях $\Delta\Phi$				
	0.1	0.2	0.3	0.4	0.5
1	0.174	0.2913	0.3463	0.3709	0.3842
2	0.0792	0.1326	0.1587	0.1681	0.1685
3	0.0712	0.1271	0.1639	0.1867	0.2001
4	0.162	0.2673	0.3108	0.3224	0.3214
5	0.1803	0.342	0.4539	0.531	0.5883
6	0.0411	0.0778	0.1034	0.1207	0.1311
19	0.26425	0.15133	0.07505	0.03162	0.00937
20	0.3077	0.18427	0.09438	0.0407	0.01217
23	0.18292	0.12823	0.07653	0.03774	0.01291
24	0.05646	0.04275	0.02871	0.01583	0.00586
30	0.31064	0.18763	0.09689	0.04179	0.01238
Номер вершины	Значения компонент вектора при соответствующих значениях $\Delta\Phi$				
	0.6	0.7	0.8	0.9	1.0
1	-0.3937	-0.3998	-0.4029	-0.4011	-0.3928
2	-0.165	-0.1603	-0.1562	-0.1556	-0.1619
3	-0.2083	-0.2138	-0.2191	-0.2288	-0.2492
4	-0.3156	-0.3079	-0.2996	-0.2897	-0.2769
5	-0.6325	-0.668	-0.6973	-0.7209	-0.7386
6	-0.1348	-0.1335	-0.1286	-0.1206	-0.1095

19	0.001	0.00504	0.00605	0.0058	0.00526
20	0.00136	0.00711	0.00896	0.00881	0.00758
23	0.00153	0.00842	0.01078	0.01108	0.01062
24	0.00077	0.00461	0.00648	0.00709	0.00693
30	0.00136	0.00686	0.00833	0.00793	0.00671

Для нашей модели при разных значениях ДФ, $0 < \delta \leq 1$, получаемые решения отображают масштабированный результат силы действия всех внутрифакторных связей системы. Изменение ДФ в уравнении (1) можно понимать следующим образом. Если $\delta=0.1$, то влияние косвенных воздействий системы в целом проявляется только на 0.1 от заявленной мощности системы в целом; если $\delta=1$, то межфакторные связи работают в полную силу, косвенное влияние в максимальной степени передается внутри системы. Заметим, что наблюдаются неудовлетворительные результаты при значениях $0.6 \leq \delta \leq 1$. т.е. наблюдается эффект "перевернутого стола". То есть сила косвенного влияния такова, что "имеется достаточно энергии, чтобы перевернуть стол". При этом достигается сходимость решения при всех значениях ДФ. Таким образом, наличие сходимости еще не гарантирует правильности результата.

Следовательно, требуется найти такие области допустимых значений для получаемых решений, при которых (i) соблюдается условие непротиворечивости результата; и (ii) изменение рангов вершин носит медленный характер. Под непротиворечивостью результата мы понимаем удовлетворение правилам работы системы в целом. Эти правила могут выражаться в наложении ограничений на статус вершин и на знаки воздействий и откликов.

2.2. Определение области допустимых значений. В принятом формате ограничения для входных данных соответственно разделяются на 3 вида:

1. Равенство нулю компоненты (воздействий, отклика), $x_i = 0, u_i = 0$.
2. Неотрицательность компоненты (воздействий, отклика), $x_i \geq 0, u_i \geq 0$.
3. Неположительность компоненты (воздействий, отклика), $x_i \leq 0, u_i \leq 0$.

Первая группа ограничений представляет узлы модели, являющиеся неуправляемыми факторами (например, факторы внешней среды, на которые не может быть оказано непосредственное управляющее воздействие). Вторая и третья группы определяют факторы, характеристики которых могут иметь только положительные или только отрицательные изменения характеристик, в том числе и связанные. Поскольку все ограничения вводятся в систему уравнений, то решение также будет удовлетворять требуемым соотношениям знаков воздействий и откликов.

Пусть $G = \langle V, E \rangle$ граф социально-экономической системы, где $V = \{V_1, V_2, \dots, V_N\}$ – множество вершин и $E \in V \times V$ – множество ребер. Этот граф определяется матрицей смежности $A^T = \|a_{i,j}\|_n$, где $a_{i,j}$ – вес ребра от вершины V_i к вершине V_j на интервале $-1 \leq a_{i,j} \leq 1$. Рассматривается задача управления для модели роста с простой зависимостью воздействия и отклика:

$$(I - \delta A^T)x = u, \quad (1)$$

где I – единичная матрица; $x = (x_1 \ x_2 \ \dots \ x_n)^T$ – вектор влияний, называемый откликом системы; $u = (u_1 \ u_2 \ \dots \ u_n)^T$ – вектор управляющих воздействий, δ – ДФ, задаваемый в интервале $0 < \delta \leq 1$. Эта задача может быть решена либо в отношении влияний, либо в отношении откликов. Рассматривается решение задачи относительно влияний.

Задача максимизации

$$x^T x / u^T u \rightarrow \max$$

при ограничении $x^T x = 1$ (или $u^T u = 1$) эквивалентна задаче

$$x^T B x \rightarrow \min, \quad (2)$$

где $B = (I - \delta A)(I - \delta A^T)$ [20]. Заметим, что при $\delta = 0$ выполняется $x = u$, и $u^T x > 0$, т.е. сонаправлен и даже тождественен воздействию.

Любую матрицу можно представить в виде суммы симметричной $\mathbf{A}_s = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T)$ и антисимметричной $\mathbf{A}_a = \frac{1}{2}(\mathbf{A} - \mathbf{A}^T)$ матриц

$$\mathbf{A} = \mathbf{A}_s + \mathbf{A}_a = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T) + \frac{1}{2}(\mathbf{A} - \mathbf{A}^T).$$

В общем случае

$$\mathbf{u}^T \mathbf{x} = \mathbf{x}^T (\mathbf{I} - \delta \mathbf{A}) \mathbf{x} = \mathbf{x}^T \mathbf{x} - \delta \mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{x} - \frac{1}{2} \delta \mathbf{x}^T (\mathbf{A} + \mathbf{A}^T) \mathbf{x}.$$

В отношении симметричной матрицы $\mathbf{A}_s$ можно утверждать, что все её собственные числа вещественные, и

$$\lambda_{\min} \mathbf{x}^T \mathbf{x} \leq \frac{1}{2} \mathbf{x}^T (\mathbf{A} + \mathbf{A}^T) \mathbf{x} \leq \lambda_{\max} \mathbf{x}^T \mathbf{x},$$

где $\lambda_{\min}$ и $\lambda_{\max}$ – минимальное и максимальное собственные числа матрицы $\mathbf{A}_s$. Утверждать, что все собственные числа одного знака можно только для знакоопределённых симметричных матриц. Но если $\lambda_{\max} > 0$, то при $\delta = \lambda_{\max}^{-1}$ происходит смена знака с $\mathbf{u}^T \mathbf{x} > 0$ на $\mathbf{u}^T \mathbf{x} < 0$, т.е. сонаправленные компоненты воздействия и отклика становятся противоположно направленными. Для задачи безусловной оптимизации (2) и под моделью (1) это произойдёт при единственном значении δ . Тогда индикатором перехода может служить знак формы $\mathbf{u}^T \mathbf{x}$.

Определение 1. ДФ (δ) – это коэффициент, выражающий уровень косвенного воздействия сравнительно с уровнем прямого воздействия.

Определение 2. Существует допустимый диапазон области ДФ $Im(DF) = [0, \delta_{adm}]$, в границах которого гарантированно преобладает сонаправленность по компонентам векторов отклика и воздействия, где $\delta_{adm} = 1/\lambda_{\max \mathbf{A}_s}$.

Определение 3. Существует область резонанса ДФ $Im(DF) = [\delta_{adm}, 1]$, в границах которой наблюдается стремительная перестройка рангов вершин по компонентам векторов отклика и воздействия и разворот ранжирования.

Определение 4. Существует значение ДФ, называемое резонансным ДФ (δ_{res}), при котором происходит резонансный всплеск значений целевой функции задачи максимизации влияния. Именно в точке резонансного всплеска происходит смена знака с $\mathbf{u}^T \mathbf{x} > 0$ на $\mathbf{u}^T \mathbf{x} < 0$.

Замечание. Экстремальное решение в постановке безусловной оптимизационной задачи (2) приходится на один из собственных векторов матрицы $\mathbf{B}$. Поскольку $\mathbf{x}$ и $\mathbf{u}$ связаны своей матрицей, то они могут различаться направлениями даже для задачи безусловной оптимизации, т.е. вектора $\mathbf{x}$ и $\mathbf{u}$ не являются сонаправленными. Заметим, что при $\delta = 0$ отсутствуют косвенные воздействия и имеет место совпадение направлений $\mathbf{x}$ и $\mathbf{u}$. При увеличении доли косвенного воздействия происходит расхождение направлений.

При определенном значении возможно появление нулевого (или близкого к нулю) собственного числа матрицы $(\mathbf{I} - \delta \mathbf{A}^T)$. Это означает возникновение резонанса, когда в системе на ничтожно малое воздействие возникает сравнительно большой отклик. Причем этот отклик определяется исключительно внутренними свойствами системы. Это означает, что различные малые воздействия могут вызывать большие резонансные отклики одного направления, т.е. сонаправленность между резонансным откликом и вызывающим его воздействием отсутствует.

3. Эксперимент. Реализуемость предлагаемого подхода, определяющего области допустимых значений, была проверена на тестовом примере, рассмотренном в разделе 2.1. Определение области допустимых значений решений, при которых соблюдается условие непротиворечивости работы модели, рассматривается при следующей системе ограничений для входных данных:

- а) Равенство нулю компоненты воздействий $x_i = 0$ для $i = 1, \dots, 6$ (неуправляемые факторы).
- б) Неотрицательность компоненты отклика $u_i \geq 0$ для $i = 7, \dots, 13, 15, 18, 27$.
- с) Неположительность компоненты отклика $u_i \leq 0$ для $i = 1, \dots, 6, 14, 16, 17, 19, \dots, 26, 28, \dots, 30$.

Первая группа ограничений представляет узлы модели, являющиеся неуправляемыми факторами (например, факторы внешней среды, на которые не может быть оказано непосредственное управляющее воздействие). Вторая и третья группы определяют факторы, характеристики которых могут иметь только положительные или только отрицательные изменения характеристик, в том числе и связанные. Поскольку все ограничения вводятся в систему уравнений, то решение также будет удовлетворять требуемым соотношениям знаков воздействий и откликов.

Процесс выявления эффективных управлений с учетом допустимого диапазона области значений ДФ включает следующие шаги:

Шаг 1. Определение области допустимых значений ДФ на основе вычисления максимального собственного числа матрицы смежности модели.

Шаг 2. Формализация условий работы модели в виде системы ограничений.

Шаг 3. Верификация и валидация модели.

Шаг 4. Вычисление компонент векторов отклика и воздействия в области допустимых решений.

Шаг 5. Отбор вершин, оказывающих наибольшее влияние (вершин воздействия), и вершин, получивших наибольшее влияние (вершин отклика).

Как показано в табл. 2, значения $\mathbf{u}^T \mathbf{x}$, вычисленные на шаге 3, меняют знак с “плюса” на “минус” при $\delta_{res} = 0.585$. Соответствующее значение границы допустимых значений ДФ равно $\delta_{adm} = 1/\lambda_{max_{A_5}} = 0.2550648$. Отметим, что $\delta_{res} > \delta_{adm}$, что согласуется с определением 2 и определением 3.

Таблица 2

Значения $\mathbf{u}^T \mathbf{x}$ с шагом 0.1.

Тип модели	Значения $\mathbf{u}^T \mathbf{x}$ для ДФ					
	0.1	0.2	0.3	0.4	0.5	0.6
Модель без ограничений	0.618237	0.306535	0.116972	0.032865	0.005236	-0.00075
Модель с ограничениями	0.620998	0.281798	0.100844	0.028337	0.004801	-0.000825
Модель	Значения $\mathbf{u}^T \mathbf{x}$ для ДФ					
	0.7	0.8	0.9	1.0		
Модель без ограничений	-0.00066	0.000278	0.000864	0.001045		
Модель с ограничениями	0.044707	0.027549	0.017148	0.010729		

Соответствующее поведение $1/\lambda_i$ при $0 < \delta \leq 1$ для НКК в модели определения клинического риска при администрировании назначений медицинских препаратов из статьи [19] показано на рис. 2. Продемонстрировано существование резонансного значения ДФ, $\delta_{res} = 0.585$, когда возникает резонансный всплеск значения целевой функции задачи максимизации влияния. Именно в точке резонансного всплеска знак переключается с $\mathbf{u}^T \mathbf{x} > 0$ на $\mathbf{u}^T \mathbf{x} < 0$, как следует из данных, приведенных в табл. 2. В данном примере наблюдается резонансный всплеск в форме разрыва функции.

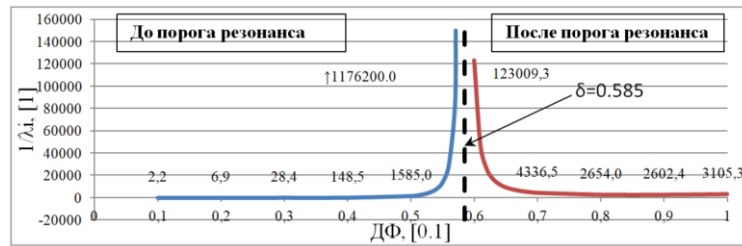


Рис. 2. Изменение $1/\lambda_i$ при изменениях ДФ, $0 < \delta \leq 1$

4. Анализ полученных результатов. Оценка предложенного решения про-изводилась по следующим вопросам:

- (i) В какой степени ранжирование, полученное в результате работы алгоритма эффективных управлений, зависит от изменений ДФ?
- (ii) Какую характеристику выражает ДФ?
- (iii) Как определить области допустимых значений ДФ?
- (iv) Какая характеристика определяет выбор ДФ в пределах области допустимых значений.

Проверка устойчивости ранга показала, что наблюдается существенное изменение рангов узлов для любых дискретных вариаций ДФ. Серия проведенных экспериментов показала, что интервал значений распадается на 3 типа интервалов:

- ♦ 1-го типа – интервалы медленного изменения рангов вершин (или сохранения рангов) при непротиворечивости результатов;
- ♦ 2-го типа – интервалы стремительной перестройки рангов вершин; и
- ♦ 3-го типа – интервалы медленного изменения рангов вершин (или сохранения рангов) при противоречивости результатов.

Очевидно, что интервалы 2-го и 3-го типов должны быть исключены из рассмотрения.

Нахождение границ интервалов 1-го типа предписывается теоремой 1 и сводится к нахождению максимального собственного числа матрицы A_s , что представляет собой очень простую задачу сложности $O(n^3)$. Тогда эта граница покажет интервал без стремительных изменений рангов и с непротиворечивым результатом, что достаточно для лица, принимающего решения (ЛПР). Для рассматриваемого примера это значение равно $\delta = 1/\lambda_{\max A_s} = 0.2550648$.

Интервалы 2-го и 3-го типов обнаруживаются несложными алгоритмами с четко формализуемыми критериями. Исключение интервалов 2-го типа основывается на нахождении значений ДФ с фактической сменой знака с $\mathbf{u}^T \mathbf{x} > 0$ на $\mathbf{u}^T \mathbf{x} < 0$. Это сводится к решению задачи максимизации влияния для некоторых характерных значений ДФ с использованием метода дихотомии. Интервалы 3-го типа следуют сразу за интервалами 2-го типа и не представляют интереса для ЛПР. Фактическая смена знака с $\mathbf{u}^T \mathbf{x} > 0$ на $\mathbf{u}^T \mathbf{x} < 0$ является точкой резонанса системы.

Для ЛПР вход системы в резонанс означает, что отклик системы определяется внутренними свойствами системы и перестает зависеть от управляющего воздействия. Тогда резонансный отклик вызывают различные управляющие воздействия. Причем отношение уровня отклика к уровню воздействия очень велико. Можно сказать, что ЛПР лишается возможности управления системой, а получает только возможность возбуждения ее резонансного отклика, который может приобрести и взрывной характер.

Тогда значение резонансного ДФ является разумной границей, по которой можно судить об устойчивости исследуемой системы. Для рассматриваемого примера это значение равно если $\delta = 0.585$. В этом случае можно сказать, что запас прочности такой системы позволяет надеяться, что система будет устойчиво работать и “не пойдет в разнос” до достижения силы действия всех внутрифакторных связей системы до достижения $\delta = 0.585$. Тогда ЛПР должен понимать, что при продолжающемся воздействии на систему (после достижения этого предела) система войдет в резонанс и может разрушиться. Поэтому требуются другие действия, которые изменят саму систему и создадут возможность исключения вредного действия резонанса. Например, создание других факторов и связей в системе. При получении для системы низких значений ДФ можно сделать вывод о небольшом запасе прочности системы и необходимости принятия срочных действий ЛПР.

Полученный нами результат нахождения резонансной границы имеет аналогию с результатами, полученными в исследованиях критической точки и ее природы [2, 9, 21]. В этой критической точке, именуемой в разных источниках также, как коэффициент затухания [13], происходит делокализация эпидемического распространения влияний в сети, представленной вещественной и симметричной матрицей. В этих исследованиях приводятся модели, в которых определяющее значение имеет главный собственный вектор и критическая точка определяется как $1/\lambda_{max}$ (α -centrality, PageRank и др.). Однако этот подход не может быть использован в случае, если хотя бы одно собственное число является комплексно-значным.

Таким образом, наш вклад для рассматриваемой модели эффективных управлений на направленных взвешенных знаковых графах заключается в следующем:

1. Определено, что выбор ДФ влияет не только на значение целевой функции задачи максимизации влияния, но и на вектор эффективного управления, на котором это решение достигается.
2. В области ДФ существуют границы быстрого изменения целевой функции и быстрого изменения решения, которое проявляется в инверсии рангов вершин.
3. Если собственное число матрицы смежности, определяющее это значение ДФ, является действительным, то резонансный всплеск имеет форму разрыва функции. Если это число комплексное, то целевая функция будет иметь конечный максимум.
4. Для графов реального мира выбор ДФ до его дальней границы может основываться на реалистичности значения целевой функции, показывающей соотношение воздействия и отклика системы.
5. Значение резонансного ДФ можно интерпретировать как предел возможной управляемости системы, т.е. предел потенциальной возможности воздействия на систему без причинения ей вреда.

Заключение. В настоящем исследовании мы показываем, что для рассматриваемой модели эффективных управлений на направленных взвешенных знаковых графах:

- (i) демпинг-фактор выражает масштабирование косвенного влияния, передаваемого по дугам графа;
- (ii) существует граница области допустимых значений полученного решения, для $Im(DF) = [0, 1/\lambda_{max_{As}}]$, в пределах которого гарантированно не произойдет стремительная перестройка рангов вершин и разворот ранжирования;
- (iii) значение ДФ в точке резонанса, при смене знака с $\mathbf{u}^T \mathbf{x} > 0$ на $\mathbf{u}^T \mathbf{x} < 0$, является разумной границей при принятии решений и означает своеобразный “запас прочности” системы;
- (iv) выбор ДФ в границах допустимых значений может основываться на реалистичности значения целевой функции.

Поддержка. Работа выполнена при поддержке гранта Российского Фонда фундаментальных исследований № 19-01-00109.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Page L., Brin S., Motwany R., and Winograd T. The PageRank citation ranking: Bringing order to the Web, in *7th International World Wide Web Conference*, 1998, pp. 161-172.
2. Goltsev A.V., Dorogovtsev S.N., Oliveira J.G., and Mendes J.F.F. Localization and Spreading of Diseases in Complex Networks, *Phys. Rev. Lett.*, Sep. 2012, Vol. 109, No. 12, pp. 128702.
3. Kosko B. Fuzzy cognitive maps, *Int. J. Man. Mach. Stud.*, Jan. 1986, Vol. 24, No. 1, pp. 65-75.
4. Knight C.J.K., Lloyd D.J.B., and Penn A.S. "Linear and sigmoidal fuzzy cognitive maps: An analysis of fixed points," *Appl. Soft Comput.*, Feb. 2014, Vol. 15, pp. 193-202.
5. Hobbs B.F., Ludsins S.A., Knight R.L., Ryan P.A., Biberhofer J., and Ciborowski J.J.H. Fuzzy cognitive mapping as a tool to define management objectives for complex ecosystems, *Ecol. Appl.*, 2002, Vol. 12, No. 5, pp. 1548-1565.
6. Tselykh A.N., Vasilev V.S., Tselykh L.A., and Barkovskii S.A. Method Maximizing the Spread of Influence in Directed Signed Weighted Graphs, *Adv. Electr. Electron. Eng.*, Jul. 2017, Vol. 15, No. 2.
7. Tselykh A., Vasilev V., and Tselykh L. Management of Control Impacts Based on Maximizing the Spread of Influence, *Int. J. Autom. Comput.*, Feb. 2019.
8. Cai G., Wang R., and Qiang B. Online Social Network Evolving Model Based on Damping Factor, *Procedia Comput. Sci.*, 2012, Vol. 9, pp. 1338-1344.
9. De Arruda G.F., Rodrigues F.A., and Moreno Y. Fundamentals of spreading processes in single and multilayer complex networks, *Phys. Rep.*, Oct. 2018, Vol. 756, pp. 1-59.
10. Son S.-W., Christensen C., Grassberger P., and Paczuski M. PageRank and rank-reversal dependence on the damping factor, *Phys. Rev. E*, Dec. 2012, Vol. 86, No. 6, pp. 066104.
11. Boldi P., Santini M., and Vigna S. PageRank as a function of the damping factor, in *Proceedings of the 14th international conference on World Wide Web - WWW '05*, 2005, 557 p.
12. Bressan M. and Peserico E. Choose the damping, choose the ranking?, *J. Discret. Algorithms*, Jun. 2010, Vol. 8, No. 2, pp. 199-213.
13. Bonacich P. and Lloyd P. Eigenvector-like measures of centrality for asymmetric relations, *Soc. Networks*, Jul. 2001, Vol. 23, No. 3, pp. 191-201.
14. Katz L. A new status index derived from sociometric analysis, *Psychometrika*, Mar. 1953, Vol. 18, No. 1, pp. 39-43.
15. Salmeron J. Fuzzy cognitivemaps for artificial emotions forecasting, *Applied Soft Computing*, 2012, No. 12 (12), 779, pp. 3704-3710.
16. Kok K. The potential of fuzzy cognitive maps for semi-quantitative scenario development, with an example from Brazil, *Global Environmental Change*, 2009, Vol. 19 (1), pp. 122-133.
17. Pei S., Wang J., Morone F., & Makse H. Influencer identification in dynamical complex systems, *Journal of Complex Networks*, 2019. Available at: <https://doi.org/10.1093/comnet/cnz029>.
18. Tikhonov A. and Arsenin V. Solutions of Ill-Posed Problems. New York: Wiley, 1977.
19. Bevilacqua M., Ciarapica F.E., and Mazzuto G. Fuzzy cognitive maps for adverse drug event risk management, *Saf. Sci.*, Feb. 2018, Vol. 102, pp. 194-210.
20. Tselykh A., Vasilev V., Tselykh L., and Fernando A.F. Ferreira. Influence control method on directed weighted signed graphs with deterministic causality. *Annals of Operations Research*, 2020. Available at: <https://doi.org/10.1007/s10479-020-03587-8>.
21. Boguñá M., Castellano C., and Pastor-Satorras R. Nature of the Epidemic Threshold for the Susceptible-Infected-Susceptible Dynamics in Networks, *Phys. Rev. Lett.*, Aug. 2013, Vol. 111, No. 6, p. 068701.

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

Целых Александр Николаевич – Южный федеральный университет; e-mail: ant@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: +79185562047; кафедра ИАСБ; д.т.н.; профессор.

Васильев Владислав Сергеевич – e-mail: vsvasilev@sfedu.ru; кафедра ИАСБ; к.т.н.; доцент.

Целых Лариса Анатольевна – Таганрогский институт имени А.П. Чехова (филиал) Ростовского государственного экономического университета (РИНХ); e-mail: l.tselykh58@gmail.com; 347936, г. Таганрог, ул. Инициативная, 48; тел.: +79185695760; кафедра экономики и предпринимательства; к.э.н.; доцент.

Tselykh Alexander Nikolaevich – Southern Federal University; e-mail: ant@sfedu.ru; 44, Nekrasovskiy, Taganrog, 347928, Russia; phone: +79185562047; the department IASB; dr. of eng. sc.; professor.

Vasilev Vladislav Sergeevich – e-mail: vsvasilev@sfedu.ru; the department IASB; dr. of eng. sc.; the senior lecturer.

Tselykh Larisa Anatolievna – Chekhov Taganrog Institute (branch) of Rostov State University of Economics; e-mail: l.tselykh58@gmail.com; Initsiativnaya, 48, Taganrog, 347936, Russia; phone: +79185695760; the department of Economics and business; dr. of ec. sc.; the senior lecturer.

УДК 629.113 (075)

DOI 10.18522/2311-3103-2020-2-123-133

А.Л. Береснев, М.А. Береснев

РАЗРАБОТКА МАТЕМАТИЧЕСКОЙ МОДЕЛИ ДЕТОНАЦИИ В СОВРЕМЕННОМ ДВС

Рассматриваются недостаточно изученные аспекты управления двигателем внутреннего сгорания, такие как использование детонации. При детонации вместо постоянного фронтального пламени в зоне сгорания образуется детонационная волна, несущаяся со сверхзвуковой скоростью. В волне сжатия детонируют топливо и окислитель. Этот процесс, с точки зрения термодинамики, повышает коэффициент полезного действия двигателя. В двигателях внутреннего сгорания существуют два различных режима распространения горения топлива воздушной смеси: дефлаграция и детонация. Двигатели работающие в режиме детонационного сгорания в настоящее время не используются и их возможности наиболее интересны. Современные двигатели внутреннего сгорания плохо переносят режим детонации, но возможность непродолжительного сгорания части топливовоздушной смеси с детонацией заложена в их конструкции. Ситуация возникновения детонации в настоящее время постоянно изучается, к ней относятся, как к вредной составляющей процесса горения, для борьбы с которой необходимо совершенствовать двигатель, его управление и использование современного топлива. Предлагается использовать этот, считающийся случайным, процесс для повышения крутящего момента и мощности двигателя внутреннего сгорания. Рассматривается возможность использования детонационного сгорания топливно-воздушной смеси в двигателе внутреннего сгорания, как полезной части рабочего процесса и делается предположение о возможности управления горением топливно-воздушной смеси в смешанном режиме, позволяющем улучшить индикаторные показатели. Высказаны предположения для создания модели и произведено моделирование давления в цилиндре на этапе горения топливовоздушной смеси с частичной детонацией. Произведено определение способа расчета тепловыделения, который является одним из важнейших этапов создания математической модели, поскольку от этого зависит точность и адекватность рассчитываемых параметров, как для режимов дефлаграционного горения топливовоздушной смеси, так и с использованием детонационного сгорания. Предложенная процедура расчета параметров рабочего цикла ДВС позволяет проводить вычисления в реальном времени и учесть влияние состава бинарного топлива на показатели мощности, экономичности, механической и динамической нагрузки на детали кривошипно-шатунного механизма, а также термической нагрузки на двигатель.

Двигатель внутреннего сгорания; детонация; модель.

A.L. Beresnev, M.A. Beresnev

DEVELOPMENT OF KNOCK MATHEMATICAL MODEL FOR MODERN IC ENGINE

The paper examines under-studied aspects of internal combustion engine management, such as the use of knock. In knock, instead of a constant frontal flame, a detonation wave is formed in the combustion zone, carrying at supersonic speed. Fuel and oxidant are detonated in compression wave. This process, from the point of view of thermodynamics, increases the efficiency of the engine. In internal combustion engines, there are two different modes of combustion propagation of air mixture fuel: deflagration and detonation. Engines operating in detonation combustion mode are currently not used and their capabilities are most interesting. Modern internal combustion engines do not carry the detonation mode well, but the possibility of short-term combustion of part of the fuel-air mixture with detonation is embedded in their design. The situation of detonation is currently constantly being studied, it is considered as a harmful component of the combustion process, which requires improvement of the engine, its control and use of modern fuel. It is proposed to use this, considered random, process to increase the torque and power of the internal combustion engine. The possibility of using detonation combustion of fuel-air mixture in internal combustion engine as a useful part of the working process is considered, and the possibility of controlling combustion of fuel-air mixture in a mixed mode is assumed, which allows to improve indicator parameters. Assumptions have been made to create a model and the cylinder pressure has been simulated during the combustion phase of the partially detonated fuel-air mixture. The method of heat generation calculation has been determined, which is one of the most important stages of mathematical model creation, as this determines accuracy and adequacy of calculated parameters, both for deflagration combustion modes of fuel-air mixture and using detonation combustion. Proposed procedure for calculation of internal combustion engine operating cycle parameters makes it possible to carry out real-time calculations and account for influence of binary fuel composition on parameters of power, economy, mechanical and dynamic load on parts of crank mechanism, as well as thermal load on engine.

Internal combustion engine; knock; model.

Введение. Советский физик Я.Б. Зельдович предложил идею детонационного двигателя в статье «Об энергетическом использовании детонационного сгорания» [1]. Значительное увеличение эффективности такого двигателя по сравнению с обычным достигается благодаря тому, что процесс детонационного горения протекает на два порядка быстрее дефлаграции. В результате при постоянном объеме быстро возрастает давление. Выделение полученного таким образом количества энергии, к сожалению, вполне способно разрушить двигатель автомобиля, поэтому такой процесс часто ассоциируется со взрывом. Работы в данном направлении ведутся во всем мире с переменным успехом [2–5]. В настоящее время использование детонации достигло конкретных результатов, в августе 2016 года российским учёным удалось создать впервые в мире полноразмерный жидкостный реактивный двигатель, работающий на принципе детонационного сгорания топлива [6].

Детонация в современном ДВС. При детонации вместо постоянного фронтального пламени в зоне сгорания образуется детонационная волна, несущаяся со сверхзвуковой скоростью. В волне сжатия детонируют топливо и окислитель. Этот процесс, с точки зрения термодинамики, повышает КПД двигателя благодаря компактности зоны [1].

В двигателях внутреннего сгорания (ДВС) существуют два различных режима распространения горения топливо воздушной смеси (ТВС).

ДВС работающий в режиме детонационного сгорания в настоящее время изучены недостаточно [6–9].

Необходимо учитывать: современные двигатели внутреннего сгорания плохо переносят режим детонации, но возможность непродолжительного сгорания части ТВС с детонацией заложена в их конструкции.

Таким образом ситуации возникновения детонации в настоящее время постоянно изучаются, к ним как правило относятся, как к вредной составляющей процесса горения, для борьбы с которой необходимо совершенствовать ДВС, управление им и совершенствование топлива [7, 11, 12].

Известен эксперимент [9] подтверждающий увеличение мощности и момента при использовании частичного сгорания ТВС в режиме детонации, он позволяет определить разновидности индикаторных диаграмм и их влияние на внешние скоростные характеристики и подтверждает гипотезу возможности использования части топлива для детонационного горения.

Для изучения возможности использования детонационного горения части ТВС в ДВС и управления этим процессом, необходимо разработать математическую модель, которая в дальнейшем позволит выбрать управляющие параметры и определить диапазоны их изменения. В связи с недостаточной повторяемостью параметров горения ТВС в камере сгорания, как между цилиндрами одного двигателя, так и в одном цилиндре от цикла к циклу, что вызвано, как объективными так и субъективными причинами существующими в современном проектировании ДВС и уровне двигателестроения в целом, необходима возможность простого адаптирования модели к конкретным условиям. В то же время из-за плохой межцикловой повторяемости индикаторной диаграммы, нет необходимости добиваться высокой повторяемости расчетов давления внутри цилиндра. Наиболее важным является время расчетов и возможность прогнозирования в дальнейшем количества диаграмм определенных в эксперименте [9].

Моделирование внутрицилиндровых процессов в двигателе внутреннего сгорания с помощью математической модели позволит описать эффективные показатели ДВС при двух режимной работе. Далее модель может использоваться для расчета ТВС и создания алгоритма определения управляющих параметров.

Определение входных и выходных данных. На выходе модели должен формироваться набор параметров, позволяющий определить оптимальный УОЗ для данных входных условий в реальном масштабе времени. Для этого следует оценить влияние топлива различного состава и значения угла опережения зажигания на основные показатели ДВС. Поскольку основным процессом в ДВС является рабочий цикл, от которого зависят мощность, экономичность и надежность работы двигателя [11], то в качестве выходных данных модели могут быть приняты следующие основные параметры и показатели рабочего цикла ДВС [23]:

- ◆ среднее индикаторное давление p_{ind} – показатель удельной работы;
- ◆ индикаторный коэффициент полезного действия η_{ind} – показатель экономичности;
- ◆ максимальное давление рабочего тела p_{max} и быстрота нарастания давления в процессе сгорания w_p – показатели механической и динамической нагрузки на детали кривошипно-шатунного механизма;
- ◆ максимальная температура газов T_{max} и температура газов в конце расширения T_b – показатели термической нагрузки;
- ◆ давление p_a и температура T_a газов на момент окончания впуска – показатели состояния рабочего тела в момент впуска;
- ◆ непрерывно изменяющиеся давление $p(\varphi)$ и температура $T(\varphi)$ рабочего тела в течение рабочего цикла.

В качестве входных данных использовались константы, определяющие конструкцию ДВС и характеристики топлива, переменные рабочего процесса, а также значения показателей окружающей среды:

- ♦ конструкция двигателя: объем камеры сгорания V_c , объем цилиндра V_h , степень сжатия ε , радиус кривошипа r , длина шатуна L_{sh} , ход поршня H_p ;
- ♦ переменные рабочего процесса: скорость вращения коленчатого вала rpm , коэффициент наполнения цилиндра η_v , температура подогрева свежего заряда ΔT , давление и коэффициент остаточных газов p_r и γ соответственно, момент зажигания G_{uoz} ;
- ♦ параметры бинарного топлива: фактор теплоемкости K_t , коэффициент избытка воздуха α_{vb} , соотношение бензин/сжиженный газ в бинарном топливе L , низшая теплотворная способность бензина и сжиженного газа H_{ub} и H_{upb} соответственно, стехиометрическое соотношение бензина и сжиженного газа λ_{vb} и λ_{vpb} соответственно, молекулярная масса бензина, сжиженного газа и воздуха m_b , m_{pb} и m_v соответственно;
- ♦ показатели окружающей среды: атмосферное давление p_0 , температура воздуха T_0 , плотность воздуха AD ;

В том случае, где это было возможно, параметры работы двигателя выбирались с возможностью достоверного измерения штатными датчиками современных автомобилей. В противном случае использовались общепризнанные формулы.

Формирование требований к модели. Выводу математической модели двигателя внутреннего сгорания посвящено множество работ, в т.ч.: [11–18, 20], причем количество моделей неуклонно растет, а существующие модели постоянно модернизируются. Это объясняется, с одной стороны, чрезвычайно сложными физико-химическими процессам, происходящими при работе двигателя, а с другой стороны – изменением его конструкции и используемых топлив. Анализ литературы не выявил готовой модели для четырехтактного ДВС с внешним смесеобразованием при частичном сгорании топлива в зоне детонации, что определило необходимость ее разработки.

К моделям рабочих процессов двигателей внутреннего сгорания предъявляются следующие требования:

- 1) модель должна быть формально правильной (корректной с физической точки зрения), т.е. свободной от противоречий;
- 2) модель по возможности должна точно описывать реальный процесс и одновременно быть математически реализуемой;
- 3) требуемые материальные и временные затраты на разработку и реализацию модели должны находиться в разумных пределах;
- 4) уровень модели должен соответствовать требованию: по возможности простая и по необходимости сложная.

Сопоставление различных моделей показывает, что многозонные модели позволяют точнее рассчитать температуру и поэтому показывают лучшие результаты при решении задач, связанных с определением токсичности, теплообмена, диссоциации и т.д. При этом расчет давления в цилиндре во всех классах моделей выполняется по одинаковым схемам и в этом плане многозонные модели не обеспечивают большую точность. Кроме того однозонные модели проще и соответственно удобней для расчета рабочего процесса. Таким образом, для определения давления в цилиндре при работе ДВС в зоне частичной детонации с необходимой точностью и скоростью расчета, на современном этапе развития можно считать достаточной однозонную модель. В дальнейшем при развитии и совершенствовании электронных блоков управления и алгоритмов, возможно использующих не четкое управление скорее всего появится возможность использования многозонной модели с выделением зон детонации, как во времени, так и в пространстве.

Выбор базового способа расчета тепловыделения. Определение способа расчета тепловыделения является одним из важнейших этапов создания математической модели, поскольку от этого зависит точность и адекватность рассчитываемых параметров. В настоящий момент не выявлен известный способа расчета для описываемого случая.

Существует несколько различных подходов к расчету процесса горения в двигателях с воспламенением от электрической искры [17]. В основу первого направления положена аппроксимация линии сгорания индикаторной диаграммы, т.е. линии давления газов в функции изменения объема или угла поворота коленчатого вала, математическим уравнением. Недостаток способов этого направления заключается в том, что поскольку форма линии сгорания определяется не только скоростью и полнотой сгорания, но и закономерностями термодинамики и отчасти теплопередачи, а также значениями углов опережения зажигания, то теоретическое выявление раздельного влияния каждого из этих факторов на линию сгорания и, следовательно, на весь цикл и его показатели в рамках данного подхода является весьма затруднительным.

Второй подход основывается на уравнениях термодинамики и скорости тепловыделения. Раскрыть закономерности протекания тепловыделения в двигателе довольно проблематично, поскольку линия сгорания является сложной кривой, характер которой обусловлен взаимодействием факторов, разных по своей природе. Это и механика для учета характера подвода теплоты и изменения объема газов, и химическая кинетика для рассмотрения процесса превращения химической энергии в тепловую и, в какой-то степени, законы теплопередачи, учитывающие потерю части тепла в стенки. Но этот подход позволяет анализировать изменение скорости сгорания во времени и учитывать влияние на него различных факторов, что необходимо для проведения всестороннего математического анализа.

Для расчета скорости сгорания в рамках второго подхода предложено множество способов [15, 20–22], и интерес к этой теме не утихает с зари автомобилестроения. Ряд исследователей используют для расчета скорости сгорания в бензиновых двигателях скорость фронта пламени. В основу всех этих методов расчета положен доказанный Михельсоном В.А. постулат Гюи, согласно которому количество горючей газовой смеси, воспламеняемой единицей поверхности пламени в единицу времени, постоянно по своей величине. Как отмечает Вибе И.И. эти способы не вызывали бы возражений, если бы соблюдались следующие условия:

- 1) реакция сгорания завершилась, как и в ламинарном потоке, в очень узкой зоне пламени, что позволило бы практически рассматривать поверхность фронта пламени, как реакционную поверхность;
- 2) экспериментально и теоретически можно было бы определить изменение этой поверхности с перемещением фронта пламени;
- 3) можно было бы определить нормальную скорость перемещения фронта пламени, от которой зависит массовая скорость выгорания.

В действительности же, первое условие не соблюдается, и поток газа в двигателе не может быть охарактеризован как ламинарный. Турбулентное горение фронта пламени имеет размытые контуры и значительную толщину. Поэтому понятие поверхности фронта пламени становится неопределенным, а постулат Гюи теряет свою силу.

Известен способ [18], учитывающий влияние турбулентности на скорость сгорания в двигателях с воспламенением от электрической искры. Его авторы принимают гипотезу о том, что сгорание в турбулентном потоке газа происходит так же, как в ламинарном потоке – в тонком слое фронта пламени. При этом крупномасштабная турбулентность деформирует поверхность горения, вызывая много-

численные искривления и увеличивая этим размер фронта пламени, а мелкомасштабная турбулентность, соизмеримая с толщиной фронта пламени, одновременно увеличивает нормальную скорость сгорания в этой зоне. Кроме того принимается допущение, что на единице поверхности фронта пламени сгорает постоянное количество топлива в единицу времени, которое называется массовой скоростью сгорания в турбулентном потоке, причем эта скорость изменяется по ходу горения. Таким образом, общее количество сгоревшего в единицу времени топлива пропорционально произведению общей поверхности фронта пламени на массовую скорость сгорания. Недостатком этого способа являются трудности формализации основных понятий метода «фронт пламени» и «скорость распространения фронта», которые оставлены авторами для дальнейших исследований, кроме того не представляется возможным учесть количество бензина окисленного с помощью детонации.

Вибе И.И. предложил способ расчета скорости сгорания в двигателях с воспламенением от электрической искры [20] без использования понятия линейной скорости пламени, а основанный на том, что горение топлива в двигателе является одним из основных процессов, определяющих качество его работы, в т.ч. мощность, удельный расход топлива, долговечность и экологические показатели.

Закономерности развития процессов сгорания во времени представлены Вибе полуэмпирическим уравнением общего характера, которое может быть использовано для широкого спектра условий, т.к. при его выводе не делалось никаких предположений относительно специальных средств, регламентирующих систему (природа исходных веществ и их физическое состояние, конкретный механизм реакции, температура, давление, вихревые движения газов, теплопередача, зарождение эффективных центров в разных точках объема и т.д.)

Уравнения Вибе позволяют описать физическое развитие процесса сгорания во времени, как в бензиновых двигателях, так и в газовых, и в дизельных, что позволяет предположить возможность использования такого подхода при решении задачи горения топливовоздушной смеси в разных режимах. По всей видимости оно позволит не только построить диаграмму цикла и его показатели, но и провести исследования по выявлению зависимостей показателя цикла от характеризующих их параметров. Формула Вибе содержит два определяющих параметра процесса сгорания m и φ_z , от правильного подбора которых зависит степень соответствия экспериментальных и расчетных параметров рабочего процесса. Для подбора параметров формулы Вибе может использоваться как статистический материал по тепловыделению в двигателях, подобных изучаемому, так и особенности физико-химических процессов в цилиндре.

Процедура расчета при использовании детонационного сгорания части топлива, в рассматриваемой ситуации примет вид уравнение сгорания топлива [8]. (при условии выгорания 99,9 %):

$$x = 1 - e^{-6.908\left(\frac{t_1}{t_{21}}\right)^{m_1+1}} + 1 - e^{-6.908\left(\frac{t_1}{t_{22}}\right)^{m_2+1}}, \quad (1)$$

где индексом «1» обозначены параметры, относящиеся к обычному горению, а индексом «2» детонационному.

При разработке процедуры расчета параметров рабочего цикла двигателя были приняты следующие допущения:

- ◆ рассматривается замкнутый цикл с реальным рабочим телом постоянной массы, имеющей переменный состав и теплоемкость, не зависящую от температуры;
- ◆ цикл состоит из политропического сжатия, сгорания, расширения продуктов сгорания с учетом выделения тепла вследствие рекомбинации диссоциированных молекул и отвода тепла при постоянном объеме;

- ♦ сжатие происходит по политропе с показателем, принятым на основании экспериментальных данных;
- ♦ рабочее тело – совершенный газ (идеальный в термодинамическом смысле), который в разные моменты времени состоит из двух компонентов: свежего заряда и продуктов сгорания, разделенных на соответствующие зоны;
- ♦ реальный процесс горения топлива заменен эквивалентным подводом тепловой энергии по закону Вибе И.И;
- ♦ давление во всем объеме камеры сгорания принимается одинаковым;
- ♦ температура стенок цилиндра, каналов систем пуска и выпуска, а также глушителя, принимается постоянной.

Расчет выходных данных модели расчета УОЗ в зависимости от состава ТВС, угла опережения зажигания и параметров окружающей среды состоит из следующих этапов [23]:

- 1) с учетом параметров окружающей среды и параметров двигателя производится расчет показателей состояния рабочего тела в момент впуска и на этапе сжатия;
- 2) на основании данных этапа сжатия рассчитывается давление в процессе горения топлива в цилиндре, температура рабочего тела в цилиндре, а также определяются показатели механической и динамической нагрузки;
- 3) после расчета горения проводится расчет давления и температуры в цилиндре при расширении;
- 4) по полученной индикаторной диаграмме выполняется определение показателей удельной работы, экономичности, рассчитываются остальные выходные параметры.

Расчет давления в цилиндре на этапе горения. Расчет параметров рабочего тела в такте сжатия можно произвести в соответствии с предложенным авторами подходом [23], учитывая тот факт, что наполнение цилиндров в такте впуска не зависят от возможности работы двигателя частичной детонацией. Кроме того использование такого же подхода для моделирования горения бинарного топлива и гашения возникающей детонации [7] позволит распространить подобное математическое моделирование для процессов управления детонацией в случае использования пропанобутановой смеси. Расчет этапа горения начинается с точки y (рис. 1), соответствующей подаче искры и поджогу топливно-воздушной смеси и оканчивается в точке z – точке окончания горения.

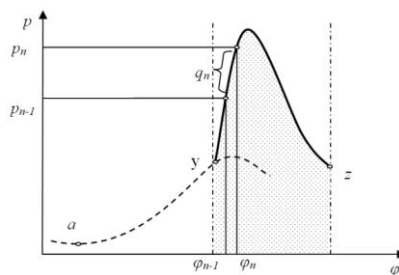


Рис. 1. Этап горения

Для расчета индикаторной диаграммы после поджога топливно-воздушной смеси необходимо провести моделирование динамики окисления топлива. Для этого весь процесс сгорания разобьем на отдельные небольшие интервалы, выражаемые в градусах поворота коленчатого вала. Интервал задается границами $[\varphi_{n-1}, \varphi_n]$. Удельная теплота сгорания q_n на участке n определяется:

$$q_n = q_z(x_n - x_{n-1}),$$

где q_z – полная удельная теплота сгорания топлива; x_{n-1} – доля топлива, сгоревшая к началу участка n ; x_n – доля топлива, сгоревшая к концу участка n .

С целью оптимизации угла опережения зажигания предполагается использовать добавление сжиженного газа. Для расчета полной удельной теплоты сгорания q_z при работе на бинарном топливе предлагается следующее выражение:

$$q_z = L_b q_{zb} + L_{pb} q_{zpb},$$

где L_b – доля бензина в бинарном топливе; q_{zb} – полная удельная теплота сгорания бензина; L_{pb} – доля сжиженного газа в бинарном топливе; q_{zpb} – полная удельная теплота сгорания сжиженного газа.

Полная удельная теплота сгорания бензина q_{zb} определяется по формуле

$$q_{zb} = H_{smb} \times \frac{g\xi}{(1 + \gamma)(1 + \alpha_{vt}\lambda_{vb})},$$

где H_{smb} – объемная теплотворность горючей смеси; α_{vt} – коэффициент избытка воздуха; λ_{vb} – соотношение воздух/бензин; ξ – коэффициент эффективности сгорания.

В двигателях с внешним смесеобразованием при $\alpha \geq 1$ коэффициент ξ находится в пределах 0,9-0,95 [21]. Величина H_{smb} для бензо-воздушной смеси вычисляется по формуле

$$H_{smb} = \frac{H_{ub}}{\alpha_{vt} \frac{\lambda_{vb}}{m_v} + \frac{1}{m_b}},$$

где H_{ub} – низшая теплотворность бензина; m_v – молярная масса воздуха; m_b – молярная масса бензина.

Доля топлива, сгоревшего за время поворота коленчатого вала от φ_{n-1} до φ_n , без возникновения детонации, рассчитывается по полуэмпирическому уравнению выгорания топлива в двигателях внутреннего сгорания, полученному И.И. Вибе [20]:

$$\Delta x_n = x_n - x_{n-1} = e^{-6,908 \times \left(\frac{\varphi_n}{\varphi_z}\right)^{m+1}} - e^{-6,908 \times \left(\frac{\varphi_{n-1}}{\varphi_z}\right)^{m+1}}, \quad (2)$$

где m – показатель характера сгорания; φ_z – продолжительность сгорания топлива, выраженная в градусах поворота коленчатого вала.

В случае сгорания части топлива в режиме детонации расчет определяется формулой (1)

Род применяемого топлива существенно влияет на развитие начального очага горения, в то время как в основной фазе скорость распространения пламени подчиняется законом крупномасштабного турбулентного горения, мало зависящим от физико-химических свойств топливно-воздушной смеси [23, 24]. Изменение показателя сгорания учитывается с помощью соотношения, успешно опробованного в работах [25]:

$$m = m_0 - 0,1(Uoz - Uoz_0),$$

где m_0 – показатель сгорания, соответствующий начальным условиям, полученный на опорном режиме работы ДВС при $rpm=2000$ об/мин, угле зажигания $Uoz_0=25$ °пКВ, коэффициенте избытка воздуха 0,96.

В то же время совершенно очевидно, что в случае частичного детонационного горения показатель m_0 будет несколько выше и зависимость его увеличения будет соответствовать доли индикаторных диаграмм второго типа [9].

Выводы. Проведенный обзор работ не выявил существующей модели для расчета УОЗ при использовании частичного детонационного сгорания топлива, что определило необходимость ее разработки.

Исходя из совокупности требований к модели и набора выходных данных в качестве базового способа расчета тепловыделения был принят способ, предложенный Вибе И.И.

Предложенная процедура расчета параметров рабочего цикла ДВС, в случае продолжения вышеописанных работ, позволит проводить вычисления в реальном времени и учесть влияние состава бинарного топлива на показатели мощности, экономичности, механической и динамической нагрузки на детали кривошипно-шатунного механизма, а также термической нагрузки на двигатель.

В дальнейшем для полного описания процессов горения необходимо определить влияние на продолжительность горения и коэффициент молекулярного изменения.

Ранее авторами выявлено, что при использовании детонационного горения части топлива в диапазоне оборотов коленчатого вала от 1800 об/мин. до 2800 об/мин., происходит 116 рабочих тактов, из них 64 такта относятся к частичной детонации, что позволяет ДВС увеличить крутящий момент на 10 %. Полученные данные будут использованы для идентификации разрабатываемой модели и увеличения тактов с частичной допустимой детонацией, что позволит значительно увеличить крутящий момент и мощность на КД ДВС.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Зельдович Я.Б., Компанец А.С. Теория детонации. – М.: Гос. изд-во техн.-теорет. лит., 1955. – 268 с.
2. Wang Z., Liu H., Reitz R.D. Knocking combustion in spark-ignition engines // Progress in Energy and Combustion Science. – 2017. – Vol. 61. – P. 78-112.
3. Gowthaman S., Sathiyagnanam A.P. Performance and emission characteristics of homogeneous charge compression ignition engine—a review // International Journal of Ambient Energy. – 2017. – Vol. 38, No. 7. – P. 672-684.
4. Митрофанов В.В. Теория детонации. – Новосибирск: НГУ, 1982. – 92 с.
5. Caton J.A. The interactions between IC engine thermodynamics and knock // Energy Conversion and Management. – 2017. – Vol. 143. – P. 162-172.
6. Береснев А.Л., Береснев М.А. Увеличение эффективности подвижных объектов с использованием детонации в ходовом двигателе внутреннего сгорания // Перспективные системы и задачи управления: Матер. XIII всероссийской научно-практической конференции. – Владивосток, 2018. – С. 131-138.
7. Береснев М.А., Береснев А.Л. Управление составом бинарного топлива для улучшения показателей ДВС // АвтоГазоЗаправочный комплекс + Альтернативное топливо. – 2012. – № 3 (63). – С. 7-10.
8. Береснев А.Л., Береснев М.А., Быстрицкий А.В. Возможности управления детонационным горением различных типов в ДВС // Известия ЮФУ. Технические науки. – 2019. – № 3 (205). – С. 198-207.
9. Береснев А.Л., Береснев М.А., Быстрицкий А.В. Исследование возможности управления ДВС во время детонационного горения части топливо-воздушной смеси // Известия ЮФУ. Технические науки. – 2018. – № 6 (200). – С. 164-174.
10. Береснев М.А., Береснев А.Л. Особенности контроля детонации при использовании бинарного топлива // Сб. научных трудов SWorld. – 2011. – Т. 1, № 2. – С. 56-58.
11. Орлин А.С. Теория поршневых и комбинированных двигателей / под ред. А.С. Орлина и М.Г. Круглова. – М.: Машиностроение, 1983. – 372 с.
12. Льюис Б., Эльбе Г. Горение, пламя и взрывы в газах. – М.: Мир, 1968. – 592 с.
13. Глаголев Н.М. Рабочие процессы двигателей внутреннего сгорания. Новый метод расчета. – М.: Изд-во машиностр. лит., 1950. – 481 с.
14. Гирявец А.К. Теория управления автомобильным бензиновым двигателем. – М.: Стройиздат, 1997. – 161 с.
15. Кавтарадзе Р.З. Теория поршневых двигателей. Специальные главы. – М.: Изд-во МГТУ им Н.Э. Баумана, 2008. – 720 с.
16. Иващенко Н.А., Кавтарадзе Р.З. Многозонные модели рабочего процесса двигателей внутреннего сгорания. – М.: Изд-во МГТУ им. Н.Э. Баумана, 1997. – 57 с.

17. Варнатц Ю., Маас У., Диббл Р. Основы горения. Физические и химические аспекты, моделирование, эксперименты. Образование загрязняющих веществ: пер. с англ. Г.Л. Агафонова / под. ред. П.А. Власов. – М.: Физматлит, 2003. – 352 с.
18. Иноземцев Н.В., Кошкин В.К. Процессы сгорания в двигателях. – М.: Машгиз, 1949. – 344 с.
19. Глаголев Н.М. Рабочие процессы двигателей внутреннего сгорания. – М.: Машгиз, 1950. – 479 с.
20. Вибе И.И. Новое о рабочем цикле двигателя. – М.: Машгиз, 1962. – 271 с.
21. Шароглазов Б.А., Фарафонов М.Ф., Клементьев В.В. Двигатели внутреннего сгорания: теория, моделирование и расчет процессов. – Челябинск: Изд-во ЮУрГУ, 2004. – 344 с.
22. Елагин М.Ю., Кузьмина И.В. Математическая модель и теоретические исследования рабочих процессов многоцилиндровых двигателей внутреннего сгорания // Известия ТулГУ. Серия «Автомобильный транспорт». Вып. 3. – Тула: ТулГУ, 1999. – С. 104-107.
23. Beresnev M.A., Beresnev A.L. Mathematical model of gasoline-liquified gas mixture combustion for IC engine control // Modern Fundamental and Applied Researches. – 2011. – No. 3. – P. 58-61.
24. Кузьмин А.В. Показатели и регулировки битопливного двигателя при переводе его с бензина на сжиженный углеводородный газ: дисс. ... канд. техн. наук. – Волгоград: ВолгГТУ, 2008. – 116 с.
25. Федянова Н.А. Использование математического моделирования рабочего процесса для разработки и обоснования концепции эффективного управления углом опережения зажигания в ДВС: дисс. ... канд. техн. наук. – Волгоград: ВолгПИ, 1992. – 139 с.

REFERENCES

1. Zel'dovich Ya.B., Kompaneets A.S. Teoriya detonatsii [Theory of detonation]. Moscow: Gos. izd-vo tekhn.-teoret. lit., 1955, 268 p.
2. Wang Z., Liu H., Reitz R.D. Knocking combustion in spark-ignition engines, *Progress in Energy and Combustion Science*, 2017, Vol. 61, pp. 78-112.
3. Gowthaman S., Sathiyagnanam A.P. Performance and emission characteristics of homogeneous charge compression ignition engine—a review, *International Journal of Ambient Energy*, 2017, Vol. 38, No. 7, pp. 672-684.
4. Mitrofanov V.V. Teoriya detonatsii [Theory of detonation]. Novosibirsk: NGU, 1982, 92 p.
5. Caton J.A. The interactions between IC engine thermodynamics and knock, *Energy Conversion and Management*, 2017, Vol. 143, pp. 162-172.
6. Beresnev A.L., Beresnev M.A. Uvelichenie effektivnosti podvizhnykh ob'ektov s ispol'zovaniem detonatsii v khodovom dvigatele vnutrennego sgoraniya [Increasing the efficiency of mobile objects using detonation in a running internal combustion engine], *Perspektivnye sistemy i zadachi upravleniya: Mater. XIII vserossiyskoy nauchno-prakticheskoy konferentsii* [Perspective systems and management tasks: Materials of the XIII all-Russian scientific and practical conference]. Vladivostok, 2018, pp. 131-138.
7. Beresnev M.A., Beresnev A.L. Upravlenie sostavom binarnogo topliva dlya uluchsheniya pokazateley DVS [Managing the composition of binary fuel to improve the performance of internal combustion engines], *AvtoGazoZapravochnyy kompleks + Al'ternativnoe toplivo* [Gas station complex + Alternative fuel], 2012, No. 3 (63), pp. 7-10.
8. Beresnev A.L., Beresnev M.A., Bystritskiy A.V. Vozmozhnosti upravleniya detonatsionnym goreniem razlichnykh tipov v DVS [Features of controlling detonation Gorenje of various types in the internal combustion engine], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences]. – 2019. – № 3 (205). – С. 198-207.
9. Beresnev A.L., Beresnev M.A., Bystritskiy A.V. Issledovanie vozmozhnosti upravleniya DVS vo vremya detonatsionnogo goreniya chasti toplivo-vozdushnoy smesi [Investigation of the possibility of controlling the internal combustion engine during the detonation Gorenje part of the fuel-air mixture], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences]. 2018, No. 6 (200), pp. 164-174.
10. Beresnev M.A., Beresnev A.L. Osobennosti kontrolya detonatsii pri ispol'zovanii binarnogo topliva [Features of detonation control when using binary fuel], *Sb. nauchnykh trudov SWorld* [Collection of scientific papers SWorld], 2011, Vol. 1, No. 2, pp. 56-58.
11. Orlin A.S. Teoriya porshnevnykh i kombinirovannykh dvigateley [Theory of piston and combined engines], ed. by A.S. Orlina i M.G. Kruglova. Moscow: Mashinostroenie, 1983, 372 p.

12. L'yuis B., El'be G. Gorenje, plama i vzryvy v gazakh [Combustion, flames and explosions of gases]. Moscow: Mir, 1968, 592 p.
13. Glagolev N.M. Rabochie protsessy dvigateley vnutrennego sgoraniya. Novyy metod rascheta [Working processes of internal combustion engines. A new method of calculation]. Moscow: Izd-vo mashinostr. lit., 1950, 481 p.
14. Giryavets A.K. Teoriya upravleniya avtomobil'nyim benzinovym dvigatелеm [Theory of automobile gasoline engine control]. Moscow: Stroyizdat, 1997, 161 p.
15. Kavtaradze R.Z. Teoriya porshnevnykh dvigateley. Spetsial'nye glavy [Theory of piston engines. Special chapters]. Moscow: Izd-vo MGTU im N.E. Bauman, 2008, 720 p.
16. Ivashchenko N.A., Kavtaradze R.Z. Mnogozonnye modeli rabocheho protsessa dvigateley vnutrennego sgoraniya [Multi-zone models of the internal combustion engine workflow]. Moscow: Izd-vo MGTU im. N.E. Bauman, 1997, 57 p.
17. Varnatts Yu., Maas U., Dibbl R. Osnovy gorenija. Fizicheskie i khimicheskie aspekty, modelirovanie, eksperimenty. Obrazovanie zagryznyayushchikh veshchestv [Basics of Gorenje. Physical and chemical aspects, modeling, experiments], transl. from engl. G.L. Agafonova, ed. by P.A. Vlasov. Moscow: Fizmatlit, 2003, 352 p.
18. Inozemtsev N.V., Koshkin V.K. Protsessy sgoraniya v dvigatelyakh [Combustion processes in engines]. Moscow: Mashgiz, 1949, 344 p.
19. Glagolev N.M. Rabochie protsessy dvigateley vnutrennego sgoraniya [Working processes of internal combustion engines]. Moscow: Mashgiz, 1950, 479 p.
20. Vibe I.I. Novoe o rabochem tsikle dvigatelya [New about the working cycle of the engine]. Moscow: Mashgiz, 1962, 271 p.
21. Sharoglazov B.A., Farafontov M.F., Klemen'tev V.V. Dvigateli vnutrennego sgoraniya: teoriya, modelirovanie i raschet protsessov [Internal combustion engines: theory, modeling and calculation of processes]. Chelyabinsk: Izd-vo YuUrGU, 2004, 344 p.
22. Elagin M.Yu., Kuz'mina I.V. Matematicheskaya model' i teoreticheskie issledovaniya rabochikh protsessov mnogotsilindrovnykh dvigateley vnutrennego sgoraniya [Mathematical model and theoretical studies of working processes of multi-cylinder internal combustion engines], *Izvestiya TulGU. Seriya «Avtomobil'nyy transport»* [Izvestiya of the Tula state University. Series "Road transport"], Issue 3. Tula: TulGU, 1999, pp. 104-107.
23. Beresnev M.A., Beresnev A.L. Mathematical model of gasoline-liquified gas mixture combustion for IC engine control, *Modern Fundamental and Applied Researches*, 2011, No. 3, pp. 58-61.
24. Kuz'min A.V. Pokazateli i regulirovki bitoplivnogo dvigatelya pri perevode ego s benzina na szhizhenny uglevodorodnyy gaz: diss. ... kand. tekhn. nauk [Indicators and adjustments of the bi-fuel engine when converting it from gasoline to liquefied petroleum gas: cand. of eng. sc. diss.]. Volgograd: VolGTU, 2008, 116 p.
25. Fedyanova N.A. Ispol'zovanie matematicheskogo modelirovaniya rabocheho protsessa dlya razrabotki i obosnovaniya kontseptsii effektivnogo upravleniya uglom operezheniya zazhiganiya v DVS: diss. ... kand. tekhn. nauk [Using mathematical modeling of the workflow to develop and justify the concept of effective control of the ignition advance angle in the internal combustion engine: cand. of eng. sc. diss.]. Volgograd: VolgPI, 1992, 139 p.

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

Береснев Алексей Леонидович – Южный федеральный университет; e-mail: beresneval@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: +78634371694; кафедра электротехники и мехатроники; к.т.н.; доцент.

Береснев Максим Алексеевич – ООО ИК «Медотрейд»; e-mail: beresnev@medotrade.ru; 347922, г. Таганрог, ул. Чехова, 104-А; к.т.н.; инженер-программист.

Beresnev Aleksey Leonidovich – Southern Federal University; e-mail: beresneval@sfedu.ru; 44, Nekrasovsky, Taganrog, 347928, Russia; phone: +78634371694; the department of electrotechnics and mechatronics; cand. of eng. sc.; associate professor.

Beresnev Maksim Alekseevich – Medotrade Ltd; e-mail: beresnev@medotrade.ru; 104-A Chekhov street, Taganrog, 347922, Russia; cand. of eng. sc.; engineer-programmer.

В.А. Павский, К.В. Павский**МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДЛЯ РАСЧЕТА ПОКАЗАТЕЛЕЙ
НАДЕЖНОСТИ МАСШТАБИРУЕМЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ
С УЧЕТОМ ВРЕМЕНИ ПЕРЕКЛЮЧЕНИЯ***

Основной особенностью масштабируемых вычислительных систем является модульность. Нарращивание производительности в таких системах достигается за счет увеличения однотипных элементов, элементарных машин (например, вычислительный узел). В результате отказов, производительность системы изменяется. Таким образом, масштабируемость вычислительных систем (ВС), с одной стороны, увеличивает производительность, но с другой стороны, наращивание ресурса обостряет проблему надежности и увеличивает сложность организации эффективного функционирования. Анализ надежности и потенциальных возможностей вычислительных систем по-прежнему остается актуальной задачей. Для количественного анализа функционирования масштабируемых вычислительных систем используют показатели живучести, связанные с надежностью. Например, показатели потенциальной живучести ВС учитывают то обстоятельство, что при решении задач используются все исправные элементарные машины, количество которых изменяется во времени в результате отказов и восстановлений. При анализе надежности популярными в теории вычислительных систем, являются модели, основанные на теории марковских процессов и теории массового обслуживания (ТМО). Большинство аналитических моделей ТМО не учитывают время переключения (реконфигурации) в отдельном параметре, ввиду сложности решения. Обычно ограничиваются тем, что время восстановления и переключения объединяют в один параметр. В работе, на примере одной модели ТМО, получены аналитические решения системы дифференциальных уравнений с тремя параметрами (отказ, восстановление и переключение) для расчета показателей надежности и потенциальной живучести. Тем самым предоставляется возможность пользователю самому определить, стоит ли учитывать времена переключения. Показано, что решения трехпараметрической модели сводятся к решениям двух параметрической модели, если время переключения не берется в рассмотрение.

Вычислительные системы; масштабируемость; отказы; время переключения; математическая модель; анализ; показатели надежности и живучести; аналитические решения.

V.A. Pavsky, K.V. Pavsky**MATHEMATICAL MODEL FOR CALCULATING RELIABILITY INDICES
OF SCALABLE COMPUTER SYSTEMS WITH SWITCHING TIME**

The main feature of scalable computer systems is modularity. Increasing performance in such systems is achieved by increasing the same type of elements, elementary machines (EM, for example, a computing node). As a result of failures, the system performance is changed. Thus, scalability of computer systems (CS), on the one hand, increases performance, but on the other hand, computer resource growth exacerbates the problem of reliability and increases the complexity of organizing effective functioning. Analysis of reliability and potential capabilities of computing systems is still an urgent problem. For quantitative analysis of the functioning of scalable computing systems, robustness indices related to reliability are used. For example, indices of potential robustness of CS take into account the fact that all operable elementary machines are used in solving tasks, the number of which (EM) changes over time as a result of failures and recoveries. When analyzing reliability, models based on the theory of Markov processes and Queuing theory (QT) are popular in the theory of computing systems. Most QT analytical models do not consider the switching time (reconfiguration) in a separate parameter, due to the complexity of the

* Работа выполнена при поддержке, РФФИ (грант № 20-07-00039), программы фундаментальных исследований СО РАН (ГЗ 0306-2019-0019).

solution. Usually, models are simplified by the fact that the recovery time and switch combined in a single parameter. Analytical solutions of a system of differential equations with three parameters (failure, recovery, and switching) for calculating reliability and potential robustness are obtained on the example of the QT model. This allows the user to determine whether the switching time should be taken into account. Also it is shown that solutions of the three-parameter model are reduced to solutions of the two-parameter model if the switching time is not taken into consideration.

Computer systems; scalability; failures; switching time; mathematical model; analysis; reliability and robustness indices; analytical solutions.

Введение. Масштабируемость вычислительных систем (ВС), с одной стороны, обеспечивает высокую доступность сервиса [1, 2], но с другой стороны, наращивание вычислительных ресурсов обостряет проблему надежности и увеличивает сложность организации эффективного функционирования [3, 4]. Анализ надежности и потенциальных возможностей ВС по-прежнему остается актуальной задачей.

Для количественного анализа функционирования масштабируемых вычислительных систем (ВС) используют показатели живучести, связанные с надежностью [1]. Например, показатели потенциальной живучести вычислительных систем (ВС) учитывают то обстоятельство, что при решении задач используются все исправные элементарные машины, количество которых изменяется во времени в результате отказов и восстановлений. При определении показателей живучести следует учитывать, что параллельные программы сложных задач при их реализации на живучих ВС способны задействовать суммарную производительность всех работоспособных элементарных машин (ЭМ, узлов) [1], число которых не является фиксированным.

Наиболее распространенными в теории вычислительных систем, при анализе надежности и функционировании, являются модели, основанные на теории марковских процессов [5–11] и теории массового обслуживания (ТМО) [12–17]. Большинство аналитических моделей ТМО не учитывает время переключения (реконфигурации) как отдельный параметр, ввиду сложности решения. Обычно ограничиваются тем, что время восстановления и переключения объединяют в один параметр [17]. Для расчета показателей надежности и потенциальной живучести, разработана модель системы массового обслуживания, описываемая системой дифференциальных уравнений с тремя параметрами (отказ, восстановление и переключение). Приводятся аналитические решения для расчета состояний системы, среднего числа машин в отказе и переключении. Итак, рассматриваем ВС как стохастический объект. При расчете показателей надежности и потенциальной живучести за основу берётся стохастическая модель функционирования ВС [1], представленная на рис. 1.

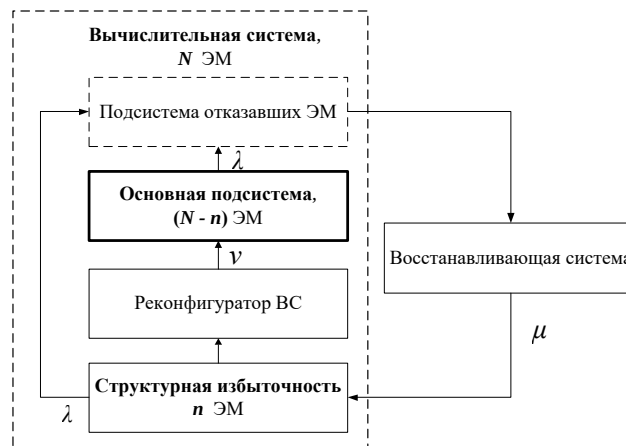


Рис. 1. Модель функционирования ВС

Основная подсистема, непосредственно занимается решением задач, состоит из исправных машин. Каждая исправная машина отказывает с интенсивностью λ , после чего переходит в восстанавливающую систему. В восстанавливающей системе элементарная восстанавливается с интенсивностью μ . При наличии структурной избыточности, её исправные машины позволяют поддерживать заданную производительность системы. В системе реконфигурации происходит настройка ЭМ для работы и включение её в состав основной системы с определенной интенсивностью.

1. Математическая модель. Пусть вычислительная система состоит из N ЭМ. Интенсивность выхода из строя любой ЭМ – λ (параметр входящего потока), а восстановления – μ (параметр исходящего потока; рассматриваем простейшие потоки). Как только машина восстановилась, она сразу же включается в систему. В рамках ТМО, описанная модель соответствует процессу рождения и гибели [13, 14], граф схема которого представлена на рис. 2, где C_k – состояние системы, т.е. имеем k заявок (в случае ВС, можно, например, понимать число отказавших или исправных машин) на обслуживании.

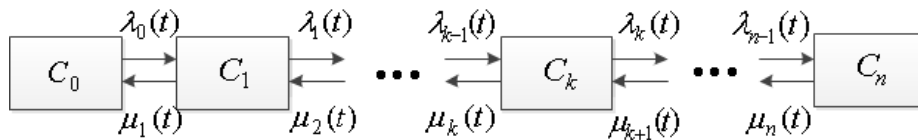


Рис. 2. Граф-схема процесса рождения и гибели

Для двух параметрической модели восстановление каждой машины можно отразить граф-схемой, представленной на рис. 3, где C_0 – соответствует исправному состоянию ЭМ, а C_1 – состояние в отказе.

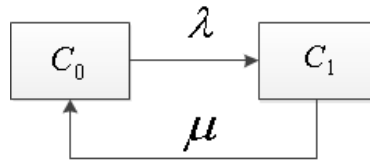


Рис. 3. Граф-схема состояний элементарной машины для двух-параметрической модели

Пусть $R(t)$ – математическое ожидание числа работающих ЭМ в момент времени t , $t \in [0, \infty)$, $M(t)$ – математическое ожидание числа отказавших ЭМ в момент времени t . В работе [1] приводятся решения для расчета показателей потенциальной живучести в случае двух параметрической модели (в соответствии с рис. 1, 2) для переходного и стационарного режимов работы ВС. Например для восстанавливающей системы высокой производительности в стационарном режиме

$$R = \frac{N\mu}{\lambda + \mu}, \quad M = \frac{N\lambda}{\lambda + \mu}, \quad (1)$$

где R – математическое ожидание числа рабочих ЭМ; M – математическое ожидание числа отказавших ЭМ.

Рассмотрим трехпараметрическую модель, где восстановленная после отказа ЭМ не включается сразу в систему, а происходит её настройка, т.е. переключение идет с определенной интенсивностью. Таким образом, каждая ЭМ в любой момент

времени t , $t \in [0, \infty)$, может находиться в одном из трех несовместных состояний C_i , $i = 1, 2, 3$: рабочем (C_0), отказа (C_1) и переключения (C_0^*).

На рис. 4 представлена граф-схема процесса восстановления ЭМ до рабочего состояния, где ν – интенсивность переключения каждой ЭМ.

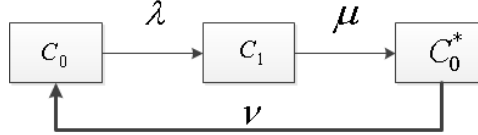


Рис. 4. Граф-схема состояний элементарной машины для трех-параметрической модели

Приведем аналитические решения для распределения вероятностей в переходном режиме, являющегося базовым для расчета набора показателей эффективности (надежности и потенциальной живучести) ВС [1] в случае трехпараметрической модели.

2. Показатели потенциальной живучести ВС. Пусть $r(t)$ – вероятность того, что в момент времени t ЭМ находится в состоянии отказа,

$l(t)$ – вероятность того, что в момент времени t , ЭМ находится в состоянии переключения,

$\varphi(t)$ – вероятность того, что в момент времени t ЭМ находится в рабочем состоянии, $t \in [0, \infty)$.

Тогда для любой ЭМ и $t \in [0, \infty)$, по заданному графу состояний (рис. 3), составляем следующую систему уравнений

$$\begin{cases} \frac{d}{dt} r(t) = -\mu \cdot r(t) + \lambda \cdot \varphi(t), \\ \frac{d}{dt} \varphi(t) = -\lambda \cdot \varphi(t) + \nu \cdot l(t), \\ \frac{d}{dt} l(t) = -\nu \cdot l(t) + \mu \cdot r(t), \\ 1 = r(t) + \varphi(t) + l(t). \end{cases} \quad (2)$$

Для решения системы (2), достаточно взять любые три уравнения. Общее решение системы (2) имеет вид

$$\begin{cases} r(t) = \frac{\lambda \nu}{\alpha_1 \alpha_2} + C_1 \cdot e^{\alpha_1 t} + C_2 \cdot e^{\alpha_2 t}, \\ \varphi(t) = \frac{\mu \nu}{\alpha_1 \alpha_2} + \frac{(\mu + \alpha_1) C_1}{\lambda} e^{\alpha_1 t} + \frac{(\mu + \alpha_2) C_2}{\lambda} e^{\alpha_2 t}, \\ l(t) = \frac{\lambda \mu}{\alpha_1 \alpha_2} - \frac{(\lambda + \mu + \alpha_1) C_1}{\lambda} e^{\alpha_1 t} - \frac{\nu(\lambda + \mu + \alpha_2) C_2}{\lambda} e^{\alpha_2 t}, \end{cases} \quad (3)$$

$$\text{где } \alpha_{1,2} = -\frac{1}{2} \left(\lambda + \mu + \nu \pm \sqrt{\lambda^2 + \mu^2 + \nu^2 - 2(\lambda\mu + \mu\nu + \lambda\nu)} \right).$$

Для получения частных решений (3), рассмотрим возможные начальные условия

- 1) $r(0) = 0, \varphi(0) = 1, l(0) = 0$;
- 2) $r(0) = 1, \varphi(0) = 0, l(0) = 0$;
- 3) $r(0) = 0, \varphi(0) = 0, l(0) = 1$.

Пусть $r_i(t) = r(t)$ при начальном условии $i, i = 1, 2, 3$, тогда частные решения имеют вид:

а) для п. 1)

$$\begin{cases} r_1(t) = \frac{\lambda v}{\alpha_1 \alpha_2} - \frac{\lambda(v + \alpha_1)}{\alpha_1(\alpha_2 - \alpha_1)} e^{\alpha_1 t} + \frac{\lambda(v + \alpha_2)}{\alpha_2(\alpha_2 - \alpha_1)} e^{\alpha_2 t}, \\ \varphi_1(t) = \frac{\mu v}{\alpha_1 \alpha_2} - \frac{(v + \alpha_1)(\mu + \alpha_1)}{\alpha_1(\alpha_2 - \alpha_1)} e^{\alpha_1 t} + \frac{(v + \alpha_2)(\mu + \alpha_2)}{\alpha_2(\alpha_2 - \alpha_1)} e^{\alpha_2 t}, \\ l_1(t) = \frac{\lambda \mu}{\alpha_1 \alpha_2} - \frac{\lambda \mu}{\alpha_1(\alpha_2 - \alpha_1)} e^{\alpha_1 t} - \frac{\lambda \mu}{\alpha_2(\alpha_2 - \alpha_1)} e^{\alpha_2 t}, \end{cases} \quad (4)$$

б) для п. 2)

$$\begin{cases} r_2(t) = \frac{\lambda v}{\alpha_1 \alpha_2} + \frac{\mu(\lambda + v + \alpha_1)}{\alpha_1(\alpha_2 - \alpha_1)} e^{\alpha_1 t} - \frac{\mu(\lambda + v + \alpha_2)}{\alpha_2(\alpha_2 - \alpha_1)} e^{\alpha_2 t}, \\ \varphi_2(t) = \frac{\mu v}{\alpha_1 \alpha_2} - \frac{\mu v}{\alpha_1(\alpha_2 - \alpha_1)} e^{\alpha_1 t} + \frac{\mu v}{\alpha_2(\alpha_2 - \alpha_1)} e^{\alpha_2 t}, \\ l_2(t) = \frac{\mu \lambda}{\alpha_1 \alpha_2} - \frac{\mu(\alpha_1 + \lambda)}{\alpha_1(\alpha_2 - \alpha_1)} e^{\alpha_1 t} + \frac{\mu(\alpha_2 + \lambda)}{\alpha_2(\alpha_2 - \alpha_1)} e^{\alpha_2 t}, \end{cases} \quad (5)$$

в) для п. 3)

$$\begin{cases} r_3(t) = \frac{\lambda v}{\alpha_1 \alpha_2} - \frac{\lambda v}{\alpha_1(\alpha_2 - \alpha_1)} e^{\alpha_1 t} + \frac{\lambda v}{\alpha_2(\alpha_2 - \alpha_1)} e^{\alpha_2 t}, \\ \varphi_3(t) = \frac{\mu v}{\alpha_1 \alpha_2} - \frac{v(\mu + \alpha_1)}{\alpha_1(\alpha_2 - \alpha_1)} e^{\alpha_1 t} + \frac{v(\mu + \alpha_2)}{\alpha_2(\alpha_2 - \alpha_1)} e^{\alpha_2 t}, \\ l_3(t) = \frac{\mu \lambda}{\alpha_1 \alpha_2} + \frac{v(\lambda + \mu + \alpha_1)}{\alpha_1(\alpha_2 - \alpha_1)} e^{\alpha_1 t} - \frac{v(\lambda + \mu + \alpha_2)}{\alpha_2(\alpha_2 - \alpha_1)} e^{\alpha_2 t}. \end{cases} \quad (6)$$

Рассмотрим случай, когда переключение осуществляется мгновенно, т.е. $v \rightarrow \infty$. Тогда

$$\begin{aligned} \lim_{v \rightarrow \infty} \alpha_i &= -\frac{1}{2} \lim_{v \rightarrow \infty} \left(\lambda + \mu + v - \sqrt{\lambda^2 + \mu^2 + v^2 - 2(\lambda\mu + \mu v + \lambda v)} \right) = [\infty - \infty] = \\ &= -\frac{1}{2} \lim_{v \rightarrow \infty} \frac{(\lambda + \mu + v)^2 - \left(\sqrt{\lambda^2 + \mu^2 + v^2 - 2(\lambda\mu + \mu v + \lambda v)} \right)^2}{\lambda + \mu + v + \sqrt{\lambda^2 + \mu^2 + v^2 - 2(\lambda\mu + \mu v + \lambda v)}} = \\ &= -\frac{1}{2} \lim_{v \rightarrow \infty} \frac{4(\lambda\mu + \mu v + \lambda v)/v}{(\lambda + \mu)/v + 1 + \sqrt{(\lambda^2 + \mu^2)/v^2 + 1 - 2(\lambda\mu + \mu v + \lambda v)/v^2}} = \\ &= -\frac{1}{2} \frac{4(\lambda + \mu)}{1 + 1} = -(\lambda + \mu). \end{aligned}$$

Коэффициенты у функции $e^{\alpha_i t}$, $i=1,2$, вычисляются аналогично.

Например, вычислим $r(t) = r(t, \lambda, \mu, \infty)$. Имеем

$$r(t) = \lim_{\nu \rightarrow \infty} \left(\frac{\lambda \nu}{\lambda \nu + \lambda \mu + \mu \nu} - \frac{\lambda(\nu + \alpha_1)}{\alpha_1(\alpha_2 - \alpha_1)} e^{\alpha_1 t} + \frac{\lambda(\nu + \alpha_2)}{\alpha_2(\alpha_2 - \alpha_1)} e^{\alpha_2 t} \right),$$

или

$$r(t) = \frac{\lambda}{\lambda + \mu} + \lim_{\nu \rightarrow \infty} \left(- \frac{\lambda(\nu + \alpha_1) e^{\alpha_1 t} - \lambda(\nu + \alpha_2) e^{\alpha_2 t}}{\lambda \mu + \lambda \nu + \mu \nu - \alpha_1^2} \right).$$

Разделив числитель и знаменатель на ν и переходя к пределу, получим

$$r(t) = \frac{\lambda}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} \cdot e^{-(\lambda + \mu)t}.$$

Аналогично

$$\varphi(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \cdot e^{-(\lambda + \mu)t}.$$

Для стационарного режима имеем

а) с учетом переключения

$$\begin{cases} r = \lim_{t \rightarrow \infty} r(t) = \frac{\lambda \nu}{\lambda \mu + \lambda \nu + \mu \nu}, \\ \varphi = \lim_{t \rightarrow \infty} \varphi(t) = \frac{\mu \nu}{\lambda \mu + \lambda \nu + \mu \nu}, \\ l = \lim_{t \rightarrow \infty} l(t) = \frac{\lambda \mu}{\lambda \mu + \lambda \nu + \mu \nu}, \end{cases}$$

б) без учета переключения ($\nu \rightarrow \infty$):

$$r = \frac{\lambda}{\lambda + \mu}, \quad \varphi = \frac{\mu}{\lambda + \mu}.$$

При расчете показателей надежности по формулам (3) корни α_1, α_2 характеристического уравнения для системы (2) существуют, если

$$\lambda^2 + \mu^2 + \nu^2 \geq 2(\lambda \mu + \lambda \nu + \mu \nu). \quad (7)$$

Общее решение системы (2) для случая, когда (7) не выполняется, имеет вид

$$\begin{cases} r(t) = \frac{\lambda \nu}{\alpha_1 \alpha_2} + e^{\alpha} (C_1 \cdot \cos \beta t + C_2 \cdot i \sin \beta t), \\ \varphi(t) = \frac{\mu \nu}{\alpha_1 \alpha_2} + e^{\alpha} \left(\frac{(\mu + \alpha) C_1 + \beta \cdot C_2}{\lambda} \cos \beta t + i \frac{(\mu + \alpha) C_2 - \beta \cdot C_1}{\lambda} \sin \beta t \right), \\ l(t) = 1 - r(t) - \varphi(t) \end{cases}$$

где $\alpha_1 \alpha_2 = \lambda \mu + \lambda \nu + \mu \nu$, $\alpha = -(\lambda + \mu + \nu) / 2$;

$$\beta = -\sqrt{|\lambda^2 + \mu^2 + \nu^2 - 2(\lambda \mu + \mu \nu + \lambda \nu)|} / 2.$$

Теперь перейдем к показателям потенциальной живучести ВС.

Пусть $Z(t)$ – математическое ожидание числа переключаемых ЭМ в момент времени t .

Для случая, если в начальный момент времени $t = 0$ в ВС все ЭМ исправны, имеем $r(0) = 0$, $\varphi(0) = 1$, $l(0) = 0$ и следовательно

$$R(t) = N\varphi(t) = N\left(\frac{\mu\nu}{\alpha_1\alpha_2} - \frac{(\nu + \alpha_1)(\mu + \alpha_1)}{\alpha_1(\alpha_2 - \alpha_1)}e^{\alpha_1 t} + \frac{(\nu + \alpha_2)(\mu + \alpha_2)}{\alpha_2(\alpha_2 - \alpha_1)}e^{\alpha_2 t}\right),$$

$$M(t) = Nr(t) = N\left(\frac{\lambda\nu}{\alpha_1\alpha_2} - \frac{\lambda(\nu + \alpha_1)}{\alpha_1(\alpha_2 - \alpha_1)}e^{\alpha_1 t} + \frac{\lambda(\nu + \alpha_2)}{\alpha_2(\alpha_2 - \alpha_1)}e^{\alpha_2 t}\right),$$

$$Z(t) = Nl(t) = N\left(\frac{\lambda\mu}{\alpha_1\alpha_2} - \frac{\lambda\mu}{\alpha_1(\alpha_2 - \alpha_1)}e^{\alpha_1 t} + \frac{\lambda\mu}{\alpha_2(\alpha_2 - \alpha_1)}e^{\alpha_2 t}\right),$$

$$R(t) + M(t) + Z(t) = N.$$

На рис. 5 представлена зависимость количества отказавших ЭМ $M(t)$ и машин в переключении $Z(t)$ от времени.

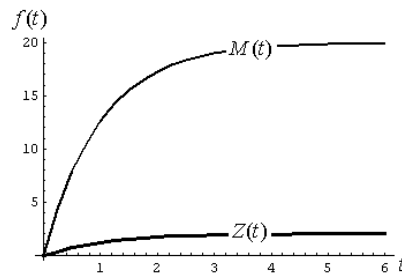


Рис. 5. Зависимость среднего числа отказавших и машин в переключении от времени: $N = 20000$, $\lambda = 0,001$ 1/ч., $\mu = 1$ 1/ч., $\nu = 10$ 1/ч.

Для стационарного режима имеем:

$$R = \frac{N\mu\nu}{\lambda\mu + \mu\nu + \lambda\nu}, \quad M = \frac{N\lambda\nu}{\lambda\mu + \mu\nu + \lambda\nu}, \quad Z = \frac{N\lambda\mu}{\lambda\mu + \mu\nu + \lambda\nu}. \quad (8)$$

Видно, что если пренебречь временем переключения, то решения (8) трансформируются в решения (1) двухпараметрической системы.

3. Вероятности состояний для трехпараметрической системы. Требуется найти $P_{k,l}(t)$ – вероятность того, что в момент времени t , число k ЭМ находятся в состоянии отказа и l ЭМ в состоянии переключения при условии, что в начальный момент времени, $t = 0$, в состоянии отказа находилось i ЭМ и в состоянии переключения j ЭМ, $k, l, i, j \in E_0^N = \{0, 1, \dots, N\}$, $k + l \leq N$, $i + j \leq N$. Считаем, что все параметры пуассоновские.

Для трех параметрической системы граф схема, представленная на рис. 2, трансформируется в трехпараметрическую граф схему СМО, фрагмент которой представлен на рис. 5.

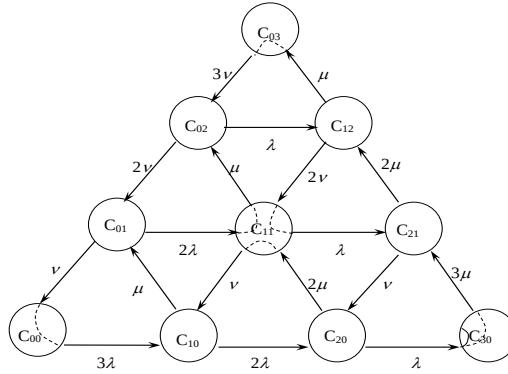


Рис. 6. Фрагмент граф-схемы при $N=3$

В соответствие рис. 6, составляем систему дифференциальных уравнений [13–15] с неизвестными функциями $P_{k,l}(t)$. Состоянию СМО – C_{kl} соответствует вероятность $P_{k,l}(t)$.

$$\left\{ \begin{array}{l} \frac{d}{dt} P_{k,0}(t) = -[(N-k)\lambda + k\mu]P_{k,0}(t) + \lambda(N-k+1)P_{k-1,0}(t) + \nu P_{k,1}(t), \quad k \in E_0^N; \\ \frac{d}{dt} P_{0,l}(t) = -[l\nu + (N-l)\lambda]P_{0,l}(t) + \mu P_{1,l-1}(t) + \nu(l+1)P_{0,l+1}(t), \quad l \in E_0^N; \\ \frac{d}{dt} P_{k,l}(t) = -[l\nu + (N-k-l)\lambda + k\mu]P_{k,l}(t) + \lambda(N-k-l+1)P_{k-1,l}(t) + \\ + (k+1)\mu P_{k+1,l-1}(t) + \nu(l+1)P_{k,l+1}(t), \quad k+l < N; \\ \frac{d}{dt} P_{k,N-k}(t) = -[(N-k)\nu + k\mu]P_{k,N-k}(t) + (k+1)\mu P_{k+1,N-k-1}(t) + \\ + \lambda P_{k-1,N-k}(t), \quad k+l = N; \end{array} \right. \quad (9)$$

с условием нормировки $\sum_{k=0}^N \sum_{l=0}^{N-k} P_{k,l}(t) = 1$ и начальными условиями $P_{i,j}(0) = 1$,

$P_{k,l}(0) = 0$, для $k \neq i$, $l \neq j$, $k, l, i, j \in E^N$, $(k+l) \leq N$, $(i+j) \leq N$. Если $k+l > N$ и $k, l < 0$, то $P_{k,l}(t) \equiv 0$, $l \in E^N$.

Для решения системы (9) используем аппарат производящих функций [18, 19]. Введем производящую функцию

$$F(x, z, t) = \sum_{k=0}^N \sum_{l=0}^{N-k} z^k x^l P_{k,l}(t),$$

тогда

$$P_{k,l}(t) = \frac{1}{k!l!} \cdot \frac{\partial^{k+l}}{\partial x^l \cdot \partial z^k} F(0,0,t). \quad (10)$$

Применяя к системе (9) производящую функцию $F(x, z, t)$, получаем следующее уравнение в частных производных

$$\frac{\partial}{\partial t} F + (\lambda z^2 + (\mu - \lambda)z - \mu x) \frac{\partial}{\partial z} F + (\lambda z x + (\nu - \lambda)x - \nu) \frac{\partial}{\partial x} F = -N\lambda(1-z)F, \quad (11)$$

где $F = F(z, x, t)$. Начальные условия имеют

$$F(x, z, 0) = x^j z^i. \quad (12)$$

Решение уравнения (11), при начальных условиях (12), позволяет, учитывая (10), записать решение системы (9) в виде [20–21]

$$P_{k,l}(t) = \sum_{a=0}^k \sum_{b=0}^l \sum_{c=0}^k \sum_{d=0}^l I(a, b, c, d) r_1^{k-a-d}(t) l_1^c(t) \varphi_1^{N-i-j-k-l+A}(t) r_2^a(t) l_2^{l-b-c}(t) \varphi_2^{i-a-c}(t) \cdot r_3^d(t) l_3^{l-b-c}(t) \varphi_3^{j-b-d}(t), \quad (13)$$

где $A = a + b + c + d$,

$$I(a, b, c, d) = \frac{i! j! (N - i - j)}{(N - i - j - k + a + b - c)! a! b! c! d! (i - a - l + c + d)! (j - b - d)! (k - a - b)! l! b - c)!}$$

В частном случае, когда в начальный момент времени t все ЭМ работали, имеем

$$P_{k,l}(t) = \frac{N! r_1^k(t) l_1^l(t) \varphi_1^{N-k-l}(t)}{k! l! (N - k - l)!}, \quad (14)$$

где $r_i(t)$, $\varphi_i(t)$, $l_i(t)$, $i = 1, 2, 3$, удовлетворяют формулам (4)–(6).

На рис. 7 представлена зависимость вероятностей $P_{k,l}(t)$ при $N = 2000$ ЭМ; в начальный момент $t = 0$, в состоянии отказа находилось $i = 0$ ЭМ и в состоянии переключения $j = 0$ ЭМ; $\mu = 1$ 1/ч.; $\nu = 10$ 1/ч. Стационарный режим достигается достаточно быстро $t < 7$ час. При экспресс анализе функционирования ВС можно использовать простые формулы стационарного режима.

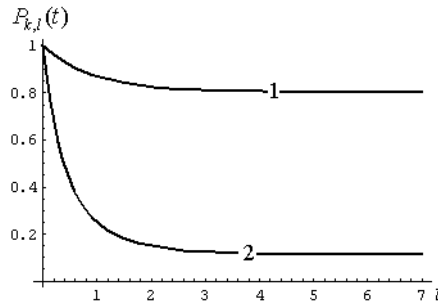


Рис. 7. Зависимость наиболее вероятного значения функций $P_{0,0}(t)$ от времени t :

$$1 - \lambda = 10^{-4} \text{ 1/ч.; } 2 - \lambda = 10^{-3} \text{ 1/ч.}$$

Для вероятностей $P_{k,l}$ функционирования ВС в стационарном режиме, имеем

а) с учетом времени переключения

$$P_{k,l} = \lim_{t \rightarrow \infty} P_{k,l}(t) = \frac{N! \lambda^{k+l} \mu^{N-k} \nu^N}{k! l! (N - k - l)! (\lambda \mu + \lambda \nu + \mu \nu)^N}; \quad (15)$$

б) без учета времени переключения ($\nu \rightarrow \infty$)

$$\begin{aligned}
 P_k &= \lim_{\substack{t \rightarrow \infty \\ \nu \rightarrow \infty}} P_k(t) = \frac{N!}{k!(N-k)!} \lim_{\substack{t \rightarrow \infty \\ \nu \rightarrow \infty}} (\varphi_i(t) + l_i(t))^{N-k} (1 - \varphi_i(t) - l_i(t))^k = \\
 &= \lim_{\substack{t \rightarrow \infty \\ \nu \rightarrow \infty}} \frac{N!}{k!(N-k)!} \left(\frac{\lambda\mu + \mu\nu}{\lambda\mu + \lambda\nu + \mu\nu} \right)^{N-k} \left(1 - \frac{\lambda\mu + \mu\nu}{\lambda\mu + \lambda\nu + \mu\nu} \right)^k = \\
 &= \frac{N!}{k!(N-k)!} \frac{\mu^{N-k} \lambda^N}{(\lambda + \mu)^N},
 \end{aligned}$$

то есть

$$P_k = \binom{N}{k} \left(\frac{\mu}{\lambda + \mu} \right)^N \left(\frac{\lambda}{\mu} \right)^k, \quad k \in E_0^N.$$

Таким образом, решение (15), при $\nu \rightarrow \infty$, совпадает с решением, представленным в работе [1], для двухпараметрической модели.

Заключение. Основной особенностью масштабируемых вычислительных систем является модульность. Нарастание производительности в таких системах достигается за счет увеличения однотипных элементов, элементарных машин (например, вычислительный узел). В результате отказов, производительность системы изменяется. Рассматривая ВС как объект стохастический, при его анализе, можно оценить потенциальные возможности вычислительной системы.

В рамках теории массового обслуживания рассмотрена модель функционирования масштабируемых вычислительных систем при отказах с учетом времени переключения. В работе предложены аналитические решения для распределения вероятностей состояний системы и для расчета показателей потенциальной живучести в случае модели с тремя параметрами. На примере расчетов, показано как соотносится количество машин на восстановлении с количеством машин в состоянии переключения. Тем самым предоставляется возможность пользователю самому определить, стоит ли пренебречь временем переключения. Показано, что решения трехпараметрической модели сводятся к решениям двух параметрической модели, если не принимать во внимание время переключения.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Хорошевский В.Г.* Архитектура вычислительных систем. – М.: МГТУ им. Баумана, 2008. – 520 с.
2. TOP500 Supercomputers Official Site. TOP500 Lists. – Режим доступа: <http://www.top500.org> (дата обращения 25.03.2020).
3. *Gupta S., Patel T., Engelmann C., Tiwari D.* Failures in large scale systems: long-term measurement, analysis, and implications // SC '17: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, Article No. 44, Denver, Colorado – November 12-17, 2017.
4. *Schroeder B., Gibson Garth.* A large-scale study of failures in high-performance computing systems // Proceedings of the International Conference on Dependable Systems and Networks (DSN2006), Philadelphia, PA, USA, June 25-28, 2006. – 10 p.
5. *Вишневский В.М.* Теоретические основы проектирования компьютерных сетей. – М.: Техносфера, 2003. – 512 с.
6. *Хорошевский В.Г.* Модели анализа и организации функционирования больших распределенных вычислительных систем // Электронное моделирование. – Киев, 2003. – Т. 25, № 6.
7. *Blischke W.R., Murthy D.N.P.* Reliability. – New York: Wiley, 2000.
8. *Hoyland A., Rausand M.* System reliability theory. – New York: Wiley, 1994.

9. Xie M., Dai Y.S., Poh K.L. Computing system reliability: models and analysis. – New York: Kluwer academic publishers, 2004.
10. Kuo W., Zuo M.J. Optimal reliability modeling: principles and applications. – New York: Wiley, 2003.
11. Чечельницкий А.А., Кучеренко О.В. Стационарные характеристики параллельно функционирующих систем обслуживания с двумерным входным потоком // Сб. научных статей. – Минск, 2009. – Вып. 2. – С. 262-268.
12. Назаров А.А., Терпугов А.Ф. Теория массового обслуживания. – Томск: Изд-во НТЛ, 2010. – 228 с.
13. Саати Т.Л. Элементы теории массового обслуживания и ее приложения. – 3-е изд. – М.: Книжный дом «ЛИБРОКОМ», 2010. – 520 с.
14. Клейнрок Л. Теория массового обслуживания. – М.: Машиностроение, 1979. – 432 с.
15. Боровков А.А. Вероятностные процессы в теории массового обслуживания. – М.: Наука, 1972. – 368 с.
16. Венцель Е.С. Теория случайных процессов и ее инженерные приложения. – М.: Наука, 1991. – 384 с.
17. Mor Harchol-Balter Performance Modeling and Design of Computer Systems: Queueing Theory in Action. – Cambridge University Press, 2013.
18. Феллер В. Введение в теорию вероятностей и ее приложения: в 2-х т. – Т. 1. – М.: «ЛИБРОКОМ», 2010. – 528 с.
19. Павский В.А., Павский К.В., Хорошевский В.Г. Вычисление показателей живучести распределенных вычислительных систем и осуществимости решения задач // Искусственный интеллект. – 2006. – № 4. – С. 28-34.
20. Такач Л. Комбинаторные методы в теории случайных процессов. – М.: Мир, 1971. – 264 с.
21. Райдер Г.Дж. Комбинаторная математика. – М.: Мир, 1966. – 154 с.

REFERENCES

1. Khoroshevskiy V.G. Arkhitektura vychislitel'nykh sistem [Architecture of computing systems]. Moscow: MGTU im. Bauman, 2008, 520 p.
2. TOP500 Supercomputers Official Site. TOP500 Lists. Available at: <http://www.top500.org> (accessed 25 March 2020).
3. Gupta S., Patel T., Engelmann C., Tiwari D. Failures in large scale systems: long-term measurement, analysis, and implications, SC '17: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, Article No. 44, Denver, Colorado – November 12-17, 2017.
4. Schroeder B., Gibson Garth. A large-scale study of failures in high-performance computing systems, Proceedings of the International Conference on Dependable Systems and Networks (DSN2006), Philadelphia, PA, USA, June 25-28, 2006, 10 p.
5. Vishnevskiy V.M. Teoreticheskie osnovy proektirovaniya komp'yuternykh setey [Theoretical foundations of computer network design]. Moscow: Tekhnosfera, 2003, 512 p.
6. Khoroshevskiy V.G. Modeli analiza i organizatsii funktsionirovaniya bol'shemasshtabnykh raspredelennykh vychislitel'nykh sistem [Models of analysis and organization of functioning of large-scale distributed computing systems], Elektronnoye modelirovaniye [Electronic simulation]. Kiev, 2003, Vol. 25, No. 6.
7. Blischke W.R., Murthy D.N.P. Reliability. New York: Wiley, 2000.
8. Hoyland A., Rausand M. System reliability theory. New York: Wiley, 1994.
9. Xie M., Dai Y.S., Poh K.L. Computing system reliability: models and analysis. New York: Kluwer academic publishers, 2004.
10. Kuo W., Zuo M.J. Optimal reliability modeling: principles and applications. New York: Wiley, 2003.
11. Chechel'nitskiy A.A., Kucherenko O.V. Statsionarnye kharakteristiki parallel'no funktsioniruyushchikh sistem obsluzhivaniya s dvumernym vkhodnym potokom [Stationary characteristics of parallel functioning service systems with two-dimensional input flow], Sb. nauchnykh statey [Collection of scientific articles]. Minsk, 2009, Issue 2, pp. 262-268.
12. Nazarov A.A., Terpugov A.F. Teoriya massovogo obsluzhivaniya [Queueing theory]. Tomsk: Izd-vo NTL, 2010, 228 p.

13. Saati T.L. Elementy teorii massovogo obsluzhivaniya i ee prilozheniya [Elements of Queuing theory and its applications]. 3 ed. Moscow: Knizhnyy dom «LIBROKOM», 2010, 520 p.
14. Kleynrok L. Teoriya massovogo obsluzhivaniya [Theory of Queuing]. Moscow: Mashinostroenie, 1979, 432 p.
15. Borovkov A.A. Veroyatnostnye protsessy v teorii massovogo obsluzhivaniya [Probabilistic processes in the theory of Queuing]. Moscow: Nauka, 1972, 368 p.
16. Venttsel' E.S. Teoriya sluchaynykh protsessov i ee inzhenernye prilozheniya [Theory of random processes and its engineering applications]. Moscow: Nauka, 1991, 384 p.
17. Mor Harchol-Balter Performance Modeling and Design of Computer Systems: Queueing Theory in Action. Cambridge University Press, 2013.
18. Feller V. Vvedenie v teoriyu veroyatnostey i ee prilozheniya [Introduction to probability theory and its applications]: in 2 vol. Vol. 1. Moscow: «LIBROKOM», 2010, 528 p.
19. Pavskiy V.A., Pavskiy K.V., Khoroshevskiy V.G. Vychislenie pokazateley zhivuchesti raspredelennykh vychislitel'nykh sistem i osushchestvimosti resheniya zadach [Calculating the survivability indicators of distributed computing systems and the feasibility of solving problems], *Iskusstvennyy intellekt* [Artificial intelligence], 2006, No. 4, pp. 28-34.
20. Takach L. Kombinatornye metody v teorii sluchaynykh protsessov [Combinatorial methods in the theory of random processes]. Moscow: Mir, 1971, 264 p.
21. Rayder G.Dzh. Kombinatornaya matematika [Combinatorial mathematics]. Moscow: Mir, 1966, 154 p.

Статью рекомендовал к опубликованию д.т.н., профессор Г.М. Лебедев.

Павский Валерий Алексеевич – Кемеровский государственный университет; e-mail: pavva46@mail.ru; 50043, г. Кемерово, ул. Красная, 6; тел.: +73842396832; кафедра общей математики и информатики; д.т.н.; профессор.

Павский Кирилл Валерьевич – Институт физики полупроводников им. А.В. Ржанова СО РАН; e-mail: pkv@isp.nsc.ru; 630090, г. Новосибирск, пр. Ак. Лаврентьева, 13; тел.: +73833332171, 3305626; д.т.н.; доцент; зав. Лабораторией вычислительных систем; профессор кафедры вычислительных систем Сибирского государственного университета телекоммуникаций и информатики.

Pavsky Valery Alekseevich – Kemerovo State University; e-mail: pavva46@mail.ru; 6, Krasnaya Street, Kemerovo, 650043, Russia; phone: +73842396832; the department of general mathematics and informatics; dr. of eng. sc.; professor.

Pavsky Kirill Valerievich – Rzhzanov Institute of Semiconductor Physics Siberian Branch of Russian Academy of Sciences; e-mail: pkv@isp.nsc.ru; 13, Ak. Lavrentiev Ave., Novosibirsk, 630090, Russia; phone: +73833332171, 3305626; dr. of eng. sc.; head of computer systems laboratory; professor of computer systems department, SibSUTIS.

Раздел III. Эволюционное моделирование

УДК 004.896

DOI 10.18522/2311-3103-2020-2-146-157

Б.К. Лебедев, О.Б. Лебедев, Е.О. Лебедева

ЭВОЛЮЦИОННЫЙ АЛГОРИТМ РАЗБИЕНИЯ МЕТОДОМ КРИСТАЛЛИЗАЦИИ РОССЫПИ АЛЬТЕРНАТИВ*

Работа алгоритма разбиения базируется на использовании коллективной эволюционной памяти, под которой подразумевается информация, отражающая историю поиска решения и хранится независимо от индивидуумов. Алгоритм, связанный с эволюционной памятью, стремится к запоминанию и многократному использованию способов достижения лучших результатов. Коллективная эволюционная память алгоритма разбиения состоит из некоторого количества статистических индикаторов, отображающих для каждого выполненного варианта число θ его вхождений в состав лучших решений на выполненных генерациях алгоритма и число, δ определяющее насколько полезна реализованная альтернатива при формировании результатов на прошлых генерациях алгоритма. Коллектив не имеет централизованного управления, и в связи с этим используется не прямой обмен информацией. Непрямой обмен состоит в выполнении неких действий, в разное время, при которых происходит изменение некоторых частей эволюционной памяти одним агентом. В дальнейшем происходит использование этой измененной информации другими агентами, в этих частях. Вначале на каждой итерации конструктивным алгоритмом формируется n_k решений Q_k . Каждое решение Q_k является отображением $F_k = V \rightarrow X$, представляется в виде двудольного подграфа D^k и формируется путем последовательного назначения элементов в узлы. Формирование каждого решения Q_k выполняется множеством агентов A , посредством вероятностного выбора каждым агентом a_i узла v_j . Процесс назначения элемента в узел включает две стадии. На первой стадии выбирается агент a_i , а на второй стадии – узел v_j . При этом должно выполняться ограничение: каждому агенту множества A соответствует один единственный узел множества V . Рассчитывается оценка ξ_k решения Q_k и оценка полезности δ_k множества альтернатив, реализованных агентами в решении Q_k . На втором этапе агенты увеличивают в интегральной россыпи альтернатив R^* интегральную полезность множества альтернатив на величину δ_k . На третьем этапе осуществляется снижение оценок полезности δ_k интегральной россыпи альтернатив на величину μ . В работе используется циклический метод формирования решений. В этом случае наращивание оценок интегральной полезности δ_k множества позиций P выполняется после полного формирования множества решений Q на итерации I . Экспериментальные исследования проводились на основе сформированных тестовых примеров с полученным ранее оптимальным решением. Полученные результаты сравнивались с результатами полученными другими известными алгоритмами разбиения схем на части. Для сравнения был сформирован набор стандартных бенчмарков. Проанализировав полученные результаты, можно сделать вывод, что предложенный метод позволяет получать на 4–5 % решения качественнее, чем его аналоги.

Роевой интеллект; адаптация; кристаллизация; СБИС; разбиение; россыпь альтернатив; двудольный граф.

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

B.K. Lebedev, O.B. Lebedev, E.O. Lebedeva

EVOLUTION ALGORITHM FOR PARTITION BY METHOD OF CRYSTALLIZATION OF ALTERNATIVES FIELD

The operation of the partitioning algorithm is based on the use of collective evolutionary memory, which means information that reflects the history of the search for a solution and is stored independently of individuals. The algorithm associated with evolutionary memory seeks to memorize and reuse ways to achieve better results. The collective evolutionary memory of the partitioning algorithm is a set of statistical indicators that reflect, for each implemented alternative, the number θ of its occurrences in the best solutions at previous iterations of the algorithm and the number δ indicating the usefulness of the implemented alternative when constructing solutions at previous iterations of the algorithm. The team does not have centralized management, and its features are the presence of indirect exchange of information. Indirect exchange consists in performing certain actions, at different times, during which some parts of evolutionary memory change by one agent. In the future, this changed information is used by other agents in these parts. First, at each iteration, a constructive algorithm generates n_k solutions Q_k . Each solution Q_k is a mapping $F_k = V \rightarrow X$, is represented as a bipartite subgraph D^k and is formed by sequentially assigning elements to nodes. The formation of each solution Q_k is performed by the set of agents A , by means of the probabilistic choice by each agent a_i of the node v_j . The process of assigning an element to a node involves two stages. In the first stage, agent a_i is selected, and in the second stage, the node. In this case, the restriction must be fulfilled: each agent of the set A corresponds to one unique node of the set V . The estimate ζ_k of the solution Q_k and the utility estimate δ_k of the set of alternatives implemented by the agents in the solution Q_k are calculated. At the second stage, the agents increase the integral utility of the set of alternatives in the integral placer of alternatives R^* by the value δ_k . At the third stage, the utility estimates δ_k of the integral placer of alternatives are reduced by μ . The paper uses the cyclic method of forming decisions. In this case, the building up of estimates of the integral utility δ_k of the set of positions P is performed after the complete formation of the set of solutions Q at iteration l . Experimental studies were carried out on the basis of formed test cases with the optimal solution obtained earlier. The results obtained were compared with the results obtained by other well-known algorithms for dividing circuits into parts. For comparison, a set of standard benchmarks was formed. After analyzing the results, we can conclude that the proposed method allows you to get 4–5 % better solutions than its analogues.

Swarm intelligence; adaptation; crystallization; VLSI; partitioning; placer alternatives; dicotyledonous graph.

Введение. При проектировании топологии СБИС выполняется ряд этапов связанных с решением оптимизационных задач. Одной из таких задач является задача разбиения схем на части. Поиск оптимального решения задачи разбиения схем на части, имеет важное значение, т.к. от найденного результата напрямую зависит качество выполнения другой задачи проектирования – трассировки межсоединений, что приводит к эффективной работоспособности всей схемы [1–2]. Задача разбиения – это класс NP -полных задач. В задачах такого класса невозможно определить решение точными алгоритмами [1, 3, 4]. К методам, позволяющим решать оптимизационные задачи такого типа, можно отнести алгоритмы с использованием процедур роевого интеллекта, например метод моделирования поведения муравьиной колонии [4–10].

В работе 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]. В процессе эволюционной коллективной адаптации методами дискриминантного анализа формируются оценки приспособленности альтернатив. Приспособленность альтернатив рас-

считается как вероятность ее использования в формируемом решении. 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).

Достоинство предложенного подхода состоит в возможности определять решение высокого качества, для задач большой размерности. Методика вычисления целевой функции отличается простотой, что позволяет легко оперировать с видоизмененными решениями в процессе поиска.

1. Постановка задачи разбиения. Пусть дан граф $G(X, U)$, где X – множество вершин, $|X|=n$, U – множество ребер. Необходимо разбить множество X на m непустых и непересекающихся подмножества X_j , $\cup X_j = X$, $X_i \cap X_j = \emptyset$, $X_j \neq \emptyset$. На формируемые узлы накладываются ограничения: $|X_j|=n_j$, $\sum n_j = n$. Будем считать, что множество вершин X_i помещается в узел v_i . Критерий оптимизации – суммарное число связей D между узлами. Цель оптимизации – минимизация критерия D [14–16].

Решить задачу разбиения равносильно построению всюду определенного, сюръективного и инъективного соответствия $F=V \rightarrow X$, где X – множество вершин, V – множество узлов.

Соответствие $F=V \rightarrow X$ представляется в виде двудольного графа $H=(V \cup X, E)$, где $V=\{v_j | j=1, 2, \dots, m\}$ – множество вершин (первая доля), соответствующих множеству узлов $V=\{v_j | j=1, 2, \dots, m\}$, а $X=\{x_i | i=1, 2, \dots, n\}$ – множество вершин (вторая доля), соответствующих множеству вершин графа $G(X, U)$, E – множество ребер (x_i, v_j) связывающих вершины $x_i \in X$ с вершинами множества V . $|E|=nm$, рис. 1.

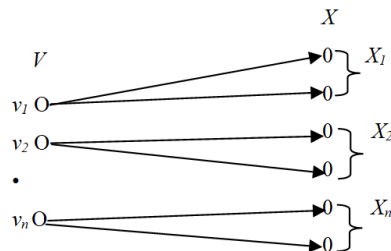


Рис. 1. Модель задачи разбиения

Образом узла v_j является множество вершин $X_j \subset X$, включенных в узел v_j , $F(v_j)=\{X_j\}$. Прообразом вершины $x_i \in X_j$ является узел v_j , $F^{-1}(v_j)=\{x_i\}$.

2. Разбиение схем на части с использованием метода кристаллизации россыпи альтернатив. В методе КРА [10–14] каждый результат Q_k синтезируется множеством представителей коллектива a_i отражает некоторое количество возможных состояний $C_i=\{c_{ij} | j=1, 2, \dots, s\}$. Каждый a_i может быть в одном из возможных состояний. Решение Q_k будет сформировано набором возможных состояний множества представителей коллектива.

Представление данной задачи полагает, что членами сообщества (агентами) выступают вершины графа, а альтернативами – множество узлов $V=\{v_j | j=1, 2, \dots, m\}$, где m – число узлов, в которые помещаются вершины. Можно сказать, что члены сообщества $A=\{a_i | i=1, 2, \dots, n\}$, сопоставляются множеству всех вершин. Варианты возможных состояний $C_i=\{c_{ij} | j=1, 2, \dots, s\}$ члена сообщества a_i сопоставляются множеству возможных вариантов распределения агента a_i . Вариант c_{ij} соответствует вхождению агента a_i в узел v_j [8].

В качестве графа поиска решений используется полный двудольный граф $H=(V \cup X, E)$, где $X=\{x_i | i=1, 2, \dots, n\}$ – множество вершин (первая доля), соответствующих множеству вершин графа $H(X, U)$, а $V=\{v_j | j=1, 2, \dots, m\}$, – множество вершин (вторая доля), соответствующих множеству узлов $|V|=m$. E – множество ребер (x_i, v_j) связывающих вершины $x_i \in X$ с вершинами множества V . $|E|=nm$.

Задача алгоритма разбиения сводится к поиску на полном двудольном графе $H=(V \cup X, E)$ двудольного подграфа $H^k=(V \cup X, E^k)$, $E^k=\{E^k_j | j=1, 2, \dots, m\}$ для которого оценка D имеет минимальное значение. Двудольный подграф $H^k=(V \cup X, E^k)$, обладает ограничениями:

- ♦ число ребер множества E^k_j , инцидентных вершине v_j , равно числу вершин множества X^k_j , $|X^k_j|=n_j$;
- ♦ каждое ребро $(x_i, v_j) \in E^k_j$, с одной стороны, инцидентно вершине $v_j \in V$, с другой стороны – инцидентно одной и только одной вершине $x_i \in X_j$; $U X_j = X$;
- ♦ локальная степень любой вершины $x_i \in X$ равна единице, $\rho(x_i)=1$;
- ♦ локальная степень вершины v_j равна мощности множества X_j , $\rho(v_j)=|X_j|$.

Представим структуру данных для отображения одного решения Q_k , формируемого множеством агентов A , в виде матрицы $R_k = ||r_{kji}||_{m \times n}$. Столбцы матрицы соответствуют агентам, строки – узлам, в которые может назначаться агент (состояниям агентов).

Размерность строки (вектора $R_{kj} = \{r_{kji} | i=1, 2, \dots, n\}$) определяется числом n агентов. Размерность столбца (вектора $R_{ki} = \{r_{kji} | j=1, 2, \dots, m\}$) – число m альтернатив состояний агента a_i .

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

Пусть для решения задачи разбиения построено множество решений $Q = \{Q_k | k=1, 2, \dots, n_k\}$. Для каждого решения Q_k рассчитаны оценки ζ_k и δ_k и сформирована индивидуальная россыпь альтернатив R_k .

Представим структуру данных для отображения b решений $Q_1 - Q_k$, сформированных множеством агентов A , в виде совокупности матриц $R_1 - R_k$, которую назовем россыпью альтернатив (РА).

Пример. Пусть имеется решение Q_k синтезированное десятью агентами. Агенты имеют 3 варианта состояния. В решении Q_k агентами реализовано разбиение множества вершин X , $|X|=n=10$ на 3 подмножества – X_1, X_2, X_3 : $|X_1|=4, |X_2|=3, |X_3|=3$. Решения имеют следующие оценки полезности: $\delta_1=4, \delta_2=6, \delta_3=2, \delta_4=8$.

Россыпь альтернатив для решений R_k имеет вид, представленный на рис. 2.

Множество индивидуальных россыпей альтернатив $R = \{R_k | k=1, 2, \dots, n_k\}$ формируется на базе сгенерированного множества решений $Q = \{Q_k | k=1, 2, \dots, n_k\}$.

Платформой для организации эволюционной процедуры поиска результата выступает интегральная россыпь альтернатив (ИРА) – R^* , получаемая при учете всех россыпей альтернатив:

$$R^* = ||r^*_{ji}||_{m \times n}, \text{ где } r^*_{ji} = \sum_k (r_{kji}), k = \{1, n_k\}.$$

По сути r^*_{ji} – это общая оценка полезностей результатов, в которых агентом a_j был выполнен вариант c_{ji} . Интегральная россыпь альтернатив используется в качестве коллективной эволюционной памяти и служит базой для формирования новых решений.

$R_1=$		агенты											
	блоки	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	n_j	
	v_1	4	0	0	4	0	0	4	0	4	0	4	x_1, x_4, x_7, x_9
	v_2	0	4	4	0	0	0	0	0	0	4	3	x_2, x_3, x_{10}
	v_3	0	0	0	0	4	4	0	4	0	0	3	x_5, x_6, x_8

$R_2=$		агенты											
	блоки	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	n_j	
	v_1	6	0	6	0	0	0	0	6	0	6	4	x_1, x_3, x_8, x_{10}
	v_2	0	6	0	0	6	0	6	0	0	0	3	x_2, x_5, x_7
	v_3	0	0	0	6	0	6	0	0	6	0	3	x_4, x_6, x_9

$R_3=$		агенты											
	блоки	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	n_j	
	v_1	2	2	0	0	0	0	2	0	0	2	4	x_1, x_2, x_7, x_{10}
	v_2	0	0	2	2	0	0	0	0	2	0	3	x_3, x_4, x_9
	v_3	0	0	0	0	2	2	0	2	0	0	3	x_1, x_3, x_8, x_{10}

$R_4=$		агенты											
	блоки	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	n_j	
	v_1	8	0	8	0	0	8	0	8	0	0	4	x_1, x_3, x_6, x_8
	v_2	0	8	0	8	8	0	0	0	0	0	3	x_2, x_4, x_5
	v_3	0	0	0	0	0	0	8	0	8	8	3	x_3, x_9, x_{10}

Рис. 2. Россыпи альтернатив R_k

Пример. На рис. 3. представлена интегральная россыпь альтернатив, R^* , объединяющая россыпи альтернатив $R_1 - R_4$

$R^*=$		агенты												
	блоки	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	n_j	max	min
	v_1	20	2	14	4	0	8	6	14	4	8	4	20	2
	v_2	0	18	6	10	14	0	7	0	2	10	3	18	2
	v_3	0	0	0	6	6	12	8	6	14	8	3	14	6

Рис. 3. Интегральная россыпь альтернатив R^*

Лучшее значение интегральной оценки полезности решений $f^{\#}=20$, худшее значение $f^0=2$. Выбирается параметр ε , лежащий в границах $f^0 \leq \varepsilon \leq f^{\#}$. Пусть $\varepsilon=1$.

Работа алгоритма разбиения базируется на использовании коллективной эволюционной памяти (КЭП), под которой подразумевается информация, отражающая историю поиска решения, и хранится независимо от индивидуумов. Алгоритм, связанный с эволюционной памятью, стремится к запоминанию и многократному использованию способов достижения лучших результатов. Коллективная эволюционная память алгоритма разбиения – это набор индикаторов, показывающих для каждой альтернативы число θ ее вхождений в наилучшие решения на предшествующих этому генерациях работы алгоритма и число δ отражающее полезность реализованной альтернативы при формировании результата на прошлых генерациях работы алгоритма. В основе такой системы лежит процесс самоорганизации, который позволяет получить общий результат коллектива, при помощи низкоуровневых совместных действий. У данного сообщества не имеется центрального управления, в связи, с чем присутствует не прямой обмен информацией. Такой обмен информацией, выглядит как распределенные во времени взаимные действия, т.е. такие действия, когда один из членов сообщества производит изменения в каких-то областях ко-эволюционной памяти, а другие члены этого сообщества в дальнейшем применяют эту видоизмененную информации содержащуюся в областях [8, 9].

Процесс поиска решений итерационный. На первом этапе каждой итерации конструктивным алгоритмом формируется n_k решений Q_k . Каждое решение Q_k является отображением $F_k = V \rightarrow X$, представляется в виде двудольного подграфа D^k и формируется путем последовательного назначения элементов в узлы. Формирование каждого решения Q_k выполняется множеством агентов A , посредством вероятностного выбора каждым агентом a_i узла v_j . Процесс назначения элемента в узел включает две стадии. На первой стадии выбирается агент a_i , а на второй стадии – узел v_j . При этом должно выполняться ограничение: каждому агенту множества A соответствует один единственный узел множества V . Рассчитывается оценка ξ_k решения Q_k и оценка полезности δ_k множества альтернатив, реализованных агентами в решении Q_k . На втором этапе агенты увеличивают в интегральной россыпи альтернатив R^* интегральную полезность множества альтернатив на величину δ_k . На третьем этапе осуществляется снижение оценок полезности δ_k интегральной россыпи альтернатив на величину μ . В работе используется циклический метод формирования решений. В этом случае наращивание оценок интегральной полезности δ_k множества позиций P выполняется после полного формирования множества решений Q на итерации l [12, 13].

После создания начальной ИРА некоторые полезности вариантов могут быть равны нулю. Следовательно, вероятность того, что эти варианты будут выбраны, также равна нулю. Следовательно, такие варианты сразу будут «отброшены» из списка рассматриваемых вариантов и не будут участвовать в процессе поиска решений. Для того, чтобы такие варианты не были удалены из рассматриваемого списка, выполняется корректировка начальной ИРА. Корректировка заключается в присвоении элементам имеющим значение «ноль», значения $f^0 \leq \varepsilon \leq f^\#$.

3. Алгоритм разбиения методом кристаллизации россыпи альтернатив.

Введем некоторые обозначения:

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

$\rho_j(t)$ – локальная степень узла v_j на шаге t .

$V^*(t) \subset V$ – множество узлов на шаге t таких, что для каждого узла $v_j \in V^*(t)$ его текущая локальная степень $\rho_j(t) < n_j$.

Представим алгоритм разбиения методом кристаллизации россыпи альтернатив.

1. Задается объем обучающей популяции решений – n_k .

2. Вначале формируется популяция решений $Q = \{Q_k | k=1, 2, \dots, n\}$ с помощью самопроизвольного определения вариантов c_{ij} . Вычисление параметров ξ_k и δ_k для всех решений Q_k популяции Q .

3. Уменьшение сформированной популяции решений Q к начальному размеру $|Q|=n_k$ с помощью выбора n_k наилучших результатов (с наибольшей обозначенной оценкой полезности δ_k).

Поиск в полученной популяции Q решения $Q^\#$, имеющего наилучшую оценку $\xi^\#$ и решения Q^0 , имеющего наихудшую оценку – ξ^0 .

4. Синтез исключительной россыпи альтернатив R_k для каждого решения $Q_k \in Q$.

5. Синтез вводной интегральной россыпи альтернатив $R^*(1)$ при помощи сведения в совокупность всех россыпей альтернатив R_k :

$$r_{ji}^*(1) = \sum_k (r_{kji}), k = \{1, n_k\}, \quad R^*(1) = ||r_{ji}^*||_{m \times n}.$$

6. Корректировка вводной интегральной россыпи альтернатив $R^*(1)$ при помощи возрастания элементов r_{ji}^* с нулевым значением на величину ε , $f^0 \leq \varepsilon \leq f^\#$.

7. Задается число итераций – n_l . $l=1$. (l – номер итерации).

8. $k=1$. (k – номер решения).

9. Обнуление всех россыпей альтернатив R_k . $t=1$. (t – номер шага).

10. Приводится в начальное состояние служебная память системы: $A(t)=A$, $\forall j(\rho_j(t)=0$. (Отметим, что $\rho_j(t)$ растет с ростом t). $V(t)=V$.

11. Первая стадия. На первой стадии шага t для каждого $a_i \in A(t)$ формируется множество $V^*(t)$ узлов таких, что для каждого узла $v_j \in V^*(t)$ его текущая локальная степень $\rho_j(t) < n_j$.

12. Для каждого $a_i \in A(t)$ среди узлов множества $V^*(t)$ отыскивается узел с максимальным значением полезности назначения в него a_i , которое назовем стоимостью $\pi_i(t)$ агента a_i на шаге t .

13. Среди $a_i \in A(t)$ с вероятностью $P_{ik}(t) = \pi_i(t) / \sum_i (\pi_i(t))$, пропорциональной стоимости $\pi_i(t)$ агента a_i выбирается агент a^*_i .

14. Вторая стадия. Среди множества узлов $V^*(t)$, доступных на шаге t агенту a^*_i , с вероятностью $P_{ij}(t) = r_{ji} / \sum_j (r_{ji})$ ($j | v^*_j \in V^*(t)$), пропорциональной полезности r_{ji} назначения a^*_i в узел v^*_j , выбирается узел $v^*_j \in V^*(t)$.

15. Агент a^*_i назначается в узел v^*_j .

16. $t = t + 1$ локальная степень узла $v^*_j \in$ увеличивается на 1, $\rho_j(t) = \rho_j(t-1) + 1$; a^*_i исключается из $A(t-1)$, $A(t) = A(t-1) \setminus a^*_i$.

17. Если все элементы a_i назначены в узлы, т.е. $A(t) = \emptyset$, то переход к 16, иначе переход к 9.

18. Расчет оценок ζ_k решения Q_k , $\delta_k = f(\zeta_k)$.

19. Запоминание лучшей оценки. Если оценка ζ_k решения Q_k лучше оценки $\zeta^{\#}$ решения $Q^{\#}$, то $\zeta^{\#} = \zeta_k$, а $Q^{\#} = Q_k$.

20. Формируется индивидуальная россыпь альтернатив R_k для Q_k .

21. Если $(k < n_k)$ то $k = k + 1$ и переход к 9, иначе переход к 22.

22. Если $(l < N_l)$, то $l = l + 1$ и переход к 23, иначе переход к 25.

23. Суммирование оценок полезности интегральной россыпи альтернатив $R(l-1)^*$, построенной на предыдущей итерации $(l-1)$, с соответствующими оценками полезности всех индивидуальных россыпей альтернатив R_k , сформированных на итерации l .

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

25. Завершение работы алгоритма. Фиксация и вывод лучшего решения $Q^{\#}$.

Чтобы увеличить вероятность выбора альтернативы, разработана (модификация) следующая методика. В ИРА отыскивается вариант имеющий минимум $(r^*_{ij})_{\min}$. Отыскивается параметр $q < (r^*_{ij})_{\min}$. После модификации формула примет вид:

$$p_{ij} = (r^*_{ij} - q) / (\sum_j (r^*_{ij} - q)).$$

4. Проведение тестовых испытаний. Создана программа Crystallization of alternatives field partitioning (CAFP), в основу работы которой заложены новые разработанные метаэвристики, позволяющие находить оптимальные решения.

Экспериментальные исследования проводились на основе сформированных тестовых примеров с полученным ранее оптимальным решением. Полученные результаты сравнивались с результатами полученными другими известными алгоритмами разбиения схем на части. Для сравнения был сформирован набор стандартных бенчмарков. Проанализировав полученные результаты, можно сделать вывод, что предложенный метод позволяет получать на 4–5 % решения качественнее, чем его аналоги [14, 17–20].

Для исследования синтезировались примеры, содержащие до 1000 вершин. Вес вершин и вес всех рёбер принимался равным единице. При этом графы «разбивались» на подграфы с разным количеством вершин в каждом подграфе.

В качестве оценки качества используется безразмерная величина $F_{\text{опт}}/F$, где F – оценка решения полученного, полученного с помощью разработанного алгоритма. Выполнив тестовые испытания и проведя их анализ, построена графическая зависимость, представленная на рис.4 и отражающая процесс получения качест-

венного решения от количества выполненных генераций. Также, на рис. 5 приведена графическая зависимость, отражающая процесс получения качественного решения от размера популяции. Установлено что начальное количество оценки полезности должно быть в 16 раз больше среднего количества значения оценки полезности, добавляемого в интегральную россыпь альтернатив на каждой итерации. Коэффициент обновления интегральной россыпи альтернатив $\rho=0,88$.

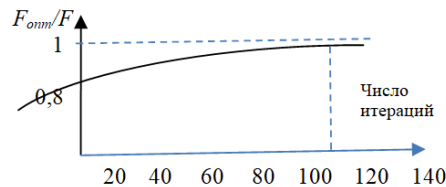


Рис. 4. Зависимость качества решений алгоритма CAFP от числа итераций

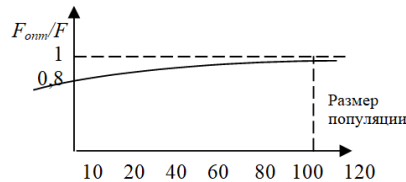


Рис. 5. Зависимость качества решений алгоритма CAFP от размера популяции

В ходе тестовых испытаний определено, что имея множество решений равным сто, схождение разработанного алгоритма разбиения схем на части, происходит примерно на сто десятой генерации. Отметим, что у ряда экспериментов число итераций меньше среднего составляло до 8 %, а больше с среднего до 30 %. Временная сложность процедуры на одной выполняемой генерации алгоритма равна $O(n)$.

Сравнение полученных результатов алгоритма разбиения схем на части, с разработанной метаэвристикой интеллектуальной оптимизации, и известными в литературе другими алгоритмами разбиения, осуществлялся с помощью набора стандартных бенчмарков [14–16]. Эти тестовые бенчмарки являются используемыми в промышленности фрагментами схем, отдельными блоками, сверхбольшими интегральными схемами и т.д. Параметрами, используемыми для сравнения выступали:

- ♦ Cut – количество межсоединений, оказавшихся в разрезе после разбиения схемы на части;

- ♦ CPU – время работы, требуемое для нахождения результата.

Выполнена сравнительная оценка работы разработанных алгоритмов на бенчмарках cordic(881), misex3(1349), X3(1369). c6288(2435) s15850(4321), frisc(4400), elliptic(4711), (в скобках приведено число вентиляей) с оценками известных из литературы алгоритмов (оценки получены на тех же бенчмарках): Pure hMetis [15]; Net&Path Based [16]; PPFF [16]; VPR [15]. Результаты показаны в табл. 1.

Проведя тестовые испытания и выполнив анализ полученных решений, можно сказать, что алгоритм разбиения схем на части с встроенной в него разработанной новой метаэвристикой интеллектуальной оптимизации, позволяет получать результаты не хуже, а в ряде случаев лучше имеющихся известных алгоритмов на 4–5 %.

Например, разработанный алгоритм дает значительно лучшие результаты, чем hMetis. Проведенный тест схемы Cordic показал, что решение, полученное в результате выполнения алгоритмом CAFP, лучше на 5,1 % чем решение, полученное в результате выполнения алгоритмом hMetis, который считается алгоритмом позволяющим находить оптимальные решения.

Таблица 1

Сравнительная оценка работы алгоритмов

		Cordic	misex3	X3	c6288	s15850	Frisc	Elliptic
Pure hMetis	Cut	327	543	211	182	617	838	508
	CPU	2	5	4	31	28	34	16
Net& Path	Cut	278	614	261	216	671	869	619
	CPU	5	13	6	80	37	61	22
PPFF	Cut	313	540	228	201	615	851	510
	CPU	6	14	8	79	43	66	25
VPR	Cut	319	543	216	204	622	843	513
	CPU	10	18	13	85	47	71	27
CAFP	Cut	270	600	251	207	621	812	517
	CPU	4	11	6	74	300	46	14

Заключение. В работе выполнен обзор известных в литературе алгоритмов, применяемых для решения задачи разбиения схем на части. Проведя анализ этих алгоритмов, выбран метод интеллектуальной оптимизации, на основе которого строился представленный алгоритм.

Разработан новый способ решения оптимизационных задач с использованием интеллектуальных процедур. Этот способ заключается в применении выбранных лучших из найденных вариантов решений. В основу этого метода положено построение моделей коллективного интеллекта. Набор данных о вариантах реализации и их целевых функциях, является совокупностью вариантов. Идея метода заключается в следующем. Каждое решение строится агентами, при этом у агента имеется множество альтернатив состояний или же он находится в какой то конкретной альтернативе. Формирование решения происходит на основе возможных вариантов состояний всех агентов.

Выполнен анализ различных вариантов состояний, возникающих в ходе коэволюционной адаптации. Этот процесс назван кристаллизацией, т.к. является аналогом вычленения объектов (формирования кристаллов). Разработанный метод достаточно эффективен при выполнении поиска решений в оптимизационных задачах, особенно в тех задачах, где возможно их представление в виде россыпи альтернатив.

Проведенные тестовые испытания позволили сделать вывод, что применяя алгоритмы на основе разработанного метода, для решения оптимизационных задач, полученные результаты дают улучшение целевой функции до 5%. Временная зависимость работы алгоритма с встроенной в него разработанной метаэвристикой, находится в интервале $O(n^2)$ - $O(n^3)$.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Немудров В., Мартин Г. Системы-на-кристалле. Проектирование и развитие. – М.: Техносфера, 2004. – 157 с.
2. Казеннов Г.Г. Основы проектирования интегральных схем и систем. – М.: Бином. Лаборатория знаний, 2005. – 295 с.
3. Alpert C.J., Mehta D.P., Sapatnekar S.S. Handbook of Algorithms for Physical Design Automation. Boston. – MA: Auerbach, 2009. – 245 p.
4. Карпенко А.П. Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой: учеб. пособие. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2014. – 448 с.

5. Курейчик В.М., Лебедев Б.К., Лебедев О.Б. Гибридный алгоритм разбиения на основе природных механизмов принятия решений // Искусственный интеллект и принятие решений. – М.: Изд-во Институт системного анализа РАН, 2012. – С. 3-15.
6. Лебедев Б.К., Лебедев О.Б. Муравьиные алгоритмы разбиения, использующие представление задачи, отличное от канонического // Вестник Ростовского государственного университета путей сообщения. – 2016. – № 2 (62). – С. 71-77.
7. Лебедев Б.К., Лебедев О.Б. Разбиение на основе многоуровневой параллельной эволюционной адаптации // Проблемы разработки перспективных микро- и наноэлектронных систем – 2008: Сб. научных трудов / под ред. А.Л. Стемпковского. – М.: ИПИМ РАН, 2008. – С. 36-41.
8. Ворбьева Е.Ю., Карпенко А.П., Селиверстов Е.Ю. Ко-гибридизация алгоритмов роя частиц // Наука и жизнь. – 2012. – № 4.
9. Zhou Y., Pei Sh. A hybrid co-evolutionary particle swarm optimization algorithm for solving constrained engineering design problems // Journal of computers (JCP). – 2010. – Vol. 5, No. 6. – P. 965-972.
10. Лебедев Б.К., Лебедев В.Б. Оптимизация методом кристаллизации россыпи альтернатив // Известия ЮФУ. Технические науки. – 2013. – № 7 (144). – С. 11-17.
11. Clerc M. Particle Swarm Optimization. – ISTE, London, UK, 2006. – 198 p.
12. Poli R. Analysis of the publications on the applications of particle swarm optimization // Journal of Artificial Evolution and Applications, Article ID 685175, 2008.
13. Лебедев Б.К., Лебедев В.Б. Построение кратчайших связывающих соединений методом кристаллизации россыпи альтернатив // МЭС-2014. VI Всероссийская научно-техническая конференция «Проблемы разработки перспективных микро- и наноэлектронных систем – 2014»: Сб. трудов / под общ. ред. академика РАН А.Л. Стемпковского. – М.: ИПИМ РАН, 2014. Ч. I. – С. 177-183.
14. 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.
15. Selvakkumaran N., Karypis G. Multi-Objective Hypergraph Partitioning Algorithms for Cut and Maximum Subdomain Degree Minimization // ICCAD. – 2003. – P. 234-249.
16. Ababei C., Selvakkumaran N., Bazargan K., Karypis G. Multi-objective Circuit Partitioning for Cutsizes and Path-Based Delay Minimization // ICCAD. – 2002. – P. 118-137.
17. Агасиев Т.А., Карпенко А.П. Современные техники глобальной оптимизации // Информационные технологии. – 2018. – № 6. – С. 370-386.
18. Норенков И.П., Уваров М.Ю. Поддержка принятия решений на основе паттернов проектирования // Наука и образование: научное издание. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2011. – № 9. – С. 25-32.
19. Божко А.Н. Методы структурного анализа сложных изделий в интегрированных CFD/CAM-системах // Информационные технологии. – 2018. – № 8. – С. 499-506.
20. Столяров А.И., Донецкая Ю.В., Гатчин Ю.А. Особенности САПР при проектировании комплекса // Вестник Иркутского государственного технического университета. – 2018. – № 7 (138). – С. 88-95.

REFERENCES

1. Nemudrov V., Martin G. Sistemy-na-kristalle. Proektirovanie i razvitie [Systems-on-a-chip. Design and development]. Moscow: Tekhnosfera, 2004, 157 p.
2. Kazennov G.G. Osnovy proektirovaniya integral'nykh skhem i system [Fundamentals of design of integrated circuits and systems]. Moscow: Binom. Laboratoriya znaniy, 2005, 295 p.
3. Alpert C.J., Mehta D.P., Sapatnekar S.S. Handbook of Algorithms for Physical Design Automation. Boston. MA: Auerbach, 2009, 245 p.
4. Karpenko A.P. Sovremennyye algoritmy poiskovoy optimizatsii. Algoritmy, vdokhnovlennyye prirodoy: ucheb. posobie [Modern search engine optimization algorithms. Algorithms inspired by nature]. Moscow: Izd-vo MGTU im. N.E. Bauman, 2014, 448 p.
5. Kureychik V.M., Lebedev B.K., Lebedev O.B. Gibridnyy algoritm razbieniya na osnove prirodnikh mekhanizmov prinyatiya resheniy [Hybrid partition algorithm based on natural decision-making mechanisms], *Iskusstvennyy intellekt i prinyatie resheniy* [Artificial intelligence and decision making]. Moscow: Izd-vo Institut sistemnogo analiza RAN, 2012, p. 3-15.

6. Lebedev B.K., Lebedev O.B. Murav'inye algoritmy razbieniya, ispol'zuyushchie predstavleniye zadachi, otlichnye ot kanonicheskogo [Ant partitioning algorithms using a task representation other than the canonical], *Vestnik Rostovskogo gosudarstvennogo universiteta puty soobshcheniya* [Bulletin of the Rostov State University of Railways], 2016, No. 2 (62), pp. 71-77.
7. Lebedev B.K., Lebedev O.B. Razbienie na osnove mnogourovnevnoy parallel'noy evolyutsionnoy adaptatsii [Partitioning on the basis of multi-level parallel evolutionary adaptation], *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem* [Problems of developing promising micro- and nanoelectronic systems], 2008: *Sb. nauchnykh trudov* [Problems of developing promising micro- and nanoelectronic systems – 2008: Collection of scientific papers], ed. by A.L. Stempkovskogo. Moscow: IPPM RAN, 2008, pp. 36-41.
8. Vorb'eva E.Yu., Karpenko A.P., Seliverstov E.Yu. Ko-gibridizatsiya algoritmov roya chastits [Co-hybridization of particle swarm algorithms], *Nauka i zhizn'* [Science and Life], 2012, No/ 4.
9. Zhou Y., Pei Sh. A hybrid co-evolutionary particle swarm optimization algorithm for solving constrained engineering design problems, *Journal of computers (JCP)*, 2010, Vol. 5, No. 6, pp. 965-972.
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. Clerc M. Particle Swarm Optimization. ISTE, London, UK, 2006, 198 p.
12. Poli R. Analysis of the publications on the applications of particle swarm optimization, *Journal of Artificial Evolution and Applications*, Article ID 685175, 2008.
13. Lebedev B.K., Lebedev V.B. Postroenie krachayshikh svyazyvayushchikh soedineniy metodom kristallizatsii rossypi al'ternativ [Construction of the shortest binding compounds by the method of crystallization of alternatives field], *MES-2014. VI Vserossiyskaya nauchno-tekhnicheskaya konferentsiya «Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem – 2014»*: *Sb. trudov* [MES-2014. VI All-Russian scientific and technical conference "Problems of developing promising micro- and nanoelectronic systems – 2014": Proceedings], under the total. ed. akademika RAN A.L. Stempkovskogo. Moscow: IPPM RAN, 2014. Part I, pp. 177-183.
14. 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.
15. Selvakkumaran N., Karypis G. Multi-Objective Hypergraph Partitioning Algorithms for Cut and Maximum Subdomain Degree Minimization, *ICCAD*, 2003, pp. 234-249.
16. Ababei C., Selvakkumaran N., Bazargan K., Karypis G. Multi-objective Circuit Partitioning for Cutsizes and Path-Based Delay Minimization, *ICCAD*, 2002, pp. 118-137.
17. Agasiev T.A., Karpenko A.P. Sovremennye tekhniki global'noy optimizatsii, *Informatsionnye tekhnologii*, 2018, No. 6, pp. 370-386.
18. Norenkov I.P., Uvarov M.Yu. Podderzhka prinyatiya resheniy na osnove patternov proektirovaniya [Decision support based on design patterns], *Nauka i obrazovanie: nauchnoe izdanie* [Science and Education: Scientific publication]. Moscow: Izd-vo MGTU im. N.E. Bauman, 2011, No. 9, pp. 25-32.
19. Bozhko A.N. Metody strukturnogo analiza slozhnykh izdeliy v integrirovannykh CFD/CAM-sistemakh [Structural analysis methods for complex products in integrated CAD], *Informatsionnye tekhnologii* [CAM-systems. Information Technologies]. Moscow: Izd-vo MGTU im. N.E. Bauman, 2018, No. 8, pp. 499-506.
20. Stolyarov A.I., Donetskaya Yu.V., Gatchin Yu.A. Osobennosti SAPR pri proektirovanii kompleksa [Features of CAD in the design of the complex], *Vestnik Irkutskogo gosudarstvennogo tekhnicheskogo universiteta* [Bulletin of the Irkutsk State Technical University], 2018, No. 7 (138), pp. 88-95.

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.

Лебедева Екатерина Олеговна – e-mail: lbedevakate@mail.ru; тел.: 89289591426; кафедра систем автоматизированного проектирования; аспирант.

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.

Lebedeva Ekaterina Olegovna – e-mail: lbedevakate@mail.ru; phone: +79289591426; the department of computer aided design; graduate student.

УДК 004.021

DOI 10.18522/2311-3103-2020-2-157-169

А.А. Могилев, В.М. Курейчик

**МОДИФИЦИРОВАННЫЙ ГЕНЕТИЧЕСКИЙ АЛГОРИТМ
ПЛАНИРОВАНИЯ ПРОЕКТОВ, РЕАЛИЗОВАННЫЙ
С ИСПОЛЬЗОВАНИЕМ ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ***

Предложена структура модифицированного генетического алгоритма для решения задачи построения расписания проекта с учетом ограниченности ресурсов, реализованного с использованием облачных вычислений, проведен вычислительный эксперимент, в ходе которого было произведено сравнение результатов работы предложенного алгоритма с лучшими из известных, на данный момент, результатами. Исходя из результатов эксперимента был сделан вывод, о том, что предложенный алгоритм может быть использован для планирования работ реальных проектов, так как с его помощью возможно составлять расписания для проектов с количеством работ $n = 90$ за приемлемый промежуток времени. При планировании проектов с количеством работ $n = 30$, $n = 60$, $n = 90$, 120 время выполнения предложенного алгоритма было меньше, чем время выполнения стандартного генетического алгоритма в 2.8, в 4, в 5.5 и 6.8 раз соответственно. В связи с тем, что задача построения расписания проекта с учетом ограниченности ресурсов является NP-трудной, проблема создания новых и модификации существующих методов её решения по-прежнему остается актуальной. Для планирования проектов с большим количеством работ целесообразно использовать облачные вычисления, так как планирование таких проектов может потребовать много времени и вычислительных ресурсов. Использование облачных вычислений позволит сократить время выполнения генетического алгоритма за счет предоставления поставщиком облачного сервиса больших вычислительных ресурсов. В связи с этим, предложенный в данной работе алгоритм отличается от уже имеющихся использованием облачных вычислений для распределения нагрузки между рабочими станциями, на которых одновременно выполняется данный алгоритм. Применение в генетическом алгоритме модифицированных операторов, а также использование облачной инфраструктуры как услуги для реализации генетического алгоритма при решении задачи планирования проектов определяет научную новизну исследования.

Теория расписаний; задача построения расписания для проекта с учетом ограничений на ресурсы; генетический алгоритм.

* Работа выполнена при финансировании гранта РФФИ №18-07-00050.

A.A. Mogilev, V.M. Kureichik

MODIFIED GENETIC PROJECT PLANNING ALGORITHM IMPLEMENTED WITH THE USE OF CLOUD COMPUTING

The paper proposes a structure of a modified genetic algorithm for solving resource constrained project scheduling problem implemented with the use of cloud computing, a computational experiment was conducted, during which the results of the proposed algorithm were compared with the best known, at the moment, results. Based on the results of the experiment, it was concluded that the proposed algorithm can be used to plan the work of real projects, since it is possible to draw up schedules for projects with the number of works $n = 90$ for an acceptable period of time. When planning projects with the number of jobs $n = 30, n = 60, n = 90, 120$, the execution time of the proposed algorithm was less than the execution time of the standard genetic algorithm by 2.8, 4, 5.5 and 6.8 times, respectively. Due to the fact that the task of constructing a project schedule taking into account limited resources is NP-difficult, the problem of creating new and modifying existing methods for solving it remains relevant. For planning projects with a large number of works, it is advisable to use cloud computing, since planning such projects can require a lot of time and computing resources. In this regard, the algorithm proposed in this paper differs from the existing ones by using cloud computing to distribute the load between workstations on which this algorithm is simultaneously running. The use of modified operators in the genetic algorithm, as well as the use of cloud infrastructure as a service for implementing a distributed genetic algorithm, determines the scientific novelty of the study.

Scheduling theory; resource constrained project scheduling problem; genetic algorithm.

Введение. Задача оптимального планирования проекта при наличии ресурсных ограничений известна как задача планирования расписания для проекта с учетом ограниченности ресурсов – *RCPSP (resource constrained project scheduling problem)*. Эта задача остается актуальной среди различных исследователей на протяжении трех десятков лет, так как ее решения могут иметь практическое применение во многих отраслях. Различают три основных категории методов решения задачи *RCPSP* [1]: точные методы [2, 3], эвристические методы [4–6] и метаэвристические [7–14]. Предложенный в данной работе генетический алгоритм можно отнести к метаэвристическим методам.

Задача *RCPSP* является NP-трудной даже без дополнительных ограничений. Лучший из известных точных алгоритмов Брукера за приемлемое время может решать примеры размерности не больше шестидесяти работ [15]. В связи с этим, для достижения близких к оптимальным результатов решения задачи за приемлемое время в данной работе предлагается структура генетического алгоритма решения задачи *RCPSP*, основанная на использовании облачных вычислений. Использование параллельной структуры позволит достигать близких к оптимальным результатов за счет распределения альтернативных решений между тремя независимыми рабочими станциями.

Использование облачных вычислений позволит сократить время выполнения генетического алгоритма за счет предоставления поставщиком облачного сервиса больших вычислительных ресурсов.

В рамках концепции облачных систем принято выделять три модели обслуживания: облачное программное обеспечение как услуга (*Cloud Software as a Service, SaaS*), облачная платформа как услуга (*Cloud Platform as a Service, PaaS*) и облачная инфраструктура как услуга (*Cloud Infrastructure as a Service, IaaS*) [16].

Для апробации предложенного в данной работе алгоритма была выбрана модель «облачная инфраструктура как услуга». При использовании данной модели обслуживания потребителям предоставляются по подписке фундаментальные информационно-технологические ресурсы – виртуальные серверы с заданной вычислительной мощностью, операционной системой (чаще всего – предустановленной провайдером из шаблона) и доступом к сети.

Принципиальным отличием предложенного генетического алгоритма от уже имеющихся решений является использование распределенной структуры для выполнения в «облаке», а также использование модифицированных генетических операторов таких как параметризованный кроссинговер, мутация на основе дихотомии и транслокация. Проведенный вычислительный эксперимент показал, что использование модифицированных генетических операторов и облачных вычислений позволяет достигать близких к оптимальным результатов при планировании проектов с количеством работ $n = 90$ за приемлемый промежуток времени.

Постановка задачи построения расписания для проекта с учетом ограниченности ресурсов. В задаче построения расписания для проекта необходимо построить оптимальное расписание проекта (выполнения работ проекта) с учетом сетевого графика (отношений предшествования между работами) и с учетом необходимых/доступных ресурсов, при котором будет оптимизирована некоторая целевая функция. Самая популярная целевая функция – общее время выполнения проекта (*makespan* или $C_{\max}$) [15].

Постановка задачи построения расписания для проекта имеет следующий вид [15]. Дано множество работ (требований) $N = \{1, \dots, n\}$ и R возобновляемых ресурсов $r = 1, \dots, R$. В каждый момент времени t доступно Q_r единиц ресурса r . Заданы продолжительности обслуживания $d_i \geq 0$ для каждого требования $i = 1, \dots, n$. Во время обслуживания требования i требуется $q_{ir} \leq Q_r$ единиц ресурса $r = 1, \dots, R$. После завершения обслуживания требования, освобожденные ресурсы в полном объеме могут быть мгновенно назначены на обслуживание других требований.

Между некоторыми парами требований заданы ограничения предшествования: $i \rightarrow j$ означает, что обслуживание требования j начинается не раньше окончания обслуживания требования i . Обслуживание требований начинается в момент времени $t = 0$. Прерывание при обслуживании требований запрещены. Необходимо определить моменты времени начала обслуживания требований S_i , $i = 1, \dots, n$, так, чтобы минимизировать время выполнения всего проекта, т.е. минимизировать значение

$$C_{\max} = \max\{C_i\}, i = 1, \dots, n, \quad (1)$$

где $C_i = S_i + d_i$. При этом должны быть соблюдены следующие ограничения:

1) в каждый момент времени $t \in [0, C_{\max}]$ должно выполняться

$$\sum_{i=1}^n q_{ir} \phi_i(t) \leq Q_r, r = 1, \dots, R, \quad (2)$$

где $\phi_i(t) = 1$, если требование i обслуживается в момент времени t и $\phi_i(t) = 0$, в противном случае. То есть требования в процессе своего обслуживания должны быть полностью обеспечены ресурсами;

2) не нарушаются отношения предшествования между требованиями, т.е.

$$S_i + d_i \leq S_j, \text{ если } i \rightarrow j \text{ для } i, j \in N.$$

На рис. 1 представлен пример сетевого графа проекта. Вершинами графа являются работы проекта, над вершинами изображены продолжительность выполнения соответствующих работ d_i и необходимое количество ресурсов для выполнения работы r_i . Количество возобновляемых ресурсов проекта $R = 4$. На рис. 2 представлена диаграмма возможного расписания данного проекта, по вертикали указано количество ресурсов, по горизонтали – время выполнения работ.

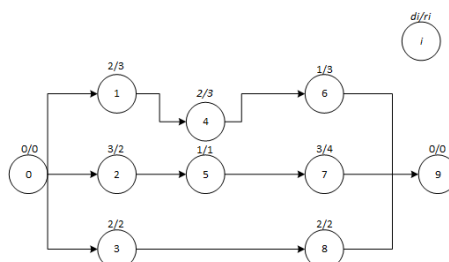


Рис. 1. Сетевой граф проекта

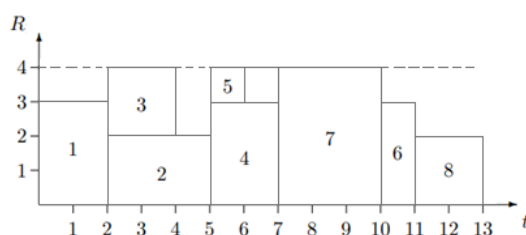


Рис. 2. Пример расписания для проекта

Модифицированный генетический алгоритм. Генетический алгоритм выполняется на четырех вычислительных машинах (Мастер, ПК1, ПК2 и ПК3), расположенных в «облаке» и состоит из следующих этапов:

- ♦ «Мастер» генерирует начальную популяцию хромосом, разделяет ее на три части и передает каждую из частей ПК1, ПК2 и ПК3;
- ♦ На ПК1, ПК2 и ПК3 выполняется модифицированный генетический алгоритм, представленный на рис. 4;
- ♦ Мастер выбирает наилучшее из трех решений в соответствии с целевой функцией.

Схема выполнения генетического алгоритма представлена на рис. 3.

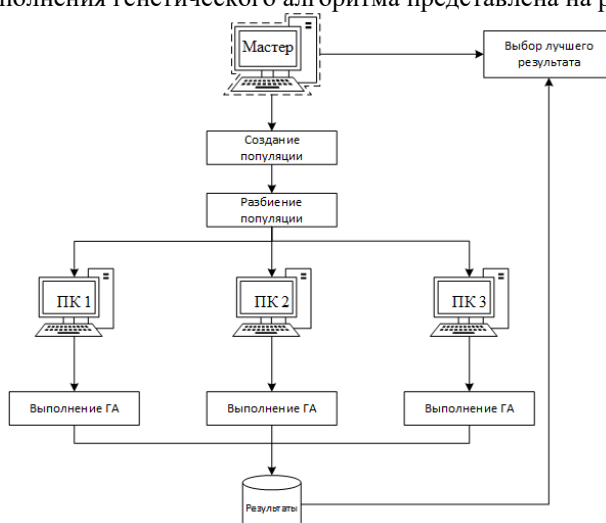


Рис. 3. Схема выполнения генетического алгоритма

Структурная схема генетического алгоритма для решения *RCPSP*-задачи, выполняемого на рабочих станциях ПК1, ПК2 и ПК3 представлена на рис. 4. Временная сложность данного алгоритма зависит от размера начальной популяции N и длины набора альтернативных решений L и определяется формулой:

$$T = O(NL). \quad (3)$$



Рис. 4. Структурная схема генетического алгоритма

Генерация начальной популяции. Процесс выполнения любого генетического алгоритма начинается с генерации набора альтернативных решений (хромосом) [17], который называется популяцией. Альтернативное решение характеризуется набором параметров, которые называют генами. В альтернативном решении ген соответствует элементу, концептуальному узлу, связи и т.п. [18]

При генерации начальной популяции необходимо учитывать два типа ограничений: ограничения предшествования (работа не может быть начата, прежде чем закончатся все её предшественники) и ресурсные ограничения (число ресурсов, выделенных на работы проекта в единицу времени, не может превышать максимальное доступное число ресурсов).

Популяция в предложенном генетическом алгоритме представляет собой набор альтернативных решений (рис. 5), каждый ген которых содержит: t_j – время начала работы, d_j – продолжительность работы, r_j – необходимое число ресурсов, где j – номер работы проекта. Таким образом решение C представляет собой потенциально возможное решение и состоит из последовательности работ:

$$C = \{j_1, j_2, j_3, \dots, j_n\}. \quad (4)$$

j_1	j_2	j_3	...	j_n
$t_{j_1}, d_{j_1}, r_{j_1}$	$t_{j_2}, d_{j_2}, r_{j_2}$	$t_{j_3}, d_{j_3}, r_{j_3}$	...	$t_{j_4}, d_{j_4}, r_{j_4}$

Рис. 5. Представление альтернативных решений

Селекция. После генерации начальной популяции необходимо выбрать пары хромосом для скрещивания. Выбор пар хромосом производится следующим образом. В качестве родителей берутся сперва наборы хромосом с наибольшим значением функции приспособленности, далее два следующих по значению и так далее пока в исходной популяции есть хромосомы, не прошедшие процесс скрещивания.

Качество популяции в генетических алгоритмах принято оценивать с помощью функции приспособленности. Так как задача построения расписания для проекта предполагает нахождение решения с минимально возможным временем выполнения проекта, функция приспособленности выбрана обратно пропорциональной сумме времени выполнения всех работ проекта [19]:

$$F(x) = \frac{1}{\max\{t_i + d_i | i=1, 2, \dots, n\}}. \quad (5)$$

Вероятность выживания особи зависит от значения данной функции.

Кроссинговер. Для построения потомков хромосом был использован оператор параметризованного равномерного кроссинговера [20]. Рассмотрим реализацию параметризованного равномерного кроссинговера на примере из рис. 2. Пусть даны две родительские хромосомы, выбранные случайным образом из всей популяции. Данные хромосомы представляют собой последовательности работ проекта $P_1 = \{1, 2, 3, 4, 5, 6, 7, 8\}$ и $P_2 = \{3, 1, 2, 5, 4, 8, 7, 6\}$. Для каждого нового гена в хромосоме потомке P_3 подбрасывается абстрактная «монета» для того, чтобы определить хромосому родителя, из которой будет браться ген для хромосомы потомка. Если выпадает «орел», то берется ген из хромосомы P_1 , если «решка» - из P_2 . Если в хромосоме потомке уже имеется выбранная работа, то выбирается ген из другого родителя. В таблице 1 представлен результат выполнения данного оператора, с вероятностью выпадения «орла» $P = 0.6$.

Таблица 1

Результат выполнения оператора кроссинговера

P_1	1	2	3	4	5	6	7	8
P_2	3	1	2	5	4	8	7	6
Случайное число	0.25	0.5	0.7	0.85	0.3	0.9	0.1	0.35
Вероятность	< 0.6	< 0.6	> 0.6	> 0.6	< 0.6	> 0.6	< 0.6	< 0.6
Повтор	-	-	+	-	+	-	-	+
P_3	1	2	3	5	4	8	7	6

Мутация. В качестве оператора мутации был выбран оператор мутации на основе дихотомии [21]. Его нечеткий алгоритм имеет вид:

1. Задать родительскую хромосому длины L .
2. Разделить хромосому пополам (при нечетном размере в любую часть берется большее число).
3. Определить точку мутации метода дихотомии через точку разрыва.
4. По правилам построения одноточечного оператора мутации получить новую хромосому-потомок.
5. Каждую половину хромосомы-потомка снова разделить на две части и процесс расчета продолжить по исходной схеме до тех пор, пока не будет получено заданное число хромосом-потомков.
6. Конец работы алгоритма.

Приведем пример реализации оператора мутации дихотомии. Пусть задана родительская хромосома, представляющая собой последовательность работ $P1 = \{1, 2, 3, 4, 5, 6, 7, 8\}$, в соответствии с рис. 1. Необходимо выполнить три итерации алгоритма и получить три хромосомы-потомка: $P2, P3, P4$. Согласно алгоритму, получим:

$$\begin{array}{l}
 P1: 1\ 2\ 3\ 4\ |\ 5\ 6\ 7\ 8 \\
 \hline
 P2: 1\ 2\ 3\ 5\ |\ 4\ 6\ 7\ 8' \\
 P2: 1\ 2\ |\ 3\ 5\ 4\ 6\ |\ 7\ 8 \\
 \hline
 P3: 1\ 3\ |\ 2\ 5\ 4\ 7\ |\ 6\ 8' \\
 P3: 1\ |\ 3\ 2\ 5\ 4\ 7\ 6\ |\ 8 \\
 \hline
 P4: 3\ |\ 1\ 2\ 5\ 4\ 7\ 8\ |\ 6'
 \end{array}$$

Применение данного оператора позволяет минимизировать количество потомков, не удовлетворяющих функции пригодности.

Транслокация. На последнем этапе генетического алгоритма выполняется оператор транслокации [22]. Данный оператор представляет собой языковую конструкцию, позволяющую на основе скрещивания и инвертирования из пары родительских хромосом создавать две хромосомы потомков. Алгоритм выполнения оператора транслокации имеет следующий вид:

1. Из заданной популяции случайным образом выбрать две альтернативных хромосомы;
2. Случайным образом выбрать точку разрыва в заданных хромосомах.
3. Выполнить скрещивание для получения хромосом-потомков.
4. В хромосомах-потомках выполнить инверсию правой части до точки разрыва;
5. Конец работы алгоритма.

Рассмотрим пример выполнения оператора транслокации. Пусть для транслокации задана последовательность работ $P_1 = \{1, 2, 3, 4, 5, 6, 7, 8\}$ и $P_2 = \{3, 1, 2, 5, 4, 8, 6, 7\}$. Необходимо получить две хромосомы-потомка P'_1 и P'_2 на основе транслокации.

$$\begin{array}{l}
 P_1: 1\ 2\ 3\ 4\ 5\ |\ 6\ 7\ 8 \\
 P_2: 3\ 1\ 2\ 5\ 4\ |\ 8\ 6\ 7 \\
 \hline
 P'_1: 1\ 2\ 3\ 4\ 5\ 7\ 6\ 8 \\
 P'_2: 3\ 1\ 2\ 5\ 4\ 8\ 7\ 6
 \end{array}$$

Использование приведенных выше модифицированных генетических операторов позволит обеспечить высокую выживаемость лучших альтернативных решений и получать близкие к оптимальным результаты за отведенное время.

Выбор облачного сервиса. Риски использования облачных вычислений. Для предприятий, выполняющих планирование своих проектов при помощи облачных вычислений важно оценить риски использования различных облачных сервисов.

С точки зрения клиента облачных сервисов принято выделять следующие типы рисков:

- ♦ риски, связанные с неправомерным доступом к данным;
- ♦ риск, связанный с потерей данных;
- ♦ риски, связанные с доступностью данных.

Таким образом, поставщик облачных услуг должен обеспечивать следующие условия для минимизации рисков:

- ♦ Конфиденциальность и безопасность данных. Поставщик облачных услуг должен обеспечивать шифрование данных, а также автоматически создавать резервные копии данных.
- ♦ Управление доступом. Поставщик облачных услуг должен предоставлять возможность аутентификации пользователей и управления доступом.
- ♦ Планирование ресурсов. Поставщик облачных услуг должен обеспечивать контроль над планированием вычислительных ресурсов.

Вероятность возникновения рисков в облачной среде зачастую связана с определенными событиями, вызывающими изменения в работе облачной среды, и определяется следующей формулой [23]:

$$(r|E) = \frac{P(E|r)}{P(E)} * P(r), \quad (6)$$

где E – событие, $P(E)$ – вероятность наступления события, $P(r)$ – вероятность возникновения риска r до события E , $P(r|E)$ – вероятность возникновения риска r во время события E , $P(E|r)$ – вероятность того, что событие E приведет к возникновению риска r .

В качестве поставщика облачных услуг был выбран *Amazon Web Services* (AWS), так как AWS имеет стратегический бизнес-план, который включает идентификацию рисков и внедрение мер контроля для снижения или управления рисками.

Amazon Web Services обеспечивает защиту глобальной инфраструктуры, которая управляет всеми услугами, предлагаемых в облаке AWS. Эта инфраструктура включает в себя оборудование, программное обеспечение, сети и средства, работающие с сервисами AWS.

Результаты вычислений. Генетический алгоритм был разработан с использованием языка программирования C++. Для эксперимента были арендованы четыре виртуальные машины *t3.2xlarge* (CPU: 8 и RAM: 32 Gb). Набор данных для апробации генетического алгоритма был взят из открытой библиотеки *PSPLIB datasets*. Для задачи RCPSP в данной библиотеке содержатся примеры разной степени сложности как для точных методов, так и для приближенных методов [24]. В библиотеке имеется по 480 примеров для количества работ $n = 30, 60, 90$. При апробации генетического алгоритма были использованы наборы данных *j301_1 – j301_10*, содержащие по 30 работ, *j601_1 – j601_10*, содержащие по 60 работ, *j901_1 – j901_10* и *j1201_1 – j1201_10*, содержащие по 90 работ. Результаты вычислений и сравнение их с лучшими значениями, по версии PSPLIB, представлены на рис. 6–9. На данных графиках по горизонтали указаны номера проектов, по вертикали – время выполнения проекта (меньше – лучше).

В табл. 2 представлено сравнение времени выполнения работы в секундах генетического алгоритма, запущенного в облаке и генетического алгоритма, запущенного на одной рабочей станции с процессором AMD Ryzen 7 2700x 3.7 ГГц и 16 Гб оперативной памяти.

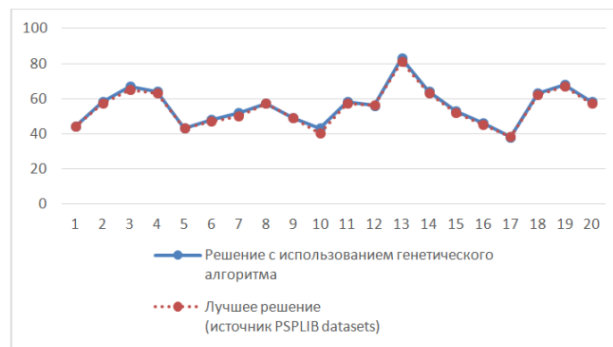


Рис. 6. Результаты вычислений при количестве работ $n = 30$

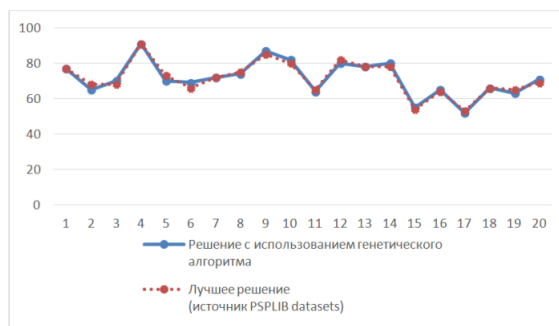


Рис. 7. Результаты вычислений при количестве работ $n = 60$

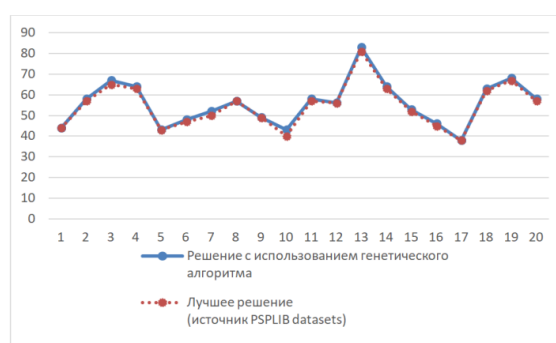


Рис. 8. Результаты вычислений при количестве работ $n = 90$

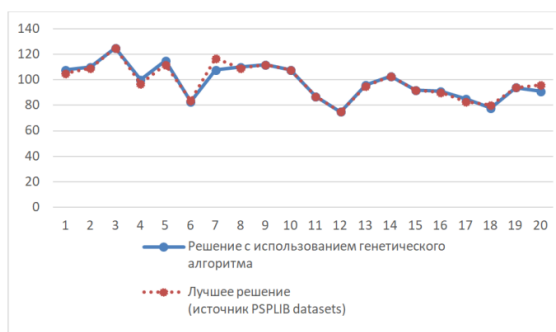


Рис. 9. Результаты вычислений при количестве работ $n = 120$

Таблица 2

Сравнение времени выполнения алгоритмов в секундах

Количество работ	МГА на основе облачных вычислений	МГА
30	0.65	1.85
60	1.1	4.4
90	1.45	7.92
120	2.1	14.2

Из анализа таблиц и графиков видно, что использование предложенного модифицированного генетического алгоритма при планировании расписания для проекта может гарантировать получение близких к оптимальным результатов за меньшее время чем модифицированный генетический алгоритм, не использующий облачных вычислений. Однако, использование предложенного алгоритма имеет смысл при общем количестве работ в проекте $n > 30$, так как время выполнения алгоритма при таком количестве работ существенно не отличается от времени выполнения модифицированного генетического алгоритма, выполненного на одной рабочей станции.

Заключение. В работе был представлен модифицированный генетический алгоритм для решения актуальной задачи построения расписания проекта с учетом ограниченности ресурсов.

Принципиальным отличием предложенного генетического алгоритма от уже имеющихся решений является использования модифицированных генетических операторов, а также использование облачных вычислений при планировании работ проектов. Данными факторами также характеризуется научная новизна исследования.

Применение модифицированных генетических операторов дает возможность оптимизировать генетический алгоритм по времени. Использование данных операторов позволяет получать близкие к оптимальным результаты за отведенное время.

Применение облачных вычислений позволяет выполнять планирование проектов с большим объемом необходимых для выполнения работ со значительно меньшей вычислительной стоимостью, в сравнении с алгоритмами, выполненными на одной рабочей станции.

Эффективность алгоритма была подтверждена экспериментом, в результате которого было установлено, что результаты вычислений, полученных с использованием предложенного генетического близки к лучшим результатам опубликованных на портале *PSPLIB datasets*, результаты вычисления некоторого количества проектов превосходят результаты, опубликованные на данном портале, а время выполнения алгоритма значительно превосходит генетический алгоритм, реализованный без применения облачных вычислений. При планировании проектов с количеством работ $n = 30$, $n = 60$, $n = 90$, $n = 120$ время выполнения предложенного алгоритма было меньше, чем время выполнения стандартного генетического алгоритма [25] в 2.8, в 4, в 5.5 и 6.8 раз соответственно. Таким образом, предложенный алгоритм может быть использован при планировании реальных проектов с большим количеством работ.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Abdolshah M.* A Review of Resource-Constrained Project Scheduling Problems (RCPSP) Approaches and Solutions/ *M. Abdolshah* // International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies. – 2014. – No. 1. – P. 253-286.
2. *Shirzadeh Chaleshtarti A., Shadrokh S., & Fathi Y.* Branch and bound algorithms for resource constrained project scheduling problem subject to nonrenewable resources with prescheduled procurement // Mathematical Problems in Engineering. – 2014.
3. *Azizoglu M., Çetinkaya F.C., & Pamir S.K.* LP relaxation-based solution algorithms for the multi-mode project scheduling with a non-renewable resource // European Journal of Industrial Engineering. – 2015. – No. 9 (4). – P. 450-469.
4. *Daniel Morillo-Torres, Luis Fernando Moreno-Velásquez & Francisco Javier Díaz-Serna.* A branch and bound hybrid algorithm with four deterministic heuristics for the resource constrained project scheduling problem (RCPSP). – 2014.
5. *Zhenyuan L., & Hongwei W.* Heuristic algorithm for RCPSP with the objective of minimizing activities' cost // Journal of Systems Engineering and Electronics. – 2006. – Vol. 17 (1). – P. 96-102.
6. *Zhu X., Ruiz R., Li S., & Li X.* An effective heuristic for project scheduling with resource availability cost // European Journal of Operational Research. – 2017. – Vol. 257 (3). – P. 746-762.

7. *Gutierrez-Franco E.* A Genetic Algorithm for the Resource Constrained Project Scheduling Problem (RCPSP) // *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies*. – 2014. – No. 1. – P. 13-22.
8. *Yuan X., Liu J., & Wimmers M.O.* A multi-agent genetic algorithm with variable neighborhood search for resource investment project scheduling problems // In 2015 IEEE Congress on Evolutionary Computation (CEC). – 2015, May. – P. 23-30. IEEE.
9. *Vanucci S.C., Carrano E.G., Bicalho R., & Takahashi R.H.* A modified NSGA-II for the Multiobjective Multi-mode Resource-Constrained Project Scheduling Problem // In 2012 IEEE Congress on Evolutionary Computation. – 2012, June. – P. 1-7. IEEE.
10. *Vartouni A.M., & Khanli L.M.* A hybrid genetic algorithm and fuzzy set applied to multimode resource-constrained project scheduling problem // *Journal of Intelligent & Fuzzy Systems*. – 2014. – Vol. 26 (3). – P. 1103-1112.
11. *Xiao J., Wu Z., Hong X.X., Tang J.C., & Tang Y.* Integration of electromagnetism with multi-objective evolutionary algorithms for RCPSP // *European Journal of Operational Research*. – 2016. – Vol. 251 (1). – P. 22-35.
12. *Sebt M.H., Afshar M.R., & Alipouri Y.* Hybridization of genetic algorithm and fully informed particle swarm for solving the multi-mode resource-constrained project scheduling problem // *Engineering Optimization*. – 2017. – Vol. 49 (3). – P. 513-530.
13. *Gomes H.C., das Neves F.D.A., & Souza M.J.F.* Multi-objective metaheuristic algorithms for the resource-constrained project scheduling problem with precedence relations // *Computers & Operations Research*. – 2014. – Vol. 44. – P. 92-104.
14. *Kadam S.U., & Kadam N.S.* Solving resource-constrained project scheduling problem by genetic algorithm // In *Business and Information Management (ICBIM), 2014 2nd International Conference on*. – 2014, January. – P. 159-164. IEEE.
15. *Лазарев А.А., Гафаров Е.Р.* Теория расписаний. Задачи и алгоритмы: учеб. пособие. – М.: Московский государственный университет им. М.В. Ломоносова, 2011. – 222 с.
16. *Вильчинский Н.* От хранения данных к управлению информацией. – Питер, 2016. – 544 с.
17. *Гладков Л.А., Зинченко Л.А., Курейчик В.В. и др.* Методы генетического поиска. – Таганрог: Изд-во ТРТУ, 2002.
18. *Курейчик В.В., Сороколетов П.В.* Концептуальная модель представления решений в генетических алгоритмах // *Известия ЮФУ. Технические науки*. – 2008. – № 9 (86). – С. 7-12.
19. *Шевляков А.О., Матвеев М.Г.* Решение RCPSP при нечетких трудозатратах выполнения операций // *Вестник ВГУ*. – 2015. – № 4. – С. 119-123.
20. *Gonçalves José & Resende Mauricio.* A multi-population genetic algorithm for a constrained two-dimensional orthogonal packing problem // *J. Comb. Optim.* – 2011. – Vol. 22. – P. 180-201. – DOI 10.1007/s10878-009-9282-1.
21. *Гладков Л.А., Курейчик В.В., Курейчик В.М.* Генетические алгоритмы: учеб, пособие. – 2-е изд. – М.: Физматлит. 2006. – 320 с.
22. *Kureichik V.V., Kureichik V.M.* Genetic search-based control // *Automation and Remote Control*. – 2001. – Vol. 62, No. 10. – P. 1698-1710.
23. *Samad J., Loke S.W. and Reed K.* Quantitative Risk Analysis for Mobile Cloud Computing: A Preliminary Approach and a Health Application Case Study // 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, Melbourne, VIC, 2013. – P. 1378-1385. – Doi: 10.1109/TrustCom.2013.166.
24. *Habibi, Farhad & Barzinpour, Farnaz & Sadjadi, Seyed.* Resource-constrained project scheduling problem: review of past and recent developments // *Journal of Project Management*. – 2018. – No. 3. – P. 55-88. 10.5267/j.jpm.2018.1.005.
25. *Могилев А.А.* Обзор методов решения задач теории расписаний // *Информатика, вычислительная техника и инженерное образование*. – 2019. – Т. 37, № 2.

REFERENCES

1. *Abdolshah M.* A Review of Resource-Constrained Project Scheduling Problems (RCPSP) Approaches and Solutions/ M. Abdolshah, *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies*, 2014, No. 1, pp. 253-286.
2. *Shirzadeh Chaleshtarti A., Shadrokh S., & Fathi Y.* Branch and bound algorithms for resource constrained project scheduling problem subject to nonrenewable resources with prescheduled procurement, *Mathematical Problems in Engineering*, 2014.

3. Azizoglu M., Çetinkaya F.C., & Pamir S.K. LP relaxation-based solution algorithms for the multi-mode project scheduling with a non-renewable resource, *European Journal of Industrial Engineering*, 2015, No. 9 (4), pp. 450-469.
4. Daniel Morillo-Torres, Luis Fernando Moreno-Velásquez & Francisco Javier Díaz-Serna. A branch and bound hybrid algorithm with four deterministic heuristics for the resource constrained project scheduling problem (RCPSP). – 2014.
5. Zhenyuan L., & Hongwei W. Heuristic algorithm for RCPSP with the objective of minimizing activities' cost, *Journal of Systems Engineering and Electronics*, 2006, Vol. 17 (1), pp. 96-102.
6. Zhu X., Ruiz R., Li S., & Li X. An effective heuristic for project scheduling with resource availability cost, *European Journal of Operational Research*, 2017, Vol. 257 (3), pp. 746-762.
7. Gutierrez-Franco E. A Genetic Algorithm for the Resource Constrained Project Scheduling Problem (RCPSP), *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies*, 2014, No. 1, pp. 13-22.
8. Yuan X., Liu J., & Wimmers M.O. A multi-agent genetic algorithm with variable neighborhood search for resource investment project scheduling problems, *In 2015 IEEE Congress on Evolutionary Computation (CEC)*, 2015, May, pp. 23-30. IEEE.
9. Vanucci S.C., Carrano E.G., Bicalho R., & Takahashi R.H. A modified NSGA-II for the Multiobjective Multi-mode Resource-Constrained Project Scheduling Problem, *In 2012 IEEE Congress on Evolutionary Computation*, 2012, June, pp. 1-7. IEEE.
10. Vartouni A.M., & Khanli L.M. A hybrid genetic algorithm and fuzzy set applied to multimode resource-constrained project scheduling problem, *Journal of Intelligent & Fuzzy Systems*, 2014, Vol. 26 (3), pp. 1103-1112.
11. Xiao J., Wu Z., Hong X.X., Tang J.C., & Tang Y. Integration of electromagnetism with multi-objective evolutionary algorithms for RCPSP, *European Journal of Operational Research*, 2016, Vol. 251 (1), pp. 22-35.
12. Sebt M.H., Afshar M.R., & Alipouri Y. Hybridization of genetic algorithm and fully informed particle swarm for solving the multi-mode resource-constrained project scheduling problem, *Engineering Optimization*, 2017, Vol. 49 (3), pp. 513-530.
13. Gomes H.C., das Neves F.D.A., & Souza M.J.F. Multi-objective metaheuristic algorithms for the resource-constrained project scheduling problem with precedence relations, *Computers & Operations Research*, 2014, Vol. 44, pp. 92-104.
14. Kadam S.U., & Kadam N.S. Solving resource-constrained project scheduling problem by genetic algorithm, *In Business and Information Management (ICBIM), 2014 2nd International Conference on*, 2014, January, pp. 159-164. IEEE.
15. Lazarev A.A., Gafarov E.R. Teoriya raspisaniy. Zadachi i algoritmy: ucheb. posobie [he scheduling theory. Problems and algorithms: textbook]. Moscow: Moskovskiy gosudarstvennyy universitet im. M.V. Lomonosova, 2011, 222 p.
16. Vil'chinskiy N. Ot khraneniya dannykh k upravleniyu informatsiy [From data storage to information management]. Piter, 2016, 544 p.
17. Gladkov L.A., Zinchenko L.A., Kureychik V.V. i dr. Metody geneticheskogo poiska [Genetic search methods]. Taganrog: Izd-vo TRTU, 2002.
18. Kureychik V.V., Sorokoletov P.V. Kontseptual'naya model' predstavleniya resheniy v geneticheskikh algoritмах [A conceptual model for representing solutions in genetic algorithms], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2008, No. 9 (86), pp. 7-12.
19. Shevlyakov A.O., Matveev M.G. Reshenie RCPSP pri nechetkikh trudozatratakh vypolneniya operatsiy [RcpSP solution for fuzzy labor costs of performing operations], *Vestnik VGU* [Vestnik VSU], 2015, No. 4, pp. 119-123.
20. Gonçalves José & Resende Mauricio. A multi-population genetic algorithm for a constrained two-dimensional orthogonal packing problem, *J. Comb. Optim.*, 2011, Vol. 22, pp. 180-201. Doi 10.1007/s10878-009-9282-1.
21. Gladkov L.A., Kureychik V.V., Kureychik V.M. Geneticheskie algoritmy: ucheb. posobie [Genetic algorithms: a textbook]. 2nd ed. Moscow: Fizmatlit. 2006, 320 p.
22. Kureichik V.V., Kureichik V.M. Genetic search-based control, *Automation and Remote Control*, 2001, Vol. 62, No. 10, pp. 1698-1710.
23. Samad J., Loke S.W. and Reed K. Quantitative Risk Analysis for Mobile Cloud Computing: A Preliminary Approach and a Health Application Case Study, *12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, Melbourne, VIC, 2013*, pp. 1378-1385. Doi: 10.1109/TrustCom.2013.166.

24. Habibi, Farhad & Barzinpour, Farnaz & Sadjadi, Seyed. Resource-constrained project scheduling problem: review of past and recent developments, *Journal of Project Management*, 2018, No. 3, pp. 55-88. 10.5267/j.jpjpm.2018.1.005.
25. Mogilev A.A. Obzor metodov resheniya zadach teorii raspisaniy [Review of methods for solving problems in the theory of schedules], *Informatika, vychislitel'naya tekhnika i inzhenernoe obrazovanie* [Informatics, computer engineering and engineering education], 2019, Vol. 37, No. 2.

Статью рекомендовал к опубликованию д.т.н., профессор Н.И. Витиска.

Могилев Алексей Алексеевич – Южный федеральный университет; e-mail: alexmogilev08@mail.ru; 347928, г. Таганрог, Некрасовский, 44; тел.: 89897258619; кафедра систем автоматизированного проектирования; аспирант.

Курейчик Виктор Михайлович – e-mail: vmkureychik@sfedu.ru; тел.: 88634311487; кафедра системы автоматизированного проектирования; д.т.н.; профессор.

Mogilev Alexey Alexeevich – Southern Federal University; e-mail: alexmogilev08@mail.ru; 44, Nekrasovskiy, Taganrog, 347928, Russia; phone: +79897258619; the department of computer aided design; postgraduate student.

Kureichik Viktor Mikhaylovich – e-mail: vmkureychik@sfedu.ru; phone: +78634311487; the department of computer aided design; dr. of eng. sc.; professor.

УДК 004.896

DOI 10.18522/2311-3103-2020-2-169-179

Э.В. Кулнев, А.А. Лежебоков, Семенова М.М., В.А. Семенов

ПОДХОД К КОДИРОВАНИЮ РЕШЕНИЙ В ЭВОЛЮЦИОННЫХ МЕТОДАХ ДЛЯ СОЗДАНИЯ ИНСТРУМЕНТАЛЬНОЙ ПЛАТФОРМЫ ПРОЕКТИРОВАНИЯ*

Рассмотрены актуальные вопросы и проведен анализ проблемы трехмерной интеграции и трехмерного моделирования, возникающей на этапе конструкторского проектирования в ходе решения задачи оптимального планирования компонентов больших и сверхбольших интегральных схем и корпусных устройств электронной вычислительной аппаратуры. Представлены и достаточно детально описаны основные преимущества применения принципов трехмерной интеграции, позволяющие эффективно организовывать производство персонализированной электроники, оптимально планировать конфигурацию больших и сверхбольших интегральных схем с учетом тепловых и энергетических характеристик. В ходе выполнения исследований авторами разработан подход к кодированию решений на основе интеллектуального механизма, который характеризуется наличием встроенных средств контроля допустимых решений. Одним из таких средств, экспериментально доказавших свою эффективность, является встроенный механизм «смертельных мутаций», учитывающий статусы генов и заранее заданные ограничения на итоговую конфигурацию корпуса проектируемого устройства. В работе предложен ряд общих подходов и конкретных алгоритмов решения задачи планирования, основывающихся на результатах исследований авторского коллектива и современных подходах к решению NP-полных задач. Важнейшим практически значимым результатом исследований обозначенной проблемы является разработанная программно-инструментальная платформа проектирования на современном кроссплатформенном языке программирования Java. Выбранная технология разработки позволяет использовать все основные достоинства современных многоядерных и многопроцессорных архитектур, по использованию программной многопоточности для реализации параллельных схем решения комбинаторных задач. Программно-инструментальная платформа обладает дружественным интерфейсом, что позволяет эффективно управлять процессом решения задачи планирования компонентов

* Работа выполнена при финансовой поддержке РФФИ, проекты № 18-01-00041 и № 18-07-00055.

больших и сверхбольших интегральных схем трехмерной интеграции, путем визуализации ключевых показателей работы алгоритмов на графиках и в блоках текстовой статистики. Разработанное прикладное программное обеспечение позволило провести серию вычислительных экспериментов, на основе наборов случайных данных также, как и наборах открытых данных бенчмарков для подобного рода задач. Результаты экспериментальных исследований позволили подтвердить теоретические оценки временной сложности и эффективности предложенных подходов и алгоритмов, в том числе генетического алгоритма, который использует предложенный в работе новый механизм кодирования решений.

Трехмерное моделирование; трехмерная интеграция; размещение; БИС; СБИС; генетический алгоритм; эволюция; биоинспирированный алгоритм.

E.V. Kuliev, A.A. Lezhebokov., M.M. Semenova, V.A. Semenov

SOLUTIONS' ENCODING IN EVOLUTIONARY METHODS FOR INSTRUMENTAL DESIGN PLATFORM

The article considers current issues and analyzes the problems of three-dimensional integration and three-dimensional modeling that arise at the design stage during the solution of the problem of optimal planning of components of large and extra-large integrated circuits and case devices of electronic computing equipment. The main advantages of applying the principles of three-dimensional integration are presented and described in sufficient detail, which allow efficiently organizing the production of personalized electronics, optimally planning the configuration of large and ultra-large integrated circuits, taking into account thermal and energy characteristics. In the course of research, the authors developed an approach to encoding decisions based on an intelligent mechanism, which is characterized by the presence of built-in means of control of acceptable decisions. One of such tools that have experimentally proven their effectiveness is the built-in mechanism of "deadly mutations", which takes into account the status of genes and predetermined restrictions on the final configuration of the housing of the designed device. A series of general approaches and specific algorithms for solving the planning problem based on the results of research by the author's team and modern approaches to solving NP-complete problems are proposed. The most important practically significant result of the research of the indicated problem is the developed software and instrumental design platform in the modern cross-platform Java programming language. The selected development technology allows you to use all the main advantages of modern multi-core and multi-processor architectures, to use software multi-threading to implement parallel schemes for solving combinatorial problems. The software and tool platform has a user-friendly interface, which allows you to effectively manage the process of solving the problem of planning the components of large and ultra-large integrated circuits of three-dimensional integration by visualizing key performance indicators of algorithms on graphs and in text statistics blocks. The developed application software made it possible to carry out a series of computational experiments based on random data sets, as well as on open-data boron benchmarks for such tasks. The results of experimental studies have confirmed the theoretical estimates of the time complexity and effectiveness of the proposed approaches and algorithms, including the genetic algorithm, which uses the new decision coding mechanism proposed in the work.

Three-dimensional modeling; three-dimensional integration; placement; LSI; VLSI; genetic algorithm; evolution; bioinspired algorithm.

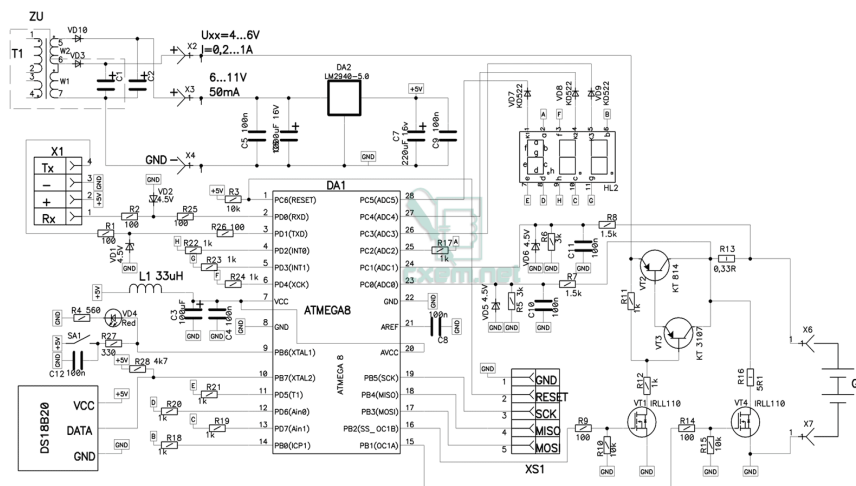
Введение. Устройства трехмерной интеграции и современные сверхбольшие интегральные схемы (СБИС) представляют собой попытку преодолеть ограниченность плоских устройств путем интеграции нескольких слоев кристаллов в единую монолитную структуру. Преимуществом такого вертикального расширения является лучшее использование площади чипа или объема корпуса устройства ЭВА, уменьшение общей длины проводников, оптимизация критических связей и увеличение количества транзисторов на единицу площади, это приводит к росту быстродействия и повышению энергоэффективности [1].

Предлагаемый подход трехмерной интеграции актуален и в случае проектирования монолитных корпусных устройств вычислительной аппаратуры. В настоящее время задача проектирования объемных устройств актуальна, так как в

процессе исторического развития компьютерной техники и информационных технологий стала востребованной задача оперирования объемными представлениями (моделями) реальных объектов в процессе автоматизированного проектирования. Появлению современных методов трехмерного моделирования объектов способствовал бурный рост производительности вычислительной техники. Большинство из существующих сегодня электронных устройств представляют собой двумерную (плоскую) печатную плату, вложенную в трёхмерный (объёмный) корпус [2–3]. В настоящее время активно развиваются технологии аддитивного производства, которые позволяют производить более сложные электронные устройства за одну технологическую операцию. Данные технологии размещают компоненты по всему внутреннему объёму корпуса и соединив их объемными цепями, что обеспечивает появление новых технологических преимуществ:

- ◆ компактность изделия – максимально эффективное использование пространства;
- ◆ сверхсложные геометрические формы – аддитивные технологии позволяют произвести изделие произвольной геометрической формы;
- ◆ возможность изготавливать невозможные по форме для классических технологий изделия;
- ◆ хорошие механические, временные и тепловые характеристики;
- ◆ производство персонализированной носимой электроники (например, наушники, оптимально подходящие под ушные раковины конкретного потребителя, удобные персонализированные роботизированные протезы конечностей, различные биологические датчики);
- ◆ снижение стоимости проектирования и изготовления – уменьшается как цикл разработки, так и длина производственных цепочек, что, в том числе, сокращает расходы на логистику;
- ◆ возможность спроектировать и изготовить изделие, не требующее дополнительной операции сборки.

1. Проблема трехмерного размещения. В качестве примера практической задачи рассмотрим задачу планирования конфигурации устройства из исходных данных, таких как принципиальная электрическая схема (рис. 1), геометрия корпуса устройства (рис. 2) и дополнительных ограничений (например, специальной геометрии для дальнейшей подготовки к печати на 3D-принтере). В качестве моделируемого устройства рассмотрим зарядное устройство-анализатор NiMh/NiCd аккумуляторов.



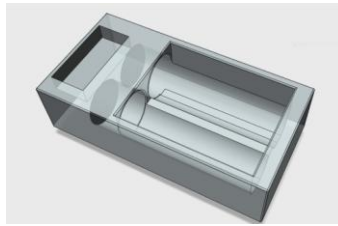


Рис. 2. Пример исходной геометрии корпуса устройства.

При малом количестве элементов схемы и отсутствии жестких требований по компактности схемы инженер-конструктор имеет возможность с помощью классических инструментов проектирования разместить элементы в корпусе и провести соединяющие трассы [4–7]. При повышении степени интеграции СБИС или увеличении количества размещаемых элементов ЭВА и их взаимосвязей – оптимальное размещение и соединение элементов в корпусе становится задачей требующей большого числа ресурсов для её решения. Известные алгоритмы, решающие NP-полные задачи, в частности, задачу размещения, имеют экспоненциальную временную сложность, что эквивалентно полному перебору всех возможных вариантов решений. Следовательно, актуальной и востребованной остается задача разработки новых методов и алгоритмов планирования и размещения элементов СБИС трехмерной интеграции на основе проверенных современных эвристик и подтвержденных гипотез об эффективности методов эволюционного моделирования и генетического поиска.

2. Механизм кодирования решений. Новизна подхода к кодированию решения в эволюционных методах (генетических, эволюционных и бионических алгоритмах) для задачи планирования конфигурации устройств трехмерной интеграции заключается в использовании модели представления объема с помощью октантного дерева (octree) [9-15]. Октантное дерево (октодерево) является структурой, которая описывает последовательное разбиение объема изображения на элементы уменьшающегося объема. Формирование октодерева удобнее всего описать в рекурсивной форме. В корневом узле октодерева находится весь объем изображения. Далее каждый узел может иметь восемь потомков, если необходимый элемент объема, или блок (прямоугольный параллелепипед), разделен на подблоки (которые, как правило, берутся равными по размеру и форме). Степенью детализации представления можно управлять, изменяя глубину рекурсии процедуры разбиения [8].

Хромосома представляет собой вектор, каждый ген кодируется объектом класса Gene, UML диаграмма которого представлена на рис. 3.

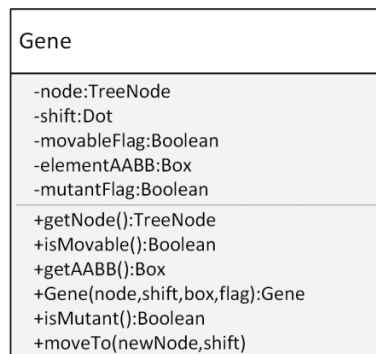


Рис. 3. Структура класса Gene (один ген хромосомы).

JSON – это специальная библиотека, которая используется для записи в файл или передачи по сети в системе. Она позволяет, реализовать интерфейс JsonSerializer<?>, а так же сериализовать любой объект в JSON (англ. JavaScript Object Notation). Эта текстовая, человеко-читаемая нотация – стандарт в кодировании и передаче объектов в сети интернет.

Рассмотрим пример, иллюстрирующий процесс кодирования и декодирования хромосомы. В качестве примера исходных данных используется геометрическая форма, представленная на рис. 4.

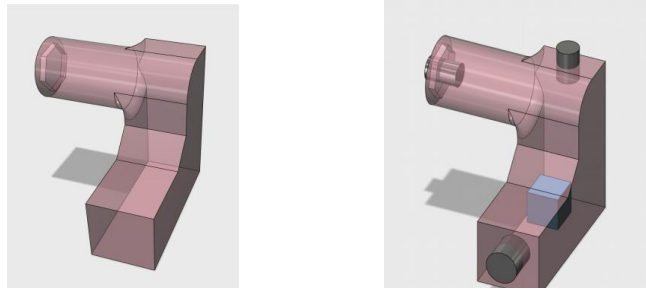


Рис. 4. Геометрия тестового корпуса и пример размещения элементов

Описание хромосомы.

Целевая функция, рассчитываемая данным алгоритмом выражается в условных единицах, но для удобства восприятия приведена к миллиметрам в кубе. Пример хромосомы может быть представлен вектором:

$$\langle A1, m1, A2, m2, \dots, An, mn \rangle,$$

где элементы A – это элементы цепи, а элементы m это модификаторы изменения направления движения трассы в пространстве. Объем всех трасс/цепей характеризует проектируемый объект по заданным характеристикам (теплопроводность и эффективность, трассируемость и т.п.)

При глубине рекурсии алгоритма построения равному пяти, дерево этого корпуса имеет 8857 узлов (или листов). В этом корпусе необходимо разместить 4 элемента. В схеме три элемента являются терминальными (или выходными), то есть это элементы, которые нельзя передвигать. На рис. 4 показано возможное расположение элементов, в виде трех цилиндров и куба. Цилиндры являются неподвижными элементами, а куб подвижным.

3. Алгоритм расчета целевой функции. На рис. 5 представлен алгоритм расчета целевой функции хромосомы. Геометрия элементов (на рис. представлена в виде куба и цилиндров), список цепей, проложенные трассы между размещенными элементами визуализация этих данных представлена на рис.6. Как было отмечено ранее, для простоты примера куб будет единственным изменяемым (перемещаемым) элементом. Поэтому в примере хромосомы будут отличаться только этим геном [9].

Описание алгоритма.

Алгоритм имеет линейную структуру с единственным циклом по числу цепей. Для каждой цепи, значение вычисленного объема кэшируется и сохраняется в специальном хранилище, которая позволяет не тратить ресурсы на перерасчет целевой функции для цепей, конфигурация которых не поменялась. В случае если конфигурация цепи меняется незначительно, можно отдельно пересчитать функцию для участка цепи, но это требуется дополнительного усложнения алгоритма,

поэтому на данном этапе исследование деление цепи на участки не предусмотрено. Алгоритм пересчета объема одной цепи учитывает каждую точку изменения направления трехмерной трассы и все необходимые точки контактных площадок.

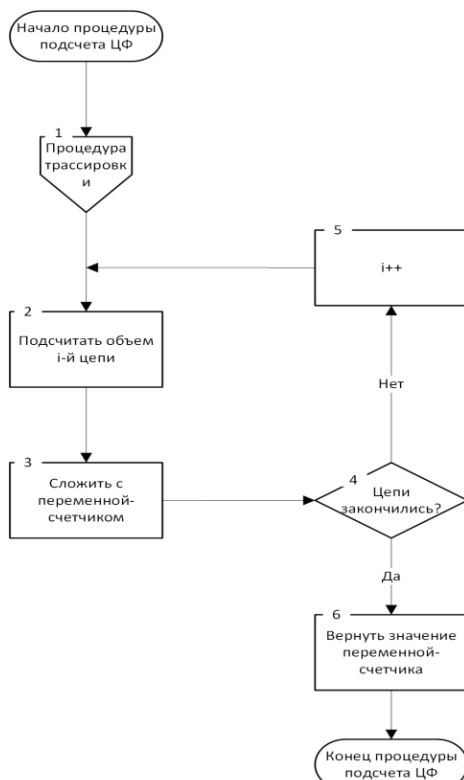


Рис. 5. Алгоритм расчета целевой функции хромосомы

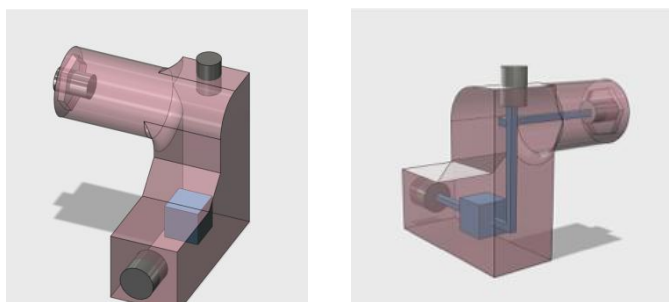


Рис. 6. Модель объема и конфигурация устройства для хромосомы 1

Вычисленная по приведенному алгоритму оценки объема первой хромосомы – 46381.194 mm^3 , второе – 54721.409 mm^3 .

В ходе выполнения генетического алгоритма и применения различных генетических операторов (кроссинговер, мутация, инверсия) используется новый механизм, при котором в алгоритм вводится понятие «смертельных мутаций». Суть этого механизма в следующем: если мутация (изменение значения) происходит в гене, который помечен специальным флагом, то все потомки этой хромосомы, несущие этот ген, считаются «мертвыми» (недопустимыми).

Модификация информационной модели гена состоит в следующем: ген несет специальные флаги – флаг, который запоминает мутировал ли ген в нынешнем поколении и флаг не изменения. Модификация оператора мутации же заключается в том, что он взводит флаг мутации гена, если она в нем происходит. Дальнейшие модификации алгоритмов связаны с реализацией различных методик использования этих флагов (указателей).

4. Результаты вычислительных экспериментов. В ходе выполнения исследований была разработана программа для ЭВМ, предназначенная для решения задачи трехмерного размещения элементов ЭВА. В качестве основного подхода к исследованиям разработанного механизма кодирования была выбрана теория эволюционных вычислений, на основе которой разработаны классический генетический и новый биоинспирированный алгоритмы [9–10]. Суть эволюционного подхода заключается в последовательном развитии (эволюции) набора решений согласно выбранной модели эволюции. В результате эволюции в популяции (наборе) решений закрепляются особи (решения) с наиболее удачными характеристиками, что позволяет находить оптимальные и квазиоптимальные решения оптимизационных задач за приемлемое время. Разработанная программа имеет встроенный конфигуратор генетического алгоритма, который позволяет гибко подстраивать алгоритм к конкретным исходным данным. Для использования полученных результатов другими программными комплексами реализованная функция экспорта результатов. Программный продукт является кроссплатформенным, так как реализован на языке высокого уровня Java. То есть он запустится на любой операционной системе (ОС), где реализована среда исполнения Java (JRE). К этим ОС относятся все наиболее популярные, такие как Windows, Linux, MAC OS X и другие менее распространенных.

Основная цель разработанного программного продукта, состоит в реализации алгоритма, который эффективно решает прикладную задачу планирования конфигурации объемных устройств. В качестве объекта исследования был использован биоинспирированный алгоритм размещения элементов схем ЭВА в трехмерном пространстве с учетом трассируемости. Сравнение разработанного алгоритма проводилось со случайным и генетическим алгоритмами, так как они являются итерационными и используют схожие вероятностные механизмы генерации новых решений. Исследуемые сравниваемые алгоритмы также имеют однотипный набор управляющих параметров, что позволяет повысить точность анализа результатов экспериментов [15–20]. Эффективность алгоритма размещения заключается в подборе параметров используемого генетического алгоритма таких как очередность и вероятности генетических операторов. Анализ эффективности разработанного алгоритма размещения состоит в определении временной сложности алгоритма (BCA).

Для проведения экспериментальных исследований, авторами были выбраны схемы с различным набором выходных данных. Экспериментальные исследования проводились над сверхбольшими интегральными схемами с различным количеством размещаемых элементов, от 10000 до 100000. В данной работе временная сложность рассчитывается исходя из зависимости времени работы алгоритмов от количества размещаемых ими элементов [20]. Было проведено сравнение производительности предложенного алгоритма с несколькими другими методами оптимизации на основе популяции, которые представляют собой генетический, муравьиный алгоритмы (ГА, МА).

Полученные в ходе проведенных экспериментов данные представлены в табл. 1 и на рис. 7.

Таблица 1

Оценка временной сложности алгоритмов

Количество элементов	Алгоритм		
	Генетический (мин)	Случайный (мин)	Биоинспирированный (мин)
10000	1,07	2,22	1,87
20000	2,61	3,89	3,77
30000	3,81	5,28	5,18
40000	5,04	6,11	6,01
50000	7,18	8,51	8,17
60000	9,2	10,4	9,7
70000	12,6	13,5	12,9
80000	16,5	18,0	17,77
90000	18,7	20,2	19,8
100000	21,2	23,1	21,9

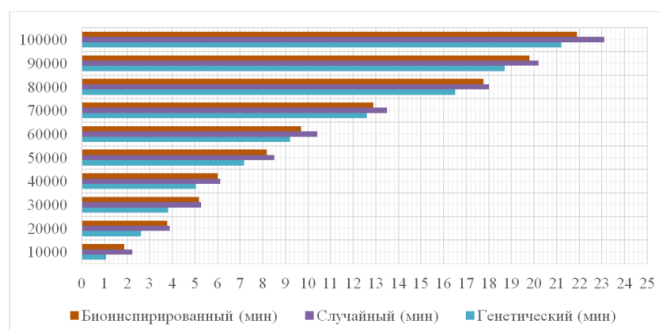


Рис. 7. График зависимости времени работы от количества элементов

В результате, наиболее эффективным является разработанный алгоритм размещения элементов в трехмерном пространстве с учетом трассируемой. Он в среднем на 7–9 % эффективнее случайного и на 9–13 % эффективнее генетического алгоритма. Временная сложность разработанного алгоритма равна $O(n^2)$.

Заключение. В результате проведенного исследования был выполнен анализ проблемы трехмерной интеграции и трехмерного моделирования для задачи оптимального планирования компонентов СБИС и корпусных устройств ЭВА. Установлены основные преимущества применения трехмерной интеграции, позволяющие эффективно организовывать производство персонализированной электроники, оптимально планировать конфигурацию СБИС с учетом тепловых и энергетических характеристик. Разработанный механизм кодирования решений отличается наличием встроенных средств контроля допустимых решений, путем использования механизма «смертельных мутаций», учитывающего статусы генов и заданную конфигурацию корпуса проектируемого устройства.

Предложенный подход и алгоритм были реализованы в виде программы на языке Java, которая является кроссплатформенной, обладает дружелюбным интерфейсом и позволяет управлять процессом решения задачи планирования компонентов СБИС трехмерной интеграции. Разработанное программное обеспечение позволило провести вычислительный эксперимент, подтвердивший теоретические оценки временной сложности и эффективности генетического алгоритма с предложенным механизмом кодирования решений. Количественно эффективность

предложенного подхода к кодированию решений можно определить на уровне 7–9 % при использовании данной схемы в биоинспирированном алгоритме по сравнению с классическим генетическим алгоритмом. Данные показатели могут быть дополнительно улучшены за счет применения схем распараллеливания вычислений на современных многоядерных и многопроцессорных системах.

Дальнейшие исследования будут связаны с развитием теории биоинспирированного поиска, направлены на разработку новых методов и алгоритмов решения оптимизационных задач автоматизированного проектирования СБИС трехмерной интеграции с учетом предложенного механизма кодирования решений. Гипотеза дальнейших исследований базируется на предположении о том, что интеграция нового механизма кодирования с принципами теории принятия решений и методами искусственного интеллекта позволит повысить эффективность процесса проектирования сложного оборудования.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Норенков И.П.* Основы автоматизированного проектирования. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2010. – 364 с.
2. *Карпенко А.П.* Современные алгоритмы поисковой оптимизации // Алгоритмы, вдохновленные природой: учеб. пособие. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2014. – 474 с.
3. *Курейчик В.В., Курейчик В.М., Родзин С.И.* Теория эволюционных вычислений. – М.: Физматлит, 2012. – 260 с.
4. *Курейчик В.В., Запорожец Д.Ю.* Современные проблемы при размещении элементов СБИС // Известия ЮФУ. Технические науки. – 2011. – № 7 (120). – С. 68-73.
5. *Лежебоков А.А., Кравченко Ю.А., Пащенко С.В.* Особенности использования технологии дополненной реальности для поддержки образовательных процессов // Открытое образование. – 2014. – № 3 (104). – С. 49-54.
6. *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.
7. *Курейчик В.М., Курейчик В.В.* Генетический алгоритм разбиения графа // Известия Российской академии наук. Теория и системы управления. – 1999. – № 4. – С. 79-87.
8. *Кулиев Э.В., Кравченко Ю.А., Логинов О.А., Запорожец Д.Ю.* Метод интеллектуального принятия эффективных решений на основе биоинспирированного подхода // Известия Кабардино-Балкарского научного центра РАН. – 2017. – № 6-2 (80). – С. 162-169.
9. *Кулиев Э.В., Лежебоков А.А., Кравченко Ю.А.* Роевой алгоритм поисковой оптимизации на основе моделирования поведения летучих мышей // Известия ЮФУ. Технические науки. – 2016. – № 7 (180). – С. 53-62.
10. *Кулиев Э.В., Лежебоков А.А.* О гибридном алгоритме размещения компонентов СБИС // Известия ЮФУ. Технические науки. – 2012. – № 11 (136). – С. 188-192.
11. *Bova V.V., Lezhebokov A.A., Gladkov L.A.* Problem-oriented algorithms of solutions search based on the methods of swarm intelligence // World Applied Sciences Journal. – 2013. – Vol. 27 (9). – P. 1201-1205.
12. *Zaruba D., Zaporozhets D., Kureichik V.* VLSI placement problem based on ant colony optimization algorithm // Advances in Intelligent Systems and Computing. – 2016. – Vol. 464. – P. 127-133.
13. *Kureichik V., Kureichik V., Bova V.* Placement of VLSI fragments based on a multilayered approach // Advances in Intelligent Systems and Computing. – 2016. – Vol. 464. – P. 181-190.
14. *Kureichik V.V., Zaruba D.V.* The bioinspired algorithm of electronic computing equipment schemes elements placement // Advances in Intelligent Systems and Computing. – 2015. – Vol. 347. – P. 51-58.
15. *Гладков Л.А., Гладкова Н.В., Гордиенко В.Н.* Модифицированный генетический алгоритм решения задачи компоновки блоков ЭВА // Информатика, вычислительная техника и инженерное образование. – 2015. – № 4 (24). – С. 18-27.
16. *Курейчик В.В., Курейчик Вл.Вл.* Биоинспирированный алгоритм разбиения схем при проектировании СБИС // Известия ЮФУ. Технические науки. – 2013. – № 7 (144). – С. 23-29.

17. Курейчик В.В., Курейчик В.М., Гладков Л.А., Сороколетов П.В. Биоинспирированные методы в оптимизации. – М.: Физмалит, 2009. – 384 с.
18. Курейчик В.В., Заруба Д.В., Запорожец Д.Ю. Биоинспирированный алгоритм компоновки блоков ЭВА на основе модифицированной раскраски графа // Известия ЮФУ. Технические науки. – 2015. – № 4 (165). – С. 6-14.
19. Курейчик В.В., Курейчик В.В., Бова В.В. Биоинспирированный поиск в задачах конструкторского проектирования и оптимизации // Информационные технологии в науке, образовании и управлении / под ред. проф. Е.Л. Глорнозова. – 2015. – С. 427-432.
20. Кулиев Э.В., Лежебоков А.А., Дуккардт А.Н. Подход к исследованию окрестностей в роевых алгоритмах для решения оптимизационных задач // Известия ЮФУ. Технические науки. – 2014. – № 7 (156). – С. 15-25.
21. Кулиев Э.В., Лежебоков А.А. Исследование характеристик гибридного алгоритма размещения // Известия ЮФУ. Технические науки. – 2013. – № 3 (140). – С. 255-261.

REFERENCES

1. Norenkov I.P. Osnovy avtomatizirovannogo proektirovaniya [Fundamentals of computer-aided design]. Moscow: Izd-vo MGTU im. N.E. Bauman, 2010, 364 p.
2. Karpenko A.P. Sovremennyye algoritmy poiskovoy optimizatsii [Modern search engine optimization algorithms], *Algoritmy, vdokhnovlennyye prirodoy: ucheb. posobie* [Algorithms inspired by nature: a textbook]. Moscow: Izd-vo MGTU im. N.E. Bauman, 2014, 474 p.
3. Kureychik V.V., Kureychik V.M., Rodzin S.I. Teoriya evolyutsionnykh vychisleniy [Theory of evolutionary computation]. Moscow: Fizmatlit, 2012, 260 p.
4. Kureychik V.V., Zaporozhets D.Yu. Sovremennyye problemy pri razmeshchenii elementov SBIS [Current problems in placing VLSI elements], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2011, No. 7 (120), pp. 68-73.
5. Lezhebokov A.A., Kravchenko Yu.A., Pashchenko S.V. Osobennosti ispol'zovaniya tekhnologii dopolnennoy real'nosti dlya podderzhki obrazovatel'nykh protsessov [Features of using augmented reality technology to support educational processes], *Otkrytoe obrazovanie* [Open education], 2014, No. 3 (104), pp. 49-54.
6. 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 tsentra RAN* [Izvestiya Kabardino-Balkar scientific center of the Russian Academy of Sciences], 2019, No. 4 (90), pp. 14-23.
7. Kureychik V.M., Kureychik V.V. Geneticheskyy algoritm razbieniya grafa [Genetic algorithm for graph splitting], *Izvestiya Rossiyskoy akademii nauk. Teoriya i sistemy upravleniya* [Proceedings of the Russian Academy of Sciences. Theory and control systems], 1999, No. 4, pp. 79-87.
8. Kuliev E.V., Kravchenko Yu.A., Loginov O.A., Zaporozhets D.Yu. Metod intellektual'nogo prinyatiya effektivnykh resheniy na osnove bioinspirirovannogo podkhoda [Method of intelligent decision-making based on a bioinspired approach], *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN* [Izvestiya Kabardino-Balkar scientific center of the Russian Academy of Sciences], 2017, No. 6-2 (80), pp. 162-169.
9. Kuliev E.V., Lezhebokov A.A., Kravchenko Yu.A. Roevoy algoritm poiskovoy optimizatsii na osnove modelirovaniya povedeniya letuchikh myshey [Swarm algorithm search engine optimization is based on modeling the behavior of bats], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2016, No. 7 (180), pp. 53-62.
10. Kuliev E.V., Lezhebokov A.A. O gibridnom algoritme razmeshcheniya komponentov SBIS [About a hybrid VLSI component placement algorithm], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2012, No. 11 (136), pp. 188-192.
11. Bova V.V., Lezhebokov A.A., Gladkov L.A. Problem-oriented algorithms of solutions search based on the methods of swarm intelligence, *World Applied Sciences Journal*, 2013, Vol. 27 (9), pp. 1201-1205.
12. Zaruba D., Zaporozhets D., Kureichik V. VLSI placement problem based on ant colony optimization algorithm, *Advances in Intelligent Systems and Computing*, 2016, Vol. 464, pp. 127-133.
13. Kureichik V., Kureichik V., Bova V. Placement of VLSI fragments based on a multilayered approach, *Advances in Intelligent Systems and Computing*, 2016, Vol. 464, pp. 181-190.
14. Kureichik V.V., Zaruba D.V. The bioinspired algorithm of electronic computing equipment schemes elements placement, *Advances in Intelligent Systems and Computing*, 2015, Vol. 347, pp. 51-58.

15. Gladkov L.A., Gladkova N.V., Gordienko V.N. Modifitsirovannyi geneticheskiy algoritm resheniya zadachi komponovki blokov EVA [Modified genetic algorithm for solving the problem of EVA block layout], *Informatika, vychislitel'naya tekhnika i inzhenernoe obrazovanie* [Computer science, computer engineering and engineering education], 2015, No. 4 (24), pp. 18-27.
16. Kureychik V.V., Kureychik V.I. Bioinspirovannyi algoritm razbieniia skhem pri proektirovanii SBIS [Bioinspired algorithm for splitting circuits in VLSI design], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2013, No. 7 (144), pp. 23-29.
17. Kureychik V.V., Kureychik V.M., Gladkov L.A., Sorokoletov P.V. Bioinspirovannye metody v optimizatsii [Bioinspired methods in optimization]. Moscow: Fizmatlit, 2009, 384 p.
18. Kureychik V.V., Zaruba D.V., Zaporozhets D.Yu. Bioinspirovannyi algoritm komponovki blokov EVA na osnove modifitsirovannoy raskraski grafa [Bioinspired algorithm for EVA block layout based on modified graph coloring], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, No. 4 (165), pp. 6-14.
19. Kureychik V.V., Kureychik V.I., Bova V.V. Bioinspirovannyi poisk v zadachakh konstruktorskogo proektirovaniia i optimizatsii [Bioinspired search in the problems of design design and optimization], *Informatsionnye tekhnologii v nauke, obrazovanii i upravlenii* [Information technologies in science, education and management], ed. by prof. E.L. Gloriozova, 2015, pp. 427-432.
20. Kuliev E.V., Lezhebokov A.A., Dukkardt A.N. Podkhod k issledovaniyu okrestnostey v roevykh algoritmakh dlya resheniya optimizatsionnykh zadach [Approach to neighborhood research in swarm algorithms for solving optimization problems], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2014, No. 7 (156), pp. 15-25.
21. Kuliev E.V., Lezhebokov A.A. Issledovanie kharakteristik gibridnogo algoritma razmeshcheniya [The study of the characteristics of hybrid positioning algorithm], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2013, No. 3 (140), pp. 255-261.

Статью рекомендовал к опубликованию д.т.н., профессор Н.И. Витиска.

Кулиев Эльмар Валерьевич – Южный федеральный университет; e-mail: ekuliev@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371651; доцент.

Лежебоков Андрей Анатольевич – e-mail: aalezhebokov@sfedu.ru; доцент.

Семенова Марина Максимовна – e-mail: bogomolova@sfedu.ru; студент.

Семенов Виктор Алексеевич – e-mail: viksemenov@sfedu.ru; студент.

Kuliev Elmar Valerievich – Southern Federal University; e-mail: ekuliev@sfedu.ru; 44, Nekrasovsky, Taganrog, 347928, Russia; phone: +78634371651; associate professor.

Lezhebokov Andrey Anatolevich – e-mail: aalezhebokov@sfedu.ru; associate professor.

Semenova Marina Maximovna – e-mail: bogomolova@sfedu.ru; student.

Semenov Victor Alekseevich – e-mail: viksemenov@sfedu.ru; student.

УДК 658.512

DOI 10.18522/2311-3103-2020-2-179-192

Е.В. Данильченко, В.И. Данильченко, В.М. Курейчик

КЛАССИФИКАЦИЯ И АНАЛИЗ ЭВОЛЮЦИОННЫХ МЕТОДОВ КОМПОНОВКИ БЛОКОВ ЭВА

В настоящее время наблюдается большой рост потребности в проектировании и разработке радиоэлектронных устройств. Это связано с повышающимися требованиями к радиоэлектронным системам, а также появлением новых поколений полупроводниковых приборов. В этой связи возникает необходимость в разработке новых средств автоматизированного компоновки блоков ЭВА. Перед компоновкой блоков ЭВА существует ряд проблем, которые усложняют реальное представление знаний в САПР и вероятно разрешимы на нынешнем уровне развития когнитивных наук. Проблема стереотипа и проблема огруб-

ления - взаимосвязаны и нуждаются в создании гибридных моделей представления. В работе рассмотрена проблема решения задачи компоновки блоков ЭВА при проектировании радиоэлектронной аппаратуры. **Цель данной работы** заключается в нахождении путей оптимизации планирования компоновки блоков ЭВА с применением генетического алгоритма. **Актуальность работы** состоит в том, что генетический алгоритм позволяет повысить качество планирования компоновки. Рассматриваемые алгоритмы позволяют повысить качество и скорость планирования компоновки. **Научная новизна** заключается в поиске и анализе эффективных методов компоновки блоков ЭВА с помощью генетических алгоритмов. **Принципиальное отличие** от известных сравнений в анализе новых перспективных алгоритмов компоновки блоков ЭВА. **Результаты работы.** В работе указаны недостатки традиционных алгоритмов поиска субоптимального плана ЭВА. Приведены описания современных моделей эволюционных и других вычислений. Генетические алгоритмы обладают рядом важных преимуществ – это приспособляемость к изменяющейся окружающей среде, при эволюционном подходе есть возможность анализировать, дополнять и изменять базу знаний в зависимости от изменяющихся условий, а также быстрое создание оптимальных решений. Если применять генетические алгоритмы и эвристику предварительной обработки, чтобы обеспечить оптимальные начальные решения, то можно достичь более продуктивного использования алгоритмов. Известные генетические алгоритмы быстро сходящиеся, но при этом они теряют разнообразие популяции, что влияет на снижение качества решения. Для балансировки данных решение выправляют с помощью эффективных операторов или устойчивой мутацией.

Генетический алгоритм; компоновка блоков ЭВА; методы оптимизации; графы и гиперграфы; эволюционные вычисления.

E.V. Danilchenko, V.I. Danilchenko, V.M. Kureychik

CLASSIFICATION AND ANALYSIS OF EVOLUTIONARY METHODS OF EVA BLOCK LAYOUT

Currently, there is a large increase in the need for the design and development of radio-electronic devices. This is due to increasing requirements for radio-electronic systems, as well as the emergence of new generations of semiconductor devices. In this regard, there is a need to develop new tools for automated layout of EVA blocks. There are a number of problems that complicate the actual representation of knowledge in CAD and are probably solvable at the current level of cognitive science development. The problem of stereotyping and the problem of coarsening are interrelated and need to create hybrid models of representation. The paper deals with the problem of solving the problem of EVA block layout in the design of radio-electronic equipment. **The purpose of this work** is to find ways to optimize the planning of EVA block layout using a genetic algorithm. **The relevance of the work** is that the genetic algorithm can improve the quality of layout planning. These algorithms allow you to improve the quality and speed of layout planning. **The scientific novelty lies** in the search and analysis of effective methods for composing EVA blocks using genetic algorithms. **The main difference** from the known comparisons is in the analysis of new promising algorithms for composing EVA blocks. **Result of work.** The paper shows the disadvantages of traditional algorithms for searching for a suboptimal EVA plan. Descriptions of modern models of evolutionary and other calculations are given. Genetic algorithms have a number of important advantages – adaptability to a changing environment, the evolutionary approach makes it possible to analyze, Supplement and change the knowledge base depending on changing conditions, as well as quickly create optimal solutions. If you apply genetic algorithms and pre-processing heuristics to provide optimal initial solutions, you can achieve more productive use of algorithms. Known genetic algorithms converge quickly, but they lose population diversity, which affects the quality of the solution. To balance data, the solution is corrected using efficient operators or stable mutation.

Genetic algorithm; EVA block layout; optimization methods; graphs and hypergraphs; evolutionary calculations.

Введение. В данной работе освещаются актуальные способы для решения задачи компоновки блоков ЭВА. Задачи конструкторского проектирования относятся к классу комбинаторных оптимизационных задач, которые осуществляют с по-

мощью теории графов, гиперграфов, а также на их матричных и списковых эквивалентах. Огромный плюс в моделировании коммутационной схемы посредством графа – это эффективное применение в решении задач компоновки. От четкости структуры графа и его данных о решаемых задачах, соответственно будет зависеть правильность результатов и качество проектирования.

При рассмотрении задачи с точки зрения математической модели, она представляет собой сложно-характерную задачу выбора оптимального или квазиоптимального решения из огромного числа представленных вариантов.

В работе рассмотрены положительные и отрицательные стороны традиционных алгоритмов компоновки блоков ЭВА. Приведены описания современных моделей эволюционных и других вычислений. Сделан вывод о перспективных на данный момент алгоритмах компоновки ЭВА.

Постановка задачи. *Задача работы* – классификации известных методов компоновки блоков ЭВА, определение оптимального варианта компоновки блоков ЭВА.

Генерация методов компоновки блоков ЭВА или разбиение коммутационной схемы на части связана с необходимостью обрабатывания колоссальных объемов информации. Такие задачи относят к классу комбинаторно-логических задач и считаются NP-полными. При помощи алгоритма уменьшается порядок графа, а следовательно, увеличивается объем обрабатываемой информации. *В связи с этим фактом разработка новых методов компоновки схем ЭВА – важная и актуальная задача.*

Классификация и анализ методов компоновки блоков ЭВА. Рассмотрим графовые, гиперграфовые и ультраграфовые модели блоков. Решение задач компоновки блоков ЭВА осуществляется посредством перехода от блоков и схем соединений ЭВА к графам, гиперграфам и др. Чтобы качественно решить задачу разбиения блоков ЭВА необходимо разработать алгоритмы и математические методы для ЭВМ. Поэтому на практике удобнее использовать не схемы соединений ЭВМ, а их абстрактные математические модели (графы).

Условная классификация алгоритмов компоновки показана на рис. 1 [2].

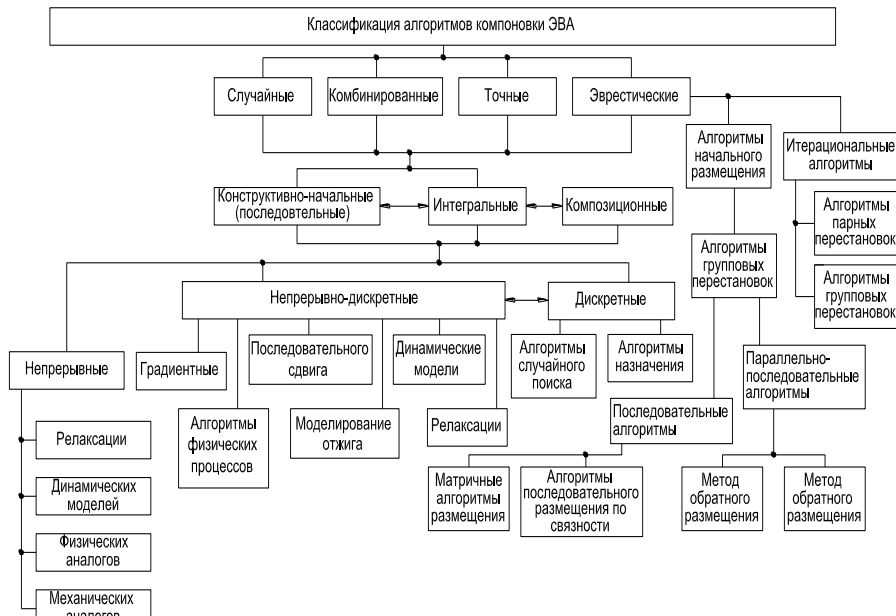


Рис. 1. Условная классификация алгоритмов компоновки блоков ЭВА

В этой классификации объединены методы, учитывающие различные требования. Рассмотрим основные алгоритмы, кратко обозначив плюсы и минусы. Самыми применимыми алгоритмами в САПР считаются: алгоритм Кернигана – Лиана, а также его улучшенные эвристики, алгоритм Федучио – Маттеуса с модификациями, кластерные алгоритмы, алгоритмы формирования минимальных массивов, алгоритм моделирования отжига и модификации, матричные и списковые алгоритмы, алгоритмы случайных назначений, генетические и эволюционного моделирования.

Отметим, что данную классификацию будем считать условной в связи с тем, что многие методы можно учитывать в разных группах и на основе их объединения. Это позволяет автору выделить следующие классы методов на основе характеристик объекта проектирования: конструктивные (последовательные), итерационные, реализующие математические или физические аналоги задачи компоновки, декомпозиционные, поисковые, случайные, эволюционные и биоинспирированные.

Рассмотрим *непрерывные методы* [3]. Основным его представителем является метод силовой релаксации, который основан на механической аналогии. Данные методы эффективны для компоновки небольшого количества элементов $n < 100$, но очень зависим от начальных условий задачи, и имеет медленную скорость получения результата. $BSCA = O(n^2) \div O(n^4)$.

Рассмотрим далее группу *непрерывно-дискретные*. Гильотинный план методов оптимизации [8]. Одним из представителей этой группы является метод градиента. В этом методе поиск осуществляется в направлении наискорейшего возрастания ЦФ, т.е. в направлении градиента.

В отличие от метода градиента в данном алгоритме:

Направление поиска определяется не градиентом функции цели, а наиболее удачным из пробных шагов, выполненным из k -й точки в случайных направлениях; при этом координаты текущей изображающей точки в пространстве нормированных параметров вычисляются по формуле [9]:

$$X'_{i,k} = X'_{i,k-1} + \frac{\delta x'_i \cdot h}{\delta},$$

где $\delta x'_i$ – проекция наиболее удачного пробного шага на ось x_i , d – размер пробного шага;

Для формирования направлений пробных шагов здесь так же, как и в методе статистических испытаний, нужен датчик случайных чисел; это обстоятельство определяет особенности построения алгоритма реализации данного метода в сравнении с методом градиента; в общем случае количество необходимых пробных шагов определяется из требования равномерного исследования ϵ -окрестности вокруг изображающей точки, часто оно выбирается равным 5.

В остальном процесс поиска экстремума функции цели здесь идентичен поиску по методу градиента.

На начальном этапе температуру принимают высокой, а затем ее ступенчато снижают. При каждой температуре выполняют серию пробных переборов решений, и после каждой перестановки подсчитывается значение целевой функции. Лучшие решения принимаются с вероятностью 1, а «плохие», для которых значение целевой функции ухудшается, принимаются с некоторой вероятностью. Такой вероятностный механизм дает возможность, принимая в качестве исходных некоторые «плохие» решения, проскакивать через локальные оптимумы и находить глобальные.

Если пробное изменение привело к улучшению целевой функции F , то это изменение фиксируется. Если пробное изменение привело к ухудшению F на величину ΔF , то рассчитывается вероятность сохранения изменения по формуле

$$P = \exp\left(-\frac{\Delta F}{kT}\right).$$

где k – константа.

Следующим представителем данной группы являются генетические методы. Данные методы являются поисковыми и основаны на механизмах натуральной селекции и генетики. Здесь также реализуется принцип "выживание сильнейших". Причем эти методы позволяют использовать информацию, накопленную в процессе поиска [12].

Генетические алгоритмы. Генетические алгоритмы (ГА) оперируют с популяцией решений. С одной стороны, это позволяет быстрее находить лучшие решения, но с другой стороны требуется больше памяти для хранения информации о популяции решений. Тем не менее, последние исследования, связанные с использованием генетических методов оптимизации в различных областях, показали их высокую эффективность.

Каждой индивидуальности соответствует оценка качества. После репродукции новых членов популяции на основе селекции осуществляется усечение (уменьшение) популяции до начального размера. Эта процедура соответствует методу естественного отбора или методу выживания сильнейшего.

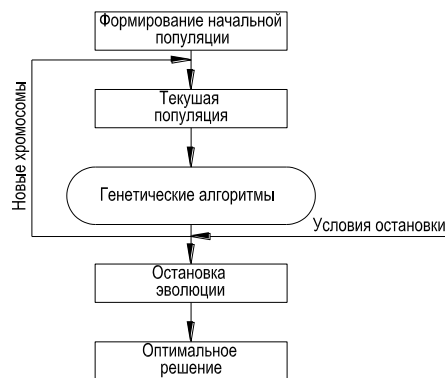


Рис. 2. Структура генетического алгоритма

Селекция – это процесс, посредством которого хромосомы, имеющие более высокое функциональное значение, получают большую возможность для репродукции, чем «слабые» хромосомы. Элементы, выбранные для селекции, обмениваются генетическим материалом, создавая потомков. Существует несколько основных видов селекции [14].

Основным итогом анализа генетического алгоритма стал выбор триады генетических операторов (ГО): кроссингвер – мутация – инверсия. Воздействуя с некоторой вероятностью на генотипы родительских особей, каждый из них, с одной стороны, обеспечивает передачу потомству важных признаков, а, с другой, – поддерживает на протяжении эволюционно значимого периода достаточно высокий уровень его изменчивости. Определение в потомстве новых, отличных от родительских, фенотипических (совокупность всех внешних и внутренних признаков) признаков открывает для популяции дополнительные возможности для адаптации.

Мутация – генетическое изменение, приводящее к качественно новому проявлению основных свойств генетического материала: дискретности, непрерывности и линейности. Таким образом, мутации представляют собой изменения генов, хромосом.

Основной недостаток методов генетического поиска – большой объем вычислений на каждой итерации.

Муравьиный алгоритм. Идея муравьиного алгоритма (ACO – ant colony optimization) заключается в моделировании поведения муравьев, который связан с их умением быстро находить наиболее краткий маршрут от муравейника к источнику пищи.

Поиск решения задачи планирования осуществляется коллективом муравьев $Z = \{z_k | k = 1, 2, \dots, 1\}$. На каждой итерации муравьиного алгоритма каждый муравей z_k строит индивидуальное решение задачи планирования. Решением является маршрут в графе $G = (X, U)$, включающий n_M вершин, принадлежащих множествам X_1 или X_2 , n_R вершин, принадлежащих множествам X_3 или X_4 .

Моделирование поведения муравьев в задаче планирования связано с распределением феромона на ребрах графа G . На начальном этапе на всех ребрах графа G откладывается небольшое количество феромона. Процесс поиска решений итерационный. Каждая итерация включает три этапа. На первом этапе муравей находит решение, на втором этапе откладывает феромон, на третьем этапе осуществляется испарение феромона [17].

Метод роя частиц. Рой частиц можно рассматривать как многоагентная система, каждый агент действует независимо по простым правилам. Совокупное поведение системы получается разумным. Оптимизация методом роя частиц (particle swarm optimization – PSO) была предложена Кеннеди и Эберхартом в 1995 г.

На каждой итерации частицы перемещаются в новые позиции. Новая позиция определяется по формуле:

$$x_i(t+1) = x_i(t) + v_i(t+1),$$

где $v_i(t+1)$ – скорость перемещения частицы из позиции $x_i(t)$ в позицию $x_i(t+1)$.

Формула расчета скорости состоит из трех слагаемых. Предыдущая скорость частицы служит инерционным компонентом. Второе слагаемое прямо пропорционально текущему расстоянию частицы от ее наилучшей позиции, которая была найдена с первой итерации. Второе слагаемое является когнитивным компонентом. Третье слагаемое, называемое социальным компонентом, прямо пропорционально текущему расстоянию частицы от наилучшей позиции роя в текущий момент времени t . Благодаря социальному компоненту, частица может передвигаться в оптимальные позиции, найденные другими частицами [18].

Декомпозиционные методы заключаются в последовательном, параллельном, параллельно-последовательном (последовательно параллельном) разбиении математической модели КС и модели монтажного пространства.

Декомпозиция основана на иерархии. В общем случае термин иерархия означает соподчиненность. К таким объектам относятся структурные элементы системы: иерархические структуры представляют собой декомпозицию системы в пространстве [15].

Поисковые методы компоновки для получения эффективного решения используют иерархические приемы с разбиением исходной задачи на ряд подзадач и решения каждой из них перебором на дереве решений. Поисковые методы позволяют заменить полный перебор вариантов перебором с отсечениями неэффективных решений, обеспечивая при этом высокую точность решения.

Методы компоновки по связности относятся к последовательным методам. Здесь вводится t -шаговый процесс принятия решений. При этом, на каждом шаге выбирается один из не размещенных элементов и помещается в одну из свободных позиций.

Рассматриваются позиции, соседние с уже скомпонованными элементами, например: (здесь x – уже размещенные элементы, v – оцениваемые позиции-кандидаты). Пусть L – множество незанятых позиций-кандидатов, а его элементы – l_i . Элемент, для которого ищется позиция e_i .

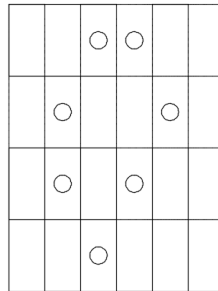


Рис. 3. Оценка позиции

Последующий выбор элементов основан на анализе связности оставшихся скомпонованных. Временная сложность алгоритмов равна $O(\alpha_n)$.

Методы связывания пар относятся к группе последовательных алгоритмов. Процесс начинается с выбора в качестве начальной пары элементов, которые имеют, например, наибольшую сумму общих связей (ребер ММ, цепей КС).

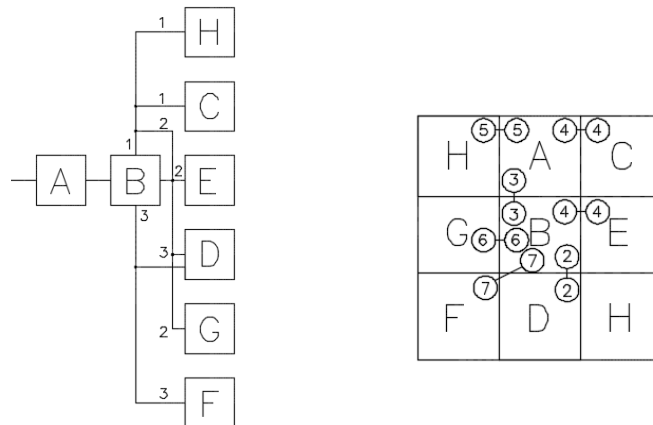


Рис. 4. Алгоритм разрешения связывания пар

Происходит последовательное подсоединение элементов с выполнением первоначально введенного правила до полной компоновки всех элементов. Временная сложность алгоритма равна $O(\alpha_n)$.

Методы расширения ядра относятся к последовательным алгоритмам. Здесь для каждого не размещенного элемента определяется среднее суммарное число ожидаемых связей с скомпонованными элементами на основе анализа покрывающих деревьев.

При использовании метода восходящего проектирования в первую очередь определяются вспомогательные функции, которые могут потребоваться для проектирования программы.

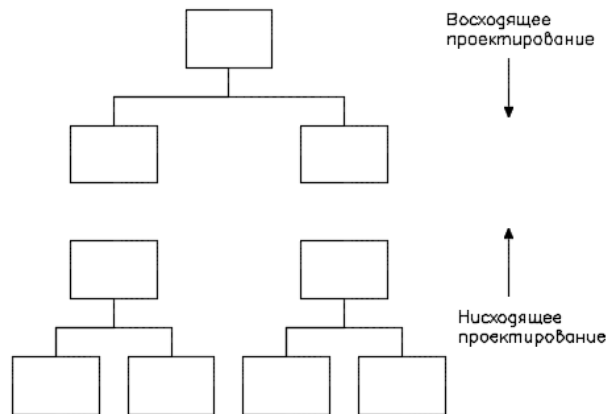


Рис. 5. Проектируются модули нижнего уровня, а затем программа проектируется одновременно нисходящими и восходящими методами

Для компоновки выбирается элемент с наибольшим значением и располагается во взвешенном центре тяжести групп уже размещенных элементов. Далее процесс продолжается итеративно. Временная сложность алгоритма равна $O(\alpha_n)$.

Матричные методы компоновки относятся к конструктивным алгоритмам. Выбор элемента и позиции для него на γ -м шаге выполняется по специальной матрице компоновки $B = \|b_{ij}\|$, каждый элемент которой b_{ij} соответствует «цене назначения» элемента x_i в позицию p_j при условии, что $\gamma-1$ элементов уже размещены. Структура алгоритма основана на различных способах определения «цен назначений» и методиках выбора. Временная сложность алгоритма лежит в диапазоне от $O(\alpha_n)$ до $O(\alpha_n^2)$.

Методы обратной компоновки относятся к последовательным алгоритмам. Осуществляют предварительную оценку каждого из размещаемых элементов $x_1, x_2, \dots, x_n$ и каждой свободной позиции $p_1, p_2, \dots, p_n$. Далее производится упорядочивание элементов по возрастанию или убыванию введенных характеристик, и затем все элементы размещаются одновременно [13–15].

В методе осуществляется предварительная оценка каждого из размещаемых элементов $x_1, x_2, \dots, x_n$ и свободной позиции $l_1, l_2, \dots, l_n$. После все элементы размещаются одновременно.

Позиции в центре коммутационного поля имеют меньшее значение d_i , чем крайние. Т.е. в центре коммутационного поля размещены элементы с большим значением g_i (сильно связанные элементы).

Метод прост, но не всегда дает наилучший вариант компоновки (используется, как начальный вариант компоновки). Временная сложность алгоритма равна $O(\alpha_n)$.

Метод разбиения, используются как последовательные, так и итерационные алгоритмы [14]. Для улучшения качества алгоритма компоновки разбиение можно выполнять на горизонтальные и вертикальные линейки и в общем случае на области произвольной конфигурации.

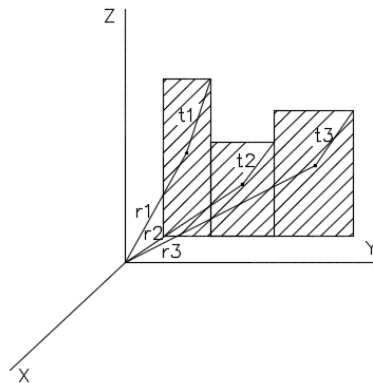


Рис. 6. Пример плоской фигуры

Плоскую фигуру можно разбить на три части, центры тяжести которых C_1 , C_2 , C_3 известны. Они находятся на пересечении диагоналей прямоугольников. Их радиусы-векторы обозначим r_1 , r_2 , r_3 и площади S_1 , S_2 , S_3 . Общая площадь сложной фигуры будет $S = S_1 + S_2 + S_3$.

$$r = \frac{S_1 r_1 + S_2 r_2 + S_3 r_3}{S_1 + S_2 + S_3}.$$

Дополненная формула имеет ту же структуру, что и формула, определяющая радиус-вектор центра тяжести тела при разбиении его на элементарные частицы, только в нее входят величины для конечных частей тела. Временная сложность алгоритма равна $O(\alpha_n^2)$.

Методы перестановок. Являются простейшими, их временная сложность составляет ориентировочно $O(n^2)$. Методы групповых перестановок. Одним из основных является метод Штейнберга, который использует алгоритм нахождения независимых подмножеств графа.

Алгоритм Штейнберга анализирует независимые подмножества графа и осуществляет перестановку вершин между ними [11]. Получить независимые подмножества при $n > 1000$ невозможно, что является основным недостатком этого метода.

Перестановки используются также в методе, основанном на применении маршрутов Гамильтона.

Расшифровка производится в обратном порядке. В соответствии с ключом выбирается маршрут и заполняется таблица согласно этому маршруту.

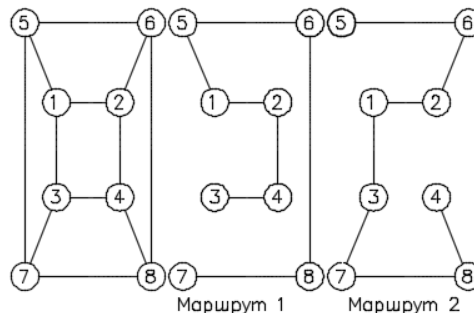


Рис. 7. Вариант 8-элементной таблицы и маршрутов Гамильтона

Из таблицы символы считываются в порядке следования номеров элементов. *Достоинства*: они всегда обеспечивают получение локального оптимума. *Недостатки*: итерационные алгоритмы часто попадают в локальные оптимумы и не могут выйти из них. В итерационных алгоритмах ВСА лежит в пределах от $O(n \log n)$ до $O(n^2)$.

Экспериментальные исследования. Алгоритм как объект разработки с одной стороны и средство решения задачи с другой является таким же предметом исследования, как те задачи, которые с помощью него решаются. С прикладной точки зрения разработчика алгоритма интересуют такие вопросы, как:

- ◆ качество решения, предоставляемого алгоритмом;
- ◆ скорость решения поставленной задачи (время работы).

Определим теоретическую оценку временной сложности композитного метода размещения. Начальное решение задачи можно получить последовательным алгоритмом, временная сложность которого имеет линейный характер $O(\alpha N)$, или алгоритмом «слепого» поиска, временная сложность которого составляет $O(\beta IN)$, где I – это количество итераций (шагов) выполняемых алгоритмом. В результате анализа выходных данных, авторами отмечается, что временная сложность разработанного модифицированного алгоритма не выходит за пределы полиномиальной зависимости, и может быть выражена формулой: $O(\alpha N^2) - O(\beta N^3)$, где N – число элементов схемы (размер решаемой задачи).

Проведем анализ зависимости времени работы от количества итераций алгоритмов.

Таблица 1

Зависимость времени работы алгоритмов от размера схемы

Алгоритмы	Число элементов схемы									
	100	200	300	400	500	600	700	800	900	1000
Генетический	0,2	0,6	0,9	1,2	1,5	1,9	2,3	3,3	4,7	7,5
Эволюционный	1,5	2,3	2,3	3,3	5,2	9,8	11,7	16,4	28,6	56,3
Роевой	3,6	5,6	8	9,4	13,1	17,8	28,6	32,8	46,4	87,7
Гибридный	8,5	10,1	12,6	15,5	24,3	35,2	50,2	63,3	82	108,8

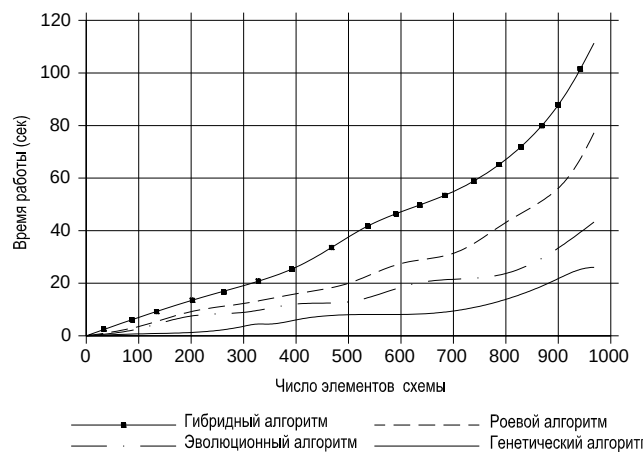


Рис. 8. Графики зависимости времени работы алгоритмов от количества элементов схемы

Для определения эффективности разработанного метода были проведены исследования качества решения на нескольких тестовых примерах. Под эффективностью алгоритма будем понимать качество решения, получаемого с его помощью (табл. 2). На рис. 9 приведена гистограмма сравнения качества решения, основываясь на данных из табл. 2.

Таблица 2

Определение эффективности алгоритмов

Алгоритмы	Число элементов схемы				
	1000	2500	5000	7500	10000
Генетический	0,2	0,6	0,9	1,2	1,5
Эволюционный	1,5	2,3	2,3	3,3	5,2
Роевой	3,6	5,6	8	9,4	13,1
Гибридный	8,5	10,1	12,6	15,5	24,3

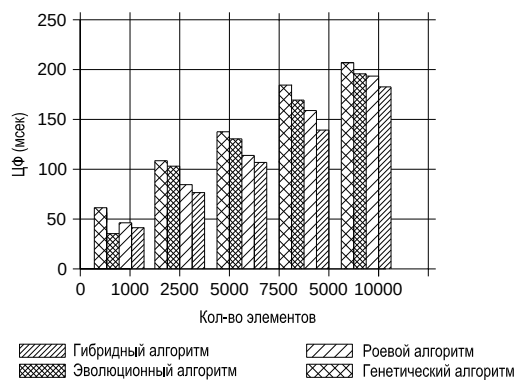


Рис. 9. Гистограмма сравнения качества решения

Проанализировав полученные данные из таблицы (см. табл. 2) и гистограммы (см. рис. 9), можно сделать вывод, что применение генетических и эволюционных алгоритмов целесообразнее на задачах малой размерности (до 2000 элементов), а при решении задачи компоновки в размерности, приближенной к промышленным объемам (более 5000 элементов), эффективным является гибридный алгоритм.

Закключение. Исходя из данного анализа становится ясно, что нужно разрабатывать новые нестандартные архитектуры ЭМ, гибридные ГА с эвристическими методами поиска и модифицированных генетических операторов. ГА – это эффективное решение для выхода и локальных оптимумов. Хорошие результаты достигаются при использовании в решении комбинаторно-логических задач компоновки в САПР. Однако, созданные ГА не до конца решают проблему выходов из локальных оптимумов. Чтобы улучшить качество разбиения и сократить затраты времени алгоритмов, следует разработать новые последовательно-параллельные и параллельно-последовательные структуры поиска.

Рассмотрев известные методы и проведя их анализ, можно отметить следующее, что большое число рассмотренных методов не позволяет эффективно решать НР-сложные задачи. Поэтому для повышения эффективности решаемых задач автор предлагает создавать новые архитектуры поиска, основанные на гибридных и комбинированных подходах, учитывающих все достоинства перспективной технологии эволюционного моделирования и биоинспирированных методов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Курейчик В.В., Курейчик В.М., Родзин С.И. Концепция эволюционных вычислений, инспирированных природными системами // Известия ЮФУ. Технические науки. – 2009. – № 4 (93). – С. 16-24.
2. Данильченко В.И., Курейчик В.М. Классификация и анализ методов решения задачи компоновки ЭВА // Информатика, вычислительная техника и инженерное образование. – 2018. – № 1 (32). – С. 3-23.
3. 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).
4. Jackuline Moni D. and Arumugam S. and Gracia Nirmala Rani D. VLSI Floor Planning relying on Differential Evolution Algorithm // ICGST International Journal on Artificial Intelligence and Machine Learning. – 2007. – Vol. 7, No. 1. – P. 62-67.
5. Лебедев О.Б. Планирование ЭВА на основе метода муравьиной колонии // Известия ЮФУ. Технические науки. – 2010. – № 7 (108). – С. 67-73.
6. Tsung-Ying Sun, Sheng-Ta Hsieh, Hsiang-Min Wang, Cheng-Wei Lin. Floorplanning based on particle swarm optimization // ISVLSI '06 Proceedings of the IEEE Computer Society Annual Symposium on Emerging VLSI Technologies and Architectures. – 2006. – P. 5-10.
7. 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.
8. 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.
9. Wu Y., Jiang H., Zou M. The Research on Fuzzy PID Control of the Permanent Magnet Linear Synchronous Motor // Physics Procedia. – 2012. – Vol. 24. – P. 1311-1318.
10. Abbasi E., Mahjoob M. J., Yazdanpanah R. Controlling of Quadrotor UAV Using a Fuzzy System for Tuning the PID Gains in Hovering Mode // Fourth International Conference on Advances in Computer Engineering – ACE 2013. – Frankfurt, Germany, 2013. Int. j. adv. robot. syst. – 2013. – Vol. 10, 380-2013.
11. Kai Ou, Ya-Xiong Wang, Zhen-Zhe Li, Yun-De Shen, Dong-Ji Xuan. Feedforward fuzzy-PID control for air flow regulation of PEM fuel cell system // International journal of hydrogen energy. – 21 September 2015. – Vol. 40, Issue 35. – P. 11686-11695.
12. Ahmet Karli, Vasfi Emre Omurlu, Utku Buyuksahin, Remzi Artar, 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).
13. Hamed Beirami, Ali Zargar Shabestari, Mohammad Mahdi Zerafat. Optimal PID plus fuzzy controller design for a PEM fuel cell air feed system using the self-adaptive differential evolution algorithm // International journal of hydrogen energy. – 10 August 2015. – Vol. 40, Issue 30. – P. 9422-9434.
14. Jahedi G., Ardehali M.M. Genetic algorithm-based fuzzy-PID control methodologies for enhancement of energy efficiency of a dynamic energy system // Energy Conversion and Management. – 2011. – Vol. 52. – P. 725-732.
15. Лебедев Б.К., Лебедев В.Б. Планирование на основе роевого интеллекта и генетической эволюции // Известия ЮФУ. Технические науки. – 2009. – № 4 (93). – С. 25-33.
16. Кулиев Э.В. Задача размещения элементов ЭВА с использованием генетического алгоритма и алгоритма пчелиной колонии // Тр. конгресса по интеллектуальным системам и информационным технологиям «IS-IT'12». Научное издание в 4-х т. Т. 3. – М.: Физматлит, 2012. – С. 99-104.
17. Кришук В.Н., Карпов Л.М., Шило Г.Н., Фармонов А.Ю., Артюшенко Б.А. Применение генетического алгоритма в задачах допускового синтеза микрополосковых устройств // Техника сверхвысоких частот. – Киев: Научное издание НПП «Сатурн», 2007. – С. 29-32.
18. Кришук В.Н., Шило Г.Н., Артюшенко Б.А. Генетический алгоритм назначения интервальных допусков // Всероссийское совещание по интервальному анализу и его приложениям. – Петергоф, Россия. – 2006. – С. 5-8.
19. 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. – 2012. – Т. 26, № 2. – С. 104-112.
20. Abrie P.L.D. Design of RF and microwave amplifiers and oscillators. – London; Boston: Artech House, 2000. – 480 p.

REFERENCES

1. Kureychik V.V., Kureychik V.M., Rodzin S.I. Kontseptsiya evolyutsionnykh vychisleniy, inspirirovannykh prirodnymi sistemami [Concept of evolutionary computing inspired by natural systems], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2009, No. 4 (93), pp. 16-24.
2. Danil'chenko V.I., Kureychik V.M. Klassifikatsiya i analiz metodov resheniya zadachi komponovki EVA [Classification and analysis of methods for solving the problem of EVA layout], *Informatika, vychislitel'naya tekhnika i inzhenernoe obrazovanie* [Computer science, computer engineering and engineering education], 2018, No. 1 (32), pp. 3-23.
3. 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).
4. Jackuline Moni D. and Arumugam S. and Gracia Nirmala Rani D. VLSI Floor Planning relying on Differential Evolution Algorithm, *ICGST International Journal on Artificial Intelligence and Machine Learning*, 2007, Vol. 7, No. 1, pp. 62-67.
5. Lebedev O.B. Planirovanie EVA na osnove metoda murav'inoi kolonii [EVA planning based on the ant colony method], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2010, No. 7 (108), pp. 67-73.
6. Tsung-Ying Sun, Sheng-Ta Hsieh, Hsiang-Min Wang, Cheng-Wei Lin. Floorplanning based on particle swarm optimization, *ISVLSI '06 Proceedings of the IEEE Computer Society Annual Symposium on Emerging VLSI Technologies and Architectures*, 2006, pp. 5-10.
7. 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.
8. 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.
9. Wu Y., Jiang H., Zou M. The Research on Fuzzy PID Control of the Permanent Magnet Linear Synchronous Motor // *Physics Procedia*. – 2012. – Vol. 24. – P. 1311-1318.
10. Abbasi E., Mahjoob M.J., Yazdanpanah R. Controlling of Quadrotor UAV Using a Fuzzy System for Tuning the PID Gains in Hovering Mode, *Fourth International Conference on Advances in Computer Engineering – ACE 2013*. Frankfurt, Germany, 2013. Int. j. adv. ro-bot. syst., 2013, Vol. 10, 380:2013.
11. Kai Ou, Ya-Xiong Wang, Zhen-Zhe Li, Yun-De Shen, Dong-Ji Xuan. Feedforward fuzzy-PID control for air flow regulation of PEM fuel cell system, *International journal of hydrogen energy*, 21 September 2015, Vol. 40, Issue 35, pp. 11686-11695.
12. Ahmet Karli, Vasfi Emre Omurlu, Utku Buyuksahin, Remzi Artar, 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).
13. Hamed Beirami, Ali ZargarShabestari, Mohammad Mahdi Zerafat. Optimal PID plus fuzzy controller design for a PEM fuel cell air feed system using the self-adaptive differential evolution algorithm, *International journal of hydrogen energy*, 10 August 2015, Vol. 40, Issue 30, pp. 9422-9434.
14. Jahedi G., Ardehali M.M. Genetic algorithm-based fuzzy-PID control methodologies for enhancement of energy efficiency of a dynamic energy system, *Energy Conversion and Management*, 2011, Vol. 52, pp. 725-732.
15. 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.
16. Kuliev E.V. Zadacha razmeshcheniya elementov EVA s ispol'zovaniem geneticheskogo algoritma i algoritma pchelinoi kolonii [The problem of placing EVA elements using a genetic algorithm and the algorithm of a bee colony], *Tr. kongressa po intellektual'nym sistemam i informatsionnym tekhnologiyam «IS-IT'12»* [Proceedings of the Congress on intelligent systems and information technologies "IS-IT'12"]. Scientific publication in 4 vols. Vol. 3. Moscow: Fizmatlit, 2012, pp. 99-104.
17. Krishchuk V.N., Karpov L.M., Shilo G.N., Faramonov A.Yu., Artyushenko B.A. Primenenie geneticheskogo algoritma v zadachakh dopuskovogo sinteza mikropoloskovykh ustroystv [Application of the genetic algorithm in the problems of tolerance synthesis of microstrip devices], *Tekhnika sverkhvysokikh chastot* [Ultrahigh frequency technology]. Kiev: Nauchnoe izdanie NPP «Saturn», 2007, pp. 29-32.

18. *Krishchuk V.N., Shilo G.N., Artyushenko B.A.* Geneticheskiy algoritm naznacheniya interval'nykh dopuskov [Genetic algorithm for assigning interval tolerances], *Vserossiyskoe soveshchanie po interval'nomu analizu i ego prilozheniyam* [All-Russian meeting on interval analysis and its applications]. Petergof, Rossiya, 2006, pp. 5-8.
19. *Kalent'ev A.A., Garays D.V., Dobush I.M., Babak L.I.* Strukturno-parametricheskii 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 TUSURA* [Reports Of TUSUR], 2012, Vol. 26, No. 2, pp. 104-112.
20. *Abrie P.L.D.* Design of RF and microwave amplifiers and oscillators. London; Boston: Artech House, 2000, 480 p.

Статью рекомендовал к опубликованию к.т.н., доцент 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: vdanilchenko@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, Taganrog, 347928, Russia; phone: +79525691761; the department of computer-aided design; postgraduate.

Danilchenko Vladislav Ivanovich – e-mail: vdanilchenko@sfedu.ru; phone: +79526088561; the department of computer-aided design; postgraduate.

Kureichik Viktor Mikhailovich – e-mail: vmkureychik@sfedu.ru; phone: +79282132730; the department of computer-aided design; dr. of eng. sc.; professor.

Раздел IV. Информационные системы и технологии

УДК 004.94

DOI 10.18522/2311-3103-2020-2-193-200

В.Н. Гридин, В.И. Анисимов, С.А. Васильев

МЕТОДЫ ПОВЫШЕНИЯ ПРОИЗВОДИТЕЛЬНОСТИ СОВРЕМЕННЫХ ВЕБ-ПРИЛОЖЕНИЙ*

Рассматриваются существующие и развивающиеся подходы построения современных веб-приложений. Определяются основные формы и направления развития современных веб-приложений, а также методы для повышения производительности обмена данными клиент-серверных системах. Освещаются разновидности и принципы установления каналов связи в распределенной клиент-серверной среде. Приводятся основные преимущества комбинированных методов взаимодействия с применением асинхронных полнодуплексных протоколов обмена данными для обеспечения высокой скорости передачи данных, предоставления информации своевременно, снижения нагрузки на серверную составляющую, снижения избыточности передаваемых данных. Указываются технологии децентрализации state management в одностраничных приложениях, взаимосвязь современных методик для обеспечения высокой степени интерактивности клиентской составляющей. Производится сравнительный анализ интеллектуального механизма обработки запросов, декларирования структуры данных и методов доступа к ним с ориентированным на работу с данными REST API, обеспечивающий различные вариации базовых CRUD операций. Освещаются основные достоинства подхода GraphQL по организации распределенного state management на основе предоставления клиентскому приложению графо-подобных структур неопределенного уровня вложенности, и возможности организации подписки на изменения в интересующем наборе данных. Приводятся проблемы традиционных систем хранения данных в современных информационных условиях, геометрическом накоплении сложно структурированных данных. Описываются основные подходы к хранению данных в разрезе концепции NoSQL. Рассматриваются преимущества использования модели ключ-значения в информационных системах. Определяются преимущества использования и принципы работы баз данных, использующих в качестве хранилища оперативную память. Рассматриваются недостатки указанных технологий хранения данных и предлагаются возможные пути их минимизации на основе коллаборации методов. В качестве выводов приводится схема зависимостей технологий эффективного обмена данными в современных веб-приложениях для обеспечения высокой степени интерактивности клиент-серверных веб-приложений.

Клиент-серверное приложение; WebSocket; GraphQL; SignalR; NoSQL; BASE; ACID; Redis.

V.N. Gridin, V.I. Anisimov, S.A. Vasilev

METHODS OF IMPROVING PERFORMANCE OF MODERN WEB-APPLICATIONS

Existing and developing approaches to build modern web applications are considered. The main forms and directions of development modern web applications are determined, as well as methods for increasing the productivity of data exchange client-server systems. The varieties and principles of establishing communication channels in a distributed client-server environment are highlighted. The main advantages to use combined methods of interaction made with asynchronous full-duplex data exchange protocols are provided to ensure a high data transfer rate, provide

* Работа выполнена в рамках темы № 0071-2019-0001.

information in a timely manner, reduce the load on the server component, and reduce the redundancy of transmitted data. The technologies of state management decentralization in single-page applications, the interconnection of modern technologies to ensure a high degree of interactivity of the client component are indicated. A comparative analysis of the intelligent query processing mechanism, declaring the data structure and access methods with data-oriented REST APIs is carried out, which provides variations of the basic CRUD operations. The main advantages of GraphQL's approach to organize distributed state management based on providing graph-like structures with an indefinite level of nesting to the client application and the possibility of subscribing to change in the data set of interest are highlighted. The problems of traditional data storage systems in modern information conditions, the geometric accumulation of complex structured data are presented. The basic approaches to data storage in the context of the NoSQL concept are described. The advantages of using the key-value model in information systems and the principles of operation databases using RAM as storage are considered. The disadvantages of these technologies for data storage are considered and possible ways to minimize them based on a collaboration of methods are proposed. Thus, a dependency diagram of technologies for efficient data exchange in modern web applications is provided to ensure a high degree of interactivity client-server web applications.

Client-server application; WebSocket; GraphQL; SignalR; NoSQL; BASE; ACID; Redis.

Введение. В настоящее время наиболее распространён вид предоставления сервисов и услуг посредством веб-приложений через сеть интернет. Существуют два принципа построения веб-приложений – традиционный и одностраничный (SPA – Single Page Application и PWA – Progressive Web Application). В первом случае бизнес логика выполняется на серверной составляющей, основные преимущества такого подхода весьма архаичны, это минимальные технические требования к клиентской стороне, к hardware и software вплоть до отсутствия поддержки JavaScript. Одностраничные приложения являются более требовательными к клиентской составляющей, как с точки зрения hardware, так и software несмотря на то, что основная бизнес логика может быть также размещена на серверной части, интерфейс целиком определяется и формируется на клиентской стороне, используя сервер как поставщик данных, заполняющих сформировавшийся шаблон. Сохранение состояния сервиса, предоставляемого пользователю, может осуществляться именно на клиентской стороне, при этом взаимодействие с сервером осуществляется главным образом через веб-API. Также имеются гибридные вариации приложений, например Blazor, однако на текущий момент они не являются востребованными [1].

Современное восприятие пользовательского интерфейса сдвигает парадигму построения веб-приложений в сторону одностраничной модели. Коммерческая конкурентоспособность такого метода построения веб-приложений достигается путем имитации поведения нативного программного обеспечения. В некоторых случаях веб-приложение преобразуется в нативное программное обеспечение, ярким примером разработки графических приложений для основных операционных систем (MacOS, Windows, Linux) на основе веб-технологий является фреймворк Electron, способный к рендеренгу html, css и JavaScript на вышеуказанных платформах. Для обеспечения такой работы клиентской стороне необходим высокоскоростной обмен данными, который может быть достигнут путем эффективной подготовки данных на сервере, а также быстрым каналом связи, обеспечивающим своевременной информацией как клиентскую, так и серверную составляющую [2].

Каналы связи. В качестве канала связи рекомендуется использовать полнодуплексный протокол, обеспечивающий своевременную доставку данных клиенту. В случае выполнения ресурсоемких вычислений, сервер инициирует отправку данных по их готовности. Такой подход повышает эффективность канала связи, а также снижает нагрузку на сервер, т.к. клиентское приложение ожидает ответ в отличие от стандартной модели, циклично формирующей запросы на сервер с заданным интервалом до момента получения результата. В качестве примера можно

привести протокол WebSocket, отвечающий вышеуказанным требованиям, в дополнении к этому следует отметить, что производительность данного канала связи выше в сравнении с протоколом HTTP, используемым в традиционной модели за счет существенного снижения объема служебной информации в сообщениях, например, отсутствия заголовков сообщений. Для организации работы WebSocket соединения требуется лишь один HTTP-запрос и HTTP-ответ (рукопожатие) для установления связи. Впоследствии передаваемые данные ограничиваются всего двумя байтами служебной информации при каждом запросе и ответе соответственно.

Преимуществом использования протокола WebSocket является увеличение продолжительности жизни канала связи, который может существовать значительно дольше в сравнении с другими псевдо-асинхронными технологиями из-за особенностей формата запросов и алгоритмов фильтрации прокси-серверов. Соответственно соединение может находиться в открытом состоянии сколь угодно долго и быть неактивным, а значит не требовать ресурсов для обработки. При этом в любой момент времени имеется возможность передавать данные по заранее созданному соединению.

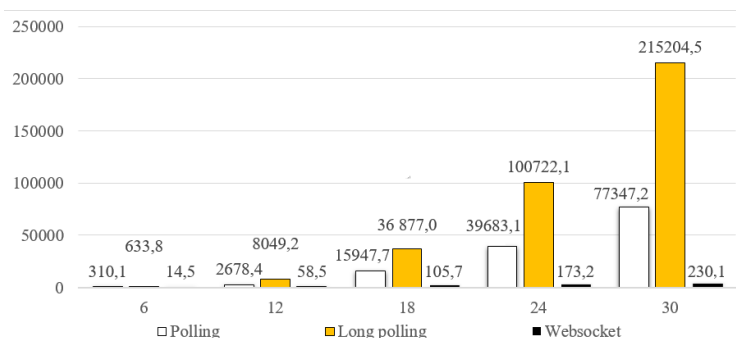


Рис. 1. Зависимость объема переданных данных при использовании различных подходов от времени

Для оценки эффективности полнодуплексной асинхронной технологии WebSocket по отношению к другим псевдо-асинхронным технологиям проведено испытание по сравнению количества передаваемых пакетов для поддержания актуальной информации в одностороннем приложении. В течении 30 минут производилось наблюдение следующих процессов: формирование данных на сервере каждые 5 минут, отправление polling запроса каждые 3 минуты, измерение трафика, проходящего через сервер каждые 6 минут. На рис. 1 представлена диаграмма, отражающая зависимость количества переданных пакетов от времени.

Стоит отметить, что использование протокола WebSocket налагает обязательства на построение клиент-серверных систем, например, обязательная поддержка технологии WebSocket браузерами, в качестве альтернативного подхода следует рассмотреть решение SignalR, обеспечивающее наиболее эффективный транспорт из доступных в конкретной ситуации, в этот пакет входят WebSocket, Events и LongPolling [3].

Подготовка данных. К эффективной подготовке данных относится их хранение на стороне сервера, т.е. насколько быстро их можно получить по запросу. Сперва следует уделить внимание технологиям middleware, оптимизирующим доступ и компоновку данных. Стандартным решением для односторонних веб-приложений является REST API, отличительная черта которого – ориентация на работу с данными с обеспечением различных вариаций базовых CRUD (create,

read, update, delete) операций. Каждый endpoint (точка входа) позволяет провести одну или несколько указанных операций над определенным набором данных. Узкое место при использовании данной концепции проявляется в случаях необходимости получения части определенного набора данных или измененного набора данных. Решения – два, в первом случае клиентское приложение получает избыточный набор данных, нагружая базу данных и канал связи, увеличивая скорость реакции клиентской части для пользователя; во-втором случае производится дублирование endpoint на сервере с изменениями, обеспечивающими передачу необходимого набора данных, но усложняющих модернизацию, эксплуатацию и сопровождение серверной составляющей [4].

Для обеспечения эффективной работы endpoint, отправки данных на клиентское приложение без избытков и отсутствия дублирования самих endpoint, следует использовать интеллектуальный механизм декларирования структуры данных и методов доступа к ним. Для достижения описанного результата клиентское приложение должно отправить специальный запрос на именной endpoint описывающий необходимые данные. Сервер при получении такого запроса применяет resolver функцию, обеспечивающую поиск, изменение и компоновку необходимых данных для отправки их на клиентскую составляющую. Механизмом работы с запросами по данному принципу является GraphQL, выпущенный компанией Facebook в 2015 году для обеспечения эффективной обработки запросов своей широкой аудиторией. На текущий момент существует большое количество библиотек для различных платформ, обеспечивающих интеллектуальную работу с запросами на сервере. Одна из отличительных особенностей подхода GraphQL – наличие возможности обеспечения ответа на запросы с подзапросами за счет их рекурсивной обработки с учетом возможности добавления аргументов в самих запросах, предоставляя клиентскому приложению графо-подобных структур неопределенного уровня вложенности [5]. Также преимуществом является возможность мутации данных для обеспечения всех CRUD операций на одном endpoint. При этом существует механизм организации подписки на данные – постоянной связи между клиентским и серверным приложениями, реагирующий на изменения в указанном наборе данных. В совокупности своих функций подход GraphQL позволяет отказаться от использования state management на клиентской стороне для SPA или существенно снизить масштаб реализации для PWA приложений. Существуют специальные библиотеки, способные работать с децентрализованным state management на клиентской и серверной составляющих, например Loona, представляющие связующее звено для state management на сервере в виде подписок на структуру данных GraphQL и локального state management Apollo на клиентской стороне, для распространенных фреймворков Angular и React специализирующихся на создании SPA и PWA приложений. Также следует отметить, что использование данного middleware позволяет поддерживать взаимодействие с неограниченным количеством источников данных [6].

Хранение данных. К эффективной подготовке данных относятся непосредственно средства хранения – базы данных. Начало 2000-х годов было ознаменовано теоремой CAP (Принцип был предложен Эриком Брюером), согласно которой распределённая система может обеспечить не более двух принципов из трех: согласованность (Consistency), доступность (Availability) и устойчивость к разделению (Partition tolerance), и чуть позже появлению термина Big Data, определяющего данные с тремя признаками: отсутствие структурированности, гипертрофированность физического объема и наличие динамики их прироста. Эти события привели к необходимости изменения систем хранения данных в пользу технологий, обеспечивающих высокую степень масштабируемости (Partition tolerance), являю-

шейся слабой стороной реляционной концепции наряду с модификацией её схемы ввиду строгой структурированности данных. Впоследствии появляется новый подход к построению распределенных систем BASE (Basically Available, Soft-state, Eventually consistent) – базовая доступность, неустойчивое состояние, согласованность в конечном счёте. Этот подход описывает свойства присущие NoSQL базам данных, основной целью которых является решение проблемы масштабирования данных и доступность в той или иной степени за счет атомарности и согласованности [7].

Подходы NoSQL также имеют несколько разновидностей: хранилище ключ-значение, документно-ориентированное, колоночное хранилище, хранилище на основе графов. Большого внимания в отношении информационных систем заслуживает разделяющая современный формат обмена данными в сети интернет, модель ключ-значение. При запросе клиентское приложение может получать объект данных (слабоструктурированные данные), целиком отвечающий решению конкретной задачи, при этом достигается высокая скорость предоставления информации по ключу, уменьшается количество запросов для получения разных структур данных, т.к. они не разделены по таблицам, а находятся внутри объекта, уменьшается объем передаваемых данных при условии отсутствия какого-либо свойства объекта эта информация исключается из самого объекта (ключ и значение) [8].

Следующая стадия оптимизации подготовки данных – использование баз данных, работающих с оперативной памятью (In-memory). Не теряя всех вышеописанных преимуществ NoSQL подхода, использование In-memory баз данных позволяет добиться увеличения скорости выполнения запроса до пяти порядков, при этом обеспечивает сохранность данных в условиях нестабильной работы сервера, отключения питания и т.п. Несмотря на то, что данные хранятся в оперативной памяти и доступ к ним осуществляется предельно быстро, транзакции (операции изменяющие данные) дополнительно дублируются в лог в ПЗУ, обеспечивая их сохранность. Для оптимизации лога транзакций производятся периодические снимки состояния (snapshot), являющиеся копией базы данных, позволяющие очищать лог транзакций до момента времени предшествующему снимку. Ярким примером данного решения может служить резидентная система управления базами данных Redis, являющаяся наиболее востребованной на текущий момент по версии ресурса db-engines [9].

Подход NoSQL также имеет и ряд недостатков, в первую очередь – это отсутствие возможности соответствия требованиям ACID (атомарность, согласованность, изолированность) к транзакциям. Существуют попытки создания оптимальной системы хранения данных, сочетающей преимущества обоих подходов (NewSQL), однако на текущий момент они не увенчались успехом. В случае решения задачи обеспечения высокой производительности для ряда данных и надежности транзакций для другой группы данных следует применять коллаборативные решения на основе традиционного и описанного подходов.

Вывод. Для поддержания коммерческой конкурентоспособности современных веб-приложений следует обеспечить высокую степень интерактивности их интерфейса на уровне нативного приложения. Чтобы обеспечить такой уровень быстродействия необходимо иметь эффективно организованный распределенный state management для снижения нагрузки на клиентское приложение. Для обеспечения своевременной поставки необходимой информации, по мере появления данных, без избыточности следует использовать полнодуплексные или комбинированные подходы для установления каналов связи. Сопряжение указанных подходов с интеллектуальным механизмом декларирования структуры данных и методов доступа к ним, предоставляющих возможность установления подписки на необходимые наборы данных, а также формирование графо-подобных структур не-

определенного уровня вложенности, позволяют добиться высоко уровня производительности обмена данными. Для обеспечения эффективности работы с информацией приведенный middleware должен предусматривать работу с коллаборативными масштабируемыми системами хранения слабосвязанных данных, использующих в качестве структурных единиц объектную модель, а в качестве носителя информации оперативное запоминающее устройство.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Валитова Н.Л., Кремлева Э.Ш., Кашафутдинов Р.К.* Перспективы применения технологии рwa для расширения e-learning систем на мобильные платформы // Образовательные технологии и общество. – 2020. – № 1. – С. 115-124.
2. *Гавриленко Ю.Ю., Саада Д.Ф., Ильющин Е.А., Намиот Д.Е.* Разработка прогрессивного web-приложения для системы управления push-уведомлениями // International Journal of Open Information Technologies. – 2018. – № 9.
3. *Долгов А.Н., Нуруллин Р.Ю.* Программная платформа node. JS // Достижения науки и образования. – 2016. – № 12 (13).
4. *Пустобаев А.И.* О сервисе рассылки push-уведомлений // International Journal of Open Information Technologies. – 2015. – № 6. – С. 13-20.
5. *Ильин Д.Ю., Никульчев Е.В., Колясников П.В.* Выбор технологических решений для разработки программного обеспечения распределенных информационных систем // Современные информационные технологии и ИТ-образование. – 2018. – № 2. – С. 344-354
6. *Анисимов В.И., Васильев С.А., Гридин В.Н.* Высокоскоростной полнодуплексный метод обмена данными для распределенных САПР // Известия ЮФУ. Технические науки. – 2018. – № 4 (198). – С. 38-47.
7. *Фереферов Е.С., Бычков И.В., Хмельнов А.Е.* Технология разработки приложений баз данных на основе декларативных спецификаций // Вычислительные технологии. – 2014. – № 5. – С. 85-100.
8. *Анисимов В.И., Васильев С.А., Тарасова О.Б., Чернов А.Н.* Сравнительный обзор методологий для создания интерактивных веб-приложений // Тр. конгресса по интеллектуальным системам и информационным технологиям «IS&IT'18». – 2018. – Т. 1. – С. 36-39.
9. *Игнатьев А.Ю.* Обзор технологии GraphQL // Молодой ученый. – 2019. – № 15 (253). – С. 22-24.
10. *Бодров М.Ю.* Современный подход к разработке одностраничных веб приложений // Современные тенденции развития науки и технологий. – 2015. – № 3-1. – С. 43-45.
11. *Erik Wittern, Alan Cha1, James C. Davis, Guillaume Baudart, Louis Mandel.* An Empirical Study of GraphQL Schemas // ICSOC19. – 2019. – P. 1-16.
12. *Романенко Е.В.* Место Big Data в современной социально-экономической жизни общества // Инновационная наука. – 2016. – № 4-3 (16). – С. 143-145.
13. *Абдыкаримова А.Т.* Big Data: проблемы и технологии // Международный журнал гуманитарных и естественных наук. – 2019. – № 5-1. – С. 16-18.
14. *Brewer Eric A.* A Certain Freedom: Thoughts on the CAP Theorem (англ.) // Proceeding of the XXIX ACM SIGACT-SIGOPS symposium on Principles of distributed computing. – N. Y.: ACM, 2010.
15. *Новиков Б.А.* Сравнительный анализ производительности SQL И NOSQL СУБД // КИО. – 2017. – № 4. – С. 48-63.
16. *Прамодкумар Дж. Садаладж, Мартин Фаулер.* NoSQL. Новая методология разработки нереляционных баз данных. – Вильямс, 2017.
17. *Савоськин И.В., Фирсов А.О.* Исследование способов применения nosql и реляционных баз данных // E-Scio. – 2019. – № 6 (33). – С. 41-49.
18. *Якушин А.Ю., Муковозов А.М., Исмоилов М.И.* Сравнительный анализ реляционной базы данных и документоориентированной NoSQL базы данных в разрезе их применения при создании локального чата/мессенджера // Инновационная наука. – 2018. – № 4. – С. 73-82.
19. *DB-Engines Ranking of Key-value Stores* // <https://db-engines.com> URL: <https://db-engines.com/en/ranking/key-value+store> (дата обращения: 12.05.2020).
20. *Соболь А.С.* Построение и адаптация NewSQL СУБД в частном «Облаке» // Сибирский журнал науки и технологий. – 2013. – № 4 (50). – С. 75-80.

REFERENCES

1. Valitova N.L., Kremleva E.Sh., Kashafutdinov R.K. Perspektivy primeneniya tekhnologii pwa dlya rasshireniya e-learning sistem na mobil'nye platformy [Prospects for applying pwa technology to expand e-learning systems to mobile platforms], *Obrazovatel'nye tekhnologii i obshchestvo* [Educational technologies and society], 2020, No. 1, pp. 115-124.
2. Gavrilenko Yu.Yu., Saada D.F., Il'yushin E.A., Namiot D.E. Razrabotka progressivnogo web-prilozheniya dlya sistema upravleniya push-uedomleniyami [Development of a progressive web application for push notification management system], *International Journal of Open Information Technologies*, 2018, No. 9.
3. Dolgov A.N., Nurullin R.Yu. Programmnaya platforma node. JS [The software platform of the node. JS], *Dostizheniya nauki i obrazovaniya* [Achievements of science and education], 2016, No. 12 (13).
4. Pustobaev A.I. O servise rassylki push-uedomleniy [About the push notification mailing service], *International Journal of Open Information Technologies*, 2015, No. 6, pp. 13-20.
5. Il'in D.Yu., Nikul'chev E.V., Kolyasnikov P.V. Vybory tekhnologicheskikh resheniy dlya razrabotki programmnogo obespecheniya raspredelennykh informatsionnykh sistem [Selection of technological solutions for software development of distributed information systems], *Sovremennye informatsionnye tekhnologii i IT-obrazovanie* [Modern information technologies and IT education], 2018, No. 2, pp. 344-354.
6. Anisimov V.I., Vasil'ev S.A., Gridin V.N. Vysokoskorostnoy polnodupleksnyy metod obmena dannymi dlya raspredelennykh SAPR [High-speed full-duplex data exchange method for distributed CAD systems], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2018, No. 4 (198), pp. 38-47.
7. Fereferov E.S., Bychkov I.V., Khmel'nov A.E. Tekhnologiya razrabotki prilozheniy baz dannykh na osnove deklarativnykh spetsifikatsiy [Technology for developing database applications based on declarative specifications], *Vychislitel'nye tekhnologii* [Computing technologies], 2014, No. 5, pp. 85-100.
8. Anisimov V.I., Vasil'ev S.A., Tarasova O.B., Chernov A.N. Sravnitel'nyy obzor metodologiy dlya sozdaniya interaktivnykh veb-prilozheniy [Comparative review of methodologies for creating interactive web applications], *Tr. kongressa po intellektual'nyim sistemam i informatsionnym tekhnologiyam «IS&IT'18»* [Proceedings of the Congress on intelligent systems and information technologies "IS&IT'18"], 2018, Vol. 1, pp. 36-39.
9. Ignat'ev A.Yu. Obzor tekhnologii GraphQL [Overview of GraphQL technology], *Molodoy uchenyy* [Young scientist], 2019, No. 15 (253), pp. 22-24.
10. Bodrov M.Yu. Sovremennyy podkhod k razrabotke odnostranichnykh veb prilozheniy [Modern approach to developing single page web applications], *Sovremennye tendentsii razvitiya nauki i tekhnologii* [Modern trends in the development of science and technology], 2015, No. 3-1, pp. 43-45.
11. Erik Wittern, Alan Cha1, James C. Davis, Guillaume Baudart, Louis Mandel. An Empirical Study of GraphQL Schemas, *ICSOC19*, 2019, pp. 1-16.
12. Romanenko E.V. Mesto Big Data v sovremennoy sotsial'no-ekonomicheskoy zhizni obshchestva [The place of Big Data in the modern socio-economic life of society], *Innovatsionnaya nauka* [Innovative science], 2016, No. 4-3 (16), pp. 143-145.
13. Abdykarimova A.T. Big Data: problemy i tekhnologii [Big Data: problems and technologies], *Mezhdunarodnyy zhurnal gumanitarnykh i estestvennykh nauk* [International journal of Humanities and natural Sciences], 2019, No. 5-1, pp. 16-18.
14. Brewer Eric A. A Certain Freedom: Thoughts on the CAP Theorem (англ.), *Proceeding of the XXIX ACM SIGACT-SIGOPS symposium on Principles of distributed computing*. N. Y.: ACM, 2010.
15. Novikov B.A. Sravnitel'nyy analiz proizvoditel'nosti SQL I NOSQL SUBD [Comparative analysis of SQL and NOSQL DBMS performance], *KIO* [KIO], 2017, No. 4, pp. 48-63.
16. Pramodkumar Dzh. Sadaladzh, Martin Failer. NoSQL. Novaya metodologiya razrabotki nerelyatsionnykh baz dannykh [NoSQL. A new methodology for developing non-relational databases]. Vil'yams, 2017.
17. Savos'kin I.V., Firsov A.O. issledovanie sposobov primeneniya nosql i relyatsionnykh baz dannykh [Research of ways to use nosql and relational databases], *E-Scio*, 2019, No. 6 (33), pp. 41-49.

18. Yakushin A.Yu., Mukovozov A.M., Ismoilov M.I. Sravnitel'nyy analiz relyatsionnoy bazy dannykh i dokumentoorientirovannoy NoSQL bazy dannykh v razreze ikh primeneniya pri sozdanii lokal'nogo chata/messendzhera [Comparative analysis of a relational database and a document-oriented NoSQL database in terms of their use in creating a local chat/messenger], *Innovatsionnaya nauka* [Innovative science], 2018, No. 4, pp. 73-82.
19. DB-Engines Ranking of Key-value Stores. Available at: <https://db-engines.com> URL: <https://db-engines.com/en/ranking/key-value+store> (accessed 12 May 2020).
20. Sobol' A.S. Postroenie i adaptatsiya NewSQL SUBD v chastnom «Oblake» [Building and adapting a NewSQL DBMS in a private "Cloud"], *Sibirskiy zhurnal nauki i tekhnologii* [Siberian journal of science and technology], 2013, No. 4 (50), pp. 75-80.

Статью рекомендовал к опубликованию д.ф.-м.н. А.С. Бугаев.

Гридин Владимир Николаевич – ФГБУН Центр информационных технологий в проектировании РАН; e-mail: info@ditc.ras.ru; 143003, Одинцово, ул. маршала Бирюзова, 7а; тел.: +74955960219; д.т.н.; профессор; научный руководитель.

Анисимов Владимир Иванович – Санкт-Петербургский государственный электротехнический университет; e-mail: vianisimov@inbox.ru; 197376, Санкт-Петербург, ул. Профессора Попова, 5; тел.: +78122343675; д.т.н.; профессор; г.н.с.

Васильев Сергей Алексеевич – e-mail: venom-gt@list.ru; к.т.н.; ассистент.

Gridin Vladimir Nikolayevich – Design Information Technology Center RAS (DITC RAS); e-mail: info@ditc.ras.ru; 7a, Marshal Biryuzov street, Odintsovo, 143003, Russia; phone: +74955960219; dr. of eng. sc.; professor; scientific director.

Anisimov Vladimir Ivanovic – Saint Petersburg State Electrotechnical University; e-mail: vianisimov@inbox.ru; 5, Prof. Popov street, Saint Petersburg, 197376, Russia; phone: +78122343675; dr. of eng. sc.; professor; chief researcher.

Vasilev Sergey Alexievich – e-mail: venom-gt@list.ru; assistant.

УДК 004.051

DOI 10.18522/2311-3103-2020-2-200-209

Д.Е. Чикрин, А.А. Егорчев, Д.В. Ермаков

МЕТОДОЛОГИЯ S.M.A.R.T.E.S.T. Н-GQM ДЛЯ КОНТРОЛИРУЕМОЙ ЭВОЛЮЦИИ СИСТЕМ ADAS

Вывод на массовый рынок транспортных средств (легковых и грузовых автомобилей) с высокой степенью автоматизации – уровня ADAS 3+ – ожидается с начала 2020-х годов. На текущий момент абсолютным большинством крупных автопроизводителей ведутся исследования и разработки в данном направлении, достаточно большое количество прототипов, предсерийных и серийных систем¹ уже продемонстрировано. Системы автоматизированного управления автомобилем – ADAS (advanced driver assistance systems) – представляют собой сложные аппаратно-программные комплексы, особенность которых состоит в неизменности ядра аппаратной платформы на протяжении одного или нескольких поколений автомобилей. При этом требуется обеспечить возможность обновления (эволюции) системы для исправления ошибок и расширения функциональности, особенно в условиях активно развивающихся сенсорных периферийных систем и программных алгоритмов. Для оценки и сопровождения разработки сложных систем применяется методология GQM (Goal, Question, Metric – цель, вопрос, метрика) и её модификации. Однако, область их применения ограничена исключительно программными продуктами; также не рассматриваются явно вопросы применения методологии GQM для анализа и сопровождения процессов эволюции сложных технических систем. В статье предлагается методоло-

¹ Tesla, Audi, Volvo, Komatsu и другие производители.

гия H-GQM (Hardware GQM) для проведения контролируемой эволюции сложных аппаратно-программных систем современной автомобильной техники. Представленная методология H-GQM базируется на методологии GQM и предназначена для аппаратно-программных комплексов с монолитным аппаратным ядром, модифицируемым программным ядром и периферией, удовлетворяющей принципу атомарности. Доказана применимость методологии GQM для анализа программно-аппаратных систем ADAS путем проведения процедуры гармонизации сущностей системы. Для формирования эволюционных целей предложена концепция целеполагания S.M.A.R.T.E.S.T, расширяющая методику формирования целей бизнес-процессов S.M.A.R.T. путем введения ограничений, полученных в результате гармонизации сущностей и описывающих требования к эволюционной способности системы. Формулирование фреймворка планов H-GQM рассматривается на примере систем ADAS, в рамках предложенной методологии сформирован масштабируемый шаблон целей, учитывающий специфику систем ADAS.

GQM; цель-вопрос-метрика; S.M.A.R.T.; ADAS; контролируемая эволюция технических систем; разработка сложных технических систем; аппаратно-программные системы.

D.E. Chickrin, A.A. Egorchev, D.V. Ermakov

EVOLVABLE ADAS: H-GQM S.M.A.R.T.E.S.T. APPROACH

The introduction to the mass market of vehicles with an ADAS 3+ level of automation is expected in the early 2020s. Currently, the vast majority of automakers conduct research in this field, a fairly large number of prototypes, pre-production and production systems have already been demonstrated. ADAS (advanced driver assistance systems) are complex hardware & software systems, the feature of which is that the core hardware platform remains unchanged for one or even several generations of vehicles (5–7 years). At the same time, the system should be able to transform and evolve to correct errors and expand functionality, especially due to active development of sensory peripheral systems and software algorithms. The GQM methodology and its modifications are used to support the development process of complex systems and evaluate them. However, these methodologies are limited exclusively to software products. Also, authors of these methodologies are not addressing explicitly the issues of applying the GQM methodology for analyzing and tracking the process of evolution of complex technical systems. This paper presents H-GQM (Hardware GQM) methodology for controllable evolution of complex automotive hardware & software systems. The H-GQM methodology is based on GQM and is aimed at hardware-software systems with a monolithic hardware core, a modifiable software core and atomic peripherals. Entity harmonization process is described to prove the applicability of the GQM for software-and-hardware systems analysis. S.M.A.R.T.E.S.T goal-setting concept is proposed for choice of evolutionary goals. This concept is based on S.M.A.R.T. criteria for the setting objectives of business processes and extended with harmonization and evolvability restrictions. The formulation of the H-GQM plan framework is provided using ADAS as an example. Within the framework of the proposed methodology, an ADAS-specific scalable target template has been formed.

GQM; S.M.A.R.T.; ADAS; controllable evolution of complex technical systems; complex technical systems development; software and hardware systems.

I. Введение. Все существующие и перспективные системы ADAS для массового автомобильного рынка имеют ту же специфику, что и любые другие сложные аппаратно-программные системы автомобильной техники [1, 2]. Указанная специфика заключается в аппаратной неизменности (**монолитности**) на протяжении одного или нескольких поколений (в горизонте 5–7 лет) центральной вертикально-интегрированной² платформы, максимально широко используемой автопроизводителем. В случае систем ADAS такого рода платформой является вычислительный блок – аппаратное ядро ADAS.

² Вертикальная интеграция – объединение в единой сущности блоков и подсистем различных уровней абстракции (операционной деятельности).

Задачи расширения функционала и исправления ошибок решаются за счет первоначальной избыточности возможностей и производительности монолитного аппаратного ядра путем:

1. Модификации программного ядра – «прошивки» вычислительного блока ADAS. Данное программное ядро относится к т. н. программно-интенсивным системам с критическим уровнем надежности (software intensive safety critical systems [3]).

2. Широкой варьированности характеристик и номенклатуры различных устройств и подсистем, представляющих собой горизонтально-интегрированную периферию³ платформы ADAS.

Требования по наличию нескольких взаимозаменяемых поставщиков на каждый тип устройства из периферии при работе с монолитным аппаратным ядром приводят также к:

1. Необходимости максимальной унификации устройств с периферии ADAS с точки зрения интерфейсов взаимодействия и подключения.

2. Функционированию каждого элемента из периферии ADAS по атомарному принципу – в виде «черного ящика» – для обеспечения заданных характеристик с максимальным абстрагированием от внутреннего устройства и образа функционирования изделия.

Смешанная методология оценки и сопровождения эволюционирующих аппаратно-программных систем H-GQM⁴, основывающаяся на методологии GQM [4, 5] и предлагается в указанной статье.

II. Эволюционирующие системы ADAS и методология H-GQM. Для фиксированного состава периферии и монолитного аппаратного ядра задачи внесения контролируемых изменений и дополнений в функционал различных подсистем ADAS могут быть проанализированы с помощью GQM, как программно-интенсивная система с критическим уровнем надежности [6]. В то же время процессы эволюции требований к системам ADAS⁵ в рамках одной и той же монолитной аппаратной платформы приводят к необходимости экстенсивного развития периферии.

Сам процесс изменения контролируемых изменений и дополнений одновременно в функционал программного ядра ADAS и состав периферии и представляет собой **процесс эволюции** ADAS между основными ревизиями основной вычислительной платформы – между сменами архитектуры и состава монолитного аппаратного ядра.

Таким образом, эволюционирующая система ADAS представляет собой совокупность фиксированного монолитного аппаратного ядра (вычислительной платформы ADAS); изменяемого программного ядра ADAS – программно-интенсивной системы с критическим уровнем надежности, и периферию с изменяемой номенклатурой и конфигурацией (в пределах определенных граничных условий) с атомарными блоками, подключаемыми по стандартным интерфейсам.

Оригинальная методология GQM представляет собой смешанную методологию оценки и сопровождения разработки программного обеспечения, использующую механизмы «сверху-вниз» с определением исходной цели исследуемого продукта и/или процесса как первоначального этапа проектирования и механизмы «снизу-вверх» для интерпретации анализируемой информации с точки зрения ранее определенного набора вопросов и метрик [4, 7]. Существует значительное ко-

³ Горизонтальная интеграция – объединение в единой сущности блоков и подсистем одного уровня абстракции (операционной деятельности).

⁴ H-GQM - Hardware GQM – предлагаемая методология GQM-типа, используемая для аппаратно-программных систем эволюционирующего типа.

⁵ Например, в процессе формирования различной конфигурации опций автопилотирования или обновления семейств автомобильной техники

личество модификаций и адаптаций методологии GQM для различных видов сложных – но только программных – систем [3, 6–10]. Также в литературе явно не рассматриваются вопросы применения методологии GQM для анализа и сопровождения именно процессов эволюции сложных технических систем.

Предлагаемая нами методология Н-GQM представляет собой методологию GQM, адаптированную и расширенную для использования в случае аппаратно-программных систем с изменяемым программным ядром, монолитным аппаратным ядром и атомарной периферией.

III. Гармонизация сущностей исследования Н-GQM. Как было указано ранее, рассматриваемые эволюционирующие системы ADAS представляют собой аппаратно-программную систему – совокупность модифицируемого программного ядра, монолитного аппаратного ядра (на котором выполняется программное ядро) и атомарной периферии.

Как было показано в [3], методология GQM полностью применима для систем, к классу которых принадлежит модифицируемое программное ядро ADAS – программно-интенсивным системам с критическим уровнем надежности. Для обеспечения правомерности применения методологии GQM (т.е. формирования предлагаемой методологии Н-GQM) необходимо:

1. Провести гармонизацию⁶ сущностей монолитного аппаратного ядра и атомарной периферии с классом программно-интенсивных систем критического уровня надежности.
2. Определить ограничения изучения гармонизируемых систем с учетом проведенных процедур гармонизации.

Обобщенная процедура гармонизации

По определению [3, 11], программно-интенсивными называются программные системы, взаимодействующие с совокупностью других систем, устройств и системных операторов. В свою очередь, системами критического уровня надежности⁷ называются системы, результат работы которых может прямо или косвенно определять безопасность жизни и здоровья человека-оператора [3]. Достаточно очевидным является вхождение систем ADAS в оба класса данных систем.

Задача гармонизации сущностей атомарной периферии и монолитного аппаратного ядра⁸ сводится к тому, чтобы они либо были также отнесены к классу программно-интенсивных систем критического уровня надежности, независимых от программного ядра – либо являлись частью единой программно-интенсивной системы критического уровня надежности на равных правах с остальными.

Последствия первого способа обеспечения гармонизации – необходимость создания многокомпонентной (векторной) цели, последствия второго – возможность GQM-анализа гармонизируемых сущностей лишь на том уровне абстракции (страте представления) [12], на котором было показано их вхождение в состав единой анализируемой системы. Гармонизация по второму принципу является более приемлемой, т. к. многокомпонентность цели противоречит принципу однозначности концепции S.M.A.R.T.

Гармонизация монолитного аппаратного ядра

Рассмотрим стратегию представления монолитного аппаратного ядра в виде вычислительной платформы, на которой запускается программный код модифицируемого программного ядра. В этом случае монолитное аппаратное ядро А является неотъемлемой частью программно-интенсивной системы критического уровня

⁶ Показать включение данных сущностей в анализируемый класс систем.

⁷ В оригинале - safety-critical computer systems.

⁸ По отношению к модифицируемому программному ядру, уже могущему быть проанализированным по методологии GQM.

надежности – программного ядра S – т.к. данная система не существует сама по себе, независимо от вычислительной платформы. Проведенная гармонизация накладывает ограничения, связанные с необходимостью обязательного учета интегральной аппаратной специфики вычислительной платформы A – производительность платформы, перечень имеющихся интерфейсов, доступность средств программной отладки, перечень допустимых языков программирования, условия эксплуатации – и невозможностью учета ее аппаратно-неспецифичных аспектов (например, возможностей отдельных элементов аппаратного ядра).

Гармонизация атомарной периферии

Рассмотрим стратегию представления атомарной периферии в виде системы высокой степени аддитивности, состоящей из элементов – черных ящиков, обеспечивающих прием и передачу стандартизованных информационных потоков и имеющих векторные управляющие воздействия стандартизованного типа. Данная система высокой степени аддитивности является идентичной совокупности аппаратно-неспецифичных⁹ программных функций F , имеющих фиксированные типы выходных данных *output*, входных данных *input* и управляющих переменных $x_1 \dots x_n$:

$$output = F(input, x_1, x_2, \dots, x_n).$$

Модифицируемое программное ядро, выполняемое на монолитной вычислительной платформе, может быть рассмотрено в эквивалентной стратегии – как совокупность программных функций FS . Таким образом, итоговая совокупность модифицируемого программного ядра, и атомарной периферии может быть представлена в виде программно-интенсивной системы критического уровня надежности с расширенным набором функций – $F + FS$ – и имеющей аппаратную специфичность A . Дополнительные условия, накладываемые данной гармонизацией (с учетом предыдущей) – необходимость рассмотрения чистого логического функционала периферии с общей аппаратной спецификой, определяемой аппаратной спецификой и сценариями применения аппаратного ядра¹⁰.

IV. S.M.A.R.T.E.S.T. - концепция формализация целей Н-GQM. Наиболее важным – т. н. концептуальным [4] – этапом формирования планов и фреймворков планов GQM является этап формирования целей. Цели GQM формируются из [5, 7]:

- ♦ Объекта целеполагания (object) – то, что изучается методологией GQM – какой продукт или процесс.
- ♦ Назначения исследования (purpose) – для чего проводится GQM-анализ – для целей лучшего понимания свойств объекта, для его сравнения с аналогами, для определения его эффективности и пр.
- ♦ Акцента исследования (quality focus) – какое конкретно свойство (совокупность свойств) объекта должны рассматриваться – например, цена, надежность и пр.
- ♦ Аспекта рассмотрения (viewpoint) – с чьей точки зрения должны быть получены ответы на вопросы исследования, кто будет интерпретировать результаты исследования.
- ♦ Контекста рассмотрения – окружающей среды для объекта целеполагания (environment) – в каких условиях рассматривается объект целеполагания и какие внешние факторы являются существенными для процесса исследований.

⁹ В связи с представлением отдельных блоков периферии в виде черных ящиков.

¹⁰ Т. е. в процессе Н-GQM-анализа вследствие проведенного процесса гармонизации могут должны быть определены общие эксплуатационные и физические ограничения для аппаратного ядра и атомарной периферии.

Принципиальной особенностью целей для GQM-подобных методик применительно к промышленным системам любого типа является требование по их однозначной и понятной интерпретации и легкой измеримости [13]. В случае нарушения данного требования применение методологии для оценки и сопровождения развития реальных систем становится невозможным в связи с противоречивостью получаемых результатов [7, 8]. Для корректного формирования целей, удовлетворяющих данным требованиям, авторами предлагается расширение методики формирования целей бизнес-процессов S.M.A.R.T. [13-15]:

- ◆ S. – Specific – Однозначность формулировки того, что должно быть достигнуто в процессе исследований.
- ◆ M. – Measurable – Может ли цель быть объективно и понятно измеренной.
- ◆ A. – Attainable – Может ли цель реально быть достигнутой.
- ◆ R. – Realistic – Будет ли реально проводиться работа по достижению цели, целесообразно ли ее достижение.

- ◆ T. – Timely – приемлемо ли время, в которое может быть достигнута цель.

Расширение концепции S.M.A.R.T. до S.M.A.R.T.E.S.T. обеспечивается за счет введения ограничений, полученных в результате гармонизации сущностей и анализа в процессе эволюции:

- E. – Evolvable – система рассматривается в процессе эволюции.

S. – Set of functions – программное обеспечение системы рассматривается с функциональной точки зрения: как ансамбль программных функций F, имеющих фиксированные типы выходных данных output, входных данных input и управляющих переменных $x_1 \dots x_n$:

$$output = F(input, x_1, x_2, \dots, x_n)$$

T. – Tolerance limits – рассматриваемая система должна рассматриваться в пределах тех физических и эксплуатационных ограничений, которые являются общими для монолитного аппаратного ядра и атомарной периферии.

V. Шаблон формирования целей H-GQM для ADAS. На базе сформулированной концепции S.M.A.R.T.E.S.T. и специфики систем ADAS [1, 2] в табл. 1 построен масштабируемый шаблон целей H-GQM с примерами основных целевых компонент:

Таблица 1

Масштабируемый шаблон целей H-GQM

S.M.A.R.T.E.S.T.	Object	Purpose	Quality Focus	Viewpoint	Environment
Шаблоны целевых компонент	Система ADAS Подсистемы ADAS, удовлетворяющие концепции H-GQM: подсистемы SLAM, V2X, Powertrain control, Autopilot, HMI и т. д.	Для сравнения в пределах эволюционного цикла (между ревизиями аппаратного ядра ADAS): 1. Стоимость 2. Надежности 3. Технологичности 4. Эффективности и т. д.	Что конкретно сравнивается в рамках purpose: 1. Если стоимость - то себестоимость, добавляемая ценность, ориентировочная выручка за опции и т. д. 2. Если эффективность - то вероятность принятия правильного решения, объем множества альтернатив решения, задержка принятия решения и т.д.	С чьей точки зрения: 1. Инженера-разработчика 2. Инженера-испытателя 3. Продавца 4. Покупателя 5. Оператора системы (водителя) и т. д.	В рамках чего рассматривается эволюция системы: Времени жизни платформы. Проекта по модернизации платформы. Проекта по оптимизации производства. Долгосрочного плана продаж и т. д.

S.M.A.R.T.E.S.T.	Object	Purpose	Quality Focus	Viewpoint	Environment
S - Specific	Должны быть однозначно определены - функционалом, структурой и спецификацией.	Должен быть выбран единственный показатель.	Должен быть выбран единственный показатель.	Должна быть выбрана единственная точка зрения.	Должен быть однозначно определен контекст и граничные условия среды (контекста) рассмотрения.
M - Measurable	Не применяется.	Показатель должен быть измеримым.	Показатель должен быть измеримым.	Не применяется.	Не применяется.
A - Attainable	Система должна содержать в себе все критически важные элементы для обеспечения функционирования.	Показатель должен реально использоваться и регулярно вычисляться.	Показатель должен реально использоваться и регулярно вычисляться.	Цель исследования должна быть актуальна для целевой аудитории.	Цель исследования должна быть актуальна в рамках контекста рассмотрения.
R - Realistic	В структуру системы должны входить только элементы с полным описанием и присутствующие на рынке.			Представители целевой аудитории должны участвовать в формировании плана H-GQM.	Контекст и граничные условия среды должны быть реалистичными.
T - Timely	Должны быть определены точные временные границы эволюционного процесса.	Должны быть определены точные временные границы исследований и получения результатов.		Цели и задачи целевой аудитории (применительно к исследованию) должны оставаться в большей части постоянными.	Временные границы эволюционного процесса должны соответствовать дорожной карте контекста рассмотрения.
E - Evolvable	Аппаратное ядро системы (включая интерфейсы) должно быть неизменно, изменяться могут лишь программное ядро, состав и спецификации периферии	Должен быть определен конкретный изменяемый параметр, например: Стоимость компонент. Трудоемкость производства. Качество распознавания. Дальность связи. Точность навигации и т. д.	Параметр должен быть конкретизирован, например: Стоимость компонент в партии от 1000 шт. Процент распознавания объектов в ночное время суток. Точность навигации в благоприятных климатических условиях и т. д.	Purpose должен иметь значение для целевой аудитории (например, изменение перечня опций и характеристик оборудования)	Purpose должен иметь значение в рамках контекста рассмотрения.
S - Set of functions	Функционал системы должен быть описан в виде списка функций (например - ACC, ABS, CDW и т. д.)	Измеряемому параметру должно соответствовать подмножество списка функций.		Исследуемое подмножество списка функций должно быть известно и иметь значение для целевой аудитории (например - перечень опций и возможности оборудования)	Исследуемое подмножество списка функций должно рассматриваться в рамках контекста рассмотрения (например - продаж автомобилей в максимальной комплектации)
T - Tolerance limits	Должны быть определены сценарии эксплуатации исследуемой системы (легковые автомобили, грузовая техника, карьерная техника и т.д.)	Должны быть определены унифицированные характеристики сценариев эксплуатации, имеющие значение в рамках рассматриваемой цели (климатические требования, ресурсные требования, энергопотребление и т. д.)	Должны быть детализированы унифицированные характеристики (температура от -30 до 55 гр. С; нагрузка на отказ 10000 моточасов и пр.)	Определенные детализированные унифицированные характеристики должны иметь значение для целевой аудитории и быть с ней согласованы.	Определенные детализированные унифицированные характеристики должны иметь значение в рамках контекста рассмотрения.

VI. Рекомендации по формированию ансамблей вопросов и метрик H-GQM. Для формирования ансамблей вопросов после формирования цели H-GQM авторами предлагается использование двух способов, рассмотренных ранее в литературе:

1. Использование т. н. листа абстракций GQM [7, 16], позволяющего сформировать вопросы путем опроса эксперта и целевой группы, определяемой аспектом рассмотрения.

2. Использование диаграммы онтологий анализируемой проблемы – POD¹¹ [17], обеспечивающую формирование полного множества вопросов по полностью сформированной цели GQM.

Выбор метрик по сформированным вопросам определяется из существующих обзоров и каталогов¹² в соответствии с экспертными оценками эффективности их сбора и репрезентативности.

Заключение. В представленной статье нами был представлен адаптированный к эволюционирующим аппаратно-программным системам и комплексам подход цель-вопрос-метрика – методология H-GQM. Для использования указанного подхода необходимо, чтобы аппаратное ядро системы в процессе эволюции было неизменно, а аппаратная периферия представлялась в виде «черных ящиков» со стандартизированными интерфейсами подключения. Ярким примером подобных систем являются системы автоматизированного управления автомобилями – системы ADAS, для которых и построен пример масштабируемого шаблона целей H-GQM, формализованный в таблице 1.

Центральным тезисом представленной методологии является разработка подхода формирования и верификации целей S.M.A.R.T.E.S.T., для формирования ансамблей вопросов и метрик предлагается использование листов абстракций GQM, диаграммы онтологий анализируемой проблемы и каталогов метрик по областям применения.

Дальнейшим направлением исследований в рамках серии публикаций планируется разработка и анализ полных планов H-GQM по отдельным подсистемам ADAS.

Благодарности. Работа выполнена за счет средств субсидии, выделенной в рамках государственной поддержки Казанского (Приволжского) федерального университета в целях повышения его конкурентоспособности среди ведущих мировых научно-образовательных центров.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Jo K. et al. Development of autonomous car–Part I: Distributed system architecture and development process // IEEE Transactions on Industrial Electronics. – 2014. – Vol. 61, No. 12. – P. 7131-7140.
2. Heinecke H. Automotive system design-challenges and potential // Design, Automation and Test in Europe. – IEEE, 2005. – P. 656-657.
3. Jayasri K., Seetharamaiah P. A GQM Based Approach towards the Development of Metrics for Software Safety // Journal of Computer Science. – 2015. – Vol. 11, No. 6. – P. 813.
4. Caldiera G, Basili V. R., Rombach H. D. The goal question metric approach // Encyclopedia of software engineering. – 1994. – P. 528-532.
5. Dow H. Goal Question Metric (GQM) and Software Quality. – 2007.
6. Olsson T., Runeson P. V-GQM: A feed-back approach to validation of a GQM study // Proceedings Seventh International Software Metrics Symposium. – IEEE, 2001. – P. 236-245.
7. Differding C., Hoisl B., Lott C. M. Technology package for the goal question metric paradigm. – 1996.
8. Lindstrom B. A software measurement case study using GQM // J. Lund Univ., USA. – 2004.
9. Basili V. et al. GQM⁺ strategies--aligning business strategies with software measurement // First international symposium on empirical software engineering and measurement (ESEM 2007). – IEEE, 2007. – P. 488-490.

¹¹ POD - Problem Ontology Diagram.

¹² Например, приведенных в источниках [18-20].

10. Solingen, van R., Berghout E. W. The Goal/Question/Metric Method: a practical guide for quality improvement of software development. – McGraw-Hill, 1999.
11. Alberico D. et al. Software System Safety Handbook-A Technical Managerial Team Approach // Joint Services Computer Resources Management Group. – 1999.
12. Мако Д., Месарович М., Такахара И. Теория иерархических многоуровневых систем. – М.: Мир, 1973.
13. Saar C., Safirstein M. Defining Goal-based Project Metrics. – Allstate Insurance Company, 2009.
14. TOWES Goal Planning Essentials. – Bow Valley College, 2015.
15. Doran G.T. There's a SMART way to write management's goals and objectives // Management review. – 1981. – Vol. 70, No. 11. – P. 35-36.
16. Solingen, van R. et al. Application of software measurement at Schlumberger RPS: Towards enhancing GQM // Proceedings of the 6th European Software Control and Metrics (ESCOM) Conference. – 1995. – P. 17-19.
17. Seidita V. et al. Simulation goals and metrics identification // 2016 Federated Conference on Computer Science and Information Systems (FedCSIS). – IEEE, 2016. – P. 1491-1494.
18. Babu P. C., Prasad A. N., Sudhakar D. Software Complexity Metrics: A Survey // International Journal. – 2013. – Vol. 3, No. 8.
19. Stephens R. Beginning software engineering. – John Wiley & Sons, 2015.
20. Cardoso J. Quality Metrics for Business Processes. – 2011.

REFERENCES

1. Jo K. et al. Development of autonomous car-Part I: Distributed system architecture and development process, *IEEE Transactions on Industrial Electronics*, 2014, Vol. 61, No. 12, pp. 7131-7140.
2. Heinecke H. Automotive system design-challenges and potential, *Design, Automation and Test in Europe*. IEEE, 2005, pp. 656-657.
3. Jayasri K., Seetharamaiah P. A GQM Based Approach towards the Development of Metrics for Software Safety, *Journal of Computer Science*, 2015, Vol. 11, No. 6, pp. 813.
4. Caldiera G, Basili V. R., Rombach H. D. The goal question metric approach, *Encyclopedia of software engineering*, 1994, pp. 528-532.
5. Dow H. Goal Question Metric (GQM) and Software Quality, 2007.
6. Olsson T., Runeson P. V-GQM: A feed-back approach to validation of a GQM study, *Proceedings Seventh International Software Metrics Symposium*. IEEE, 2001, pp. 236-245.
7. Differding C., Hoisl B., Lott C. M. Technology package for the goal question metric paradigm, 1996.
8. Lindstrom B. A software measurement case study using GQM, *J. Lund Univ., USA*, 2004.
9. Basili V. et al. GQM⁺ strategies-aligning business strategies with software measurement, *First international symposium on empirical software engineering and measurement (ESEM 2007)*. IEEE, 2007, pp. 488-490.
10. Solingen, van R., Berghout E. W. The Goal/Question/Metric Method: a practical guide for quality improvement of software development. McGraw-Hill, 1999.
11. Alberico D. et al. Software System Safety Handbook-A Technical Managerial Team Approach, *Joint Services Computer Resources Management Group*, 1999.
12. Mako D., Mesarovich M., Takahara I. Teoriya ierarkhicheskikh mnogourovnevnykh system [Theory of hierarchical multilevel systems]. Moscow: Mir, 1973.
13. Saar C., Safirstein M. Defining Goal-based Project Metrics. Allstate Insurance Company, 2009.
14. TOWES Goal Planning Essentials. Bow Valley College, 2015.
15. Doran G.T. There's a SMART way to write management's goals and objectives, *Management review*, 1981, Vol. 70, No. 11, pp. 35-36.
16. Solingen, van R. et al. Application of software measurement at Schlumberger RPS: Towards enhancing GQM, *Proceedings of the 6th European Software Control and Metrics (ESCOM) Conference*, 1995, pp. 17-19.
17. Seidita V. et al. Simulation goals and metrics identification, *2016 Federated Conference on Computer Science and Information Systems (FedCSIS)*. IEEE, 2016, pp. 1491-1494.

18. Babu P. C., Prasad A. N., Sudhakar D. Software Complexity Metrics: A Survey, *International Journal*, 2013, Vol. 3, No. 8.
19. Stephens R. Beginning software engineering. John Wiley & Sons, 2015.
20. Cardoso J. Quality Metrics for Business Processes. 2011.

Статью рекомендовал к опубликованию к.т.н. П.А. Кокунир.

Чикрин Дмитрий Евгеньевич – Казанский (Приволжский) федеральный университет; e-mail: dmitry.kfu@gmail.com; 420008, Казань, ул. Кремлевская, 18; к.т.н.; доцент.

Егорчев Антон Александрович – e-mail: eanton090@gmail.com; научный сотрудник.

Ермаков Дмитрий Владимирович – e-mail: DVErmakov@kpfu.ru; аспирант.

Chickrin Dmitriy Evgen'evich – Kazan Federal University; e-mail: dmitry.kfu@gmail.com; 18, Kremlyovskaya street, Kazan, 420008, Russia; cand. of eng. sc.; associate professor.

Egorchev Anton Alexandrovich – e-mail: eanton090@gmail.com; researcher.

Ermakov Dmitriy Vladimirovich – e-mail: DVErmakov@kpfu.ru; postgraduate student.

УДК 654.02

DOI 10.18522/2311-3103-2020-2-209-218

А.И. Рыбаков, Р.Е. Кротов, С.А. Кокин

АДАПТАЦИЯ ИНФОРМАЦИОННО-ТЕХНИЧЕСКИХ ХАРАКТЕРИСТИК К ПОСТОЯННО ИЗМЕНЯЮЩИМСЯ ПАРАМЕТРАМ ИОНОСФЕРНОГО РАСПРОСТРАНЕНИЯ

Целью исследовательской работы явилось изучение и выбор существующих вариантов адаптации по параметрам передачи, для снижения влияния недостатков коротковолновой радиолонии, целесообразно максимально эффективно использовать методы цифровой обработки сигналов. По результатам характеристик аналогово-цифровых преобразователей (АЦП), стало исследования доступных аппаратных средств, для построения протяженных радиолоний, был сделан вывод о том, что с ростом производительности ПЛИС, на которых реализуется цифровая обработка сигналов и технических представляется возможной реализация технологии создания активной антенной решетки (ААР), состоящей из N -го количества независимых антенных модулей, что и является концептуальной задачей в решении вопроса адаптации информационно-технических характеристик к постоянно изменяющимся параметрам ионосферного прохождения, для более энергоэффективного подхода к проектированию системы ионосферной радиосвязи. Повышение производительности радиосистемы путём совершенствования протоколов связи, решение вопроса оптимального по загруженности канала от времени формирования и приема сигналов. Основная идея такой ААР состоит в оцифровке или генерации высокочастотного сигнала в непосредственной близости от антенны, в составе антенных модулей. Указанные результаты позволяют заменить отдельно настраиваемые радиоприемники и трансиверы, построенные по сложной супергетеродинной схеме, на ограниченное число доступных аппаратных блоков, работающих под управлением ПО модели программно-конфигурируемого радиоканала. В следующей работе планируется провести исследования по оценке прохождения сигналов OFDM через многолучевые каналы связи с замираниями Релея и Райса. Получаемая модель позволит оценить помехоустойчивость при различной длине циклического префикса OFDM символа и пронаблюдать за поведением сигнального созвездия при воздействии различных нестабильностей.

Радиотрасса; радиосвязь; дециметровые волны; цифровая обработка сигналов; аналогово-цифровой преобразователь (АЦП); активная антенная решетка (ААР); коротковолновый (КВ) диапазон; уровень сигнала; ионосфера; ионосферное прохождение; модель распространения.

A.I. Rybakov, R.E. Krotov, S.A. Kokin

ADAPTATION OF INFORMATION AND TECHNICAL CHARACTERISTICS TO THE CONSTANTLY CHANGING PARAMETERS OF IONOSPHERIC PROPAGATION

The aim of the research work was to study and select the existing adaptation options for transmission parameters, in order to reduce the influence of shortcomings of the short-wave radio line, it is advisable to use the methods of digital signal processing as efficiently as possible. According to the results of the characteristics of analog-to-digital converters (ADCs), it became a study of the available hardware for constructing extended radio links, it was concluded that with an increase in the performance of FPGAs on which digital signal processing is implemented, it is possible to implement the technology of creating an active antenna array (AAR), consisting of the Nth number of independent antenna modules, which is a conceptual task in solving the issue of adapting information and technical characteristics to constantly changing parameters of ionospheric transmission, for a more energy-efficient approach to designing an ionospheric radio communication system. Improving the performance of the radio system by improving communication protocols, solving the problem of optimal channel load from the time of formation and reception of signals. The main idea of such an AAR is to digitize or generate a high-frequency signal in the immediate vicinity of the antenna, as part of the antenna modules. The indicated results make it possible to replace separately tuned radios and transceivers built according to a complex super-heterodyne circuit with a limited number of available hardware units operating under software model of a software-configurable radio channel. In the next work, it is planned to conduct studies to assess the passage of OFDM signals through multipath communication channels with fading of Rayleigh and Rice. The resulting model will allow us to evaluate the noise immunity at different lengths of the cyclic prefix of the OFDM symbol and to observe the behavior of the signal constellation under the influence of various instabilities.

Radio path; radio communication; decameter waves; digital signal processing; analog-to-digital converter (ADC); active antenna array (AAR); short-wave (HF) range; signal level; ionosphere; ionospheric propagation; propagation model.

Введение. В продолжительном прогнозе, развитие экономики Российской Федерации важную роль отводится на освоение территорий Арктики, Сибири и Дальнего Востока, имеющих огромный ресурсный потенциал.

При этом, важным элементом экономических районов остается инфокоммуникационная инфраструктура [1].

На сегодняшний день крупнейшим оператором магистральных сетей связи является ПАО «Ростелеком», обладающий мощной магистральной сетью связи в стране, по которому можно судить о развитии инфраструктуры в целом.

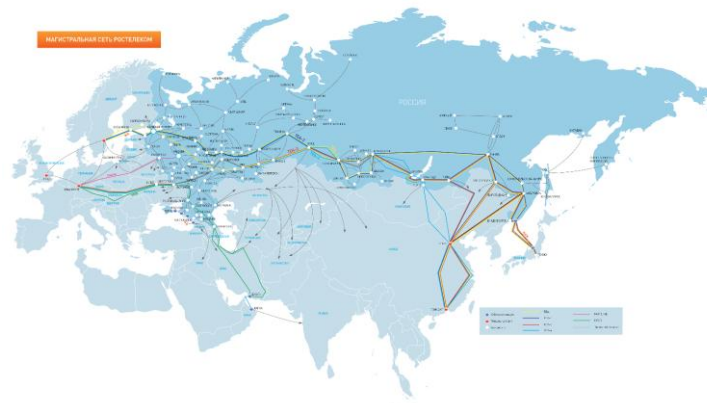


Рис. 1. Магистральная сеть ПАО «Ростелеком» [2]

Можно отметить, что процедуры адаптации по рабочей частоте, мощности РПДУ, выбору сигнально-кодовой конструкции не представляют из себя ничего нового и описаны в зарубежных стандартах связи, таких, как MIL-STD-188-110(A,B,C), MIL-STD-188-141C(A,B,C), HFDL (ARINC 635), STANAG 4538. В частности, реализованы они на таком оборудовании, как: Rohde & Schwarz M3SR Series4100, Codan 3012, 3112, 3212, RM50e HF Data modems, Rockwell Collins MDM Q9604, Harris Corp. RF-5800H-MP, RF-7800H-MP, AN/PRC-150(C), Rapid Mobile RM6, RM6-A, RM8, TC4.

В Российской Федерации единых стандартов адаптивных линий связи нет, среди отечественных разработок можно выделить следующие адаптивные комплексы радиосвязи: ААКТС «Пирс» (ПАО «РИМР»), «МКТС-1» (АО ОНИИП), «КТС ААРС» (ФГУП НИИР), «Ангара-5М» (АО «Егоршинский радиозавод»), «Нептун» (АО НИИ АСикС «Нептун»), «Антей» (АО «Концерн «Созвездие»). [10]

По результатам исследования доступных аппаратных средств, для построения протяженных радиолиний, можно сделать вывод, что с ростом производительности ПЛИС, на которых реализуется цифровая обработка сигналов и технических характеристик аналогово-цифровых преобразователей (АЦП), представляется возможным реализация технологии создания активной антенной решетки (ААР), состоящей из N-го количества независимых антенных модулей, что и является концептуальной задачей в решении вопроса адаптации информационно-технических характеристик к постоянно изменяющимся параметрам ионосферного прохождения.

Основная идея такой ААР состоит в оцифровке или генерации высокочастотного сигнала в непосредственной близости от антенны, в составе антенных модулей.

Приёмный сегмент комплекса радиосвязи представляется следующим образом (рис. 3):

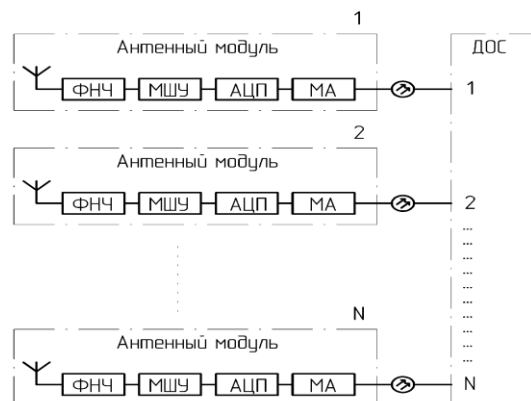


Рис. 3. Приёмный сегмент комплекса радиосвязи

Получается, ВЧ сигнал, поступающий с антенны, фильтруется ФНЧ от бытовых помех и мощных вещательных станций, находящихся на частоте менее 1,5 МГц.

Малошумящий усилитель (МШУ) усиливает сигнал, наведенный на антенну.

Блок АЦП оцифровывает всю полосу от 1,5 МГц до 30 МГц. [11].

После чего, оцифрованный сигнал, при помощи медиаконвертера (МА), преобразуется в оптический и по ВОЛС передается в диаграмма-образующую систему (ДОС), основным вычислительным средством которого является ПЛИС [12].

Поскольку оцифрованный ВЧ сигнал передается по независимым каналам, математический аппарат позволяет сдвигать фазу сигнала, поступающего с каждого антенного элемента, тем самым, формировать диаграмму направленности всей антенной решетки.

При таком решении задачи, мы получим следующие возможности:

- 1) реализацию антенной системы с многолепестковой диаграммой направленности [13];
- 2) использование разнесенного по частотам и пространству приёма [14, 15];
- 3) оцифровывая весь КВ диапазон, возможно оперативно оценить качество радиолинии по ряду доступных частот [16];
- 4) реализацию приёма по всему запасу доступных частот, для значительного повышения скорости приёма данных;
- 5) исключает потери в аналоговых трактах, коммутаторах и разъёмных соединениях, минимизирует регламентные работы по поддержанию открытых электрических соединений;
- 6) минимизация воздействия электромагнитных помех на весь аналоговый приёмный тракт [17];
- 7) в случае выхода из строя части элементов сети, программными методами возможно адаптировать АР под новые условия и приёмный комплекс останется в работоспособном состоянии;
- 8) в случае реализации антенных с двумя поляризациями, возможно добиться снижения вероятности битовой ошибки (E) за счет поляризационного разнесения, тем самым, снизив влияние рэлеевских замираний. На рис. 4 представлена вероятность битовой ошибки, в зависимости от отношения сигнал/шум (ОСШ). [18–20].
- 9) в режиме работы приёмных антенных модулей, в качестве самостоятельных антенн, реализуема система разнесенного приёма. На рис. 5 представлена вероятность битовой ошибки (E), в зависимости от ОСШ, для количества ветвей разнесения (M) 1...8. [21, 22].

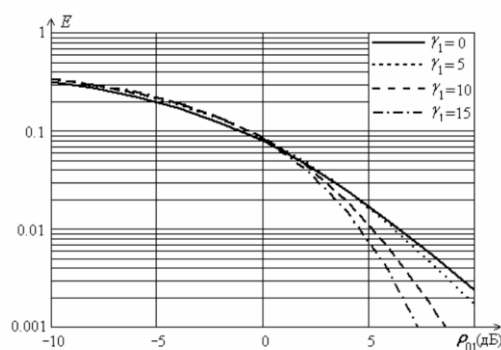


Рис. 4. Вероятность битовой ошибки (E), в зависимости от ОСШ (ρ_{01})

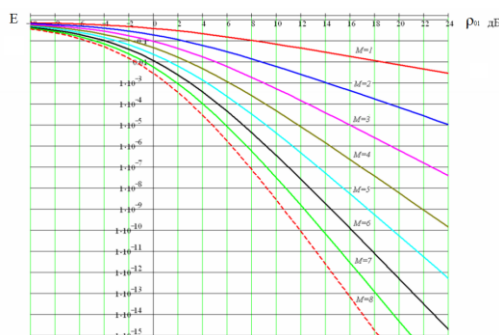


Рис. 5. Соотношения вероятностей

Передающий сегмент радиоцентра:

Передающая антенная решетка реализуется методом прямого цифрового синтеза с формированием высокочастотного сигнала, непосредственно, в каждом передающем антенном модуле. Передающие антенные модули строятся по схеме, представленной на рис. 6:

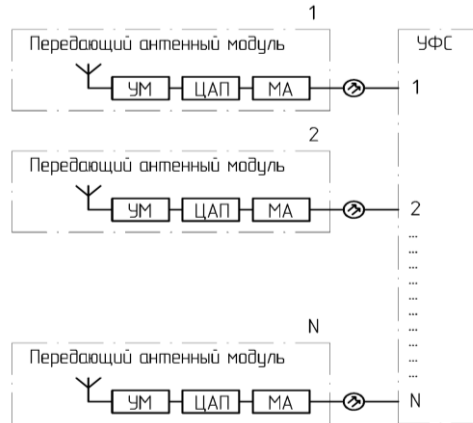


Рис. 6. Передающий сегмент радиоцентра

Для каждого передающего антенного модуля сигнал формируется устройстве формирования сигналов (УФС) индивидуально, с заданной фазой, и передается по ВОЛС. В медиаконвертере (МА) оптический сигнал преобразуется в цифровой и передается блоку ЦАП. Блок ЦАП формирует аналоговый высокочастотный сигнал, усиливаемый усилителем мощности (УМ), подается непосредственно на антенну [23].

При формировании высокочастотного сигнала, непосредственно, в каждом антенном модуле, предоставляются следующие преимущества:

- 1) исключение потерь ВЧ сигнала на протяженных фидерных линиях [24];
- 2) с экономической точки зрения, освобождает от необходимости постройки дорогостоящих мощных фидерных трактов;
- 3) передающие антенные модули могут использовать усилители малой мощности и «набирать» мощность, в направлении корреспондента, путем пространственного сложения мощностей [25];
- 4) формирование нескольких каналов связи с разными рабочими частотами и направлениями излучения [26].

Заключение. Экономическая целесообразность выражается в достаточной реализации комплексов связи, в масштабах, требуемых для решения задач обеспечения цифровой связью конкретного региона. Количество приёмных и передающих антенных элементов выбирается, исходя из основных требуемых параметров радиоцентра [27].

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Гурлев И.В. Развитие волоконно-оптических линий связи как средства управления и обеспечения национальной безопасности // Вестник Евразийской науки. – 2018. – № 4 (10). – С. 42.
2. Воробьев О.В., Рыбаков А.И. Вариант реализации двунаправленной связи в СМС (системе метеорной связи). Описание программно-аппаратного комплекса СМС // Матер. VII Международную научно-технической и научно-методической конференции «Актуальные проблемы инфокоммуникаций в науке и образовании». – 2017. – С. 128-136.

3. Bilitza D., Altadill D., Zhang Y., Mertens Ch., Truhlik V., Richards Ph., McKinnell L-A., Reinisch B. The International Reference Ionosphere 2012 – a model of international collaboration // Journal of Space Weather and Space Climate. – 2014. – Vol. 4. A07. – P. 1-12.
4. Березин Ю.В., Вылегжанин И.С. Декаметровые ионосферные линии радиосвязи с высокой пропускной способностью // Радиотехника. – 2005. – № 1. – С. 6-12.
5. Иванов В.С., Никитин Б.К., Пирмагомедов Р.Я. Строительство ВОЛС. // Современные технологии и организация. Ч. 1. – СПб.: СПбГУТ им. М.А. Бонч-Бруевича, 2015. – 71 с.
6. Богданов [и др.]. Современное состояние и перспективы развития радиопередающих устройств и радиомодемов профессиональной ДКМВ радиосвязи в ОНННП // Успехи современной радиоэлектроники. – 2011. – № 11. – С. 24-31.
7. Браницкий А.В., Ким В.Ю., Полиматиди В.П., Пучков В.А. Методика измерения доплеровского смещения частоты многолучевого сигнала // Матер. VI Международную научно-технической и научно-методической конференции «Актуальные проблемы инфокоммуникаций в науке и образовании». – 2016. – С. 126-132.
8. Климов И.З., Котысов А.Н., Чувашов А.М. Исследование вариантов построения широкополосных систем связи // Цифровая обработка сигналов и ее применение - DSPA-2012: Тр. Российского научно-технического общества радиотехники, электроники и связи им. А.С. Попова. Доклады 14-й Международной конференции. – 2012. – С. 435-439.
9. Смаль М.С. Бестестовые способы оценивания состояния коротковолнового радиоканала в адаптивных радиоприемах: дисс. ... канд. техн. наук: 05.12.13; Место защиты: ГУАП. – СПб., 2018. – 147 с.
10. Березин Ю.В., Вылегжанин И.С., Якушева М.А. Адаптивная по поляризации сеть коротковолновой ионосферной радиосвязи с селективным возбуждением электромагнитных волн в ионосфере // Тр. X Всероссийской школы-семинара «Физика и применение микроволн». – 2005. – С. 29-31.
11. The Art of Electronics [Horowitz, Paul, Hill, Winfield]. – 2015. – P. 900-910.
12. Никитин М.Л. Особенности построения широкополосного программно-определяемого радиомодема с использованием аппаратных возможностей ПЛИС // Интеллектуальные системы в производстве. – 2015. – № 3 (27). – С. 59-62.
13. Лучин Д.В., Плотников А.М., Трофимов А.П., Юдин В.В. Компактные приемные антенны для поляризационно-избирательного приема в составе систем радиомониторинга // Электросвязь. – 2015. – № 8.
14. Ермолаев В.Т., Маврычев Е.А. Флакман А.Г. Эффективность систем связи с антенными решетками в условиях рассеивающей среды // Успехи современной радиоэлектроники. – 2003. – № 3. – С. 41-48.
15. Luchin D., Plotnikov A., Trofimov D., Filippov D. Problems of implementation of ground biorthogonal and triorthogonal antenna systems // XI Междунар. IEEE Сибирская конф. по управлению и связи (SIBCON-2015). – Омск, 2015.
16. Флакман А.Г. Адаптивная пространственная обработка сигналов в многоканальных информационных системах: дисс. ... д-ра наук. – Нижний Новгород: Нижегородский госуниверситет им. Н.И. Лобачевского, 2004.
17. Simon K., Alouini M.-S. Digital Communication over Fading Channels: A Unified Approach to Performance Analysis. – N. Y.: John Wiley&Sons, 2000. – 544 p.
18. Ступницкий М.М., Лучин Д.В. Потенциал КВ-радиосвязи - для создания цифровой экосистемы России // Электросвязь. – 2018. – С. 49-54.
19. Ермолаев В.Т., Маврычев Е.А. Флакман А.Г. Уменьшение вероятности битовой ошибки при параллельной передаче информации в ММО системе // Известия Вузов. Радиофизика. – 2003. – Т. 46, № 3. – С. 251-260.
20. Svischenko Nikolay V. Special Integral Functions Used in Wireless Communications Theory. – Singapore: World Scientific Publishing Company, 2014.
21. Савищенко Н.В. и др. Расчет вероятности битовой и символьной ошибок для канала связи, при приеме сигнальных конструкций стандарта DVB-S2 Н. В. // Информация и Космос. – 2015. – № 1. – С. 9-15.
22. Digital Communications: Fundamentals and Applications. – 2nd ed. [Bernard Sklar]. – P. 961-975.

23. Ступницкий М.М., Лучин Д.В. Потенциал КВ-радиосвязи - для создания цифровой эко-системы России // Электросвязь. – 2018. – С. 49-54.
24. Богданов А.В. [и др.]. Об оптимизации требований к передающим комплексам радиоли-ний высокоскоростной передачи данных диапазона ДКМВ // Успехи современной ра-диоэлектроники. – 2011. – № 7. – С. 10-16.
25. Жидяев А.В., Копысов А.Н., Богданов А.А., Савельев А.В., Никитин М.Л. Исследование энер-гетических характеристик сигналов, применяемых для передачи данных по декаметровому каналу // Вестник ИжГТУ им. М.Т. Калашникова. – 2015. – № 3 (67). – С. 85-88.
26. Bridger Wray W., Ruiz Mark D. Advisors: Aruna Apte, James B. Greene. Naval Postgraduate School Monterey, California "Total Ownership Cost Reduction Case Study: AEGIS Radar Phase Shifters" December 2006.
27. Ступницкий М.М., Харитонов Н.И., Девяткин Е.Е. Инфокоммуникационная инфра-структура цифровой экономики: задачи отраслевого института // Электросвязь. – 2018. – № 4. – С. 70-76.

REFERENCES

1. Gurlev I.V. Razvitie volokonno-opticheskikh liniy svyazi kak sredstva upravleniya i obespecheniya natsional'noy bezopasnosti [Development of fiber-optic communication lines as a means of managing and ensuring national security], *Vestnik Evraziyskoy nauki* [Bulletin of Eurasian science], 2018, No. 4 (10), pp. 42.
2. Vorob'ev O.V., Rybakov A.I. Variant realizatsii dvunapavlennoy svyazi v SMS (sisteme meteornoj svyazi). Opisanie programmno-apparatnogo kompleksa sms [A variant of imple-menting bidirectional communication in SMS (meteor communication system). Description of the SMS software and hardware complex], *Mater. VII Mezhdunarodnuyu nauchno-tekhnicheskoy i nauchno-metodicheskoy konferentsii «Aktual'nye problemy infokommunikatsiy v nauke i obrazovanii»* [Materials of the VII International scientific-technical and scientific-methodical conference "Actual problems of Infocommunications in science and education"], 2017, pp. 128-136.
3. Bilitza D., Altadill D., Zhang Y., Mertens Ch., Truhlik V., Richards Ph., McKinnell L-A., Reinisch B. The International Reference Ionosphere 2012 – a model of international collabora-tion, *Journal of Space Weather and Space Climate*, 2014, Vol. 4. A07, pp. 1-12.
4. Berezin Yu.V., Vylegzhanin I.S. Dekametrovye ionosfernye linii radiosvyazi s vysokoy propusknoy sposobnost'yu [Decameter ionospheric radio communication lines with high throughput], *Radiotekhnika* [Radiotechnics], 2005, No. 1, pp. 6-12.
5. Ivanov V.S., Nikitin B.K., Pirmagomedov R.Ya. Stroitel'stvo VOLS [Construction of VOLS], *Sovremennye tekhnologii i organizatsiya* [Modern technologies and organization]. Part 1. Saint Petersburg: SPbGUT im. M.A. Bonch-Bruевича, 2015, 71 p.
6. Bogdanov [i dr.]. Sovremennoe sostoyanie i perspektivy razvitiya radioperedayushchikh ustroystv i radiomodemov professional'noy DKMV radiosvyazi v ONIP [Current state and prospects of development of radio transmitting devices and radio modems of professional dcmv radio communication in onip], *Uspekhi sovremennoy radioelektroniki* [Success of mod-ern radio electronics], 2011, No. 11, pp. 24-31.
7. Branitskiy A.V., Kim V.Yu., Polimatidi V.P., Puchkov V.A. Metodika izmereniya doplerovskogo smeshcheniya chastoty mnogoluchevogo signala [Method for measuring the Doppler frequency offset of a multipath signal], *Mater. VI Mezhdunarodnuyu nauchno-tekhnicheskoy i nauchno-metodicheskoy konferentsii «Aktual'nye problemy infokommunikatsiy v nauke i obrazovanii»* [Materials of the VI International scientific-technical and scientific-methodical conference "Actual problems of Infocommunications in science and education"], 2016, pp. 126-132.
8. Klimov I.Z., Kopysov A.N., Chuvashov A.M. Issledovanie variantov postroeniya shiroko-polosnykh sistem svyazi [Research of variants of construction of wide-band communication systems], *Tsifrovaya obrabotka signalov i ee primeneniye - DSPA-2012: Tr. Rossiyskogo nauchno-tekhnicheskogo obshchestva radiotekhniki, elektroniki i svyazi im. A.S. Popova. Doklady 14-y Mezhdunarodnoy konferentsii* [Digital signal processing and its application- DSPA-2012: Proceedings of the Russian scientific and technical society of radio engineering, electronics and communications named after A. S. Popov. Reports of the 14th International conference], 2012, pp. 435-439.

9. Smal' M.S. Bestestovye sposoby otsenivaniya sostoyaniya korotkovolnovogo radiokanala v adaptivnykh radiolinnykh: disc. ... kand. tekhn. nauk: 05.12.13 [Best-selling methods for evaluating the state of a short-wave radio channel in adaptive radio lines: cand. of eng. dc. diss.]; Place of protection: GUAP. Saint Petersburg, 2018, 147 p.
10. Berezin Yu.V., Vylegzhanin I.S., Yakusheva M.A. Adaptivnaya po polarizatsii set' korotkovolnovoy ionosfernoy radiosvyazi s selektivnym vzbuzhdeniem elektromagnitnykh voln v ionosfere [Adaptive polarization network of short-wave ionospheric radio communication with selective excitation of electromagnetic waves in the ionosphere], *Tr. X Vserossiyskoy shkoly-seminara «Fizika i primeneniye mikrovoln»* [Proceedings of the X all-Russian school-seminar "Physics and application of microwaves"], 2005, pp. 29-31.
11. The Art of Electronics, Horowitz, Paul, Hill, Winfield, 2015, pp. 900-910.
12. Nikitin M.L. Osobennosti postroyeniya shirokopolosnogo programmno-opredelyaemogo radiomodema s ispol'zovaniem apparatnykh vozmozhnostey PLIS [Features of building a broadband software-defined radio modem using hardware capabilities of the FPGA], *Intellektual'nye sistemy v proizvodstve* [Intelligent systems in production], 2015, No. 3 (27), pp. 59-62.
13. Luchin D.V., Plotnikov A.M., Trofimov A.P., Yudin V.V. Kompaktnye priemnye anteny dlya polarizatsionno-izbiratel'nogo priema v sostave sistem radiomonitoringa [Compact receiving antennas for polarizing selective reception as part of radio monitoring systems], *Elektrosvyaz'* [Telecommunications], 2015, No. 8.
14. Ermolaev V.T., Mavrychev E.A. Flaksman A.G. Effektivnost' sistem svyazi s antennymi reshetkami v usloviyakh rasseivayushchey sredy [Efficiency of communication systems with antenna arrays in a scattering medium], *Uspekhi sovremennoy radioelektroniki* [Success of modern Radioelectronics], 2003, No. 3, pp. 41-48.
15. Luchin D., Plotnikov A., Trofimov D., Filippov D. Problems of implementation of ground biorthogonal and triorthogonal antenna systems, *XI Mezhdunar. IEEE Sibirskaya konf. po upravleniyu i svyazi (SIBCON-2015)* [XI international IEEE Siberian conference. management and communications (SIBCON-2015)]. Omsk, 2015.
16. Flaksman A.G. Adaptivnaya prostranstvennaya obrabotka signalov v mnogokanal'nykh informatsionnykh sistemakh: diss. ... d-ra nauk [Adaptive spatial signal processing in multi-channel information systems: dr. of sc. diss.], Nizhniy Novgorod: Nizhegorodskiy gos-universitet im. N.I. Lobachevskogo, 2004.
17. Simon K., Alouini M.-S. Digital Communication over Fading Channels: A Unified Approach to Performance Analysis. N. Y.: John Wiley&Sons, 2000, 544 p.
18. Stupnitskiy M.M., Luchin D.V. Potentsial KV-radiosvyazi - dlya sozdaniya tsifrovoy ekosistemy Rossii [Potential of KV-Radiocommunication-for creating a digital ecosystem in Russia], *Elektrosvyaz'* [Telecommunications], 2018, pp. 49-54.
19. Ermolaev V.T., Mavrychev E.A. Flaksman A.G. Umen'shenie veroyatnosti bitovoy oshibki pri parallel'noy peredache informatsii v MIMO sisteme [Reducing the probability of bit error when transmitting information in parallel in the MIMO system], *Izvestiya Vuzov. Radiofizika* [Proceedings of Universities. Radiophysics], 2003, Vol. 46, No. 3, pp. 251-260.
20. Savishchenko Nikolay V. Special Integral Functions Used in Wireless Communications Theory. Singapore: World Scientific Publishing Company, 2014.
21. Savishchenko N.V. i dr. Raschet veroyatnosti bitovoy i simvol'noy oshibok dlya kanala svyazi, pri prieme signal'nykh konstruktivnykh standartov DVB-S2 N. V. [Calculation of the probability of bit and character errors for the communication channel, when receiving signal structures of the DVB-S2 standard N. V.], *Informatsiya i Kosmos* [Information and Space], 2015, No. 1, pp. 9-15.
22. Digital Communications: Fundamentals and Applications. 2nd ed. Bernard Sklar, pp. 961-975.
23. Stupnitskiy M.M., Luchin D.V. Potentsial KV-radiosvyazi - dlya sozdaniya tsifrovoy ekosistemy Rossii [Potential of KV-Radiocommunication-for creating a digital ecosystem in Russia], *Elektrosvyaz'* [Telecommunications], 2018, pp. 49-54.
24. Bogdanov A.V. [i dr.]. Ob optimizatsii trebovaniy k peredayushchim kompleksam radiolinii vysokoskorostnoy peredachi dannykh diapazona DKMV [About optimization of requirements for transmitting complexes of high-speed data transmission lines in the dcmv range], *Uspekhi sovremennoy radioelektroniki* [Success of modern radio electronics], 2011, No. 7, pp. 10-16.

25. Zhidyayev A.V., Kopysov A.N., Bogdanov A.A., Savel'ev A.V., Nikitin M.L. Issledovanie energeticheskikh kharakteristik signalov, primenyaemykh dlya peredachi dannykh po dekametrovomu kanalu [Investigation of energy characteristics of signals used for data transmission over a decameter channel], *Vestnik IzhGTU im. M.T. Kalashnikova* [Bulletin of Kalashnikov ISTU], 2015, No. 3 (67), pp. 85-88.
26. Bridger Wray W., Ruiz Mark D. Advisors: Aruna Apte, James B. Greene. Naval Postgraduate School Monterey, California "Total Ownership Cost Reduction Case Study: AEGIS Radar Phase Shifters" December 2006.
27. Stupnitskiy M.M., Kharitonov N.I., Devyatkin E.E. Infokommunikatsionnaya infrastruktura tsifrovoy ekonomiki: zadachi otraslevogo instituta [Information and communication infrastructure of the digital economy: challenges for industry Institute], *Elektrosvyaz'* [Telecommunications], 2018, No. 4, pp. 70-76.

Статью рекомендовал к опубликованию к.т.н., профессор О.В. Воробьев.

Рыбаков Алексей Игоревич – Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича; e-mail: lexeus.r1@gmail.com; 193232, Санкт-Петербург, пр. Большевиков, 22; кафедра радиопередающих устройств и средств подвижной связи; аспирант.

Кротов Роман Евгеньевич – e-mail: ub1cag@yandex.ru; кафедра радиопередающих устройств и средств подвижной связи; аспирант.

Кокин Сергей Алексеевич – e-mail: sergeikokins@gmail.com; кафедра радиопередающих устройств и средств подвижной связи; магистр.

Rybakov Aleksei Igorevich – St. Petersburg state University of telecommunications. prof. M.A. Bonch-Bruевич; e-mail: lexeus.r1@gmail.com; 22 Bolshevikov Ave., Saint Petersburg, 193232, Russia; the department of radio transmitting devices and means of mobile communication; post-graduate student.

Krotov Roman Evgen'evich – e-mail: ub1cag@yandex.ru; the department of radio transmitting devices and means of mobile communication; post-graduate student.

Kokin Sergey Alexandrovich – e-mail: sergeikokins@gmail.com; the department of radio transmitting devices and means of mobile communication; master.

УДК 004.422

DOI 10.18522/2311-3103-2020-2-218-227

Л.К. Бабенко, И.Д. Русаловский

БИБЛИОТЕКА ПОЛНОСТЬЮ ГОМОМОРФНОГО ШИФРОВАНИЯ ЦЕЛЫХ ЧИСЕЛ*

Рассматривается одно из новых направлений криптографии – гомоморфная криптография. Его отличительной особенностью является то, что данный вид криптографии позволяет обрабатывать зашифрованные данные без их предварительной расшифровки таким образом, что результат операций над зашифрованными данными эквивалентен после расшифровки результату операции над открытыми данными. В работе приведены основные области применения гомоморфного шифрования. Выполнен анализ существующих разработок в области гомоморфного шифрования. Анализ показал, что существующие реализации библиотек позволяют обрабатывать только биты или массивы бит и не поддерживают операцию деления. Однако для решения прикладных задач необходима поддержка выполнения целочисленных операций. В результате анализа была выявлена необходимость реализации операции гомоморфного деления, а также актуальность разработки собственной реализации библиотеки гомоморфного шифрования над целыми числами. Возможность выполнения четырех операций (сложение, разность, умножение и деление) над

* Работа выполнена при финансовой поддержке РФФИ в рамках проекта № 18-07-01347.

зашифрованными данными позволит расширить области прикладного использования гомоморфного шифрования. Предложен метод гомоморфного деления, позволяющий выполнять операцию деления над гомоморфно зашифрованными данными. Предложена архитектура библиотеки полностью гомоморфных операций над целыми. Библиотека поддерживает основные гомоморфные операции над целыми числами, а также операцию деления, благодаря методу гомоморфного деления. На базе предложенных метода гомоморфного деления и архитектуры библиотеки была выполнена реализация библиотеки гомоморфных операций над целыми. В статье также приведены замеры времени, необходимого на совершение определенных операций над зашифрованными данными и выполняется анализ эффективности работы разработанной реализации библиотеки. Приводятся выводы и возможные пути дальнейшего развития.

Гомоморфное шифрование; криптографическая защита; криптографическая библиотека; C++.

L.K. Babenko, I.D. Rusalovsky

THE LIBRARY OF FULLY HOMOMORPHIC ENCRYPTION OVER THE INTEGERS

The article discusses one of the new directions of cryptography, a homomorphic cryptography. Its distinctive feature is that this type of cryptography allows you to process encrypted data without first decrypting it 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 paper describes the main areas of application of homomorphic encryption. The analysis of existing developments in the field of homomorphic encryption is performed. The analysis showed that existing library implementations only allow processing of bits or arrays of bits and do not support the division operation. However, to solve applied problems, support for performing integer operations is necessary. The analysis revealed the need to implement the operation of homomorphic division, as well as the relevance of developing your own implementation of a library of homomorphic encryption over integers. The ability to perform four operations (addition, difference, multiplication and division) on encrypted data will expand the field of application of homomorphic encryption. A method of homomorphic division is proposed, which allows performing the division operation on homomorphically encrypted data. A library architecture of completely homomorphic operations on integers is proposed. The library supports the basic homomorphic operations on integers, as well as the division operation, thanks to the method of homomorphic division. Based on the proposed method of homomorphic division and library architecture, a library of homomorphic operations on integers was implemented. The article also provides measurements of the time required to perform certain operations on encrypted data and analyzes the effectiveness of the developed library implementation. Conclusions and possible ways of further development are given.

Homomorphic encryption; cryptographic protection; cryptographic library; C++.

Введение. Криптография с незапамятных времен обеспечивают безопасную передачу информации в небезопасной среде, сохраняя пересылаемые данные в секрете. Эта наука непрерывно развивается. Не так давно зародилось новое направление – гомоморфная криптография. Его отличительной особенностью является то, что данный вид криптографии позволяет обрабатывать зашифрованные данные без их предварительной расшифровки таким образом, что результат операций над зашифрованными данными эквивалентен после расшифровки результату операции над открытыми данными [1–7]. При этом решается одна из проблем криптографии – генерации, хранения и распространения общих сеансовых ключей. При этом 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, несмотря на все свои достоинства, порождает ряд проблем, которые в будущем будет необходимо решить:

- ♦ необходимость обеспечения целостности пересылаемых данных;
- ♦ высокая трудоемкость операций над зашифрованными данными;
- ♦ быстрый рост коэффициентов после выполнения операций над данными.

Развитие и совершенствование гомоморфного шифрования актуально в наше время [8–10]. Улучшение алгоритмов и схем данного направления позволит заменить его во многих сферах, где в данное время используется симметричное шифрование:

- ♦ Облачные вычисления.
- ♦ Облачная обработка фотографий.
- ♦ Электронные голосование (выборы).
- ♦ Защищенный поиск информации.

Анализ актуальности. На данный момент уже выполнено несколько реализаций библиотек для гомоморфного шифрования [11–19]. Наиболее серьезными реализациями, доступными для общего пользования, можно считать две:

- ♦ библиотека HElib, сделанная Шаем Хавели и Виктор Шоуп которая реализует криптосистему BGV с GHS оптимизацией;
- ♦ библиотека FHEW сделанная 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. Обе реализации выполнены на языке программирования C++. Однако, учитывая области применения гомоморфного шифрования, вышеперечисленные библиотеки имеют низкую практическую ценность, так как позволяют выполнять обработку только битов (либо массивов бит), но не целочисленных значений. Также в представленных реализациях отсутствует операция деления над зашифрованными данными. Опираясь на проведенный анализ, сделан вывод о необходимости разработки библиотеки полностью гомоморфного шифрования, позволяющей работать с целыми числами и выполнять над ними все математические операции (сложение, разность, умножение и деление).

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 возможности по выполнению математических операций над ним. Для этого необходимо выделить два уровня представления данных – криптографический и математический. На математическом уровне данные будут представлены в виде простых дробей, а операция деления будет реализована как операция деления простых дробей. Данное решение было разработано в результате анализа [20] и будет подробнее описано далее в статье. При реализации библиотеки с предложенной архитектурой при соблюдении низкой связности модулей и общего интерфейса можно добиться того, чтобы без труда заменять алгоритмы шифрования криптографического уровня, но при этом сохранять весь функционал, который предоставляет математический уровень.

Криптографический уровень представляет собой модуль, основным типом данных которого является гомоморфно зашифрованное 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 шифрования (операция расшифрования данных), а также основные математические операции над шифротекстами – сложение (а также разность) и умножение.

Математический уровень является надстройкой над криптографическим. Для реализации операции деления математический уровень будет реализован как простая дробь – он содержит экземпляры объектов криптографического уровня в виде делимого и делителя и повторяет интерфейс криптографического уровня, дополняя его операцией деления. Все операции над объектом математического уровня выполняются как операции над простыми дробями и использует все возможности криптографического уровня.

Архитектура библиотеки. Реализация библиотеки выполнена на языке C++. Для поддержки выполнения множественных операций над зашифрованными числами и сведения к минимуму неточностей вычисления (округления значений), используется библиотека больших чисел NTL.

При реализации разработанной библиотеки перед ней стояли задачи:

- ◆ Возможность обработки целых чисел.
- ◆ Полностью гомоморфное шифрование.
- ◆ Поддержка всех математических операций, включая операцию деления.

Для поддержки операции деления архитектура библиотеки реализована на основе приведенного выше метода гомоморфного деления. Библиотека представлена классами криптографическим, математическим и классом, отвечающим за ключевую информацию. Архитектура библиотеки представлена на рис. 1.

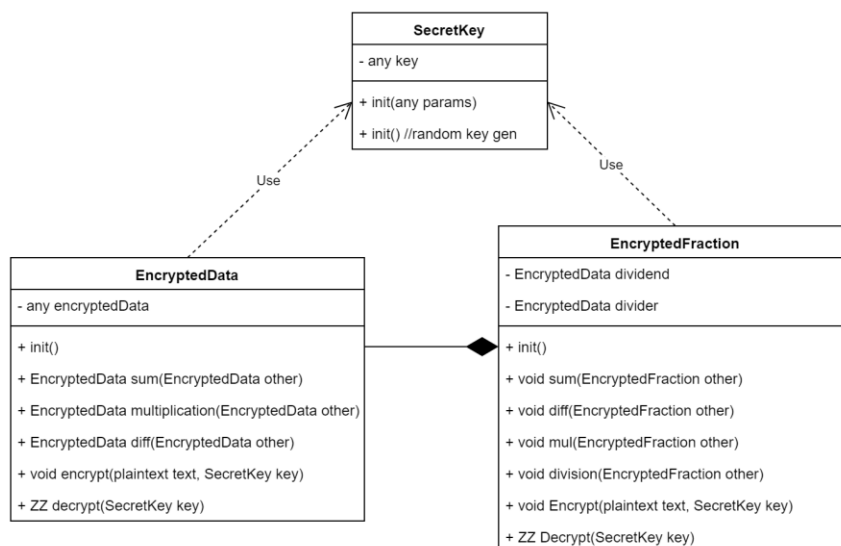


Рис. 1. Архитектура библиотеки

Secret Key. Класс оперирует информацией о секретном ключе, используемом в криптографическом алгоритме. Предоставляет возможности по созданию, случайной генерации нового ключа и его использованию. В текущей реализации библиотеки для генерации ключей и полиномов используется генератор случайных чисел из стандартной библиотеки с автоматической рандомизацией относительно текущего времени. В статье приведено описание библиотеки относительно симметричного шифрования, однако небольшое изменение библиотеки позволит работать и с асимметричным шифрованием. Интерфейс класса представлен на рис. 2.

SecretKey
- any key
+ init(any params)
+ init() //random key gen

Рис. 2. Интерфейс класса SecretKey

Encrypted Data. Класс, определяющий основной тип данных криптографического уровня. Шифрование и расшифровка возможны с использованием заранее сгенерированного ключа, либо с помощью передачи секретных параметров. Также класс реализует все необходимые математические операции криптографического уровня – сумму, разность и умножение. Интерфейс класса представлен на рис. 3.

EncryptedData
- any encryptedData
+ init()
+ EncryptedData sum(EncryptedData other)
+ EncryptedData multiplication(EncryptedData other)
+ EncryptedData diff(EncryptedData other)
+ void encrypt(plaintext text, SecretKey key)
+ ZZ decrypt(SecretKey key)

Рис. 3. Интерфейс класса EncryptedData

Encrypted Fraction. Зашифрованная дробь – основной объект данных математического уровня. Содержит в себе делимое и делитель – объекты данных криптографического уровня (EncryptedData). Реализует все основные операции над дробями – сумму, разность, умножение и деление. Позволяет шифровать данные и расшифровывать полученную дробь на секретном ключе (в ходе операции расшифровки расшифровываются делимое и делитель, а над результатами расшифровки производится операция деления). Архитектура класса представлена на рис. 4.

EncryptedFraction
- EncryptedData dividend
- EncryptedData divider
+ init()
+ void sum(EncryptedFraction other)
+ void diff(EncryptedFraction other)
+ void mul(EncryptedFraction other)
+ void division(EncryptedFraction other)
+ void Encrypt(plaintext text, SecretKey key)
+ ZZ Decrypt(SecretKey key)

Рис. 4. Интерфейс класса EncryptedFraction

Примеры использования библиотеки. Для использования библиотеки необходимы базовые навыки программирования на языке C++. Ниже приведен пример использования библиотеки. В примере фигурируют полиномы, так как в на-

чальной реализации библиотеки именно в виде полиномов представлялись зашифрованные данные, но при замене криптографического уровня библиотеки изменится и представление зашифрованных данных.

Первоначально создается секретный ключ шифрования. Секретный ключ может быть создан на основе заданных параметров или сгенерирован случайным образом.

// Генерация секретного ключа шифрования

SecretKey secKey = SecretKey();

После создания секретного ключа необходимо создать объекты шифротекстов и выполнить их начальную инициализацию. В данном примере создаются начальные шифрующие полиномы 10 степени с коэффициентами в диапазоне (0, 10000).

// Начальная инициализация шифротекстов

int range = 10000;

int power = 10;

EncryptedFraction firstPoly = EncryptedFraction(power, range);

EncryptedFraction secondPoly = EncryptedFraction(power, range);

На начально сгенерированных объектах можно выполнить шифрование данных. Для этого необходимо вызвать соответствующий метод и передать ему в качестве параметров открытый текст и секретный ключ шифрования. В качестве примера приводится шифрование чисел 10 и 20.

// Шифрование данных

// ZZ – основной тип данных библиотеки NTL – большое целое

firstPoly.Encrypt(ZZ(10), secKey);

secondPoly.Encrypt(ZZ(20), secKey);

Над зашифрованными данными можно выполнять различные математические операции. После выполнения операций над данными можно их расшифровать на секретном ключе и получить результат операций над данными. В примере все операции выполняются последовательно.

// Математические операции над данными и последующая расшифровка

firstPoly.sum(secondPoly);

ZZ sum = firstPoly.Decrypt(secKey);

// sum = 30

firstPoly.diff(secondPoly); // Из первого полинома

ZZ diff = firstPoly.Decrypt(secKey);

// diff = 10

firstPoly.mul(secondPoly);

ZZ mul = firstPoly.Decrypt(secKey);

// mul = 200

firstPoly.division(secondPoly);

ZZ division = firstPoly.Decrypt(secKey);

// division = 10

Анализ эффективности. Разработанная библиотека тестировалась на производительность. Тесты проводились на ноутбуке со следующими характеристиками:

- ◆ ОС – Windows 10 Pro;

- ♦ процессор – Intel Core i5-3210M, 4 ядра, 2,5 ГГц (библиотека не имеет параллельных операций, все вычисления происходят на одном ядре);

- ♦ оперативная память – 8 Гб, DDR3.

В ходе тестов измерялось время, необходимое на последовательное умножение или сложение числа самого с собой. Тесты выполнялись следующим образом – выбирались начальные параметры шифрования (степень полинома и диапазон, в котором генерировались коэффициенты полинома), на основе выбранных параметров генерируется начальный полином, далее полученный полином складывался с самим собой или умножался сам на себя определенное число раз (1000 раз). Каждые 100 итераций производились замеры времени, необходимого на выполнение этих операций. На основе выполненных замеров оценивалась зависимость времени, необходимого на выполнение последовательных операций сложения или умножения для различного числа операций в серии.

Тесты умножения. На рис. 5 приведены графики зависимости времени обработки данных с помощью библиотеки от числа последовательных операций над данными для операции умножения.

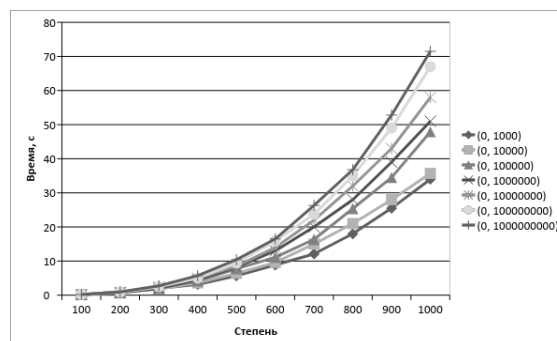


Рис. 5. Возведение в степень полинома 2-го порядка

Для операции умножения наблюдается экспоненциальная сложность выполнения операций. Это обусловлено быстрым ростом шума при выполнении данной операции (после каждой операции умножения объем шифротекста возрастает примерно в 2 раза).

Тесты сложения. На рис. 6 приведены графики зависимости времени обработки данных с помощью библиотеки от числа последовательных операций над данными для операции сложения.

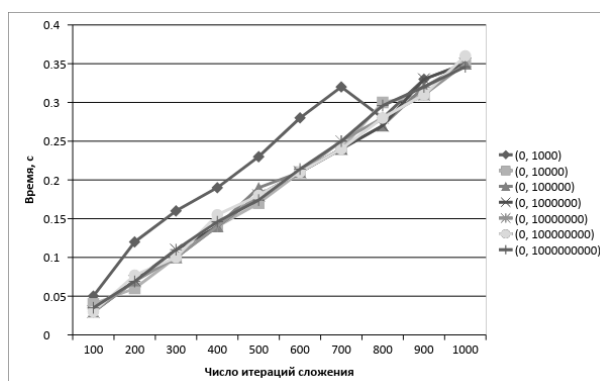


Рис. 6. Сложение полинома 2-го порядка с самим собой

Как видно из представленных графиков, операция сложения имеет линейную сложность от количества последовательных операций сложения в серии. Это связано с низкой скоростью увеличения размера шифротекста в результате операции сложения.

Направление дальнейшей работы. Результатом работы является функциональная библиотека полностью гомоморфного шифрования целых чисел, разработанная на языке C++. Библиотека имеет удобный интерфейс и неплохое быстродействие, как это показали проведенные тесты производительности. Библиотека может использоваться для разработки более сложных продуктов, которые в своем составе используют гомоморфное шифрование. В дальнейшем библиотека будет совершенствоваться и улучшаться. Планируется замена ее криптографического ядра на более защищенный и эффективный алгоритм – реализацию алгоритма Джентри для целых чисел.

Заключение. В ходе проделанной работы был произведен анализ гомоморфного шифрования, выявлены его слабые и сильные стороны и возможности, по его практическому использованию. Был выполнен анализ существующих реализаций библиотек гомоморфного шифрования, который показал, что на данный момент нет реализации, позволяющей работать с целыми числами и иметь полноценный математический аппарат для работы над ними. На основе данного анализа была выявлена необходимость спроектировать и разработать собственную библиотеку полностью гомоморфного шифрования целых чисел, которая позволила бы корректно выполнять криптографические преобразования над целочисленными данными и выполнять над зашифрованными данными все математические операции – сложения, умножения, разности и деления. Для реализации гомоморфного деления также был сформулирован и предложен метод, позволяющий реализовать данную операцию для любого полностью гомоморфного алгоритма.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. 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.
2. Бабенко Л.К., Буртыка Ф.Б., Макаревич О.Б., Трепачева А.В. Полностью гомоморфное шифрование (обзор) // Вопросы защиты информации. – 2015. – №. 3. – С. 3-26.
3. Бабенко Л.К., Буртыка Ф.Б., Макаревич О.Б., Трепачева А.В. Защищенные вычисления и гомоморфное шифрование // III Национальный суперкомпьютерный форум (25-27 ноября 2014, г. Переславль-Залесский). ИПС имени А.К. Айламазяна РАН. – 2014.
4. Макаревич О.Б., Буртыка Ф.Б. Защищенная облачная база данных с применением гомоморфной криптографии // Тез. докладов 6-й Российской мультikonференции «Информационные технологии в управлении» (ИТУ-2014). – СПб., 2014. – С. 567-572.
5. Буртыка Ф.Б. Пакетное симметричное полностью гомоморфное шифрование на основе матричных полиномов // Тр. Института системного программирования РАН. – 2014. – Т. 26, № 5. – С. 99-116.
6. Буртыка Ф.Б. Симметричное полностью гомоморфное шифрование с использованием неприводимых матричных полиномов // Известия ЮФУ. Технические науки. – 2014. – № 8. – С. 107-122.
7. Трепачева А.В. Криптоанализ симметричных полностью гомоморфных линейных криптосистем на основе задачи факторизации чисел // Известия ЮФУ. Технические науки. – 2015. – № 5 (166). – С. 89-102.
8. Diffie W. and Hellman M. New directions in cryptography // IEEE Transactions on Information Theory. – 1976. – Vol. IT-22. – P. 644-654.
9. Гомоморфное шифрование. – URL: <https://habrahabr.ru/post/255205/> (дата обращения: 01.06.2020).
10. Гомоморфное шифрование своими руками. – URL: <https://habrahabr.ru/post/150067/> (дата обращения: 01.06.2020).

11. Gentry. Fully homomorphic encryption using ideal lattices // STOC. – 2009. – P. 169-178.
12. Gentry Craig, A fully homomorphic encryption scheme // A dissertation submitted to the department of computer science and the committee on graduate students of Stanford University. – 2009.
13. Regev O. New lattice-based cryptographic constructions // J. ACM. – 2004. – Vol. 51, No. 6. – P. 899-942.
14. Regev O. On lattices, learning with errors, random linear codes, and cryptography // STOC. – 2005. – P. 84-93.
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. Helib. – URL: <https://github.com/homenc/HElib> (дата обращения: 01.06.2020).
17. FHEW. – URL: <https://github.com/lducas/FHEW> (дата обращения: 01.06.2020).
18. Варновский Н.П., Шокуров А.В. Гомоморфное шифрование // Тр. Института системного программирования РАН. – 2007. – № 12. – С. 27-36.
19. Варновский Н.П., Мартишин С.А., Храпченко М.В., Шокуров А.В. Пороговые системы гомоморфного шифрования и защита информации в облачных вычислениях // Программирование. – 2015. – № 4. – С. 47-51.
20. Яковлев М.О. Защищенный калькулятор. Разработка клиентского компонента. – URL: <http://pdf.knigi-x.ru/21informatika/429422-1-kafedra-sistem-informatiki-vipusknaya-kvalifikacionnaya-rabota-bakalavra-yakovlev-mihail-olegovich-za.php> (дата обращения: 01.06.2020).

REFERENCES

1. Babenko L.K., Burtyka F.B., Makarevich O.B., Trepacheva A.V. Metody polnost'yu gomomorfno shifrovaniya na osnove matrichnykh polinomov [Methods of fully homomorphic encryption based on matrix polynomials], *Voprosy kiberbezopasnosti* [Cybersecurity issues], 2015, No. 1, pp. 17-20.
2. Babenko L.K., Burtyka F.B., Makarevich O.B., Trepacheva A.V. Polnost'yu gomomorfno shifrovaniye (obzor) [Fully Homomorphic Encryption (Overview)], *Voprosy zashchity informatsii* [Information Security Issues], 2015, No. 3, pp. 3-26.
3. Babenko L.K., Burtyka F.B., Makarevich O.B., Trepacheva A.V. Zashchishchennyye vychisleniya i gomomorfno shifrovaniye [Secure computing and homomorphic encryption], *III Natsional'nyy superkomp'yuternyy forum (25-27 noyabrya 2014, g. Pereslavl'-Zalesskiy). IPS imeni A.K. Aylamazyan RAN* [III National supercomputer forum. November, 25-27 of 2014], 2014.
4. Makarevich O.B., Burtyka F.B. Zashchishchennaya oblachnaya baza dannykh s primeneniem gomomorfnoy kriptografii [Secure cloud database using homomorphic cryptography], *Tez. dokladov 6-y Rossiyskoy mul'tikonferentsii «Informatsionnye tekhnologii v upravlenii» (ITU-2014)* [Proceedings of 6th Russian multiconference «Information Technologies in Control» (ITU-2014)]. Saint Petersburg, 2014, pp. 567-572.
5. Burtyka F.B. Paketnoe simmetrichnoe polnost'yu gomomorfno shifrovaniye 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.
6. Burtyka F.B. Simmetrichnoe polnost'yu gomomorfno shifrovaniye 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, pp. 107-122.
7. 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.
8. Diffie W. and Hellman M. New directions in cryptography, *IEEE Transactions on Information Theory*, 1976, Vol. IT-22, pp. 644-654.
9. Gomomorfno shifrovaniye [Homomorphic encryption]. Available at: <https://habrahabr.ru/post/255205/> (accessed 01 June 2020).

10. Gomomorfnoe shifrovanie svoimi rukami [DIY homomorphic encryption]. Available at: <https://habrahabr.ru/post/150067/> (accessed 01 June 2020).
11. Gentry. Fully homomorphic encryption using ideal lattices, *STOC*, 2009, pp. 169-178.
12. Gentry Craig. A fully homomorphic encryption scheme, *A dissertation submitted to the department of computer science and the committee on graduate students of Stanford University*, 2009.
13. Regev O. New lattice-based cryptographic constructions, *J. ACM*, 2004, Vol. 51, No. 6, pp. 899-942.
14. Regev O. On lattices, learning with errors, random linear codes, and cryptography, *STOC*, 2005, pp. 84-93.
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. Helib. Available at: <https://github.com/homenc/HElib> (accessed 01 June 2020).
17. FHEW. Available at: <https://github.com/lducas/FHEW> (accessed 01 June 2020).
18. Varnovskiy N.P., Shokurov A.V. Gomomorfnoe shifrovanie [Homomorphic encryption], *Tr. Instituta sistemnogo programmirovaniya RAN* [Proceedings of the Institute for System Programming RAS], 2007, No. 12, pp. 27-36.
19. Varnovskiy N.P., Martishin S.A., Khrapchenko M.V., Shokurov A.V. Porogovye sistemy gomomorfno shifrovaniya i zashchita informatsii v oblachnykh vychisleniyakh [Threshold systems of homomorphic encryption and information security in cloud computing], *Programmirovaniye* [Programming], 2015, No. 4, pp. 47-51.
20. Yakovlev M.O. Zashchishchenny kal'kulyator. Razrabotka klientskogo komponenta [Protected calculator. Client component development]. Available at: <http://pdf.knigi-x.ru/21informatika/429422-1-kafedra-sistem-informatiki-vipusknaya-kvalifikacionnaya-rabota-bakalavra-yakovlev-mihail-olegovich-za.php> (accessed 01 June 2020).

Статью рекомендовал к опубликованию д.э.н., профессор Е.Н. Тищенко

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 Il'ya Dmitrievich – e-mail: ilya.rusalovskiy@mail.ru; phone: +79885526701; the department of information technologies security; postgraduate student.

Раздел V. Нанотехнологии

УДК 621.38-022.532

DOI 10.18522/2311-3103-2020-2-228-237

А.А. Гелдаш, Л.Э. Левенец, Е.Ю. Гусев, В.Н. Джуплин

ИССЛЕДОВАНИЯ РЕЖИМОВ ФОРМИРОВАНИЯ НАНОКРИСТАЛЛИЧЕСКИХ ПЛЁНОК ZnO:Ga МЕТОДОМ МАГНЕТРОННОГО РАСПЫЛЕНИЯ

Основной целью данной работы является исследование режимов формирования тонких нанокристаллических плёнок ZnO:Ga методом магнетронного распыления на постоянном токе. Основной задачей исследования является получение тонких (~ 300 нм), прозрачных, проводящих плёнок со значением удельного сопротивления менее $5 \cdot 10^{-3}$ Ом·см, которые могут быть применимы в качестве контактов для наноструктур фоточувствительных элементов, а также исследование технологических параметров оборудования магнетронного распыления и мишеней оксидов металлов. Проведено исследование морфологии получаемых тонких плёнок ZnO:Ga. Выявлено, что поверхность плёнок состоит из отдельных кристаллов, объединившихся между собой в процессе осаждения материала. Эти кристаллы имеют ярко выраженные грани и вершины. При увеличении значения мощности источника постоянного тока, кристаллы на поверхности плёнки увеличиваются в несколько раз, пропорционально увеличению мощности, и возрастает толщина плёнки, вследствие увеличения скорости распыления материала мишени на подложку. Далее исследованы электрические характеристики полученных плёнок и выведены зависимости влияния мощности (толщины) плёнки на концентрацию носителей, их подвижность, а также удельное сопротивление. При увеличении толщины плёнки с 320 нм до 340 нм подвижность носителей тока увеличивается с $3,027 \text{ см}^2/(\text{В} \cdot \text{с})$ до $3,228 \text{ см}^2/(\text{В} \cdot \text{с})$, и при увеличении толщины плёнки с 800 нм до 1200 нм возрастает с $6,511 \text{ см}^2/(\text{В} \cdot \text{с})$ до $6,547 \text{ см}^2/(\text{В} \cdot \text{с})$. При увеличении толщины плёнки с 320 нм до 340 нм концентрация носителей тока уменьшается с $1,571 \cdot 10^{20} \text{ см}^{-3}$ до $1,489 \cdot 10^{20} \text{ см}^{-3}$, и при увеличении толщины плёнки с 800 нм до 1200 нм концентрация носителей тока также уменьшается с $2,481 \cdot 10^{20} \text{ см}^{-3}$ до $1,653 \cdot 10^{20} \text{ см}^{-3}$. При увеличении толщины плёнки с 320 нм до 340 нм удельное сопротивление увеличивается с $1,303 \cdot 10^{-2} \text{ Ом} \cdot \text{см}$ до $1,385 \cdot 10^{-2} \text{ Ом} \cdot \text{см}$, и при увеличении толщины плёнки с 800 нм до 1200 нм удельное сопротивление также увеличивается с $3,851 \cdot 10^{-2} \text{ Ом} \cdot \text{см}$ до $5,779 \cdot 10^{-2} \text{ Ом} \cdot \text{см}$.

Тонкие; прозрачные; проводящие плёнки; магнетронное распыление; ZnO:Ga.

A.A. Geldash, L.E. Levenets, E.Y. Gusev, V.N. Dzhuplin

STUDY OF THE MODES OF FORMING ZnO:Ga NANOCRYSTALLINE FILMS BY MAGNETRON SPUTTERING

The main goal of this work is to study the modes of formation of thin nanocrystalline ZnO:Ga films by direct current magnetron sputtering. The main objective of the study is to obtain thin (~ 300 nm), transparent, conductive films with a resistivity value of less than $5 \cdot 10^{-3}$ Ohm cm, which can be used as contacts for nanostructures of photosensitive elements, as well as studying the technological parameters of magnetron sputtering equipment and metal oxide targets. The morphology of the resulting thin ZnO:Ga films was studied. As a result of the study, it was revealed that the surface of the films consists of individual crystals, united together during the deposition of the material. These crystals have pronounced faces and peaks. It was also revealed that with an increase in the power of a direct current source, the crystals on the film surface increase several times in proportion to the increase in power, and the film thickness increases due to an

increase in the speed of sputtering of the target material on the substrate. Further, the electrical characteristics of the obtained films were investigated and the dependences of the influence of the power (thickness) of the film on the carrier concentration, their mobility, and also resistivity were derived. Thus, it was revealed that with an increase in the film thickness from 320 nm to 340 nm, the mobility of current carriers increases from $3.027 \text{ cm}^2/(\text{V}\cdot\text{s})$ to $3.228 \text{ cm}^2/(\text{V}\cdot\text{s})$, and with an increase in the film thickness from 800 nm to 1200 nm of current carriers increases from $6.511 \text{ cm}^2/(\text{V}\cdot\text{s})$ to $6.547 \text{ cm}^2/(\text{V}\cdot\text{s})$. With an increase in the film thickness from 320 nm to 340 nm, the concentration of current carriers decreases from $1.571 \cdot 10^{20} \text{ cm}^{-3}$ to $1.489 \cdot 10^{20} \text{ cm}^{-3}$, and with an increase in the film thickness from 800 nm to 1200 nm, the concentration of current carriers decreases from $2.481 \cdot 10^{20} \text{ cm}^{-3}$ to $1.653 \cdot 10^{20} \text{ cm}^{-3}$. As the film thickness increases from 320 nm to 340 nm, the resistivity increases from $1.303 \cdot 10^{-2} \Omega\cdot\text{cm}$ to $1.38 \cdot 10^{-2} \Omega\cdot\text{cm}$, and as the film thickness increases from 800 nm to 1200 nm, the resistivity increases from $3.851 \cdot 10^{-2} \Omega\cdot\text{cm}$ to $5.779 \cdot 10^{-2} \Omega\cdot\text{cm}$.

Thin; transparent; conductive films; magnetron sputtering; ZnO:Ga.

Введение. Проводимые прозрачные слои широко применяются в электронных приборах, таких как солнечные батареи, светодиоды, и другие типов электронных приборов, параметры которых зависят от прохождения излучения через эти слои [1–3]. Поиск и исследование альтернативных прозрачных проводимых слоев в настоящее время является широко распространенной задачей, так как с каждым годом мировой спрос на источники альтернативной энергии увеличивается [4, 5].

ZnO:Ga является прозрачным проводниковым материалом. Этот материал является альтернативой дорогостоящему ИТО. Оксид металла ZnO:Ga, является одновременно и прозрачным и проводящим. Эти свойства присущи и другим оксида металлов, например ZnO:In или ИТО. Поэтому в большинстве случаев, при использовании современных прозрачных проводников (слоев, пленок), они сформированы именно из оксидов металлов. На электропроводность и оптические свойства таких материалов влияют прежде всего их природа (количества и атомного расположения катионов металла в структуре оксида), а также наличия собственных или введенных дефектов кристаллической структуры [6–8]. Тем временем как электрическая проводимость обусловлена легированием, кислородными вакансиями или введением примеси. Легирование оксида цинка галлием усиливает генерацию свободных носителей в пленках ZnO [7–9]. Применение ZnO:Ga обусловлено низкими материальными затратами и относительно низкими температурами осаждения [10], в данной работе ZnO:Ga рассматривается как одна из альтернатив применению ИТО слоев или пленок. До недавнего времени ИТО слои являлись неким «стандартом» прозрачных проводящих пленок со значением удельного сопротивления порядка $\sim 10^{-4} \text{ Ом}\cdot\text{см}$ и коэффициентом пропускания более 80 %. Но из-за его недостатков и разрастающегося применения таких слоев, сегодня остро стоит задача в поиске и исследовании альтернативных оксидов металлов близкими ИТО электрическими и оптическими параметрами. Основным недостатком ИТО является высокая стоимость индия, входящего в его соединение, а также ограничение оптической настройки и плохой механической гибкости.

Для напыления тонких пленок проводящих, резистивных и диэлектрических материалов широко используется магнетронное распыление [11–13]. Данный метод позволяет проводить контролируемое распыление тонких слоев с требуемыми параметрами. Сама технология основывается на нанесении тонкой пленки на подложку с помощью катодного распыления мишени в плазме магнетронного разряда – диодного разряда в скрещенных полях [14–17]. Метод магнетронного распыления на сегодняшний день является одним из наиболее перспективных и широко применяемых методов формирования тонких нанокристаллических пленок различных материалов для применения в микро- и нанoeлектронике. С помощью данного метода можно проводить контролируемое распыление тонких слоев с тре-

буемыми параметрами, но с учетом преждевременной проработки режимов, включающих в себя как исследование возможностей конкретного технологического оборудования, так и конкретной мишени (материала мишени), а также исследование влияния множества параметров в процессе нанесения слоев [18]. Все эти подготовительные операции крайне важны для получения слоев с требуемыми параметрами, а также для дальнейшей воспроизводимости подобных слоев или пленок. Именно эти операции служат в качестве «отправной точки» в исследовании возможности получения требуемых (возможных) слоев или пленок для различного назначения. Данный метод формирования тонких пленок развивается благодаря тому, что все больше требований предъявляется к параметрам формируемых слоев и пленок, особенно к их морфологическим и электрическим характеристикам [19–22]. При воспроизводимости тонких слоев или пленок методом магнетронного распыления, стоит учитывать очень важную особенность самой мишени – это образование зоны эрозии («круговой канавки» мишени, формируемой между центром мишени и ее краями, в процессе магнетронного распыления), так как эта особенность влияет на скорость распыления материала мишени и следовательно на саму толщину пленки. А толщина пленки, при магнетронном распылении, напрямую влияет на величину удельного электрического сопротивления. Так при уменьшении толщины пленки оно увеличивается, а при увеличении толщины уменьшается. Это в свою очередь приведет к невозможности требуемых характеристик и параметров тонких пленок. С помощью метода магнетронного распыления можно формировать тонкие проводящие, резистивные или диэлектрические слои.

В качестве оборудования для формирования тонких пленок ZnO:Ga выступает установка для магнетронного и электронно-лучевого осаждения серии AUTO 500 на постоянном токе. Она оснащена высокочастотным магнетроном и магнетроном постоянного тока. При помощи магнетронного распыления могут быть нанесены пленки практически из всех металлов, сплавов и соединений. Стехиометрический состав пленки, полученной магнетронным распылением, остается неизменным для большинства материалов, в отличие от термического испарения, при котором соединение распадается на отдельные фракции. Установка, так же, имеет линию подачи кислорода и, следовательно, присутствует возможность нанесения широкого спектра оксидов металлов и полупроводников. Электронный луч с 4 тиглями позволяет напылять сложные структуры в вакууме. Для контроля толщины нанесенных пленок используется кварцевый толщиномер. Основными достоинствами установки являются: Возможность объединить несколько источников в связку (в круг), для направленного распыления; Наличие системы магнетронного распыления как на постоянном токе ПТ, так и на токе высокой частоты ВЧ; системы ПТ и ВЧ травления; Возможность комбинирования систем ПТ, ВЧ, Резистивного и электронно-лучевого испарения; Возможность подачи небольшого ВЧ-напряжения на подложку в процессе напыления. Установка AUTO 500 также характеризуется наличием устройства распределения мощности для систем ВЧ и ПТ травления, водоохлаждаемого, изолированного по ВЧ напряжению подложкодержателя, автоматической системы управления напуском технологического газа, системы автоматического управления технологическим процессом, а также системы косвенного нагрева подложек.

Методическая часть. В данной работе для получения тонких прозрачных проводящих пленок применялся метод магнетронного распыления на постоянном токе. В качестве подложек, на которых формировались пленки ZnO:Ga, выступали стекло, ситалл и Si p-типа с концентрацией носителей тока порядка 10^{19} см^{-3} . Все подложки перед нанесением подвергались химической очистке от загрязнений и окислов. Формирование пленки ZnO:Ga на три типа подложек осуществлялось в одном технологическом режиме, но в две экспериментальные сессии.

Перед началом процесса магнетронного распыления, было осуществлен не-прямой подогрев подложкодержателей с образцами до температуры 250°C , с удержанием этой температуры на время 10–20 минут, для того чтобы во время процесса распыления значение температуры не падало, так как процесс происходит без параллельного нагрева. Далее была осуществлена проработка мишени, продолжительностью около 15 минут. Проработка подразумевает снятия окисла с поверхности мишени, и выхода на её рабочий режим распыления. При завершении процесса проработки свечение от мишени становится голубым, что косвенно подтверждает проработку мишени и распыление «чистого» материала. Основные параметры при магнетронном распылении таковы: время осаждения – 18 минут, мощность источника постоянного тока – 100 Вт для первого образца и 170 Вт для второго образца, поток рабочего газа Ar – порядка 10 см^3 в минуту, диапазон значений постоянного напряжения 406–434 В, диапазон значения тока 223–238 А. Температура в процессе нанесения падала со 143°C до 114°C , со скоростью $1\text{--}2^{\circ}\text{C}$ в минуту.

Результаты и их обсуждение. На рис. 1 представлены РЭМ изображения двух образцов с пленкой ZnO:Ga, сформированной на поверхности Si, методом магнетронного распыления на постоянном токе. РЭМ изображения получены с помощью растрового электронного микроскопа Nova Nanolab 600. На рис. 1,а представлено РЭМ изображение образца с пленкой ZnO:Ga на Si, сформированной при мощности источника постоянного тока 100 Вт. На рис. 1, б представлено РЭМ изображение образца с пленкой ZnO:Ga на Si, сформированной при мощности источника постоянного тока 170 Вт.

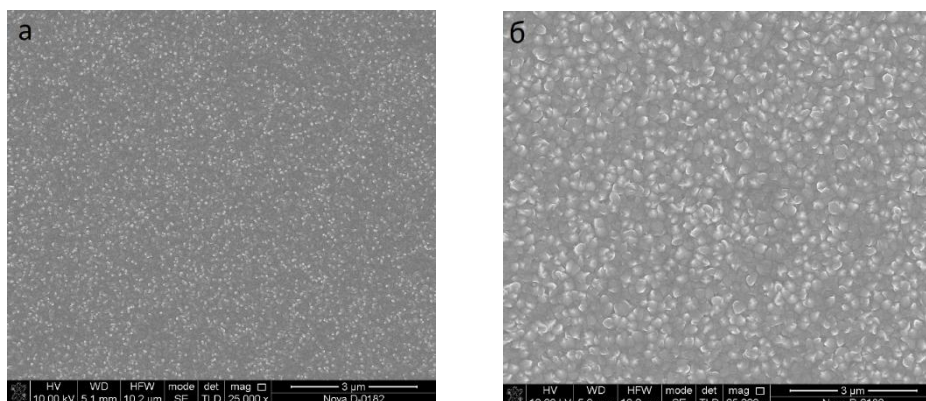


Рис. 1. РЭМ изображения поверхности тонких пленок ZnO:Ga: а – пленка ZnO:Ga на Si, сформированной при мощности источника постоянного тока 100 Вт; б – пленка ZnO:Ga на Si, сформированной при мощности 170 Вт

Как видно из рис. 1,а тонкая пленка ZnO:Ga представляет собой не гладкую поверхность, а состоящую из кристаллов, объединенных между собой. Такой характер морфологии поверхности характерен для пленок, сформированных магнетронным распылением. Он обуславливается прежде всего высокими скоростями распыления материала мишени на подложку.

На рис. 2 представлены РЭМ изображения двух образцов с пленкой ZnO:Ga, сформированной на поверхности Si, методом магнетронного распыления на постоянном токе, при увеличении $\times 250000$.

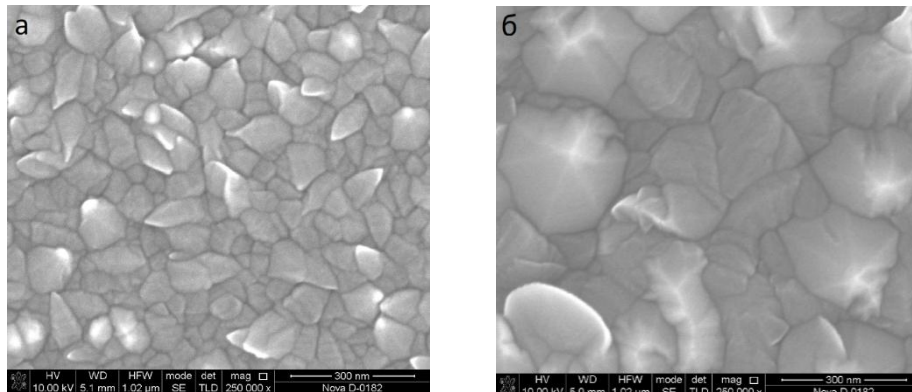


Рис. 2. РЭМ изображения поверхности тонких пленок ZnO:Ga, сформированных методом магнетронного распыления (при увеличении $\times 250000$): а – образец, сформированной при мощности источника постоянного тока 100 Вт; б – образцы, сформированной при мощности 170 Вт

Как видно из рис. 2,а поверхность пленки ZnO:Ga сформирована кристаллами плоской и заостренной формы. Они имеют разные геометрические размеры по всей поверхности пленки. Однако просматривается единый характер их построения на поверхности. На рис. 2,б видно, что кристаллы на поверхности пленки ZnO:Ga в несколько раз больше кристаллов на поверхности пленки ZnO:Ga на рис. 2,а. Что также следует из того, что мощности источника постоянного тока в обоих случаях различались. И следовательно, при большей мощности, скорости распыления больше, при одинаковом времени распыления, и объему подаваемого рабочего газа Ar. Стоит также отметить, что на рис. 2,б более отчетливо видны вершины кристаллов, сформировавшиеся на поверхности пленки, или являющиеся одновременно и самой пленкой ZnO:Ga. У этих кристаллов просматриваются также грани, их 6.

На рис. 3 представлены РЭМ изображения двух образцов с пленкой ZnO:Ga, сформированной на поверхности Si, методом магнетронного распыления на постоянном токе, но уже под наклоном. Эти РЭМ изображения были получены для первичной оценки размеров кристаллов на поверхности пленки ZnO:Ga.

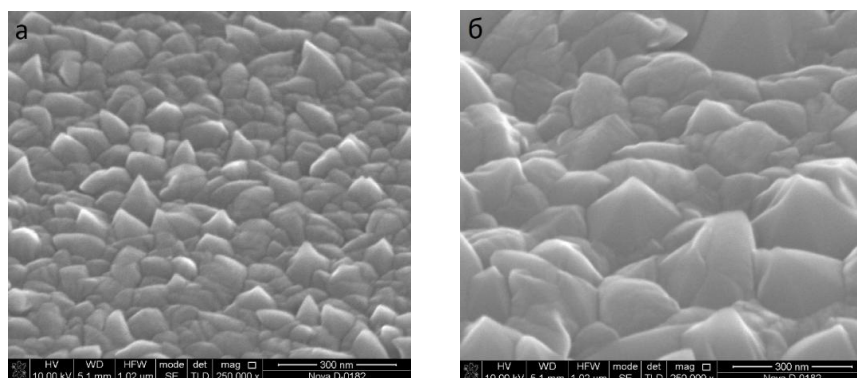


Рис. 3. РЭМ изображения поверхности тонких пленок ZnO:Ga, сформированных методом магнетронного распыления: а – образец с пленкой ZnO:Ga на Si, сформированный при мощности источника постоянного тока 100 Вт; б – образец с пленкой, сформированный при мощности 170 Вт

Из рис. 3 видно, что при увеличении мощности источника постоянного тока, а следовательно скорости распыления материала на подложку, уменьшается количество кристаллов на поверхности пленки, но увеличиваются размеры самих кристаллов, на величину схожую с величиной пары или нескольких пар кристаллов. Вероятно, при более высоких скоростях роста, кристаллы формируются с большими размерами.

Исследованы зависимости концентрации носителей тока от толщины пленки ZnO:Ga, при различных значениях мощности источника постоянного тока. На рис. 4, а показана зависимость концентрации носителей тока от толщины пленки ZnO:Ga при 100 Вт. При увеличении толщины пленки с 320 нм до 340 нм концентрация носителей тока уменьшается с $1,571 \cdot 10^{20} \text{ см}^{-3}$ до $1,489 \cdot 10^{20} \text{ см}^{-3}$. Уменьшение концентрации происходит не значительно. На рис. 4, б показана зависимость концентрации носителей тока от толщины пленки ZnO:Ga при 170 Вт. При увеличении толщины пленки с 800 нм до 1200 нм концентрация носителей тока уменьшается с $2,481 \cdot 10^{20} \text{ см}^{-3}$ до $1,653 \cdot 10^{20} \text{ см}^{-3}$. Уменьшение концентрации также происходит не значительно.

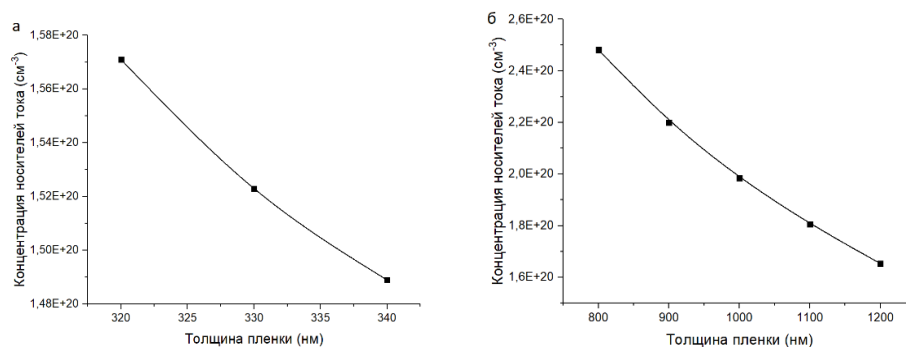


Рис. 4. Зависимости концентрации носителей тока от толщины пленок ZnO:Ga: а – сформированной при 100 Вт; б – при 170 Вт

Определены зависимости подвижности носителей тока от толщины пленки ZnO:Ga, при различных значениях мощности источника постоянного тока. На рис. 5, а показана зависимость подвижность носителей тока от толщины пленки ZnO:Ga при 100 Вт. При увеличении толщины пленки с 320 нм до 340 нм подвижность носителей тока увеличивается с $3,027 \text{ см}^2/(\text{В} \cdot \text{с})$ до $3,228 \text{ см}^2/(\text{В} \cdot \text{с})$. Увеличение подвижность происходит не значительно. На рис. 5, б показана зависимость подвижность носителей тока от толщины пленки ZnO:Ga при 170 Вт. При увеличении толщины пленки с 800 нм до 1200 нм подвижность носителей тока увеличивается с $6,511 \text{ см}^2/(\text{В} \cdot \text{с})$ до $6,547 \text{ см}^2/(\text{В} \cdot \text{с})$. Увеличение подвижность носителей тока также происходит не значительно.

Также определены зависимости удельного сопротивления от толщины пленки ZnO:Ga, при различных значениях мощности источника постоянного тока. На рис. 6, а показана зависимость удельного сопротивления от толщины пленки ZnO:Ga при 100 Вт. При увеличении толщины пленки с 320 нм до 340 нм удельное сопротивление увеличивается с $1,303 \cdot 10^{-2} \text{ Ом} \cdot \text{см}$ до $1,385 \cdot 10^{-2} \text{ Ом} \cdot \text{см}$. Увеличение удельного сопротивления происходит не значительно. На рис. 5, б показана зависимость удельного сопротивления от толщины пленки ZnO:Ga при 170 Вт. При увеличении толщины пленки с 800 нм до 1200 нм удельное сопротивление увеличивается с $3,851 \cdot 10^{-2} \text{ Ом} \cdot \text{см}$ до $5,779 \cdot 10^{-2} \text{ Ом} \cdot \text{см}$. Увеличение удельного сопротивления также происходит не значительно, если учитывать его на отдельных образцах.

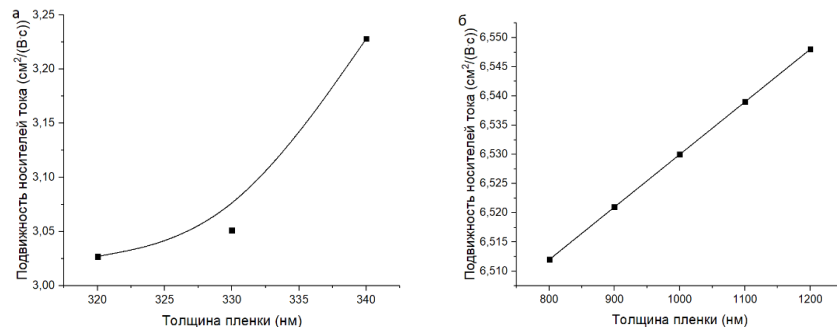


Рис. 5. Зависимости концентрации носителей тока от толщины пленок ZnO:Ga: а – сформированной при 100 Вт; б – при 170 Вт

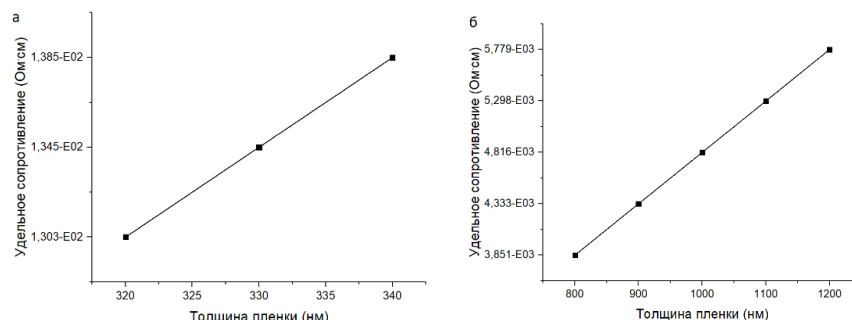


Рис. 6. Зависимости удельного сопротивления от толщины пленок ZnO:Ga: а – сформированной при 100 Вт; б – при 170 Вт

Заключение. В работе проведены экспериментальные исследования по получению тонких нанокристаллических пленок ZnO:Ga методом магнетронного распыления на постоянном токе, а также исследование электрических характеристик в зависимости от режимов формирования пленок. Исследована морфология тонких пленок ZnO:Ga. Было выявлено, что при увеличении мощности источника постоянного тока, размер зерен на поверхности пленок увеличивается в несколько раз, пропорционально увеличению мощности. Также значение мощности источника постоянного тока напрямую влияет на скорость распыления материала мишени, и следовательно, на итоговую толщину пленки. Одной из трудностей в формировании таких пленок является неравномерность осаждения материала на подложку, из-за особенностей технологического оборудования. В заключении исследования морфологии поверхности пленок, стоит выявить явное отличие самих кристаллов на поверхности, так как они различаются геометрическим построением в зависимости от скорости осаждения материала, и следовательно, от мощности источника постоянного тока. Мы показали, что в зависимости от мощности источника постоянного тока, форма поверхности слоя и его структура характерны или мелкими или крупными кристаллами. То есть, прозрачность подобного слоя для проникающего излучения будет разной для разных длин волн. Это один из вариантов создания селективных фильтров проникающего или отраженного света для видимого, инфракрасного или ультрафиолетового диапазонов.

Что касается исследования электрических характеристик сформированных пленок ZnO:Ga, то они были проведены преимущественно с помощью метода Холла. Было выявлено, что концентрация носителей тока в пленке уменьшается с увеличением ее толщины, плотность носителей заряда при этом увеличивалась, а удельное сопротивление возрастало.

Только пленки ZnO:Ga с толщинами порядка 800-1000 нм соответствуют минимальным требованиям, которые предъявлялись в планировании исследования. Так как основная задача состояла в том, чтобы получить тонкие (порядка 300 нм), прозрачные, проводимые пленки со значением удельного сопротивления менее 10^{-3} Ом·см. Которые в дальнейшем будут использоваться в качестве контактных слоев к фоточувствительным структурам. Дальнейшая работа по данной тематике будет нацелена на дальнейшее исследование и оптимизация технологического процесса получения тонких, прозрачных проводящих слоев, то дальнейшая работа по данной тематике может быть направлена на оптимизацию технологического процесса получения слоев, а также на увеличение концентрации Ga в исходной мишени.

Благодарности. Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 19-37-90139, № 19-38-60052, № 18-28-11019 мк. Результаты были получены с использованием оборудования Научно-образовательного центра и Центра коллективного пользования «Нанотехнологии» Южного федерального университета.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Агеев О.А., Коноплев Б.Г. и др. Нанотехнологии в микроэлектронике. – М.: Наука, 2019. – С. 112-165.
2. Kelly P.J., Arnell R.D. Магнетронное распыление: обзор последних разработок и применений // *Vacuum*. – 2000. – Т. 56. – С. 159-172.
3. Климин В.С., Резван А.А., Агеев О.А. Исследование применения плазменных методов для формирования полевых эмиттеров на основе углеродных наноразмерных структур // *Journal of Physics: Conference Series*. – 2018. – Т. 1124. – С. 071020.
4. Томинов Р.В., Смирнов В.А., Авилов В.И., Федотов А.А., Климин В.С., Черненко Н.Е. Формирование мемристорных структур ZnO методом сканирующей зондовой нанолитографии // *IOP Conf. Series: Materials Science and Engineering*. – 2018. – Т. 443. – С. 012036.
5. Murdick D.A., Wadley H.N., Zhou X.W. Механизмы конденсации пара, обогащенного мышьяком, на поверхности GaAs (001) // *Phys. Rev. B*. – 2007. – Т. 75. – С. 125318.
6. Вакулов З.Е., Ивонин М., Замбург Е.Г., Климин В.С., Волик Д.П., Голосов Д.А., Завадский С.М., Достанко А.П., Миаконких А.В., Клементе И.Е., Руденко К.В., Агеев О.А. Размерные эффекты в тонких пленках LiNbO₃, полученных методом импульсного лазерного напыления // *Journal of Physics: Conference Series*. – 2018. – Т. 1124. – С. 022032.
7. Балакирев С.В., Еременко М.М., Михайлин И.А., Климин В.С., Солодовник М.С. Капельная эпитаксия наноструктур In/AlGaAs на As-стабилизированной поверхности // *Journal of Physics: Conference Series*. – 2018. – Т. 1124. – С. 022018.
8. Tseng A. Последние разработки в области микрообработки с использованием технологии сфокусированных ионных пучков // *Journal of Micromechanics and Microengineering*. – 2004. – Т. 14. – С. 15.
9. Han J., Lee H., Min B., Lee S. Прогнозирование топологии наноматериалов с использованием двумерного фокусированного ионного фрезерования с интервалами облучения пучком // *Microelectronic Engineering*. – 2010. – Т. 87. – С. 1-9.
10. Климин В.С., Солодовник М.С., Лисицин С.А., Резван А.А., Балакирев С.В. Формирование наноразмерных структур на поверхности арсенида галлия при локальном анодном окислении и плазмохимическом травлении // *Journal of Physics: Conference Series*. – 2018. – Т. 1124. – С. 041024.
11. Adachi S. Свойства полупроводников группы IV, III-V и II-VI. – John Wiley & Son, 2005. – 408 с.
12. Cao G. Наноструктуры и наноматериалы. Синтез, свойства и применение. – Imperial College Press, 2004. – 448 с.
13. Асеев А.Л. Нанотехнологии в полупроводниковой электронике. – Новосибирск: Изд-во СО РАН, 2004. – 368 с.
14. Jha N.K., Chen D. Проектирование наноэлектронных схем design. – Springer, 2011. – 290 p.
15. Piccin P. et al. Рост методом молекулярно-лучевой эпитаксии и электрическая характеристика нановискеров GaAs // *Physica E*. – 2007. – Т. 37. – С. 134-137.

16. Черненко Н.Е., Балакирев С.В., Еременко М.М., Солодовник М.С. Межфазное взаимодействие в системе Ga-As-ZnO в условиях молекулярно-лучевой эпитаксии // Известия ЮФУ. Технические науки. – 2019. – № 2 (204). – С. 184-192.
17. Балакирев С.В., Блинов Ю.Ф., Солодовник М.С. Модель начальной стадии гомоэпитаксиального роста GaAs методом МЛЭ с учетом соотношения потоков ростовых компонент // Известия ЮФУ. Технические науки. – 2014. – № 9 (158). – С. 94-105.
18. Chu C-P., Arafin S., Nie T. Наноразмерный рост GaAs на узорчатых подложках Si (111) методом молекулярно-лучевой эпитаксии // Crystal Growth & Design. – 2014. – Т. 14. – С. 593-598.
19. Гусев Е.Ю., Житяева Ю.Ю., Агеев О.А. Влияние условий PECVD на механическое напряжение пленок кремния // Materials Physics and Mechanics. – 2018. – Т. 37, № 1. – С. 67-72.
20. Гусев Е.Ю., Житяева Ю.Ю., Коломийцев А.С., Гамалеев В.А., Коц И.Н., Быков А.В. Исследование режимов жидкостного травления жертвенного слоя SiO₂ для формирования микромеханических структур на основе Si⁺/SiO₂/Si // Известия ЮФУ. Технические науки. – 2015. – № 2 (163). – С. 236-245.
21. Sze S. Физика полупроводниковых приборов. – Т. 1. – М.: Мир, 1984. – С. 258-261.
22. Епифанов Г.И. Физика Основы микроэлектроники. – М.: Советское радио, 1971. – С. 145-156.

REFERENCES

1. Ageev O.A., Konoplev B.G. i dr. Nanotekhnologii v mikroelektronike [Nanotechnology in Microelectronics]. Moscow: Nauka, 2019, pp. 112-165.
2. Kelly P.J., Arnell R.D. Magnetronnoe raspylenie: obzor poslednikh razrabotok i primeneniya [Magnetron sputtering: a review of recent developments and applications], *Vacuum*, 2000, Vol. 56, pp. 159-172.
3. Klimin V.S., Rezvan A.A., Ageev O.A. Issledovanie primeneniya plazmennyykh metodov dlya formirovaniya polevykh emitterov na osnove ugleodnykh nanorazmernyykh struktur [Research of using plasma methods for formation field emitters based on carbon nanoscale structures], *Journal of Physics: Conference Series*, 2018, Vol. 1124, pp. 071020.
4. Tominov R.V., Smirnov V.A., Atilov V.I., Fedotov A.A., Klimin V.S., Chernenko N.E. Formirovanie memristornyykh struktur ZnO metodom skaniruyushchey zondovoy nanolitografii [Formation ZnO memristor structures by scratching probe nanolithography], *IOP Conf. Series: Materials Science and Engineering*, 2018, Vol. 443, pp. 012036.
5. Murdick D.A., Wadley H.N., Zhou X.W. Mekhanizmy kondensatsii para, obogashchennogo mysh'yakom, na poverkhnosti GaAs (001) [Condensation mechanisms of an arsenic-rich vapor on GaAs (001) surfaces], *Phys. Rev. B*, 2007, Vol. 75, pp. 125318.
6. Vakulov Z.E., Ivonin M., Zamburg E.G., Klimin V.S., Volik D.P., Golosov D.A., Zavadskiy S.M., Dostanko A.P., Miakonkikh A.V., Klemente I.E., Rudenko K.V., Ageev O.A. Razmernye efekty v tonkikh plenkakh LiNbO₃, poluchennykh metodom impul'snogo lazernogo napyleniya [Size effects in LiNbO₃ thin films fabricated by pulsed laser deposition], *Journal of Physics: Conference Series*, 2018, Vol. 1124, pp. 022032.
7. Balakirev S.V., Eremenko M.M., Mikhaylin I.A., Klimin V.S., Solodovnik M.S. Kapel'naya epitaksiya nanostruktur In/AlGaAs na Asstabilizirovannoy poverkhnosti [Droplet epitaxy of In/AlGaAs nanostructures on the As-stabilized surface], *Journal of Physics: Conference Series*, 2018, Vol. 1124, pp. 022018.
8. Tseng A. Poslednie razrabotki v oblasti mikroobrabotki s ispol'zovaniem tekhnologii sfokusirovannykh ionnykh puchkov [Recent developments in micromilling using focused ion beam technology], *Journal of Micromechanics and Microengineering*, 2004, Vol. 14, pp. 15.
9. Han J., Lee H., Min B., Lee S. Prognozirovanie topologii nanomaterialov s ispol'zovaniem dvumernogo fokusirovannogo ionnogo frezerovaniya s intervalami oblucheniya puchkom [Prediction of nanopattern topography using two-dimensional focused ion beam milling with beam irradiation intervals], *Microelectronic Engineering*, 2010, Vol. 87, pp. 1-9.
10. Klimin V.S., Solodovnik M.S., Lisitsin S.A., Rezvan A.A., Balakirev S.V. Formirovanie nanorazmernyykh struktur na poverkhnosti arsenida galliya pri lokal'nom anodnom okislenii i plazmokhimicheskom travlenii [Formation of nanoscale structures on the surface of gallium arsenide by local anodic oxidation and plasma chemical etching], *Journal of Physics: Conference Series*, 2018, Vol. 1124, pp. 041024.

11. *Adachi S.* Svoystva poluprovodnikov gruppy IV, III-V i II-VI [Properties of group-IV, III-V and II-VI semiconductors]. John Wiley & Son, 2005, 408 p.
12. *Cao G.* Nanostruktury i nanomaterialy. Sintez, svoystva i primeneniye [Nanostructures & nanomaterials. Synthesis, properties & applications]. Imperial College Press, 2004, 448 p.
13. *Aseev A.L.* Nanotekhnologii v poluprovodnikovoy elektronike [Nanotechnology in semiconductor electronics]. Novosibirsk: Izd-vo SO RAN, 2004, 368 p.
14. *Jha N.K., Chen D.* Proektirovaniye nanoelektronnykh skhem design [Nanoelectronic circuit design]. Springer, 2011, 290 p.
15. *Piccin P. et al.* Rost metodom molekulyarno-luchevoy epitaksii i elektricheskaya kharakteristika nanoviskerov GaAs [Growth by molecular beam epitaxy and electrical characterization of GaAs nanowires], *Physica E*, 2007, Vol. 37, pp. 134-137.
16. *Chernenko N.E., Balakirev S.V., Eremenko M.M., Solodovnik M.S.* Mezhfaznoye vzaimodeystvie v sisteme Ga-As-ZnO v usloviyakh molekulyarno-luchevoy epitaksii [Interfacial interaction in the Ga-As-ZnO system under molecular beam epitaxy], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2019, No. 2 (204), pp. 184-192.
17. *Balakirev S.V., Blinov Yu.F., Solodovnik M.S.* Model' nachal'noy stadii gomoepitaksial'nogo rosta GaAs metodom MLE s uchetom sootnosheniya potokov rostovykh komponent [The model of the initial stage of GaAs homoepitaxial growth by the MBE method taking into account the ratio of the fluxes of growth components], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2014, No. 9 (158), pp. 94-105.
18. *Chu C-P., Arafin S., Nie T.* Nanorazmernyy rost GaAs na uzorchatykh podlozhkakh Si (111) metodom molekulyarno-luchevoy epitaksii [Nanoscale Growth of GaAs on Patterned Si (111) Substrates by Molecular Beam Epitaxy], *Crystal Growth & Design*, 2014, Vol. 14, pp. 593-598.
19. *Gusev E.Yu., Zhityaeva Yu.Yu., Ageev O.A.* Vliyaniye usloviy PECVD na mekhanicheskoye napryazheniye plenok kremniya [Effect of PECVD conditions on mechanical stress of silicon films], *Materials Physics and Mechanics*, 2018, Vol. 37, No. 1, pp. 67-72.
20. *Gusev E.Yu., Zhityaeva Yu.Yu., Kolomiitsev A.S., Gamaleev V.A., Kots I.N., Bykov A.V.* Issledovaniye rezhimov zhidkostnogo travleniya zhertvennogo sloya SiO₂ dlya formirovaniya mikromekhanicheskikh struktur na osnove Si*/SiO₂/Si [Study of liquid etching modes of a sacrificial SiO₂ layer for the formation of micromechanical structures based on Si*/SiO₂/Si], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, No. 2 (163), pp. 236-245.
21. *Sze S.* Fizika poluprovodnikovyykh priborov [Physics of semiconductor devices]. Vol. 1. Moscow: Mir, 1984, pp. 258-261.
22. *Epifanov G.I.* Fizika Osnovy mikroelektroniki [Physics Fundamentals of Microelectronics]. Moscow: Sovetskoye radio, 1971, pp. 145-156.

Статью рекомендовал к опубликованию д.т.н. доц. И.Е. Лысенко

Гелдаш Андрей Александрович – Южный федеральный университет; e-mail: andrewgeldash@gmail.com; 347922, г. Таганрог, ул. Шевченко, 2; тел.: 89515242876; кафедра нанотехнологий и микросистемной техники; аспирант.

Левенец Лев Эдуардович – кафедра нанотехнологий и микросистемной техники; студент.

Гусев Евгений Юрьевич – кафедра нанотехнологий и микросистемной техники; к.т.н.; доцент.

Джуплин Владимир Николаевич – e-mail: dzhuplin@mail.ru; кафедра нанотехнологий и микросистемной техники; к.т.н.; доцент.

Geldash Andrei Aleksandrovich – Southern Federal University; e-mail: andrewgeldash@gmail.com; 2, Shevchenko street, Taganrog, 347922, Russia; phone: +79515242876; the department of nanotechnologies and microsystems; graduate student.

Levenets Lev Eduardovich – the department of nanotechnologies and microsystems; student.

Gusev Evgenii Yrevich – the department of nanotechnologies and microsystems; senior researcher.

Dzhuplin Vladimir Nikolaevich – e-mail: dzhuplin@mail.ru; the department of nanotechnologies and microsystems; senior researcher.

И.В. Писаренко, Е.А. Рындин**ФОТОДЕТЕКТОР С УПРАВЛЯЕМОЙ ПЕРЕДИСЛОКАЦИЕЙ:
ДИФфуЗИОННО-ДРЕЙФОВАЯ МОДЕЛЬ И ИСПОЛЬЗОВАНИЕ
В ОПТИЧЕСКИХ МЕЖСОЕДИНЕНИЯХ**

Ранее для реализации оптических межсоединений в интегральных схемах были предложены инжекционные лазеры с двойными $A^{III}B^V$ наногетероструктурами, выполняющие функции источников и модуляторов излучения. Чтобы преобразовать короткие оптические импульсы, генерируемые лазерами-модуляторами, в электрические сигналы, необходимы технологически совместимые фотодетекторы с субпикосекундным временем отклика. Поскольку традиционные конструкции фоточувствительных полупроводниковых приборов не удовлетворяют предъявляемым требованиям, разработан перспективный метод построения быстродействующих фотодетекторов на основе принципа управляемой передислокации максимумов плотности носителей заряда в специально организованных квантовых областях. Данные оптоэлектронные приборы содержат продольный фоточувствительный p-i-n переход и поперечную управляющую гетероструктуру, в которую входят слои, выращенные методом молекулярно-лучевой эпитаксии при низких температурах, и два управляющих перехода. До наступления среза оптического импульса фотодетектор работает аналогично классическому p-i-n фотодиоду. Поперечное электрическое поле включается только во время заднего фронта лазерного импульса и передислоцирует максимумы плотности электронов и дырок из области поглощения в области с низкой подвижностью и коротким временем жизни, в результате чего время отклика сокращается до субпикосекундной величины. В предыдущих исследованиях быстродействие рассматриваемого фотодетектора оценивалось с помощью квантово-механической комбинированной модели, которая не учитывала некоторые важные аспекты протекающих в нем физических процессов. В данной статье предложена двумерная нестационарная диффузионно-дрейфовая модель, позволяющая провести детальный анализ транспорта носителей заряда в структуре фотодетектора с управляемой передислокацией с точки зрения полуклассического подхода. Для реализации представленной модели разработаны методика конечно-разностного численного моделирования на основе явного метода и прикладные программные средства. Полученные результаты диффузионно-дрейфового моделирования выявили необходимость использования дифференциального принципа подключения для компенсации токов смещения в цепи питания прибора. С учетом данной особенности разработаны фотоприемная схема, обеспечивающая как формирование результирующего электрического сигнала, так и требуемый режим подачи управляющего напряжения на контакты фотодетектора, и драйвер для лазеров-модуляторов.

Оптические межсоединения в интегральных схемах; лазеры-модуляторы; фотодетектор с управляемой передислокацией; диффузионно-дрейфовое приближение; численная модель; фотоприемная схема; драйвер.

I.V. Pisarenko, E.A. Ryndin**PHOTODETECTOR WITH CONTROLLED RELOCATION:
DRIFT-DIFFUSION MODEL AND APPLICATION IN OPTICAL
INTERCONNECTIONS**

Previously, we proposed an injection laser with a double $A^{III}B^V$ nanoheterostructure for the generation and modulation of light in optical interconnections for integrated circuits. To convert short optical pulses generated by the laser-modulator into electrical signals, a technologically compatible photodetector with subpicosecond response time is needed. Traditional designs of photosensitive semiconductor devices do not meet the specified requirements. Therefore, we developed a promising concept of a high-speed photodetector with controlled relocation of carrier density peaks within specially organized quantum regions. This optoelectronic device includes a longitudinal photosensitive p-i-n junction and a transverse control heterostructure, which contains

two low-temperature-grown layers and two control junctions. Before the trail of an optical pulse, the photodetector operates as a classical p-i-n photodiode. Transverse electric field is activated only during the back edge of a laser pulse. It relocates the peaks of electron and hole densities from the absorbing region to the regions with low carrier mobility and short lifetime. This process leads to the decrease in response time to a subpicosecond value. In our previous papers, we estimated the performance of the considered device using a quantum mechanical combined model that had not taken into account certain aspects of charge carrier transport in its structure. This paper is aimed at a proper semiclassical analysis of transients in the photodetector with controlled relocation by means of a two-dimensional drift-diffusion model. For the numerical implementation of the model, we develop a finite difference simulation technique based on the explicit method and applied software. According to the obtained results, it is reasonable to use the differential connection principle in order to compensate displacement currents in the supply circuit of the device. In view of this feature, we propose a circuit of optical receiver that provides the generation of resultant electrical signal as well as the required mode of the control voltage application to the photodetector contacts, and a driver circuit for the lasers-modulators.

Optical interconnections in integrated circuit; lasers-modulators; photodetector with controlled relocation; drift-diffusion approximation; numerical model; optical receiver circuit; driver circuit.

Введение. Физическое масштабирование транзисторов до нанометровых размеров и увеличение степени их интеграции на кристалле приводят к ухудшению характеристик металлических межэлементных соединений в интегральных схемах (ИС) [1, 2]. Оптоэлектронный подход к решению представленной проблемы предполагает замену отдельных металлических проводников с критическими параметрами (например, межъядерных соединений в многоядерных ИС) на конструктивно и технологически интегрированные системы оптической коммутации [3–5]. В рамках рассмотренного подхода разработаны инжекционные лазеры с двойными $A^{III}B^V$ наногетероструктурами, которые совмещают функции источников излучения и быстродействующих оптических модуляторов [6–8]. Для детектирования коротких лазерных импульсов, генерируемых лазерами-модуляторами, необходимы фотодетекторы с субпикосекундным временем отклика, которые могут быть монолитно интегрированы с другими элементами оптических межсоединений. Исследования, проведенные в работах [8, 9], показали, что стандартные конструкции интегральных фоточувствительных приборов не удовлетворяют предъявляемым требованиям. Поэтому был предложен перспективный метод построения фотодетекторов на основе принципа управляемой передислокации максимумов плотности носителей заряда в специально организованных квантовых областях [10], позволяющий существенно сократить длительность заднего фронта импульса фототока в цепи питания прибора [11, 12]. Разработка физикотопологических моделей нового типа фотодетектора для анализа протекающих в нем физических процессов представляет собой актуальную научную задачу.

Предложенная ранее квантово-механическая комбинированная модель фотодетектора с управляемой передислокацией (ФДУП) [11] основана на системе уравнений Шредингера-Пуассона и не учитывает влияние ряда полуклассических факторов (токи смещения, диффузию носителей заряда и двумерный характер их переноса), что требует реализации одной из численных моделей полуклассического подхода [13]. Поэтому первой целью данной статьи является разработка двумерной диффузионно-дрейфовой (ДД) модели ФДУП, а также соответствующих прикладных средств численного моделирования.

В интегральной системе оптической коммутации ФДУП функционирует в составе фотоприемной схемы, требования к которой сформулированы в данной работе на основе анализа результатов ДД моделирования ФДУП. В связи с этим вторая цель данной статьи – разработка принципов схемной реализации драйвера лазеров-модуляторов и фотоприемника на основе ФДУП, которые обеспечат создание оптических межсоединений с субпикосекундным временем отклика.

Диффузионно-дрейфовая модель. Как правило, полуклассическое моделирование транспорта носителей заряда в полупроводниковых приборах осуществляется в рамках ДД, квазигидродинамического (КГД) или гидродинамического (ГД) приближений [13]. Длительность переходных процессов в ФДУП на порядок меньше среднего времени релаксации энергии в материалах типа $A^{III}B^V$, поэтому запаздывание изменения подвижности носителей заряда относительно изменения электрического поля, описываемое КГД моделью, не оказывает существенного влияния на характеристики прибора, и зависимостью подвижности носителей заряда от времени можно пренебречь. Анализ баллистического транспорта носителей заряда в рамках ГД приближения также не актуален: область поглощения фотодетектора имеет длину не менее 200 нм в продольном направлении, что исключает возможность баллистического пролета, а поперечный транспорт ограничен потенциальными барьерами и осуществляется через область толщиной несколько нанометров, после чего электроны и дырки попадают в рекомбинационные слои с высокой концентрацией дефектов и быстро претерпевают рассеяние. Поэтому для моделирования ФДУП выбрано ДД приближение, в рамках которого возможен учет различных механизмов генерации, рекомбинации и рассеяния носителей заряда, а также двумерного характера транспортных процессов.

В общем случае ДД модель включает электронное и дырочное уравнения непрерывности и уравнение Пуассона, которые для структур с гетеропереходами могут быть представлены следующим образом [13, 14]:

$$\frac{\partial n}{\partial t} = -\nabla \cdot \{ \mu_n [n \cdot \nabla (\varphi + V_n) - \varphi_T \cdot \nabla n] \} + G - R_n; \quad (1)$$

$$\frac{\partial p}{\partial t} = \nabla \cdot \{ \mu_p [p \cdot \nabla (\varphi - V_p) + \varphi_T \cdot \nabla p] \} + G - R_p; \quad (2)$$

$$\nabla (\varepsilon \cdot \nabla \varphi) = -\frac{q}{\varepsilon_0} \cdot (p - n + N_d - N_a); \quad (3)$$

где n, p – концентрации электронов и дырок; t – время; μ_n, μ_p – подвижности электронов и дырок; φ – электростатический потенциал; V_n, V_p – гетероструктурные потенциалы в зоне проводимости и валентной зоне; φ_T – температурный потенциал; G – скорость генерации электронно-дырочных пар; R_n, R_p – скорости рекомбинации электронов и дырок; ε – относительная диэлектрическая проницаемость полупроводникового материала; q – элементарный заряд; ε_0 – электрическая постоянная; N_d, N_a – концентрации ионизированных доноров и акцепторов.

При заданных управляющих напряжениях на контактах граничные условия для системы ДД уравнений (1)–(3) имеют следующий вид [8]:

$$n = \left\{ \frac{N_d - N_a}{2} + \sqrt{\left(\frac{N_d - N_a}{2} \right)^2 + n_i^2} \right\} \cdot \exp \left(-\frac{\varphi_{Sch}}{\varphi_T} \right); \quad (4)$$

$$p = \left\{ -\frac{N_d - N_a}{2} + \sqrt{\left(\frac{N_d - N_a}{2} \right)^2 + n_i^2} \right\} \cdot \exp \left(\frac{\varphi_{Sch}}{\varphi_T} \right); \quad (5)$$

$$\varphi = \varphi_T \ln \left(\frac{n}{n_i} \right) + U - \varphi_{Sch} = -\varphi_T \ln \left(\frac{p}{n_i} \right) + U - \varphi_{Sch}; \quad (6)$$

$$\frac{\partial n}{\partial \xi} = 0; \quad \frac{\partial p}{\partial \xi} = 0; \quad \frac{\partial \varphi}{\partial \xi} = 0, \quad (7)$$

где ϕ_{Sch} – высота барьера Шоттки (для омических контактов $\phi_{\text{Sch}} = 0$); n_i – собственная концентрация носителей заряда; U – напряжение, прикладываемое к контакту в рассматриваемый момент времени; ξ – нормаль к поверхности контакта. Выражения (4)–(6) используются на омических контактах и контактах с барьерами Шоттки, а уравнения (7) – на свободных от контактов границах.

Начальные условия для модели (1)–(3) определяются численным решением системы ДД уравнений в стационарной форме при начальных значениях нестационарных параметров модели, определяемых исследуемым режимом работы прибора. При моделировании ФДУП наибольший интерес представляет переходной процесс, протекающий при освещении активной области прибора и последующем включении управляющего напряжения. Следовательно, начальные условия вычисляются при рабочем значении напряжения питания и нулевых мощности оптического излучения и управляющем напряжении.

Для описания процессов генерации и рекомбинации носителей заряда в данном исследовании применяются следующие аналитические модели [8, 15]:

модель, вычисляющая скорости электронной и дырочной рекомбинации R_n , R_p через времена жизни τ_n , τ_p и концентрации неравновесных носителей заряда Δn , Δp в момент времени t :

$$R_n(t) = \Delta n(t) / \tau_n; \quad R_p(t) = \Delta p(t) / \tau_p; \quad (8)$$

1) модель скорости биполярной оптической генерации G в волноводном фотодетекторе со встроенным оптическим резонатором:

$$G(t) = \frac{P_{\text{in}}(t) \cdot Q}{V_{\text{res}} \cdot E_{\text{ph}}} \quad (9)$$

$$Q = (1 - R_1) [1 - \exp(-\alpha L)] [1 + R_2 \cdot \exp(-\alpha L)] [1 - \sqrt{R_1 R_2} \cdot \exp(-\alpha L)]^{-2}, \quad (10)$$

где $P_{\text{in}}(t)$ – мощность падающего оптического излучения в момент времени t ; Q – квантовая эффективность фотодетектора; V_{res} – объем резонатора (в случае ФДУП – i -области поглощения излучения); E_{ph} – энергия фотонов, равная ширине запрещенной зоны в области поглощения; R_1 , R_2 – коэффициенты отражения полу- и полностью отражающего зеркал оптического резонатора; α – коэффициент поглощения полупроводника; L – длина резонатора, соответствующая условию $\lambda = 2L / m$, где λ – резонансная длина волны оптического излучения, m – натуральное число.

Одним из фундаментальных свойств LT слоев, входящих в состав управляющей гетероструктуры ФДУП, является низкая подвижность носителей заряда в сравнении с областями, выращенными при стандартных температурных условиях. В данной работе используются экспериментальные значения подвижностей для полупроводниковых материалов различных областей прибора, приведенные в работах [16, 17].

Таким образом, ДД модель ФДУП включает базовые дифференциальные уравнения (1)–(3), граничные условия (4)–(7), начальные условия, получаемые из решения системы уравнений (1)–(3) в стационарной форме, и модели процессов рекомбинации и генерации (8)–(10).

Методика и программные средства моделирования. Предлагаемая в данном исследовании методика ДД численного моделирования объединяет эффективную схему дискретизации и экономичный вычислительный алгоритм, необходимые для устойчивого и сходящегося решения системы дифференциальных уравнений в частных производных, и имеет следующие характерные особенности:

1. Уравнения (1)–(3) рассматриваются в комбинированном базисе переменных $\{n, p, \varphi, F_n, F_p\}$, где F_n, F_p – переменные Слотбума [20].
2. Все переменные, входящие в базовые уравнения ДД модели, нормируются стандартными коэффициентами [20].
3. Решение дифференциальных уравнений осуществляется численным методом конечных разностей [21].
4. Схема дискретизации уравнений непрерывности основана на явном [21] и противопоточном [22] разностных методах.
5. Уравнение Пуассона дискретизируется стандартным способом, представленным в работе [23].
6. Решение системы дискретных ДД уравнений осуществляется модифицированным итерационным методом Гуммеля [23]. Поскольку уравнения непрерывности дискретизированы явным методом, на каждом шаге временной сетки реализуется цикл по невязке электрического потенциала φ , в котором на каждой итерации рассчитываются концентрации носителей заряда n и p , а затем решается система линейных алгебраических уравнений относительно переменной φ . Перед переходом на следующую итерацию или на следующий шаг по времени концентрации пересчитываются в величины F_n, F_p .
7. Для расчета начальных условий стационарная задача в базисе переменных $\{F_n, F_p, \varphi\}$ решается итерационным методом Ньютона [23].

Применение явного и противопоточного методов дискретизации в комплексе с итерационным методом Гуммеля позволяет свести задачу реализации нестационарной диффузионно-дрейфовой модели к решению на каждом шаге по времени (либо итерации) одной СЛАУ с размерностью, соответствующей рассматриваемой координатной сетке, и выполнению ряда сопутствующих математических операций. Очевидно, что подобная методика характеризуется существенно меньшими затратами временных и вычислительных ресурсов на получение результатов моделирования по сравнению с полунявным или неявным методами.

Для реализации рассмотренных выше диффузионно-дрейфовой модели и методики конечно-разностного моделирования разработан пакет прикладных программ в свободной среде GNU Octave, использующей MATLAB-совместимый язык высокого уровня [24]. Представленные программные средства позволяют моделировать переходные и стационарные процессы в структурах быстродействующих $A^{III}B^V$ фотодетекторов с различными электрофизическими и конструктивно-технологическими параметрами в широком диапазоне управляющих воздействий.

Результаты моделирования. Конфигурация структуры ФДУП, исследованная в данной работе в рамках ДД приближения полуклассического подхода, представлена на рис. 1. Продольная фоточувствительная p - i - n структура образована высоколегированными p^+ - и n^+ -GaAs областями длиной по 100 нм и расположенной между ними 200-нм i -InGaAsSb областью поглощения оптического излучения. Направление распространения лазерного излучения перпендикулярно плоскости показанного сечения. Поперечная управляющая гетероструктура включает верхний управляющий переход Шоттки с n -AlGaAs областью толщиной 25 нм, верхний и нижний 10-нм LT-слои с малым временем жизни и низкой подвижностью носителей заряда, 5-нм i -область и нижний управляющий переход Шоттки с p -AlGaAs областью толщиной 25 нм. Управляющее напряжение $(U_{c,1} - U_{c,2}) = 1,5$ В прикладывается к верхнему и нижнему контактам Шоттки, а напряжение питания $(U_{s,1} - U_{s,2}) = -0,5$ В – к боковым омическим контактам, соединенным с высоколегированными областями.

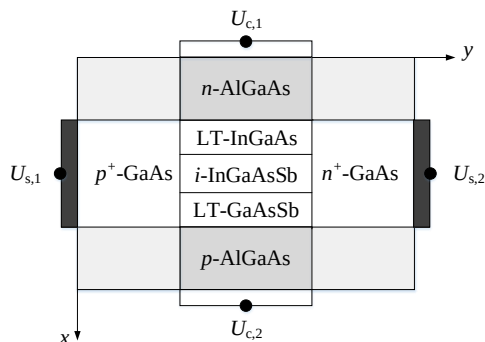


Рис. 1. Схематическое представление структуры ФДУП с двумя управляющими переходами Шоттки n- и p-типов проводимости; области диэлектрика выделены штриховкой

ФДУП имеет несимметричную двумерную структуру, поэтому эквивалентные емкости его полупроводниковых переходов и выпрямляющих контактов металл-полупроводник не сбалансированы. Поэтому достаточно быстрое включение управляющего напряжения во время заднего фронта оптического импульса вызывает протекание токов перезарядки емкостей через питающие и управляющие контакты. Результаты нестационарного диффузионно-дрейфового моделирования, которые иллюстрируют данный негативный эффект, приведены на рис. 2, а. Как следует из представленных кривых, амплитуда тока смещения в 3,75 раза превышает амплитуду фототока на правом контакте и почти в 10 раз – на левом контакте, что неприемлемо с точки зрения требований к измерительному сигналу фотодетектора. С одной стороны, вычислительные эксперименты показали, что амплитуда токов смещения может быть снижена до адекватных значений (порядка 5 мкА) путем задания определенных разностей потенциалов между контактами прибора. В таком случае обеспечивается баланс между эквивалентными емкостями переходов структуры. С другой стороны, практическая реализация данного подхода достаточно проблематична, так как исследуемый фотодетектор предназначен для работы в составе оптических межсоединений ИС и прецизионная настройка управляющих воздействий не представляется возможной.

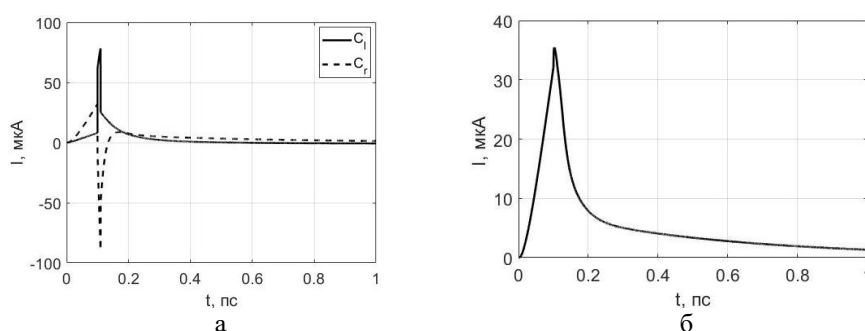


Рис. 2. Всплески тока на левом (C_l) и правом (C_r) контактах питания ФДУП при включении управляющего напряжения во время заднего фронта оптического импульса длительность 0,1 пс (а); фототок в цепи питания ФДУП при компенсации токов смещения за счет использования дифференциального принципа подключения (б)

Для решения рассмотренной проблемы и получения адекватного измерительного сигнала предлагается использовать дифференциальный принцип, в соответствии с которым измерительный сигнал формируется двумя ФДУП, подключенными к быстродействующему дифференциальному усилителю. Обе фоточувствительные структуры имеют одинаковые конструктивно-технологические параметры и управляются одинаковыми напряжениями, но только одна из структур соединена с оптическим волноводом и освещается лазерными импульсами. При обработке сигналов дифференциальным усилителем токи перезарядки емкостей одинаковых фотодетекторов компенсируют друг друга, и на выходе схемы формируется выделенный полезный сигнал – фототок. Результаты ДД моделирования, подтверждающие возможность использования дифференциального принципа для компенсации токов смещения в цепи питания ФДУП, приведены на рис. 2, б. На приведенном графике длительность импульсного отклика, измеренная по половине его амплитуды, равна около 0,08 пс (при длительности оптического импульса 0,1 пс).

Принципы схемной реализации оптических межсоединений. Подводя итоги физико-топологического моделирования ФДУП в рамках квантово-механического и полуклассического подходов, следует отметить следующие особенности, которые необходимо учитывать при разработке схем приемной и передающей частей интегральных систем оптической коммутации:

- 1) в отличие от традиционных устройств, схема фотоприемника с ФДУП должна не только усиливать выходной сигнал фотодетектора, но и своевременно подавать импульсы напряжения на управляющие контакты;
- 2) импульсы управляющего напряжения и оптические импульсы должны иметь фиксированную длительность;
- 3) для компенсации токов перезарядки емкостей в цепи питания ФДУП две структуры с одинаковыми параметрами, одна из которых освещается лазерным излучением, подключаются ко входам дифференциального каскада фотоприемника;
- 4) чтобы обеспечить высокое быстродействие оптических межсоединений, фотоприемная схема и драйвер должны быть построены на основе электронной элементной базы с субпикосекундным временем задержки.

Предлагаемая в данной статье принципиальная схема драйвера для функционально-интегрированных инжекционных лазеров-модуляторов [6–8] изображена на рис. 3. Устройство построено с использованием логических элементов на основе комплементарных туннельно-связанных наногетероструктур (КТСН) с управляемой передислокацией максимумов плотности носителей заряда, обеспечивающих требуемое быстродействие и подробно рассмотренных в работе [7]. Поскольку лазеры-модуляторы и КТСН-элементы управляются противофазными сигналами, схема имеет прямой и инверсный входы V_{IN} , $\overline{V_{IN}}$. Функциональный узел на элементах DD1.1 – DD5.1 формирует по каждому переднему и каждому заднему фронтам сигнала V_{IN} субпикосекундные импульсы (V_R и V_F , соответственно), длительность которых определяется временем задержки КТСН-инвертора $t_{NOT} < 0,2$ пс [7]. Временные диаграммы сигналов V_{IN} , V_R и V_F показаны на рис. 4.

Соответствующие инвертированные сигналы $\overline{V_R}$ и $\overline{V_F}$ формируются аналогичным образом функциональным узлом на элементах DD1.2 – DD5.2. При этом конструктивно два рассмотренных функциональных узла представляют собой единый интегральный блок вследствие функциональной и структурной интеграции противофазных инверторов DD1 – DD3 и элементов 2ИЛИ-НЕ/2И-НЕ DD4, DD5, схемы которых представлены в работе [7].

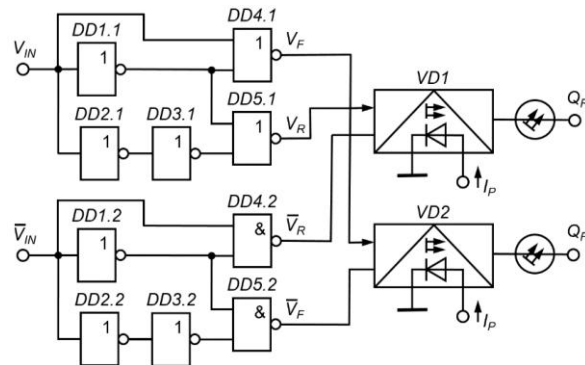


Рис. 3. Принципиальная схема передающей части интегральной системы оптической коммутации, построенная на основе быстродействующих лазеров-модуляторов и логических элементов с КТСН

Субпикосекундные противофазные импульсы V_R , $\overline{V_R}$ и V_F , $\overline{V_F}$ подаются на управляющие входы функционально интегрированных лазеров-модуляторов VD1, VD2, формирующих соответствующие оптические импульсы Q_R и Q_F при неизменном во времени токе накачки I_p . Временные диаграммы оптических сигналов Q_R и Q_F , приведенные на рис. 4, отражают переходные процессы в инжекционных лазерах. Скорость затухания оптических импульсов в каждой серии определяется плотностью тока накачки и временем жизни фотонов в резонаторе лазера-модулятора [6–8].

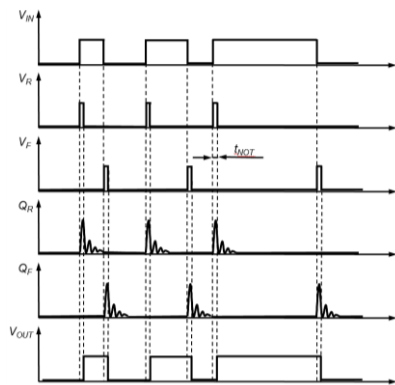


Рис. 4. Временные диаграммы, поясняющие принцип работы интегральной системы оптической коммутации

Приемная часть интегральной системы оптической коммутации на основе ФДУП, принципиальная схема которой приведена на рис. 5, позволяет восстановить входные сигналы V_{IN} , $\overline{V_{IN}}$ на выходных контактах V_{OUT} , $\overline{V_{OUT}}$ с использованием оптических импульсов Q_R и Q_F . При этом необходимо решить следующие основные задачи:

1) обеспечить изменение выходных сигналов V_{OUT} , $\overline{V_{OUT}}$ только по первому в каждой серии затухающих оптических импульсов Q_R и Q_F , как показано на рис. 4;

2) устранить влияние токов перезаряда емкостей управляющих цепей фотоприемников на импульсы фототока и, соответственно, на выходные сигналы V_{OUT} , $\overline{V_{OUT}}$.

Для решения первой задачи в качестве основного функционального узла приемной части системы оптической коммутации используется быстродействующий триггер на КТСН VT11 – VT14 и управляющих транзисторах VT10, VT15 (рис. 5).

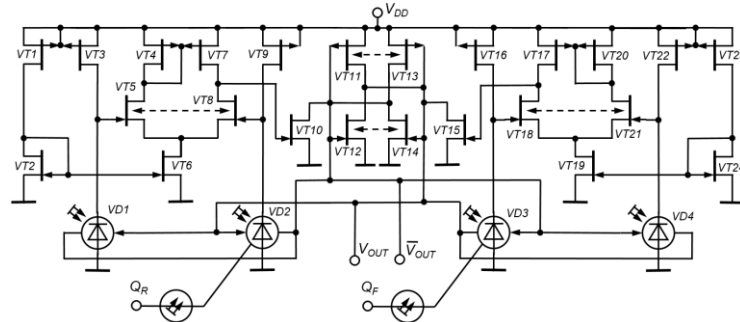


Рис. 5. Принципиальная схема быстродействующего фотоприемника на основе ФДУП

Для решения второй задачи помимо основных фотоприемников VD2, VD3, на которые по интегральным оптическим линиям подаются лазерные импульсы Q_R и Q_F , используются дополнительные фотоприемники VD1, VD4, на которые лазерные импульсы не подаются, а также дифференциальные усилители VT4 – VT8 и VT17 – VT21. При этом управляющие электрические сигналы поступают параллельно на основные и дополнительные фотоприемники с выходов триггера V_{OUT} , $\overline{V_{OUT}}$ (см. рис. 5). Для повышения коэффициента усиления и коэффициента подавления синфазных сигналов дифференциальных каскадов используются токовые зеркала (VT4, VT7), (VT2, VT6), (VT17, VT20), (VT19, VT24) и источники стабильного тока на транзисторах VT1, VT3, VT9, VT16, VT22, VT23.

При отсутствии входных оптических импульсов Q_R и Q_F разности потенциалов на затворах КТСН VT5, VT8 и КТСН VT18, VT21 равны нулю, в результате чего на выходах дифференциальных усилителей (стоки VT8 и VT18) формируются низкие уровни напряжения, транзисторы VT10 и VT15 закрыты, и КТСН-триггер VT11 – VT14 находится в режиме хранения текущего состояния.

Пусть текущее состояние триггера определяется выходными уровнями $V_{OUT} \approx 0$, $\overline{V_{OUT}} \approx V_{DD}$, обеспечивающими пространственную передислокацию носителей заряда в фотоприемниках VD1, VD2 в активную область, а в фотоприемниках VD3, VD4 – в LT области. При этом VD2 находится в состоянии готовности к приему оптического импульса, а VD3 – в состоянии формирования заднего фронта фототока. При поступлении оптического импульса Q_R на фотоприемник VD2, на выходе дифференциального усилителя (сток VT8) формируется усиленный импульс напряжения, отпирающий транзистор VT10 и устанавливающий триггер через время задержки, примерно равное t_{NOT} , в состояние $V_{OUT} \approx V_{DD}$, $\overline{V_{OUT}} \approx 0$, инициирующее пространственную передислокацию фотогенерированных носителей в VD2 из активной области в LT области, что приводит к формированию заднего фронта импульса фототока длительностью порядка t_{NOT} с одновременным переходом фотоприемника VD3 в состояние готовности к приему оптического импульса. При этом последующие в текущей серии оптические импульсы Q_R не будут изменять установленное первым импульсом состояние триггера (см. рис. 4).

Аналогично, при поступлении оптического импульса Q_F на фотоприемник VD3, на выходе дифференциального усилителя (сток VT18) формируется усиленный импульс напряжения, отпирающий транзистор VT15 и устанавливающий триггер через время задержки, примерно равное t_{NOT} , в состояние $V_{OUT} \approx 0$, $\overline{V_{OUT}} \approx V_{DD}$, инициирующее пространственную передислокацию фотогенерированных носителей в VD3 из активной области в LT области, что приводит к формированию заднего фронта импульса фототока длительностью порядка t_{NOT} с одновременным переходом фотоприемника VD2 в состояние готовности к приему очередного оптического импульса Q_R .

Для усиления импульса фототока и подавления токов перезаряда емкостей в предложенной на рис. 5 схеме используются дифференциальные усилители, входы которых (затворы VT5, VT8 и VT18, VT21) подключены к двум парам фотоприемников VD1, VD2 и VD3, VD4, питаемым от источников тока VT3, VT9 и VT16, VT21. Поскольку управляющие электрические импульсы подаются на каждую пару фотоприемников параллельно, а оптические импульсы Q_R и Q_F поступают лишь на фотоприемники VD2 и VD3, дифференциальные каскады усиливают отклик фотоприемников VD2 и VD3 на оптические импульсы, а синфазные импульсы, обусловленные перезарядом емкостей, эффективно подавляют. В результате на выходах V_{OUT} , $\overline{V_{OUT}}$ приемной части системы оптической коммутации (см. рис. 5) воспроизводятся сигналы V_{IN} , $\overline{V_{IN}}$ передающей части (см. рис. 3) с временной задержкой порядка $t_{NOT} < 0,5$ пс.

Заключение. Таким образом, в данной статье разработаны двумерная ДД модель быстродействующего оптического сенсора нового типа – ФДУП, а также средства численного моделирования (методика и пакет прикладных программ), которые позволили исследовать протекающие в структуре переходные процессы. Согласно полученным результатам моделирования, для компенсации токов смещения в цепи питания прибора, возникающих при подаче коротких импульсов управляющего напряжения, целесообразно использовать дифференциальный принцип, реализованный в данной работе в виде схемы, в которой две одинаковые фоточувствительные структуры подключаются к дифференциальному усилителю, причем только один фотодетектор освещается оптическим излучением. При дифференциальном подключении ФДУП с двумя управляющими переходами Шоттки обеспечивает импульсный фотоотклик с амплитудой более 30 мкА и длительностью (измеренной по полуамплитуде) около 0,08 пс при длительности оптического импульса 0,1 пс. На основе анализа полученных результатов моделирования предложены принципы схемной реализации драйвера лазеров-модуляторов и быстродействующего фотоприемника для интегральных систем оптической коммутации, обеспечивающего эффективную компенсацию токов смещения и временную задержку передачи сигнала порядка 0,5 пс.

Поддержка. Результаты получены в рамках задания № FENW-2020-0022 на выполнение научных исследований, осуществляемых за счет средств федерального бюджета.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Clark R., Tapily K., Yu K.-H., Hakamata T., Consiglio S., O'Meara D., Wajda C., Smith J., Leusink G. New Process Technologies Required for Future Devices and Scaling // APL Materials. – 2018. – Vol. 6. – P. 058203.
2. Brain R. Interconnect Scaling: Challenges and Opportunities // Proceedings of 2016 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA. – IEEE, 2016. – P. 9.3.1-9.3.4.

3. Chen Z., Gu H., Yang Y., Bai L., Li H. A Power Efficient and Compact Optical Interconnect for Network-on-Chip // IEEE Computer Architecture Letters. – 2014. – Vol. 13, No. 1. – P. 5-8.
4. Bashir J., Peter E., Sarangi S.R. A Survey of On-Chip Optical Interconnects // ACM Computing Surveys. – 2019. – Vol. 51, No. 6. – P. 115.
5. Miller D.A.B. Optical interconnects to electronic chips // Applied Optics. – 2010. – Vol. 49, No. 25. – P. F59-F70.
6. Ryndin E.A., Konoplev B.G. Active Region of a Functionally Integrated Laser-Modulator // Quantum Electronics. – 2019. – Vol. 49, No. 6. – P. 563-569.
7. Konoplev B.G., Ryndin E.A., Denisenko M.A. Components of Integrated Microwave Circuits Based on Complementary Coupled Quantum Regions // Russian Microelectronics. – 2015. – Vol. 44, No. 3. – P. 190-196.
8. Pisarenko I., Ryndin E. Drift-Diffusion Simulation of High-Speed Optoelectronic Devices // Electronics. – 2019. – Vol. 8, No. 1. – P. 106.
9. Pisarenko I.V., Ryndin E.A. Numerical Simulation of High-Speed $A^{III}B^V$ Photodetectors Within Drift-Diffusion Approximation // Journal of Physics: Conference Series. – 2019. – Vol. 1368, No. 4. – P. 042052.
10. Webb K.J., Cohen E.B., Melloch R.E. Fabrication and Operation of a Velocity Modulation Transistor // IEEE Transactions on Electron Devices. – 2001. – Vol. 48, No. 12. – P. 2701-2709.
11. Pisarenko I., Ryndin E. Photodetector with Controlled Relocation of Carrier Density Peaks: Concept and Numerical Simulation // Photonics. – 2020. – Vol. 7, No. 1. – P. 21.
12. Pisarenko I.V., Ryndin E.A. High-speed photodetector with controlled relocation of carrier density peaks // Journal of Physics: Conference Series. – 2020. – Vol. 1482. – P. 012037.
13. Абрамов И.И. Проблемы и принципы физики и моделирования приборных структур микро- и наноэлектроники. Ч. II. Модели полуклассического подхода // Нано- и микро-системная техника. – 2006. – № 9. – С. 26–36.
14. Абрамов И.И. Основы моделирования элементов микро- и наноэлектроники. – Saarbrücken: LAP LAMBERT Academic Publishing, 2016. – 434 с.
15. Шалимова К.В. Физика полупроводников. 4-е изд. – СПб.: Изд-во «Лань», 2010. – 400 с.
16. Palankovski V., Quay R. Analysis and Simulation of Heterostructure Devices. – Wien: Springer-Verlag, 2004. – 257 p.
17. Currie M. Low-Temperature Grown Gallium Arsenide (LT-GaAs) High-Speed Detectors // Photodetectors: Materials, Devices and Applications. Ed. by B. Nabet. – Cambridge: Woodhead Publishing, 2016. – Chapter 5. – P. 121-155.
18. Analyze Semiconductor Devices at the Fundamental Level with the Semiconductor Module // COMSOL Website. – URL: <https://www.comsol.ru/semiconductor-module> (дата обращения: 01.05.2020).
19. Genius: 3D Parallel Device Simulator // Cogenda Website. – URL: <http://www.cogenda.com/article/Genius> (дата обращения: 01.05.2020).
20. Vasiliska D., Goodnick S.M., Klimeck G. Computational Electronics: Semiclassical and Quantum Device Modeling and Simulation. – Boca Raton: CRC Press, 2010. – 782 p.
21. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы. 7-е изд. – М.: БИНОМ. Лаборатория знаний, 2017. – 636 с.
22. Куликова И.В., Лысенко И.Е., Приступчик Н.К., Лысенко А.С. Численное решение нестационарной фундаментальной системы уравнений полупроводника в диффузионно-дрейфовом приближении // Известия ЮФУ. Технические науки. – 2014. – № 9 (158). – С. 106-111.
23. Абрамов И.И. Моделирование физических процессов в элементах кремниевых интегральных микросхем. – Мн.: БГУ, 1999. – 189 с.
24. GNU Octave: Scientific Programming Language // GNU Octave Website. – URL: <https://www.gnu.org/software/octave/> (дата обращения: 01.05.2020).

REFERENCES

1. Clark R., Tapily K., Yu K.-H., Hakamata T., Consiglio S., O'Meara D., Wajda C., Smith J., Leusink G. New Process Technologies Required for Future Devices and Scaling, *APL Materials*, 2018, Vol. 6, pp. 058203.
2. Brain R. Interconnect Scaling: Challenges and Opportunities, *Proceedings of 2016 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA*. IEEE, 2016, pp. 9.3.1-9.3.4.

3. Chen Z., Gu H., Yang Y., Bai L., Li H. A Power Efficient and Compact Optical Interconnect for Network-on-Chip, *IEEE Computer Architecture Letters*, 2014, Vol. 13, No. 1, pp. 5-8.
4. Bashir J., Peter E., Sarangi S.R. A Survey of On-Chip Optical Interconnects, *ACM Computing Surveys*, 2019, Vol. 51, No. 6, pp. 115.
5. Miller D.A.B. Optical interconnects to electronic chips, *Applied Optics*, 2010, Vol. 49, No. 25, pp. F59-F70.
6. Ryndin E.A., Konoplev B.G. Active Region of a Functionally Integrated Laser-Modulator, *Quantum Electronics*, 2019, Vol. 49, No. 6, pp. 563-569.
7. Konoplev B.G., Ryndin E.A., Denisenko M.A. Components of Integrated Microwave Circuits Based on Complementary Coupled Quantum Regions, *Russian Microelectronics*, 2015, Vol. 44, No. 3, pp. 190-196.
8. Pisarenko I., Ryndin E. Drift-Diffusion Simulation of High-Speed Optoelectronic Devices, *Electronics*, 2019, Vol. 8, No. 1, pp. 106.
9. Pisarenko I.V., Ryndin E.A. Numerical Simulation of High-Speed $A^{III}B^V$ Photodetectors With- in Drift-Diffusion Approximation, *Journal of Physics: Conference Series*, 2019, Vol. 1368, No. 4, pp. 042052.
10. Webb K.J., Cohen E.B., Melloch R.E. Fabrication and Operation of a Velocity Modulation Transistor, *IEEE Transactions on Electron Devices*, 2001, Vol. 48, No. 12, pp. 2701-2709.
11. Pisarenko I., Ryndin E. Photodetector with Controlled Relocation of Carrier Density Peaks: Concept and Numerical Simulation, *Photonics*, 2020, Vol. 7, No. 1, pp. 21.
12. Pisarenko I.V., Ryndin E.A. High-speed photodetector with controlled relocation of carrier density peaks, *Journal of Physics: Conference Series*, 2020, Vol. 1482, pp. 012037.
13. Abramov I.I. Problemy i printsipy fiziki i modelirovaniya pribornykh struktur mikro- i nanoelektroniki. Ch. II. Modeli poluklassicheskogo podkhoda [Problems and Principles of Physics and Simulation of Device Structures for Micro- and Nanoelectronics. Part II. Models of Semiclassical Approach], *Nano- i mikrosistemnaya tekhnika* [Nano- and Microsystems Technology], 2006, No. 9, pp. 26-36.
14. Abramov I.I. Osnovy modelirovaniya elementov mikro- i nanoelektroniki [Fundamentals of Micro- and Nanoelectronic Elements' Simulation]. Saarbrücken: LAP LAMBERT Academic Publishing, 2016, 434 p.
15. Shalimova K.V. Fizika poluprovodnikov [Semiconductor Physics]. 4 ed. Saint Petersburg: Izd-vo «Lan'», 2010, 400 p.
16. Palankovski V., Quay R. Analysis and Simulation of Heterostructure Devices. Wien: Springer-Verlag, 2004, 257 p.
17. Currie M. Low-Temperature Grown Gallium Arsenide (LT-GaAs) High-Speed Detectors, *Photodetectors: Materials, Devices and Applications*. Ed. by B. Nabet. Cambridge: Woodhead Publishing, 2016, Chapter 5, pp. 121-155.
18. Analyze Semiconductor Devices at the Fundamental Level with the Semiconductor Module // COMSOL Website. Available at: <https://www.comsol.ru/semiconductor-module> (accessed 01 May 2020).
19. Genius: 3D Parallel Device Simulator // Cogenda Website. Available at: <http://www.cogenda.com/article/Genius> (accessed 01 May 2020).
20. Vasileska D., Goodnick S.M., Klimeck G. Computational Electronics: Semiclassical and Quantum Device Modeling and Simulation. Boca Raton: CRC Press, 2010, 782 p.
21. Bakhvalov N.S., Zhidkov N.P., Kobel'kov G.M. Chislennyye metody [Numerical Methods]. 7 ed. Moscow: BINOM. Laboratoriya znaniy, 2017, 636 p.
22. Kulikova I.V., Lysenko I.E., Pristupchik N.K., Lysenko A.S. Chislennoe reshenie nestatsionarnoy fundamental'noy sistemy uravneniy poluprovodnika v diffuzionno-dreyfovom priblizhenii [Numerical Solution of Fundamental Equation System for Semiconductor in Drift-Diffusion Approximation], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2014, No. 9 (158), pp. 106-111.
23. Abramov I.I. Modelirovanie fizicheskikh protsessov v elementakh kremnievykh integral'nykh mikroskhem [Simulation of Physical Processes in Elements of Silicon Integrated Circuits]. Minsk: BGU, 1999, 189 p.
24. GNU Octave: Scientific Programming Language, *GNU Octave Website*. Available at: <https://www.gnu.org/software/octave/> (accessed 01 May 2020).

Статью рекомендовал к опубликованию д.т.н. И.Е. Лысенко.

Писаренко Иван Вадимович – Южный федеральный университет; e-mail: ivan123tgn@yandex.ru; 347942, г. Таганрог, ул. Нижняя Линия, 8; тел.: +79185978535; аспирант.

Рындин Евгений Адальбертович – Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина); e-mail: rynenator@gmail.com; 197376, г. Санкт-Петербург, ул. Профессора Попова, 5а, 5-й корпус; тел.: +78122343164; д.т.н.; профессор.

Pisarenko Ivan Vadimovich – Southern Federal University; e-mail: ivan123tgn@yandex.ru; 8, Nizhnyaya Liniya street, Taganrog, 347942, Russia; phone: +79185978535; postgraduate student.

Ryndin Eugeny Adal'bertovich – Saint Petersburg Electrotechnical University "LETI"; e-mail: rynenator@gmail.com; 5a, Professora Popova street, build. 5, Saint Petersburg, 197376, Russia; phone: +78122343164; dr. of eng. sc.; professor.

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 ссылки 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/>.