



№5-2023

ISSN 1999-9429

ИЗВЕСТИЯ ЮФУ

ТЕХНИЧЕСКИЕ НАУКИ

- Алгоритмы обработки информации
- Анализ данных и моделирование
- Электроника, приборостроение и радиотехника

ИЗВЕСТИЯ ЮФУ. ТЕХНИЧЕСКИЕ НАУКИ IZVESTIYA SFedU. ENGINEERING SCIENCES

Свидетельство о регистрации средства массовой информации

ПИ № ФС77-28889 от 12.07.2007

Федеральная служба по надзору в сфере массовых коммуникаций, связи
и охраны культурного наследия

Научно-технический и прикладной журнал

Издается с 1995 года, до середины 2007 года под названием «Известия ТРТУ»

Подписной индекс ПС704

№ 5 (235). 2023 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

Курейчик В.В. (гл. редактор); Кравченко Ю.А. (зам. гл. редактора); Курейчик В.М. (зам. гл. редактора); Бородинский И.М. (ученый секретарь); Абрамов С.М.; Агеев О.А.; Бабенко Л.К.; Боженюк А.В.; Веселов Г.Е.; Гайдук А.Р.; Горбанёва О.И.; Каляев И.А.; Коноплев Б.Г.; Касьянов А.О.; Левин И.И.; Куповых Г.В.; Лысенко И.Е.; Малюков С.П.; Медведев М.Ю.; Мельник Э.В.; Никитов С.А.; Обуховец В.А.; Панич А.Е.; Петров В.В.; Пшихопов В.Х.; Редько В.Г.; Румянцев К.Е.; Семенихина Д.В.; Сергеев Н.Е.; Стемпковский А.Л.; Сухинов А.И.; Тополов В.Ю.; Турулин И.И.; Угольницкий Г.А.; Усов А.Б.; Федосов В.П.; Целых А.Н.; Юханов Ю.В.

Учредитель Южный федеральный университет.

Издатель Южный федеральный университет.

Ответственный за выпуск Самойлов А.Н.

Технический редактор Ярошевич Н.В.

Оригинал-макет выполнен Ярошевич Н.В.

Дата выхода в свет 08.12. 2023 г. Формат 70×108 $\frac{1}{16}$. Бумага офсетная.

Офсетная печать. Усл. печ. л. – 18,7. Уч.-изд. л. – 13,5.

Заказ № . Тираж 250 экз.

Адрес издателя: 344090, г. Ростов-на-Дону, пр. Стачки, 200/1, тел. 8(863)243-41-66.

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)243-41-66.

Адрес редакции: 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, 2023

СОДЕРЖАНИЕ

РАЗДЕЛ I. АЛГОРИТМЫ ОБРАБОТКИ ИНФОРМАЦИИ

С.М. Гушанский, В.И. Божич, В.С. Потапов ИССЛЕДОВАНИЕ И РАЗРАБОТКА СХЕМ С ОПТИМИЗАЦИЕЙ ПО ГЛУБИНЕ В АЛГОРИТМЕ КВАНТОВОЙ ПРИБЛИЖЕННОЙ ОПТИМИЗАЦИИ.....	6
Ю.А. Заргарян, Е.В. Заргарян, В.И. Кошенский, К.О. Кирсанов РАЗРАБОТКА СИСТЕМЫ ИЗБИРАТЕЛЬНОГО ПОДАВЛЕНИЯ БЕСПРОВОДНЫХ СИГНАЛОВ	16
В.А. Костюков, М.Ю. Медведев, В.Х. Пшихопов АЛГОРИТМ ПЛАНИРОВАНИЯ ПУТИ В ДВУХМЕРНОЙ СРЕДЕ С ПОЛИГОНАЛЬНЫМИ ПРЕПЯТСТВИЯМИ НА КЛАССЕ КУСОЧНО-ЛОМАНЫХ ТРАЕКТОРИЙ	34
Д.Б. Шокарев, Р.Ж. Чочаев, А.Н. Щелоков, С.В. Гаврилов РАЗРАБОТКА АЛГОРИТМА ДЕТАЛЬНОГО РАЗМЕЩЕНИЯ НА ПЛИС	48
В.В. Курейчик, В.В. Бова, А.Ю. Халенков МНОГОУРОВНЕВЫЙ ПОДХОД ДЛЯ РЕШЕНИЯ ЗАДАЧИ ДВУМЕРНОЙ УПАКОВКИ ГЕОМЕТРИЧЕСКИХ ФИГУР СЛОЖНЫХ ФОРМ.....	58
А.А. Олейникова, В.В. Золотарев КОНЦЕПЦИЯ УПРАВЛЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТЬЮ НА ОСНОВЕ ЦИКЛА НЕПРЕРЫВНОГО ДЕТЕКТИРОВАНИЯ И РЕАГИРОВАНИЯ НА ИНЦИДЕНТЫ БЕЗОПАСНОСТИ ИНФОРМАЦИИ.....	66
В.С. Усатюк, С.И. Егоров, А.П. Локтионов, Е.А. Титенко, И.Е. Чернецкая АРХИТЕКТУРА НЕЙРОННЫХ СЕТЕЙ НА ОСНОВЕ КОДОВ НА ГРАФАХ.....	81
Я.В. Куликова, В.А. Литвиненко РАЗРАБОТКА МЕТОДА ОПРЕДЕЛЕНИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА С ПОМОЩЬЮ ЧАТ-БОТА.....	93
И.И. Левин, К.Н. Алексеев ПРЕОБРАЗОВАНИЕ СОРТИРУЮЩИХ СЕТЕЙ ДЛЯ РАЗНОЙ СТЕПЕНИ ПАРАЛЛЕЛИЗМА	104

РАЗДЕЛ II. АНАЛИЗ ДАННЫХ И МОДЕЛИРОВАНИЕ

А.В. Боженюк, М.В. Князева, О.В. Косенко, Е.Е. Косенко ИСПОЛЬЗОВАНИЕ ПЕРИОДИЧЕСКИХ НЕЧЕТКИХ ГРАФОВ В ЗАДАЧАХ УПРАВЛЕНИЯ ПРОИЗВОДСТВЕННЫМ ОБОРУДОВАНИЕМ	119
В.В. Гилка, А.С. Кузнецова, А.А. Молдовская, Д.Ф. Эль-Аит ПРОВЕРКА РАБОТОСПОСОБНОСТИ МОДЕЛИ И МЕТОДА УДАЛЕННОГО МОНИТОРИНГА СОСТОЯНИЯ ЗДОРОВЬЯ НА ПРИМЕРЕ ОТКЛОНЕНИЙ ПОКАЗАТЕЛЕЙ ТЕМПЕРАТУРЫ ТЕЛА ЧЕЛОВЕКА.....	127
Л.А. Гладков, Н.В. Гладкова, В.М. Курейчик ПОДСИСТЕМА АВТОМАТИЧЕСКОГО АННОТИРОВАНИЯ ТЕКСТОВ НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ.....	138
Д.В. Тимошенко ОСОБЕННОСТИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ЗАДАЧ ЭЛЕКТРОДИНАМИКИ ПРИЗЕМНОГО СЛОЯ.....	152
Э.В. Мельник, И.Б. Сафроненкова, А.Ю. Таранов ОНТОЛОГИЧЕСКИЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ ПЕРЕРАСПРЕДЕЛЕНИЯ ВЫЧИСЛИТЕЛЬНОЙ НАГРУЗКИ В РАСПРЕДЕЛЕННОЙ СИСТЕМЕ МОНИТОРИНГА С МОБИЛЬНЫМИ КОМПОНЕНТАМИ НА БАЗЕ РАСПРЕДЕЛЁННОГО РЕЕСТРА.....	163

**РАЗДЕЛ III. ЭЛЕКТРОНИКА, ПРИБОРОСТРОЕНИЕ
И РАДИОТЕХНИКА**

В.В. Бурлака, С.В. Гулаков, А.Ю. Головин, Д.С. Мироненко СИСТЕМА БЕСПРОВОДНОЙ ПЕРЕДАЧИ ЭЛЕКТРИЧЕСКОЙ ЭНЕРГИИ.....	174
А.А. Жук БЫСТРОДЕЙСТВУЮЩИЕ ВЫХОДНЫЕ КАСКАДЫ ОПЕРАЦИОННЫХ УСИЛИТЕЛЕЙ С ДИФФЕРЕНЦИРУЮЩЕЙ ЦЕПЬЮ КОРРЕКЦИИ ПЕРЕХОДНОГО ПРОЦЕССА.....	183
Б.М. Антипин, Е.М. Виноградов, О.А. Симонина, С.Л. Фёдоров АНАЛИЗ ТРЕБОВАНИЙ К ПАРАМЕТРАМ ИЗЛУЧЕНИЙ РАДИОТЕХНИЧЕСКОГО ОБОРУДОВАНИЯ БЕСПИЛОТНЫХ АВИАЦИОННЫХ СИСТЕМ.....	192
В.В. Бурлака, С.В. Гулаков, А.Ю. Головин, Д.С. Мироненко УНИВЕРСАЛЬНЫЙ ДВУНАПРАВЛЕННЫЙ DC-AC ПРЕОБРАЗОВАТЕЛЬ.....	204

CONTENT

SECTION I. INFORMATION PROCESSING ALGORITHMS

S.M. Gushanskiy, V.I. Bozhich, V.S. Potapov RESEARCH AND DEVELOPMENT OF DEPTH OPTIMIZED CIRCUITS IN QUANTUM APPROXIMATE OPTIMIZATION ALGORITHM.....	6
Yu.A. Zargaryan, E.V. Zargaryan, V.I. Koshensky, K.O. Kirsanov DEVELOPMENT OF A SYSTEM OF SELECTIVE SUPPRESSION OF WIRELESS SIGNALS	16
V.A. Kostyukov, M.Yu. Medvedev, V.Kh. Pshikhopov AN ALGORITHM FOR PATH PLANNING IN A TWO-DIMENSIONAL ENVIRONMENT WITH POLYGONAL OBSTACLES ON A CLASS OF PIECEWISE POLYLINE TRAJECTORIES	35
D.B. Shokarev, R.Zh. Chochaev, A.N. Schelokov, S.V. Gavrilov DEVELOPMENT OF THE DETAILED PLACEMENT ALGORITHM FOR FPGAS	49
V.V. Kureichik, V.V. Bova, A.Yu. Khalenkov MULTILEVEL APPROACH TO TWO-DIMENSIONAL PACKING PROBLEM FOR GEOMETRIC FIGURES OF COMPLEX SHAPES.....	58
A.A. Oleynikova, V.V. Zolotarev THE CONCEPT OF INFORMATION SECURITY MANAGEMENT BASED ON A CYCLE OF INFORMATION SECURITY INCIDENTS CONTINUOUS DETECTION AND RESPONSE.....	67
V.S. Usatyuk, S.I. Egorov, A.P. Loktionov, E.A. Titenko, I.E. Chernetskaya NEURAL NETWORK ARCHITECTURE BASED ON GRAPH CODES.....	82
Ya.V. Kulikova, V.A. Litvinenko DEVELOPMENT OF A METHOD FOR DETERMINING THE EMOTIONAL STATE OF A PERSON USING A CHAT BOT.....	93
I.I. Levin, K.N. Alekseev TRANSFORMATION OF SORTING NETWORKS FOR DIFFERENT DEGREES OF PARALLELISM	105

SECTION II. DATA ANALYSIS AND MODELING

A.V. Bozhenyuk, M.V. Knyazeva, O.V. Kosenko, E.E. Kosenko USING PERIODIC FUZZY GRAPHS IN PRODUCTION EQUIPMENT CONTROL PROBLEMS	119
V.V. Gilka, A.S. Kuznetsova, A.A. Moldovskaya, D.F. El-Ait VERIFICATION OF THE MODEL AND METHOD FUNCTIONALITY FOR REMOTE HEALTH MONITORING ILLUSTRATED BY THE DEVIATIONS IN HUMAN BODY TEMPERATURE INDICATORS	127
L.A. Gladkov, N.V. Gladkova, V.M. Kureichik SUBSYSTEM FOR AUTOMATIC TEXT ANNOTATION BASED ON MACHINE LEARNING METHODS	138
D.V. Timoshenko FEATURES OF MATHEMATICAL MODELING IN THE GROUND LAYER ELECTRODYNAMICS	152
E.V. Melnik, I.B. Safronenkova, A.Yu. Taranov ONTOLOGICAL APPROACH TO SOLVING THE WORKLOAD RELOCATION PROBLEM IN A DISTRIBUTED MONITORING SYSTEM WITH MOBILE COMPONENTS BASED ON A DISTRIBUTED LEDGER.....	163

**SECTION III. ELECTRONICS, INSTRUMENT ENGINEERING
AND RADIO ENGINEERING**

V.V. Burlaka, S.V. Gulakov, A.Y. Golovin, D.S. Mironenko THE WIRELESS ELECTRIC POWER TRANSFER SYSTEM.....	174
A.A. Zhuk HIGH-SPEED OUTPUT STAGES OF OPERATIONAL AMPLIFIERS WITH DIFFERENCING CIRCUIT CORRECTION OF TRANSITION PROCESS	183
B.M. Antipin, E.M. Vinogradov, O.A. Simonina, S.L. Fedorov THE ANALYSIS OF REQUIREMENTS FOR RADIATION PARAMETERS OF RADIO EQUIPMENT IN UNMANNED AERIAL SYSTEMS.....	193
V.V. Burlaka, S.V. Gulakov, A.Y. Golovin, D.S. Mironenko UNIVERSAL BIDIRECTIONAL DC-AC CONVERTER.....	205

Раздел I. Алгоритмы обработки информации

УДК 004.032

DOI 10.18522/2311-3103-2023-5-6-16

С.М. Гушанский, В.И. Божич, В.С. Потапов

ИССЛЕДОВАНИЕ И РАЗРАБОТКА СХЕМ С ОПТИМИЗАЦИЕЙ ПО ГЛУБИНЕ В АЛГОРИТМЕ КВАНТОВОЙ ПРИБЛИЖЕННОЙ ОПТИМИЗАЦИИ

Одной из основных проблем, с которой сталкиваются исследователи в области квантовых вычислений, является проблема шума в квантовых системах. Шум может существенно ограничивать производительность квантовых алгоритмов. Именно в этом контексте актуализируется наше исследование, направленное на разработку и оптимизацию квантовых алгоритмов с фокусом на глубине. Глубина квантовых цепей – это один из критически важных параметров в разработке квантовых алгоритмов. Оптимизированные схемы с улучшенной глубиной имеют потенциал существенно снизить влияние шума, что, в свою очередь, должно привести к повышению эффективности. Мы стремимся предложить решения, которые не только учитывают технические ограничения, но и предоставляют практически применимые результаты для квантовых вычислений в контексте оптимизационных задач. В рамках данного исследования проводится анализ применения квантового алгоритма приближенной оптимизации для решения сложных задач комбинаторной оптимизации. Однако в процессе использования данного алгоритма сталкиваемся с серьезным ограничением – шумом в квантовой системе, что существенно снижает его эффективность. Для преодоления влияния шума и повышения эффективности квантовых алгоритмов, было предложено несколько методов. В данной статье представлен жадный эвристический алгоритм, направленный на уменьшение воздействия шума. Основная цель этого алгоритма заключается в поиске остова дерева минимальной высоты. Это, в свою очередь, приводит к сокращению общей глубины квантовых схем и минимизации количества вентилей CNOT, что является ключевым моментом в оптимизации квантовых вычислений. Через проведение численного анализа было продемонстрировано, что предложенный жадный эвристический алгоритм способен существенно увеличить вероятность успешного завершения каждой итерации в задаче поиска максимального разреза в графе в 10 раз. Кроме того, исследование подтверждает, что средняя глубина квантовой схемы, созданной предложенным эвристическим алгоритмом, все еще линейно зависит от размера входных данных, но угол наклона этой линейной зависимости снижается с 1 до 0,11 благодаря использованию предложенного метода.

Моделирование; квантовый алгоритм; кубит; модель квантового компьютера; запутанность; суперпозиция; квантовый оператор.

S.M. Gushanskiy, V.I. Bozhich, V.S. Potapov

RESEARCH AND DEVELOPMENT OF DEPTH OPTIMIZED CIRCUITS IN QUANTUM APPROXIMATE OPTIMIZATION ALGORITHM

One of the main challenges faced by researchers in the field of quantum computing is the problem of noise in quantum systems. Noise can significantly limit the performance of quantum algorithms. It is in this context that our research, aimed at the development and optimization of quantum algorithms with a focus on depth, is updated. The depth of quantum circuits is one of the critical parameters in the development of quantum algorithms. Optimized circuits with improved depth have the

potential to significantly reduce the impact of noise, which in turn should lead to improved efficiency. We aim to provide solutions that not only address technical constraints, but also provide practical results for quantum computing in the context of optimization problems. This study analyzes the use of a quantum approximate optimization algorithm for solving complex combinatorial optimization problems. However, in the process of using this algorithm we encounter a serious limitation – noise in the quantum system, which significantly reduces its efficiency. To overcome the influence of noise and improve the efficiency of quantum algorithms, several methods have been proposed. This paper presents a greedy heuristic algorithm aimed at reducing the impact of noise. The main goal of this algorithm is to find a spanning tree of minimum height. This, in turn, reduces the overall depth of quantum circuits and minimizes the number of CNOT gates, which is key to optimizing quantum computing. Through numerical analysis, it was demonstrated that the proposed greedy heuristic algorithm is capable of significantly increasing the probability of successful completion of each iteration in the problem of finding the maximum cut in a graph by 10 times. Moreover, the study confirms that the average depth of the quantum circuit generated by the proposed heuristic algorithm is still linearly dependent on the size of the input data, but the slope of this linear dependence is reduced from 1 to 0.11 by using the proposed method.

Modeling; quantum algorithm; qubit; model of a quantum computer; entanglement; superposition; quantum operator.

Введение. Квантовые алгоритмы приближенной оптимизации (КАПО) представляет собой гибридный метод, который объединяет квантовые и классические вычисления для решения задач комбинаторной оптимизации. Этот алгоритм работает на основе специальных математических структур, таких как проблемный гамильтониан и миксерный гамильтониан H_p , которые используются для кодирования задачи оптимизации. В процессе выполнения КАПО применяются две параметризованные унитарные операции $U(\theta, \gamma)$ и $U(\theta, \beta)$, которые последовательно воздействуют на начальное квантовое состояние $U(H_p; \gamma) = \exp(-i\gamma H_p)$ и $U(H_M; \beta) = \exp(-i\beta H_M)$. Параметры этих операций инициализируются случайным образом, и после нескольких итераций, они обновляются с использованием классического оптимизатора. Этот процесс повторяется с обновленными параметрами, и целью является приближение к оптимальному решению задачи оптимизации.

В литературе также рассматривались методы смягчения, которые позволяют уменьшить воздействие шума на выполнение КАПО и других гибридных квантово-классических алгоритмов. Эти методы включают в себя изменения в гамильтониане миксера и функции стоимости, направленные на снижение шума и увеличение скорости сходимости алгоритма. Одним из ключевых аспектов улучшения производительности квантовых схем является уменьшение ошибок, вызванных вентилями CNOT. Эти вентили являются источником ошибок в современных квантовых устройствах, и методы оптимизации с целью уменьшения их воздействия на шум играют важную роль в совершенствовании квантовых алгоритмов. Таким образом, КАПО исследует различные аспекты оптимизации в квантовых вычислениях, и его развитие может привести к созданию более эффективных квантовых алгоритмов, что делает его перспективным направлением для будущих исследований в области квантовой вычислительной технологии.

1. КАПО на основе максимального разреза графа и поиска в глубину. Традиционный подход к параметризации квантовых схем (ПАКПО) основан на двух параметризованных унитарных операторах: $U(H_p; \gamma) = \exp(-i\gamma H_p)$ и $U(H_M; \beta) = \exp(-i\beta H_M)$. При работе с графом $G = (V, E)$, где $|V| = n$, квантовая схема инициализируется суперпозицией n кубитов, что представляет собой один шаг инициализации с глубиной 1. Точно так же оператор $U(\theta, \beta)$ применяется ко всем кубитам одновременно. Оператор $U(\theta, \gamma)$ зависит от гамильтониана задачи и может быть представлен уравнением (2).

Схема для реализации оператора, соответствующего ребру (j, k), представлена на рис. 1. На каждом шаге может выполняться несколько таких операторов одновременно, соответствующих непересекающимся рёбрам. Например, на рисунке 2 операторы $U(\theta, \alpha)$ и $U(\theta, \beta)$ выполняются одновременно на одном шаге.

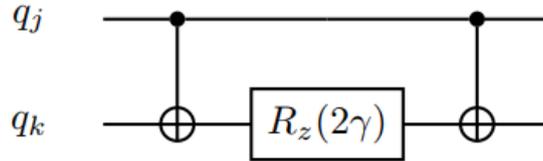


Рис. 1. Схема для оператора $U(\theta, \gamma)$

В оставшейся части статьи термин "ПАКПО" будет использоваться для обозначения параметризованных квантовых схем для задачи максимального разреза графа. В данном контексте граф $G = (V, E)$ является связным, неориентированным и невзвешенным, где $|V| = n$ и $|E| = m$. Для полноты изложения на рис. 2 представлена схема для случая $p = 1$, соответствующая 2-регулярному графу с четырьмя вершинами.

$$U(H_p, \gamma) = \prod_{(j,k) \in E} U(H^{(j,k)}_p) = \prod_{(j,k) \in E} \exp(-i\gamma(\frac{I - Z_j Z_k}{2})) \quad (1)$$

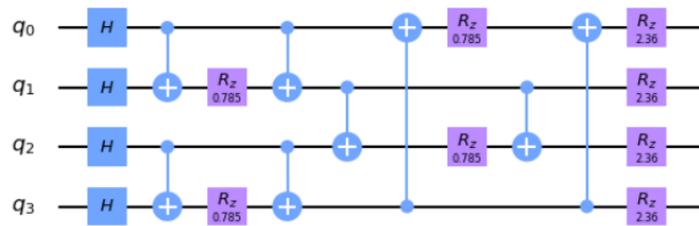


Рис. 2. Схема АКПО для $p = 1$, соответствующая 2-регулярному графу с четырьмя вершинами

2. Оптимизация квантовой схемы на основе поиска в глубину. В данной работе рассматривается метод оптимизации выполнения квантовых схем, основанный на деревьях и операторах Адамара-Коэнси. Данный метод предназначен для уменьшения количества вентилях CNOT (контроля над соседним кубитом) в квантовых схемах. Метод начинается с выбора корневой вершины в графе и создания дерева, которое представляет собой ациклический подграф графа, содержащий все вершины и $n - 1$ ребро. Затем схема Адамара-Коэнси разделяется на две части: одну с операторами, связанными с ребрами в дереве, и другую с операторами, связанными с другими ребрами. Операторы, связанные с ребрами в дереве, должны быть выполнены в определенном порядке, учитывая их положение в дереве. Если ребро появляется раньше другого в дереве, то оператор, соответствующий этому ребру, должен быть выполнен раньше. Для ребер, которые включены в дерево, операторы применяются с условиями управления и цели вентиля CNOT, где управление – вершина u , инцидентная ребру, уже находящемуся в дереве, и цель – вершина v . Это обеспечивает оптимизацию выполнения операторов для квантовых схем, снижая количество вентилях CNOT. Однако этот метод может быть неэффективным для больших графов, так как он увеличивает глубину схемы, что может привести к превышению времени когерентности оборудования.

Увеличение глубины также делает схему более уязвимой к декогеренции, что требует баланса между уменьшением числа вентилях CNOT и увеличением глубины схемы. Важным результатом работы является выявление ограничений метода оптимизации на основе деревьев для случаев, когда увеличение глубины схемы может стать проблемой.

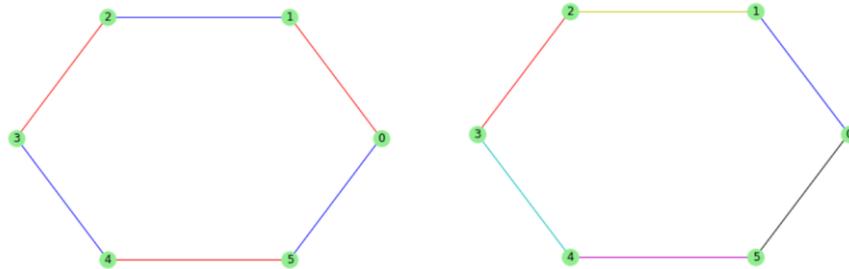


Рис. 3. Этапы работы для $U(H_p; \gamma)$ для традиционного АКПО и оптимизированного на основе поиска в глубину

Этот метод оптимизации, направленный на сокращение числа вентилях CNOT, может быть применен к любому корневому остовному дереву, при условии выполнения двух указанных выше условий относительно упорядочения ребер в дереве. Метод поиска в глубину, фактически, представляет собой способ построения корневого остовного дерева для графа. Далее, вместо того чтобы специально искать конкретное дерево, мы сосредотачиваемся на поиске корневого остовного дерева для графа G , что соответствует описанной цели.

3. Формулировка задачи оптимизации. Сосредоточимся исключительно на оптимизации оператора $U(H)$, который связан с каждым ребром. Под термином "схема" или "глубина" мы будем иметь в виду исключительно схему, соответствующую оператору $U(H)$. Максимальная высота дерева с n вершинами равна $n - 1$, что также соответствует глубине соответствующей схемы. Мы можем утверждать, что остовное дерево с более низкой высотой может привести к схеме с меньшей глубиной. В этом случае, использование поиска в ширину (BFS) может предоставить остовное дерево с минимальной высотой. Однако на рис. 4,а,б показаны два дерева с разной высотой, но с одинаковой глубиной схемы. На этих рисунках значения, связанные с ребрами, показывают уровень, на котором можно применять оператор, соответствующий данному ребру, для оптимизации (т.е. уменьшения количества вентилях CNOT). Схема, соответствующая обоим деревьям, как показано на рис. 5, остается одинаковой. Эти два дерева и соответствующая им схема наглядно демонстрируют, что простое снижение высоты дерева не гарантирует получения схемы с меньшей глубиной. Таким образом, использование поиска в ширину (BFS) вместо поиска в глубину (DFS) не всегда приведет к созданию схемы с меньшей глубиной.

На иллюстрации, изображенной на рис. 4, представлены два дерева разной высоты, и каждое ребро дерева имеет целочисленную метку, которая указывает на шаг, на котором можно применить оператор $U(H)$ к этому ребру. Максимальное значение этих меток соответствует глубине контура. Высота первого дерева равна 3, а второго – 2. Однако важно отметить, что оба дерева приводят к одной и той же схеме, изображенной на рис. 5.

Причина, по которой дерево в правой части с высотой 2 на рис. 4 не способно сократить глубину схемы, заключается в том, что вентилях CNOT, соответствующие двум соседним ребрам, не могут выполняться на одном и том же шаге. Следо-

вательно, простое уменьшение высоты остоного дерева недостаточно для сокращения глубины схемы. Кроме того, на рис. 6 представлены два дерева с одинаковой высотой, но разной длиной цепей.

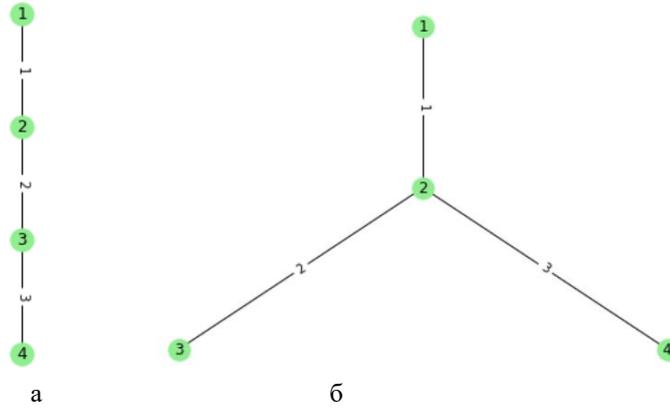


Рис. 4. (а) Остоное дерево высоты 3; (б) Остоное дерево высоты 2

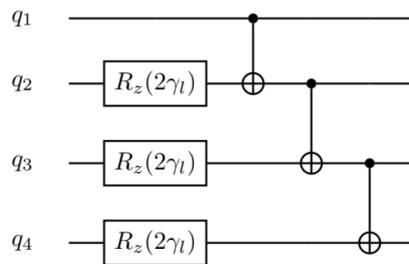


Рис. 5. Квантовая схема $U(H_p; \gamma)$, соответствующая обоим деревьям на рис. 4

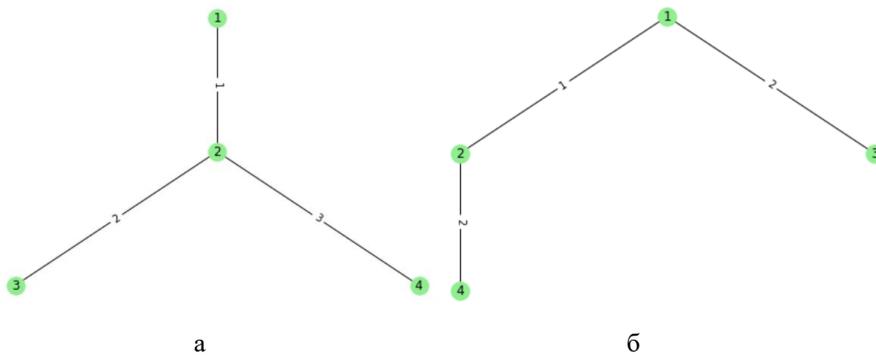


Рис. 6. (а) Пример дерева высоты 2; (б) Еще один пример дерева высоты 2

Теперь определим несколько терминов, чтобы легко уточнить требования к корневому остоному дереву, которое приведет к схеме с меньшей глубиной.

1. Коэффициент ветвления [16]. Коэффициент ветвления вершины v определяется как количество вершин, обнаруженных в корневом остоном дереве из v .

2. Уровень: если вершина v обнаружена в корневом остоном дереве из вершины u , то уровень вершины v равен уровню вершины $u + 1$. Уровень корневой вершины равен 0.

3. Отложенный старт: Отложенный старт [17] определяется как явление, при котором вершины v_1, \dots, v_k обнаруживаются в остоном дереве из одной вершины v и принадлежат одному уровню. Это связано с тем, что смежные ребра имеют общую вершину, и одновременные операции CNOT невозможны с общим управляющим или целевым кубитом [18]. Поэтому операция, соответствующая ребру (v, v_i) , задерживается до тех пор, пока все операции, соответствующие ребрам (v, v_j) не заполнены. Оператор $U(H^{(v, v_i)}_p)$ можно использовать раньше всего на уровне $v + i$.

4. Реализация функции стоимости. Предложим функцию стоимости, учитывающую следующие наблюдения:

1) Если коэффициенты ветвления вершин очень высоки, то соответствующий контур [19] будет страдать от задержки старта, что приведет к увеличению глубины. С другой стороны, если коэффициент ветвления вершин очень мал, то высота дерева, а значит, и глубина цепи будут увеличиваться.

2) Между двумя вершинами u и v лучше разветвить вершину на более низком уровне дерева, чтобы у ребер в этой ветви все еще была возможность выполняться параллельно с другими ребрами на более высоком уровне даже после задержки.

3) Для графов с меньшим количеством вершин коэффициент ветвления должен быть низким, чтобы избежать увеличения глубины из-за отложенного старта. Однако по мере увеличения числа вершин необходимо допускать более высокий коэффициент ветвления, чтобы уменьшить высоту дерева. Принимая во внимание все три критерия, указанные выше, предложим функцию стоимости C_v , которая будет связана с каждой вершиной v . Пусть n будет количеством вершин в графе, l_v и v_{bf} будут уровнем и текущим коэффициентом ветвления вершины v соответственно, и B – максимальный коэффициент ветвления, определенный для любой вершины остоного дерева, тогда $C_v = (n - l_v)(B - v_{bf})$. При выращивании остоного дерева из корневой вершины к дереву добавляется то ребро $(v; w)$, для которого функция стоимости C_v максимальна. Заметим здесь, что для нового ребра $(v; w)$ функция стоимости не зависит от вершины w , а на вершине v , из которой это ребро обнаружено. Предлагаемый алгоритм избегает ветвления в вершине, для которой $v_{bf} \geq B$.

На самом деле, когда $v_{bf} = B$, функция стоимости имеет вклад 0. Таким образом, если для двух вершин $u \neq v$, $v_{bf} = v_{bf} < B$, алгоритм выбирает для ветвления ту вершину, которая имеет более низкий уровень. Это гарантирует, что отложенный запуск будет ближе к корню, так что у этих ветвей все еще есть возможность для параллельного выполнения с некоторыми ветвями более высокого уровня. Кроме того, если $v_{bf} > B$, произведение с $(n - 1)$ приводит к значительно низким значениям для низкого l . Это сильно препятствует ветвлению более чем B на более низких уровнях дерева, чтобы предотвратить чрезмерно отложенный запуск. Алгоритм 1 представляет, как сгенерировать остоное дерево [20], которое имеет компромисс между высотой дерева и коэффициентом ветвления.

Алгоритм 1. Генерация корневого связующего дерева на основе функции стоимости

Вход: граф $G = (V; E)$, $|V| = n$, $|E| = m$; максимальный коэффициент ветвления B .

Результат: корневое остовное дерево T графа G .

$T = \{\}$.

$u_{bf} \leftarrow 0$ для всех вершин u .

$r \leftarrow$ случайно выбранная начальная вершина.

Посещенная вершина = $\{r\}$.

$r_{bf} = r_{bf} + 1$.

ребра_для_добавления = $\text{neigh}(r)$.

while $|\text{Посещенная_вершина}| < n$ do

$e = \text{ребра_для_добавления}[0]$.

$c = 0$.

 для всех ребер $= (u; v) \in$ массиву добавления ребер выполнить

$$C_v = (n - l_v)(B - v_{bf}).$$

 если $C_v > c$, то

$$c = C_v.$$

$e = \text{край}$.

 конец, если

конец цикла

$$T = T \cup \{e\}$$

Посещенная_вершина = $\text{Посещенная_вершина} \cup \{y\}$, где $e = (x; y)$.

$$x_{bf} = x_{bf} + 1$$

Удалите все ребра формы $(*; q)$ из ребер, которые нужно добавить.

для всех ребер $= (p; q) \in \text{neigh}(y)$ выполнить

 если $q \notin \text{Посещенная_вершина}$, то

$$\text{Ребра_для_добавления} = \text{ребра_для_добавления} \cup \{\text{edge}\}$$

 конец цикла если

конец цикла

конец

Давайте проиллюстрируем работу метода поиска в глубину и предлагаемого алгоритма на примере графа, изображенного на рис. 7. Сначала на рис. 8 мы представим традиционную схему с вероятностью КАПО (первого достижения цели) $p = 1$ для данного графа. Затем на рис. 9 мы представим два остовных дерева этого графа. Первое остовное дерево на рис. 9,а создается с использованием метода поиска в глубину, в то время как второе дерево на рис. 9,б создается с использованием алгоритма 1 с параметром $B = 3$. На рис. 10,а,б мы демонстрируем оптимизированные схемы для КАПО с вероятностью $p = 1$ для графа на рис. 7. Оптимизированные схемы создаются с использованием метода поиска в глубину и алгоритма 1 соответственно. Значения параметров γ и β выбираются случайным образом.

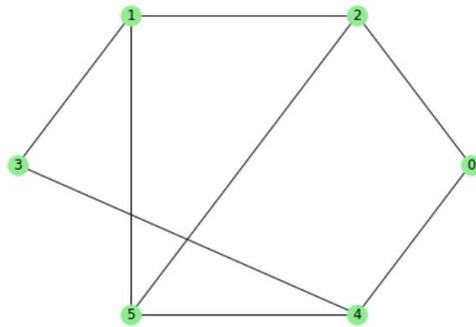


Рис. 7. Пример графа, состоящего из шести вершин

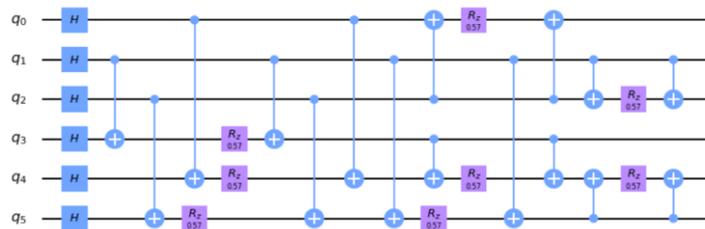


Рис. 8. Традиционная схема КАПО $p = 1$, соответствующая $U(H_p; \gamma)$ для графа на рис. 7

Глубина схем на рис. 8, а также на рис. 10,а,б составляет 11, 14 и 12, соответственно. Эти значения были получены с использованием функции `.depth()` [21] в рамках программы Qiskit [22].

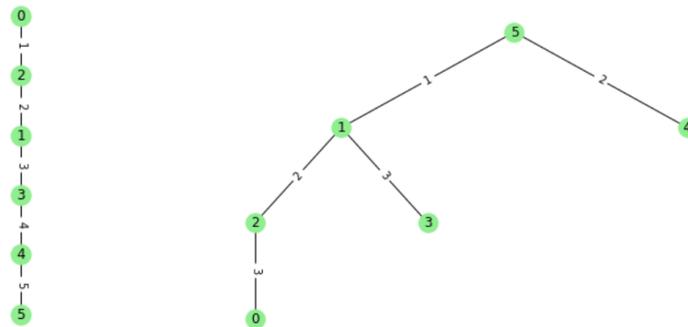


Рис. 9. Два остовных дерева графа. Левое дерево было создано с применением метода поиска в глубину, в то время как правое дерево сформировано с применением алгоритма 1 с параметром $B = 3$

Подчеркнем, что как в случае оптимизированных схем [23], так и в сравнении с традиционным методом Адамара-Каши, оба оптимизированных варианта требуют большей глубины. Тем не менее, можно утверждать, что оптимизированная схема, представленная на рис. 10,б, можно считать более предпочтительной, так как она требует на 5 вентилях CNOT меньше, чем схема на рис. 8, при увеличении глубины всего лишь на 1.

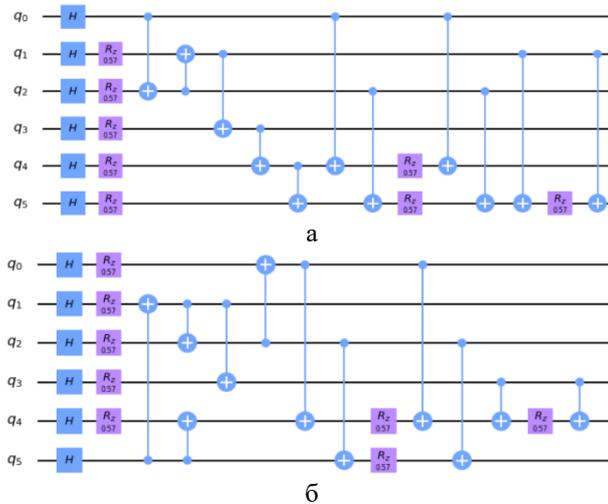


Рис. 10 (а) Оптимизированная схема с использованием метода поиска в глубину, (б) Оптимизированная схема с использованием предложенного алгоритма 1

Заключение. На данный момент активно разрабатывается теория квантовых вычислений [24], несмотря на то, что идея квантовых компьютеров была предложена еще Р. Фейнманом в 1982 году. Научные исследования в этой области продолжаются. В данной статье предлагается эвристический алгоритм с полиномиальным временем, который находит остоное дерево с корнем [25], сохраняя сокращение числа вентилей CNOT, но при этом ограничивая увеличение глубины. Этот метод способен снизить увеличение глубины по сравнению с аналогичными методами, что, в свою очередь, повышает вероятность успешного выполнения алгоритма. Это обусловлено сохранением сокращения элементов CNOT и снижением увеличения глубины, что делает схему менее чувствительной к ошибкам релаксации.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Farhi E., Goldstone J., and Gutmann S. A quantum approximate optimization algorithm, *arXiv preprint arXiv:1411.4028*, 2014.
2. Cerezo M., Arrasmith A., Babbush R., Benjamin S., Endo S., Fujii K., McClean J., Mitarai K., Yuan X., Cincio L., et al. Variational quantum algorithms, *arXiv preprint arXiv:2012.09265*, 2020.
3. Linke N.M., Gutierrez M., Landsman K.A., et al. Fault-tolerant quantum error detection, *Science Advances*, 2017, 3 (10): e1701074. Available from: <https://doi.org/10.1126/sciadv.1701074>.
4. Vuillot C. Is error detection helpful on IBM 5q chips?, *Quantum Information and Computation*, 2018, Vol. 18, No. 11-12, pp. 0949-0964.
5. Barron G. and Wood C. Measurement error mitigation for variational quantum algorithms, *arXiv preprint arXiv:2010.08520*, 2020.
6. Endo S., Benjamin S., and Li Y. Practical quantum error mitigation for near-future applications, *Physical Review X*, 2018, 8 (3):031027.
7. Endo S., Cai Z., Benjamin S., and Yuan X. Hybrid quantum-classical algorithms and quantum error mitigation, *Journal of the Physical Society of Japan*, 2021, 90 (3):032001.
8. Zhu L., Tang H., Barron G., Calderon-Vargas F., Mayhall N., Barnes E., and Economou S. An adaptive quantum approximate optimization algorithm for solving combinatorial problems on a quantum computer, *arXiv preprint arXiv:2005.10258*, 2020.

9. Larkin J., Jonsson M., Justice D., and Guerreschi G. Evaluation of quantum approximate optimization algorithm based on the approximation ratio of single samples, *arXiv e-prints, pages arXiv-2006*, 2020.
10. Barkoutsos P., Nannicini G., Robert A., Tavernelli I., and Woerner S. Improving variational quantum optimization using cvar, *Quantum*, 2020, 4:256.
11. Harper R, Flammia S.T. Fault-tolerant logical gates in the IBM quantum experience, *Phys. Rev. Lett.*, 2019, 122:080504. Available from: <https://link.aps.org/doi/10.1103/PhysRevLett.122.080504>.
12. Hales S. Hallgren An improved quantum Fourier transform algorithm and applications, *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, November 12–14, 2000, pp. 515.
13. Guzik V., Gushanskiy S., Polenov M., Potapov V. Complexity Estimation of Quantum Algorithms Using Entanglement Properties, *16th International Multidisciplinary Scientific GeoConference, Bulgaria*, 2016, pp. 20-26.
14. Guzik V., Gushanskiy S., Polenov M., Potapov V. Models of a quantum computer, their characteristics and analysis, *9th International Conference on Application of Information and Communication Technologies (AICT)*. Institute of Electrical and Electronics Engineers, 2015, pp. 583-587.
15. Collier D. The Comparative Method. In: Finifter A.W. (ed.) *Political Sciences: The State of the Discipline II*. American Science Association. Washington, DC, 1993, pp. 105-119.
16. Olukotun K. Chip Multiprocessor Architecture – Techniques to Improve Throughput and Latency. Morgan and Claypool Publishers, San Rafael, 2007.
17. Raedt K.D., Michielsen K., De Raedt H., Trieu B., Arnold G., Marcus Richter, Th Lip-pert, Watanabe H., and Ito N. Massively parallel quantum computer simulator, *Computer Physics Communications*, 176, pp. 121-136.
18. Williams C.P. Explorations in Quantum Computing, *Texts in Computer Science*, Chapter 2. Quantum Gates. Springer, 2011, pp. 51-122.
19. Potapov V., Gushanskiy S., Guzik V., Polenov M. The Computational Structure of the Quantum Computer Simulator and Its Performance Evaluation, *In: Software Engineering Perspectives and Application in Intelligent Systems. Advances in Intelligent Systems and Computing*. Springer, 2019, Vol. 763, pp. 198-207.
20. Rahman M. and Kaykobad M. Complexities of some interesting problems on spanning trees, *Information Processing Letters*, 2005, 94 (2), pp. 93-97.
21. Bennett C.H., Shor P.W., Smolin J.A., Thapliyal A.V. Entanglement-assisted Capacity of a Quantum Channel and the Reverse Shannon Theorem, *IEEE Transactions on Information Theory*, 2002, 48, 2637.
22. Hector A. et al. Qiskit: An open-source framework for quantum computing, 2019.
23. Milner R.G. A Short History of Spin., *In: Contribution to the XV International Workshop on Polarized Sources, Targets, and Polarimetry. Charlottesville, Virginia, USA, September 9 – 13, 2013*. arXiv:1311.5016 (2013);
24. Boneh D., Zhandry M. Quantum-secure message authentication codes, *In: Proceedings of Eurocrypt*, 2013, pp. 592-608;
25. Potapov V., Gushansky S., Guzik V., Polenov M. Architecture and Software Implementation of a Quantum Computer Model, *In: Advances in Intelligent Systems and Computing*. Springer, 2016, Vol. 465, pp. 59-68.

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

Гушанский Сергей Михайлович – Южный федеральный университет; e-mail: smgushanskiy@sfn.edu.ru; г. Таганрог, Россия; тел.: 88634371656; кафедра вычислительной техники; к.т.н.; доцент.

Потапов Виктор Сергеевич – e-mail: vitya-potapov@rambler.ru; тел.: 88634371656; кафедра вычислительной техники; ассистент.

Божич Владимир Иванович – Таганрогский институт имени А.П. Чехова (филиал) ФГБОУ ВО "РГЭУ (РИНХ)", e-mail: vladimir.bozhich@gmail.com; г. Таганрог, Россия; тел.: 88634601812; кафедра биолого-географического образования и здоровьесберегающих дисциплин; д.т.н.; профессор.

Gushanskiy Sergey Mikhailovich – Southern Federal University; e-mail: smgushanskiy@sfnu.ru; Taganrog, Russia; phone: +78634371656; the department of computer engineering; cand. of eng. sc.; associate professor.

Potapov Victor Sergeevich – e-mail: vitya-potapov@rambler.ru; phone: +78634371656; the department of computer engineering; assistant.

Bozhich Vladimir Ivanovich – FSBEI HE “RSEU (RINH)”, Taganrog Institute A.P. Chekhov, e-mail: vladimir.bozhich@gmail.com; Taganrog, Russia; phone: +78634601812; the department of biological and geographical education and health preserving disciplines; dr. of eng. sc.; professor.

УДК 004.738.2

DOI 10.18522/2311-3103-2023-5-16-34

Ю.А. Заргарян, Е.В. Заргарян, В.И. Кошенский, К.О. Кирсанов
РАЗРАБОТКА СИСТЕМЫ ИЗБИРАТЕЛЬНОГО ПОДАВЛЕНИЯ
БЕСПРОВОДНЫХ СИГНАЛОВ

Использование устройств с беспроводной передачей данных в современном мире давно стало обыденностью. К подобным устройствам относятся смарт-часы, беспроводные наушники и гарнитуры. Такие устройства обеспечивают мобильность и удобство отсутствием проводов и высокой функциональностью. Одним из существенных недостатков использования беспроводных устройств является снижение эффективности восприятия окружающего мира человеком, в частности это касается слухового канала информации. Особенно остро такая проблема возникает при переходе железнодорожных путей пешеходами в наушниках. В данной работе предлагается решение такой проблемы, рассматривается система избирательного подавления беспроводной связи, а именно сигналов Bluetooth и Wi-Fi. Такая система на базе уникальных алгоритмов имеет возможность обнаружения приближения пешеходов к железнодорожным переходам с помощью видеокамер, а также на основе сканирования и анализа эфира окружающих беспроводных сетей может избирательно подавлять сигналы Bluetooth и Wi-Fi в зоне действия установки. Таким образом, система отключает беспроводные устройства пешеходов вблизи железнодорожных переходов, тем самым привлекает их внимание к запрещающему сигналу светофора, светомушовой сигнализации и приближающемуся поезду. Система работает на основе главного управляющего микроконтроллера, который с помощью реле контролирует подачу электропитания на включение подавителя только при обнаружении в радиусе досягаемости беспроводных сигналов, которые необходимо подавить, или при наличии пешеходов в поле зрения видеокамер. Кроме того, система подразумевает метод оценки паттернов поведения пешеходов при подавлении беспроводной связи на основе нейросетей по распознаванию образов с последующей интерпретацией полученных данных. Это позволит собрать статистическую информацию с целью анализа реакции людей на работу системы вблизи железнодорожных переходов.

Избирательное подавление; беспроводные сигналы; сигналы Bluetooth и Wi-Fi; уникальные алгоритмы; обнаружение пешеходов; сканирование эфира; повышение безопасности; железнодорожный переход; паттерны поведения.

Yu.A. Zargaryan, E.V. Zargaryan, V.I. Koshensky, K.O. Kirsanov
DEVELOPMENT OF A SYSTEM OF SELECTIVE SUPPRESSION
OF WIRELESS SIGNALS

The use of devices with wireless data transmission has long become commonplace in today's world. Such devices include smart watches, wireless headphones and headsets. Such devices provide mobility and convenience with the absence of wires and high functionality. One of the significant disadvantages of using wireless devices is a decrease in the efficiency of human perception of the surrounding world, in particular it concerns the auditory channel of information. Especially

acute such a problem arises when pedestrians crossing railroad tracks wearing headphones. In this paper a solution to such a problem is proposed, a system of selective suppression of wireless communication, namely Bluetooth and Wi-Fi signals is considered. Such a system based on unique algorithms has the ability to detect the approach of pedestrians to railroad crossings using video cameras, and based on scanning and analyzing the airwaves of surrounding wireless networks can selectively suppress Bluetooth and Wi-Fi signals in the area of the installation. Thus, the system disables pedestrians' wireless devices near railroad crossings, thereby drawing their attention to the prohibited traffic signal, light and noise signaling, and approaching train. The system is based on a master control microcontroller, which controls, via relays, the power supply to turn on the suppressor only when wireless signals to be suppressed are detected in range or when pedestrians are in the field of view of video cameras. In addition, the system involves a method for estimating patterns of pedestrian behavior during wireless suppression based on neural network-based pattern recognition, followed by interpretation of the resulting data. This will make it possible to collect statistical information in order to analyze the reaction of people to the operation of the system near railway crossings.

Suppressor; selective suppression; wireless signals; Bluetooth and Wi-Fi signals; unique algorithms; pedestrian detection; air scanning; safety enhancement; railroad crossing; behavior patterns.

Введение. В современном мире очень распространены устройства беспроводной связи: часы, наушники, гарнитуры – всё это придаёт не только удобство и мобильность современному человеку, разнообразие функциональных возможностей и контента по типу мобильной связи, музыки, аудиокниг, видео, но и несёт вместе с тем потенциальную опасность жизни и здоровью человека, поскольку оказывают негативное влияние на его бдительность и зачастую мешают оперативной оценке обстановки и окружения в условиях оживлённой городской среды. Значительно снижаются возможности зрительного и слухового каналов восприятия информации. Особенно остро это касается автодорог и железнодорожных переходов. В рамках данной статьи речь пойдёт о проблемных ситуациях и методах их разрешения применительно к инфраструктуре РЖД [1–3].

Обеспечение безопасности людей в различных сферах деятельности является актуальной задачей в настоящее время, которая требует эффективных и надёжных решений. Данный вопрос не обходит стороной и сферу железнодорожного транспорта. Железная дорога является зоной повышенной опасности, в частности это касается железнодорожных переходов, переездов и других потенциально опасных участков путей, где может оказаться человек. Данная проблема очень актуальна, так как ситуацию зачастую осложняют сами люди, которые не соблюдают минимальные правила безопасности и дорожного движения, теряют бдительность и подвергают себя риску. Одним из ярких примеров несоблюдения минимальных правил безопасности является использование беспроводных аудиоустройств и иной носимой электроники вблизи железнодорожных переходов и переездов, что не позволяет вовремя заметить приближающийся поезд, запрещающие сигналы светофора и светового сигнала, что часто приводит к трагедиям [4–6].

Статистика несчастных случаев. По информации ОАО «РЖД» только за последние 5 лет в период 2017–2022 гг. ситуация с травматизмом и несчастными случаями развивается в худшую сторону, наблюдается стремительное увеличение динамики производственного травматизма на железнодорожных переходах, что говорит о негативном развитии уровня безопасности пешеходов. Только по Северо-Кавказской железной дороге в данный период травмировано более 500 человек, в том числе 48 на пешеходных железнодорожных переходах. При этом, по судебным искам ОАО «РЖД» было выплачено компенсаций на сумму свыше 72 млн. руб. Согласно исследованиям 20% смертельных случаев приходится на пешеходов в беспроводных наушниках, которые не слышат приближающегося поезда и пре-

дупреждающих сигналов световых сигнализаций. Для исключения смертельных исходов в этих случаях требуется оснащение нерегулируемых переходов системами избирательного подавления беспроводной связи. В марте 2023 был получен прямой запрос от представителей филиала ОАО «РЖД» Северо-Кавказской железной дороги на проведение исследований, направленных на возможность избирательного подавления беспроводных сигналов для инфраструктуры РЖД. Данная разработка будет иметь социальную значимость и направлена на создание инновационной системы подавления беспроводных сигналов для носимой электроники с целью повышения бдительности пешеходов при переходе через железнодорожные пути.

Анализ существующих решений и патентный поиск. В рамках данной разрабатываемой системы на первом этапе работы, согласно поставленным целям, были проведены патентный поиск и исследования существующих решений. Прямое решение проблемы непроизводственного травматизма на железнодорожных переходах в открытых источниках отсутствует. Патентный поиск выполнялся по отечественным и зарубежным базам. Было проанализировано более 1500 патентов, по результатам патентного поиска было отобрано 90 и найдено 5 ближайших аналогов [7]:

1. US20170117974A1 Signal Separation And Suppression System And Method (2015, USA).
2. CN102684812B Multifunctional signal suppression system (2012, China).
3. US20110223851A1 Multi-band jammer including airborne systems (2010, USA).
4. RU2669775C1 Способ подавления канала передачи непрерывной информации путём воздействия на него организованной импульсной помехой (2017, Россия).
5. RU182622U1 Адаптивный подавитель пассивных помех (2018, Россия).

По результатам анализа патентного поиска можно сделать вывод, что разрабатываемая система не имеет аналогов на внутреннем рынке современных технологичных решений Российской Федерации и будет обладать следующими существенными отличительными признаками:

1. В состав системы будет включён управляющий контроллер, позволяющий осуществлять сканирование окружающего пространства на предмет выявления внешних сигналов беспроводной связи.
2. В состав системы будут включены видеокамеры для отслеживания и анализа паттернов поведения пешеходов при подавлении сигналов беспроводных устройств.

Указанные существенные отличительные признаки обеспечат повышенную надёжность системы и возможность сбора статистической информации, а также значительно увеличат срок службы подавителя благодаря его периодической работе [8].

Статистическая информация о паттернах поведения пешеходов в зоне действия подавителя позволит найти пути дальнейшего совершенствования системы, а также позволит разработать новые направления по предупреждению смертельных случаев на нерегулируемых железнодорожных переходах.

Разработка системы избирательного подавления. Подавитель беспроводных сигналов представляет собой аппарат, препятствующий передаче информации в беспроводной персональной сети между двумя сопряжёнными устройствами. Изначально подобное оборудование предназначалось с целью подавления радиосвязи противника и защиты от средств прослушивания или слежения на военных или стратегических объектах. В настоящее время подаватели беспроводной связи используются [9]:

1. В компаниях для сохранения конфиденциальной информации и борьбы с промышленным шпионажем.
2. В самолётах, лечебных заведениях и исследовательских лабораториях для бесперебойной работы чувствительного оборудования.
3. В метро для предотвращения террористических актов.
4. В исправительных учреждениях для предупреждения передачи информации криминальными элементами своим подельникам.
5. В учебных заведениях при проверке знаний, вовремя сдаче экзаменов.
6. В местах, где использование мобильной радиосвязи будет мешать проведению мероприятия.

Принцип работы подавителя беспроводной связи напрямую зависит от технологии беспроводной связи, на подавление которой он рассчитан. Например, подавитель Bluetooth-сигналов основан на базовых понятиях технологии беспроводной передачи данных Bluetooth, которая осуществляется с помощью дециметровых радиоволн в диапазоне от 2402 МГц до 2480 МГц между двумя или несколькими устройствами посредством установленных в них специальных модулей или чипов. При этом возможно создание пикосети, в которой может существовать лишь одно главное устройство и до семи вспомогательных. Для обособленной работы Bluetooth-пары, её устойчивости к помехам и защите передаваемых данных в технологии применяется алгоритм Frequency Hopping Spread Spectrum, другими словами, скачкообразная перестройка частоты в спектре распространения сигнала. Суть данного алгоритма в том, что частота сигнала Bluetooth меняется скачкообразно 1600 раз в секунду, поэтому, чтобы помешать подобной передаче данных, существуют два способа:

1. Воздействовать на Bluetooth-сигнал DDoS-атакой.
2. Глушители и подавители Bluetooth-сигнала.

Если говорить о первом способе, то речь идёт о разрыве сопряжения между двумя связанными в одной сети устройствами при помощи специального программного обеспечения. Одним из таких ПО является Kali Linux – это операционная система (дистрибутив), основанная на ядре Linux и изначально предназначенная её создателями (Matí Aharóni и Max Moser) для проведения тестирования на безопасность от проникновения [10–12].

Воздействие с помощью Kali Linux предполагает влияние на устройство Bluetooth в виде DDoS-атаки с последующим внедрением в сигнал и само устройство, что находится за рамками юридического законодательного поля Российской Федерации. Кроме того, самые распространённые носимые устройства Bluetooth и гарнитуры работают на данной технологии от версии 4.2 до версии 5.0, в которых есть защита от подобного воздействия ещё с 2010 года.

Если рассматривать второй подход воздействия на беспроводной сигнал путём глушителей и подавителей Bluetooth-сигнала, важно разобраться в принципе работы подобных устройств и сигнала, на который они влияют. Обычно полезный сигнал представляет собой электромагнитные волны низкой частоты, которые достаточно трудно передать. Для решения этой проблемы используется модуляция. Модуляция представляет собой преобразование одной либо нескольких характеристик несущего высокочастотного электромагнитного колебания в соответствии с параметрами низкочастотного модулирующего сигнала. Пример амплитудной модуляции приведён на рис. 1.

На рис. 1 на первом графике изображён пример длинноволнового сигнала, который необходимо передать. На втором графике приведён несущий сигнал высокой частоты, который и позволит осуществить передачу информационного сигнала. На третьем графике изображён промоделированный сигнал. Амплитуда высокочастотного ко-

лебания изменяется по закону полезного сигнала. Для передачи полезного сигнала несущая волна может изменять амплитуду, частоту и фазу. Из этого следует, что существует три вида модуляции: амплитудная, частотная и фазовая. В результате такого преобразования длинноволновый полезный сигнал переносит свой спектр в область достаточно высоких частот, чтобы их было возможно передать без использования чрезмерно больших антенн. Диапазоны несущих частот стандартизированы. Каждый из таких диапазонов имеет свою сферу применения. Наиболее распространёнными стандартами несущих частот являются: FM – 88-108 МГц, GSM900 – 925-960 МГц, GSM1800 – 1805-1880 МГц, GPS L1 – 1570-1620 МГц, GPS и ГЛОНАСС L2 – 1200-1310 МГц, GPS и ГЛОНАСС L3 – 1380-1410 МГц, 3G – 2110-2170 МГц, 4G WIMAX – 2570-2690 МГц, 4G LTE – 791-820 МГц, Wi-Fi и Bluetooth – 2400-2500 МГц. Количество устройств, использующих технологию радиосвязи, достаточно велико.

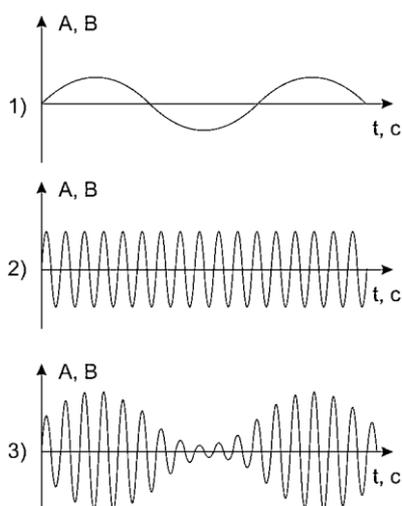


Рис. 1. Амплитудная модуляция сигнала

Довольно распространены устройства, управление которыми осуществляется при помощи беспроводной связи, например, беспилотные летательные аппараты или электронные замки. Помимо них имеют широкое распространение устройства, обеспечивающие коммуникацию между людьми при помощи беспроводной связи, например, мобильные телефоны. В ряде случаев, например, на совещании или в библиотеке, применение устройств, вроде сотового телефона, может доставить определённый дискомфорт окружающим людям или же с помощью его можно осуществить передачу конфиденциальной информации. Иногда с помощью беспроводных устройств радиосвязи можно осуществлять деятельность, угрожающую жизни людей. Генераторы помех в сетях беспроводных связей, которые при включении делают невозможным использование конкретных или любых беспроводных устройств, позволяют решить такие проблемы. Чтобы нарушить работу беспроводного устройства, необходимо создать в области действия этого устройства помеху, не дающую возможность принять необходимый для его работы сигнал. Сигнал помехи изменяет полезный сигнал так, что принимающее устройство перестаёт его распознавать. Пример влияния помех на полезный сигнал представлен на рис. 2-5.

На рис. 2 изображён пример полезного сигнала. В данном случае у несущего сигнала изменяется амплитуда по закону информационного сигнала.

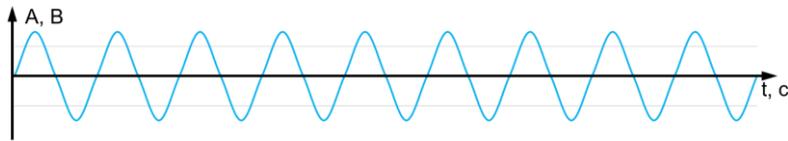


Рис. 2. Полезный сигнал

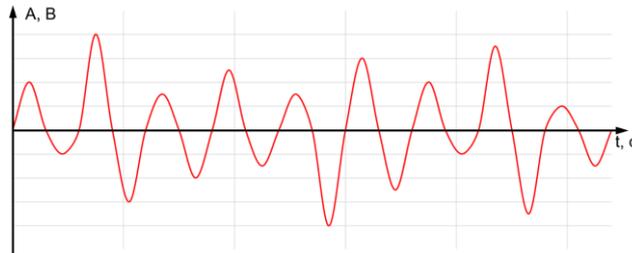


Рис. 3. Сигнал аналогичной частоты со стохастически изменяющейся амплитудой

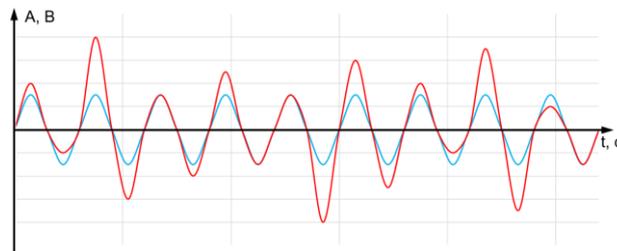


Рис. 4. Наложение двух сигналов друг на друга

Для его искажения необходимо излучать в пространство сигнал аналогичной частоты со стохастически изменяющейся амплитудой. Такой сигнал изображён на рис. 3. В результате два этих сигнала накладываются друг на друга, как показано на рис. 4, и получается искажённый сигнал, приведённый на рис. 5. Также из рис. 5 видно, что сигнал помехи должен иметь большую мощность для перекрытия полезного сигнала, иначе полезный сигнал сам перекроет помеху и сведёт на нет её воздействие [13–15].

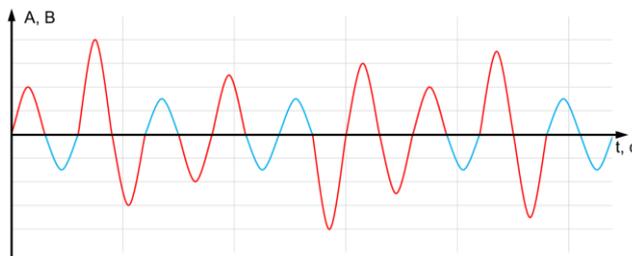


Рис. 5. Искажённый сигнал

Такой принцип работы реализован в средствах подавления радиосигналов вне зависимости от их конструкции, назначения и цены. Устройства создания по-

мех различаются, в первую очередь, по мощности излучаемых помех и по количеству подавляемых диапазонов частот. Также подавители радиосигналов делятся на мобильные устройства и стационарные станции. Стационарные бывают как мало функциональными устройствами, с радиусом действия в несколько десятков метров, подавляя при работе 6-10 диапазонов частот, так и более мощными приборами с обширным количеством своих возможностей.

Очень важно было на ранних этапах работы изучить возможные риски воздействия подавителей беспроводной связи на устройства жизнеобеспечения и угрозу создания помех для них. Проанализировав принципы работы различных видов кардиостимуляторов, слуховых аппаратов и подавителей беспроводных сигналов, было установлено, что на работу устройств поддержания жизни влияют электромагнитные поля. Важно отметить, что беспроводные сети (Bluetooth, Wi-Fi, Lora WAN, GSM и другие) не относятся к данной категории и не оказывают никакого влияния на работу кардиостимуляторов, слуховых аппаратов. Поскольку подавители беспроводных сигналов, работают по принципу создания помех на тех же частотах, что и сами беспроводные сети, и соответственно также не способны оказывать влияния на работу устройств жизнеобеспечения.

Проработка решений. В 2023 году проблеме травматизма и несчастных случаев из-за использования наушников вблизи железнодорожных переходов и поездов ОАО «РЖД» были посвящены множество тематических конкурсов, конференций и форумов, из которых стоит отметить «День инноваций ОАО «РЖД» Северо-Кавказской железной дороги, конкурс молодёжных проектов ОАО «РЖД» «Новое звено», образовательная конференция для молодёжи Северо-Кавказской железной дороги «XVI Слёт молодёжи С-КАВ». На данных мероприятиях авторами статьи была представлена концепция системы избирательного подавления беспроводных сигналов для обеспечения безопасности пешеходов на железнодорожных переходах. Общий вид системы представлен на рис. 6.



Рис. 6. Общий вид системы избирательного подавления беспроводных сигналов

На основе обратной связи от экспертной группы ОАО «РЖД» были разработаны общая структурная схема, отражающая основные элементы и связи между ними, представленная на рис. 7, и функциональная схема системы, представленная на рис. 8.

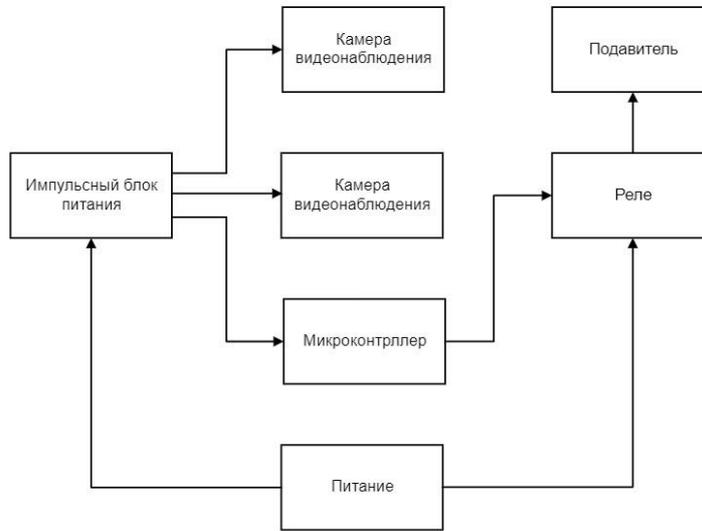


Рис. 7. Общая структурная схема системы

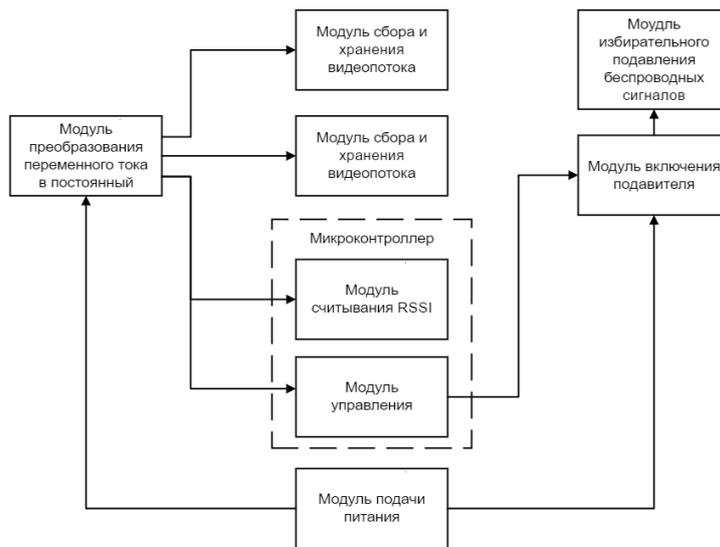


Рис. 8. Функциональная схема системы

При решении задач вышеуказанной проблемы предполагается разработка уникальных алгоритмов обнаружения приближения пешеходов к железнодорожным переходам, как показано на рис. 9. Это позволит реализовать работу системы подавления в заданные интервалы времени, с целью повышения её ресурса [16–17].

Также в системе предполагается использовать разработанный метод избирательного подавления беспроводной связи, основанный на анализе сигналов Bluetooth и Wi-Fi в зоне действия установки, как показано на рис. 10, что также нацелено на увеличение рабочего ресурса подавителя.

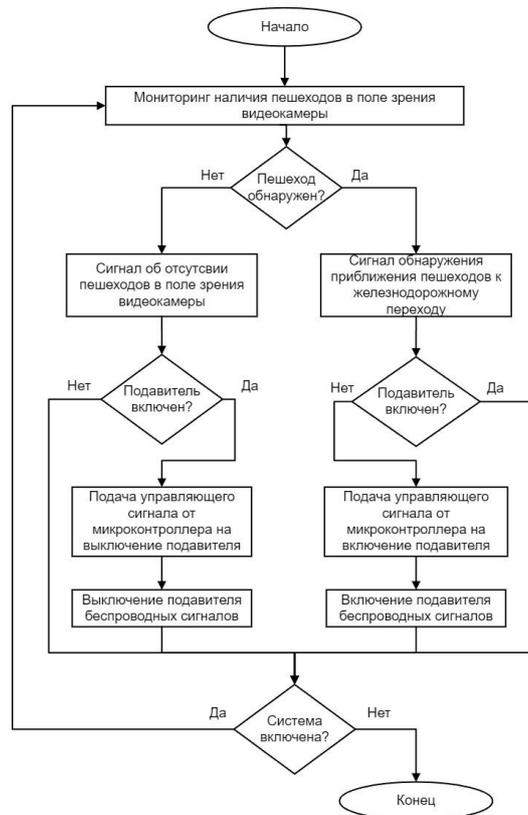


Рис. 9. Алгоритм обнаружения приближения пешеходов к железнодорожным переходам

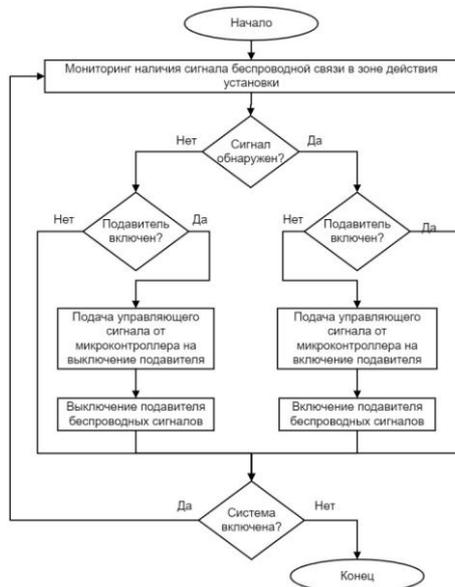


Рис. 10. Алгоритм метода избирательного подавления беспроводной связи

Дополнительно, планируется разработать метод оценки паттернов поведения пешеходов при подавлении беспроводной связи на основе нейросетей по распознаванию образов с последующей интерпретацией полученных данных. Это позволит набрать статистическую информацию поведения людей вблизи железнодорожного полотна при внезапном пропадании Bluetooth, Wi-Fi, чего не было обнаружено в открытых источниках. Алгоритм данного подхода представлен на рис. 11 [5]



Рис. 11. Алгоритм метода оценки паттернов поведения пешеходов при подавлении беспроводной связи

Моделирование. На основе общей концепции и вида системы избирательного подавления беспроводных сигналов, а также на базе структурной и функциональной схем было проведено 3D-моделирование для наглядной демонстрации внешнего вида будущей установки и компоновки всех комплектующих. Полученные 3D-модели приведены на рис. 12-13.

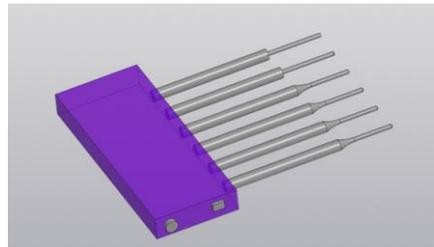


Рис. 12 3D-модель подавителя беспроводных сигналов

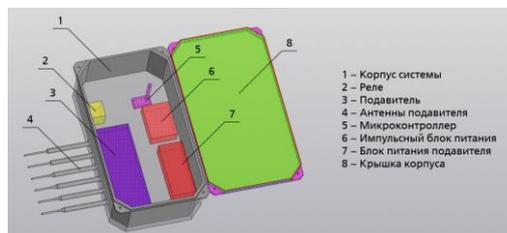


Рис. 13. 3D-модель прототипа корпуса системы избирательного подавления беспроводных сигналов

Кроме того, были составлены чертежи основных составляющих системы и некоторых компонентов с учётом доработки характеристик под требуемые параметры, как показано на рис. 14-15.

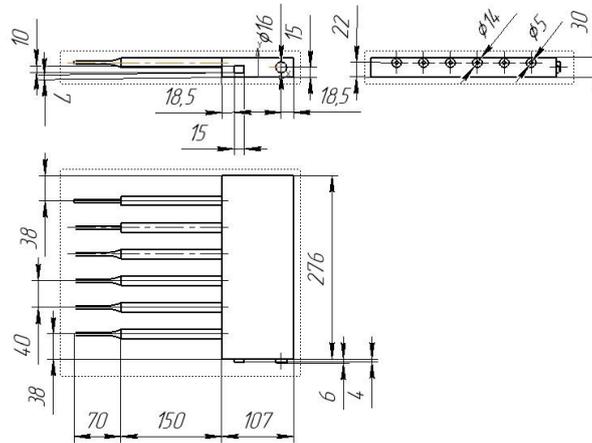


Рис. 14. Чертёж подавителя беспроводных сигналов

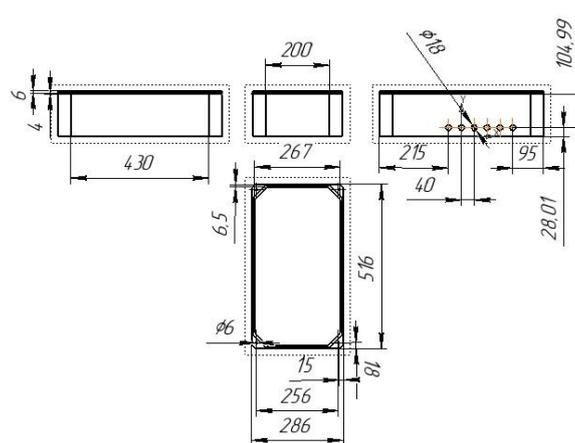


Рис. 15. Чертёж корпуса системы

На данном этапе также был разработан корпус системы из пластика. За основу был взят пластиковый щит с монтажной панелью ЩМПлл Solid-5 и глухой дверью. Подобный корпус используется для сборки разнообразных электрощитов: силовых, управления, автоматики. Одной из особенностей является возможность производить монтаж аппаратуры как модульного, так и обычного исполнения.

Проведение экспериментов. Разрабатываемый продукт находится на стадии активного тестирования базовых элементных частей.

В первую очередь были проведены эксперименты подавителя беспроводных сетей с целью проверки его работоспособности и соответствия паспортным характеристикам. Задачи, которые были поставлены во время испытаний:

1. Определить дальность работы подавителя.
2. Определить частоты подавления и подавляемые протоколы.

3. Определить время с момента включения подавителя до рабочего состояния.
4. Определить количество одновременно подавляемых устройств.

В соответствии с планом экспериментов были проведены испытания и оценены функциональные возможности радиочастотного подавителя. Выявлены отклонения в характеристиках эффективного радиуса подавления, который оказался в 1.5–2 раза меньше от заявленных значений в 40 метров и составляет около 20–25 метров, как показано на рис. 16–17.

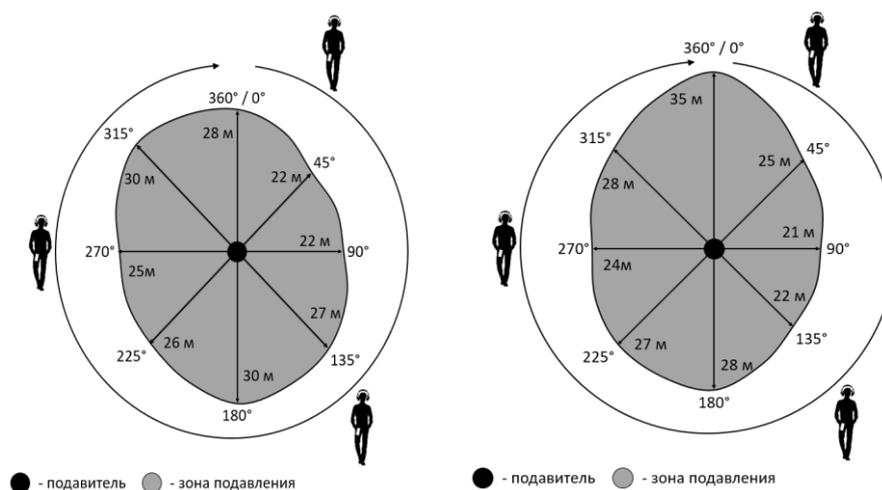


Рис. 16. Результаты экспериментов



Рис. 17. Схема зоны действия подавителя

Полученные в ходе испытаний результаты соответствуют требуемым параметрам системы, как показано в табл. 1. Также время с момента включения подавителя до рабочего состояния, количество одновременно подавляемых устройств и возможность изменять диаграмму направленности подавления удовлетворяют поставленным задачам. Устройство можно использовать для дальнейшей реализации системы [6–8].

Таблица 1

Результаты экспериментов подавителя в соответствии с требуемыми параметрами

п/п	Наименование параметра	Требуемые параметры	Полученное значение во время эксперимента	Примечание
1	Радиус подавления	не менее 20 метров	не менее 20 метров	–
2	Время с момента включения подавителя до рабочего состояния	не более 1 секунды	не более 1 секунды	–
3	Радиус подавления с изменением диаграммы направленности при использовании фольги	Есть возможность изменять диаграмму направленности подавления	17 метров	Во время данного эксперимента использовалась пищевая фольга, сложенная в 9 слоёв
4	Радиус подавления с изменением диаграммы направленности при использовании металла	Есть возможность изменять диаграмму направленности подавления	25 метров	Во время данного эксперимента использовался лист оцинкованного металла толщиной 1.5 мм
5	Количество одновременно подавляемых устройств	не менее 2 устройств	не менее 2 устройств	Во время эксперимента в поле действия подавителя находилось 2 устройства Bluetooth, на которые было успешно оказано влияние

Кроме того, были проведены эксперименты видеокамер с целью проверки их работоспособности и соответствия паспортным характеристикам. Задачи, которые были поставлены во время испытаний:

1. Определить дальность видимости камер и качество видео.
2. Определить угол обзора камер.
3. Проверить возможность камер записывать видео на съёмный носитель карту памяти MicroSD.
4. Проверить возможность камер транслировать видео в прямом эфире и записывать его в «облако».

В соответствии с планом экспериментов были проведены испытания и оценены функциональные возможности видеокамер. Радиус визуального контакта в видеопотоке с возможностью определения реакции человека на работу системы составляет не более 20 метров, как показано на рис. 18, радиус определения движения камерой в её поле зрения составляет не более 10 метров, как показано на рис. 19, угол обзора по горизонтали составляет не более 120 градусов, угол обзора по вертикали составляет не более 55 градусов, есть возможность записи видео на съёмный носитель карту памяти MicroSD, а также возможность транслировать видео в прямом эфире и записывать видео в «облако», при этом время задержки видеопотока в прямом эфире составляет не более 1 секунды [10].

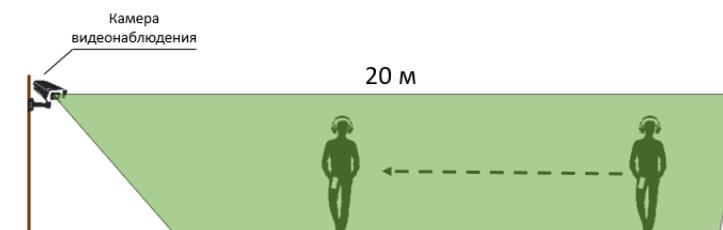


Рис. 18. Радиус визуального контакта в видеопотоке с возможностью определения реакции человека на работу системы

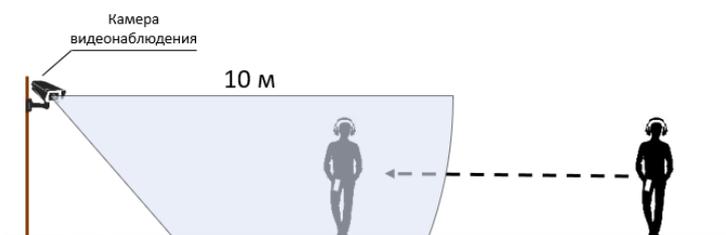


Рис. 19. Радиус определения движения человека камерой в её поле зрения

Данные результаты соответствуют требуемым параметрам системы и удовлетворяют поставленным задачам, как показано в табл. 2. Устройство можно использовать для дальнейшей реализации системы.

Таблица 2

Результаты экспериментов видеокамер в соответствии с требуемыми параметрами

п/п	Наименование параметра	Требуемые параметры	Полученное значение во время эксперимента	Примечание
1	Радиус визуального контакта в видеопотоке с возможностью определения реакции человека на работу системы	20 метров	не более 20 метров	—

Окончание табл. 2

2	Радиус определения движения камерой в её поле зрения	10 метров	не более 10 метров	–
3	Угол обзора по горизонтали	110 градусов	не более 120 градусов	–
4	Угол обзора по вертикали	50 градусов	не более 55 градусов	–
5	Возможность записи видео на съёмный носитель карту памяти MicroSD	Есть	Есть	–
6	Возможность транслировать видео в прямом эфире	Есть	Есть	–
7	Время задержки видеопотока в прямом эфире	не более 1 секунды	не более 1 секунды	–
8	Возможность записи видео в «облако»	Есть	Есть	Камеры обладают данным функционалом и имеют все необходимые настройки

Электропитание. При разработке системы на ранних этапах предполагалось использовать аккумуляторные батареи в качестве электропитания устройства.

В нашем случае устройства, которые необходимо обеспечить электропитанием от аккумулятора имеют параметры, указанные ниже.

Подавитель беспроводных сигналов:

1. Потребляемая мощность устройства: 60 Вт*ч.
2. Необходимое напряжение: 5 В.
3. Необходимая сила тока: 12А.

Из этих данных становится ясно, что для подавителя требуется 12000 мА*ч, так как для начальной проверки работоспособности системы необходимо минимум три часа автономной работы, то ёмкость аккумулятора должна составлять 36000 мА*ч.

Видеокамера:

1. Потребляемая мощность устройства: 12 Вт*ч.
2. Необходимое напряжение: 12 В.
3. Необходимая сила тока: 1А.

Проведя необходимые расчёты, можно сделать вывод, что для электропитания двух видеокамер, требуется 6000 мА*ч на три часа работы.

Таким образом, общее потребление электроэнергии составит 42000 мА*ч, с учётом саморазряда и прочих потерь энергии ёмкость аккумулятора должна составлять не менее 50000 мАч. Подобные устройства являются дорогостоящими и не могут обеспечить необходимыми параметрами зависимые устройства, поэтому было решено перейти на стационарный источник питания системы [11–21].

Вывод. Разрабатываемая система реализуется по запросу филиала ОАО «РЖД» Северо-Кавказской железной дороги, имеет социальную значимость и направлен на создание инновационной системы подавления сигналов беспроводной связи для носимой электроники на железнодорожных переходах, с целью снижения травматизма и несчастных случаев и существенного усиления существующей системы безопасности, не имеющей аналогов на отечественном рынке.

Перспективность представленного решения была высоко оценена: начальником Северо-Кавказского центра инновационного развития РЖД Концевым Д.А., главным инженером железной дороги Черномазовым А.В..

Важно отметить, что при активном взаимодействии с Северо-Кавказским центром инновационного развития филиала ОАО «РЖД» Северо-Кавказской железной дороги после представления полученных результатов было принято решение о выделении площадки для проведения полевых экспериментов в реальных условиях, в виде железнодорожного пешеходного перехода в г. Ростове-на-Дону. По результатам работы авторами статьи планируется подготовка заявок на РИД.

Потребителем разработки является ОАО «РЖД» для исключения несчастных случаев на нерегулируемых переходах из-за использования пешеходами беспроводных гарнитур, однако функционал системы не ограничен сферой железных дорог. Также разработка может быть использована при подавлении беспроводных сигналов на нерегулируемых пешеходных переходах автодорог с активным трафиком, где риск ДТП с участием пешеходов остаётся высоким из-за использования ими беспроводных гарнитур. Кроме этого, предлагаемая разработка может быть применена различными учебными заведениями для подавления беспроводной связи при проведении государственных испытаний и экзаменов, что особенно актуально из-за широкого распространения миниатюрных беспроводных гарнитур.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Губкин В.С. Заргарян Ю.А. Технологии беспроводной передачи данных в современном мире // Проблемы автоматизации. Региональное управление. Связь и акустика: Сб. трудов X Всероссийской научной конференции и молодежного научного форума в рамках мероприятий, посвященных году Науки и технологий в Российской Федерации. – Ростов-на-Дону, 2021. – С. 188-191.
2. РЖД. – URL: <https://www.rzd.ru/>.
3. Кошенский В.И., Заргарян Ю.А. Особенности функционирования BLE в микроконтроллере ESP32 // Сб. трудов международной молодёжной школы С 23 «Инженерия – XXI» (г. Новороссийск, 21-22 апреля 2022 г.) / под общ. ред. к.ф.н. доцента И.В. Чистякова. – Новороссийск: Изд-во НФ БГТУ им. В.Г. Шухова, 2022. – С. 65.
4. Zargaryan E.V., Zargaryan Y.A., Dmitrieva I.A., Sakharova O.N. and Pushnina I.V. Modeling design information systems with many criteria. Information Technologies and Engineering – APITECH - 2020 // Journal of Physics: Conference Series. – 2020. – Vol. 2085 (3). – P. 032057(1-7). – DOI: 10.1088/1742-6596/1679/3/032057.
5. Google Патенты. – URL: <https://patents.google.com/>.
6. Финаев В.И., Заргарян Ю.А., Заргарян Е.В., Соловьев В.В. Формализация групп подвижных объектов в условиях неопределённости для выбора управляющих решений // Информатизация и связь. – 2016. – № 3. – С. 56-62.
7. Безопасность в беспроводной сети Bluetooth. – URL: <https://habr.com/ru/articles/595821/>.
8. Аламир Х.С., Заргарян Е.В., Заргарян Ю.А. Модель прогнозирования транспортного потока на основе нейронных сетей для предсказания трафика на дорогах // Известия ЮФУ. Технические науки. – 2021. – № 6 (223). – С. 124-132.
9. ПАТЕНТНЫЙ ПОИСК. – URL: <https://new.fips.ru/about/vptb-otdelenie-vserossiyskaya-patentno-tehnicheskaya-biblioteka/patentnyy-poisk.php>.
10. Яндекс.Патенты. – URL: <https://yandex.ru/patents>.
11. Технологии позиционирования. – URL: <https://merusoft.ru/tehnologii-pozicionirovaniya-vnutri-pomeshhenij-rtls/>.

12. СРЕДСТВА СОЗДАНИЯ ПОМЕХ В СЕТЯХ БЕСПРОВОДНОЙ СВЯЗИ. – URL: <https://www.elibrary.ru/item.asp?id=43072956>.
13. Глушилка Bluetooth колонок. – URL: <https://spy-soft.net/bluetooth-jammer/>.
14. Заргарян Ю.А., Кошенский В.И., Кирсанов К.О., Пресняков М.С. Применение технологии Bluetooth Low Energy для контроля перемещения людей в помещениях // Известия ЮФУ. Технические науки. – 2022. – № 3 (227). – С. 103-119.
15. Все о Bluetooth. – URL: <https://radio-secure.ru/technology/bluetooth>.
16. Аламир Хайдер Сагбан Хуссейн, Заргарян Е.В., Заргарян Ю.А. Интеллектуальная система контроля заторов на дорогах с использованием контролируемого алгоритма машинного обучения на базе адаптивного ЮТН // Известия ЮФУ. Технические науки. – 2023. – № 2 (232). – С. 175-186. – https://izv-tn.tti.sfedu.ru/index.php/izv_tn/article/view/784/972.
17. Ауси Рим Мохаммед Худхейр, Заргарян Е.В., Заргарян Ю.А. Глубокое обучение методам защиты от атак // Известия ЮФУ. Технические науки. – 2023. – № 2 (232). – С. 227-239. – https://izv-tn.tti.sfedu.ru/index.php/izv_tn/article/view/788/976.
18. Заргарян Е.В., Ганцевский А.В. Анализ существующих нейронных сетей и их применение в настоящее время // Донецкие чтения 2022: образование, наука, инновации, культура и вызовы современности: Матер. VII Международной научной конференции, посвящённой 85-летию Донецкого национального университета (Донецк, 27–28 октября 2022 г.). Т. 1: Механикоматематические, компьютерные науки, управление / под общ. ред. проф. С.В. Беспаловой. – Донецк: Изд-во ДонНУ, 2022. – С 159-161.
19. АРДУИНО РОСТОВ. – URL: <https://xn---7sbhgu4ahbanfnng.xn--p1ai/>.
20. Идентификаторы // PERCo. – URL: <https://www.perco.ru/products/identifikatory/>.
21. IBEACONRUSSIA. – URL: <https://www.ibeaconrussia.ru/gateway-beacon.html>.

REFERENCES

1. Gubkin V.S. Zargaryan Yu.A. Tekhnologii besprovodnoy peredachi dannykh v sovremennom mire [Technologies of wireless data transmission in the modern world], *Problemy avtomatizatsii. Regional'noe upravlenie. Svyaz' i akustika: Sb. trudov X Vserossiyskoy nauchnoy konferentsii i molodezhnogo nauchnogo foruma v ramkakh meropriyatiy, posvyashchennykh godu Nauki i tekhnologii v Rossiyskoy Federatsii* [Problems of automation. Regional management. Communication and acoustics. proceedings of the X All-Russian scientific conference and youth scientific forum in the framework of events dedicated to the year of Science and Technology in the Russian Federation]. Rostov-on-Don, 2021, pp. 188-191.
2. RZhd [RZD]. Available at: <https://www.rzd.ru/>.
3. Koshenskiy V.I., Zargaryan Yu.A. Osobennosti funktsionirovaniya BLE v mikrokontrollere ESP32 [Features of BLE functioning in ESP32 microcontroller], *Sb. trudov mezhdunarodnoy molodezhnoy shkoly C 23 «Inzheneriya – XXI» (g. Novorossiysk, 21-22 aprelya 2022 g.)* [Proceedings of the International Youth School C 23 "Engineering - XXI" (Novorossiysk, April 21-22, 2022)], ed. by cand. of phys. sc., associate professor I.V. Chistyakov. Novorossiysk: Izd-vo NF BGTU im. V.G. Shukhova, 2022, pp. 65.
4. Zargaryan E.V., Zargaryan Y.A., Dmitrieva I.A., Sakharova O.N. and Pushnina I.V. Modeling design information systems with many criteria. Information Technologies and Engineering – APITECH – 2020, *Journal of Physics: Conference Series*, 2020, Vol. 2085 (3), pp. 032057(1-7). DOI: 10.1088/1742-6596/1679/3/032057.
5. Google Patenty [Google Patents]. Available at: <https://patents.google.com/>.
6. Finaev V.I., Zargaryan Yu.A., Zargaryan E.V., Solov'ev V.V. Formalizatsiya grupp podvizhnykh ob"ektov v usloviyakh neopredelennosti dlya vybora upravlyayushchikh resheniy [Formalization of groups of mobile objects in conditions of uncertainty for the choice of control decisions], *Informatizatsiya i svyaz' [Informatization and communication]*, 2016, No. 3, pp. 56-62.
7. Bezopasnost' v besprovodnoy seti Bluetooth [Security in Bluetooth wireless network]. Available at: <https://habr.com/ru/articles/595821/>.
8. Alimir Kh.S., Zargaryan E.V., Zargaryan Yu.A. Model' prognozirovaniya transportnogo potoka na osnove neyronnykh setey dlya predskazaniya trafika na dorogakh [Neural network based traffic flow prediction model for predicting traffic on roads], *Izvestiya YuFU. Tekhnicheskie nauki [Izvestiya SFedU. Engineering Sciences]*, 2021, No. 6 (223), pp. 124-132.

9. PATENTNYY POISK [PATENT SEARCH]. Available at: <https://new.fips.ru/about/vptb-otdelenie-vserossiyskaya-patentno-tehnicheskaya-biblioteka/patentnyy-poisk.php>.
10. Yandex.Patenty [Yandex.Patents]. Available at: <https://yandex.ru/patents>.
11. Tekhnologii pozitsionirovaniya [Positioning technologies]. Available at: <https://merusoft.ru/tehnologii-pozicionirovaniya-vnutri-pomeshhenij-rtls/>.
12. SREDSTVA SOZDANIYA POMEKH V SETYAKH BESPROVODNOY SVYAZI [EMERGENCY CONSTRUCTION IN WIRELESS NETWORKS]. Available at: <https://www.elibrary.ru/item.asp?id=43072956>.
13. Glushilka Bluetooth kolonok [Bluetooth speaker jammer]. Available at: <https://spy-soft.net/bluetooth-jammer/>.
14. Zargaryan Yu.A., Koshenskiy V.I., Kirsanov K.O., Presnyakov M.S. Primenenie tekhnologii Bluetooth Low Energy dlya kontrolya peremeshcheniya lyudey v pomeshcheniyakh [Presnyakov.Application of Bluetooth Low Energy technology for control of people movement in premises], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2022, No. 3 (227), pp. 103-119.
15. Vse o Bluetooth [All about Bluetooth]. Available at: <https://radio-secure.ru/technology/bluetooth>.
16. Alamir Khayder Sagban Khusseyn, Zargaryan E.V., Zargaryan Yu.A. Intel'ktual'naya sistema kontrolya zatorov na dorogakh s ispol'zovaniem kontroliruemogo algoritma mashinnogo obucheniya na baze adaptivnogo IOTN [Intelligent traffic congestion monitoring system using a supervised machine learning algorithm based on adaptive IOTN], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2023, No. 2 (232), pp. 175-186. Available at: https://izv-tn.tti.sfedu.ru/index.php/izv_tn/article/view/784/972.
17. Ausi Rim Mokhammed Khudkheyr, Zargaryan E.V., Zargaryan Yu.A. Glubokoe obuchenie metodam zashchity ot atak [Deep training in attack defense techniques], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2023, No. 2 (232), pp. 227-239. Available at: https://izv-tn.tti.sfedu.ru/index.php/izv_tn/article/view/788/976.
18. Zargaryan E.V., Gantsievskiy A.V. Analiz sushchestvuyushchikh neyronnykh setey i ikh primeneniye v nastoyashchee vremya [Analysis of existing neural networks and their application at the present time], *Donetskie chteniya 2022: obrazovanie, nauka, innovatsii, kul'tura i vyzovy sovremennosti: Mater. VII Mezhdunarodnoy nauchnoy konferentsii, posvyashchennoy 85-letiyu Donetskogo natsional'nogo universiteta (Donetsk, 27–28 oktyabrya 2022 g.). T. 1: Mekhanikomatematicheskie, komp'yuternye nauki, upravlenie* [Donetsk Readings 2022: Education, Science, Innovation, Culture and Modern Challenges: Proceedings of the VII International Scientific Conference dedicated to the 85th anniversary of Donetsk National University (Donetsk, October 27-28, 2022). Vol. 1: Mechanicomathematics, Computer Science, Management], edited by prof. S.V. Bespalova. Donetsk: Izd-vo DonNU, 2022, pp 159-161.
19. ARDUINO ROSTOV [ARDUINO ROSTOV]. Available at: <https://xn---7sbhgu4ahbanfng.xn--p1ai/>.
20. Identifikatory [Identifiers], *PERCo*. Available at: <https://www.perco.ru/products/identifikatory/>.
21. IBEACONRUSSIA. Available at: <https://www.ibeaconrussia.ru/gateway-beacon.html>.

Статью рекомендовал к опубликованию д.т.н., профессор К.Е. Румянцев.

Заргарян Юрий Артурович – Южный федеральный университет; e-mail: yazargaryan@sfedu.ru; г. Таганрог, Россия; кафедра систем автоматического управления; к.т.н.; доцент.

Заргарян Елена Валерьевна – e-mail: evzargaryan@sfedu.ru; кафедра систем автоматического управления; к.т.н.; доцент.

Кошенский Вадим Игоревич – e-mail: koshenskii@sfedu.ru; кафедра систем автоматического управления; бакалавр 4 года обучения.

Кирсанов Кирилл Олегович – e-mail: kkirsanov@sfedu.ru; кафедра систем автоматического управления; бакалавр 4 года обучения.

Zargaryan Yuri Arturovich – Southern Federal University, e-mail: yazargaryan@sfedu.ru; Taganrog, Russia; the department of automatic control systems; cand. of eng. sc.; associate professor.

Zargaryan Elena Valerevna – e-mail: evzargaryan@sfnedu.ru; the department of automatic control systems; cand. of eng. sc.; associate professor.

Koshensky Vadim Igorevich – e-mail: koshenskii@sfnedu.ru; the department of automatic control systems; bachelor 4 years of study.

Kirsanov Kirill Olegovich – e-mail: kkirsanov@sfnedu.ru; the department of automatic control systems; bachelor 4 years of study.

УДК 007.52:629.3.05

DOI 10.18522/2311-3103-2023-5-34-48

В.А. Костюков, М.Ю. Медведев, В.Х. Пшихопов**АЛГОРИТМ ПЛАНИРОВАНИЯ ПУТИ В ДВУХМЕРНОЙ СРЕДЕ
С ПОЛИГОНАЛЬНЫМИ ПРЕПЯТСТВИЯМИ НА КЛАССЕ
КУСОЧНО-ЛОМАНЫХ ТРАЕКТОРИЙ***

Актуальной проблемой, возникающей при разработке алгоритмов автоматического планирования пути, является рост вычислительных затрат при увеличении сложности среды функционирования. Не лишены этого недостатка графовые методы планирования, в частности, метод диаграмм видимости. Он позволяет сформировать в качестве узлов графа вершины полигональных границ каждого препятствия, а ребрами графа являются все те отрезки, соединяющие эти вершины, которые не имеют пересечений с препятствиями. При увеличении количества препятствий возрастает сложность такого графа, причем этот рост очень быстрый. Поэтому наиболее важной задачей становятся приемы сокращения сложности графа видимости. В данной статье предлагается гибридный алгоритм, строящийся на методе диаграмм видимости и семействе Vig-алгоритмов. Vig-алгоритм относится к классу локальных, поскольку каждый раз имеет дело с огибанием одного препятствия, появляющегося на пути следования робота, и этот алгоритм не может предсказать заранее, какое следующее препятствие придется обходить. Предлагаемый в данной статье метод планирования траектории движения сочетает графовый алгоритм с Vig-алгоритмами, что позволяет построить специальный граф с узлами в виде характерных точек препятствий. При этом Vig-алгоритм является шагом итерационного процесса оптимизации на графе, позволяющего за конечное число шагов прийти к оптимальному решению на классе кусочно-ломанных кривых. Предлагаемый метод решает задачу глобального поиска пути на классе кусочно-линейных траекторий с полигональными препятствиями; а в отличие от классических методов диаграмм прямой видимости, существенно снижает размерность графа за счет специального выбора ограниченного количества характерных точек соответствующих препятствий. В статье проводится разработка и теоретическое обоснование предлагаемого метода. Приводятся расчетные соотношения алгоритма, обосновывается оптимальность получаемой траектории. Аналитические соотношения подтверждаются результатами численного моделирования в различных средах, заполненных полигональными препятствиями. При этом эффективность предлагаемых алгоритмов подтверждается на примерах среды, заполненной препятствиями до 70–80%. Показано, что для прокладывания пути на сценах лабиринтного типа с одним распространенным видом препятствий рассматриваемый алгоритм на 10% превосходит оптимальный алгоритм Дейкстры.

Планирование пути; сложные полигональные препятствия; поиск на графах; Vig-алгоритм; гибридный алгоритм планирования.

* Исследование выполнено за счет гранта Российского научного фонда № 22-29-00337 «Разработка вероятностных методов оптимизации траекторий группы подвижных робототехнических комплексов в априори неопределенной среде», <https://rscf.ru/project/22-29-00337/> на базе АО «НКБ Робототехники и систем управления».

V.A. Kostyukov, M.Yu. Medvedev, V.Kh. Pshikhopov

AN ALGORITHM FOR PATH PLANNING IN A TWO-DIMENSIONAL ENVIRONMENT WITH POLYGONAL OBSTACLES ON A CLASS OF PIECEWISE POLYLINE TRAJECTORIES

An urgent problem of developing algorithms for automatic path planning is the increase in computational costs with an increase in the complexity of the operating environment. Graphic planning methods, in particular the method of visibility diagrams, are not without this drawback. As the number of obstacles increases, the complexity of a visibility graph increases, and this growth is very fast. Therefore, techniques for reducing the complexity of the visibility graph become the most important task. This article proposes a hybrid algorithm based on the method of visibility diagrams and Bug algorithms. The method of path planning proposed in this article combines a graph algorithm with Bug algorithms, which allows you to build a special graph with nodes in the form of characteristic obstacle points. In this case, the Bug algorithm is a step in the iterative optimization process on the graph, which allows for a finite number of steps to come to an optimal solution on a class of piecewise polyline curves. The proposed method solves the problem of global path search. Unlike the classical methods of line-of-sight diagrams, it significantly reduces the dimension of the graph due to a special selection of a limited number of characteristic points of the corresponding obstacles. The article develops and theoretically substantiates the proposed method. The calculated relations of the algorithm are given. The optimality of the obtained trajectory is justified. The analytical relations are confirmed by the results of numerical modeling in various environments filled with polygonal obstacles. It is shown that in order to pave the way for maze-type scenes with one common type of obstacles, the algorithm under consideration is 10% superior to the optimal Dijkstra algorithm.

Path planning; complex polygonal obstacles; graph search; Bug algorithm; hybrid planning algorithm.

Введение. Большинство существующих алгоритмов планирования пути сталкиваются с проблемой роста вычислительной сложности при увеличении сложности карты препятствий. В связи с этим актуальной является задача разработки алгоритма планирования пути, у которых сложность растет медленнее, чем у существующих алгоритмов.

Наиболее распространенные методы планирования пути используют алгоритмы поиска на графах. Каждый узел графа представляет одно состояние робота, которое может представлять собой положение, угол ориентации, скорость или ускорение робота. Переходы между состояниями отражаются ребрами графа и характеризуются функцией затрат. Это потенциально позволяет найти путь, который имеет минимальную общую стоимость достижения до целевого состояния. Для построения узлов графа может быть использован метод диаграмм видимости, который позволяет сформировать в качестве узлов вершины полигональных границ каждого препятствия, а ребрами графа являются отрезки, соединяющие эти вершины, которые не имеют пересечений с препятствиями.

При увеличении количества препятствий возрастает сложность такого графа, причем этот рост очень быстрый. Поэтому здесь наиболее важной задачей становятся приемы сокращения сложности графа видимости. Huang [1] предложил метод динамической диаграммы видимости, а Janet ввел концепцию т-вектора [2].

Другими подходами являются метод структурирования свободного пространства [3] и метод на основе диаграммы Вороного [4]. Здесь выделяются так называемые свободные звенья – отрезки, соединяющие вершины препятствий и не пересекающиеся с ними. Из свободных звеньев собираются выпуклые многогран-

ники, описывающие области свободного пространства. Эти данные оформляются в виде специального графа, а спланированный путь в обход препятствий строится алгоритмами поиска путей на графе. Использование диаграмм Вороного также приводит к построению допустимого пути, максимально удаленного от препятствий, при этом обеспечивает дополнительные возможности с точки зрения позиционирования робота. Методы структурирования свободного пространства и диаграммы Вороного дают более безопасный путь с точки зрения столкновения с препятствиями, но общая длина пути увеличивается.

К методам на основе случайного выбора относится метод быстро исследующих случайных деревьев (RRT) [5]. В этом методе процесс генерации пути состоит в построении дерева опорных точек, которое последовательно расширяется от начальной точки до целевой точки. Данный метод обладает невысокой сходимостью к оптимальному пути, поэтому направления его развития связаны с методами повышения его эффективности [6, 7].

Графовые методы поиска также используются в методах клеточной декомпозиции, предполагающих разбиение исходного пространства на дискретные ячейки [8]. Также получили распространение методы адаптивной пространственной декомпозиции [9, 10]. Пространственная декомпозиция позволяет применять методы дискретного поиска, базирующиеся на теории графов; наибольшую популярность получили алгоритмы семейств A^* и D^* [11, 12].

Алгоритм A^* является развитием алгоритма Дейкстры, в котором за счет выбора эвристической функции стоимости пути снижается вычислительная сложность. Вместе с тем, алгоритм требователен к памяти и его вычислительная сложность резко растет с ростом числа ячеек разбиения. Для снижения требовательности исходного алгоритма A^* к памяти и его вычислительной сложности разработаны различные модификации [13, 14]. Также заметим, что алгоритм Дейкстры позволяет прийти оптимальному решению с точностью до погрешности квантования, в отличие от методов A^* и D^* , чьи эвристики в некоторых случаях отклоняют поиск от оптимального решения.

Также существует семейство Bug-алгоритмов базирующихся на обходе препятствий, которые встретились на пути отрезка «прямой видимости», соединяющего начальную и конечную точки [15, 16]. При этом обход происходит по границам этих препятствий, до некоторой точки «отрыва», начиная с которой траектория уходит с указанной границы. Bug-алгоритм относится к классу локальных, поскольку каждый раз имеет дело с обходом одного препятствия, появляющегося на пути следования робота, и этот алгоритм не может предсказать заранее, какое следующее препятствие придется обходить [17].

Цель настоящей статьи заключается в разработке нового гибридного метода планирования траекторий, в котором bug-алгоритм является шагом итерационного процесса оптимизации на графе, что позволяет за конечное число шагов прийти к оптимальному решению на классе кусочно-ломаных кривых. Таким образом, предлагаемый аналитический метод планирования объединяет поиск на графах с итерационным Bug-алгоритмом, что позволяет построить специальный граф с узлами в виде характерных точек препятствий. Эти узлы соединяются кусочно-линейными отрезками. Затем решается задача поиска кратчайшего пути на построенном графе. Новизна алгоритма заключается в итерационной процедуре анализа соответствий между отрезком «прямой видимости», соединяющим текущую точку с целевой, и криволинейными полигональными границами препятствий, пересекаемых им.

В отличие от исходных Bug-алгоритмов, предлагаемый метод решает задачу глобального поиска пути на классе кусочно-линейных траекторий с полигональными препятствиями. В отличие от классических методов диаграмм прямой видимости, существенно снижает размерность графа за счет специального выбора ограниченного количества характерных точек соответствующих препятствий.

I. Постановка задачи. Рассматривается двумерная односвязная область $D \subset \mathbb{R}_2$, в которой заданы полигональные препятствия, множество которых обозначим $\{Obs\}$. Обозначим точки старта и финиша как А и В соответственно.

Под точечной границей $\partial(Obs)$ некоторого полигонального препятствия Obs , заданного как двумерное многообразие в \mathbb{R}_2 , или просто его границей, будем понимать упорядоченное множество вершин замкнутой кусочно-ломаной – реальной границы данного препятствия. Последнюю обозначим так же, что будем оговаривать всякий раз отдельно. Рассматриваются всевозможные кусочно-линейные траектории $\{Tr\}$, целиком принадлежащие области D. Каждая такая траектория Tr определяется как совокупность ее вершин, которые обозначаются в квадратных скобках $Tr=[Вершина1, Вершина2, ..]$. Тем же названием Tr обозначим саму кусочно-ломаную как одномерное многообразие. Каждый раз далее всякий раз оговаривается, какая именно интерпретация траектории имеется в виду. Если некоторая траектория Tr в качестве граничных имеет точки А и В, то назовем ее траекторией глобального пути из А в В. Множество таких траекторий обозначим $G_{AB}(Tr) \subset \{Tr\}$.

Ставится задача разработать и обосновать метод, решающий задачу глобального поиска пути на классе кусочно-линейных траекторий с полигональными препятствиями. Данный метод должен обеспечивать снижение размерности графа по сравнению с классическим методом диаграмм прямой видимости.

II. Алгоритм планирования. Предлагаемый алгоритм является итерационным. Далее приводится подробное описание указанного алгоритма.

Вначале итерационного алгоритма проводится отрезок прямой видимости $L_0=AB$, причем случай отсутствия пересечения $L_0 \cap Obs = \emptyset$ не рассматривается как тривиальный. Находится такое препятствие с номером j , что оно пересекается с прямой L_0 и $d_{min,j} = \min_i(d_{min,i})$, где $d_{min,i}$ – расстояние от начальной точки А до наиболее близкой к ней точки пересечения M_{j1} прямой L_0 и i -го препятствия. Обозначим такую точку, соответствующую j -му препятствию, как M_{j1} . Кроме того, выделяется точка M_{j2} пересечения препятствия Obs_j с прямой L_0 , наиболее удаленная от точки А из всех точек пересечения. Далее строятся два пути Tr_1 и Tr_2 : $Tr_1 = [A, M_{j1}, \partial_1(Obs_j)]$, $Tr_2 = [A, M_{j1}, \partial_2(Obs_j)]$. Точки А и M_{j2} являются граничными точками этих траекторий. Здесь $\partial_1(Obs_j)$ и $\partial_2(Obs_j)$ есть связные подмножества границы $\partial(Obs_j)$ с непрерывной нумерацией элементов и такие, что $\partial_1(Obs_j) \cup \partial_2(Obs_j) = \partial(Obs_j)$, $\partial_1(Obs_j) \cap \partial_2(Obs_j) = \{M_{j1}, M_{j2}\}$.

Далее пути Tr_1 и Tr_2 пути оптимизируются: для первого находится такое упорядоченное подмножество точек $Tr_1^* \subseteq Tr_1$, что для соответствующей кусочно-ломаной выполняются условия: а) она начинается с точки А; б) не пересекается ни с одним из препятствий сцены, отличным от Obs_j ; в) пересекая границу $\partial(Obs_j)$, не пересекает внутреннюю часть Obs_j ; г) имеет минимальную длину при условии, что последняя точка $Tr_1^*(end)$ множества Tr_1^* позволяет провести прямолинейный отрезок $Tr_1^*(end)B$, не имеющий пересечений с препятствием Obs_j . Аналогичным образом оптимизируется второй частичный путь Tr_2 с получением Tr_2^* . Заметим,

что оптимизированные частичные пути уже могут не включать точки пересечения M_{j1} или/и M_{j2} . При этом используется специальный метод прореживания кусочно-ломаной, рассмотренный ниже.

На рис. 1 приведен пример построения исходных Tr_1, Tr_2 и оптимизированных путей Tr_1^* и Tr_2^* для сцены из трех препятствий. Здесь $Tr_1=[A, M_{j1}, A_{j1}, A_{j2}, M_{j2}], Tr_2=[M_{j1}, A_{j4}, A_{j3}, M_{j2}], Tr_1^*=[A, A_{j1}, A_{j2}], Tr_2^*=[A, A_{j4}]$.

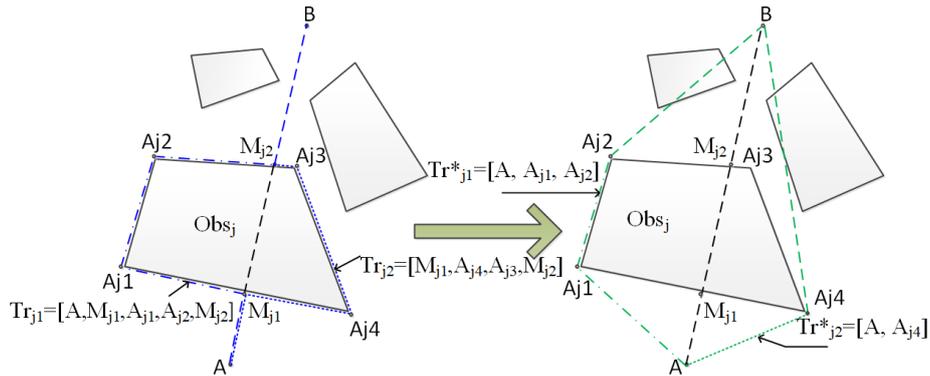


Рис. 1. Построение частичных путей на 1-й итерации

На рис. 2 показано построение частичных оптимизированных путей на двух итерациях алгоритма планирования.

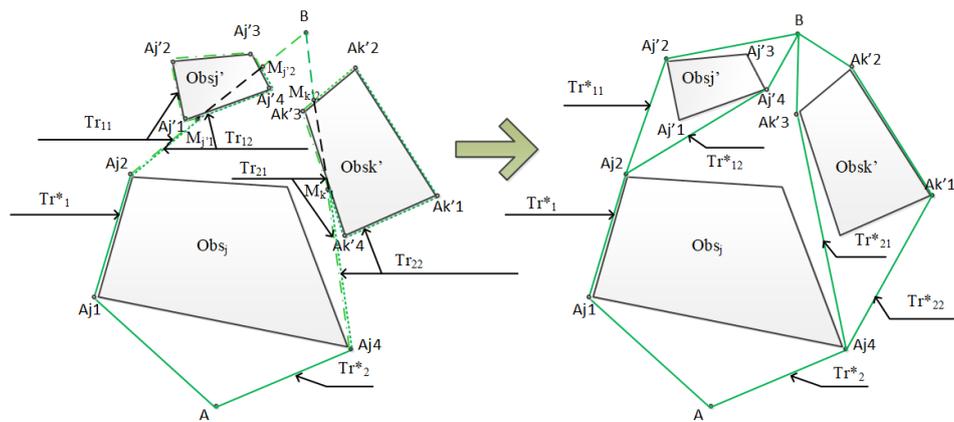


Рис. 2. Построение частичных путей на 2-й итерации

После первой итерации возникает ветвление траекторий, соответствующее двум найденным частичным путям Tr_1^* и Tr_2^* .

Введем понятие глобального пути, соответствующего конечной точке P некоторого частичного оптимизированного пути, достигнутого на u -й итерации. Оно обозначает объединение всех частичных оптимизированных путей до этой точки и отрезка PB . Каждый такой путь обозначим как $GTr_p^{(u)}$. Полную длину всех частичных оптимизированных путей до точки P обозначим как $L(P)$, тогда длина $GTr_p^{(u)}$ есть $|GTr_p^{(u)}| = L(P) + |PB|$.

Будем считать, что соответствующие глобальные пути после первой итерации $u = 1$ есть $GTr_1^{(1)} = [Tr_1^*, B]$ и $GTr_2^{(1)} = [Tr_2^*, B]$. Если отрезок $L_1 = Tr_1^*(end)B$ не пересекается ни с одним из препятствий, то $GTr_1^{(1)}$ – путь, который мы отнесем во множество $GTr^{(1)}$ отобранных глобальных путей на 1-й итерации. Точно так же поступим с $GTr_2^{(1)}$, если только отрезок $L_1 = Tr_2^*(end)B$ не пересекает ни одно из препятствий сцены. При этом, если множество $GTr^{(1)}$ имеет два элемента, то минимальный по длине путь из двух и дает искомое решение. Иначе, если мощность множества не более 1, то выбирается тот из глобальных путей, который дает пересечение с $\{Obs\}$. Предположим, что $|GTr^{(1)}| = 0$ и рассмотрим ветвление обоих глобальных путей.

Пусть $Obs_{j'}$ есть препятствие, с которым пересекается отрезок L_1 . Заметим, что по построению L_1 не может пересечься с Obs_j , поэтому $j' \neq j$. Теперь повторим рассуждения, которые мы применили ранее к точке A и отрезку L_0 , но теперь – в отношении точки $Tr_1^*(end)$ и отрезка L_1 . Находим частичные пути $Tr_{11} = [Tr_1^*(end), M_{j'1}, \partial_1(Obs_{j'})]$ и $Tr_{12} = [Tr_1^*(end), M_{j'1}, \partial_2(Obs_{j'})]$, где $Tr_1^*(end)$ и $M_{j'2}$ есть крайние, ближайшая и дальняя точки пересечения прямолинейного отрезка L_1 с препятствием $Obs_{j'}$; $\partial_1(Obs_{j'})$ и $\partial_2(Obs_{j'})$ есть соответствующие непрерывные по нумерации своих элементов связные подмножества точечной границы $\partial(Obs_{j'})$, удовлетворяющие условиям: $\partial_1(Obs_{j'}) \cup \partial_2(Obs_{j'}) = \partial(Obs_{j'})$, $\partial_1(Obs_{j'}) \cap \partial_2(Obs_{j'}) = \{M_{j'1}, M_{j'2}\}$. Далее строится оптимизированный частичный путь Tr_{11}^* , удовлетворяющий условиям, аналогичным указанным в пунктах а) – г) для Tr_1^* . Здесь и далее обозначение «end», фигурирующее в круглых скобках при некоторых, в общем случае, частичных, траекториях типа Tr_{11} , Tr_1^* , Tr_{12} , и т.п., означает число узлов соответствующей траектории.

Теперь можно объединить Tr_1^* и Tr_{11}^* с получением первого глобального пути на второй итерации:

$$GTr_1^{(2)} = [Tr_1^*(1: end - 1), Tr_{11}^*, B]. \quad (1)$$

Точно так же оптимизируем второй частичный путь с получением множества точек Tr_{12}^* , после чего можно построить начало второго глобального пути на второй итерации:

$$GTr_2^{(2)} = [Tr_1^*(1: end - 1), Tr_{12}^*, B]. \quad (2)$$

Далее берем частичный путь Tr_2^* , полученный на первой итерации, и находим для него частичные пути, отвечающие ветвлению в конечной точке $Tr_2^*(end)$ первой итерации. Вновь исключаем тривиальный случай, и находим ближайшее к этой точке препятствие по направлению отрезка $Tr_2^*(end)B$ - пусть $Obs_{k'}$ с точками пересечения $M_{k'1}, M_{k'2}$, тогда получим частичные огибающие траектории $Tr_{21} = [Tr_2^*(end), M_{k'1}, \partial_1(Obs_{k'})]$ и $Tr_{22} = [Tr_2^*(end), M_{k'1}, \partial_2(Obs_{k'})]$.

Соответствующие глобальные пути, порожденные рассматриваемой процедурой ветвления в точке $Tr_2^*(end)$, есть:

$$GTr_3^{(2)} = [Tr_2^*(1: end - 1), Tr_{21}^*], \quad GTr_4^{(2)} = [Tr_2^*(1: end - 1), Tr_{22}^*]. \quad (3)$$

Очевидно, что мощность множества отобранных глобальных путей на второй итерации может быть не более четырех.

На рис. 2 приведены построенные частичные, исходные и оптимизированные, пути для рассматриваемого примера на 2-й итерации: Tr_{11} , Tr_{12} , Tr_{21} , Tr_{22} и Tr_{11}^* , Tr_{12}^* , Tr_{21}^* , Tr_{22}^* .

Рассмотренная процедура ветвления приводит к постоянному зарождению новых путей, причем ветвление в данной точке какого-либо частичного пути на текущей u -й итерации прекращается, если эта точка совпала с целевой B . Такой путь считается глобальным завершённым.

Для структурного описания и анализа получающегося итерационного процесса может быть построен ориентированный граф, в узлах которого располагаются концевые точки каждого частичного оптимизированного пути; его ребра обозначают оптимизированные кусочно-линейные пути, соединяющие соответствующие узлы, причем длины таких путей будем считать весами обозначающих их ребер. Например, на рис. 3 приведен такой граф для рассмотренного выше примера; здесь, в частности, A'_{11} есть точка A_{j_2} , а также выполнены следующие соответствия (см. рис. 1): $A'_{12} = A_{j_4}$, $A'_{111} = A_{j_2}$, $A'_{112} = A_{j_4}$, $A'_{121} = A_{k/3}$, $A'_{122} = A_{k/2}$.

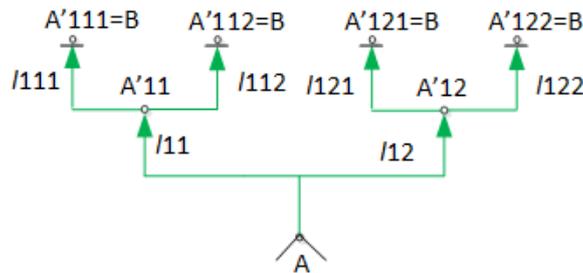


Рис. 3. Граф для случая без замыканий

Численный анализ разработанного алгоритма показал, что возможны случаи, когда одна или более пар разных узлов такого графа совпадают на некоторой итерации. В этом случае итерационный процесс приводит к пересечению двух или более разных путей, в некоторой одной характерной точке какого-либо препятствия. Например, на рис. 4 рассмотрен более сложный итерационный процесс, мощность множества всех его глобальных путей, без учета динамического отсеивания (см. далее), составляет 10. При этом узел A'_{12211} совпадает с A'_{11111} , аналогично, A'_{11211} - с A'_{122} . При каждом таком совпадении, по сути, происходит поглощение либо исключение определенного подмножества глобальных путей на текущей итерации по следующему правилу. Пусть узел \bar{A}' совпал с узлом \hat{A}' на u -й итерации. Для определенности будем считать, что узел \hat{A}' получен на некоторой предшествующей итерации $\hat{u} < u$. Пусть $GTr_{\bar{A}'}^{(u)}$ и $GTr_{\hat{A}'}^{(\hat{u})}$ есть глобальные пути этих двух узлов, полученные на u -й и \hat{u} -й итерациях, соответственно. Без учета замыкающих прямых отрезков $\bar{A}'B$ и $\hat{A}'B$ длины этих путей есть $L(\bar{A}')$ и $L(\hat{A}')$, соответственно. Если выполнено равенство $\hat{A}' = \bar{A}'$ и условие

$$L(\bar{A}') \leq L(\hat{A}'), \tag{4}$$

тогда необходимо заменить все глобальные пути, порожденные точкой \hat{A}' и полученные на всех итерациях от $\hat{u} + 1$ до u , на новые по правилу:

$$GTr_{\hat{A}'}^{(u)} \rightarrow (GTr_{\hat{A}'}^{(u)} \setminus \hat{A}'B) \cup [GTr_{\bar{A}'}^{(u)} \setminus (GTr_{\hat{A}'}^{(\hat{u})} \setminus \hat{A}'B)], \tag{5}$$

означающему, что одинаковые части $GTr_{\hat{A}'}^{(\hat{u})} \setminus \hat{A}'B$ всех таких путей до точки \hat{A}' заменяются на кусочно-ломаную $GTr_{\bar{A}'}^{(u)} \setminus \hat{A}'B$, порожденную узлом \bar{A}' . Если же указанное условие не выполняется, т.е. $L(\bar{A}') > L(\hat{A}')$, то, наоборот, глобальный путь,

порождаемый узлом \bar{A}' на u -й итерации – исключается, как проигрышный по длине по сравнению с ранее полученным вариантом альтернативного движения до точки \hat{A}' . Понятно, что это исключает все ветвления из этого узла, которые ранее считались потенциально возможными (до фиксации совпадения $\hat{A}' = \bar{A}'$ и условия $L(\bar{A}') > L(\hat{A}')$). Здесь “ \setminus ” означает операцию вычитания одного множества из другого.

Граф для любого итерационного процесса, приводящего к совпадению двух или более пар (троек, четверок и т.д.) узлов, имеет дополнительные циклы, порожденные указанными перекрестными связями вершин, порождаемых разными родительскими вершинами. Для рассмотренного на рис. 4 случая граф имеет два таких дополнительных цикла.

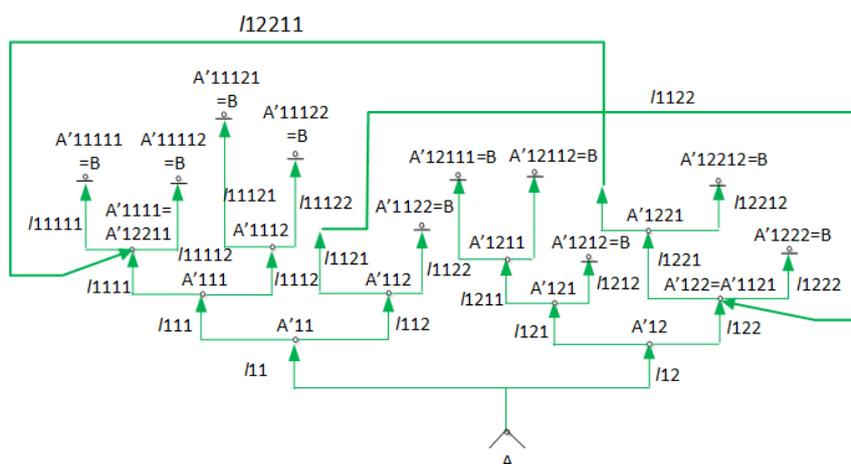


Рис. 4. Граф для случая с 2-я замыканиями

Будем использовать следующий сквозной порядок нумерации путей, которые возникают/продолжаются на каждой итерации. Первая точка ветвления порождает обход первого встретившегося препятствия слева – этот путь считаем первым, и соответствующий обход справа вторым; далее, нумеруем возникающие на каждой итерации новые пути таким образом, чтобы для каждого разветвления правый путь, если он еще не исключен условиями (5) или (6), имел на единицу больший номер, чем левый; если номер пути, подходящего к некоторому узлу M , есть n_M , то левый путь из точки M будет иметь номер $n_M + 1$, а правый - $n_M + 2$. Также потребуем, чтобы любой путь, порожденный разветвлениями $n_M + 1$ -го пути, имел номер, меньший любого пути, порожденного разветвлением $n_M + 2$ -го пути. Поскольку часть путей будет исключена по условиям (5) и (6), то на каждом u -м шаге, при $u > 1$ для некоторых соседних путей в построенной упорядоченной последовательности номера могут отличаться более чем на единицу. На u -й итерации последовательность путей есть: $\{GTr_{n(l,u)}^{(u)}\}$, $l = 1, 2, \dots, l_{max}$, где $n(l, u)$ есть номер l -го по порядку следования пути в упорядоченном наборе путей, существующих на данной итерации, в сквозной нумерации.

Заметим, что нет необходимости доводить всякий раз все ветви графа до логического завершения – достижения целевой точки. Действительно, рассмотрим способ отсеивания глобальных путей, значительно увеличивающий быстродействие алгоритма. Как только процедура ветвления на некотором своем u -м шаге получает набор из одного или нескольких завершенных путей, тогда среди них вы-

бирается путь с минимальной длиной и фиксируется также эта длина как $L_{curr}(u)$. Тогда на следующем $u+1$ -м шаге производится сравнение приведенной длины каждого l -го глобального пути с величиной $L_{curr}(u)$ с отбрасыванием всех тех глобальных путей, которые получили превышение по длине над $L_{curr}(u)$. А именно, для всех $l = 1, 2, \dots, l_{max}(u+1)$ проверяются условия:

$$L(P_l) + |P_l B| = |GT r_n^{(u+1)}| \geq L_{curr}(u), \quad (6)$$

где P_l есть предпоследняя точка глобального пути $GT r_l^{(u+1)}$, полученного на u -й итерации. Напомним, что конечной точкой для любого глобального пути всегда является целевая. Здесь $l_{max}(u+1)$ есть число всех глобальных путей, полученных на $u+1$ -й итерации и не отсеянных ранее.

При выполнении для некоторого l неравенства (6) далее l -й глобальный путь не должен производить ветвлений, поскольку каждый из производных путей, полученных экстраполяцией ветвления l -го пути вплоть до их завершения, будет иметь длину, не меньшую чем левая часть неравенства (6).

Предположим, что вследствие итерационного процесса найден путь $GT r_{n_{opt}}^{(u_{max})}$ с минимальной длиной на множестве всех глобальных путей:

$$\left| GT r_{n_{opt}}^{(u_{max})} \right| = \min_{1 \leq l \leq l_{max}(u_{max})} \left| GT r_{n(l, u_{max})}^{(u_{max})} \right|, \quad n_{opt} = n(l_{opt}, u_{max}). \quad (7)$$

Рассмотренную выше последовательность действий, приводящую к построению $GT r_{n_{opt}}^{(u_{max})}$, назовем алгоритм построения глобального пути первого приближения, или первым этапом алгоритма планирования.

Второй этап подразумевает корректировку описанного выше алгоритма. Покажем, что отсеиваемые по критериям (4) и (6) глобальные пути могут быть в общем случае оптимизированы по длине, и поэтому получение результирующего пути алгоритма сопряжено с учетом не только оптимального по длине пути $GT r_{n_{opt}}^{(u_{max})}$ на множестве $GT r_{n(l, u_{max})}^{(u_{max})}$, но и, в общем случае, всех путей этого множества.

Действительно, при обходе некоторого препятствия П1 по принципу построения виртуальной прямой видимости может возникнуть ситуация, когда в некоторой окрестности этого препятствия могут быть другие, которые увеличивают длину участка кусочно-ломаной, идущей до характерной точки препятствия П1. Подобная ситуация показана на рис. 3. Препятствие П2 не допускает проложить локальный маршрут AA_3 , и для обхода П2 первоначально строится более длинный локальный путь AA_2A_3 . Вместе с тем, наличие препятствия П2 позволяет решить локальную задачу его обхода при движении от точки A к A_3 , с тем, чтобы сократить длину этого участка пути. В результате получатся два пути $AA_5A_6A_3$ и AA_7A_3 длины которых меньше, чем у пути AA_2A_3 . Заметим, что препятствий, подобных П2 по влиянию на возможность быстрого достижения точки A_5 , может быть более одного. В этом случае целесообразно применить последовательность действий первого этапа алгоритма, рассмотренную выше, ко всем тройкам узлов построенной на этом этапе кусочно-ломаной $Tr(1: N) = [A_1, A_2, \dots, A_N]$ с числом узлов N .

Суть второго этапа планирования заключается в следующем. Для каждой тройки узлов $Tr_k(1: 3) \equiv [A_k, A_{k+1}, A_{k+2}]$, $k=1: N-2$, в цикле по k применяется рассмотренный выше алгоритм первого этапа, причем точка A_k считается начальной, а A_{k+2} – конечной. Если длина $|Tr'_k|$ полученной кусочно-ломаной $Tr'_k(1: N_k) \equiv [A_k, A'_{k1}, A'_{k2}, \dots, A'_{kN_k-1}, A_{k+2}]$, $N_k \geq 2$, меньше длины $|Tr_k|$, то производится замена участка $[A_k, A_{k+1}, A_{k+2}]$ на $Tr'_k(1: N_k)$. При этом производится инкрементация $k:=k+2$ и действия в цикле повторяются для новой тройки с начальной точ-

кой, совпадающей с конечной предыдущей тройки узлов. Если же $|Tr'_k| \geq |Tr_k|$, тогда производится инкрементация $k:=k+1$ и действия в цикле повторяются для новой тройки, у которой два первых узла совпадают с последними двумя предыдущей тройки. В результате проходит набор троек точек, полностью покрывающих множество узлов траектории $Tr(1:N)$. Поскольку указанный перебор троек может оставить незадействованными в корректировке ряд узлов, в которых некоторые тройки стыкуются, то необходимо применить рассмотренную корректировку к траектории $Tr'_k(1:N_k)$, т.е. провести вторую итерацию алгоритма. Кроме того, последующие итерации необходимы также для того, чтобы учесть препятствия, которые появляются на линиях прямой видимости некоторых троек точек очередной скорректированной траектории. Начиная с некоторой итерации, новых препятствий появляться не будет, в силу их конечного числа. Далее рассматривается завершающая корректировка: учесть узлы, которые не были подвержены корректирующему перемещению, если таковые имеются. Пусть есть некоторая точка с номером A'_l , $l \geq 3$, которая не была подвержена корректировке. И пусть она является границей двух троек, которые перебираются указанным выше образом для кусочно-ломаной на некоторой итерации. Отсутствие новых препятствий, которые необходимо учитывать при корректировке выделенных троек точек, означает, что тройка $[A_{l-2}, A_{l-1}, A_l]$ не будет более корректироваться, но тогда будет произведено инкрементирование на единицу номера $k=l-1$ и будет рассмотрена тройка $[A_{l-1}, A_l, A_{l+1}]$, что позволит произвести корректирующее смещение точки A_l , если это необходимо. Если даже корректировка ранее не рассмотренных точек приведет к необходимости учета новых препятствий, то в силу их конечного числа получим, что на некоторой итерации исчерпаются как незадействованные точки, так и препятствия, и траектория далее меняться не будет, так как *любая* тройка соседних ее точек будет более не улучшаема по длине.

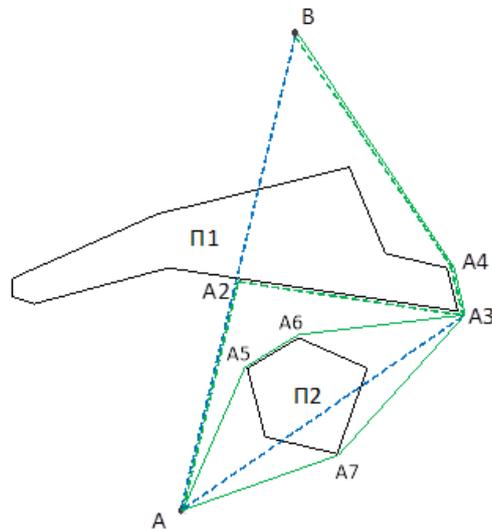


Рис. 5. К обоснованию необходимости второго этапа алгоритма

Рассмотренной выше корректировке необходимо подвергнуть каждый путь из множества $GTr_{n(l,umax)}^{(umax)}$, а далее выбрать наименьший по длине, чем и завершается алгоритм планирования перемещения в 2d-среде с полигональными препятствиями.

III. результаты численного исследования. Результаты численного моделирования представлены на рис. 6 и 7. При моделировании использован программный комплекс, разработанный ранее для исследования различных алгоритмов планирования и управления движением [18, 19].

На рис. 6 и 7 представлены результаты расчета целевых траекторий для двух сцен с помощью двух алгоритмов: рассматриваемого и Дейкстры. Каждая сцена здесь представляет собой сложный лабиринт с препятствиями, каждое из которых есть связанное объединение некоторой совокупности ячеек прямоугольной равномерной сетки. Для первой сцены длина траектории, построенной по рассматриваемому алгоритму и показанной сплошной кривой, составляет $L_A = 1.535e+03$ условных единиц длины, при времени расчета $T_A = 5,3$ с, а для траектории, построенной по алгоритму Дейкстры и показанной пунктирной кривой, – $L_D = 1.564741996484947e+03$ при времени расчета $T_D = 14,4$ с. Во втором случае $L_A = 1.446e+03$ при времени расчета $T_A = 7,5$ с, $L_D = 1.467e+03$ при времени расчета $T_D = 8,7$ с.

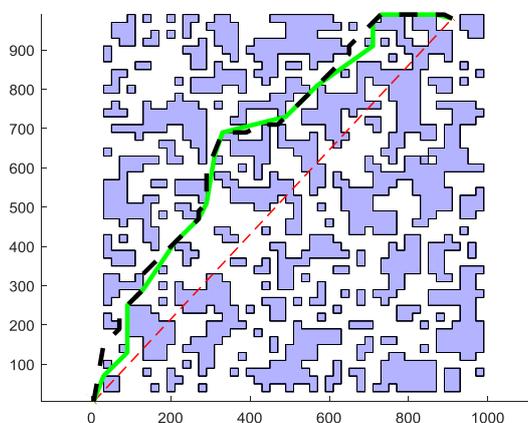


Рис. 6. Результаты моделирования на случайно сгенерированной сцене 1

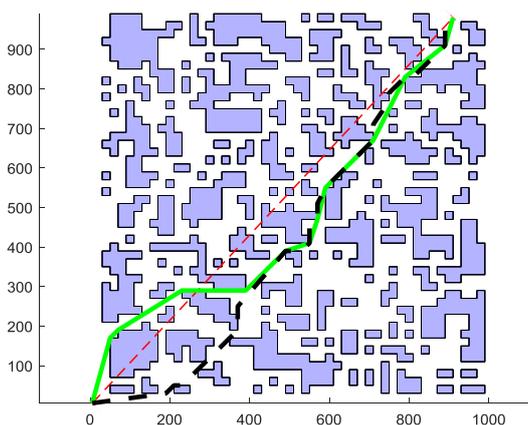


Рис. 7. Результаты моделирования на случайно сгенерированной сцене 2

На рис. 6 и 7 сплошной линией показаны результаты моделирования разработанного в данной статье алгоритма, а пунктирной линией – результаты моделирования алгоритма Дейкстры.

Проведенные численные исследования на различных сценах, при случайном генерируемых картах и положениях начальной и целевой точек показали, что разработанный аналитический алгоритм выигрывает у алгоритма Дейкстры, как по длине найденного пути (в среднем, на 4%), так и по времени расчета (в среднем, на 10%).

Численный анализ показал наличие следующей тенденции: с ростом сложности препятствий рассматриваемый алгоритм выигрывает по быстродействию у алгоритмов клеточной декомпозиции на один-два порядка и более, потому что достаточно точный учет формы этих препятствий требует соответствующего детального клеточного разбиения, а это резко снижает быстродействие алгоритмов Дейкстры, A* и других аналогичных алгоритмов.

С другой стороны, окрестности препятствий прямоугольного типа, препятствий, полученных объединением прямоугольных ячеек, или приближенных к ним, могут быть аппроксимированы меньшим количеством ячеек сетки, чем для препятствий со сложной полигональной границей при одной и той же точности приближения.

Заметим, что разработанный алгоритм может быть эффективно использован и при планировании перемещений группы роботов в средах с источниками помеховых возмущений со сложными по форме геометрическими областями действия; для этого дополнительно может быть использован аппарат характеристических вероятностных функций и/или функций риска с соответствующими методами решения задач поиска целевой траектории с наибольшей вероятностью успешного прохождения [20, 21].

Заключение. В работе представлена итерационная процедура, в которой использован bug-алгоритм локального планирования. В ходе разработки указанной процедуры как шага специального итерационного процесса оптимизации на графе, отражающем связь характерных точек полигональных препятствий кусочно-ломаными траекториями, bug-алгоритм обобщен до глобального алгоритма планирования.

Особенностью алгоритма является итерационный процесс анализа соответствий между текущим отрезком «прямой видимости», соединяющим текущую точку с целевой, и криволинейными полигональными границами препятствий, пересекаемых им.

В отличие от исходных bug-алгоритмов, предлагаемый метод решает задачу глобального поиска пути на классе заданных траекторий с полигональными препятствиями; а в отличие от классических методов диаграмм прямой видимости, существенно снижает размерность графа за счет специального выбора ограниченного количества характерных точек соответствующих препятствий.

Показано, что для прокладывания пути на сценах лабиринтного типа со сложными препятствиями, являющимися связным объединением некоторых ячеек прямоугольного сеточного разбиения, рассматриваемый алгоритм на 10% превосходит оптимальный алгоритм Дейкстры. В том случае, когда препятствия будут иметь более сложную полигональную форму, разработанный алгоритм будет еще более превосходить алгоритм Дейкстры по вычислительной эффективности.

Превосходство методов клеточной декомпозиции над разработанным методом прогнозируется только в тех случаях, когда число препятствий сцены начнет сильно превосходить число свободных ячеек клеточного разбиения.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Han-Pang Huang, Shu-Yun Chung*. Dynamic visibility graph for path planning // IEEE-RSJ International Conference on Intelligent Robots and Systems. Sendai, Japan, 2004. – Vol. 3. – P. 2813-2818.
2. *Janet J.A., Luo R.C., Kay M.G.* The essential visibility graph: An approach to global motion planning for autonomous mobile robots // IEEE International Conference on Robotics and Automation. Nagoya, Japan, 1995. – Vol. 2. – P. 1958-1963.
3. *Habib M.K., Asama H.* Efficient method to generate collision free paths for an autonomous mobile robot based on new free space structuring approach // IEEE-RSJ International Conference on Intelligent Robots and Systems. Osaka, Japan, 1991. – Vol. 2. – P. 563-567.
4. *Wallgrun J.O.* Voronoi graph matching for robot localization and mapping // Transactions on computational science. – Springer. 2010. – P. 76-108.
5. *LaValle S.M., Kuffner J.* Randomized kinodynamic planning // Int. Journal of Robotics Research. – 2001. – Vol. 20 (5). – P. 378-400.
6. *Пишихов В.Х., Медведев М.Ю., Бросалин Д.О., Васильева М.А., Гуренко Б.В., Хамдан Низар*. Исследование методов планирования движения в двумерных картографированных средах // Известия ЮФУ. Технические науки. – 2022. – № 3 (227). – С. 170-192.
7. *Пишихов В.Х., Медведев М.Ю., Костюков В.А., Хуссейн Ф., Кадим А.* Алгоритмы планирования траекторий в двумерной среде с препятствиями // Информатика и автоматизация. – 2022. – Т. 21 (3). – С. 459-492.
8. *Сухарев А.Г.* Об оптимальных стратегиях поиска экстремума // Журнал вычислительной математики и математической физики. – 1971. – Т. 11 (4). – С. 910-924.
9. *Hornung A., Wurm K.M., Bennet M., Stachniss C., Burgard W.* OctoMap: An efficient probabilistic 3D mapping framework based on octrees // Autonomous Robots. – 2013. – Vol. 34 (3). – P. 189-206.
10. *Janson L., Ichter B., Pavone M.* Deterministic sampling-based motion planning: Optimality, complexity, and performance // International Journal of Robotics Research. – 2018. – Vol. 37 (1). – P. 46-61.
11. *Hart P.E., Nilsson N.J., Raphael B.A.* Formal Basis for the Heuristic Determination of Minimum Cost Paths // IEEE Transactions on Systems Science and Cybernetics. – 1968. – Vol. 2. – P. 100-107.
12. *Stentz A.* Optimal and efficient path planning for partially known environments // In Intelligent Unmanned Ground Vehicles. – Springer, Boston, MA, USA, 1997. – P. 203-220.
13. *Koenig S., Likhachev M., Furcy D.* Lifelong Planning A* // Artificial Intelligence. – 2004. – Vol. 155 (1-2). – P. 93-146.
14. *Казаков К.А., Семенов В.А.* Обзор современных методов планирования движения // Тр. ИСП РАН. – 2016. – Т. 28 (4). – С. 241-294.
15. *Ng J., Braunl T.* Performance comparison of bug navigation algorithms // Journal of Intelligent and Robotic Systems. – 2007. – Vol. 50 (1). – P. 73-84.
16. *Lumelsky V., Stepanov A.* Dynamic path planning for a mobile automaton with limited information on the environment // IEEE Transactions on Automatic Control. – 1986. – Vol. 31 (11). – P. 1058-1063.
17. *Yufka A., Parlaktuna O.* Performance comparison of the BUG's algorithms for mobile robots // Proc. of International Symposium on Innovations in Intelligent Systems and Applications. Trabzon, Turkey, 2009. – P. 416-421.
18. *Костюков В.А., Медведев М.Ю., Пишихов В.Х.* Планирование движения наземных роботов в среде с препятствиями: алгоритмы построения траекторий в группе при заданном шаблоне // Мехатроника, автоматизация, управление. – 2023. – Т. 24. № 1. – С. 33-45.
19. *Костюков В.А., Медведев М.Ю., Пишихов В.Х.* Алгоритмы планирования сглаженных индивидуальных траекторий движения наземных роботов // Мехатроника, автоматизация, управление. – 2022. – Т. 23 (11). – С. 585-595.
20. *Kostjukov V., Pshikhopov V., Medvedev M.* Optimization of mobile robot movement on a plane with finite number of repeller sources // SPIIRAS Proceedings. – 2020. – Vol. 19 (1). – P. 43-78.
21. *Zabarankin M., Uryasev S., Pardalos P.* Optimal Risk Path Algorithms // Cooperative Control and Optimizat. Dordrecht: Kluwer Acad., 2002. – P. 271-303.

REFERENCES

1. Han-Pang Huang, Shu-Yun Chung. Dynamic visibility graph for path planning, *IEEE-RSJ International Conference on Intelligent Robots and Systems. Sendai, Japan, 2004*, Vol. 3, pp. 2813-2818.
2. Janet J.A., Luo R.C., Kay M.G. The essential visibility graph: An approach to global motion planning for autonomous mobile robots, *IEEE International Conference on Robotics and Automation. Nagoya, Japan, 1995*, Vol. 2, pp. 1958-1963.
3. Habib M.K., Asama H. Efficient method to generate collision free paths for an autonomous mobile robot based on new free space structuring approach, *IEEE-RSJ International Conference on Intelligent Robots and Systems. Osaka, Japan, 1991*, Vol. 2, pp. 563-567.
4. Wallgrun J.O. Voronoi graph matching for robot localization and mapping, *Transactions on computational science*. Springer. 2010, pp. 76-108.
5. LaValle S.M., Kuffner J. Randomized kinodynamic planning, *Int. Journal of Robotics Research*, 2001, Vol. 20 (5), pp. 378-400.
6. Pshikhopov V.Kh., Medvedev M.Yu., Brosalin D.O., Vasil'eva M.A., Gurenko B.V., Khamdan Nizar. Issledovanie metodov planirovaniya dvizheniya v dvumernykh kartografirovannykh sredakh [Study of path planning methods in two-dimensional mapped environments], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2022, No. 3 (227), pp. 170-192.
7. Pshikhopov V.Kh., Medvedev M.Yu., Kostyukov V.A., Khusseyn F., Kadim A. Algoritmy planirovaniya traektoriy v dvumernoy srede s prepyatstviyami [Algorithms for trajectory planning in a two-dimensional environment with obstacles], *Informatika i avtomatizatsiya* [Informatics and Automation], 2022, Vol. 21 (3), pp. 459-492.
8. Sukharev A.G. Ob optimal'nykh strategiyakh poiska ekstremuma [About optimal strategies for finding the extremum], *Zhurnal vychislitel'noy matematiki i matematicheskoy fiziki* [Journal of Computational Mathematics and Mathematical Physics], 1971, Vol. 11 (4), pp. 910-924.
9. Hornung A., Wurm K.M., Bennewitz M., Stachniss C., Burgard W. OctoMap: An efficient probabilistic 3D mapping framework based on octrees, *Autonomous Robots*, 2013, Vol. 34 (3), pp. 189-206.
10. Janson L., Ihter B., Pavone M. Deterministic sampling-based motion planning: Optimality, complexity, and performance, *International Journal of Robotics Research*, 2018, Vol. 37 (1), pp. 46-61.
11. Hart P.E., Nilsson N.J., Raphael B.A. Formal Basis for the Heuristic Determination of Minimum Cost Paths, *IEEE Transactions on Systems Science and Cybernetics*, 1968, Vol. 2, pp. 100-107.
12. Stentz A. Optimal and efficient path planning for partially known environments, *In Intelligent Unmanned Ground Vehicles*. Springer, Boston, MA, USA, 1997, pp. 203-220.
13. Koenig S., Likhachev M., Furcy D. Lifelong Planning A*, *Artificial Intelligence*, 2004, Vol. 155 (1-2), pp. 93-146.
14. Kazakov K.A., Semenov V.A. Obzor sovremennykh metodov planirovaniya dvizheniya [Overview of modern traffic planning methods], *Tr. ISP RAN* [Proceedings of ISP RAS], 2016, Vol. 28 (4), pp. 241-294.
15. Ng J., Braunl T. Performance comparison of bug navigation algorithms, *Journal of Intelligent and Robotic Systems*, 2007, Vol. 50 (1), pp. 73-84.
16. Lumelsky V., Stepanov A. Dynamic path planning for a mobile automaton with limited information on the environment, *IEEE Transactions on Automatic Control*, 1986, Vol. 31 (11), pp. 1058-1063.
17. Yufka A., Parlaktuna O. Performance comparison of the BUG's algorithms for mobile robots, *Proc. of International Symposium on Innovations in Intelligent Systems and Applications. Trabzon, Turkey, 2009*, pp. 416-421.
18. Kostyukov V.A., Medvedev M.Yu., Pshikhopov V.Kh. Planirovanie dvizheniya nazemnykh robotov v srede s prepyatstviyami: algoritmy postroeniya traektoriy v gruppe pri zadannom shablone [Planning the movement of ground robots in an environment with obstacles: algorithms for constructing trajectories in a group with a given formation], *Mekhatronika, avtomatizatsiya, upravlenie* [Mechatronics, Automation, Control], 2023, Vol. 24, No. 1, pp. 33-45.

19. Kostyukov V.A., Medvedev M.Yu., Pshikhopov V.Kh. Algoritmy planirovaniya sglazhennykh individual'nykh traektoriy dvizheniya nazemnykh robotov [Algorithms for planning smoothed individual trajectories of ground robots], *Mekhatronika, avtomatizatsiya, upravlenie* [Mechatronics, Automation, Control], 2022, Vol. 23 (11), pp. 585-595.
20. Kostjukov V. Pshikhopov V., Medvedev M. Optimization of mobile robot movement on a plane with finite number of repeller sources, *SPIIRAS Proceedings*, 2020, Vol. 19 (1), pp. 43-78.
21. Zabaranin M., Uryasev S., Pardalos P. Optimal Risk Path Algorithms, *Cooperative Control and Optimizatoin. Dordrecht: Kluwer Acad., 2002*, pp. 271-303.

Статью рекомендовал к опубликованию д.т.н., профессор В.В. Курейчик.

Костюков Владимир Александрович – НКБ «РиСУ»; e-mail: wkost-einheit@yandex.ru; г. Таганрог, Россия; тел.: 88634371694; к.т.н.; с.н.с.

Медведев Михаил Юрьевич – НИИ робототехники и процессов управления Южного федерального университета; e-mail: medvmihal@sfedu.ru; г. Таганрог, Россия; тел.: 88634371694; д.т.н.; в.н.с.

Пшихопов Вячеслав Хасанович – e-mail: pshichop@rambler.ru; тел.: 88634371694; д.т.н.; профессор; директор.

Kostyukov Vladimir Aleksandrovich – Joint-Stock Company "Robotics and Control Systems"; e-mail: wkost-einheit@yandex.ru; Taganrog, Russia; phone: 88634371694; can. of eng. sc.; senior researcher.

Medvedev Mikhail Yurjevich – R&D Institute of Robotics and Control Systems; e-mail: medvmihal@sfedu.ru; Taganrog, Russia; phone: 88634371694; dr. of eng. sc.; leading researcher.

Pshikhopov Viacheslav Khasanovich – e-mail: pshichop@rambler.ru; phone: 88634371694; dr. of eng. sc.; professor; director.

УДК 621.3.049.771.14

DOI 10.18522/2311-3103-2023-5-48-57

Д.Б. Шокарев, Р.Ж. Чочаев, А.Н. Щелоков, С.В. Гаврилов

РАЗРАБОТКА АЛГОРИТМА ДЕТАЛЬНОГО РАЗМЕЩЕНИЯ НА ПЛИС

Иерархические программируемые логические интегральные схемы (ПЛИС) состоят из множества логических блоков, объединенных в группы. Для успешной трассировки необходимо оптимальное размещение элементов в пределах групп с учётом особенностей архитектуры локальных связей. Классические алгоритмы не способны обеспечить учёт различных особенностей архитектуры. Решение данной проблемы возможно только путем разработки специализированных алгоритмов. В данной работе представлен алгоритм детального размещения, в котором для вычисления оптимальных позиций элементов в группе была разработана новая метрика, позволяющая оценить количество доступных локальных связей между элементами в группах логических блоков с учётом особенностей архитектуры связей между ними. Алгоритм детального размещения состоит из нескольких этапов. На первом этапе группа логических элементов представляется в виде ориентированного графа. На втором этапе определяется порядок размещения логических элементов в группе с помощью алгоритма поиска в ширину. На финальном этапе для каждого элемента, согласно полученному порядку, определяется оптимальное размещение в группе с учётом разработанной метрики. Если среди свободных позиций для размещения в группе нет оптимальной, то проверяются занятые позиции. Текущий элемент назначается на занятую позицию, а для замененного элемента выполняется поиск новой. Такая замена может проводиться многократно, увеличивая вероятность нахождения оптимальной конфигурации. Предложенный алгоритм был реализован и протестирован на наборах тестовых схем. На основе результатов тестирования выполнено сравнение представленного алгоритма с алгоритмом последовательного размещения. Сравнение алгоритмов показало, что приме-

нение разработанного алгоритма в маршруте проектирования в базе специализированной ПЛИС позволяет сократить в среднем на 10% количество задействованных в трассировке глобальных коммутационных шин и увеличить количество используемых локальных трассировочных ресурсов в среднем на 30%. Полученные результаты подтверждают работоспособность алгоритма и доказывают, что внедрение учета архитектуры внутренних связей ПЛИС повышает эффективность использования доступных трассировочных ресурсов.

Размещение; автоматизация проектирования; ПЛИС.

D.B. Shokarev, R.Zh. Chochaev, A.N. Schelokov, S.V. Gavrilov

DEVELOPMENT OF THE DETAILED PLACEMENT ALGORITHM FOR FPGAS

Hierarchical field-programmable gate arrays (FPGAs) consist of an array of programmable logic blocks arranged into groups. Successful routing requires optimal placement of logic elements within the groups, considering the architectural features of the local interconnections. Classical algorithms are not able to consider these features. That's why, the development of new algorithms is required. In this paper, we present a detailed placement algorithm with a new metric that allows us to estimate the number of available local interconnections inside the groups of logic blocks, considering the architectural features of the local interconnections. The detailed placement algorithm consists of several stages. At the first stage, the group of logic elements is transformed into a directed graph. Then, the placement order of logic elements in the group is determined using the breadth-first search algorithm. At the final stage, for each element, according to the obtained order, the optimal placement in the group is determined, considering the new metric. If there is no optimal position in the group among the free ones, the occupied positions are checked. The current element is placed in the occupied position, and a new position is searched for the replaced element. Such replacements can be performed repeatedly, increasing the probability of finding the optimal placement configuration. The proposed algorithm was verified on a set of benchmark circuits ISCAS'85, ISCAS'89, Cpu8080 and VGA. Experimental results show that the developed algorithm reduces the number of global interconnections used for global routing by 10% on average and increases the number of local interconnections used for detailed routing by 30% on average compared to the sequential placement algorithm. The average routing time remained unchanged.

Placement; electronic design automation (EDA); FPGA.

Введение. Программируемые логические интегральные схемы (ПЛИС) – это цифровые микросхемы, логика которых задаётся посредством программирования (проектирования) [1]. В данной работе рассматривается специализированная ПЛИС, обладающая классической иерархической архитектурой, описанной в работах [2–5]. Она состоит из множества логических блоков (ЛБ), объединенных в группы логических блоков (ГЛБ) по 16 ЛБ. Каждый ЛБ состоит из двух основных частей: таблицы преобразования (англ. *look-up table*, LUT), реализующей таблицу истинности логической функции, и триггера.

В иерархической архитектуре рассматриваемой ПЛИС существует разделение трассировочных ресурсов на локальные (в пределах ГЛБ) и глобальные (между ГЛБ). Из-за этого соединение всех ЛБ в пределах ГЛБ осуществляется через коммутатор, описание которого приведено в статье [4]. Коммутатор связывает между собой три типа шин: глобальные шины (ГШ), соединяющие различные ГЛБ, шины локальной обратной связи (ШЛОС), соединяющие ЛБ одной группы, и цепи данных, соединенные непосредственно с входными терминалами ЛБ.

Одной из особенностей архитектуры рассматриваемой ПЛИС является нестандартное устройство коммутаторов. В классических ПЛИС в коммутаторы закладывается возможность создания связи между любыми выходами и любыми входами ЛБ одной группы [6]. Такой подход дает максимальную гибкость, но име-

ет серьезный недостаток – неэффективное использование площади кристалла, так как одновременно может использоваться лишь часть заложенных трассировочных ресурсов. В рассматриваемой ПЛИС количество связей между ЛБ одной группы было значительно снижено путем редукции связей между парами ЛБ таким образом, что выход каждого ЛБ может быть соединен либо с первым или третьим, либо со вторым или четвертым входами других ЛБ той же группы.

Данное решение позволяет сократить площадь коммутаторов, но при отсутствии учета архитектуры связей лишает возможности реализовать часть локальных связей в пределах ГЛБ и приводит к неоптимальному использованию глобальных коммутационных шин. В результате наблюдается рост задержки между соседними ЛБ и перегрузка соединений между ГЛБ.

В маршруте топологического проектирования в базе иерархических ПЛИС этап размещения состоит из трех шагов [7]. На первом шаге выполняется кластеризация логических элементов (ЛЭ) проектируемой схемы в группы [8–9]. На втором шаге вычисляется глобальное размещение полученных групп логических элементов [10–12], т.е. каждой группе назначается ГЛБ в ПЛИС. На финальном шаге выполняется детальное размещение [13–14], т.е. логические элементы в группах назначаются на конкретные ЛБ в пределах выбранной ГЛБ. В данной работе представлен алгоритм детального размещения, в основе которого лежит метрика качества размещения, учитывающая архитектуру связей внутри ГЛБ.

Последовательный алгоритм детального размещения. В данном разделе описывается последовательный алгоритм, с которым выполнено сравнение разработанного алгоритма. Последовательный алгоритм не имеет каких-либо принципиальных особенностей, но полезен в случаях, когда архитектура соединений ПЛИС предусматривает наличие прямых последовательных связей между соседними ЛБ в обход коммутатора. Его ключевая идея, изложенная в работе [15], заключается в размещении наиболее длинных цепочек элементов, чтобы максимально задействовать прямые последовательные связи между соседними ЛБ. Однако, в рассматриваемой ПЛИС такие связи отсутствуют, поэтому его можно рассматривать как базовый алгоритм без учета структуры межсоединений.

Последовательный алгоритм детального размещения состоит из трех этапов: инициализации, поиска наибольшей последовательности связанных элементов и размещения найденной последовательности. При этом последние два этапа повторяются до тех пор, пока не будут размещены все ЛЭ.

На первом этапе работы алгоритма группа логических элементов, полученная после кластеризации схемы, представляется в виде ориентированного графа $G = (V, E)$, где V – множество вершин, представляющих ЛЭ, E – множество ребер, представляющих цепи между ЛЭ. Затем из полученного графа удаляются все ребра, инцидентные лишь одной вершине и представляющие цепи, проходящие через границы ГЛБ, и создаются списки *Path* и *MaxPath*, в которых будут храниться обнаруженные последовательности вершин. Для поиска наибольшей взаимосвязанной последовательности ЛЭ используется алгоритм Дейкстры, описанный в работе [16], который позволяет обнаружить наикратчайший путь от вершины-источника до каждой из вершин графа. Затем выполняется восстановление пути от самой дальней вершины до вершины-источника, после чего найденный путь сохраняется в списке *MaxPath*. Все вершины из списка *MaxPath* размещаются в ГЛБ и удаляются из графа. Поиск и размещение последовательностей ЛЭ повторяется до тех пор, пока $V \neq \emptyset$. Псевдокод данного алгоритма представлен на рис. 1.

Стоит отметить, что последовательный алгоритм, хоть и позволяет упростить будущую трассировку, но не лишен недостатков. Удаление из графа наибольшей последовательности вершин приводит к потенциальной блокировке нескольких

меньших последовательностей, размещать которые было бы выгоднее, а также к сложности внедрения в него ограничений, которые должны учитываться в ходе работы.

```

Создать  $G$ ,  $Path$  и  $MaxPath$ 
Пока  $|V| > 0$ 
     $maxLen = -1$ 
    Для  $\forall v \in V$ 
         $length = Dijkstra(G, i, Path)$ 
        Если  $length > maxLen$ 
             $maxLen = length$ 
             $MaxPath = Path$ 
    Для  $\forall v \in MaxPath$ 
        Разместить  $v$ 
        Удалить  $v$  из  $V$ 
    Очистить  $Path$  и  $MaxPath$ 
Удалить  $G$ ,  $Path$  и  $MaxPath$ 
    
```

Рис. 1. Псевдокод последовательного алгоритма детального размещения

Представление и учет архитектуры ПЛИС. Для того, чтобы представить архитектуру внутренних связей ГЛБ, удобно использовать матрицу смежности. Каждый элемент такой матрицы обозначает наличие (1) или отсутствие (0) связи между выходом i -го ЛБ и входом j -го ЛБ. В рассматриваемом случае у каждого ЛБ существует четыре входа, но доступны они попарно (первый и третий или второй и четвертый), поэтому для описания всех связей достаточно двух матриц: A – для второго и четвертого портов, и B – для первого и третьего.

Кроме того, каждая из матриц содержит дополнительный столбец, содержащий суммы элементов каждой строки S_i^a и S_i^b , представляющие собой потенциально доступное количество входных терминалов каждого типа для i -го ЛБ:

$$A = \begin{pmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,15} & S_0^a \\ a_{1,0} & a_{1,1} & \dots & a_{1,15} & S_1^a \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{15,0} & a_{15,1} & \dots & a_{15,15} & S_{15}^a \end{pmatrix} \quad B = \begin{pmatrix} b_{0,0} & b_{0,1} & \dots & b_{0,15} & S_0^b \\ b_{1,0} & b_{1,1} & \dots & b_{1,15} & S_1^b \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{15,0} & b_{15,1} & \dots & b_{15,15} & S_{15}^b \end{pmatrix}.$$

Метрика качества размещения. Для того, чтобы учесть архитектуру связей между ЛБ, необходимо оценить то, насколько будущая позиция каждого ЛЭ соответствует установленным ограничениям. В связи с этим вводится целевая функция $Cost$:

$$Cost^i = C_{place}^i + \alpha \cdot C_{out}^i,$$

где α – корректирующий коэффициент; C_{place}^i , C_{out}^i — составляющие целевой функции для i -го ЛЭ. C_{place}^i служит для учета будущего расположения элементов и вычисляется по следующей формуле:

$$C_{place}^i = C_{place}^a + C_{place}^b,$$

$$C_{place}^a = \begin{cases} F_i^a - S_i^a, & F_i^a > S_i^a \\ 0, & \text{иначе} \end{cases}, \quad C_{place}^b = \begin{cases} F_i^b - S_i^b, & F_i^b > S_i^b \\ 0, & \text{иначе} \end{cases},$$

где F_i^a – необходимое число вторых и четвёртых портов для подключения выхода i -го ЛЭ; F_i^b – необходимое число первых и третьих портов для подключения выхода i -го ЛЭ.

Вторая составляющая C_{out}^i служит для учета расположения элементов в процессе работы алгоритма:

$$C_{out}^i = C_{out}^a + C_{out}^b,$$

$$C_{out}^a = \sum_j \overline{a_{i,j}}, \quad C_{out}^b = \sum_j \overline{b_{i,j}},$$

где j – индексы уже размещенных ЛЭ, к которым подключены выходы i -го ЛЭ.

Алгоритм детального размещения с учетом связей между ЛБ. Суть разработанного алгоритма детального размещения заключается в предварительном определении наиболее оптимальных позиций всех ЛЭ с учетом связей между блоками с последующим их размещением на данные позиции. Определение оптимальных позиций происходит путем их перебора с пересчетом $Cost$ для каждого ЛЭ. Если находится размещение с $Cost = 0$, то оно считается оптимальным и поиск заканчивается. Если же подобное размещение не удастся найти, то выбирается такое размещение, значение $Cost$ для которого минимально. Разработанный алгоритм состоит из четырех этапов: инициализации, определения порядка обхода элементов, определения оптимальных позиций для всех элементов и размещения элементов на оптимальные позиции.

Данный алгоритм основывается на том же графовом представлении схемы, что и последовательный. Однако в разработанном алгоритме не создаются списки для хранения цепочек элементов, а используются массивы P и $tempP$, выступающие в роли модели ГЛБ, где индекс элементов соответствует индексу будущей позиции.

После инициализации определяется порядок обхода вершин, для чего используется модифицированный алгоритм поиска в ширину. На первом шаге из числа ранее не посещенных вершин выбирается начальная вершина s с максимальным числом исходящих ребер. Затем с помощью алгоритма поиска в ширину вычисляется кратчайшее расстояние от начальной вершины до остальных вершин. Если после обхода остаются непосещенные вершины, то весь процесс повторяется, пока каждой вершине в графе не будет задан вес w , равный кратчайшему расстоянию от s .

При дальнейшей работе алгоритма вершины обходятся в порядке роста веса w и уменьшения количества исходящих ребер. В случае равного приоритета рассматриваются два направления обхода: прямое и обратное.

Определение позиций происходит путем их перебора с пересчетом $Cost$ для каждого взаиморасположения ЛЭ. Очередная вершина добавляется в массив, рассчитывается значение $Cost$ для всех его членов и, если суммарное значение $Cost = 0$ или проверены все позиции, то лучшая конфигурация сохраняется, а в массив добавляется следующая вершина.

Сначала проверяются все свободные позиции, и если среди них нет оптимальной, то начинается проверка занятых. В таком случае новая вершина временно фиксируется на проверяемой позиции, и запускается поиск уже для старой вершины, начиная со свободных позиций. Такая операция может проводиться многократно, увеличивая вероятность нахождения оптимальной конфигурации. В данной работе использовалась глубина перебора равная трем.

После определения всех позиций, если $Cost = 0$, то выполняется размещение всех ЛЭ, в противном случае результирующая конфигурация сохраняется, а второй и третий этапы повторяются с обратным порядком обхода. Псевдокод алгоритма детального размещения с учетом связей между блоками представлен на рис. 2.

```

Создать  $G$ ,  $P$  и  $tempP$ 
 $flagINV = 0$ 
 $globalMinCost = \infty$ 
 $INV$ :
Пока  $\exists$  непосещенные  $v \in V$ 
     $s = FindMaxVer(G)$ 
     $distanceCounter(G, s, flagINV)$ 
    Пока  $\exists v \in V \setminus tempP$ 
         $minCost = \infty$ 
         $iterVertex = findNextVertex(G, P, flagINV)$ 
        Для  $\forall i \in [0,15]$ 
             $iterCost = checkPos(tempP, iterVertex, i)$ 
            Если  $iterCost = 0$  и  $\forall v \in tempP$ 
                Сохранить  $tempP$ 
                break
            Если  $iterCost < minCost$ 
                Сохранить  $tempP$ 
                 $minCost = iterCost$ 
     $minCost = CalculateCost(tempP)$ 
    Если  $minCost < globalMinCost$ 
         $P = tempP$ 
         $globalMinCost = minCost$ 
    Если  $globalMinCost \neq 0$  и  $flagINV = 0$ 
         $flagINV = 1$ 
        goto  $INV$ 
Для  $\forall v \in P$ 
    Разместить  $v$ 
    
```

Рис. 2. Псевдокод алгоритма детального размещения с учетом связей между блоками.

Результаты тестирования. Представленные алгоритмы были протестированы на наборе тестовых схем ISCAS'85 [17], ISCAS'89[18], схемах VGA [19] и Сру8080 [20]. Сравнение алгоритмов производилось по трем параметрам: количеству использованных шин локальной обратной связи, количеству использованных глобальных шин и времени, затраченному на трассировку.

Из результатов сравнения видно, что количество задействованных в трассировке шин локальной обратной связи возросло в среднем на 30% (табл. 1), использование же глобальных шин сократилось в среднем на 10% (табл. 2). Исходя из этого, можно сделать вывод, что использование алгоритма детального размещения с учетом связей между блоками позволяет значительно сократить используемое количество глобальных трассировочных ресурсов за счет использования шин локальной обратной связи.

Таблица 1

Результаты тестов. Количество использованных ШЛОС

Схема	Количество ЛЭ	Количество использованных шин локальной обратной связи		
		Последовательный	С учетом связей	Δ , %
c880	123	96	145	+51
c6288	919	1016	1232	+21
s15850	1363	1299	1591	+22
VGA	2177	1815	2318	+28
Cpu8080	2550	1881	2575	+37
s38417	4240	3839	4733	+23
Δ avg, %				+30

Таблица 2

Результаты тестов. Количество использованных ГШ

Схема	Количество ЛЭ	Количество использованных глобальных шин		
		Последовательный	С учетом связей	Δ , %
c880	123	345	296	-14
c6288	919	1954	1738	-11
s15850	1363	3533	3241	-8
VGA	2177	6990	6487	-7
Cpu8080	2550	6980	6286	-10
s38417	4240	10401	9507	-9
Δ avg, %				-10

Среднее время трассировки (таблица 3) не изменилось значительно, но имеет место существенный разброс значения параметра на различных схемах. Это стало результатом весомого усложнения трассировки в пределах ГЛБ для схем с высокой связанностью элементов, что, однако, приводит к большему росту эффективности использования трассировочных ресурсов, как видно на примере схемы Cpu 8080 [20].

Таблица 3

Результаты тестов. Время трассировки

Схема	Количество ЛЭ	Время последующей трассировки, сек		
		Последовательный	С учетом связей	Δ , %
c880	123	367	328	-11
c6288	919	787	829	+5
s15850	1363	4032	4380	+9
VGA	2177	9237	8133	-12
Cpu8080	2550	7625	9941	+30
s38417	4240	12975	12017	-7
Δ avg, %				+2

Заключение. В данной работе представлен алгоритм детального размещения на основе метрики качества, учитывающей архитектуру связей между логическими блоками ПЛИС. Разработанный алгоритм состоит из четырех этапов: инициализации и определения порядка обхода группы размещаемых элементов с помощью алгоритма поиска в ширину, определения оптимальных позиций для всех элементов и их размещения. Для вычисления оптимальных позиций элементов внутри ГЛБ была разработана новая метрика, которая, в отличие от классических, таких как длина соединений и задержки, позволяет оценить количество доступных локальных связей между логическими элементами в ГЛБ.

Разработанный алгоритм и метрика оценки могут использоваться при проектировании в базе отечественных ПЛИС. Экспериментальное сравнение разработанного алгоритма детального размещения с последовательным алгоритмом размещения показало, что применение разработанного алгоритма в маршруте проектирования в базе рассматриваемой в работе ПЛИС позволяет сократить в среднем на 10% количество задействованных в трассировке глобальных коммутационных шин при росте количества используемых внутренних трассировочных ресурсов ГЛБ в среднем на 30%.

Финансирование. Работа выполнена при финансовой поддержке Минобрнауки в рамках государственного задания FSMR-2023-0002 (Соглашение №075-03-2023-024) и государственного задания FFNZ-2021-0001 (Соглашение №075-03-2023-247).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Курский В.В., Сунка В.Я., Польшкова Е.В. Программируемые ПЛИС // Машиностроение: республиканский межведомственный сборник научных трудов: по материалам Международной научно-технической конференции «Материалы, оборудование и ресурсосберегающие технологии в машиностроении», 06-09 апреля 2010 года: в 2 т. / под ред. Б.М. Хрусталева. – Минск: БНТУ, 2012. – Вып. 26, Т. 2. – С. 255-259.
2. Farooq U., Marrakchi Z., Mehrez H. Tree-based heterogeneous FPGA architectures: application specific exploration and optimization. – Springer Science & Business Media, 2012. – P. 186.
3. Singh A., Parthasarathy G., Marek-Sadowska M. Interconnect resource-aware placement for hierarchical FPGAs // IEEE/ACM International Conference on Computer Aided Design. ICCAD 2001. IEEE/ACM Digest of Technical Papers (Cat. No. 01CH37281). – IEEE, 2001. – P. 132-136.
4. Boutros A., Betz V. FPGA architecture: Principles and progression // IEEE Circuits and Systems Magazine. – 2021. – Vol. 21, No. 2. – P. 4-29.
5. ПЛИС емкостью 21 502 вентилях 5510TC068. – АО «Микрон». – URL: <https://mikron.ru/products/high-rel-ic/programmiruemaya-logika-fpga/fpga/product/5510ts068/> (дата обращения: 29.01.2022).
6. Petelin O., Betz V. Wotan: evaluating FPGA architecture routability without benchmarks // ACM Trans. Reconfigurable Technol. Syst. – 2018. – Vol. 11, No. 2. – P. 1-23.
7. Фролова П.И., Чочаев Р.Ж., Иванова Г.А., Гаврилов С.В. Алгоритм размещения с оптимизацией быстродействия на основе матриц задержек для реконфигурируемых систем на кристалле // Проблемы разработки перспективных микро- и наноэлектронных систем (МЭС). – 2020. – Вып. 1. – С. 2-7.
8. Фролова П.И., Хватов В.М., Чочаев Р.Ж. Сравнительный анализ методов кластеризации и размещения схем для реконфигурируемых систем на кристалле // Проблемы разработки перспективных микро- и наноэлектронных систем (МЭС). – 2022. – Вып. 4. – С. 63-70.
9. Kureichik V.V., Zaporozhets D.Y., Zaruba D.V. Partitioning of VLSI fragments based on the model of glowworm's behavior // Proceedings of the 19th International Conference on Soft Computing and Measurements, SCM 2016. – 2016. – P. 268-272.
10. Li W. et al. UTPlaceF 3.0: A parallelization framework for modern FPGA global placement // 2017 IEEE/ACM International Conference on Computer-Aided Design (ICCAD). – IEEE, 2017. – P. 922-928.

11. *Chen G., et al.* RippleFPGA: Routability-driven simultaneous packing and placement for modern FPGAs // *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*. – 2017. – Vol. 37, No. 10. – P. 2022-2035.
12. *Dhar S., et al.* Detailed placement for modern FPGAs using 2D dynamic programming // 2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD). – IEEE, 2016. – P. 1-8.
13. *Nikolić S., Zgheib G., Jenne P.* Detailed Placement for Dedicated LUT-Level FPGA Interconnect // *ACM Transactions on Reconfigurable Technology and Systems*. – 2022. – Vol. 15, No. 4. – P. 1-33.
14. *Гаврилов С.В., Железников Д.А., Чочаев Р.Ж.* Разработка и сравнительный анализ методов решения задачи размещения для реконфигурируемых систем на кристалле // *Известия высших учебных заведений. Электроника*. – 2020. – Т. 25, № 1. – С. 48-57.
15. *Fan D.K., Shi P.* Improvement of Dijkstra's algorithm and its application in route planning // 2010 seventh international conference on fuzzy systems and knowledge discovery. – IEEE, 2010. – Vol. 4. – P. 1901-1904.
16. *Dijkstra E.W.* A note on two problems in connexion with graphs // *Edsger Wybe Dijkstra: His Life, Work, and Legacy*. – 2022. – P. 287-290.
17. *Bryan D.* The ISCAS'85 benchmark circuits and netlist format // *North Carolina State University*. – 1985. – Vol. 25. – P. 39.
18. *Brglez F., Bryan D., Kozminski K.* Combinational profiles of sequential benchmark circuits // 1989 IEEE International Symposium on Circuits and Systems (ISCAS). – IEEE, 1989. – P. 1929-1934.
19. 8080 Compatible CPU. – URL: <https://opencores.org/projects/cpu8080>
20. The OpenCores VGA/LCD Controller. – URL: https://opencores.org/projects/vga_lcd.

REFERENCES

1. *Kurskiy V.V., Sunka V.Ya., Polynkova E.V.* Programmiruemye PLIS [Programmable FPGAs], *Mashinostroenie: respublikanskiy mezhdunarodnyy sbornik nauchnykh trudov: po materialam Mezhdunarodnoy nauchno-tekhnicheskoy konferentsii «Materialy, oborudovanie i resursosberegayushchie tekhnologii v mashinostroenii», 06-09 aprelya 2010 goda* [Mechanical engineering: republican interdepartmental collection of scientific papers: based on the materials of the international scientific and technical conference “Materials, equipment and resource-saving technologies in mechanical engineering”, April 06-09, 2010]: In 2 vol., ed. by B.M. Khrustaleva. Minsk: BNTU, 2012, Issue 26, Vol. 2, pp. 255-259.
2. *Farooq U., Marrakchi Z., Mehrez H.* Tree-based heterogeneous FPGA architectures: application specific exploration and optimization. Springer Science & Business Media, 2012, pp. 186.
3. *Singh A., Parthasarathy G., Marek-Sadowska M.* Interconnect resource-aware placement for hierarchical FPGAs, *IEEE/ACM International Conference on Computer Aided Design. ICCAD 2001. IEEE/ACM Digest of Technical Papers (Cat. No. 01CH37281)*. IEEE, 2001, pp. 132-136.
4. *Boutros A., Betz V.* FPGA architecture: Principles and progression, *IEEE Circuits and Systems Magazine*, 2021, Vol. 21, No. 2, pp. 4-29.
5. PLIS emkost'yu 21 502 ventiley 5510TS068 [FPGA with a capacity of 21,502 gates 5510TS068]. JSC Mikron. Available at: <https://mikron.ru/products/high-relic-programmiruemaya-logika-fpga/fpga/product/5510ts068/> (accessed 29 January 2022).
6. *Petelin O., Betz V.* Wotan: evaluating FPGA architecture routability without benchmarks, *ACM Trans. Reconfigurable Technol. Syst.*, 2018, Vol. 11, No. 2, pp. 1-23.
7. *Frolova P.I., Chochaev R.Zh., Ivanova G.A., Gavrilov S.V.* Algoritm razmeshcheniya s optimizatsiey bystrodeystviya na osnove matrits zaderzhek dlya rekonfiguriruemyykh sistem na kristalle [Placement algorithm with performance optimization based on delay matrices for reconfigurable systems on a chip], *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem (MES)* [Problems in the development of advanced micro- and nanoelectronic systems (MES)], 2020, Issue 1, pp. 2-7.
8. *Frolova P.I., Khvatov V.M., Chochaev R.Zh.* Sravnitel'nyy analiz metodov klasterizatsii i razmeshcheniya skhem dlya rekonfiguriruemyykh sistem na kristalle [Comparative analysis of clustering methods and circuit placement for reconfigurable systems on a chip], *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem (MES)* [Problems of development of advanced micro- and nanoelectronic systems (MES)], 2022, Issue 4, pp. 63-70.

9. Kureichik V.V., Zaporozhets D.Y., Zaruba D.V. Partitioning of VLSI fragments based on the model of glowworm's behavior, *Proceedings of the 19th International Conference on Soft Computing and Measurements, SCM 2016*, 2016, pp. 268-272.
10. Li W. et al. UTPlaceF 3.0: A parallelization framework for modern FPGA global placement, *2017 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. IEEE, 2017, pp. 922-928.
11. Chen G., et al. RippleFPGA: Routability-driven simultaneous packing and placement for modern FPGAs, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2017, Vol. 37, No. 10, pp. 2022-2035.
12. Dhar S., et al. Detailed placement for modern FPGAs using 2D dynamic programming, *2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. IEEE, 2016, pp. 1-8.
13. Nikolić S., Zgheib G., Jenne P. Detailed Placement for Dedicated LUT-Level FPGA Interconnect, *ACM Transactions on Reconfigurable Technology and Systems*, 2022, Vol. 15, No. 4, pp. 1-33.
14. Gavrilov S.V., Zheleznikov D.A., Chochev R.Zh. Razrabotka i sravnitel'nyy analiz metodov resheniya zadachi razmeshcheniya dlya rekonfiguriruemyykh sistem na kristalle [Development and comparative analysis of methods for solving the placement problem for reconfigurable systems on a chip], *Izvestiya vysshikh uchebnykh zavedeniy. Elektronika* [News of higher educational institutions. Electronics], 2020, Vol. 25, No. 1, pp. 48-57.
15. Fan D.K., Shi P. Improvement of Dijkstra's algorithm and its application in route planning, *2010 seventh international conference on fuzzy systems and knowledge discovery*. IEEE. 2010, Vol. 4, pp. 1901-1904.
16. Dijkstra E.W. A note on two problems in connexion with graphs, *Edsger Wybe Dijkstra: His Life, Work, and Legacy*, 2022, pp. 287-290.
17. Bryan D. The ISCAS'85 benchmark circuits and netlist format, *North Carolina State University*, 1985, Vol. 25, pp. 39.
18. Brglez F., Bryan D., Kozminski K. Combinational profiles of sequential benchmark circuits, *1989 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 1989, pp. 1929-1934.
19. 8080 Compatible CPU. Available at: <https://opencores.org/projects/cpu8080>.
20. The OpenCores VGA/LCD Controller. Available at: https://opencores.org/projects/vga_lcd.

Статью рекомендовал к опубликованию д.т.н., профессор В.В. Курейчик.

Шокарев Дмитрий Борисович – Институт проблем проектирования в микроэлектронике РАН; e-mail: shokarev_d@ippm.ru; Москва, Зеленоград, Россия; тел.: +74997299890; отдел систем автоматизированного проектирования интегральных схем; инженер-исследователь.

Чочаев Рустам Жамболатович – e-mail: chochaev_r@ippm.ru; тел.: +74997299890; отдел систем автоматизированного проектирования интегральных схем; инженер-исследователь.

Щелоков Альберт Николаевич – e-mail: schan@ippm.ru; тел.: +74997299845; отдел проектирования микроэлектронных компонентов для нанотехнологий; к.ф.-м.н.; доцент; помощник директора.

Гаврилов Сергей Витальевич – e-mail: sergey_g@ippm.ru; тел.: +74997299845; д.т.н.; директор; профессор Института интегральной электроники имени академика К.А. Валиева Национального исследовательского университета «МИЭТ».

Shokarev Dmitry Borisovich – The Institute for Design Problems in Microelectronics; e-mail: shokarev_d@ippm.ru; Moscow, Zelenograd, Russia; phone: +74997299890; the department of computer-aided design of integrated circuits; research engineer.

Chochev Rustam Zhambolatovich – e-mail: chochaev_r@ippm.ru; phone: +74997299890; the department of computer-aided design of integrated circuits; research engineer.

Schelokov Albert Nikolaevich – e-mail: schan@ippm.ru; phone: +74997299845; the department of design of microelectronic components for nanotechnology; cand. of phys. and math. sc.; associate professor; director's assistant.

Gavrilov Sergey Vitalievich – e-mail: sergey_g@ippm.ru; phone: +74997299845; dr. of eng. sc.; director; professor at the Institute of Integrated Electronics, National Research University of Electronic Technology (MIET).

В.В. Курейчик, В.В. Бова, А.Ю. Халенков

МНОГОУРОВНЕВЫЙ ПОДХОД ДЛЯ РЕШЕНИЯ ЗАДАЧИ ДВУМЕРНОЙ УПАКОВКИ ГЕОМЕТРИЧЕСКИХ ФИГУР СЛОЖНЫХ ФОРМ

Рассмотрена одна из важных комбинаторных задач оптимизации – задача двумерной упаковки геометрических фигур сложных форм. Она относится к классу NP- сложных и трудных оптимизационных задач. В работе приведена и описана постановка задачи двумерной упаковки, введена комбинированная целевая функция, учитывающая все ограничения. В связи со сложностью данной задачи предлагается многоуровневый подход, заключающийся в разделении задачи двумерной упаковки на 4 подзадачи и решения каждой подзадачи последовательно 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 методах генетического поиска и биоинспирированной оптимизации. 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 двумерной упаковки на основе комбинированного биоинспирированного алгоритма. Проведен вычислительный эксперимент на тестовых примерах (бенчмарках). Качество упаковки, полученное, на основе разработанного комбинированного биоинспирированного алгоритма, в среднем на 2 % превосходит результаты упаковки, полученные с использованием известных алгоритмов при сопоставимом времени решения, 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 - $O(n^3)$.

Двухмерная упаковка; упаковка в контейнеры; многоуровневый подход; комбинированный биоинспирированный алгоритм; генетический алгоритм.

V.V. Kureichik, V.V. Bova, A.Yu. Khalenkov

MULTILEVEL APPROACH TO TWO-DIMENSIONAL PACKING PROBLEM FOR GEOMETRIC FIGURES OF COMPLEX SHAPES

The paper considers one of the important combinatorial optimization problems, namely the two-dimensional packing problem for geometric figures of complex shapes. It belongs to the class of NP-complex and difficult optimization problems. In this paper, the formulation of the two-dimensional packing problem is given and described, and a combinatorial objective function that takes into account all constraints is introduced. Due to the complexity of this problem, a multilevel approach is proposed, which consists in dividing the two-dimensional packing problem into 4 subproblems and solving each subproblem sequentially in a strict order. At the same time, for each of the subtasks a unique set of objects that are not repeated in the other subtasks is defined. To implement the multilevel approach, the authors developed a combined bioinspired algorithm based on the methods of genetic search and bioinspired optimization. This approach allows to significantly reduce the time of obtaining the result, partially solve the problem of preliminary convergence of algorithms and obtain sets of quasi-optimal solutions in polynomial time. A software package has been developed and algorithms for automated two-dimensional packing based on the combined bioinspired algorithm have been implemented. A computational experiment on test cases (benchmarks) has been carried out. The packing quality obtained on the basis of the developed combined bioinspired algorithm, on average, by 2% exceeds the packing results obtained using known algorithms at comparable solution time, which indicates the effectiveness of

the proposed approach. The conducted series of tests and experiments allowed us to refine the theoretical estimates of the time complexity of the packing algorithms. In the best case the time complexity of the algorithms is $O(n^2)$, in the worst case - $O(n^3)$.

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

Введение. На сегодняшний день проблема эффективного распределения ресурсов является актуальной и важной задачей. В настоящее время ее качественное решение может быть обеспечено за счет применения систем автоматизированного проектирования. При чем производительность этих систем значительно повышается при включении в их состав оптимизационных методов.

Множество различных оптимизационных задач распределения ресурсов со схожими начальными условиями такие, как раскрой материала, контейнерная упаковка, календарное планирование, компоновка объектов сегодня объединяются в задачи «упаковки». Для решения данной задачи, многие компании внедряют системы автоматизированного проектирования (САПР). Снижение материальных затрат, увеличение качества проектирования, снижение сроков проектирования, являются основными целями автоматизации [1].

Целью данной задачи является размещение объектов в блоки заданной вместимости, таким образом, чтобы максимально их заполнить, при этом количество этих блоков должно быть минимальным.

Упаковка представляет собой сложную комбинаторно-логическую задачу, относящуюся к классу NP-полных задач [2].

Вследствие этого одним из наиболее перспективных направлений исследований является разработка и улучшение различных приближенных, а также эвристических методов решения задач упаковки. В связи с этим, предлагается для решения задач упаковки использовать комбинированные методы, основанные на генетической и биоинспирированной оптимизации [3–13].

Постановка задачи двумерной упаковки. Задача двумерной упаковки – представляет из себя задачу по упаковке двумерных объектов в двумерный контейнер, который может быть как ограниченный, так и неограниченный. Как например упаковка прямоугольников в квадрат или упаковка в полуограниченную полосу.

Основой задачи можно считать двумерную область (ограниченную или нет) и набор двумерных объектов, форма и разнородность или однородность которых задается по условию. Таким образом, объекты необходимо разместить внутри контейнера оптимальным образом, не нарушая заданные условия. Тогда Задача двумерной упаковки формулируется как:

♦ Пусть дана область контейнер и задано некое конечное множество объектов двумерной формы. Эти объекты образуются при помощи точек, образуя замкнутые области, ограниченные замкнутой кривой [14]. При этом определенное количество геометрических фигур, заданных по условию конкретной задачи, с заданными по условию формами необходимо уложить их на поверхность контейнера, ограниченного по обоим осям (x и y), образующего замкнутый прямоугольник. Расположить заданные фигуры необходимо таким образом, чтобы фигуры были достаточно плотно уплотнены между собой [14, 15]. При этом итоговое размещение объектов должно учитывать следующие ограничения:

♦ Общая площадь фигур не должна быть больше площади самого контейнера:

$$\sum_{i=1}^n S_i \leq S_{\text{общ}}, \quad (1)$$

где S_i – площадь i -того объекта расположенного на поверхности поля (суммарная площадь занятой части), $S_{\text{общ}}$ – общая площадь поля;

- ◆ элементы ϕ не должны накладываться друг на друга:

$$f_{ij}(x_i, y_i, \phi_i, x_j, y_j, \phi_j, \dots, x_n, y_n, \phi_n) \geq 0 \quad i \neq j; \quad i, j = 1, 2, \dots, n; \quad (2)$$

- ◆ и не должны выходить за границы поля:

$$\begin{cases} x_{\min i} \geq 0; \\ y_{\min i} \geq 0; \\ x_{\min i} \leq m_x; \\ y_{\min i} \leq m_y. \end{cases} \quad (3)$$

В качестве целевой функции F чаще всего используют функцию k , значение которой равно коэффициенту использования материала [14, 15]. Так как потери материала должны быть минимальными, соответственно ЦФ необходимо максимизировать. При этом коэффициент k определяется согласно следующей формуле:

$$k = \frac{\sum_{i=1}^n S_i}{S_{\text{общ}}}, \quad (4)$$

где S_i – площадь i -того объекта расположенного на поверхности поля (суммарная площадь занятой части), $S_{\text{общ}}$ – общая площадь поля.

Сформулированная задача относится к классу NP-сложных задач. В связи с этим для ее решения предлагается использовать эвристические методы, инспирированные природными системами [3–13].

Многоуровневый подход для решения задачи двумерной упаковки геометрических фигур сложных форм. 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 [3–13]. Второй подход – это упрощение решаемых задач за счет уменьшения их размерности или их декомпозиции [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 двумерной упаковки геометрических фигур сложных форм на этапы (подзадачи) и решения каждой подзадачи в строгом порядке рис. 1 [11, 17, 18].

Такой подход позволяет значительно сократить время работы алгоритмов на каждом этапе и в итоге для всей задачи. Рассмотрим более подробно реализацию каждого этапа.



Рис. 1. Этапы решения задачи

На первом этапе выполняется симплификация – временное упрощение исходной формы. Здесь фигуры сложных форм обрамляются в более простые геометрические фигуры, как например это делают при паллетировании. Таким образом, у каждой фигуры будет свой абстрактный паллет, более простой геометрической формы, а именно прямоугольник, как показано на рис. 2.

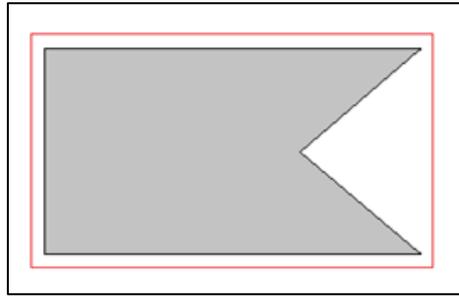


Рис. 2. Пример обрамления сложной геометрической формы в простую

Заметим, что упаковку таких объектов можно проводить быстро последовательными или итерационными алгоритмами. Опишем словесный алгоритм этого этапа. Сначала инициализируются исходные объекты, получая данные об ограничениях задачи, о размере контейнера, о количестве объектов, об их геометрических параметрах, их площади. Далее создается паллет, размеры которого изначально задаются равными размерам исходного контейнера и производится проверка вписывается ли выбранная фигура в паллет и касаются ли стороны паллета фигуры. Если да, то вычисляется площадь занятого пространства и площадь свободного пространства внутри паллеты, так как потенциал незанятого пространства в паллете нельзя игнорировать. Далее от площади паллеты отнимается площадь фигуры, подсчитанная заранее и затем в алгоритме динамически уменьшается размер обрамляющего фигуру контейнера до тех пор, пока не будет найден оптимальный размер обрамляющего паллета. Заметим, что между фигурой и обрамляющим паллетом могут оставаться небольшие зазоры, однако, они имеют условное значение и на последующих этапах будут устранены.

На втором этапе производится непосредственно упаковка полученных паллетов, т.е. задача упрощается до решения классической задачи двумерной упаковки разнородных прямоугольных форм. Поэтому на данном этапе для эффективного решения задачи предлагается использовать методы генетического поиска [3]. Данные методы позволяют достаточно быстро отбросить большое количество максимально неоптимальных решений. Окончанием поиска является условие, чем больше площадь пересечений свободного пространства подконтейнеров, тем лучшее решение получено, однако при этом необходимо учитывать ограничение, что сами фигуры не могут пересекаться. Итогом выполнения второго этапа станет получение набора квазиоптимальных решений задачи упаковки рис. 3.

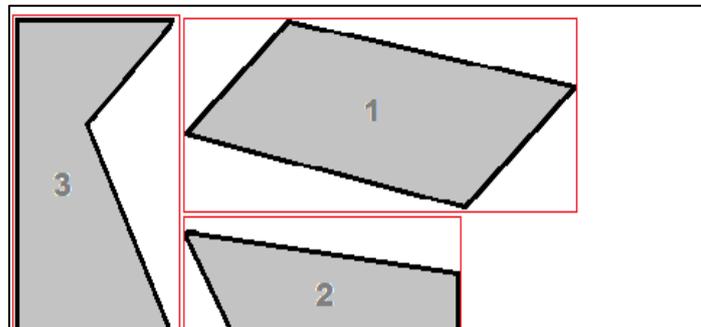


Рис. 3. Результат работы второго этапа

После выполнения второго этапа был получен результат, в котором паллеты упакованы оптимально, однако сами фигуры, как можно видеть, на рис. 3, имеют между собой достаточно большое количество свободного пространства внутри этих паллетов. В связи с этим необходимо провести концентрацию (компоновку) фигур для оптимизации свободного пространства [13, 15]. На данном этапе эффективно применяются как методы генетического поиска, так и биоинспирированной оптимизации [3, 4]. При этом алгоритм должен сдвигать фигуры таким образом, чтобы достичь как можно большего количества пересечений паллетов их обрамляющих, т.е. максимизировать площадь пересечения паллетов, без пересечений фигур друг с другом (рис. 4).

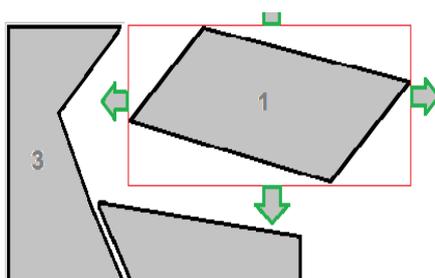


Рис. 4. Этап компоновки фигур

После завершения этапа компоновки фигур происходит переход на четвертый этап рекурсивной оптимизации и деоптимизации. Здесь наиболее близкие фигуры, попарно объединятся в один новый подконтейнер рис. 5.

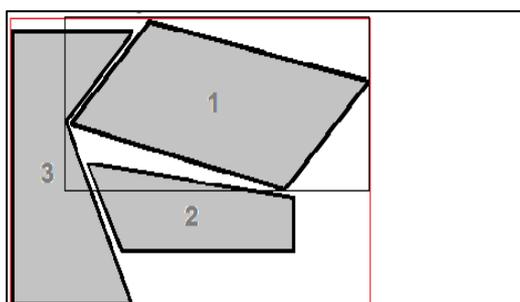


Рис. 5. Этап рекурсии и деоптимизации

После создания новых паллетов (подконтейнеров) реализуется деоптимизация [17, 18]. Здесь повторно производится упаковка новых подконтейнеров на основе проблемно-ориентированного генетического алгоритма для максимизации свободного пространства [19]. Данный шаг будет итерационно повторяться до тех пор, пока решение не перестанет улучшаться. Итоговая упаковка объекта приведена на рис. 6.

Отметим, что данный подход позволяет минимизировать потери абстрактного сырья и не затрачивать лишнее время на перебор и фильтрацию множества комбинаций фигур между собой.

Применение предложенного подхода 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.

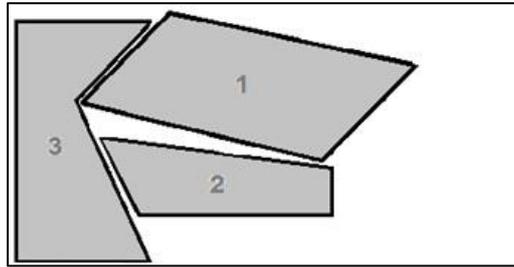


Рис. 6. Итоговая упаковка объектов

Вычислительный эксперимент. Разработана программная среда, на основе предложенного многоуровневого подхода, для решения задачи двумерной упаковки геометрических фигур сложных форм. При построении комплекса программ использовалась среда программирования Microsoft Visual Studio на языке C#. Отладка и тестирование разработанных алгоритмов выполнялось на компьютере типа IBM PC с процессором Intel Core i5 3600x с ОЗУ-16Гб. Проведен вычислительный эксперимент. Для определения эффективности разработанного подхода были проведены исследования качества решения для разного набора тестовых примеров (бенчмарк OR-Library), различающихся количеством блоков [20]. Результаты проведенных исследований представлены в табл. 1, а также на рис. 7.

Таблица 3

Сравнение разработанного БА с аналогами

Количество блоков	Ngoi et al.	Gehring и др.	Разработанный БА
шт.	%	%	%
100	61,5	63,1	62,5
200	68,2	68,5	68,8
300	74,3	75,5	77,6
400	80,6	81,4	82,5
500	88,4	90,2	92,8
Среднее значение	74,6	75,74	76,84

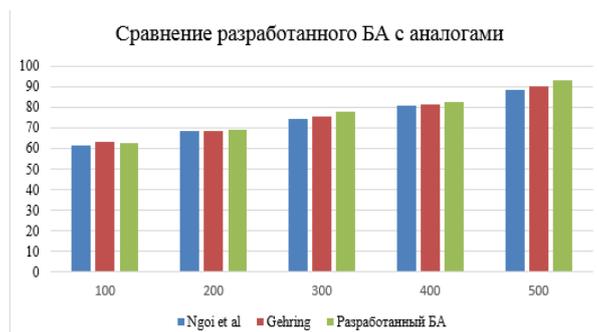


Рис. 7. Сравнение разработанного БА с аналогами

В результате анализа представленной таблицы и графика зависимостей можно сделать вывод, что разработанный многоуровневый подход на основе биоинспирированной оптимизации позволяет лучше производить упаковку блоков в

среднем на 2% использованного пространства по всем тестам, чем известные алгоритмы, при сопоставимом времени решения, что говорит об эффективности предложенного подхода.

Заключение. Предложен многоуровневый подход для решения задачи двумерной упаковки геометрических фигур сложных форм. Отличительной особенностью данного подхода является разбиение задачи на несколько подзадач (этапов), с последовательным решением каждой задачи. 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 двумерной упаковки геометрических фигур сложных форм, по сравнению с известными методами.

Качество упаковки, полученное, на основе разработанного комбинированного биоинспирированного алгоритма, в среднем на 2 % превосходит результаты упаковки, полученные с использованием известных алгоритмов при сопоставимом времени решения, что говорит об эффективности предложенного подхода.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Kacprzyk J., Kureichik V.M., Malioukov S.P., Kureichik V.V., Malioukov A.S* General questions of automated design and engineering // *Studies in Computational Intelligence*. – 2009. – Vol. 212. – P. 1-22.
2. *Кормен Т., Лейзерсон И., Ривест Р.* Алгоритмы. Построение и анализ. – М.: МЦМО, 2000.
3. *Гладков Л.А., Зинченко Л.А., Курейчик В.В., Курейчик В.М., Лебедев Б.К., Нужнов Е.В., Сорокин С.Н.* Методы генетического поиска. – Таганрог, 2002.
4. *Курейчик В.В., Родзин С.И.* Вычислительные модели эволюционных и роевых биоэвристик (Обзор) // *Информационные технологии*. – 2021. – Т. 27, № 10. – С. 507-520.
5. *Тимофеева О.П., Соколова Э.С., Милов К.В.* Генетический алгоритм в оптимизации упаковки контейнеров // *Тр. НГТУ им. П.Е. Алексеева*. – 2012. – № 4 (101). – С. 167-172.
6. *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.
7. *Kureichik V., Kureichik L., Zaruba D.* Bioinspired algorithm for 2D packing problem // *Advances in Intelligent Systems and Computing*. – 2019. – Vol. 764. – P. 39-46.
8. *Li M., Song C., Zhou Z.* Hybrid particle swarm optimization for two-dimensional irregular parts packing // *Journal of Sichuan University (Engineering Science Edition)*. – 2005. – Vol. 37, Iss. 4. – P. 134-138.
9. *Shin Y., Kita E.* Solving two-dimensional packing problem using particle swarm optimization // *Computer Assisted Mechanics and Engineering Science*. – 2012. – Vol. 19, Iss. 3. – P. 241-255.
10. *Zhang D., Dong R., Si Y.* Hybrid swarm algorithm based on ABC and AIS for 2L-HFCVRP // *Q.a View Correspondence*. – 2017. – Vol. 61. – P. 726.
11. *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.
12. *Soukaina Laabadi, Naimi M., Amri H.E., Achchab B.* A Binary Crow Search Algorithm for Solving Two-dimensional Bin Packing Problem with Fixed Orientation // *Procedia Computer Science*. – 2020. – Vol. 167. – P. 809-818.
13. *Валеева А.Ф.* Применение конструктивных эвристик в задачах раскроя-упаковки // *Информационные технологии*. – 2006. – № S11. – С. 1-24.
14. *Мухачева Э.А., Верхотуров М.А., Мартынов В.В.* Модели и методы расчета раскроя – упаковки геометрических объектов. – Уфа: УГАТУ, 1998. – 216 с.
15. *Фроловский В.Д.* Оптимальное группирование геометрических объектов при проектировании карт раскроя материалов // *Программные продукты и системы*. – 2000. – № 3. – С. 47-48.

16. *Базилевич Р.П.* Декомпозиционные и топологические методы автоматизированного конструирования электронных устройств: монография. – Львов: Вицшашкола, 1981. – 81 с.
17. *Мухачева Э.А., Мухачева А.С.* Технология блочных структур локального поиска оптимума в задачах прямоугольной упаковки // Новые технологии. Информационные технологии. Приложение. – 2004. – № 5. – С. 19-31.
18. *Fadel G. Sinha G., McKee A.* Packing optimization using a rubberband analogy // Design Engineering Technical Conference and Computers and Information in Engineering Conference, Pittsburgh, PA (ASME). – 2001. – Vol. 2. – P. 409-415.
19. *Бова В.В., Курейчик В.В., Лежебоков А.А.* Проблемно ориентированный генетический алгоритм упаковки разногабаритных элементов // Вестник Ростовского государственного университета путей сообщения. – 2014. – № 3 (55). – С. 52-59.
20. *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. *Kacprzyk J., Kureichik V.M., Malioukov S.P., Kureichik V.V., Malioukov A.S* General questions of automated design and engineering, *Studies in Computational Intelligence*, 2009, Vol. 212, pp. 1-22.
2. *Kormen T., Leyzerson I., Rivest R.* Алгоритмы. Построение и анализ [Algorithms. Construction and analysis]. Moscow: MTSMO, 2000.
3. *Gladkov L.A., Zinchenko L.A., Kureychik V.V., Kureychik V.M., Lebedev B.K., Nuzhnov E.V., Sorokin S.N.* Metody geneticheskogo poiska [Genetic search methods]. Taganrog, 2002.
4. *Kureychik V.V., Rodzin S.I.* Vychislitel'nye modeli evolyutsionnykh i roevykh bioevristik (Obzor) [Computational models of evolutionary and swarm bioheuristics (Review)], *Informatsionnye tekhnologii* [Information technologies], 2021, Vol. 27, No. 10, pp. 507-520.
5. *Timofeeva O.P., Sokolova E.S., Milov K.V.* Geneticheskii algoritim v optimizatsii upakovki konteynerov [Genetic algorithm in optimization of container packaging], *Tr. NGTU im. R.E. Alekseeva* [Proceedings of NSTU im. R.E. Alekseeva], 2012, No. 4 (101), pp. 167-172.
6. *Gehring H., Bortfeldt A.* A genetic algorithm for solving the container loading problem, *International Transactions in Operational Research*, 1997, Vol. 4, Iss. 5-6, pp. 401-418.
7. *Kureichik V., Kureichik L., Zaruba D.* Bioinspired algorithm for 2D packing problem, *Advances in Intelligent Systems and Computing*, 2019, Vol. 764, pp. 39-46.
8. *Li M., Song C., Zhou Z.* Hybrid particle swarm optimization for two-dimensional irregular parts packing, *Journal of Sichuan University (Engineering Science Edition)*, 2005, Vol. 37, Iss. 4, pp. 134-138.
9. *Shin Y., Kita E.* Solving two-dimensional packing problem using particle swarm optimization, *Computer Assisted Mechanics and Engineering Science*, 2012, Vol. 19, Iss. 3, pp. 241-255.
10. *Zhang D., Dong R., Si Y.* Hybrid swarm algorithm based on ABC and AIS for 2L-HFCVRP, *Q.a View Correspondence*, 2017, Vol. 61, pp. 726.
11. *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.
12. *Soukaina Laabadi, Naimi M., Amri H.E., Achchab B.* A Binary Crow Search Algorithm for Solving Two-dimensional Bin Packing Problem with Fixed Orientation, *Procedia Computer Science*, 2020, Vol. 167, pp. 809-818.
13. *Valeeva A.F.* Primenenie konstruktivnykh evristik v zadachakh raskroya-upakovki [Application of constructive heuristics in cutting-packing problems], *Informatsionnye tekhnologii* [Information technologies], 2006, No. S11, pp. 1-24.
14. *Mukhacheva E.A., Verkhoturov M.A., Martynov V.V.* Modeli i metody rascheta raskroya – upakovki geometricheskikh ob"ektov [Models and methods for calculating cutting and packaging of geometric objects]. Ufa: UGATU, 1998, 216 p.
15. *Frolovskiy V.D.* Optimal'noe gruppirovaniye geometricheskikh ob"ektov pri proektirovanii kart raskroya materialov [Optimal grouping of geometric objects when designing cutting charts for materials], *Programmnye produkty i sistemy* [Software products and systems], 2000, No. 3, pp. 47-48.
16. *Bazilevich R.P.* Dekompozitsionnye i topologicheskie metody avtomatizirovannogo konstruirovaniya elektronnykh ustroystv: monografiya [Decomposition and topological methods for automated design of electronic devices: monograph]. L'vov: Vishchashkola, 1981, 81 p.

17. *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], Novye tekhnologii. Informatsionnye tekhnologii. Prilozhenie [New technologies. Information Technology. Application], 2004, No. 5, pp. 19-31.*
18. *Fadel G. Sinha G., McKee A. Packing optimization using a rubberband analogy, Design Engineering Technical Conference and Computers and Information in Engineering Conference, Pittsburgh, PA (ASME), 2001, Vol. 2, pp. 409-415.*
19. *Bova V.V., Kureychik V.V., Lezhebokov A.A. Problemno orientirovannyi geneticheskiy algoritm upakovki raznogabaritnykh elementov [Problem-oriented genetic algorithm for packing multi-sized elements], Vestnik Rostovskogo gosudarstvennogo universiteta putey soobshcheniya [Bulletin of the Rostov State Transport University], 2014, No. 3 (55), pp. 52-59.*
20. *Zhukov L.A., Korchevskaya O.V. Metod ploskostey: chislennyi 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. Лисицина.

Курейчик Владимир Викторович – Южный федеральный университет; e-mail: vkur@sfedu.ru; г. Таганрог, Россия; тел.: 88634383451; кафедра систем автоматизированного проектирования; зав. кафедрой САПР; д.т.н.; профессор.

Бова Виктория Викторовна – e-mail: vbova@yandex.ru; тел.: 88634371651; кафедра систем автоматизированного проектирования; доцент.

Халенков Александр Юрьевич – e-mail: halenkov@sfedu.ru; тел.: 88634371651; кафедра систем автоматизированного проектирования; аспирант.

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

Bova Victoria Victorovna – e-mail: vbova@yandex.ru; phone: +78634371651; the department of computer aided design; associate professor.

Halenkov Alexander Yuryevich – e-mail: halenkov@sfedu.ru; phone: +78634371651; the department of computer aided design; graduate student.

УДК 004.056

DOI 10.18522/2311-3103-2023-5-66-81

А.А. Олейникова, В.В. Золотарев

КОНЦЕПЦИЯ УПРАВЛЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТЬЮ НА ОСНОВЕ ЦИКЛА НЕПРЕРЫВНОГО ДЕТЕКТИРОВАНИЯ И РЕАГИРОВАНИЯ НА ИНЦИДЕНТЫ БЕЗОПАСНОСТИ ИНФОРМАЦИИ

Для динамически изменяющихся объектов управления в задаче управления информационной безопасностью возникают новые задачи, такие как изменение подходов к сбору и анализу данных, разработка динамических сценариев реагирования на угрозы безопасности информации. Они должны быть решены через создание применимых в указанной задаче алгоритмов, моделей, методик и подходов управления безопасностью, в том числе на уровне организации процессов, работы с данными и формирования архитектуры информационной безопасности организации. Кроме того, для разработки и формирования инструментов непрерывного детектирования и реагирования необходимо предложить новые способы интеграции указанных алгоритмов в структуру объекта управления. При этом создание систем реагирования на базе новой концепции предполагает и изменение алгоритмов управления безопасностью таких систем в особых случаях, таких как децентрализованное

управление, тестирование на устойчивость, облачные сервисы безопасности и других, требующих отдельного исследования. При этом реагирование на инциденты информационной безопасности должно предполагать учет непрерывно меняющегося ландшафта угроз и реконфигурации инфраструктуры организации. Также на развитие представленной в статье новой концепции повлияла концепция объектно-ориентированного программирования в части основных положений. Настоящая работа содержит описание концепции управления на основе цикла непрерывного детектирования и реагирования, приводит некоторые алгоритмы и процессы, отличающие реализацию показанной концепции, а также примеры их реализации. Приведенные в статье практические примеры касаются таких вопросов, как формирование окрестности инцидента, и позволяют формировать контекст управления информационной безопасностью. Кроме того, показан подход к автоматизации процессов управления информационной безопасностью. Результаты работы могут быть использованы как для имитационных моделей, так и для реализации в виде набора процессов управления информационной безопасностью в практических задачах. Кроме того, полученные результаты могут быть интегрированы в средства оркестрации для систем защиты информации, что повышает эффективность реагирования на инциденты информационной безопасности.

Управление информационной безопасностью; процессный подход; алгоритм управления безопасностью; управление на основе данных; непрерывное детектирование и реагирование.

A.A. Oleynikova, V.V. Zolotarev

THE CONCEPT OF INFORMATION SECURITY MANAGEMENT BASED ON A CYCLE OF INFORMATION SECURITY INCIDENTS CONTINUOUS DETECTION AND RESPONSE

For dynamically changing management objects, new tasks arise in the task of information security management, such as changing approaches to data collection and analysis, developing dynamic scenarios for responding to information security threats. They should be solved through the creation of algorithms, models, methods and approaches of security management applicable to this task, including at the level of organizing processes, working with data and forming the organization's information security architecture. In addition, for the development and formation of continuous detection and response tools, it is necessary to propose new ways of integrating these algorithms into the structure of the control object. At the same time, the creation of response systems based on the new concept also involves changing the security management algorithms of such systems in special cases, such as decentralized management, stability testing, cloud security services and others that require separate research. At the same time, responding to information security incidents should take into account the continuously changing threat landscape and reconfiguration of the organization's infrastructure. Also, the development of the new concept presented in the article was influenced by the concept of object-oriented programming in terms of the main provisions. This work contains a description of the control concept based on a continuous detection and response cycle, provides some algorithms and processes that distinguish the implementation of the concept shown, as well as examples of their implementation. The practical examples given in the article relate to issues such as the formation of the incident neighborhood, and allow you to form the context of information security management. In addition, an approach to automation of information security management processes is shown. The results of the work can be used both for simulation models and for implementation as a set of information security management processes in practical tasks. In addition, the results obtained can be integrated into orchestration tools for information security systems, which increases the effectiveness of responding to information security incidents.

Information security management; process approach; security management algorithm; data-based management; continuous detection and response.

Введение. Управление информационной безопасностью в современных условиях – непрерывно совершенствующаяся область, содержащая как минимум средства и методы управления безопасностью инфраструктурных решений, про-

цессы и процедуры, а также соответствующие им политики и средства управления. Стандартизированные решения также предполагают опору на циклы совершенствования, контроля и анализа, а также постоянное обновление применимых политик и требований.

В настоящий момент имеет место переход от модели управления информационной безопасностью, основанной на процессной модели, к модели непрерывного детектирования и реагирования на динамически формирующийся ландшафт угроз, что приводит к изменению базовых основ управления информационной безопасностью от уровня инфраструктуры до уровня разработки сценариев реагирования на инциденты информационной безопасности. Также сформирована потребность в обширной автоматизации бизнес-процессов и применении современных подходов к интеграции. Так, эксперты ведущих компаний отмечают, что уже к 2025 году 90% рабочих процессов по обеспечению безопасности будут автоматизированы и управляться в виде кода [1].

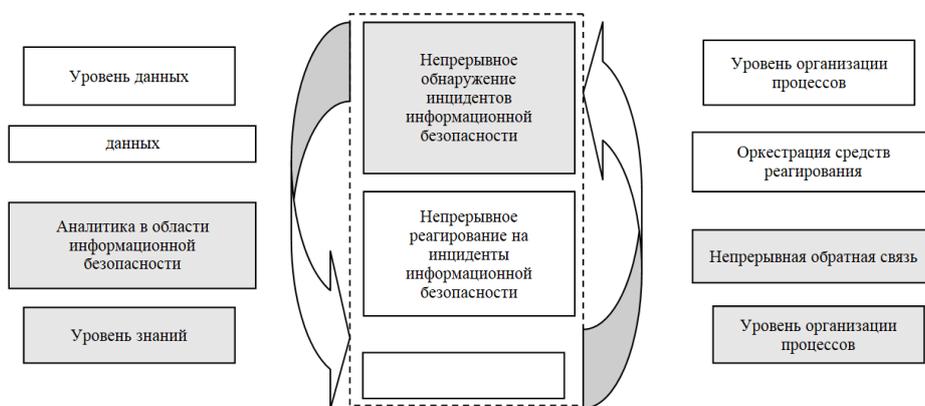


Рис. 1. Концепция CD/CR с учетом уровней развертывания

Кроме того, очевидной тенденцией является расширение области автоматизации в области информационной безопасности до фреймворков, содержащих как анализируемые объекты, так и наборы инструментов, процедур и процессов, реализующих типовые операции [2–4]. Автоматизация часто является необходимостью уже потому, что при традиционном процессном подходе количество атомарных задач реагирования на угрозы безопасности информации (задач, декомпозиция которых нецелесообразна) может достигать сотен тысяч и миллионов [5].

Внутри задач автоматизации привычным является наличие стандартных, статичных планов реагирования, которые представляют собой либо разветвленные и содержащие множество условий графы, сложные алгоритмы действий, покрывающие большое количество ситуаций, либо множество небольших алгоритмов реагирования с описанием их реализации (плейбуков), специализированных под конкретный тип инцидента.

При этом инфраструктура предприятия – живой организм, который постоянно меняется: добавляются новые устройства, сегменты сети, меняются политики работы критичных активов, добавляются новые средства защиты информации, в том числе и с уникальными характеристиками и настройками. По этой причине в организациях с развитым уровнем информационной безопасности регулярно проводятся аудит и инвентаризация, которые помогают актуализировать информационную модель предприятия.

Ключевой проблемной ситуацией в этом случае является следующее: существует ли способ в процессе инцидента подстраиваться под текущую ситуацию как в инфраструктуре, так и в способе исполнения атакующей техники?

Вопрос целесообразности реализации таких способов реагирования также зависит от уровня зрелости системы управления информационной безопасностью в организации. Здесь и далее предполагается, что организация, применяющая описанную в работе концепцию управления информационной безопасностью, как минимум стандартизировала и автоматизировала основные процессы управления информационной безопасностью, в том числе процесс управления инцидентами, и нуждается именно в оптимизации детектирования и реагирования, а не в первичной реализации этого процесса.

Следующий аспект, влияющий на эффективность защиты – динамика внешнего окружения. В текущих реалиях техники реализации атак эволюционируют, усложняются, затрагивают новые типы данных инфраструктуры, используют новые механизмы и способы посткомпрометации. Например, программы-вымогатели «ransomeware» теперь не только шифруют данные организации и требуют оплаты за восстановление доступа, но могут так же реализовать утечку данных организации, для реализации сценария шантажа организации в обмен на неразглашение информации властям, конкурентам или общественности.

В силу динамического ландшафта угроз и способов их реализации, плейбуки систем реагирования тоже должны изменяться, чтобы не устаревать на фоне меняющегося поведения злоумышленников.

В рамках развития подхода с 2022 года ведущими организациями начали на практике применяться модели, развивающие идеи непрерывного обнаружения и реагирования на инциденты информационной безопасности. В частности, операционная модель ASO [1] (рис. 1) основана на создании непрерывных циклов обратной связи в основных областях обнаружения и реагирования на инциденты. Ее совмещение с уровнями развертывания по четырехуровневой модели управления информационной безопасностью, ранее предложенной авторами, также учтено [6].

Исследователями также оценивается возможность комплексного измерения безопасности на базе оценки зрелости [7, 8], а также различные аналитические модели реагирования на угрозы безопасности, основанные на привязке к бизнес-процессам организации и возможностям оркестрации средств защиты информации [9, 10].

Следовательно, **что целью предлагаемых изменений** в целом является переход к концепции управления информационной безопасностью на основе цикла непрерывного детектирования и реагирования (CD/CR), учитывающей возможность непрерывного управления с реализацией обратной связи на основе данных от управляемого объекта, но при этом основанной на автоматизации всех рабочих процессов и быстром развертывании в инфраструктуре независимо от ее характеристик и особенностей.

Цели, задачи и ограничения. Итак, в настоящем исследовании предлагается концепция управления информационной безопасностью на основе цикла непрерывного детектирования и реагирования на инциденты информационной безопасности, базис которой составляют три условия:

- 1) Ориентация на управляемый объект;
- 2) Ориентация на данные;
- 3) Независимость от физической инфраструктуры.

Целью разработки концепции является обоснование нового подхода к автоматизации управления информационной безопасностью **на основе цикла непрерывного детектирования и реагирования на инциденты информационной**

безопасности в виде автоматизация-как-код и формирование основы для изменения политик безопасности и сценариев реагирования, учитывающих особенности этого подхода.

Проблемы, возникающие при этом, это:

- 1) **эксфльтрация данных и действий**,
- 2) создание **синтетических данных** для задач непрерывного детектирования и реагирования,
- 3) управление знаниями, а именно **сценарии динамического реагирования** на инциденты информационной безопасности.

Проблема **эксфльтрации данных** возникает при обмене данными между участниками процессов и объектами, содержащими необходимые им данные (узлами документооборота, внешними и внутренними агрегированными базами данных). При этом возникающие потоки данных могут быть нарушающими принятые политики безопасности в отношении информационного обмена, а коммуникации – скрытыми, неявными или слабоконтролируемыми [11].

При реализации управления на основе данных добавляется сложность контроля автоматизированных и автоматических коммуникаций и коннекторов, существующих как элемент автоматизации бизнес-процессов информационной безопасности и роботизированных программных решений. Эти коммуникации могут быть либо не полностью контролируемы с позиций распространения конфиденциальной информации (или иных типов информации, контроль над которыми может представлять интерес для организации), либо находиться в состоянии неполного контроля в определенных переходных состояниях организационных и информационных систем.

Проблема **эксфльтрации действий** возникает при работе автоматизированного процесса в области информационной безопасности тогда, когда требуется воздействие на объект, задействованный другими процессами, или на другие процессы. Такие действия могут быть злоумышленными, а также не полностью контролируемы, что создает опасность неконтролируемого изменения как самих процессов, так и операционной среды.

Проблема **управления знаниями** для задач автоматизации и управления в виде кода возникает из-за формирования баз знаний, содержащих сценарии с чувствительной информацией и удаленным доступом к ним, а также непрерывным изменением и распространением знаний в процессах управления информационной безопасностью. В частности, примерами проявления такой проблемы является обмен данными об индикаторах компроментации или применение сценариев реагирования на инциденты в децентрализованных системах.

Общая схема концепции управления информационной безопасностью на основе цикла непрерывного детектирования и реагирования на инциденты информационной безопасности. На уровне общей схемы и работы с данными добавляется также следующее условие: существует необходимость получить доступ к как можно большему количеству источников для сбора, анализа, валидации и верификации информации; с точки зрения информационной безопасности это будет нарушением принципа минимизации полномочий и локализации объектов защиты информации в инфраструктуре организации, если не будут приняты соответствующие усилия по формированию дополнений в политику информационной безопасности организации. Причина такого противоречия в том, что эффективность управления с увеличением количества источников данных должна возрастать [12], а сложность реализации систем защиты информации для подобных решений может становиться неприемлемой из-за возможных неполных или нецелесообразных подходов и решений. Важны также и форматы собираемых данных [13].

Итак, концепция должна на базовом уровне учитывать:

1. Работу с источниками данных.
2. Работу с политиками управления данными, применимыми в области информационной безопасности.
3. Работу с автоматизированными процессами и управляемыми объектами.

Кроме того, можно отметить, что существуют недостатки реально применяемых схем реализации детектирования и реагирования на инциденты информационной безопасности, связанные с низким уровнем зрелости процессов реагирования в организациях, даже если применены подходы по автоматизированному реагированию и сценарии реагирования на инциденты информационной безопасности реализованы в схеме автоматизация-как-код (рис. 2).

Указанная схема уже предлагает улучшения по сравнению с традиционным подходом на основе системы обеспечения информационной безопасности, основанной на процессной модели. В частности, появление сервисно-ориентированной модели управления безопасностью, шаблонов (паттернов) безопасности и отдельного управления сервисами безопасности (этап I) декларирует повышение управляемости на уровне организации процессов информационной безопасности.



Рис. 2. Недостатки схемы реагирования на низком уровне зрелости

Вместе с тем имеются следующие недостатки, которые ограничивают эффективность предлагаемой схемы:

- ♦ предполагается, что данных, полученных на этапе II, достаточно для реагирования на угрозы информационной безопасности. Это не вполне соответствует эмпирически полученным результатам. В частности, примеры извлечения информации о динамически меняющемся ландшафте угроз [14] сигнализируют о возможной неэффективности указанной схемы в случае непрерывного изменения атак и неполной информации о деструктивных действиях,

◆ предполагается, что может существовать стабильный механизм противодействия угрозам безопасности, привязанный к инфраструктуре. По-видимому, это не так, поскольку динамически меняющийся ландшафт угроз приведет к неэффективности подобного механизма, даже реализованного на этапе III в рамках сервисной модели,

◆ границы области анализа и рекомендации по визуализации данных также требуют учета динамики развития реагирования на угрозы безопасности, поскольку указанные выше недостатки говорят о необходимости формирования дополнительных (синтетических) данных для обогащения информации об объекте управления, а также постоянном мониторинге инфраструктуры с учетом полученных данных.

Основываясь на этом, можно сформировать основные **принципы** предлагаемой концепции:

0) Предварительно должно учитываться условие, что любой автоматизированный и управляемый как код процесс будет использовать множество источников данных и баз знаний и управляемый объект не должен быть причиной утечки информации во время такого использования. Следовательно, должны быть предъявлены отдельные требования к операционной среде. Кроме того, должен существовать изолированный репозиторий, содержащий все необходимые для управления данные объекта, и политика управления доступом к указанному репозиторию на всех этапах его жизненного цикла.

1) Все данные и знания, ключевые для выполнения операций над управляемым объектом, должны быть связаны с ним и храниться в общей для объекта и автоматизированного процесса управления операционной среде. Этот принцип может предполагать управление объектом как микросервисом, контейнеризацию, формирование песочниц или иных виртуальных сред, а также облачные решения. Таким образом, данные и знания должны быть инкапсулированы в объект.

В качестве дополнения необходимо отметить, что управляемый объект абстрагируется от физической инфраструктуры. Границы управляемого объекта и область управления должны быть определены согласно требованиям автоматизированного процесса управления информационной безопасностью, разворачиваемого с определенными конфигурационными параметрами, а также требований коннекторов и иных средств извлечения данных. Управление знаниями должно осуществляться в рамках границ управляемого объекта.

Внешние запросы к данным и знаниям должны быть контролируемыми и не должны изменять накопленные данные и знания внешне.

Данные и знания, необходимые для динамического изменения алгоритмов реагирования и составляющие контекст безопасности для конкретного управляемого объекта, группы объекта или системы в целом, могут и должны быть синтезированы на этапе анализа объекта и привязки к физической инфраструктуре. При этом может иметься как статический, так и динамический набор алгоритмов синтеза и обогащения контекста безопасности. Интерес может представлять также интерпретация внешних источников данных на этапе синтеза и обогащения контекста.

Также в рамках обогащения контекста могут и должны (если такая возможность существует) использоваться ретроданные, то есть данные о предыдущих состояниях управляемого объекта и (если такая возможность существует) связанных объектов.

2) Процесс может управлять группой объектов, используя общую конфигурацию автоматизированного управления. Наследование характеристик управления в рамках концепции предполагает как использование отдельных методов и процедур контроля, применимых для конкретного управляемого объекта, так и форми-

рование общей операционной модели для классов объектов, таких, например, как автоматизированные системы управления технологическими процессами. Каждый способ унификации характеристик должен предполагать безопасный алгоритм обращения к данным и знаниям, унифицированный в части источников данных и знаний, в том числе внешних, а также правил доступа к ним.

3) Разные типы данных, получаемые из источников данных объекта, должны обрабатываться универсальными способами. При этом должны поддерживаться как основанные на API способы сбора данных, так и способы, позволяющие собирать данные без прямого доступа, в том числе и свидетельства из косвенных источников. Тем не менее, даже свидетельства из косвенных источников должны быть цифровыми и собираться автоматизированно.

Кроме того, должен существовать универсальный способ обращения к ретроданным для целей обогащения контекста безопасности. Это предполагает либо существование накопленной статистики данных, либо формирование такой статистики с момента внедрения концепции, основываясь на указанных выше принципах.

Далее рассмотрена общая схема работы в рамках данной концепции, учитывающая четырехуровневую модель управления информационной безопасностью [6] (рис. 3), а также схема работы с автоматизированными бизнес-процессами управления информационной безопасностью на уровне данных и знаний (рис. 4).

Вспомогательные процессы в данной схеме должны обеспечить работу с записями, документами и регистрационными данными.

Задачей вспомогательных процессов также может быть формирование и поддержка операционных средств, включая управление облачными или виртуальными средствами, регламенты и алгоритмы их контроля, а также создание защищенных интерфейсов и баз данных для служебной (технологической) информации.



Рис. 3. Общая схема управления информационной безопасностью на основе предлагаемой концепции

Кроме того, необходимо рассмотреть высокоуровневую автоматизацию процессов управления информационной безопасностью, которая должна содержать базовые (типовые) алгоритмы и процессы реагирования, использующие только доступные извне данные управляемых объектов (рис. 4). Типовые алгоритмы и процессы могут использоваться в различных задачах, таких как быстрое развертывание в рамках инфраструктуры-как-код, обучение и демонстрационные стенды, а также для исследовательских задач и формирования новых специальных алгоритмов реагирования на угрозы информационной безопасности.

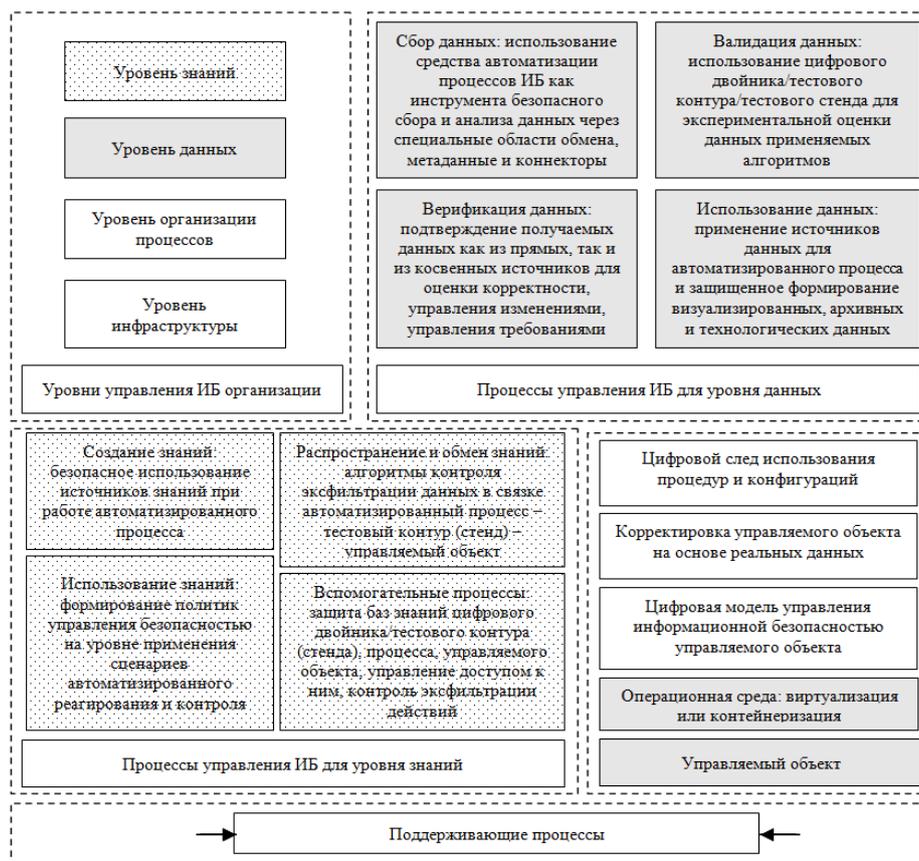


Рис. 4. Схема работы с автоматизированными бизнес-процессами управления информационной безопасностью на уровне данных и знаний

Таким образом (рис. 4) также создается и основа для быстрой разработки автоматизированных фреймворков, содержащих наборы типовых алгоритмов для различных задач и их защищенные модификации.

Пример реализации предложенной концепции. Принцип эксперимента в следующем: проверяется возможность оценки работы с объектом на описанных в концепции принципах для реальной системы класса SOAR. Для решения первой задачи (учета динамики изменений инфраструктуры) был разработан подход на основе предложенной концепции (рис. 5), который заключается в применении лучших практик CD/CR к выстраиванию процессов информационной безопасности (ИБ). Первыми об этой технологии как о перспективной разработке, как уже упоминалось выше, заявила корпорация Google [1]. В разрезе их исследований

предлагается разбить процессы и процедуры информационной безопасности до атомарных функций, которые можно переспользовать и которые являются по сути небольшими унифицированными процессами по выполнению операции. Атомарные унифицированные процессы хранятся в репозитории; функция выполняется единственным экземпляром рабочего процесса, который правится в одном месте разными заинтересованными (таким образом, соблюдается консистентность) и, когда возникает необходимость, этот процесс стандартным образом включается в плейбук за счет унифицированного input/output.

Примером унифицированного процесса может быть процедура сбора аутентификационных сессий с хоста. Она может быть выполнена разными способами и разными средствами в зависимости от типа хоста, политик использования хоста или закрывающего средства защиты информации. Но, если все эти вариации обрабатываются в рамках одного универсального процесса сбора аутентификаций с одним стандартным input/output, то строить процессы ИБ проще. Атомарный процесс, ко всему прочему, будет переиспользован в совершенно разных процессах ИБ: не только при построении динамических плейбуков в процессе управления инцидентами, но еще и в управлении активами, управлении учетными записями и т.д.

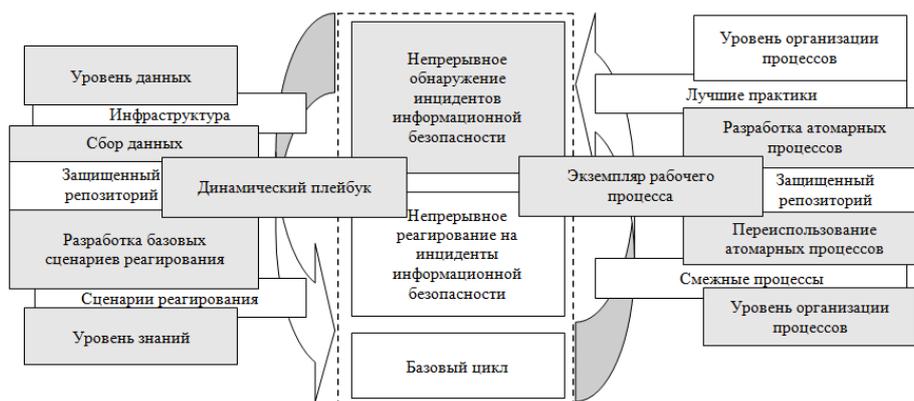


Рис. 5. Применение концепции на практике с учетом уровней развертывания

Набор таких унифицированных процессов дает нам возможность формировать уникальный подход к конкретной инфраструктуре или конкретному ее состоянию.

Выше (рис. 5) рассмотрен только способ решения первой части проблемы – учета актуального состояния инфраструктуры. Так же нам нужно решить проблему обфускации проведения атак [15], что по сути своей является уже дедуктивным расследованием, которое сложно покрыть корреляционным анализом. Дополнительно необходимо учесть, что атакующие могут применять множество дополнительных возможностей, таких как боковые перемещения, эксфильтрацию данных и прочее [16], что дополнительно осложняет детектирование и анализ. В какой-то степени отклонения от поведения можно покрыть через анализ аномалий (UEBA) [17] и все равно это не дает полной информации об инциденте: поиск аномалий в поведении устройств/человека и отклонения в исполнении атак – это разные вещи. Таким образом, на данном этапе необходимо работать над поиском признаков компрометаций устройств или учетных записей, используя технологию threat hunting (построение гипотез об угрозах безопасности информации по поиску подозрительной активности). И действительно, контролирование изменчивости про-

ведения техник атак через распространенные признаки компрометации устройств – одна из задач, решаемых аналитиками третьей линии, специализирующихся на поиске угроз.

Если использовать подходы threat hunting в применении к окрестностям инцидента, то мы можем найти дополнительные объекты, которые были скомпрометированы, несмотря на то, что ранее техника проведения атаки их не затрагивала. Например, эксплуатация уязвимости не компрометировала учетную запись или вредоносное программное обеспечение не подгружало какой-то хакерский tool set.

Алгоритм формирования окрестности инцидента, то есть автоматизированного формирования контекста инцидента, может включать (но не обязательно состоит только из них) следующие шаги:

Этап 1. Внутреннее сетевое сканирование и попытки эксплуатации уязвимостей внутри сети автоматизированными средствами анализа защищенности и/или специальными инструментальными средствами для тестирования на проникновение.

Этап 2. Детектирование подозрительной сетевой активности системных утилит (rundll32, regsvr32, mshta, certutil).

Этап 3. Детектирование запуска командных интерпретаторов (powershell, cmd, wscript, cscript) офисными приложениями (word, excel).

Этап 4. Мониторинг процесса записи в ключи реестра, отвечающие за автозагрузку либо за изменение системных настроек операционной системы.

Этап 5. Мониторинг создания служб, задач планировщика с powershell, а также LOLbins утилитами.

Этап 6. Детектирование запуска пользователем нерегламентированного ПО (утилиты администрирования, VPN-клиенты, TOR, torrent-клиенты и т.д.).

Рассмотрим также более подробно сбор данных об инциденте (рис. 5) на примере. Очевидно, что алгоритм формирования окрестности инцидента может быть дополнен, если появляется типовой способ детектирования конкретного или типового инцидента или его индикаторов компрометации.

Далее, если наложить это на практические аспекты формирования окрестности инцидента, можно заметить, что, к примеру, подозрительные процессы находятся по подозрительным родителям системных процессов, то есть логика наследования соблюдается и в области детектирования (рис. 6).

```
Команда запуска процесса: chisel client 85.192.50.11:8080 R:socks
Путь к процессу: /var/www/chisel
Родительский процесс: bash
Путь к родительскому процессу: /bin/bash
Описание:
chisel
This package contains a fast TCP/UDP tunnel, transported over HTTP, secured via SSH. Single executable including both client and server. Chisel is mainly useful for passing through firewalls, though it can also be used to provide a secure endpoint into your network.
Installed size:
```

Рис. 6. Дополнительные данные об инциденте

Расширяя пример использования, можно показать, что в командной строке можно найти код powershell с «полезной нагрузкой» в base64, что означает обфускацию кода, скрывающую дополнительный объект, потенциально используемый злоумышленником (рис. 7). Таким образом, расширяется количество данных, используемых для анализа инцидента.

жающую ландшафт атаки, растянутой по периметру, времени и способам исполнения. В действительности атаки становятся все более комплексными и сложными, из инцидентов с различными техниками, которые еще и связаны между собой неочевидными параметрами или свойствами (например, наличием в инцидентах одного и того же процесса, либо использование однотипного *named pipes*).

Подобные ситуации не покрываются стандартными планами реагирования, но, если добавить в стандартные подходы динамическую составляющую, то инциденты могут быть связаны по идентичным объектам и контексту: запуски аналогичных подозрительных процессов, сетевых аутентификаций, задействованных учетных записей, корреляция имен задействованных хостов в разных инцидентах, скорее всего, говорят о принадлежности к одному и тому же *kill chain*. При таком подходе инциденты становятся связанными, а динамические плейбуки выстраиваются в последовательность работы над атакой, то есть переходят на более высокий (стратегический) уровень планов реагирования.

Если проанализировать собранную информацию через классификацию по угрозам *Mitre Att@ck* [19] или БДУ ФСТЭК [20], то этапы *kill chain* могут быть преобразованы автоматизированно в последовательность тактик, применяемых нарушителем, на которые со стороны защищаемой организации динамически выстраивается контекстный ответ.

Заключение. Исследование сосредоточено на формировании пригодных в практике рекомендаций по построению процессной модели управления информационной безопасностью, построенной на предлагаемой концепции управления информационной безопасностью. Особенностью подхода является привязка автоматизации и детектирования угроз безопасности к управляемому объекту, а не к инфраструктуре или отдельному процессу. При текущей волатильности в информационной безопасности, сложной динамике процессов управления информационной безопасностью и контекста инцидентов, в том числе и поведения нарушителей, состояния защищаемых инфраструктур, необходимо соответствующие, симметричные динамике ситуации, динамичные подходы к детектированию и реагированию на инциденты информационной безопасности.

Представлены также некоторые условия управления информационной безопасностью при применении предлагаемой концепции, такие как контроль экс-фильтрации данных.

Подобное расширение может быть полезно при развертывании разных типов экспертных и советующих систем, систем поддержки принятия решений, ситуационных центров и особенно интересно для задач управления знаниями в области создания и улучшения динамически формируемых сценариев реагирования на угрозы информационной безопасности, опирающихся на описанную концепцию, которые возможно рассмотреть в отдельных исследованиях.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Ghanizada I.* IT prediction: the vast majority of security operations workloads will be automated. – URL: <https://cloud.google.com/blog/products/identity-security/it-prediction-vast-majority-of-security-operations-workloads-will-be-automated>.
2. *Королев И.Д., Литвинов Е.С., Маркин Д.И.* Повышение уровня автоматизации процессов сбора данных о выявленных событиях и инцидентах информационной безопасности // Инженерный вестник Дона. – 2021. – Т. 82, № 10. – С. 140-151.
3. *Котенко И.В., Саенко И.Б., Юсупов Р.М.* Новое поколение систем мониторинга и управления инцидентами безопасности // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Информатика. Телекоммуникации. Управление. – 2014. – № 3 (198). – С. 7-18.

4. *Богданов В.В., Домуховский Н.А., Савин М.В.* SOAR: автоматизация работы с инцидентами информационной безопасности // *Защита информации. Инсайд.* – 2021. – № 3 (99). – С. 13-17.
5. *Котенко И.В., Федорченко А.В., Саенко И.Б., Кушнеревич А.Г.* Технологии больших данных для корреляции событий безопасности на основе учета типов связей // *Вопросы кибербезопасности.* – 2017. – № 5 (24). – С. 2-16. – DOI: 10.21681/2311-3456-2017-5-2-16.
6. *Золотарев В.В., Лапина М.А.* Модель и алгоритм управления информационной безопасностью образовательной организации высшего образования с учетом требований управления на основе данных // *Прикаспийский журнал: управление и высокие технологии.* – 2022. – № 4 (60). – С. 107-118. – DOI: 10.54398/20741707_2022_4_107.
7. *Vailey K.* Detection Engineering Maturity Matrix. – URL: <https://detectionengineering.io/>.
8. *Велигодский С.С., Милославская Н.Г.* Подход к оценке уровня зрелости центров управления сетевой безопасностью // *Системы высокой доступности.* – 2023. – Т. 19, № 2. – С. 25-37. – DOI: 10.18127/j20729472-202302-02.
9. *Королев И.Д., Попов В.И., Коноваленко С.А.* Методика аналитической обработки распределенных во времени инцидентов информационной безопасности // *Научные исследования в космических исследованиях Земли.* – 2020. – Т. 12, № 5. – С. 53-61. – DOI: 10.36724/2409-5419-2020-12-5-53-61.
10. *Сагиров Р.А.* Применение нейронных сетей для автоматизации задач в области информационной безопасности // *Защита информации. Инсайд.* – 2019. – № 5 (89). – С. 56-59.
11. *Золотарев В.В.* Алгоритм контроля эксфильтрации данных с учетом требований управления на основе данных // *Прикаспийский журнал: управление и высокие технологии.* – 2023. – № 4 (64).
12. *Иванов А.В., Никрошкин И.В., Огнев И.А., Киселев М.А.* Применение средств экспертизы Blue Team в процессе мониторинга информационных систем на примере платформы TI (Threat Intelligence) // *Безопасность цифровых технологий.* – 2023. – № 2 (109). – С. 34-51. – DOI: 10.17212/2782-2230-2023-2-34-51.
13. *Савин М.В., Стойчин К.Л., Некрасов А.В., Комаров Н.В.* Обзор стандартов и форматов представления автоматизированных сценариев реагирования на инциденты компьютерной безопасности // *Защита информации. Инсайд.* – 2022. – № 4 (106). – С. 14-19.
14. *Rahman R., Hezaveh R., Williams L.* What Are the Attackers Doing Now? Automating Cyberthreat Intelligence Extraction from Text on Pace with the Changing Threat Landscape: A Survey // *ACM Comput. Surv.* – December 2023. – 55, 12, Article 241. – 36 p. – <https://doi.org/10.1145/3571726>.
15. *Ерохин В.* Поиск вредоносных сценариев powershell с использованием синтаксических деревьев // *Безопасность информационных технологий.* – 30 (3). – P. 77-89. – DOI: <http://dx.doi.org/10.26583/bit.2023.3.05>.
16. *Follina Exploit Leads to Domain Compromise.* – URL: <https://thefirreport.com/2022/10/31/follina-exploit-leads-to-domain-compromise/>.
17. *Salitin M.A., Zolait A.H.* The role of User Entity Behavior Analytics to detect network attacks in real time // *2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Sakhier, Bahrain, 2018.* – P. 1-5. – DOI: 10.1109/3ICT.2018.8855782.
18. *Hutchins E.M., Cloppert M.J., Amin R.M.* Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains // *Lockheed Martin Corporation.* – URL: <https://www.lockheedmartin.com/content/dam/lockheedmartin/rms/documents/cyber/LM-White-Paper-Intel-Driven-Defense.pdf>.
19. *MITRE ATT&CK.* – URL: <https://attack.mitre.org/>.
20. *Банк данных угроз безопасности информации / Федеральная служба по техническому и экспортному контролю, Государственный научно-исследовательский испытательный институт проблем технической защиты информации.* – URL: <https://bdu.fstec.ru/>.

REFERENCES

1. *Ghanizada I.* IT prediction: the vast majority of security operations workloads will be automated. Available at: <https://cloud.google.com/blog/products/identity-security/it-prediction-vast-majority-of-security-operations-workloads-will-be-automated>.

2. Korolev I.D., Litvinov E.S., Markin D.I. Povyshenie urovnya avtomatizatsii protsessov sbora dannykh o vyyavlennykh sobyitiyakh i intsidentakh informatsionnoy bezopasnosti [Increasing the level of automation of data collection processes on identified events and incidents of information security], *Inzhenernyy vestnik Dona* [Engineering Bulletin of the Don], 2021, Vol. 82, No. 10, pp. 140-151.
3. Kotenko I.V., Saenko I.B., Yusupov R.M. Novoe pokolenie sistem monitoringa i upravleniya intsidentami bezopasnosti [A new generation of security incident monitoring and management systems], *Nauchno-tehnicheskie vedomosti Sankt-Peterburgskogo gosudarstvennogo politekhnicheskogo universiteta. Informatika. Tele-kommunikatsii. Upravlenie* [Scientific and Technical Bulletin of St. Petersburg State Polytechnic University. Computer science. Telecommunications. Management], 2014, No. 3 (198), pp. 7-18.
4. Bogdanov V.V., Domukhovskiy N.A., Savin M.V. SOAR: avtomatizatsiya raboty s intsidentami informatsionnoy bezopasnosti [SOAR: automation of work with information security incidents], *Zashchita informatsii. Insayd* [Information protection. Inside], 2021, No. 3 (99), pp. 13-17.
5. Kotenko I.V., Fedorchenko A.V., Saenko I.B., Kushnerevich A.G. Tekhnologii bol'shikh dannykh dlya korrelyatsii sobyitiy bezopasnosti na osnove ucheta tipov svyazey [Big data technologies for the correlation of security events based on the types of connections], *Voprosy kiberbezopasnosti* [Issues of cybersecurity], 2017, No. 5 (24), pp. 2-16. DOI: 10.21681/2311-3456-2017-5-2-16.
6. Zolotarev V.V., Lapina M.A. Model' i algoritm upravleniya informatsionnoy bezopasnost'yu obrazovatel'noy organizatsii vysshego obrazovaniya s uchetom trebovaniy upravleniya na osnove dannykh [Model and algorithm of information security management of an educational organization of higher education taking into account the requirements of data-based management], *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: management and high technologies], 2022, No. 4 (60), pp. 107-118. DOI: 10.54398/20741707_2022_4_107.
7. Bailey K. Detection Engineering Maturity Matrix. Available at: <https://detectionengineering.io/>.
8. Veligodskiy S.S., Miloslavskaya N.G. Podkhod k otsenke urovnya zrelosti tsentrov upravleniya setevoy bezopasnost'yu [An approach to assessing the level of maturity of network security management centers], *Sistemy vysokoy dostupnosti* [High availability systems], 2023, Vol. 19, No. 2, pp. 25-37. DOI: 10.18127/j20729472-202302-02.
9. Korolev I.D., Popov V.I., Konovalenko S.A. Metodika analiticheskoy obrabotki raspredelennykh vo vremeni intsidentov informatsionnoy bezopasnosti [Methodology of analytical processing of information security incidents distributed in time], *Naukoemkie tekhnologii v kosmicheskikh issledovaniyakh Zemli* [High-tech technologies in Earth space research], 2020, Vol. 12, No. 5, pp. 53-61. DOI: 10.36724/2409-5419-2020-12-5-53-61.
10. Sagirov R.A. Primenenie neyronnykh setey dlya avtomatizatsii zadach v oblasti informatsionnoy bezopasnosti [Application of neural networks for automation of tasks in the field of information security], *Zashchita informatsii. Insayd* [Information protection. Inside], 2019, No. 5 (89), pp. 56-59.
11. Zolotarev V.V. Algoritm kontrolya eksfil'tratsii dannykh s uchetom trebovaniy upravleniya na osnove dannykh [Algorithm of data exfiltration control taking into account the requirements of data-based management], *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: management and high technologies], 2023, No. 4 (64).
12. Ivanov A.V., Nikroshkin I.V., Ognev I.A., Kiselev M.A. Primenenie sredstv ekspertizy Blue Team v protsesse monitoringa informatsionnykh sistem na primere platformy TI (Threat Intelligence) [The use of Blue Team expertise tools in the process of monitoring information systems on the example of the TI (Threat Intelligence) platform], *Bezopasnost' tsifrovyykh tekhnologiy* [Security of digital technologies], 2023, No. 2 (109), pp. 34-51. DOI: 10.17212/2782-2230-2023-2-34-51.
13. Savin M.V., Stoychin K.L., Nekrasov A.V., Komarov N.V. Obzor standartov i formatov predstavleniya avtomatizirovannykh stsensariyev reagirovaniya na intsidenty komp'yuternoy bezopasnosti [Review of standards and formats for the presentation of automated scenarios for responding to computer security incidents], *Zashchita informatsii. Insayd* [Information protection. Inside], 2022, No. 4 (106), pp. 14-19.
14. Rahman R., Hezaveh R., Williams L. What Are the Attackers Doing Now? Automating Cyberthreat Intelligence Extraction from Text on Pace with the Changing Threat Landscape: A Survey, *ACM Comput. Surv.*, December 2023, 55, 12, Article 241, 36 p. Available at: <https://doi.org/10.1145/3571726>.

15. Erokhin V. Poisk vredonosnykh stsenariiev powershell s ispol'zovaniem sintaksicheskikh derev'ev [Searching for malicious powershell scripts using syntax trees], *Bezopasnost' informatsionnykh tekhnologiy* [Information Technology Security], 30 (3), pp. 77-89. DOI: <http://dx.doi.org/10.26583/bit.2023.3.05>.
16. Follina Exploit Leads to Domain Compromise. Available at: <https://thefirreport.com/2022/10/31/follina-exploit-leads-to-domain-compromise/>.
17. Salitin M.A., Zolait A.H. The role of User Entity Behavior Analytics to detect network attacks in real time, *2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Sakhier, Bahrain, 2018*, pp. 1-5. DOI: 10.1109/3ICT.2018.8855782.
18. Hutchins E.M., Cloppert M.J., Amin R.M. Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains, *Lockheed Martin Corporation*. Available at: <https://www.lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/LM-White-Paper-Intel-Driven-Defense.pdf>.
19. MITRE ATT&CK. Available at: <https://attack.mitre.org/>.
20. Bank dannykh ugroz bezopasnosti informatsii [Data bank of information security threats], Federal'naya sluzhba po tekhnicheskomu i eksportnomu kontrolyu, Gosudarstvennyy nauchno-issledovatel'skiy ispytatel'nyy institut problem tekhnicheskoy zashchity informatsii [Federal Service for Technical and Export Control, State Research and Testing Institute of Problems of Technical Protection of Information]. Available at: <https://bdu.fstec.ru/>.

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

Олейникова Анна Алексеевна – ООО «Интеллектуальная безопасность»; e-mail: ana.oleynikova@gmail.com, г. Москва, Россия; тел.: 83912227639.

Золотарев Вячеслав Владимирович – Сибирский государственный университет науки и технологий; e-mail: amida.2@yandex.ru; г. Красноярск, Россия; тел.: 83912227639; к.т.н.; доцент.

Oleynikova Anna Alekseevna – Intellectual Security LLC; e-mail: ana.oleynikova@gmail.com; Moscow, Russia; phone: +73912227639.

Zolotarev Vyacheslav Vladimirovich – Siberian State University of Science and Technology; e-mail: amida.2@yandex.ru; Krasnoyarsk, Russia; phone: +73912227639; cand. of eng. sc.; associate professor.

УДК 004.89

DOI 10.18522/2311-3103-2023-5-81-92

В.С. Усатюк, С.И. Егоров, А.П. Локтионов, Е.А. Титенко, И.Е. Чернецкая
АРХИТЕКТУРА НЕЙРОННЫХ СЕТЕЙ НА ОСНОВЕ КОДОВ НА ГРАФАХ

Одним из важных достижений теории помехоустойчивого кодирования является открытие кодов на графах и их важного подмножества низкоплотностных кодов (LDPC-кодов). Используя проверочную матрицу кода на графе, можно получить марковское случайное поле. LDPC-код может быть вложен в модель Изинга (разновидность марковского случайного поля) путем использования топологии тора с отрицательной кривизной. При этом кодовые слова соответствуют седловым точкам (экстремумам) в модели, а треппин-сети соответствуют локальным минимумам. Использование LDPC-кодов с увеличенным кодовым расстоянием позволяет максимально разнести седловые точки, и таким образом повысить устойчивость нейронной сети к шуму и мощность представления. При этом блочная и разряженная структура, характерная для тора отрицательной кривизны, упрощает мультиплексирование и снижает число обучаемых параметров нейронной сети. Целью исследования являются снижение вычислительной сложности и увеличение точности нейронных сетей за счёт применения априорных структурных (квазициклических) разряженных графов для широкого класса задач машинного обучения на марковских случайных полях. В работе представлен новый подход, позволяющий осуществ-

лять синтез архитектур нейронных сетей на основе кодов на графах. Предложенный подход осуществляет эффективное представление марковских случайных полей за счёт применения разреженных блочных (квазициклических) матриц (тензоров). Предложенный подход позволяет снизить число обучаемых параметров и логарифмически снизить сложность мультиплексирования тензора. Полученная на основе предложенного подхода архитектура трансформера в задаче поиска пути (pathfinder) с конкурса трансформеров (long range arena) заняла пятое место по точности классификации изображений 94.95% (1.72% от первого места) при значительно меньшей сложности (число параметров (умножений) синтезированной сети меньше в более чем 5 раз). Применение предложенного подхода к задачам факторизации на плотных графах, сетевых задачах, поверхностных сетках, ковариационных матрицах позволило увеличить точность реконструкции по метрике Фробениуса (на отдельных задачах на 8 порядков) в сочетании с упрощением структуры мультиплексора в сравнение с методами усеченного сингулярного разложения TSVD и хордовой разреженной факторизации.

Нейронные сети; коды на графах; низкоплотностные коды; модель Изинга; матричная факторизация; вложение многообразия; торическая гиперболическая топология.

V.S. Usatyuk, S.I. Egorov, A.P. Loktionov, E.A. Titenko, I.E. Chernetskaya

NEURAL NETWORK ARCHITECTURE BASED ON GRAPH CODES

One of the important achievements of the theory of error-correcting coding is the discovery of graph codes and their important subset - low-density parity check codes (LDPC codes). Using the parity check matrix of the code on the graph, one can obtain a Markov random field. LDPC code can be embedded in an Ising model (a type of Markov random field) by using a torus topology with negative curvature. In this case, codewords correspond to saddle points (extrema) in the model, and trapping sets correspond to local minima. The use of LDPC codes with an increased code distance allows for maximum separation of saddle points, and thus increases the noise resistance of the neural network and the representation power. At the same time, the block and sparse structure, characteristic of a torus of negative curvature, simplifies multiplexing and reduces the number of trainable parameters of the neural network. The aim of the research is to reduce the computational complexity and increase the accuracy of neural networks through the use of a priori structural (quasi-cyclic) sparse graphs for a wide class of machine learning problems on Markov random fields. The paper presents a new approach that allows the synthesis of neural network architectures based on graph codes. The proposed approach provides an effective representation of Markov random fields through the use of QC-LDPC matrices and tensors. The proposed approach allows us to reduce the number of trainable parameters and logarithmically reduce the complexity of tensor multiplexing. The proposed approach provided an accuracy of 94.95% (1.72% to first place) of the binary classification problem "Pathfinder" of the "Long Range Arena" competition, with more than 5 times fewer parameters (multiplications). Application of the proposed approach to factorization problems on dense graphs, network problems, surface meshes, covariance matrices made it possible to increase the accuracy of reconstruction using the Frobenius metric in individual problems by more than 8 orders of magnitude in combination with simplifying the structure of the multiplexer.

Neural networks; codes on graph; LDPC codes; Ising model; matrix factorization; manifold embedding; toric hyperbolic topology.

Введение. Технологии искусственного интеллекта (ИИ) приобретают все большее значение в современном мире. Наибольший прогресс в области ИИ обеспечило применение нейронных сетей с глубоким обучением, которые в ряде приложений уже могут заменить человека. Для расширения области применения этих сетей необходимо уменьшить их сложность. В статье предлагается подход к построению нейронных сетей уменьшенной сложности на основе применения кодов на графах.

1. Коды на графах. Коды на графах были введены Таннером [1], который доказал оптимальность алгоритмов исправления ошибок методом распространения доверия (BP, belief-propagation) для кодов, определенных на графах без циклов. В классе кодов на графах особый практический интерес в силу наименьшей сложности декодирования имеют низкоплотностные коды.

Низкоплотностный (LDPC, Low Density Parity Check) код – это блочный линейный код размерностью k и длиной кодового слова n , задаваемый проверочной матрицей H размерностью $(n-k) \times n$, имеющей небольшую плотность отличных от нуля символов [2]. По определению проверочной матрицы для любого кодового слова v LDPC-кода справедливо следующее: $v \cdot H^T = 0$. Каждая строка проверочной матрицы H задает уравнение проверки на четность:

$$\sum_{l=0}^{n-1} v_l \cdot h_{j,l} = 0, \quad (1)$$

где $h_{j,l}$ – элемент проверочной матрицы, j – номер строки проверочной матрицы (номер проверочного уравнения), l – номер символа кодового слова.

Достоинством низкоплотностных кодов является возможность применения субоптимального алгоритма декодирования с мягкими решениями методом распространения доверия (BP), обладающего значительно большей помехоустойчивостью по сравнению с алгоритмами декодирования с жесткими решениями, сложность которого растет линейно относительно длины кода.

Алгоритм BP предусматривает представление LDPC-кода в виде двудольного графа Таннера (пример графа Таннера приведен на рис. 1).

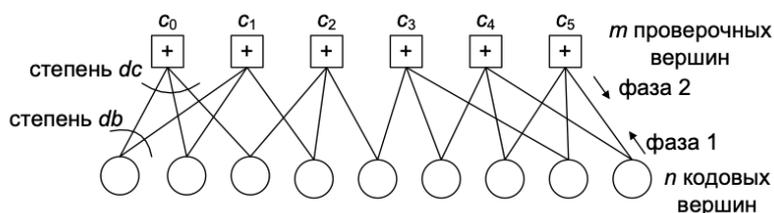


Рис. 1. Двудольный граф двоичного регулярного LDPC кода длины 9

Граф Таннера G имеет два множества вершин. Одно состоит из $m=n-k$ проверочных вершин $\{c_0, c_1, \dots, c_{m-1}\}$, соответствующих m строкам матрицы H , второе – из n кодовых вершин $\{v_0, v_1, \dots, v_{n-1}\}$, соответствующих n столбцам матрицы H . Кодовая вершина v_l соединяется ребром с проверочной вершиной c_i в том случае, если $h_{j,l} \neq 0$.

В соответствии с итеративным алгоритмом декодирования BP получение верных значений бит кодового слова осуществляется в результате многократного обмена сообщениями вершинами графа Таннера. Каждая итерация алгоритма содержит две фазы. В фазе 1 обновляются сообщения проверочных вершин на основе анализа сообщений кодовых вершин; в фазе 2 – сообщения кодовых вершин на основе анализа сообщений проверочных вершин.

На эффективность BP-декодирования отрицательно влияет наличие циклов в графе Таннера, образующих треппин-сету (Trapping set, TS,) или (a,b) -подграфы (подграфы в графе Таннера, состоящие из a символьных узлов, b из которых инцидентны проверочным узлам с нечетными степенями). Эти подграфы обуславливают ошибку BP-декодирования. В случае если сообщения проверочных узлов изменит значения символьных узлов, инцидентных нечетному числу проверок, то, вследствие неправильного подсчета условных вероятностей, обусловлен-

ного циклами графа Таннера, на символьных узлах подграфа ошибка не будет скорректирована, даже если она является корректируемой в соответствии с дистантными свойствами кода.

Важной характеристикой LDPC-кода является приближенный спектр связности (ACE Spectrum) графа Таннера. Спектр связности представляется в виде вектора:

$$ACE(H) = (ACE_{4,0}(VN), ACE_{4,1}(VN), \dots, ACE_{i,j}(VN), \dots, ACE_{k,m}(VN)),$$

где $ACE_{i,j}(VN)$ - количество символьных узлов, содержащихся в цикле длины i со значением связности j , VN - подмножество символьных узлов графа Таннера, образующих циклы длины $i \in \{4, 6, \dots\}$. Значение связности (ACE, Approximated Cycle Extrinsic Message Degree) вычисляется для символьных узлов, содержащихся в подграфе образованном циклом $v_i \in C$, $ACE(C) = \sum_{v \in C} (d(v) - 2)$, где $d(v)$ - степень инцидентности символьного узла. Таким образом, код с лучшим спектром связности при одинаковом спектре кода при использовании декодирования методом распространения доверия обеспечивает лучшую помехоустойчивость. Этому коду не будут мешать псевдокодовые слова, обусловленные TS.

На практике для уменьшения сложности декодеров LDPC-кодов используются квазициклические коды. Квазициклический регулярный (J, L) LDPC-код (J - вес столбца (число единиц в столбце) матрицы, L - вес строки) задается проверочной матрицей [3]:

$$H = \begin{bmatrix} I(p_{0,0}) & I(p_{0,1}) & \dots & I(p_{0,L-1}) \\ I(p_{1,0}) & I(p_{1,1}) & & I(p_{1,L-1}) \\ \vdots & \vdots & \ddots & \vdots \\ I(p_{J-1,0}) & I(p_{J-1,1}) & \dots & I(p_{J-1,L-1}) \end{bmatrix}, \quad (2)$$

где $0 \leq j \leq J - 1$, $0 \leq l \leq L - 1$ и $I(p_{j,l})$ - подматрица перестановки размера $z \cdot z$ (циркулянт - единичная матрица, циклически сдвинутая вправо на $p_{j,l}$ символов). Проверочная матрица регулярного кода также может содержать нулевые подматрицы размера $z \cdot z$. Если веса строк (столбцов) проверочной матрицы LDPC-кода принимают различные значения, то такой код называют нерегулярным.

2. Марковские случайные поля, описываемые кодами на графах. Марковское случайное поле (МСП) это - графовая модель, в которой множество случайных величин обладает Марковским свойством, описанным неориентированным графом. МСП широко применяются в статистической физике и обработке изображений [4, 5]. Например, МСП применяются для сглаживания, сегментации, восстановления, регистрации изображений, синтеза текстур, повышения разрешения, согласования изображений в стереопарах, аннотации, извлечения информации и решения других задач. Физическим прототипом МСП является модель Изинга намагничивания материала в статистической физике. Рассмотрим гамильтонову модель Изинга Эдварда-Андерсона, [2]:

$$H_{EA} = - \sum_{i=1, \dots, n} \sum_{a=1, \dots, m} C_{ij} J_{ij} \sigma_i \sigma_j, \quad (3)$$

где C_{ij} - элемент матрицы связности, равный 1, если два спина взаимодействуют, или 0 в противном случае, J_{ij} - вес взаимодействия между i -ым и j -ым спинами, σ_i - спин, n - количество столбцов, m - количество строк. Величины J_{ij} определяют силу двухспинового взаимодействия и обычно рассматриваются как независимые случайные величины с известным распределением вероятностей, например с J_0 средним значением, ΔJ^2 дисперсией. Тогда гамильтониан H_{EA} , можно переписать в не локальную (infinite range) модель:

$$H_{ECC} = -\sum_p \sum_{i_1, \dots, i_p=1, \dots, n} C_{i_1, \dots, i_p}^{(p)} J_{i_1, \dots, i_p} \sigma_{i_1} \dots \sigma_{i_p}, \quad (4)$$

и $J_0^{(p)}$ и $\Delta J_{(p)}^2$ – математическое ожидание и дисперсия величины J_{ij} .

Рассмотрим модель Изинга, используя код на графе с заданной проверочной матрицей H (1). Проверочная матрица LDPC-кода задает матрицу C . Тогда n будет длиной кода, m – количеством проверочных уравнений, n бит кодового слова a_i определяют значения спинов $\sigma_i = 2a_i - 1$. Закодированное сообщение соответствует элементу матрицы $J_{i_1, \dots, i_p}^0 = \sigma_{i_1} \dots \sigma_{i_p}$. Декодирование кодового слова соответствует нахождению энергетического минимума гамильтониана (4). При этом J_S – выход канала (парные спиновые взаимодействия с моментами $J_0^{(p)}$ и $\Delta J_{(p)}^2$) и J^0 – шум в канале.

Для каждого p в (4) номер координаты $z_i = \sum_{j_2, \dots, j_p} C_{i, j_2, \dots, j_p} = z$ не зависит от i и скорость кода равна $R = p/z$. При $p \rightarrow \infty$ получается модель случайной энергии, при $p = 2$ получаются спиновые стекла Шеррингтона-Киркпатрика, в которых каждый спин взаимодействует со всеми остальными спинами. Для уменьшения сложности оценки основного состояния в ВР можно предположить, что каждая кодовая вершина имеет дерево «окрестностей». Такая модель короткодействующей корреляции в спиновом стекле используется в методе нарушения симметрии реплик [2] (replica symmetry breaking method).

В работе [6] было доказано, что LDPC-код может быть вложен в модель Изинга путем использования топологии тора с отрицательной кривизной (гиперболической). При этом кодовые слова (треппин-сеты $TS(a,0)$) соответствуют седловым точкам (экстремумам) в модели, а треппин-сеты ($TS(a,b)$, $b \neq 0$) соответствуют локальным минимумам. Использование LDPC-кодов с увеличенным кодовым расстоянием позволяет максимально разнести седловые точки, и таким образом повысить устойчивость нейронной сети к шуму и мощность представления (способностью нейронной сети присваивать правильные метки конкретному экземпляру и создавать четко определенные точные границы принятия решений для этого класса). При этом блочная и разряженная структура, характерная для тора отрицательной кривизны, упрощает мультиплексирование и снижает число обучаемых параметров (умножений на веса нейронной сети).

3. Приложение предложенного подхода для построения архитектуры глубокой нейронной сети (Трансформера) в задачах обработки изображений. Предложенный подход применялся для синтеза глубокой нейронной сети архитектуры «Трансформер» с 15 слоями. Каждый из слоев описывается квазициклической проверочной матрицей размера 1024×1024 , содержащей 3072 ненулевых позиций, представленной на рис. 2 (слева). Мультиплексирование осуществляется сдвиговым регистром размера 64. Глубокая нейронная сеть, состоящая из 15 слоев, представляет собой иерархическую систему с топологией вложенных торов (до 15 торов), которые могут совпадать, частично совпадать, либо различаться (рис. 2, справа), [6].

Применительно к задаче “Pathfinder” конкурса “Long Range Arena” [7] синтезированная глубокая сеть обеспечила точность бинарной классификации 94.95%, уступая только 1,72% сети, занявшей на конкурсе первое место, при значительно меньшей сложности. Число параметров (умножений) синтезированной сети меньше в более чем 5 раз.

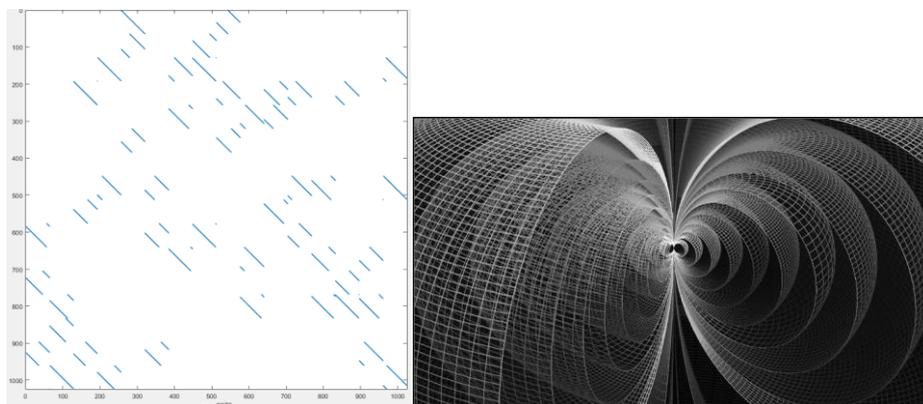


Рис. 2. Квазициклическая проверочная матрица (слева) и топология синтезированной нейронной сети (15 не пересекающихся вложенных торов (справа))

4. Приложение предложенного подхода в задачах факторизации.

Факторизация матрицы представляет ее аппроксимацию путем разложения на произведение двух или более матриц. Обычно разложение выполняется на произведение матриц меньшей размерности или на произведение разреженных матриц.

Точность матричной аппроксимации (ее ошибку) можно оценить с помощью квадрата нормы Фробениуса. Квадрат нормы Фробениуса является общей метрикой, используемой для оценки разницы между двумя матрицами и определяется как сумма квадратов разностей между соответствующими элементами двух матриц [8].

$$D(X \parallel \hat{X}) = \|X - \hat{X}\|_F^2 = \sum_{ij} (X_{ij} - \hat{X}_{ij})^2, \quad (5)$$

где X – исходная матрица и \hat{X} – ее аппроксимация.

Путем минимизации квадрата нормы Фробениуса между исходной матрицей и ее аппроксимацией можно оптимизировать факторизацию для достижения наилучшего возможного приближения исходной матрицы.

Факторизация матрицы малого ранга размерности $N \times N$ предполагает ее аппроксимацию путем разложения на произведение двух или более матриц меньшей размерности [9]: $\hat{X} \approx WH$ и $X \approx WSH$, где $W \in \mathbb{R}^{N \times r}$, $S \in \mathbb{R}^{r \times r}$, $H \in \mathbb{R}^{r \times N}$, r – ранг аппроксимации и $r \ll N$.

В работах [9, 10] показано, что минимизация квадрата нормы Фробениуса $\|X - WH\|_F^2$ относительно W и H может быть выполнена с помощью усеченного разложения по сингулярным значениям (TSVD). TSVD метод предполагает выбор диагональной матрицы Λ с r наибольшими сингулярными значениями и использование соответствующих левых и правых сингулярных векторов в качестве столбцов для формирования матриц U и V , соответственно. Тогда минимизация достигается путем выбора $W = U$ и $H = \Lambda V^T$. Масштабирующий множитель можно перенести из матрицы W в матрицу H .

Пример факторизации по методу TSVD приведен на рис. 3.

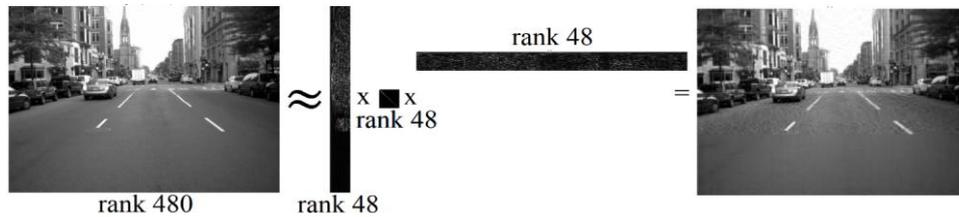


Рис. 3. Пример факторизации матрицы низкого ранга с использованием TSVD (функции svds в Matlab)

Целью малоранговой факторизации является выявление основных закономерностей и структур в больших наборах данных за счет уменьшения сложности исходной матрицы. Матричную факторизацию можно рассматривать как частный случай вложения в модель МСП (спиновых стекол) проекций низкой размерности [11]. При этом матрица раскладывается на две или более матрицы, которые представляют собой скрытые факторы или признаки. Эти скрытые факторы можно рассматривать как малоранговое вложение, которое отражает основные отношения и структуру данных. Представляя данные в виде вложения, методы матричной факторизации позволяют, среди прочего, уменьшать размерность, извлекать шаблоны и применяться в рекомендательных системах. Таким образом, факторизацию матрицы можно рассматривать как особую форму кодов вложения модели Изинга на основе кодов на графах, где матрица раскладывается на представления меньшей размерности.

Альтернативой малоранговой факторизации является факторизация матрицы путем разложения ее на произведение разреженных матриц той же размерности (SF, sparse factorization).

Для разреженной факторизации SF решается следующая задача оптимизации:

$$\min_{W^{(1)}, \dots, W^{(M)}} \|X - \prod_{m=1}^M W^{(m)}\|_F^2, \quad (6)$$

где $W^{(m)}$ – M разреженных квадратных матриц с небольшим числом ненулевых позиций. Ищутся разреженные матрицы с минимальным числом ненулевых элементов, обеспечивающие минимум квадрата нормы Фробениуса.

Пример разреженной факторизации SF Chord представлен в статье [12]. В работе для построения шаблонов разреженных матриц использовался метод хорд, основанный на протоколе P2P Chord [13]. Значения ошибки аппроксимации этого метода в сравнении с TSVD для разных типов квадратных матриц представлены в таблице.

Наш подход, использующий коды на графах, позволяет решить задачу разреженной факторизации более эффективно. В соответствии с этим подходом шаблоны разреженных матриц строятся как матрицы проверок на четность LDPC-кодов, оптимизированных с использованием PEG+ACE методов [14, 15] и методов улучшения спектра связности и кодового расстояния, предложенных в работах [16–19]. Этот метод факторизации назовем LDPC-методом.

Оптимизационная задача (6) при наличии шаблона в виде проверочной матрицы может быть решена любым методом неограниченной гладкой оптимизации. Нами использовался метод Бройдена-Флетчера-Гольдфарба-Шанно (BFGS). Минимизация ошибки аппроксимации дает в результате факторизующие матрицы $W^{(1)}, \dots, W^{(M)}$.

5. Экспериментальное исследование метода факторизации с использованием LDPC-кодов. Экспериментальное исследование предложенного метода факторизации с использованием LDPC-кодов проводилось в системе Matlab.

В рамках исследования строились шаблоны разреженных матриц $W^{(M)}$ на основе проверочной матрицы LDPC-кода. Затем решалась оптимизационная задача (6), в результате которой получались сами факторизующие матрицы и значение ошибки аппроксимации.

Полученные результаты сравнивались по точности аппроксимации с результатами, полученными для методов TSVD и SF Chord [12]. Сравнение выполнялось для одинаковых квадратных матриц из [12]. Одинаковая вычислительная сложность факторизации обеспечивалась следующим выбором параметров: для TSVD $r = \lceil (\log_2 N)^2 / 2 \rceil$, $q = 2Nr + r$; для SF и LDPC $q = N(\log_2 N)^2$ (q - общее количество ненулевых элементов для шаблонов).

LDPC шаблоны строились с использованием PEG-метода с оптимизацией спектра связности (ACE), имитацией отжига с оптимизацией степени внешнего сообщения точного цикла (EMD) [16–19] для QC-LDPC кода.

Для решения задачи (6) использовался `fminunc`-оптимизатор Matlab. Ненулевые элементы факторизующих матриц были инициализированы случайными числами в диапазоне $[K^{-1}, K^{-1} + 10^{-2}]$, $K = \log_2 N$.

`fminunc`-оптимизатор Matlab выполнялся на рабочей станции с процессором AMD Ryzen 9 3950X @ 4,0 ГГц и 128 ГБ DDR4 2400 МГц (двухканальный Kingston SK Hynix) системной памяти. Задача непараметрической оптимизации ковариации «Mfeat» с использованием BFGS Matlab, для которой требуется 137 ГБ ОЗУ, решалась на сервере с 2-мя процессорами Intel Xeon E5-2696 V4 с 256 ГБ DDR4 2133 МГц (8 каналов Cisco SK Hynix). Исходные коды использованного программного обеспечения доступны в репозитории github [20].

В качестве аппроксимированных матриц использовались изображения в оттенках серого размером 256×256 . Рис. 4 показывает шесть квадратных матриц типовых изображений, а также результирующие ошибки аппроксимации методов TSVD, SF Chord и LDPC под каждым изображением. Оценивая точность различных методов на различных типах наборов данных, можно получить представление о наиболее эффективных подходах к разреженной факторизации больших квадратных матриц.

Шахматный образ близок к низкочастотному, поскольку представляет собой черно-белую шахматную доску, имеющую всего два цвета. Поэтому TSVD работает лучше всех для этого изображения. Однако для других изображений, содержащих богатые высокочастотные детали, такие как линии и углы, TSVD не так хорош, как LDPC. Чтобы убедиться в этом, были вычислены величины градиента изображений (показано на рис. 4 справа). Матрицы градиентов, интерпретируемые как изображения, имеют нули в областях, соответствующих областям исходного изображения с постоянной интенсивностью. Ненулевые значения матриц градиентов представляют в основном высокочастотные детали исходных изображений.

Из рис. 4 видно, что TSVD дает большую ошибку аппроксимации для всех матриц градиентов, чем методы разреженной аппроксимации. Для 9 из 12 матриц изображений лучшую точность обеспечивает наш метод в сравнении с SF Chord и TSVD.

Ошибки аппроксимации по квадрату нормы Фробениуса с использованием TSVD и SF Chord [12], LDPC PEG+ACE [14,15], QC-LDPC SA+EMD [16–19] приведены в табл. 1.

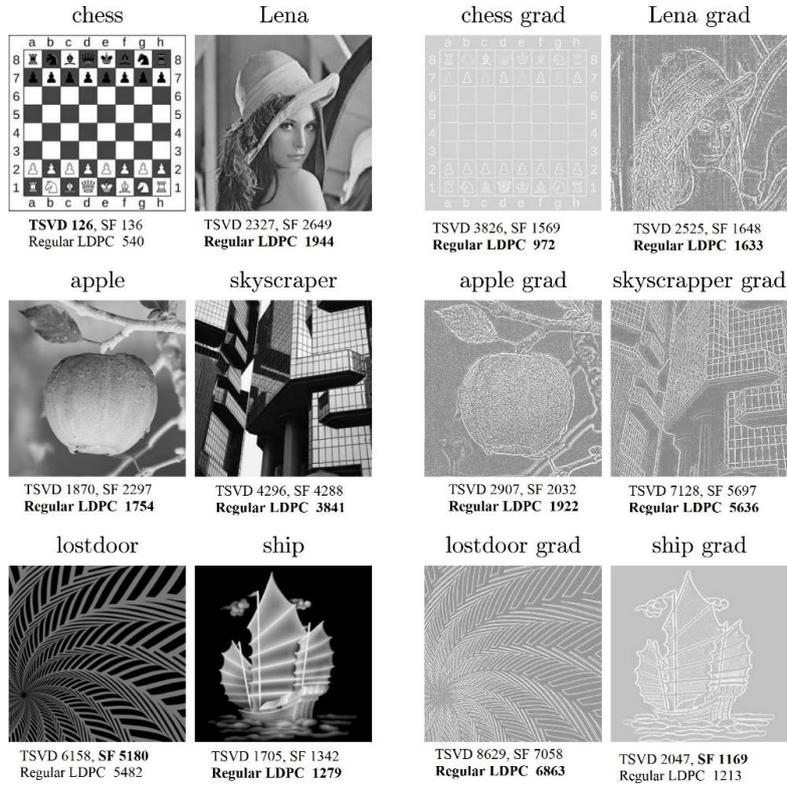


Рис. 4. Примеры квадратных матриц: исходные изображения (слева) и изображения величины градиента (справа) (отображаются после выравнивания гистограммы для лучшей наглядности)

Таблица 1

Величина ошибки аппроксимации в метрике Фробениуса различных методов факторизации квадратных матриц [6], Приложение А) (TSVD, SF Chord, LDPC)

Тип данных квадратной матрицы	Название матрицы	TSVD	SF Chord	LDPC PEG+ACE	QC-LDPC SA+EMD (MET)
Плотный граф	AuraSonar	8.54E+00	8.68E+00	7.01E+00	-
Плотный граф	Protein	1.17E+01	1.09E+01	0.74E+01	-
Плотный граф	Yeast	3.72E+01	3.61E+01	3.15E+01	-
Плотный граф	Voting	8.07E-04	1.71E+01	1.47E-02	2.1E-02
Сеть	Sawmill	3.24E+00	1.03E+00	0.26E+00	-
Сеть	Scotland	5.90E+00	3.76E+00	3.99E+00	2.42E+00
Сеть	A99 m	1.47E+01	1.01E+01	1.04E+01	0.99E+01
Сеть	Mexican Power	3.85E+00	1.71E+00	0.51E+00	-
Сеть	Strike	2.73E+00	1.04E+00	0.20E+00	-
Сеть	Webkb Cornell	6.98E+00	4.80E+00	5.48E+00	4.36E+00
Сеть	Worldtrade	8.65E+04	4.47E+04	3.54E+04	-
Поверхн. сетка	Mesh1e1	1.87E+01	9.82E+00	8.45E+00	-
Поверхн. сетка	Mesh2e1	2.48E+02	3.47E+02	1.86E+02	-
Поверхн. сетка	OrbitRaising	9.37E+01	8.53E+01	7.83E+01	-

Окончание табл. 1

Поверхн. сетка	Shuttle Entry	2.73E+03	1.86E+03	1.85E+03	-
Поверхн. сетка	AntiAngiogenesis	5.85E+01	3.29E+01	4.36E+01	2.03E+01
Ковар. матрица	Phoneme	2.80E+01	5.27E+01	2.34E+01	-
Ковар. матрица	MiniBooNE	1.04E+00	6.36E+03	4.81E+05	1.90E+03
Ковар. матрица	Coverttype	8.22E-02	1.90E-02	1.50E-02	-
Ковар. матрица	Mfeat	1.11E+03	4.01E+05	3.36E+03	0.72E+03
Ковар. матрица	OptDigits	3.28E+01	7.01E+01	1.23E+01	-
Ковар. матрица	PenDigits	4.00E+02	1.87E+02	6.00E-07	-
Ковар. матрица	Acoustic	1.36E-02	1.11E-02	5.00E-03	4.55E-03
Ковар. матрица	IJCNN	5.24E-02	3.03E-02	5.10E-03	-
Ковар. матрица	Spam Ham	1.07E-01	4.97E-02	4.70E-02	-
Ковар. матрица	TIMIT	9.64E+01	1.56E+02	8.85E+01	-
Ковар. матрица	Votes	4.00E-01	1.70E-01	2.12E-05	-

Помимо квадратных матриц изображений было выполнено сравнение применения методов TSVD, SF Chord и LDPC для других типов квадратных матриц. В табл. 1 показаны результаты сравнения. Типы данных включают матрицы плотных графов (“Плотный граф” в таблице), матрицу разреженных сетей (“Сеть”), матрицу поверхностной сетки по трехмерным объектам (“Поверхн. сетка”) и матрицу ковариации векторных данных (“Ковар. матрица”).

Типы данных описаны в приложение А статьи [6].

Ниже приведены проверочные матрицы H для задач «Антиангиогенез» (циркулянт $z = 205$) (7) и “Webkb Cornell” (циркулянт $z = 65$) (8), созданные методом QC-LDPC SA+EMD, [19]:

$$H_1 = (I^0 + I^2 + I^3 + I^{65} + I^{70} + I^{85} + I^{97} + I^{154}), \quad (7)$$

$$H_2 = \begin{pmatrix} I^1 + I^{26} + I^{50} & I^2 + I^{19} + I^{49} & I^5 + I^{13} + I^{42} \\ I^5 + I^{58} & I^5 + I^{60} & I^4 + I^{60} \\ I^4 + I^{18} + I^{48} & I^{23} + I^{28} + I^{61} & I^1 + I^4 + I^{53} \end{pmatrix}. \quad (8)$$

Из табл. 1 видно, что точность реконструкции в метрике Фробениуса метода LDPC выше почти для всех матриц (на отдельных задачах на 8 порядков) в сравнение с методами усеченного сингулярного разложения TSVD и хордовой разреженной факторизации SF Cord. Отметим, что использование квазициклических LDPC -кодов упрощает мультиплексирование.

Заключение. В работе представлен новый подход, позволяющий осуществлять синтез архитектур нейронных сетей на основе кодов на графах. Предложенный подход позволил синтезировать глубокую нейронную сеть (Transformer), обеспечивающую точность бинарной классификации 94.95% (1,72% до первого места) для задачи “Pathfinder” конкурса “Long Range Arena”, при более, чем 5-кратно меньшем числе параметров (умножений). Применение предложенного подхода к задачам факторизации на плотных графах, сетевых задачах, поверхностных сетках, ковариационных матрицах позволило увеличить точность реконструкции по метрике Фробениуса в отдельных задачах на более чем 8 порядков.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Tanner R. A recursive approach to low complexity codes // IEEE Transactions on Information Theory. – 1981. – Vol. 27, No. 5. – P. 533-547.
2. Mézard M., Montanari A. Information, Physics, and Computation. – Oxford Graduate Texts, 2009. – 569 p.
3. Tanner R., et al. LDPC block and convolutional codes based on circulant matrices // IEEE Transactions on Information Theory. – 2004. – Vol. 50, No. 12. – P. 2966-2984.
4. Сукар Л.Э. Вероятностные графовые модели. Принципы и приложения. – М.: ДМК Пресс, 2021. – 338 с.

5. *Stan Z. Li*. Markov Random Field Modeling in Image Analysis. – Springer London, 2009. – 362 p.
6. *Usatyuk V., Sapozhnikov D., Egorov S.* Spherical and Hyperbolic Toric Topology-Based Codes on Graph Embedding for Ising MRF Models: Classical and Quantum Topology Machine Learning. – 2023. – 71 p. – <https://arxiv.org/abs/2307.15778>.
7. *Yi T., et al.* Long Range Arena: A Benchmark for Efficient Transformers. – ICLR, 2021. – 15 p.
8. *Коэн М.И.* Прикладная линейная алгебра для исследователей данных. – М.: ДМК Пресс, 2023. – 328 с.
9. *Eckart C., Young G.* The Approximation of One Matrix by Another of Lower Rank // *Psychometrika*. – 1936. – Vol. 1, No. 3. – P. 211-218.
10. *Fomin F. V., et al.* Approximation schemes for low-rank binary matrix approximation problems // *ACM Transactions on Algor.* – 2019. – Vol. 16, No. 1. – P. 12:1-12:39.
11. *Camilli F., Mezard M.* Matrix factorization with neural networks // *American Physical Society, Phys. Rev. E*. – 2023. – Vol. 107, No. 6.
12. *Khalitov R., et al.* Sparse factorization of square matrices with application to neural attention modeling // *Neural Networks*. – 2022. – Vol. 152. – P. 160-168.
13. *Stoica I., et al.* Chord: A scalable peer-to-peer lookup service for internet applications // *ACM SIGCOMM Computer Communication Review*. – 2001. – Vol. 31, No 4. – P. 149-160.
14. *Hu Xiao-Yu, Eleftheriou E., Arnold D. M.* Regular and irregular progressive edge-growth Tanner graphs // *IEEE Transactions on Information Theory*. – 2005. – Vol. 51, No. 1. – P. 386-398.
15. *Tian Tao, Jones C. R., Villasenor J. D., Wesel R. D.* Selective avoidance of cycles in irregular LDPC code construction // *IEEE Transactions on Communications*. – 2004. – Vol. 52, No. 8. – P. 1242-1247.
16. *Усатюк В.С., Егоров С.И.* Построение квазициклических недвоичных низкоплотностных кодов на основе совместной оценки их дистантных свойств и спектров связности // *Телекоммуникации*. – 2016. – № 8. – С. 32-40.
17. *Усатюк В.С., Егоров С.И.* Устройство для оценки кодового расстояния линейного блочного кода методом геометрии чисел // *Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение*. – 2017. – № 4 (25). – С. 24-33.
18. *Usatyuk V., Egorov S., Svistunov G.* Construction of Length and Rate Adaptive MET QC-LDPC Codes by Cyclic Group Decomposition // *2019 IEEE East-West Design & Test Symposium (EWDTS)*. – Batumi, Georgia, 2019. – P. 1-5.
19. *Usatyuk V., Vorobyev I.* Simulated Annealing Method for Construction of High-Girth QC-LDPC Codes // *Intern. Conf. on Telecom. and Signal Proc.* – 2018. – P. 1-5.
20. *Usatyuk V.S.* Matlab implementation of TSVD, SF Chord, LDPC PEG+ACE, QC-LDPC SA+EMD (MET) codes sparse factorization. – URL: <https://github.com/Lcrypto/Classical-and-Quantum-Topology-ML-toric-spherical>.

REFERENCES

1. *Tanner R.* A recursive approach to low complexity codes, *IEEE Transactions on Information Theory*, 1981, Vol. 27, No. 5, pp. 533-547.
2. *Mézard M., Montanari A.* Information, Physics, and Computation. Oxford Graduate Texts, 2009, 569 p.
3. *Tanner R., et al.* LDPC block and convolutional codes based on circulant matrices, *IEEE Transactions on Information Theory*, 2004, Vol. 50, No. 12, pp. 2966-2984.
4. *Sukar L.E.* Veroyatnostnye grafovye modeli. Printsipy i prilozheniya [Probabilistic graph models. Principles and applications]. Moscow: DMK Press, 2021, 338 p.
5. *Stan Z. Li*. Markov Random Field Modeling in Image Analysis. Springer London, 2009, 362 p.
6. *Usatyuk V., Sapozhnikov D., Egorov S.* Spherical and Hyperbolic Toric Topology-Based Codes on Graph Embedding for Ising MRF Models: Classical and Quantum Topology Machine Learning, 2023, 71 p. Available at: <https://arxiv.org/abs/2307.15778>.
7. *Yi T., et al.* Long Range Arena: A Benchmark for Efficient Transformers. ICLR, 2021, 15 p.
8. *Koen M.I.* Prikladnaya lineynaya algebra dlya issledovateley dannykh [Applied linear algebra for data scientists]. Moscow: DMK Press, 2023, 328 p.
9. *Eckart C., Young G.* The Approximation of One Matrix by Another of Lower Rank, *Psychometrika*, 1936, Vol. 1, No. 3, pp. 211-218.

10. Fomin F. V., et al. Approximation schemes for low-rank binary matrix approximation problems, *ACM Transactions on Algor.*, 2019, Vol. 16, No. 1, pp. 12:1-12:39.
11. Camilli F., Mezard M. Matrix factorization with neural networks, *American Physical Society, Phys. Rev. E.*, 2023, Vol. 107, No. 6.
12. Khalitov R., et al. Sparse factorization of square matrices with application to neural attention modeling, *Neural Networks*, 2022, Vol. 152, pp. 160-168.
13. Stoica I., et al. Chord: A scalable peer-to-peer lookup service for internet applications, *ACM SIGCOMM Computer Communication Review*, 2001, Vol. 31, No 4, pp. 149-160.
14. Hu Xiao-Yu, Eleftheriou E., Arnold D. M. Regular and irregular progressive edge-growth Tanner graphs, *IEEE Transactions on Information Theory*, 2005, Vol. 51, No. 1, pp. 386-398.
15. Tian Tao, Jones C. R., Villasenor J. D., Wesel R. D. Selective avoidance of cycles in irregular LDPC code construction, *IEEE Transactions on Communications*, 2004, Vol. 52, No. 8, pp. 1242-1247.
16. Usatyuk V.S., Egorov S.I. Postroenie kvazitsiklicheskih nedvoichnykh nizkoplotnostnykh kodov na osnove sovместной otsenki ikh distantnykh svoystv i spektrov svyaznosti [Construction of quasi-cyclic non-binary low-density codes based on a joint assessment of their distant properties and connectivity spectra], *Telekommunikatsii* [Telecommunications], 2016, No. 8, pp. 32-40.
17. Usatyuk V.S., Egorov S.I. Ustroystvo dlya otsenki kodovogo rasstoyaniya lineynogo blochnogo koda metodom geometrii chisel [Device for estimating the code distance of a linear block code using the geometry of numbers method], *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta. Seriya: Upravlenie, vychislitel'naya tekhnika, informatika. Meditsinskoe priborostroenie* [News of the South-West State University. Series: Management, computer technology, computer science. Medical instrumentation], 2017, No. 4 (25), pp. 24-33.
18. Usatyuk V., Egorov S., Svistunov G. Construction of Length and Rate Adaptive MET QC-LDPC Codes by Cyclic Group Decomposition, *2019 IEEE East-West Design & Test Symposium (EWDTS)*. Batumi, Georgia, 2019, pp. 1-5.
19. Usatyuk V., Vorobyev I. Simulated Annealing Method for Construction of High-Girth QC-LDPC Codes, *Intern. Conf. on Telecom. and Signal Proc.*, 2018, pp. 1-5.
20. Usatyuk V.S. Matlab implementation of TSVD, SF Chord, LDPC PEG+ACE, QC-LDPC SA+EMD (MET) codes sparse factorization. Available at: <https://github.com/Lcrypto/Classical-and-Quantum-Topology-ML-toric-spherical>.

Статью рекомендовал к опубликованию д.т.н., профессор В.М. Курейчик.

Усатюк Василий Станиславович – ООО «Т8»; e-mail: l@lcrypto.com; г. Москва, Россия; к.т.н.

Егоров Сергей Иванович – Юго-Западный государственный университет; e-mail: sie58@mail.ru; г. Курск, Россия; д.т.н.; доцент; профессор кафедры вычислительной техники.

Локтионов Аскольд Петрович – e-mail: loapa@mail.ru; д.т.н.; доцент; старший научный сотрудник кафедры уникальных зданий и сооружений.

Титенко Евгений Анатольевич – e-mail: johntit@mail.ru; тел.: +79051588904; кафедра программной инженерии; к.т.н.; доцент.

Чернецкая Ирина Евгеньевна – e-mail: white@mail.ru; кафедра вычислительной техники; зав. кафедрой; д.т.н.; доцент.

Usatyuk Vasily Stanislavovich – T8 LLC; e-mail: l@lcrypto.com; Moscow, Russia; cand. of eng. sc.

Egorov Sergey Ivanovich – South-West State University; e-mail: sie58@mail.ru; Kursk, Russia; dr. of eng. sc.; associate professor; professor of the department of computer science.

Loktionov Askold Petrovich – e-mail: loapa@mail.ru; dr. of eng. sc.; associate professor; senior researcher at the department of unique buildings and structures.

Titenko Evgeny Anatolievich – e-mail: johntit@mail.ru; phone: +79051588904; the department of computer science cand. of eng. sc.; associate professor.

Chernetskaya Irina Evgenievna – e-mail: white@mail.ru; the department of computer science; head of. department; dr. of eng. sc.; associate professor.

Я.В. Куликова, В.А. Литвиненко

РАЗРАБОТКА МЕТОДА ОПРЕДЕЛЕНИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА С ПОМОЩЬЮ ЧАТ-БОТА

В настоящее время практически во всех сферах деятельности интенсивно применяются диалоговые системы (чат-боты), при этом стоит обратить внимание, что данные механизмы начали использовать не только в сфере развлечений, но и в более важных сферах – бизнес или даже медицина. В связи с этим появляется необходимость развития и совершенствования системы общения таких роботов на естественном языке. В данной статье предлагается метод определения эмоционального состояния человека с помощью чат-бота. В статье даётся определение чат-боту, рассматриваются виды чат-ботов, определяются основные принципы их работы. На основе выявленных различий в работе алгоритмов по определению текста чат-ботами описывается наиболее подходящая технология для решения поставленной задачи – работа с нейронными сетями. Выявлены и описаны недостатки существующих решений по сравнению с разрабатываемым методом. Важно отметить, что особенностью метода является соединение технологий интеллектуальных чат-ботов и дополненной реальности. В статье описан общий алгоритм работы метода определения эмоционального состояния человека с помощью чат-бота, включающий восемь основных этапов. Описаны разработанные и выделенные критерии оценки эмоций: валентность, интенсивность и стеничность. Представлена архитектура разработанного решения, которая имеет клиент-серверный подход. Серверная часть включает в себя также модуль определения тональности текста и модуль формирования ответа, в то время как клиент включает в себя чат и внешний вид «Аватара» разработанного чат-бота. Описаны основные шаги реализации, среди которых представлен первоначальный интерфейс, а также этапы разработки приложения дополненной реальности. В статье перечислены используемые средства для работы с дополненной реальностью (Unity 3D и фреймворк AR Vuforia). Представлены созданные 3D-модели. Показаны результаты обработки скелетной анимации при помощи оборудования Perception Neuron 3 и последующей обработки в Axis Studio. По итогам работы было сформулировано заключение, описывающее качественное отличие разработанной системы по отношению к существующим рассмотренным аналогам, а также приведены результаты проверки работы системы.

Чат-бот; нейронная сеть; тональность текста; эмоции.

Ya.V. Kulikova, V.A. Litvinenko

DEVELOPMENT OF A METHOD FOR DETERMINING THE EMOTIONAL STATE OF A PERSON USING A CHAT BOT

Currently, dialogue systems (chatbots) are being intensively used in almost all areas of activity, and it is worth noting that these mechanisms have begun to be used not only in the entertainment sector, but also in more important areas - business or even medicine. In this regard, there is a need to develop and improve the communication system of such robots in natural language. This article proposes a method for determining a person's emotional state using a chatbot. The article defines a chatbot, discusses the types of chatbots, and defines the basic principles of their work. Based on the identified differences in the operation of algorithms for text detection by chatbots, the most suitable technology for solving the task is described - working with neural networks. The shortcomings of existing solutions in comparison with the method being developed are identified and described. It is important to note that the peculiarity of the method is the combination of technologies of intelligent chatbots and augmented reality. The article describes the general algorithm of the method for determining the emotional state of a person using a chatbot, which includes eight main stages. The developed and identified criteria for assessing emotions are described: valence, intensity and sthenicity. The architecture of the developed solution, which has a client-server approach, is presented. The server part also includes a text sentiment detection

module and a response generation module, while the client includes the chat and the appearance of the "Avatar" of the developed chatbot. The main implementation steps are described, including the initial interface, as well as the stages of developing an augmented reality application. The article lists the tools used for working with augmented reality (Unity 3D and the AR Vuforia framework). The created 3D models are presented. The results of processing skeletal animation using Perception Neuron 3 equipment and subsequent processing in Axis Studio are shown. Based on the results of the work, a conclusion was formulated describing the qualitative difference of the developed system in relation to the existing analogues considered, and the results of testing the operation of the system were also presented.

Chatbot; neural network; text sentiment; emotions.

Введение. В современном мире все сферы жизнедеятельности связаны с внедрением информационных технологий и автоматизации рабочих процессов. Предприятия, не придерживающиеся современных тенденций, обречены на полную потерю интереса к их деятельности, поэтому им необходимо развиваться. Одним из потенциальных направлений такого развития является внедрение интеллектуальных систем, помогающих в работе, например, чат-бота как электронного помощника или консультанта в различных видах деятельности.

Чат-бот – виртуальная программа-робот собеседник. Данные программы наиболее часто используются в бизнес сфере или сфере развлечений в качестве «консультантов», помогающих наиболее быстро найти нужную для человека информацию.

Авторами предлагается описание разработанного чат-бота, распознающего текст сообщения, определяющего его эмоцию и осуществляющего подбор подходящего сообщения и соответствующего «аватара» в качестве ответа.

Актуальность темы подтверждается не только востребованностью данной технологии, но и универсальностью предлагаемого решения, так как может быть востребовано в совершенно разных сферах деятельности: от сферы развлечения до сложно рекомендательных систем в образовании. Разработка чат-бота, на данный момент, является почти необходимостью для многих организаций и пользователей, что подтверждается различными социологическими исследованиями [1, 2].

Автором были изучены основные методы автоматического определения тональности текста: с использованием словарей или на основе нейронных сетей [3, 19, 20]. Нейронные сети дают возможность обрабатывать нелинейные алгоритмы управления при недостаточном, неточном описании объекта (или даже при отсутствии описания), создавать устойчивую систему при нестабильных параметрах [4, 5]. Учитывая данную информацию, а также то, что важным критерием для разрабатываемого бота является обучаемость, наиболее подходящим методом является использование нейросети. Дополнительно, можно отметить, что формирование комфортного диалога для собеседника возможно в том случае, когда к собеседнику-роботу предъявлены следующие требования: чат-боты должны использовать естественный язык [6–8], должны быть эмоционально грамотными, должны уметь анализировать [9, 10], а также быть ориентированы на пользователя [11, 12]. Для реализации всех требований наиболее оптимальным вариантом является использование нейросети [13–15].

Важно отметить, что чат-бот – это, как правило, собеседник, которого не видно человеку, что может вызывать трудности при общении у некоторого количества людей. Поэтому, для достижения эффекта присутствия собеседника авторами было выбрано использование технологии дополненной реальности [16]. Данная технология позволит создать визуальную часть – 3-D «аватара», с которым и происходит общение собеседника чат-бота.

Таким образом, в работе описывается новый метод, который представляет собой объединение технологий интеллектуальный ботов и дополненной реальности, а также эмоциональной окраски текста сообщений.

Постановка задачи. Автором были выделены следующие задачи:

- ◆ анализ и определение основных проблем существующих решений;
- ◆ выделение основных критериев обработки текста для определения эмоционального анализа текста;
- ◆ проведение подробного анализа построения последовательной обработки сообщений;
- ◆ разработка алгоритмов работы метода и анализа, учитывающего выделенные критерии оценки;
- ◆ построение архитектуры системы;
- ◆ описание разработанного приложения.

Для оценки эмоций автором были выделены следующие основные критерии: валентность, интенсивность и стеничность.

Для определения валентности предлагается задать вес каждого положительного и отрицательного коэффициента тона и определить результат как разницу средних арифметических взвешенных сумм положительных и отрицательных коэффициентов.

Интенсивность сообщений предлагается считать как суммы средневзвешенных средних всех коэффициентов тональности каждого сообщения, деленное на количество этих сообщений.

Эмоции делятся на стенические и астенические. Стенические – положительные, астенические – отрицательные. Следовательно, отношение отрицательных к положительным эмоциям даёт стеничность текста.

Предлагаемое решение обработки потока сообщений пользователей чат-ботом. На данный момент для множества организаций использование мессенджеров и чат-ботов стало не просто удобством, а практической необходимостью для полноценной и эффективной работы. Согласно исследованию Accenture Digital [2] чат-боты позволяют более эффективно решать задачи сотрудников, ускоряют обработку запросов от клиентов, обеспечивают более быстрый и персонализированный ответ, дают возможность быстрее привлекать большее количество клиентов.

Но при использовании чат-ботов выявляются определённые проблемы [17–19].

Первое, все имеющиеся методы рассматривают обработку либо голоса, либо видео и редко учитывают обработку текста.

Второе, и самое главное, общение с ботом (переписка с ботом) не заменяет общение с живым человеком, потому что чат-бот не анализирует эмоциональную составляющую диалогов.

Если рассмотреть существующие решения по определению эмоциональной окраски текстовой информации можно отметить, что используется группа методов анализа тональности текста, однако у этих методов есть ряд ограничений:

- ◆ наиболее эффективны для большой выборки (немаленьких текстов);
- ◆ чаще всего это бинарный показатель – анализирует процент позитива и негатива;
- ◆ анализируют весь текст целиком;
- ◆ предобработка текста очищает от дополнительных слов (предлогов, частиц, союзов).

Поскольку больший интерес представляет многокритериальная оценка текста и последовательный анализ группы сообщений, а не целостного большого текста, так как разрабатываемый метод применяется в рамках внедрения концепции чат-

ботов, то для оценки эмоционального состояния пользователя необходимо использовать дополнительные инструменты, предоставляющие возможность произвести эту оценку.

Сравнение методов обработки сообщений представлено на рис. 1.



Рис. 1. Сравнение классических методов анализа тональности текста и метода оценки эмоционального состояния пользователя чат-бота

В работе автором предлагается метод последовательной обработки сообщений, позволяющий:

- ◆ составить общий анализ группы сообщений;
- ◆ автоматически связывать сообщения в группы для комплексного анализа;
- ◆ определить эмоциональное состояние пользователя, а также динамику изменения и поведения пользователя.

◆ сформировать метаинформацию на основе групп сообщений (таким образом бот может не хранить историю переписки, а только ключевые показатели, на основе которых впоследствии можно будет принимать решение о динамике изменения эмоционального состояния).

Анализ сообщений пользователя. Сообщения пользователя формируют поток коротких текстов. Сообщения распределены по времени. Одним из вариантов обработки потока сообщений может выступать метод последовательной группировки сообщений, основанный на времени прихода сообщений и поиске N ближайших соседей. Однако группировка по времени подходит для статистической обработки сообщений, но не подходит для выявления взаимосвязанных цепочек сообщений, поэтому для последовательного анализа сообщений и автоматической группировки их с целью анализа групп сообщений предложен алгоритм анализа и группировки, основанный на механизме нейронных сетей и их последовательного применения [20].

Таким образом, анализ всего диалога сводится к последовательной обработке поступающих сообщений.

Общий алгоритм работы метода представлен на рис. 2.

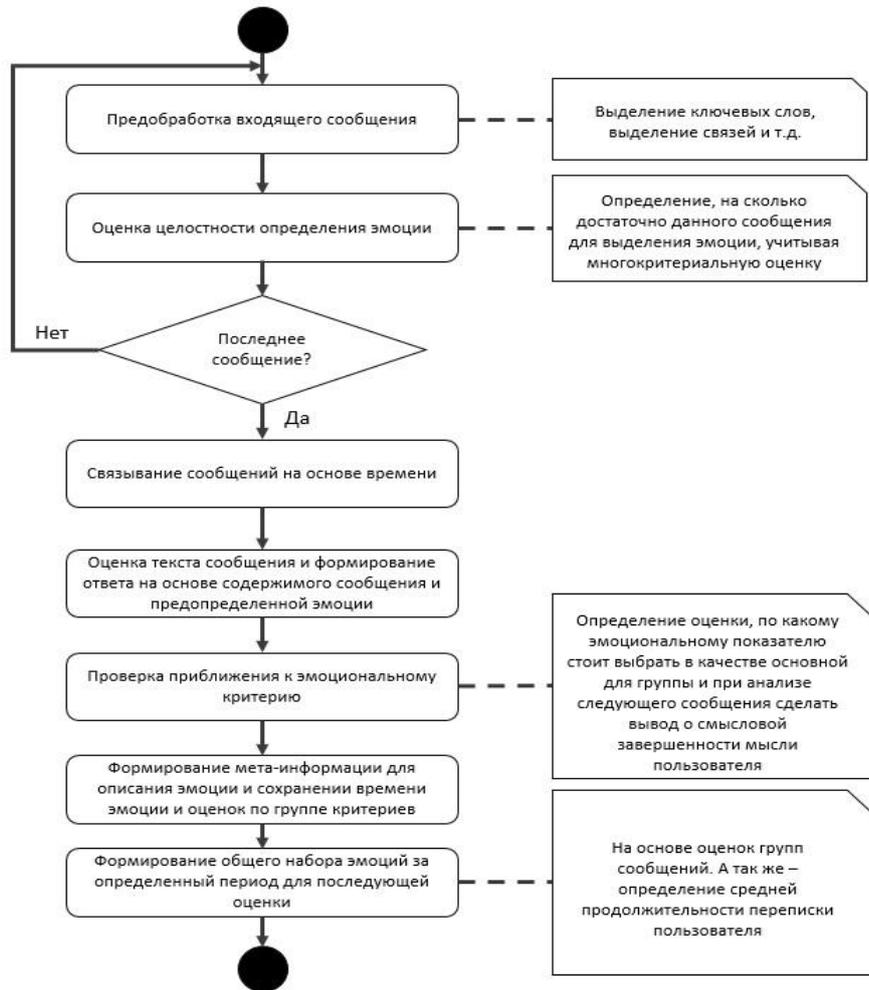


Рис. 2. Общий алгоритм работы метода

Данный метод включает в себя 8 основных этапов. Среди них имеются следующие: предварительная обработка, работа с ключевыми словами и связями; оценка целостности; проработка первых двух шагов для каждого последующего сообщения; проведение критериальной оценки и анализа сообщений; подготовка ответа; анализ смысловой завершенности сообщений; формирование мета-информации для завершенной группы сообщений; подготовка общего комплекта эмоций.

Алгоритм и схема работы нейронной сети. Нейронная сеть, лежащая в основе многофакторной классификации, работает комплексно и обрабатывает последовательно поступающую информацию в порядке обработки «смайлов», общих выражений, специальных конструкций, затем выделение особенных слов, очистка текста, классификация на основе найденных слов, а затем выделение ключевых слов и конструкций и классификация для конкретной эмоции или определения смещения к эмоциональным показателям в многокритериальном пространстве.

Алгоритм работы нейросети представлен на рис. 3.

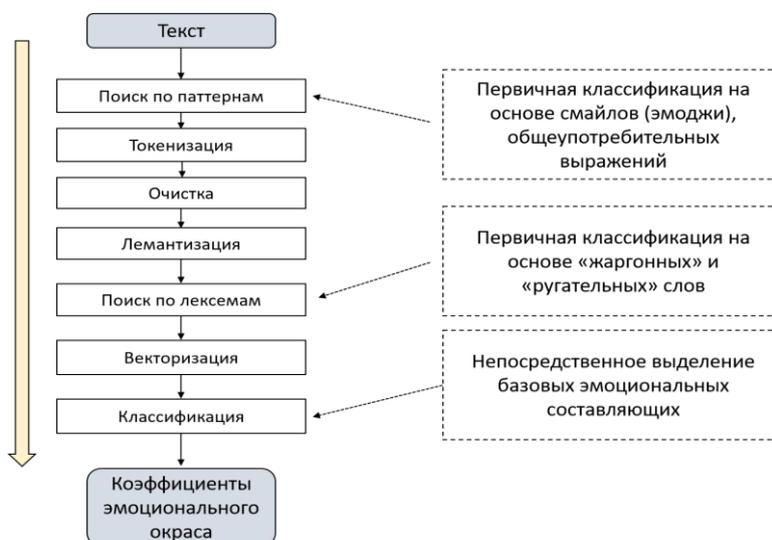


Рис. 3. Общий алгоритм работы нейронной сети оценки смещения в многокритериальном пространстве эмоциональных показателей

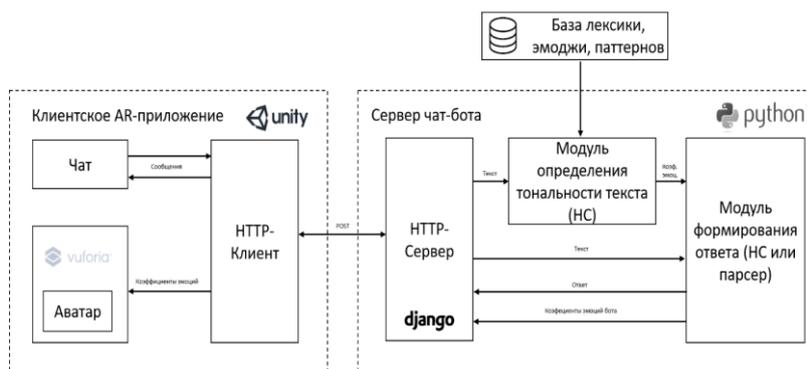


Рис. 4. Архитектура системы

На рис. 4 представлена первоначальная архитектура системы, позволяющая организовать поддержку работы интеллектуального бота. Данная архитектура основывается на классическом клиент-серверной подходе, при этом основные элементы обработки данных работают в рамках серверной части. Также серверная часть несёт ответственность за организацию многопользовательской многопоточной работы в системе. Существенной частью системы является модуль хранения информации, осуществляющий связывание текстовых сообщений и корректно описывающий изменение эмоционального фона пользователя во времени, а также хранящий информацию об эмоциональной тональности всей беседы. Последовательная обработка входящих сообщений увеличивает нагрузку на сервер и на систему коммуникаций. Однако необходима для того, чтобы корректно обрабатывать изменения во времени эмоционального фона пользователя и связывание входящих сообщений в группы сообщений по возможности выделения единой эмоции из этой группы, а также в беседе по набору определённых эмоций и по времени общения с ботом.

Одним из ключевых моментов разрабатываемого метода является формирование унифицированного интерфейса для постобработки метаинформации об эмоциональном состоянии пользователя. Данный интерфейс позволяет другим системам, программам и автоматизированным средствам в режиме реального времени или в режиме постобработки получать метаинформацию и использовать эти данные для дальнейшего прогнозирования поведения пользователя, а также комплексной оценки его эмоционального состояния при работе в сфере, в которой происходит общение с ботом. При этом на практике данный подход может применяться в любой сфере, требующей поддержки заинтересованностью или последовательного отслеживания эмоционального состояния пользователя.

Важная проблема при разработке любого метода распознавания – это оценка результатов распознавания. Учитывая, что бот должен дообучаться и переобучаться по мере общения с пользователем, а также выстраивания диалога, необходимо формировать набор метрик которые позволят оценивать качество распознавания эмоции при общении бота с пользователем.

Условно весь процесс оценки распознавания разделён на два этапа. На первом этапе мы проводим обучение нейронной сети по принципу обучения с учителем. На этом этапе обучения точность распознавания оценивается человеком и формируется индекс качество распознавания, на основе этого индекса корректируется работа нейронной сети и формируются алгоритмы до обучения и переобучения. На втором этапе нейронная сеть обучается автоматически на основе сформированных алгоритмов и индекса качества распознавания. Второй этап продолжается всё время работы бота. Таким образом, бот является самообучаемым алгоритмом анализа текста в многофакторном пространстве.

Результаты. Для наиболее полной и комплексной оценки эмоций были выделены и разработаны критерии, такие как валентность, интенсивность, стеничность. Например, для интенсивности предлагается использовать следующую формулу, в которой представлены базовые эмоции по теории Роберта Плутчика (согласно табл. 1):

$$K_n = \frac{\sum_{i=1}^n \frac{2P_i+0,5V_i+0,5Pr_i+1D_i}{4} + \frac{2\Gamma_i+0,5C_i+0,2\Pi_i+10T_i}{3,7}}{n}$$

Первый вариант разработанной программы представлял собой интерфейс чат – бота, состоящий из поля ввода сообщения, поля отображения диалога и поля отображения картинки-эмоции чат-бота (представлен на рис. 5).

Далее были проработаны детали дополненной реальности: создана 3D-модель, проработаны наборы эмоций, модель экспортирована в Unity, задана обработка анимации. 3D-модель представлена на рис. 6.

Таблица 1

Базовые эмоции Плутчика с обозначениями для предлагаемой формулы интенсивности

№	Наименование базовой эмоции	Обозначение	Вес
	Гнев	Г	2
	Страх	С	0,5
	Печаль	П	0,2
	Радость	Р	2
	Отвращение	От	1
	Удивление	У	0,5
	Предвкушение	Пр	0,5
	Доверие	Д	1

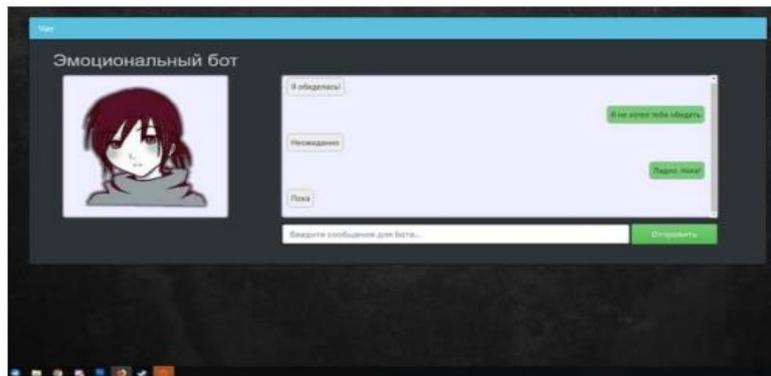


Рис. 5. Первоначальный интерфейс чат-бота



Рис. 6. 3D модель

Приложение-клиент написано на Unity с использованием фреймворка AR Vuforia (рис. 7). Создание анимаций производилось с использованием оборудования Perception Neuron 3. Обработка скелетной модели проводилась в Axis Studio, после чего переносился в Unity. Пример двигающейся модели представлен на рис. 8.



Рис. 7. 3D модель в среде Unity

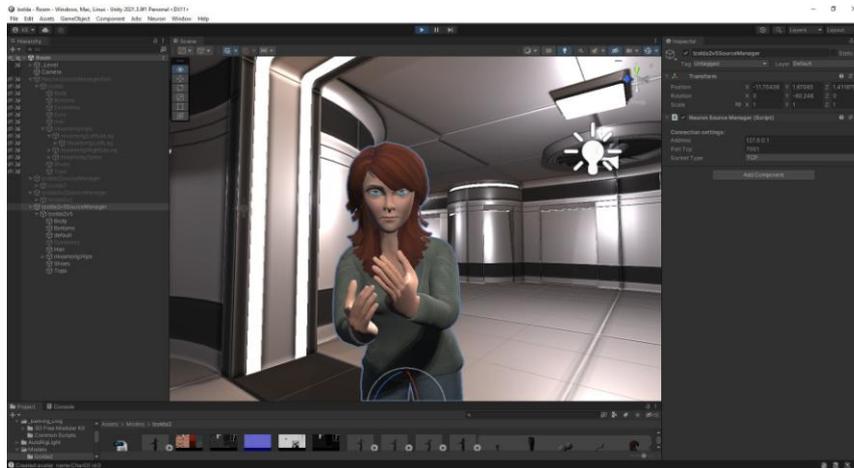


Рис. 8. «Аватар» в движении

Разработанный метод сравнивался с существующими решениями (Ассистент Google Now, Siri от компании Apple, Amazon Alexa, Microsoft Cortana и Алиса), в которых были выделены основные недостатки в виде предоставления избыточной информации, направленность на определённую ОС и, самый главный – невысокая «человечность», которая проявляется в основном функционале систем – вопросно-ответная с умением говорить некоторой базой шуток. С качественной точки зрения самое главное преимущество метод имеет именно в плане «человечности», потому что его главная задача правильно определить эмоции собеседника и корректно на них отреагировать.

В количественном отношении проверка работоспособности системы осуществлялась путем проведения тестовых экспериментов. В тестировании участвовало 120 человек.

Таблица 2

Результаты тестирования системы

Эмоция	Худший результат	Лучший результат	Достоверность
гнев	0,81	0,96	0,9
радость	0,69	0,87	0,8
отвращение	0,63	0,81	0,7
презрение	0,59	0,78	0,7
страх	0,44	0,66	0,6
удивление	0,42	0,61	0,5
печаль	0,36	0,58	0,5

Предложено было пообщаться с чат-ботом, причём с проявлением эмоций, необходимых для определения коэффициента тональности, после чего сообщить правильность определения ботом эмоций.

Заключение. Разработан новый метод, объединяющий технологии интеллектуальных чат-ботов и визуализации данных путем использования технологии дополненной реальности, а также анализа тональности текста сообщения, а не голосовых данных или изображения, отправляемого чат-боту.

Используемые технологии интеллектуальных чат-ботов, основывающиеся на последовательном применении нейронных сетей позволяют обрабатывать потоковые данные в виде цепочек пользовательских сообщений, а также определять тональность текста в сообщениях пользователя.

Метод позволяет определять тональность сообщений и формировать ответ пользователю в автоматическом режиме в зависимости от просчитанных показателей. При этом, ответ должен увеличивать эффективность восприятия информации собеседником чат-бота.

В результате выявлено, что, на данный момент, чат бот определяет эмоции со следующей точностью: радость 0.8; удивление 0.5; печаль 0.5; гнев 0.9; отвращение 0.7; презрение 0.7; страх 0.6.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Humanity in the Machine. – URL: mindshareworld.com/sites/default/files/ MINDSHARE_HUDDLE_HUMANITY_MACHINE_2016.pdf (дата обращения: 14.12.2021).
2. Digital 2022 Russian Federation // Data Sources and Methodologies. – URL: datareportal.com/ (дата обращения: 10.03.2022).
3. *Стребков Д.О., Шевчук А.В., Спирина М.О.* Самостоятельная занятость на рынке удалённой работы: распространение инновационной трудовой практики // Мониторинг общественного мнения: Экономические и социальные перемены. – 2016. – № 6. – С. 89-106.
4. *Tarasov D. [et al.]*. Language models for unsupervised acquisition of medical knowledge from natural language texts: application for diagnosis // Proceedings of the International Conference “Dialogue 2019”. – 2019. – 11 p.
5. *Razzaki S.* A comparative study of artificial intelligence and human doctors for the purpose of triage and diagnosis. – 2018.
6. *Samyak S., Gupta A., Raj T., Karnam A., Mamatha HR.* Speech Emotion Analyzer // Innovative Data Communication Technologies and Application. – 2022. – P. 113-124.
7. *Arumugam R. Shanmugamani.* Hands-On Natural Language Processing with Python. – Packt Publishing, 2018.
8. *Ryazanov VV.* Modeli, metody, algoritmy i arkhitektury sistem raspoznavaniya rechi. – Moscow: Vychislitel'nyi tsentr im. A.A. Dorodnitsyna; 2006. – 138 p. (In Russ.)
9. *Pérez J.Q., Daradoumis T., and Puig J.M.M.* Rediscovering the use of chatbots in education: A systematic literature review // Computer Applications in Engineering Education. – 2020. – Vol. 28, No. 6. – P. 1549-1565.
10. *Alm A., and Nkomo L.M.* Chatbot experiences of informal language learners: A sentiment analysis // International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT). – 2020. – Vol. 10, No. 4. – P. 51-56.
11. *Miner A., Laranjo L., Kocaballi A.B.* Chatbots in the fight against the COVID-19 pandemic // npj Digit. Med. – 2020. – Vol. 3. – P. 65.
12. *Chen H. [et al.]*. A Survey on Dialogue Systems: Recent Advances and New Frontiers // ACM SIGKDD Explorations Newsletter. Association for Computing Machinery (ACM). – 2017. – Vol. 19, No. 2. – P. 25-35.
13. *Shukla S. [et al.]*. Conversation Learner – A Machine Teaching Tool for Building Dialog Managers for Task-Oriented Dialog Systems // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations. Association for Computational Linguistics. – 2020. – P. 343-349.
14. *Xu G. [et al.]*. Optimizing Policy via Deep Reinforcement Learning for Dialogue Management // 2018 IEEE International Conference on Big Data and Smart Computing (BigComp). – IEEE, 2018. – P. 582-589.
15. *Alvaro N.E.* Implementing ChatBots using Neural Machine Translation techniques: Bachelor thesis. – Universitat Politècnica de Catalunya, 2017. – 44 p.
16. *Куликова Я.В., Куликов Е.А.* Применение технологии дополненной реальности для обучения детей с особенностями развития // Инновационные научные исследования: теория, методология, практика: Сб. статей XIII Международной научно-практической конференции: в 2 ч. Том Ч. 2. Пенза, 20 марта 2018 года. – Пенза: Наука и Просвещение (ИП Гуляев Г.Ю.), 2018. – С. 165-170. – EDN YTKTDQ.

17. Song Y. [et al.]. An Ensemble of Retrieval-Based and Generation-Based Human-Computer Conversation Systems // IJCAI. – 2018.
18. Yang L. [et al.]. A Hybrid Retrieval-Generation Neural Conversation Model // Proceedings of the 28th ACM International Conference on Information and Knowledge Management. – 2019. – P. 1341-1350.
19. Venkatesh A. [et al.]. On Evaluating and Comparing Conversational Agents // arXiv preprint arXiv:1801.03625. – 2018.
20. Serban I. [et al.]. Multiresolution Recurrent Neural Networks: An Application to Dialogue Response Generation // AACL. – 2017.

REFERENCES

1. Humanity in the Machine. Available at: mindshareworld.com/sites/default/files/MINDSHARE_HUDDLE_HUMANITY_MACHINE_2016.pdf (accessed 14 Desember 2021).
2. Digital 2022 Russian Federation, *Data Sources and Methodologies*. Available at: datareportal.com/ (accessed 10 March 2022).
3. Strebkov D.O., Shevchuk A.V., Spirina M.O. Samostoyatel'naya zanyatost' na rynke udalennoy raboty: rasprostranenie innovatsionnoy trudovoy praktiki [Self-employment in the remote work market: the spread of innovative labor practices], *Monitoring obshchestvennogo mneniya: Ekonomicheskie i sotsial'nye peremeny* [Monitoring of public opinion: Economic and social changes], 2016, No. 6, pp. 89-106.
4. Tarasov D. [et al.]. Language models for unsupervised acquisition of medical knowledge from natural language texts: application for diagnosis, *Proceedings of the International Conference "Dialogue 2019"*, 2019, 11 p.
5. Razzaki S. A comparative study of artificial intelligence and human doctors for the purpose of triage and diagnosis, 2018.
6. Samyak S., Gupta A., Raj T., Karnam A., Mamatha HR. Speech Emotion Analyzer, *Innovative Data Communication Technologies and Application*, 2022, pp. 113-124.
7. Arumugam R. Shanmugamani. Hands-On Natural Language Processing with Python. Packt Publishing, 2018.
8. Ryazanov VV. Modeli, metody, algoritmy i arkhitektury sistem raspoznavaniya rechi. Moscow: Vychislitel'nyi tsentr im. A.A. Dorodnitsyna; 2006, 138 p. (In Russ.)
9. Pérez J.Q., Daradoumis T., and Puig J.M.M. Rediscovering the use of chatbots in education: A systematic literature review, *Computer Applications in Engineering Education*, 2020, Vol. 28, No. 6, pp. 1549-1565.
10. Alm A., and Nkomo L.M. Chatbot experiences of informal language learners: A sentiment analysis, *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, 2020, Vol. 10, No. 4, pp. 51-56.
11. Miner A., Laranjo L., Kocaballi A.B. Chatbots in the fight against the COVID-19 pandemic, *npj Digit. Med.*, 2020, Vol. 3, pp. 65.
12. Chen H. [et al.]. A Survey on Dialogue Systems: Recent Advances and New Frontiers, *ACM SIGKDD Explorations Newsletter. Association for Computing Machinery (ACM)*, 2017, Vol. 19, No. 2, pp. 25-35.
13. Shukla S. [et al.]. Conversation Learner – A Machine Teaching Tool for Building Dialog Managers for Task-Oriented Dialog Systems, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations. Association for Computational Linguistics*, 2020, pp. 343-349.
14. Xu G. [et al.]. Optimizing Policy via Deep Reinforcement Learning for Dialogue Management, *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*. IEEE, 2018, pp. 582-589.
15. Alvaro N.E. Implementing ChatBots using Neural Machine Translation techniques: Bachelor thesis. Universitat Polit'ecnica de Catalunya, 2017, 44 p.
16. Kulikova Ya.V., Kulikov E.A. Primenenie tekhnologii dopolnennoy real'nosti dlya obuchenii detey s osobennostyami razvitiya [Application of augmented reality technology for teaching children with special needs], *Innovatsionnye nauchnye issledovaniya: teoriya, metodologiya, praktika: Sb. statey XIII Mezhdunarodnoy nauchno-prakticheskoy konferentsii: v 2 ch. Tom*

- СН. 2. Penza, 20 marta 2018 goda [Innovative scientific research: theory, methodology, practice: Collection of articles of the XIII International Scientific and Practical Conference: in 2 parts. Volume Part 2. Penza, March 20, 2018]. Penza: Nauka i Prosveshchenie (IP Gulyaev G.YU.), 2018, pp. 165-170. EDN YTKTDQ.
17. Song Y. [et al.]. An Ensemble of Retrieval-Based and Generation-Based Human-Computer Conversation Systems, *IJCAI*, 2018.
18. Yang L. [et al.]. A Hybrid Retrieval-Generation Neural Conversation Model, *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2019, pp. 1341-1350.
19. Venkatesh A. [et al.]. On Evaluating and Comparing Conversational Agents, *arXiv preprint arXiv:1801.03625*, 2018.
20. Serban I. [et al.]. Multiresolution Recurrent Neural Networks: An Application to Dialogue Response Generation, *AAAI*, 2017.

Статью рекомендовала к опубликованию д.т.н., профессор М.А. Бутакова.

Куликова Яна Владимировна – Волгоградский государственный технический университет; e-mail: cul.jana@yandex.ru; г. Волгоград, Россия; кафедра программного обеспечения автоматизированных систем; старший преподаватель.

Литвиненко Владимир Алексеевич – e-mail: vladimirlit00.00@mail.ru; бакалавр.

Kulikova Yana Vladimirovna – Volgograd State Technical University; e-mail: cul.jana@yandex.ru; Volgograd, Russia; the department software engineering, senior lecturer.

Litvinenko Vladimir Alekseevich – e-mail: cul.jana@yandex.ru; the department software engineering; student.

УДК 004.421.6

DOI 10.18522/2311-3103-2023-5-104-118

И.И. Левин, К.Н. Алексеев

ПРЕОБРАЗОВАНИЕ СОРТИРУЮЩИХ СЕТЕЙ ДЛЯ РАЗНОЙ СТЕПЕНИ ПАРАЛЛЕЛИЗМА

Одни известные алгоритмы сортировки могут быть эффективнее других по какому-либо из основных критериев: число выполняемых операций, время выполнения элементарных операций, объем используемой памяти, степень параллелизма, функциональная регулярность связей в информационном графе алгоритма и т.д. При этом, имеется возможность выбрать такой алгоритм сортировки, который после выполнения операции редукции производительности вычислительной структуры, будет занимать минимум аппаратного ресурса. Выбор конкретного алгоритма напрямую зависит степени его распараллеливания, заданного временем обработки данных, коэффициента редукции и латентности вычислительной структуры, количества и разрядности сортируемых элементов. Алгоритмы сортировки являются информационно-эквивалентными, так как они выполняют одну и ту же математическую функцию. Однако каждый из алгоритмов рассматривается как автономный и независимый подход к решению задачи упорядочивания данных. Известно, что алгоритмам сортировки «пузырьком», «вставками» и «выбором» соответствует одна и та же сортирующая сеть, однако переход от одного алгоритма к другому до сих пор не описан в виде математических преобразований. Можно утверждать, что в настоящее время математический аппарат для описания различных алгоритмов сортировки и сортирующих сетей не формализован в полной мере, из-за чего не существует методологических основ перехода от одного алгоритма к другому. Иным способом описания алгоритма решения задачи является его представление в виде информационного графа, где выполняемые операции являются вершинами, которые объединены дугами, отражающими информационную зависимость между операциями. Преобразование информационного графа может приводить к получению иных информационно-эквивалентных алгоритмов. Пре-

имуществом подобного подхода к описанию алгоритмов является сравнительная простота используемого понятийного аппарата. В данной работе рассмотрены правила преобразования сортирующих сетей, на основе которых выполнен переход от одной сети к другой. Каждая из полученных сортирующих сетей может быть эффективна при разных коэффициентах распараллеливания и разном темпе обработки данных, от которых напрямую зависит коэффициент редукции производительности реализуемой вычислительной структуры. Автоматизация предложенных методов преобразования может позволить использовать разные алгоритмы сортировки, полученные из единого описания задачи в виде информационного графа, и зависящие от заданной скорости обработки данных.

Алгоритмы сортировки; сортирующие сети; информационный граф; ярусно-параллельная форма; преобразование алгоритмов.

I.I. Levin, K.N. Alekseev

TRANSFORMATION OF SORTING NETWORKS FOR DIFFERENT DEGREES OF PARALLELISM

Some well-known sorting algorithms may be more efficient than others according to any of the main criteria: the number of performed operations, the execution time of elementary operations, the used memory capacity, the parallelism degree, the functional regularity of connections in the information graph of algorithm, etc. At the same time, it is possible to choose such sorting algorithm, which will take up a minimum hardware resource after performance reducing operation of computational structure. The choice of a particular algorithm directly depends on it parallelization degree, the specified by the data processing time, the coefficient of reduction and latency of the computational structure and the number and bit depth of the sorted elements. Sorting algorithms are information-equivalent, since they perform the same mathematical function. However, each of algorithms is considered as an autonomous and independent approach to solving the data ordering problem. It is known that the same sorting network correspond to the "bubble", "inserts" and "selection" sorting algorithms. However, the transition from one algorithm to another has not yet been described in the form of mathematical transformations. It can be argued that the mathematical tools for describing various sorting algorithms and sorting networks is not fully formalized for today. Because of this, there is no methodological basis for the transition from one algorithm to another. Another method to describe the algorithm for solving the problem is its representation in the form of an information graph. In it, the performed operations are vertices that are combined by arcs reflecting the information dependence between operations. Transformation of the information graph can lead to obtaining other information and equivalent algorithms. The advantage of similar approach for the algorithm description of is the comparative simplicity of the used conceptual tools. In this paper, the transformation rules of sorting networks are considered, on the basis of which the transition from one network to another is performed. Each of the resulting sorting networks can be effective at different parallelization coefficients and data processing rates, on which the performance reduction coefficient of the implemented computational structure directly depends. Automation of the proposed transformation methods can allow the use the different sorting algorithms, derived from a single problem description in the form of an information graph, and depending on the specified data processing rate.

Sorting algorithms; sorting networks; information graph; tiered-parallel form; algorithm transformation .

Введение. Сортировка данных является одной из базовых операций, которую выполняют для облегчения последующей обработки информации при решении множества прикладных вычислительно-трудоемких задач, таких как: подавление шумов в изображении с помощью медианной фильтрации сигнала; кодирование и сжатие данных различными методами, например, методом Хаффмана; обработка данных сейсмической разведки; молекулярный докинг; анализ и обработка больших массивов данных, в том числе различные операции над матрицами и графами; и т.д.

Существует множество различных подходов к решению задачи сортировки данных, причем на основе каждого подхода построено большое количество алгоритмов [1–4]. К примеру, сортировка методом «пузырька», «вставок», «выбором»

и некоторые другие известные алгоритмы используют принцип попарного сравнения элементов сортируемого множества; быстрая сортировка, сортировка слиянием и сортирующие сети Батчера помимо сравнений используют принцип «разделяй и властвуй» для упорядочивания частей множества и их последующего слияния; сортировка подсчетом использует операцию сложения для нахождения одинаковых элементов множества, а также вместо сравнения чисел между собой выполняет их сравнение с полным набором вариантов их кодов; поразрядная сортировка предполагает выполнение последовательных сравнений каждого из разрядов кодов символов, и т.д.

На практике, различные алгоритмы сортировки целесообразно использовать в зависимости от требуемой степени параллелизма или скорости обработки данных, а также от параметров сортируемого множества [5]. Так в случае сортировки небольших объемов данных, порядка десятков мегабайт, на вычислительных системах традиционной архитектуры зачастую используется последовательный алгоритм быстрой сортировки. Однако при увеличении объема входных данных, который в некоторых задачах может достигать нескольких терабайт информации, время сортировки данным алгоритмом увеличивается нелинейно [6]. Для ускорения процесса сортировки больших объемов информации используют параллельные аппаратные реализации сортирующих сетей.

На сегодняшний день существует множество разнообразных сортирующих сетей [1, 7, 8], отличающихся количеством сортирующих элементов (числом сравнений), латентностью вычислительной структуры (временем получения результата на выходе конвейерной схемы), и регулярностью структуры (организация связей между сортирующими элементами), причем эффективность каждой конкретной сети существенно зависит от требуемой скорости обработки. Так, при необходимости обеспечить наибольшую степень параллелизма, эффективной будет сеть с меньшим числом сортирующих элементов в схеме, а при возможности параллельной реализации лишь некоторой части сети – эффективно использовать сеть с функционально-регулярной структурой, где из-за наличия симметрии можно использовать меньшее число устройств коммутации [5, 9–11].

Так как использование различных алгоритмов сортировки приводит к получению упорядоченного множества, можно утверждать, что выполняемые ими математические функции эквивалентны. В тоже время все алгоритмы сортировки обычно рассматриваются как независимые и несвязанные между собой. Известно, что из алгоритма сортировки методом пузырька можно получить сортировку выбором, однако переход от одного алгоритма к другому не формализован в виде математических преобразований, а основан лишь на эмпирических наблюдениях и логике построения последовательных программ. Можно утверждать, что на сегодняшний день не существует методологических основ, позволяющих переходить от одного алгоритма к другому, в связи с чем нет возможности автоматизировать процесс получения наиболее эффективных алгоритмов в зависимости от критериев задачи.

Вместе с этим, существует иной подход к формализации и выведению параллельных алгоритмов, основанный на анализе и преобразовании информационных графов. В работах [5, 12-19] было рассмотрено преобразование графов, состоящих из ассоциативных и дистрибутивных алгебраических операций, которые приводили к получению графов, соответствующих другим известным алгоритмам решаемой задачи. Исследования показывают, что задача преобразования информационного графа обладает меньшей алгоритмической сложностью по сравнению последовательным выполнением математических преобразований, что может обеспечить возможность ее дальнейшую автоматизацию.

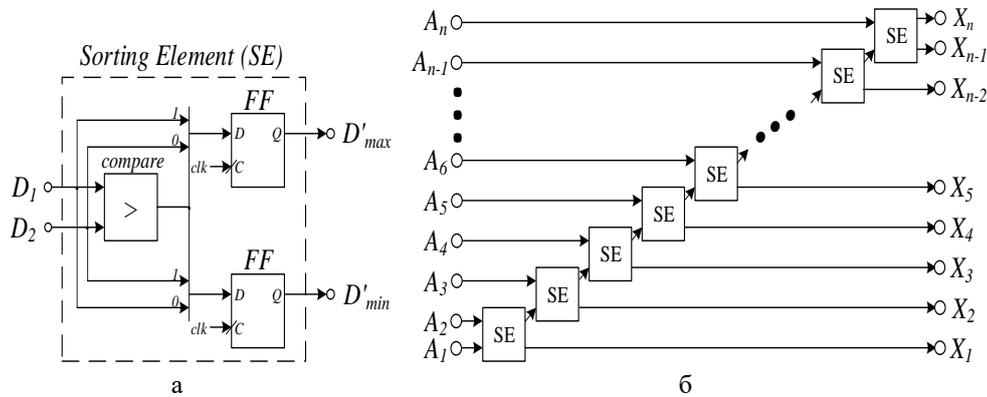


Рис. 2. а – структурная схема сортирующего элемента SE;
 б – граф, построенный путем объединения сортирующих элементов SE согласно принципу «голова-хвост»

Рассмотрим случай, когда на вход ступени подается неупорядоченное множество $\{A_1, \dots, A_k\}$ и упорядоченное подмножество $\langle A_{k+1}, \dots, A_n \rangle$, элементы которого расположены в порядке возрастания. Легко заметить, что на выходе max блока сравнения SE_{k-1} будет получен максимальный элемент неупорядоченного множества: $max(A_1, \dots, A_k)$. Все следующие блоки сравнения $SE_k \div SE_n$ можно рассматривать как ступень, на которую приходит упорядоченное подмножество $\langle A_{k+1}, \dots, A_n \rangle$ и элемент $max(A_1, \dots, A_k)$, место которого не определено. Очевидно, что результатом выполнения данной операции будет упорядоченное подмножество $\langle X_k, \dots, X_n \rangle$, при этом отношение элементов $\{X_1, \dots, X_{k-1}\}$ останется неопределенным.

Таким образом, можно утверждать, что ступень, сортирующие элементы в которой объединены согласно принципу «голова-хвост», помимо функции нахождения максимума также выполняет функцию вставки элемента A_1 в частично упорядоченное подмножество $\langle A_{k+1}, \dots, A_n \rangle$. В связи с этим, в зависимости от последовательности ступеней в сортирующей сети, при последовательном выполнении операций, она может соответствовать иным известным алгоритмам сортировки, таким как сортировка «вставками» и «выбором» (рис. 3,а,б соответственно).

Принципы объединения ступеней сети сортировки. Пусть граф, в котором ступень S_1 с меньшим числом операций следует за ступенью S_2 (рис. 4,а), описывает функцию F , а граф, в котором ступени S_1 и S_2 объединены в обратном порядке (рис. 4,б), описывает функцию F' . Докажем, что $F \Leftrightarrow F'$ при объединении ступеней S_1 и S_2 .

Очевидно, что элемент множества X_3 на рис. 4,а является максимальным элементом множества: $X_3 = max(A)$. Остальные два выхода ступени S_2 никак не относятся между собой и поступают на вход ступени S_1 . Элемент множества X_2 будет являться вторым максимальным элементом множества A . Так как после выделения двух максимумов остается только один элемент, то множество X будет полностью упорядоченным: $X_3 \geq X_2 \geq X_1$. Таким образом, ступени S_1 и S_2 объединенные в порядке уменьшения количества операций, выполняют функцию упорядочивания $F = X_3 \geq X_2 \geq X_1$.

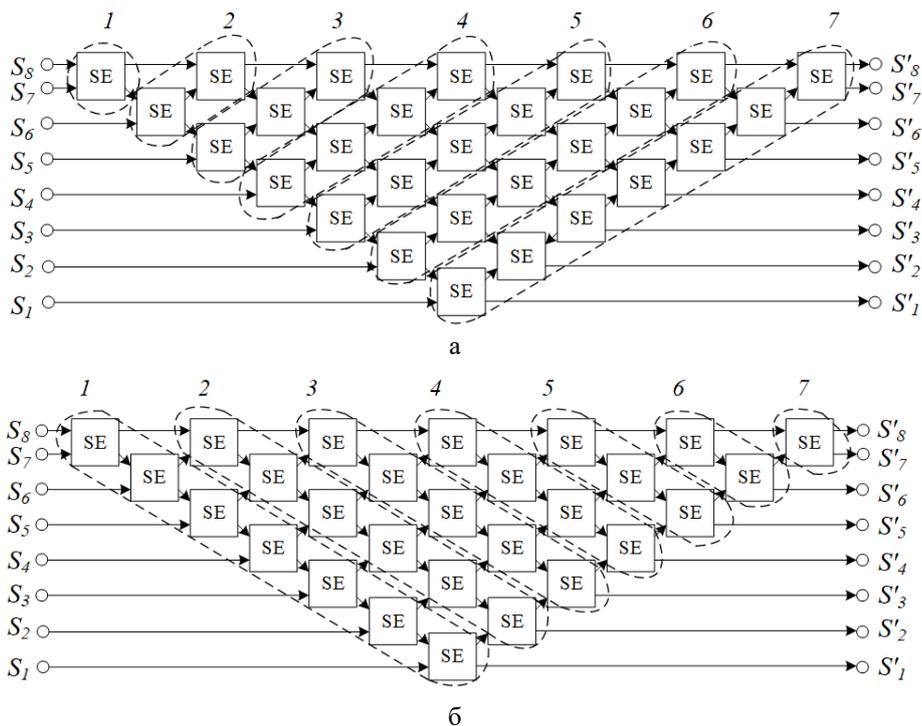


Рис. 3. Схемы сортирующих сетей, соответствующие:
 а – сортировке вставками; б – сортировке выбором

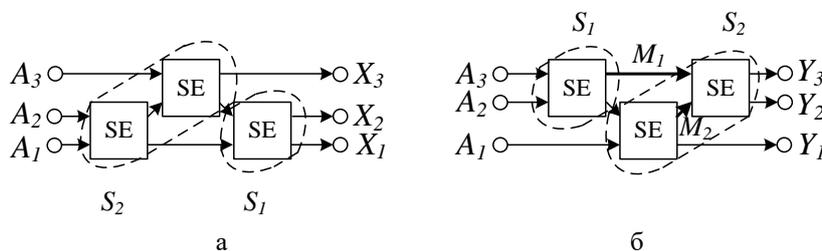


Рис. 4. Объединение ступеней S_1 и S_2 : а – в порядке уменьшения количества операций; б – в порядке увеличения количества операций

На рис. 4,б ступени S_1 и S_2 объединены в порядке увеличения количества операций. На выходе ступени S_1 будет сформировано упорядоченное подмножество, которое поступает на вход ступени S_2 . Также на вход ступени S_2 поступает элемент A_1 , место которого не определено. В этом случае на выходе ступени S_2 будет сформировано упорядоченное по возрастанию множество: $Y_3 \geq Y_2 \geq Y_1$. Можно сказать, что ступени S_1 и S_2 объединенные в порядке увеличения количества операций, выполняют функцию упорядочивания $F' = Y_3 \geq Y_2 \geq Y_1$.

Так как в результате обработки множества A и в первом и во втором варианте объединения ступеней сортирующей сети получены упорядоченные по убыванию множества, функции F и F' являются эквивалентными.

Положим, что множество $A = \{A_1, A_2, A_3, A_4\}$ является неупорядоченным. Сравним функции, выполняемую ступенями, объединенными в различном порядке и в которых на один элемент больше по сравнению с графами, приведенными на рис. 4. Докажем, что любом при подобном объединении ступеней функции F и F' являются эквивалентными.

На рис. 5,а ступени S_2 и S_3 объединены в порядке уменьшения количества операций. Выходы ступени S_3 представляют собой неупорядоченное множество $B = \{B_1, B_2, B_3\}$ и элемент $X_4 = \max(A)$. В свою очередь, выходы ступени S_2 будут представлять собой неупорядоченное множество $\{X_1, X_2\}$ и элемент $X_3 = \max(B)$. Довольно очевидно, что результирующая функция рассматриваемого объединения ступеней S_2 и S_3 будет иметь следующий вид: $F = X_4 \geq X_3 \geq \{X_1, X_2\}$.

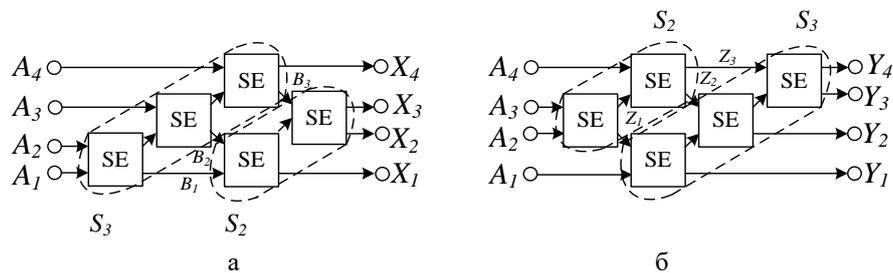


Рис. 5. Объединение ступеней S_2 и S_3 : а – в порядке уменьшения количества операций; б – в порядке увеличения количества операций

На рис. 5,б ступени S_2 и S_3 объединены в порядке увеличения количества операций. Выходы ступени S_2 представляют собой частично упорядоченное множество $Z_3 \geq \{Z_1, Z_2\}$, которое поступает в качестве входных данных на входы ступени S_3 . Также на вход ступени S_3 поступает элемент исходного множества A_1 . Тогда на вход последнего элемента SE ступени S_3 поступит два элемента: $Z_3 = \max(A_2, A_3, A_4)$ и $\max(A_1, Z_1, Z_2)$. Очевидно, что два старших выхода ступени S_3 имеют следующие отношения: $Y_4 = \max(A)$ и $Y_4 \geq Y_3$. В общем виде Y_3 можно представить как:

$$Y_3 = \min(\max(Z_2, Z_1, A_1), Z_3) = \min(\max(Z_2, Z_1, A_1), \max(A_2, A_3, A_4)),$$

или

$$Y_3 = \min\left(\max\left(\left(\{A_2, A_3, A_4\} - \max(A_2, A_3, A_4)\right), A_1\right), \max(A_2, A_3, A_4)\right).$$

Из приведенного отношения очевидно, что Y_3 является вторым по величине элементом множества A , тогда $F' = Y_4 \geq Y_3 \geq \{Y_1, Y_2\}$. В связи с этим можно утверждать, что ступени S_2 и S_3 , объединенные в различном порядке, выполняют эквивалентные преобразования: $F \Leftrightarrow F'$.

На рис. 6,а ступени S_{n-1} и S_n объединены в порядке уменьшения количества операций. Очевидно, что элемент множества X_{n+1} является максимальным элементом множества: $X_{n+1} = \max(A)$. Остальные выходы ступени S_n никак не относятся между собой и поступают на входы ступени S_{n-1} . Очевидно, что элемент множества X_n будет являться вторым максимальным элементом множества A . Тогда множество X будет частично упорядоченным, где функция упорядочивания $F = X_{n+1} \geq X_n \geq \{X_1, \dots, X_{n-1}\}$.

На рис. 6,б ступени S_{n-1} и S_n объединены в порядке увеличения количества операций. Так как на выходе ступени S_{n-1} будет найден максимальный элемент подмножества: $\max(A_2, \dots, A_{n+1})$, на вход ступени S_n поступит частично упорядоченное подмножество, а также элемент A_1 , место которого не определено. Отметим, что в качестве частично упорядоченного множества будет выступать один максимальный элемент, найденный в ступени S_{n-1} . Так как ступень S_n позволяет выделить максимальный элемент множества A , на вход Z_4 либо поступит элемент A_1 , либо второй по величине максимум из подмножества $\{A_2, \dots, A_{n+1}\}$. В любом из этих случаев на выходе ступени S_n будет сформировано частично упорядоченное множество Y , где функция упорядочивания: $F' = Y_{n+1} \geq Y_n \geq \{Y_1, \dots, Y_{n-1}\}$. Очевидно, что функции F и F' , эквивалентны.

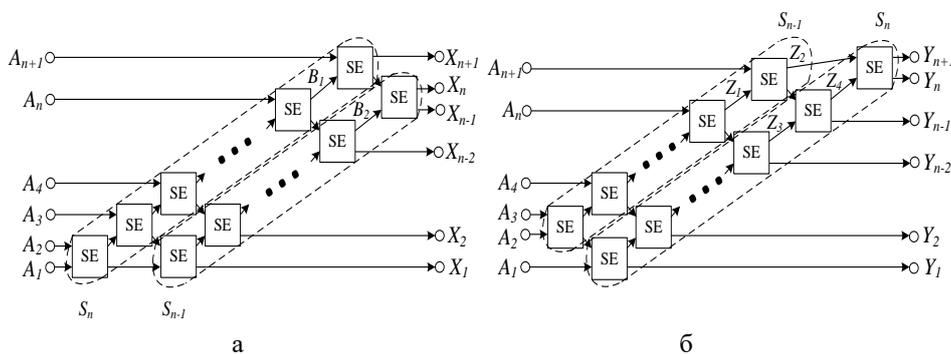


Рис. 6. Объединение ступеней S_{n-1} и S_n : а – в порядке уменьшения количества операций; б – в порядке увеличения количества операций

Таким образом доказано, что выражение $F \Leftrightarrow F'$ будет верно для ступеней, состоящих из любого числа сортирующих элементов SE при объединении ступеней S_{n-1} и S_n , и не зависит от порядка их следования. Основываясь данным свойстве, докажем, что функция, выполняемая двумя ступенями, не зависит не только от порядка их следования, но и от количества операций в них.

На рис. 7 показано два варианта последовательного объединения ступеней S_n и S_m , где $n > m$, в которых операции объединены по принципу «голова-хвост». На рис. 7,а ступени S_n и S_m объединены в порядке уменьшения количества операций. Очевидно, что элемент множества X_n является максимальным элементом множества A : $X_n = \max(A)$. Остальные выходы ступени S_n никак не относятся между собой и коммутируются ко входам ступени S_m . Для элемента множества X_{n-1} , который будет являться максимальным элементом части множества A , справедливо:

$$X_{n-1} = \max(A - \max(A)); X_{n-1} \geq \{X_{n-2}, X_{n-3}, \dots, X_{n-m+1}\}.$$

Множество X будет являться частично упорядоченным, где выделен максимальный элемент X_n и выделено подмножество $\{X_{n-2}, X_{n-3}, \dots, X_{n-m+1}\}$, для которого определен максимальный элемент X_{n-1} , причем отношения элементов $X_{n-1} > \{X_{n-2}, X_{n-3}, \dots, X_{n-m+1}\}$ и $\{X_{n-m}, \dots, X_1\}$ не определены между собой. Так, функцию F , которую выполняют ступени S_n и S_m , объединенные, как показано на рис. 5,а, можно описать как:

$$F = X_n > \{\{X_{n-1} > \{X_{n-2}, X_{n-3}, \dots, X_{n-m+1}\}\}, X_{n-m}, \dots, X_1\}.$$

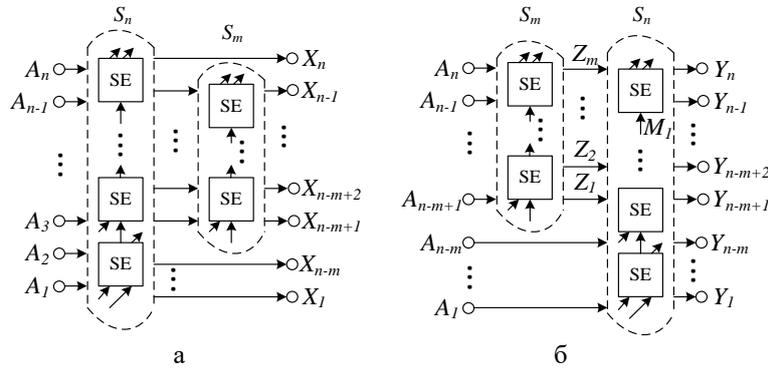


Рис. 7. Объединение ступеней S_n и S_m : а – в порядке уменьшения количества операций; б – в порядке увеличения количества операций

На рис. 7,б ступени S_n и S_m объединены в порядке увеличения количества операций. На вход ступени S_m поступают элементы исходного множества $\{A_{n-m+1}, \dots, A_n\}$, а на выходе формируется частично отсортированное множество $Z = \{Z_1, Z_2, \dots, Z_m\}$, для которого справедливо:

$$\begin{aligned} Z_m &= \max(A_{n-m+1}, \dots, A_n); \\ Z_m &> \{Z_1, Z_2, \dots, Z_{m-1}\}; \\ Z_m &> \{A_{n-m+1}, \dots, A_n\} - \max(A_{n-m+1}, \dots, A_n). \end{aligned}$$

Оставшаяся часть исходного множества $\{A_1, \dots, A_{n-m}\}$, а также частично отсортированное множество Z , поступают на вход ступени S_n . Для максимального элемента ступени S_n справедливо: $Y_n = \max(A)$, где:

$$Y_n = \max(A_1, A_2, \dots, A_{n-m}, Z_m).$$

Так как ступень S_n позволяет выделить максимальный элемент множества A , на вход M_1 либо поступит элемент $\max(A_1, A_2, \dots, A_{n-m})$, либо второй по величине максимум из подмножества $\{A_{n-m+1}, \dots, A_n\}$. В связи с этим, на выходе ступени S_n будет сформировано частично упорядоченное множество Y , со следующей функцией упорядочивания:

$$F' = Y_n > \{Y_{n-1} > \{Y_{n-2}, Y_{n-3}, \dots, Y_{n-m+1}\}, Y_{n-m}, \dots, Y_1\}.$$

Очевидно, что функции F и F' , эквивалентны. Таким образом, доказано, что результат упорядочивания входного множества A будет одинаков для обоих вариантов расположения ступеней относительно друг друга, операции в которых объединены согласно принципу «голова-хвост» и не зависит как от порядка их следования, так и от числа используемых сортирующих элементов.

В то же время стоит отметить, что для выполнения отношения $F \Leftrightarrow F'$ необходимо, чтобы ступени были объединены корректно: при порядке следования от большей ступени к меньшей, после большей ступени необходимо подавать на меньшую старшие выходы, кроме максимального элемента, тогда как при обратном порядке следования ступеней требуется подавать все выходы меньшей ступени к старшим входам большей ступени.

Построение и преобразование сортирующих сетей. Сортирующая сеть, соответствующая последовательным алгоритмам сортировки «пузырьком», «вставками» и «выбором» (рисунок 3) имеет различное количество сортирующих эле-

ментов в ступенях, из-за чего когда коэффициент редукции [20] находится в интервале $1 < r^0 < n - 1$ при редукции по операциям в итерациях, аппаратный ресурс, занимаемый вычислительной структурой задачи, будет уменьшаться нелинейно [5, 14]. Наиболее эффективной будет являться предельная редукция сортирующей сети по операциям в ступенях ($r^0 = n - 1$), как это показано на рис. 8: остается $n - 1$ изоморфных итераций, состоящих из одного сортирующего элемента.

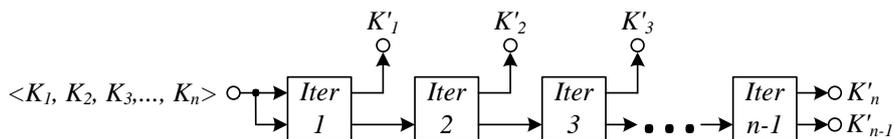


Рис. 8. Вычислительная структура сортирующей сети, соответствующая последовательным алгоритмам сортировки «пузырьком», «вставками» и «выбором», после выполнения редукции по количеству операций

Основываясь на принципах объединения ступеней нетрудно показать, что порядок их следования в структуре сортирующей сети может быть любым при корректной коммутации ступеней между собой. Так, иные архитектуры, представленные на рис. 9, также являются сортирующими сетями с одинаковым числом сортирующих элементов, а их разбиение на ступени соответствует различным известным последовательным алгоритмам сортировки.

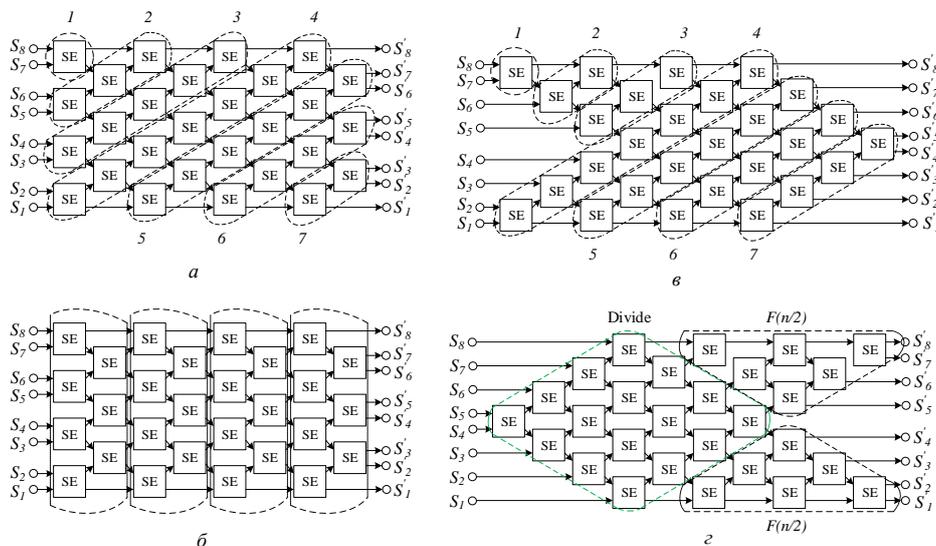


Рис. 9. Схемы иных простейших сортирующих сетей, соответствующие:
 а – смешанной сортировке «пузырьком», вставками и слиянием;
 б – нечетно-четной сортировке; в – «шейкерной» сортировке;
 г – «ведерной» сортировке

Нечетно-четная сортирующая сеть (рис. 9,а), построенная путем изменения порядка следования ступеней, имеет четкое разбиение на итерации, содержащие одинаковое количество сортирующих элементов (рис. 9,б). Когда коэффициент редукции $r^0 < n/2$, количество итераций нечетно-четной сортирующей сети со-

кращается в r^0 раз. Так как граф нечетно-четной сортирующей сети имеет регулярную структуру, при выполнении редукции по итерациям в данном случае полностью сохраняется межитерационная коммутация и возникает обратная связь по данным: обеспечивается коммутация входов первой итерации с выходами последней, а при редукции производительности будет наблюдаться линейное сокращение используемого аппаратного ресурса.

Если коэффициент редукции $r^0 = n/2$, нечетно-четную сортирующую сеть необходимо сократить до одной итерации. В этом случае для организации обратной связи необходимо обеспечить коммутацию входов единственной итерации с ее же выходами.

Все ступени сортирующих сетей, соответствующих алгоритмам «шейкерной» и «ведерной» сортировки (рис. 8, в, г), содержат различное количество сортирующих элементов, зависящее от количества обрабатываемых данных, поэтому при выполнении редукции по операциям в итерациях прослеживается нелинейная зависимость аппаратного ресурса от коэффициента редукции. В данном случае наиболее эффективной является предельная редукция сортирующей сети по числу операций в итерациях при $r^0 = n - 1$, приводящая к тому, что остается $n/2$ изоморфных итераций, состоящих из двух сортирующих элементов.

Эффективность сортирующих сетей. Эффективность применения одной из множества существующих сетей зависит от заданного темпа обработки данных, коэффициента редукции r^0 , латентности алгоритма, количества элементов в массиве n [5].

При сортировке множества массивов данных, элементы которых поступают на обработку параллельно ($r^0 = 1$), можно использовать любую из представленных сортирующих сетей, так как все они содержат одинаковое число сортирующих элементов. Однако, наиболее эффективным является использование топологии нечетно-четной сортирующей сети, которая не требует дополнительной синхронизации потоков входных и выходных данных.

Когда коэффициент редукции вычислительной структуры лежит в пределах $1 < r^0 \leq n$, наиболее эффективной также является нечетно-четная сортирующая сеть, так как позволяет обеспечить меньшее число коммутирующих элементов между итерациями.

При сортировке множества массивов данных, элементы которых поступают на обработку последовательно ($r^0 \geq n$), рационально использовать либо сортирующую сеть, соответствующую алгоритмам сортировки «пузырьком», «вставками» и «выбором», либо сети, соответствующие алгоритмам «шейкером» или «ведерной» сортировки, так как они используют значительно меньше каналов данных по сравнению с нечетно-четной сортирующей сетью. Отметим, что выбор конкретной топологии зависит от доступного аппаратного ресурса: так, если возможно организовать два входных канала данных, более эффективной является сеть с наименьшей латентностью вычислительной структуры: соответствующая алгоритмам «шейкером» или «ведерной» сортировки. Если же доступен только один канал данных, наиболее эффективной является конвейерная реализация сортирующей сети, соответствующей алгоритмам сортировки «пузырьком», «вставками» и «выбором».

Заключение. В рамках данной работы рассмотрены принципы объединения ступеней сети сортировки и способы выбора конкретной сети в зависимости от темпа поступления данных.

Отметим, что помимо рассмотренных функций, ступень сортирующей сети, сортирующие элементы в которой объединены по принципу «голова-хвост», дополнительно выполняет перенос нескольких больших элементов ближе к старшим выходам ступени. Так, в результате упорядочивания множества $A = \{7, 4, 3, 8, 6, 2, 9, 3\}$ ступенью

S_7 , состоящей из семи SE , будет получено множество $X = 9 \geq \{4,3,7,6,2,8,3\}$. Из представленных данных видно, что помимо найденного максимального элемента 9, значения 7 и 8 передвинулись на 2 позиции ближе к старшим выходам ступени.

Так как основным принципом работы каждой ступени в простейших сортирующих сетях, соответствующих последовательным алгоритмам сортировки «пузырьком» или «выбором», является именно нахождение максимума в неотсортированном множестве элементов, перемещение не максимальных элементов очевидно не имеет никакой необходимости. Также очевидно, что для сортировки элементов массива достаточно единожды переместить каждый элемент на новое место, для чего необходимо выполнить всего n операций перемещения элементов, что используется, например, в последовательном алгоритме сортировки «выбором». Исходя из вышесказанного, можно сделать вывод, что сортирующие сети, построенные путем объединения ступеней согласно принципу «голова-хвост», обладают избыточностью выполняемых операций перемещения элементов.

В следующих работах будет показано, что на основе принципов построения сортирующих сетей, а также некоторых иных базовых преобразований, можно устранить избыточность выполняемых операций и получить более эффективную нечетно-четную сеть Батчера, состоящую из значительно меньшего числа операций и имеющую функционально-регулярную структуру [1, 21]. На основе формализации принципов построения и преобразования сортирующих сетей имеется возможность разработать методику их преобразования, позволяющую оптимизировать число операционных вершин и латентность вычислительной структуры.

Потенциально, автоматизация процесса преобразования сортирующих сетей позволит увеличить эффективность используемого оборудования без больших временных затрат на разработку и отладку кода параллельной программы при изменении темпа поступления и обработки данных.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Кнут Д.Э. Искусство программирования. Т. 3. Сортировка и поиск. – 2-е изд.: пер. с англ. – М.: Издательский дом «Вильямс», 2001.
2. Дасгутта С., Пападимитриу Х., Вазирани У. Алгоритмы: пер. с англ. / под ред. А. Шеня. – М.: МЦНМО, 2014. – 320 с.
3. Кормен Томас Х. и др. Алгоритмы: построение и анализ. – 3-е изд.: пер. с англ. – М.: ООО «И.Д. Вильямс», 2013. – 1328 с.
4. Скиена С.С. Алгоритмы. Руководство по разработке. – 3-е изд.: пер. с англ. – СПб.: БХВ-Петербург, 2022. – 848 с.
5. Алексеев К.Н. Методы и средства создания параллельно-конвейерных программ для решения задач реального времени на реконфигурируемых вычислительных системах: дисс. ... канд. техн. наук, по специальности 05.13.11 “Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей”, научный руководитель: д.т.н., проф. Левин И.И. – Таганрог, 2020. – 213 с.
6. Скворцов С.В., Пюрова Т.А. Параллельные алгоритмы сортировки данных и их реализация на платформе CUDA // Вестник РГРТУ. – 2016. – № 58. – С. 42-48.
7. Сортировочные сети с особыми свойствами. – Режим доступа: https://neerc.ifmo.ru/wiki/index.php?title=%D0%A1%D0%BE%D1%80%D1%82%D0%B8%D1%80%D0%BE%D0%B2%D0%BE%D1%87%D0%BD%D1%8B%D0%B5_%D1%81%D0%B5%D1%82%D0%B8_%D1%81_%D0%BE%D1%81%D0%BE%D0%B1%D1%8B%D0%BC%D0%B8_%D1%81%D0%B2%D0%BE%D0%B9%D1%81%D1%82%D0%B2%D0%B0%D0%BC%D0%B8 (дата обращения: 01.11.2023).
8. Сортирующие сети для квадратичных сортировок. – Режим доступа: https://neerc.ifmo.ru/wiki/index.php?title=%D0%A1%D0%BE%D1%80%D1%82%D0%B8%D1%80%D1%83%D1%8E%D1%89%D0%B8%D0%B5_%D1%81%D0%B5%D1%82%D0%B8_%D0%B4%D0%BB%D1%8F_%D0%BA%D0%B2%D0%B0%D0%B4%D1%80%D0%B0%D1%82%D0%B8%D1%87%D0%BD%D1%8B%D1%85_%D1%81%D0%BE%D1%80%D1%82%D0%B8%D1%80%D0%BE%D0%B2%D0%BE%D0%BA (дата обращения: 01.11.2023).

9. *Алексеев К.Н., Левин И.И., Сорокин Д.А.* Структурно-процедурная реализация на реконфигурируемых вычислительных системах кодирования Хаффмана в реальном масштабе времени // Вестник компьютерных и информационных технологий. – 2018. – № 9. – С. 3-10. – ISSN 1810-7206. – DOI: 10.14489/vkit.2018.09.pp.003-010. – <http://www.vkit.ru/index.php/current-issue-rus/751-003-010>.
10. *Алексеев К.Н., Сорокин Д.А., Семерникова Е.Е.* Структурно-процедурная реализация кодирования алгоритмом Хаффмана в задачах управления // Матер. 10-й Всероссийской мультikonференции по проблемам управления (МКПУ-2017), 11-16 сентября 2017 г. – Ростов-на-Дону: Изд-во ЮФУ, 2017. – Т. 3. – С. 126-127.
11. *Dudnikov E.A., Alekseev K.N., Sorokin D.A.* High-Speed Huffman encoding of dense data flows on FPGA // German International Journal of Modern Science / Deutsche internationale Zeitschrift für zeitgenössische Wissenschaft. – 2021. – No. 9, Vol. 1. – P. 36-40. – <https://dizww.com/wp-content/uploads/2021/05/Deutsche-internationale-Zeitschrift-für-zeitgenössische-Wissenschaft-№9-part-1-2021-37-41.pdf>.
12. *Каляев А.В., Левин И.И.* Модульно-наращиваемые многопроцессорные системы со структурно-процедурной организацией вычислений. – М.: Янус-К, 2003. – 380 с.
13. *Каляев И.А., Левин И.И., Семерников Е.А., Шмойлов В.И.* Реконфигурируемые мультikonвейерные вычислительные структуры. – 2-е изд., перераб. и доп. / под общ. ред. И.А. Каляева. – Ростов-на-Дону: Изд-во ЮНЦ РАН, 2009. – 344 с. – ISBN 978-5-902982-61-6.
14. *Писаренко И.В., Алексеев К.Н., Мельников А.К.* Ресурснезависимое представление сортирующих сетей на языке программирования Set@1 // Вестник компьютерных и информационных технологий. – 2019. – № 11 (185). – С. 53-60. – DOI: 10.14489/vkit.2019.11.pp.053-060. – <http://vkit.ru/index.php/current-issue-rus/857-053-060>.
15. *Писаренко И.В., Касаркин А.В., Алексеев К.Н.* Рекурсивное описание графов с ассоциативными операциями на языке программирования Set@1 // Всероссийская научная конференция "Суперкомпьютерные технологии (СКТ) 2020", 05-10 октября 2020 г.: Крым, Алушта, Россия. – Ростов-на-Дону; Таганрог: Изд-во ЮФУ, 2020. – С. 79-83.
16. *Михайлов Д.В.* Методы создания параллельно-конвейерных программ для задач с последовательным информационным графом: дисс. ... канд. техн. наук, по специальности 05.13.11 "Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей", научный руководитель: д.т.н., проф. Левин И.И. – Таганрог, 2022. – 213 с.
17. *Михайлов Д.В., Левин И.И., Дордопуло А.И., Писаренко И.В.* Представление графов с ассоциативными операциями на языке программирования SET@L // Известия ЮФУ. Технические науки. – 2020. – № 3 (213). – С. 98-109. – DOI: 10.18522/2311-3103-2020-3-98-111.
18. *Михайлов Д.В.* Преобразование некоторых видов последовательных информационных графов в параллельно-конвейерную форму // Известия ЮФУ. Технические науки. – 2020. – № 7 (217). – С. 78-93. – DOI: 10.18522/2311-3103-2020-7-78-93.
19. *Михайлов Д.В.* Преобразование последовательного информационного графа метода прогонки в параллельную форму // Известия ЮФУ. Технические науки. – 2021. – № 7 (224). – С. 177-188. – DOI: 10.18522/2311-3103-2021-7-177-188.
20. *Сорокин Д.А., Дордопуло А.И.* Методика сокращения аппаратных затрат в сложных системах при решении задач с существенно-переменной интенсивностью потоков данных // Известия ЮФУ. Технические науки. – 2012. – № 4. – С. 213-219.
21. *Batcher K.E.* Sorting Networks and their Applications // Proc. AFIPS Spring Joint Comput. Conf. – 1968. – Vol. 32. – P. 307314.

REFERENCES

1. *Knut D.E.* *Iskusstvo programmirovaniya*. Т. 3. *Sortirovka i poisk* [The art of programming. Volume 3. Sorting and searching]. 2nd ed.: transl. from engl. Moscow: Izdatel'skiy dom «Vil'yams», 2001.
2. *Dasgupta S., Papadimitriou Kh., Vazirani U.* *Algoritmy* [Algorithms]: trans. from engl., ed. by A. Shenyu. Moscow: MTSNMO, 2014, 320 p.
3. *Kormen Tomas Kh. i dr.* *Algoritmy: postroenie i analiz* [Algorithms: construction and analysis]. 3rd ed.: transl. from engl. Moscow: ООО «I.D. Vil'yams», 2013, 1328 p.
4. *Skiena S.S.* *Algoritmy. Rukovodstvo po razrabotke* [Algorithms. Development Guide]. 3rd ed.: transl. from engl. Saint Petersburg: BKhV-Peterburg, 2022, 848 p.

5. *Alekseev K.N.* Metody i sredstva sozdaniya parallel'no-konveyernykh programm dlya resheniya zadach real'nogo vremeni na rekonfiguriruemyykh vychislitel'nykh sistemakh: diss. ... kand. tekhn. nauk, po spetsial'nosti 05.13.11 "Matematicheskoe i programmnoe obespechenie vychislitel'nykh mashin, kompleksov i komp'yuternykh setey", nauchnyy rukovoditel': d.t.n., prof. Levin I.I. [Methods and means for creating parallel-pipeline programs for solving real-time problems on reconfigurable computing systems: cand. of eng. sc. diss., specialty 05.13.11 "Mathematical and software support for computers, complexes and computer networks", scientific supervisor: dr. of eng. sc., professor Levin I.I.]. Taganrog, 2020, 213 p.
6. *Skvortsov S.V., Pyurova T.A.* Parallelnye algoritmy sortirovki dannykh i ikh realizatsiya na platforme CUDA [Parallel data sorting algorithms and their implementation on the CUDA platform], *Vestnik RGRU* [Vestnik of Ryazan State Radio Engineering University], 2016, No. 58, pp. 42-48.
7. Sortirovochnye seti s osobymi svoystvami [Sorting networks with special properties]. Available at: https://neerc.ifmo.ru/wiki/index.php?title=%D0%A1%D0%BE%D1%80%D1%82%D0%B8%D1%80%D0%BE%D0%B2%D0%BE%D1%87%D0%BD%D1%8B%D0%B5_%D1%81%D0%B5%D1%82%D0%B8_%D1%81_%D0%BE%D1%81%D0%BE%D0%B1%D1%8B%D0%BC%D0%B8_%D1%81%D0%B2%D0%BE%D0%B9%D1%81%D1%82%D0%B2%D0%B0%D0%BC%D0%B8 (accessed 01 November 2023).
8. Sortiruyushchie seti dlya kvadrachnykh sortirovok [Sorting networks for quadratic sorts]. Available at: https://neerc.ifmo.ru/wiki/index.php?title=%D0%A1%D0%BE%D1%80%D1%82%D0%B8%D1%80%D1%83%D1%8E%D1%89%D0%B8%D0%B5_%D1%81%D0%B5%D1%82%D0%B8_%D0%B4%D0%BB%D1%8F_%D0%BA%D0%B2%D0%B0%D0%B4%D1%80%D0%B0%D1%82%D0%B8%D1%87%D0%BD%D1%8B%D1%85_%D1%81%D0%BE%D1%80%D1%82%D0%B8%D1%80%D0%BE%D0%B2%D0%BE%D0%BA (accessed 01 November 2023).
9. *Alekseev K.N., Levin I.I., Sorokin D.A.* Strukturno-protsedurnaya realizatsiya na rekonfiguriruemyykh vychislitel'nykh sistemakh kodirovaniya Khaffmana v real'nom masshtabe vremeni [Structural and procedural implementation on reconfigurable computing systems of Huffman coding in real time], *Vestnik komp'yuternykh i informatsionnykh tekhnologiy* [Bulletin of computer and information technologies], 2018, No. 9, pp. 3-10. ISSN 1810-7206. DOI: 10.14489/vkit.2018.09.pp.003-010. Available at: <http://www.vkit.ru/index.php/current-issue-rus/751-003-010>.
10. *Alekseev K.N., Sorokin D.A., Semernikova E.E.* Strukturno-protsedurnaya realizatsiya kodirovaniya algoritmom Khaffmana v zadachakh upravleniya [Structural and procedural implementation of coding by the Huffman algorithm in control problems], *Mater. 10-y Vserossiyskoy mul'tikonferentsii po problemam upravleniya (MKPU-2017), 11-16 sentyabrya 2017 g.* [Proceedings of the 10th All-Russian Multi-Conference on Control Problems (MCPU-2017), September 11-16, 2017]. Rostov-on-Don: Izd-vo YuFU, 2017, Vol. 3, pp. 126-127.
11. *Dudnikov E.A., Alekseev K.N., Sorokin D.A.* High-Speed Huffman encoding of dense data flows on FPGA, *German International Journal of Modern Science / Deutsche internationale Zeitschrift für zeitgenössische Wissenschaft*, 2021, No. 9, Vol. 1, pp. 36-40. Available at: <https://dizww.com/wp-content/uploads/2021/05/Deutsche-internationale-Zeitschrift-für-zeitgenössische-Wissenschaft-№9-part-1-2021-37-41.pdf>.
12. *Kalyaev A.V., Levin I.I.* Modul'no-narashchivaemye mnogoprotsessornyye sistemy so strukturno-protsedurnoy organizatsiyey vychisleniy [Modularly scalable multiprocessor systems with structural and procedural organization of calculations]. Moscow: Yanus-K, 2003, 380 p.
13. *Kalyaev I.A., Levin I.I., Semernikov E.A., Shmoylov V.I.* Rekonfiguriruemyye mul'tikonveyernyye vychislitel'nye struktury [Reconfigurable multi-pipeline computing structures]. 2nd ed., under general ed. I.A. Kalyaeva. Rostov-on-Don: Izd-vo YuNTS RAN, 2009, 344 s. ISBN 978-5-902982-61-6.
14. *Pisarenko I.V., Alekseev K.N., Mel'nikov A.K.* Resursonezavisimoe predstavlenie sortiruyushchikh setey na yazyke programmirovaniya Set@l [Resource-independent representation of sorting networks in the Set@l programming language], *Vestnik komp'yuternykh i informatsionnykh tekhnologiy* [Bulletin of computer and information technologies], 2019, No. 11 (185), pp. 53-60. DOI: 10.14489/vkit.2019.11.pp.053-060. Available at: <http://vkit.ru/index.php/current-issue-rus/857-053-060>.

15. *Pisarenko I.V., Kasarkin A.V., Alekseev K.N.* Rekursivnoe opisanie grafov s assotsiativnymi operatsiyami na yazyke programmirovaniya Set@l [Recursive description of graphs with associative operations in the Set@l programming language], *Vserossiyskaya nauchnaya konferentsiya "Superkomp'yuternye tekhnologii (SKT) 2020", 05-10 oktyabrya 2020 g.: Krym, Alushta, Rossiya* [All-Russian scientific conference "Supercomputer Technologies (SCT) 2020", October 05-10, 2020: Crimea, Alushta, Russia]. Rostov-on-Don; Taganrog: Izd-vo YuFU, 2020, pp. 79-83.
16. *Mikhaylov D.V.* Metody sozdaniya parallel'no-konveyernykh programm dlya zadach s posledovatel'nyim informatsionnym grafom: diss. ... kand. tekhn. nauk, po spetsial'nosti 05.13.11 "Matematicheskoe i programmnoe obespechenie vychislitel'nykh mashin, kompleksov i komp'yuternykh setey", nauchnyy rukovoditel': d.t.n., prof. Levin I.I. [Methods for creating parallel-pipeline programs for problems with a sequential information graph: cand. of eng. sc. diss., specialty 05.13.11 "Mathematical and software support for computers, complexes and computer networks", scientific supervisor: dr. of eng. sc., professor Levin I.I.]. Taganrog, 2022, 213 p.
17. *Mikhaylov D.V., Levin I.I., Dordopulo A.I., Pisarenko I.V.* Predstavlenie grafov s assotsiativnymi operatsiyami na yazyke programmirovaniya SET@L [Representation of graphs with associative operations in the SET@L programming language], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2020, No. 3 (213), pp. 98-109. DOI: 10.18522/2311-3103-2020-3-98-111.
18. *Mikhaylov D.V.* Preobrazovanie nekotorykh vidov posledovatel'nykh informatsionnykh grafov v parallel'no-konveyernuyu formu [Transformation of some types of sequential information graphs into a parallel-pipeline form], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2020, No. 7 (217), pp. 78-93. DOI: 10.18522/2311-3103-2020-7-78-93.
19. *Mikhaylov D.V.* Preobrazovanie posledovatel'nogo informatsionnogo grafa metoda progonki v parallel'nyuyu formu [Transformation of a sequential information graph of the sweep method into parallel form], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2021, No. 7 (224), pp. 177-188. DOI: 10.18522/2311-3103-2021-7-177-188.
20. *Sorokin D.A., Dordopulo A.I.* Metodika sokrashcheniya apparatnykh zatrat v slozhnykh sistemakh pri reshenii zadach s sushchestvenno-peremennoy intensivnost'yu potokov dannykh [Methodology for reducing hardware costs in complex systems when solving problems with significantly variable intensity of data flows], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2012, No. 4, pp. 213-219.
21. *Batcher K.E.* Sorting Networks and their Applications, *Proc. AFIPS Spring Joint Comput. Conf.*, 1968, Vol. 32, pp. 307314.

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

Левин Илья Израилевич – Южный федеральный университет; e-mail: iilevin@sfedu.ru; г. Таганрог, Россия; тел.: 88634612111; зав. кафедрой; д.т.н.; профессор.

Алексеев Кирилл Николаевич – e-mail: alexseev91@mail.ru; тел.: 9283536268; к.т.н.; доцент.

Levin Ilya Izrailevich – Southern Federal University; e-mail: iilevin@sfedu.ru; Taganrog, Russia; phone: 88634612111; head of department; dr. of eng. sc.; professor.

Alekseev Kirill Nikolayevich – e-mail: alexseev91@mail.ru; phone: +79283536268; cand. of eng. sc.; professor.

Раздел II. Анализ данных и моделирование

УДК 004.896

DOI 10.18522/2311-3103-2023-5-119-126

А.В. Боженюк, М.В. Князева, О.В. Косенко, Е.Е. Косенко

ИСПОЛЬЗОВАНИЕ ПЕРИОДИЧЕСКИХ НЕЧЕТКИХ ГРАФОВ В ЗАДАЧАХ УПРАВЛЕНИЯ ПРОИЗВОДСТВЕННЫМ ОБОРУДОВАНИЕМ

Контроль износа оборудования является важной и актуальной задачей, требующей постоянного внимания и мониторинга. Это связано с тем, что износ оборудования может привести к снижению его эффективности его эксплуатации, повышению вероятности возникновения аварий или поломке. Задача замены оборудования является системной производственной задачей, при решении которой необходимо учесть множество факторов, влияющих на эффективность деятельности предприятия. В данной работе предложена концепция решения задачи своевременной замены оборудования с использованием периодических нечетких графов. Периодические нечеткие графовые модели позволяют адекватно отражать виды неопределенности, отражать специфику отношений между моделируемыми объектами, оптимизационные ограничения; решать различные задачи циклического типа, задачи сетевого планирования и управления. В данной работе применение математического аппарата периодических нечетких графов позволило задачу динамического характера, которая зависела от двух временных переменных свести к задаче, зависящей только от времени эксплуатации оборудования. При этом возраст оборудования предложено учитывать при определении коэффициента износа, определяемого степенью принадлежности к тому или иному классу износа. Данный аспект позволил свести поставленную задачу к исследованию износа всего оборудования, участвующего в технологическом процессе и учитывать покупку оборудования бывшего в употреблении (не нового). При определении функции принадлежности возможен учет факторов, которые могут повлиять на решение оптимизационной производственной задачи. Постановка задачи в нечетком виде обеспечивает возможность прогнозирования и планирования деятельности предприятия на будущие повторяющиеся периоды. В статье проведен литературный обзор, с обоснованием актуальности рассмотренной задачи. Рассмотрена задача замены оборудования с учетом динамических составляющих. Для отражения динамики производственного процесса были предложены темпоральные графы. Применение графовой модели обеспечивает наглядность отображения состояния оборудования в процессе его эксплуатации. Применение периодических нечетких графов позволяет задачу замены оборудования упростить за счет сокращения одной временной составляющей и масштабировать классическую динамическую задачу с учетом неопределенных исходных данных.

Производственный процесс; оптимизация; замена оборудования; нечеткость; функция принадлежности; периодический нечеткий граф.

A.V. Bozhenyuk, M.V. Knyazeva, O.V. Kosenko, E.E. Kosenko

USING PERIODIC FUZZY GRAPHS IN PRODUCTION EQUIPMENT CONTROL PROBLEMS

Monitoring equipment wear is an important and urgent task that requires constant attention and monitoring. This is due to the fact that wear and tear on equipment can lead to a decrease in its operating efficiency, an increased likelihood of accidents or breakdown. The task of replacing equipment is a systemic production task, in solving which it is necessary to take into account many

factors that influence the efficiency of the enterprise. This paper proposes a concept for solving the problem of timely replacement of equipment using periodic fuzzy graphs. Periodic fuzzy graph models make it possible to adequately reflect types of uncertainty, reflect the specifics of relationships between modeled objects, and optimization constraints; solve various cyclic problems, network planning and management problems. In this work, the use of the mathematical apparatus of periodic fuzzy graphs allowed a problem of a dynamic nature, which depended on two time variables, to be reduced to a problem that depended only on the operating time of the equipment. At the same time, it is proposed to take into account the age of equipment when determining the wear coefficient, determined by the degree of belonging to a particular wear class. This aspect made it possible to reduce the task to study the wear of all equipment involved in the technological process and consider the purchase of used (not new) equipment. When determining the membership function, it is possible to consider factors that may influence the solution of the optimization production problem. Setting the problem in a fuzzy form makes it possible to forecast and plan the activities of an enterprise for future recurring periods. The article contains a literature review, justifying the relevance of the problem considered. The problem of replacing equipment considering dynamic components is considered. Temporal graphs have been proposed to reflect the dynamics of the production process. The use of a graph model provides a clear display of the state of equipment during its operation. The use of periodic fuzzy graphs allows the problem of replacing equipment to be simplified by reducing one time component and to scale the classical dynamic problem considering uncertain initial data.

Production process; optimization; equipment replacement; fuzzy set; membership function; periodic fuzzy graph.

1. Введение. Организация производства предполагает скоординированное выполнение различных процессов, а именно: планирование операций, оперативное планирование и контроль. Обычно производство продукции состоит из процесса подготовки производства, из производственного процесса и процесса организации сбыта. При этом производственный процесс включает в себя множество различных операций, в том числе связанных с управлением технологического оборудования, техническим обслуживанием. При этом правильная организация производства способствует оптимизации рабочего процесса и, следовательно, повышает эффективность всего производственного процесса. Этого можно достичь за счет внедрения надлежащих систем планирования, составления графиков и контроля.

Эффективно организованная производственная система гарантирует, что каждая задача выполняется вовремя и в рамках бюджета, что приводит к более высокому уровню производительности [1, 2]. Это, в свою очередь, помогает снизить затраты на производство и максимизировать прибыль. Организация производства предполагает координацию и реализацию различных процессов. Однако эти задачи не лишены ограничений. Производственный процесс часто необходимо завершить в определенные сроки, и если эти сроки не соблюдаются, качество продукта может ухудшиться или вообще не быть законченным. Каждый производственный процесс, чтобы оптимизировать свои производственные циклы, повысить финансовые результаты, должен быть организован в соответствии с различными критериями и подходами.

Во многих публикациях отмечена возрастающая актуальность к задаче оптимизации производственных процессов. В [3–5] отмечена тенденция увеличения сложности серийных производственных линий и связанных с этим производственных затрат в современных отраслях с большими объемами производства. В [6] отмечено, что в последние годы увеличивается турбулентность внешней среды, в которой действуют компании, проявляющейся, в частности, в усилении динамичности рынков и глубоких изменениях в структуре потребительского спроса. Учитывая эти проблемы, в статье рассматриваются два отдельных пути улучшения управления производством: повышение экономической эффективности за счет

оптимизации размера партий и гибкие производственные системы за счет внедрения методов, позволяющих сократить время переналадки и замены оборудования. В [7, 8] рассмотрены системы сборочного производства. При этом авторы акцентируют на возможность объединения в рамках одного производственного процесса двух или более серийных производственных линий. Предложена оптимизационная модель на основе генетических алгоритмов. Отмечается возрастающая размерность данных задач, что накладывает ограничения на применение генетических алгоритмов. В [9] оптимизационная производственная задача определена как многокритериальная. Многоцелевая задача решается для оптимизации двух конфликтующих целей: максимизации производительности и минимизации энергопотребления при общей доступной емкости хранения. Представленные решения представляют собой компромисс между двумя целями, позволяя принимать решения, которые уравнивают максимизацию производительности с экономией энергии при проектировании производственных линий.

В данном контексте в литературе рассматриваются проблемы повышения эффективности и гибкости продукции, элементы, связанные с радикальными изменениями методов управления и технологий производства. Усовершенствование существующих и выработка новых подходов к оценке производственного процесса необходимо в условиях быстро изменяющейся номенклатуры производства. В большинстве подходов оптимизация процессов серийного производства основана на формализации задачи целесообразности замены оборудования в результате его изнашивания или устаревания. Актуальность данного направления рассмотрена в [10–15]. Модели сформулированы на основе стохастического динамического программирования, которые позволяют учесть динамику и неопределенность данной системы. При управлении промышленным производством одной из важнейших проблем является решение вопросов эффективного использования оборудования. В первом разделе рассматривается классическая задача динамического программирования по замене оборудования. Определено, что в классической постановке задача плохо масштабируется и возникают сложности, при учете факторов, влияющих на планирование производства. В данной постановке функция оптимальности зависит от стоимости оборудования и эксплуатационных затрат. Во втором разделе рассмотрена возможность применения нечетких темпоральных графов к решению задачи планирования производства в части замены оборудования. Определено, что при данном подходе задача может быть масштабируема, то есть рассматривается не одно оборудование, а несколько. Параметр – возраст оборудования, может быть комплексно учтен при определении износа оборудования через функцию принадлежности, как и другие параметры, влияющие на работоспособность оборудования. В последнем разделе рассматриваются выводы и будущие направления.

2. Постановка задачи замены оборудования. Машины и оборудование обычно представляют собой наибольшее распределение ресурсов в производственном процессе. Неправильные решения относительно замены оборудования часто могут стоить гораздо больше, чем вся экономия и снижение затрат, достигнутые во всех других областях производства и планирования. С возрастом как частота отказов, так и соответствующие затраты на техническое обслуживание значительно увеличиваются по мере того, как начинается износ и старение компонентов. Принятие оптимальных стратегических решений при планировании производственных процессов позволит обеспечить эффективность и конкурентоспособность производственной системы.

В классической постановке задача замены оборудования определена следующими параметрами [16–18]:

T – возраст оборудования, ($t=0, 1, 2, 3, \dots, n$), где $t=0$ – использование нового оборудования, $t=1$ – использование оборудования возраста одного периода и т.д.;

t_0 – начальный возраст оборудования;

n – продолжительность планового периода;

$r(t)$ – цена продукции, производимой на оборудовании возраста t за расчетный период (год);

$u(t)$ – эксплуатационные затраты за год на оборудование возраста t ;

$s(t)$ – остаточная стоимость оборудования возраста t ;

p_0 – цена нового оборудования.

Возможные управляющие воздействия:

♦ сохранить оборудование $r(t) - u(t)$;

♦ заменить оборудование (продать имеющееся оборудование и купить новое) $s(t) - p + r(0) - u(0) + F_{n-1}$.

Решение о замене имеющегося оборудования возраста t на новую принимается в том случае, когда доход от нового оборудования больше, чем от старого.

При построении модели задачи замены оборудования принято считать, что решение о замене выносится в начале каждого определенного промежутка эксплуатации (k). Возможные состояния оборудования S_k^t после каждого периода использования – либо его сохранность и продолжение эксплуатации (X^c), либо оборудование подлежит замене (X^z) (рис. 1).

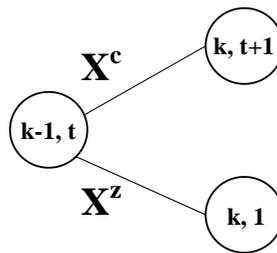


Рис. 1. Схема возможных состояний оборудования

$$S_k^t = \begin{cases} t + 1, & \text{если } X^k = X^c \\ t = 1, & \text{если } X^k = X^z \end{cases} \quad k = 1, 2, \dots, n$$

S_k^t – состояние оборудования, соответствующее возрасту t .

Критерием оптимальности данной задачи является либо прибыль от эксплуатации оборудования – задача максимизации, либо суммарные затраты на эксплуатацию в течение планируемого периода – задача минимизации. Функция определения суммарного дохода $F_n(t)$ описывается следующим образом [17, 18]:

$$F_n(t) = \max \begin{cases} r(t) - u(t) & \text{– сохранение оборудования } (X^c) \\ s(t) - p + r(0) - u(0) + F_{n-1} & \text{– замена оборудования } (X^z) \end{cases}$$

Тогда схема возможных состояний оборудования в течение некоторого эксплуатационного периода отобразится на графе следующим образом (рис. 2).

Управляющее решение по данному оборудованию определится в следующем виде: (X^c, X^c, X^c, \dots), то есть будут даны однозначные рекомендации по сохранению либо замене оборудования без учета допущений и обстоятельств, которые могут влиять на производственный процесс.

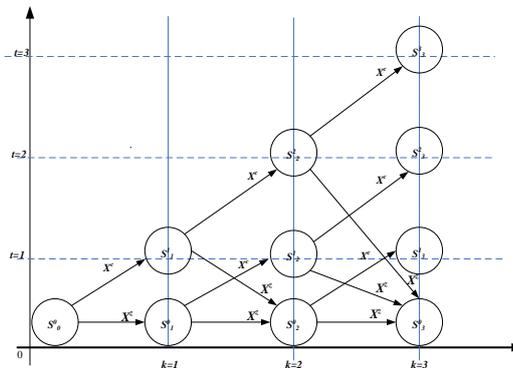


Рис. 2 Схема возможных состояний оборудования в течение некоторого эксплуатационного периода

3. Применение периодических нечетких графов к задаче замены оборудования. Рассмотренная задача относится к задачам динамического программирования. Состояние производственной системы, где объектами выступает оборудование, определяется возрастом оборудования t и фактическим временем использования оборудования k , то есть определяющим параметром является время. В статье [18] отмечено, что более половины авторов считают сетевую модель удобным инструментом для отображения жизненного цикла в производственном процессе. Применение графовой модели обеспечивает наглядность отображения состояния оборудования в процессе его эксплуатации. Состояние элементов данной модели изменяется в соответствии с некоторыми периодами времени. Для отражения динамики процесса на графах были предложены темпоральные графы. Для задачи целесообразности замены оборудования темпоральный граф запишем следующим образом:

$$G = (S_k^t, \{X_k\}, T),$$

где множество S_k^t является множеством вершин графа, множество натуральных чисел $t = \{1, 2, \dots, T\}$ определяет дискретное время, а множество $\{X_t\}$ задает семейство соответствий, которые отображают вершины S_k^t в моменты времени $t = \overline{1, T}$. Так как в динамической задаче замены оборудования обозначены две переменных, отражающих временной период (t – возраст оборудования; k – время эксплуатации данного оборудования), то определяющим моментом обозначим время k – как момент принятия решения о смене состояния используемого оборудования, либо его продолжения. Семейство соответствий (рис. 2) можно задать в виде:

$$\begin{aligned} X_1(S_0^0) &= \{S_1^0; S_1^1\}; \\ X_2(S_1^0) &= \{S_2^0; S_2^1\}; X_2(S_1^1) = \{S_2^0; S_2^2\}; \\ X_3(S_2^0) &= \{S_3^0; S_3^1\}; X_3(S_2^1) = \{S_3^0; S_3^2\}; \\ X_3(S_2^2) &= \{S_3^0; S_3^3\}; X_3(S_2^1) = \{S_3^0; S_3^2\}; X_3(S_2^2) = \{S_3^0; S_3^3\}. \end{aligned}$$

При принятии решения в момент k существует лишь два варианта: либо оборудование продолжает функционировать, либо его необходимо заменить. При стратегическом планировании производственной деятельности предприятия использование классического инструмента решения задачи замены оборудования не позволит учесть реалии производственного процесса. В современных условиях,

постоянно меняющихся логистических, ресурсных, материальных потоков, четко определить на прогнозный период необходимо ли менять оборудование или при определенных условиях оно может еще проработать – сложно. Необходимы допущения, которые смогут учесть неопределенность параметров задачи при планировании производственной деятельности.

В работах [20, 21] предложен математический аппарат, позволяющий учесть не только модель задачи в виде графа и временные отрезки, но и позволит записать исходные параметры задачи в нечетком виде. Периодический нечеткий граф $\tilde{G} = (S_k^t, \{\tilde{X}_k\}, T)$ определяется как нечеткий ориентированный граф $\tilde{G} = (S_k^t, \tilde{U}_t)$, в котором множество $\tilde{U}_t = \{\mu_t(s_i, s_j)\}$ – нечеткое множество ориентированных ребер в моменты $t = \overline{1, T}$, а функция μ_t есть функция принадлежности, определяющая степень износа оборудования в момент времени t (рис. 3). Причем, после момента $t = T$ опять наступает время $t = 1$ и период повторяется.

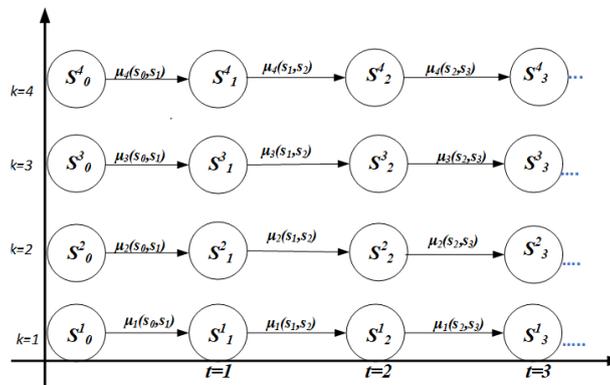


Рис. 3. Схема замены оборудования в серийном производстве

При этом функция принадлежности, определяющая степень износа оборудования, позволяет упростить задачу определения состояния производственного оборудования при организации серийного производства. В работе [17] упоминается, что к вектору управления производственными процессами необходимо добавить такой компонент как возраст устанавливаемого оборудования (не нового), что значительно увеличит размерность задачи. В нечеткой постановке такой параметр как возраст оборудования является излишним, так как $\mu_t(s_i, s_j)$ позволяет эти характеристики учесть. Тем самым, задача, зависящая от двух временных показателей, переходит в задачу так называемых «реперных точек», периодов, когда оборудование с учетом его возраста и состояния проходит оценку на пригодность. В данной постановке учитывается, что на производство может быть приобретено не только новое оборудование, но и оборудование с определенной степенью износа. В качестве управляющего решения будут выданы рекомендации по каждому периоду проверки оборудования на износ. В качестве решения могут быть определены рекомендации в соответствии с периодами использования оборудования, согласно рис. 3:

- ◆ в момент времени $t=1$: $\mu_1(s_0, s_1), \mu_2(s_0, s_1), \mu_3(s_0, s_1), \mu_4(s_0, s_1)$;
- ◆ в момент времени $t=2$: $\mu_1(s_1, s_2), \mu_2(s_1, s_2), \mu_3(s_1, s_2), \mu_4(s_1, s_2)$;
- ◆ в момент времени $t=3$: $\mu_1(s_2, s_3), \mu_2(s_2, s_3), \mu_3(s_2, s_3), \mu_4(s_2, s_3)$;

$$S_k^t = \begin{cases} \mu_t(s_i, s_j) < \delta, & \text{если } X^k = X^c \\ \mu_t(s_i, s_j) \geq \delta, & \text{если } X^k = X^z \end{cases} \quad k = 1, 2, \dots, n,$$

где δ – пороговое значение степени износа.

В зависимости от пороговых значений степени износа могут быть выданы рекомендации по каждому периоду контроля о необходимости замены оборудования либо его сохранении.

Вывод. Аппарат нечеткой математики, в том числе периодические нечеткие графы, позволили задачу замены оборудования упростить за счет сокращения второй временной составляющей. Данное обстоятельство позволяет рассматривать задачу замены оборудования не для одной, а для комплекса машин, что обеспечивает масштабируемость классической задачи динамического программирования с учетом неопределенных исходных данных.

В дальнейшем, при развитии данной темы планируется проработать математический аппарат определения устойчивости системы для каждого временного периода с учетом их неполноты данных, периодичности, рассмотреть возможность определения стратегии и обоснованность профилактического ремонта.

Подтверждения. Исследование было профинансировано Российским научным фондом, проект № 23-21-00206, <https://rscf.ru/en/project/23-21-00206> / реализовано Южным федеральным университетом.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Cochran D.S., Arinez J.F., Duda J.W., & Linck J. A decomposition approach for manufacturing system design, *Journal of manufacturing systems*, 2002, pp. 371-389.
2. Maher Rami & Hadidi Amani. Engineering project management planning and scheduling. *International Journal of Civil Engineering and Technology*, 2017, 8, pp.140-148.
3. Alkhalefah Hisham & Umer Usama & Abidi Mustufa & Elkaseer Ahmed. Development and Numerical Optimization of a System of Integrated Agents for Serial Production Lines, 2023, pp. 1-21.
4. Xi S., Chen Q., Smith J.M., Mao N., Yu, A., Zhang H. A new method for solving buffer allocation problem in large unbalanced production lines, *Int. J. Prod. Res.*, 2020, pp. 6846-6867.
5. Duan J., Li H., Zhang Q. Multiobjective optimization of buffer capacity allocation in multi-product unreliable production lines using improved adaptive NSGA-II algorithm, *Kuwait J. Sci.*, 2021, pp. 37-49.
6. Driga Imola & Dura Codruța & Isac Claudia. Possibilities of Increasing Efficiency Within Serial Production Management, *Annals of the University of Petrosani, Economics*, 2010, pp. 125-136.
7. Aboelfotoh Aaya & Süer Gürsel. Selection of Assembly Systems; Assembly Lines vs. Seru Systems, *Procedia Computer Science*, 2019, pp. 351-358.
8. Alkhalefah Hisham & Abu Qudeiri Jaber E. & Umer Usama & Abidi Mustufa & Elkaseer Ahmed. Development of an Efficient Prediction Model for Optimal Design of Serial Production Lines, *IEEE Access*, 2017, pp. 230-233.
9. Sadeghpour Houman & Tavakoli A. & Kazemi Mostafa & Pooya Alireza. A novel approximate dynamic programming approach for constrained equipment replacement problems: A case study, *Advances in Production Engineering & Management*, 2019, pp. 158-182.
10. Fan Wei & Machemehl Randy & Gemar Mason & Brown Leonard. A Stochastic Dynamic Programming Approach for the Equipment Replacement Optimization under Uncertainty, *Journal of Transportation Systems Engineering and Information Technology*, 2014, Vol. 14, pp. 76-84.
11. Zvipore David & Philimon Nyamugure & Maposa Daniel & Lesaoana Maseka. Application of the Equipment Replacement Dynamic Programming Model in Conveyor Belt Replacement: Case Study of a Gold Mining Company, *Mediterranean Journal of Social Sciences*, 2005, pp. 1123-1137.
12. Ukwu Chukwunenye. An Algorithm for Global Optimal Strategies and Returns in One Fell Swoop, for a Class of Stationary Equipment Replacement Problems with Age Transition Perspectives, Based on Nonzero Starting Ages, *Advances in Research*, 2016, Vol. 7, pp. 1-20.

13. *Sadeghpour Houman & Tavakoli A. & Kazemi Mostafa & Pooya Alireza*. A novel approximate dynamic programming approach for constrained equipment replacement problems: A case study, *Advances in Production Engineering & Management*, 2019, pp. 355-366.
14. *Abensur Eder & Santos Bruna & Bandeira Anselmo*. Optimization models as applied to equipment replacement problems: review and trends, *Gestão & Produção*, 2023, pp. 456-487.
15. *Sobczak-Piąstka Justyna & Matviishyn Yevhen & Famulyak Yuriy*. Planning of Rational Terms of the Construction Equipment Replacement, *IOP Conference Series: Materials Science and Engineering*, 2020, pp. 960-974.
16. *Sniedovich M*. Dynamic Programming: Foundations and Principles. Taylor & Francis, 2010, 624 p.
17. *Nocedal J., Wright S.J*. Numerical Optimization. Springer, 2006, 636 p.
18. *Russell S., Norvig P*. Artificial Intelligence: A Modern Approach. 3rd ed.. Prentice Hall, 2011. 1152 p.
19. *Abensur Eder & Santos Bruna & Bandeira Anselmo*. Optimization models as applied to equipment replacement problems: review and trends, *Gestão & Produção*, 2023, pp. 78-92.
20. *Gorbachev S., Bozhenyuk A., Nikashina P*. Optimization of Traffic Flow Based on Periodic Fuzzy Graphs, S. Bhattacharyya et al. (Eds.): *Intelligent Human Centered Computing*. Human 2023, Springer Tracts in Human-Centered Computing (STHC), 2023, pp. 374-383.
21. *Bozhenyuk A., Knyazeva M., Kosenko O., Rozenberg I.*: Strong Connectivity Definition of Periodic Fuzzy Graph, *Lecture Notes in Networks and Systems*, 2023, Vol. 758, pp. 168-174.

Статью рекомендовал к опубликованию д.т.н., профессор Н.Е. Сергеев.

Боженюк Александр Витальевич – Южный федеральный университет; e-mail: avbozhenyuk@sfedu.ru; г. Таганрог, Россия; тел.: +78634371743; кафедра информационно-аналитических систем безопасности имени профессора Л.С. Берштейна; д.т.н.; профессор.

Князева Маргарита Владимировна – e-mail: mknyazeva@sfedu.ru; тел.: +78634371743; кафедра информационно-аналитических систем безопасности имени профессора Л.С. Берштейна; к.т.н.; доцент.

Косенко Олеся Валентиновна – e-mail: ovkosenko@sfedu.ru; тел.: +78634371689; кафедра систем автоматического управления; к.т.н.; доцент.

Косенко Ева Евгеньевна – e-mail: kose@sfedu.ru; тел.: +7634371689; студент.

Bozhenyuk Alexander Vitalievich – Southern Federal University; e-mail: avbozhenyuk@sfedu.ru; Taganrog, Russia; phone: +78634371743; the department of information and analytical security systems named after professor Bershtein Leonid Samoylovich; dr. of eng. sc.; professor.

Knyazeva Margarita Vladimirovna – e-mail: mknyazeva@sfedu.ru; phone: +78634371743; the department of information and analytical security systems named after professor Bershtein Leonid Samoylovich; cand. of eng. sc.; associate professor.

Kosenko Olesya Valentinovna – e-mail: ovkosenko@sfedu.ru; phone: +78634371689; the department of automatic control systems; cand. of eng. sc.; associate professor.

Kosenko Eva Evgenievna – e-mail: kose@sfedu.ru; phone: +786343716-89; student.

В.В. Гилка, А.С. Кузнецова, А.А. Молдовская, Д.Ф. Эль-Аит

**ПРОВЕРКА РАБОТОСПОСОБНОСТИ МОДЕЛИ И МЕТОДА
УДАЛЕННОГО МОНИТОРИНГА СОСТОЯНИЯ ЗДОРОВЬЯ НА ПРИМЕРЕ
ОТКЛОНЕНИЙ ПОКАЗАТЕЛЕЙ ТЕМПЕРАТУРЫ ТЕЛА ЧЕЛОВЕКА**

Стремительное развитие телемедицины в сфере здравоохранения способствует активному внедрению различных методов и моделей удаленного мониторинга состояния показателей здоровья человека. В связи с этим большое внимание уделяется разработке мобильных приложений, которые способны обеспечить точный и своевременный мониторинг ключевых показателей здоровья в режиме реального времени. Цель данного исследования заключается в оценке применимости предлагаемого метода и модели для удаленного мониторинга состояния показателей здоровья человека, а также анализе эффективности разработанного мобильного приложения HelpMeTracker в выявлении отклонений в показателях температуры тела человека. Чтобы произвести оценку работоспособности предложенного способа мониторинга состояния здоровья был проведен комплексный эксперимент, который включал в себя участников из различных возрастных категорий и социальных групп. В ходе эксперимента приложение должно было анализировать и отслеживать динамику изменений в показателях температуры тела человека с применением датчиков, интегрированных в умные часы или фитнес-трекеры и своевременно оповещать об зафиксированных аномалиях для возможности быстрого реагирования на изменения в состоянии здоровья. Результаты, полученные в ходе проведенного исследования, демонстрируют, что предложенные модель и метод для удаленного мониторинга состояния показателей здоровья человека обладают высокой степенью эффективности для наблюдения за текущим состоянием здоровья. Во время эксперимента приложение HelpMeTracker надежно зафиксировало у испытуемых все отклонения в показателях температуры тела, полученных с датчиков носимых устройств и успешно проинформировало об этом всех участников процесса. На основании полученных результатов можно сделать вывод о том, что применение предложенного подхода для удаленного мониторинга достаточно, чтобы фиксировать отклонения в показателях температуры тела, отслеживать динамику изменения и формировать на основе полученной информации с датчиков устройств обоснованную комплексную оценку состояния здоровья человека.

Модель; метод; удаленный мониторинг; датчики; умные часы; фитнес-трекеры; показатели; отклонения.

V.V. Gilka, A.S. Kuznetsova, A.A. Moldovskaya, D.F. El-Ait

**VERIFICATION OF THE MODEL AND METHOD FUNCTIONALITY
FOR REMOTE HEALTH MONITORING ILLUSTRATED BY THE
DEVIATIONS IN HUMAN BODY TEMPERATURE INDICATORS**

The rapid development of telemedicine in the healthcare sector facilitates the active implementation of various methods and models for remote monitoring of human health indicators. In this regard, significant attention is paid to the development of mobile applications capable of providing accurate and timely monitoring of key health indicators in real-time. The aim of this research is to evaluate the applicability of the proposed method and model for remote monitoring of human health indicators, and to analyze the effectiveness of the developed mobile application HelpMeTracker in identifying deviations in human body temperature indicators. To assess the functionality of the proposed health monitoring method, a comprehensive experiment was conducted, which included participants from different age categories and social groups. During the experiment, the application was to analyze and track the dynamics of changes in human body temperature indicators using sensors integrated into smartwatches or fitness trackers, and timely notify about detected anomalies for the possibility of rapid response to changes in health status.

The results obtained during the conducted research demonstrate that the proposed model and method for remote monitoring of human health indicators have a high degree of effectiveness for observing the current health status. During the experiment, the HelpMeTracker application reliably detected all deviations in body temperature indicators obtained from sensors of wearable devices and successfully informed all participants of the process. Based on the obtained results, it can be concluded that the application of the proposed approach for remote monitoring is sufficient to capture deviations in body temperature indicators, track the dynamics of changes, and form a substantiated comprehensive assessment of human health based on the information received from the device sensors.

Model; method; remote monitoring; sensors; smartwatches; fitness trackers; indicators; deviations.

Введение. В сфере здравоохранения в настоящее время одним из важных направлений развития является телемедицина, позволяющая врачам консультировать, диагностировать и контролировать лечение своих пациентов, используя современные технологические решения и методы связи. Одним из перспективных направлений телемедицины является удаленный мониторинг пациентов (УМП), который предоставляет возможность на расстоянии собирать информацию о различных показателях здоровья человека в режиме реального времени с применением различных датчиков и медицинского оборудования, тем самым позволяя своевременно выявлять отклонения в самочувствии пациентов и принимать необходимые меры для предотвращения возможных осложнений [1, 2].

В настоящее время мобильные приложения (МП) для удаленного мониторинга (УМ) состояния пациентов стали неотъемлемой частью сферы здравоохранения. Согласно детальному аналитическому отчету о динамике рынка МП в области здравоохранения, наблюдается постоянное увеличение популярности и использования таких приложений во всем мире. Исследование показывает, что согласно показателю CAGR (совокупный среднегодовой темп роста), ожидается значительный прирост их применения на период с 2020 по 2027 год, составляющий 45,0% и это говорит о том, что разработка таких приложений, очень важна для мониторинга состояния здоровья [3]. Применение МП в области УМП позволяет надежно и эффективно контролировать различные физиологические показатели и параметры здоровья человека. Эти данные собираются с помощью современных высокотехнологичных носимых устройств и после передаются для дальнейшей обработки и анализа на мобильные устройства пользователей. Прогрессивное развитие этого сегмента рынка указывает на постоянно возрастающую потребность в улучшении качества и доступности медицинских услуг. Расширение возможностей дистанционного мониторинга пациентов способствует не только оптимизации процессов в медицинских учреждениях, но и улучшению самостоятельного контроля здоровья со стороны пациентов. Эти данные свидетельствуют о важности и актуальности разработки таких приложений.

Постановка задачи. Провести анализ предложенного метода и модели (УМП) [4] с целью оценки его применимости для отслеживания текущего состояния здоровья человека.

Провести экспериментальное исследование для оценки отклика разработанного мобильного приложения на различные отклонения в показателях температуры тела.

Удостовериться в работоспособности разработанного мобильного приложения «HelpMeTracker» [5–7] (на примере, одного из жизненно важных показателей (далее ЖВП) – температура тела человека, выделенного в ходе анализа возможностей современных устройств, таких как смарт-часы и фитнес-трекеры) [4].

Модель удаленного мониторинга состояния показателей здоровья человека. Разработанная модель процесса УМ с применением таких устройств как смарт-часы и фитнес-трекеры для области медицины в контексте удаленного мониторинга состояния показателей здоровья человека и представлена в соответствии с рис. 1.

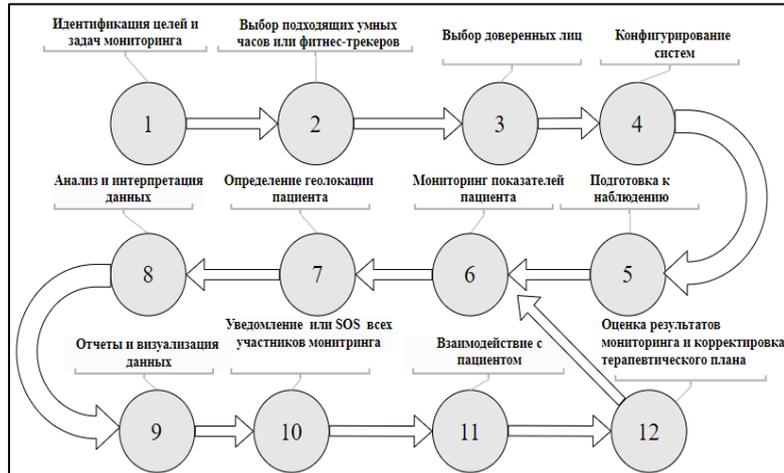


Рис. 4. Модель процесса УМ состояния показателей здоровья пациента

Научная новизна предложенной модели информационного процесса и ее отличие от существующих заключается в том, что был введен блок подготовки различных вариантов представления информации для различных классов субъектов, что дает возможность контроля состояния целевого объекта различными классами субъектов.

Представленная модель способствует решению многочисленных проблем, и обладает рядом следующих преимуществ:

- ◆ обеспечивает регулярный мониторинг физиологических параметров и состояния здоровья пациентов в режиме реального времени;
- ◆ данный тип носимых устройств использует неинвазивный подход, делая процесс мониторинга менее навязчивым;
- ◆ обладает высокой степенью точности, что обеспечивает достоверность собранных данных;
- ◆ способствует раннему выявлению потенциальных проблем в показателях здоровья;
- ◆ позволяет адаптировать подход к каждому человеку, учитывая его индивидуальные особенности и потребности;
- ◆ дает возможность предоставления медицинских услуг на расстоянии;
- ◆ способствует в обеспечении более эффективного контроля динамики изменений хронических заболеваний;
- ◆ определение геолокации человека, чтобы в случае отсутствия возможности установить с ним контакт при фиксации отклонений, отследить его текущее местоположение
- ◆ за счет систематического сбора и анализа данных о физиологических показателях и состоянии здоровья пациентов в рамках УМП создает основу для проведения научных исследований, способствуя развитию знаний в области медицины, позволяя выявлять новые закономерности и зависимости.

Метод удаленного мониторинга состояния показателей здоровья человека. Разработанный метод процесса УМ с применением таких устройств как смарт-часы и фитнес-трекеры для области медицины в контексте удаленного мониторинга состояния показателей здоровья человека и представлен в соответствии с рис. 2.

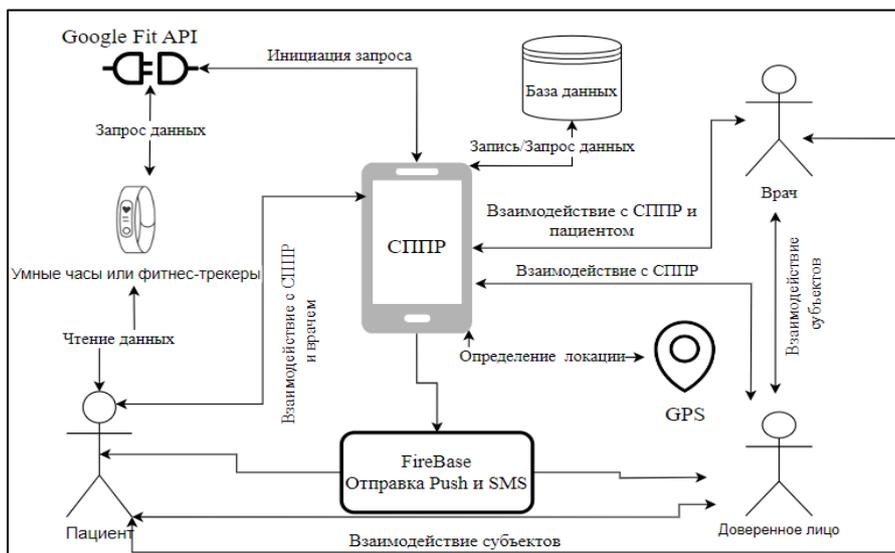


Рис. 2. Метод удаленного мониторинга пациента

Научная новизна предложенного метода сбора и анализа гетерогенной информации о состоянии целевого объекта и его отличие от известных заключается в едином источнике гетерогенных показателей, что обеспечивается введением блока формирования гетерогенных показателей - один датчик много свойств.

В представленном методе мы можем увидеть то, каким образом происходит процесс получения данных с носимых устройств таких как смарт-часы или фитнес-трекеры (браслеты). Приложением отправляется запрос к датчикам смарт-часов используя API Google Fit [8]. Датчики отправляют показания в Google Fit, мы их считываем, анализируем, перенаправляем в приложение в виде отчета, где они становятся видны пациенту и врачу. Также эти данные сохраняются на сервере. Далее, как мы можем увидеть, что приложение дублирует отчет о ЖВП наблюдателю (к примеру родственник). Отправка отчета наблюдателю будет осуществляться с использованием Firebase, который предоставляет возможность отправлять push и SMS уведомления. Такое дублирование необходимо для своевременной доставки отчета, если будет слабый сигнал LTE [9] или доступ к интернету отсутствует. Также наблюдателю доступен просмотр показателей из приложения наблюдаемого. В случае экстренной ситуации, когда полученные ЖВП выходят за пределы нормы, приложение реагирует на это, и отправляет тревожный сигнал, наблюдаемому, наблюдателю и врачу если данная опция активирована у него в приложении. На основе данных полученных с GPS [10] мы сможем определять текущее местоположение человека (к примеру, это необходимо тогда, если приложение зафиксировало критические отклонения в ЖВП и при этом мы не можем установить связь с наблюдаемым, чтобы узнать по геолокации его местоположение) [4, 5, 11].

Увидеть, внутреннюю структуру того, как это устроено в приложении можно на диаграмме классов концептуального уровня, представленной в соответствии с рис. 3.

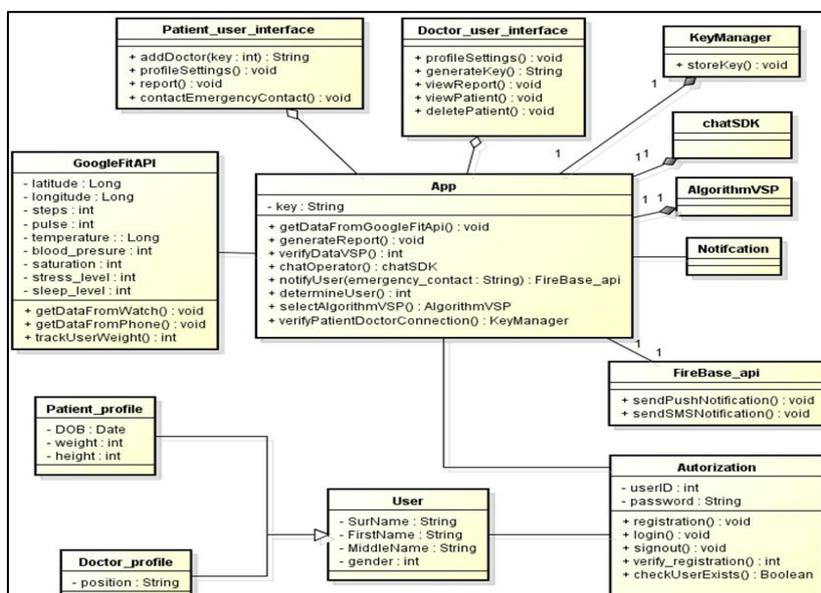


Рис. 3. Архитектура приложения в нотации UML

Детальное описание архитектуры приводится в статье [5].

Эксперимент и анализ полученных результатов. В процессе исследования, основываясь на анализе медицинской литературы и консультациях с медицинскими специалистами, были определены возрастные границы, категории показателей, а также диапазоны верхних и нижних пределов значений, используемые в разработанном приложении [12–20]. В табл. 1 представлена выделенная классификация по возрастам.

Таблица 1

Классификация возрастов

Название категории	Диапазон возрастов
Молодой возраст	18-29
Зрелый возраст	30-39
Средний возраст	40-59
Пожилой	60 +

В табл. 2 представлены категории показателей, а также диапазоны показателей уровня температуры тела верхних и нижних границ для каждой из них.

Таблица 2

Границы показателей температуры тела

Категории Возраст	Норма	В пределах нормы	Умеренные		Критическое	
			Ниж. гр.	Верх. гр.	Ниж. гр.	Верх. гр.
18+	36-36,6	35-37	34,5-35	37-39,5	34,5	39,5

На основе данных, из табл. 2 были определены конкретные значения, представленные в табл. 3 для каждой возрастной категории, при достижении которых должно осуществляться реагирование на отклонения в показателях.

Таблица 3

Значения реагирования в показателях температуры тела

Категории Возраст	Норма	В пределах нормы	Умеренные (Push)		Критическое (Push, SMS, SOS)	
			Ниж. гр.	Верх. гр.	Ниж. гр.	Верх. гр.
18+	-	-	≥ 34,6 и ≤ 34,9	≥ 37 и ≤ 39,4	≤ 34,5	≥ 39,5

Эксперимент проводился в течении 4 месяцев и для его проведения применялись браслеты «Huawei Watch GT 3», «Huawei Watch 3 Pro», «Huawei Watch D». Для корректной работы приложения и получения максимально точных данных с датчиков браслета, испытуемые и наблюдатели должны были соблюдать следующие условия:

- ◆ вести свою обычную повседневную деятельность;
- ◆ исключить физические нагрузки;
- ◆ полностью соблюдать предписания врача по приему лекарственных препаратов, если таковые имеются.

Наблюдатели во время эксперимента должны были выполнять следующие действия:

- ◆ фиксировать, когда приходит простое push уведомление или включается режим экстренного реагирования (SOS) и фиксировать показатели с временем, когда это произошло, любым из удобных способов;
- ◆ установка связи с наблюдаемым в случае, когда приложение реагирует, чтобы уточнить состояние человека [11].

Общее количество принимающих участие в тесте приложения составило 19 человек. В категорию 18-29 лет вошли 6 человек. В возрасте 30-39 лет вошли 5 человек. От 40-59 лет 4. В группе от 60 – 4 испытуемых.

На основе проведенного эксперимента были получены следующие результаты. В возрастной категории 18-29 лет аномалии были определены у трех людей, данные которых представлены в соответствии с рис. 4.

Из полученных данных человека «а» мы можем увидеть, что первое отклонение в состоянии температуры тела, было зафиксировано в послеобеденное время на уровне 37 °С и после она становилась все больше.

Установив контакт с наблюдаемым, выяснили, что он чувствует себя очень плохо. Была вызвана бригада скорой помощи, которая определила у него пищевое отравление. У испытуемого «б» скачек температуры был определен на отметке 37,7 °С. В последующие дни скачки температуры наблюдались регулярно и сопровождалась кашлем с мокротой. В ходе долгого разбирательства с врачами, было установлено, что это аллергическая реакция на увлажнителя воздуха, приобретенный им в период эксперимента. У наблюдаемого «в» отклонение в показателях было зафиксировано в ночной период времени на отметке в 37,4 °С и держалась высокой на протяжении всех суток. В следствии чего в последующие дни было установлено, что пациента covid-19.

В возрастной категории 30-39 лет отклонения были зафиксированы у одного человека, данные которого представлены в соответствии с рис. 5. У наблюдаемого «г» аномалия в отклонении температуры тела была определена на уровне 37,5 °С. В дальнейшем было установлено, что такой скачек у него стал следствием респираторной вирусной инфекцией.

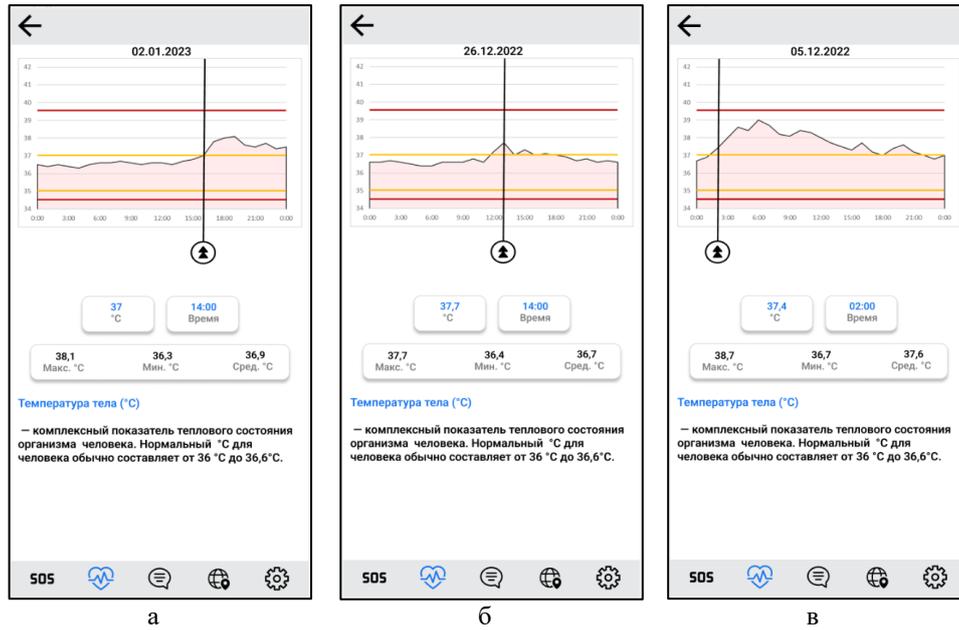


Рис. 4. Данные испытуемых из возрастной категории 18-29 лет

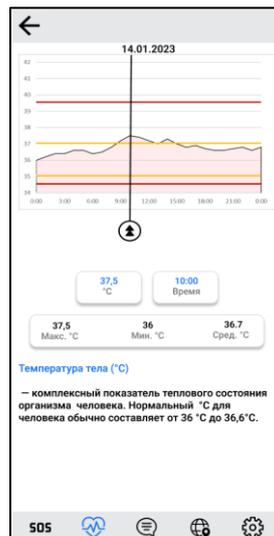


Рис. 5. Данные испытуемых из возрастной категории 30-39 лет

В категории 40-59 лет отклонения были зафиксированы у двух людей, данные которых представлены на рис. 6.

У испытуемого «д» на основе полученных сведений отклонения были зафиксированы во второй половине дня, где максимальный показатель температуры был определен на отметке в 39 °С. В ходе установленного контакта с человеком, выяснили, что он чувствует себя плохо и его лихорадит. В последующем испытуемый обратился за помощью в больницу, где ему был поставлен диагноз бронхит и оказана первая медицинская помощь. У наблюдаемого «д» отклонения начали фикси-

роваться в первой половине дня на показателе 37,2 градуса и держалась на протяжении всех суток. В ходе выяснения причины такой аномалии, врачами была определена у испытуемого респираторная вирусная инфекция.

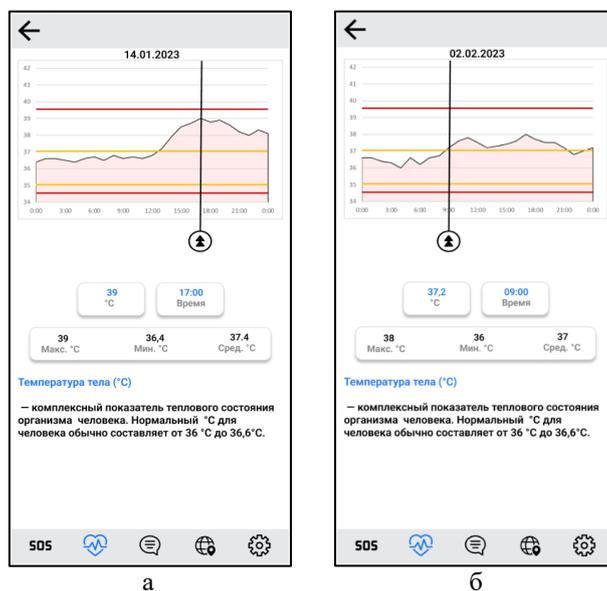


Рис. 6. Данные испытуемых из возрастной категории 40-59 лет

В категории 60+ лет приложением не были зафиксированы отклонения в показателях °C.

В ходе эксперимента, у всех категории пользователей «HelpMeTracker» были определены отклонения в состоянии температуры тела. Показатели за рамки умеренных значений не выходили и приложение успешно проинформировало об зафиксированных аномалиях наблюдателей с помощью push уведомления.

Заключение. В ходе исследования было продемонстрировано, что предложенная модель и метод успешно выполняют задачу мониторинга температуры тела человека на удаленном расстоянии. Разработанная система "HelpMeTracker" на их основе способна фиксировать отклонения и изменения в показателях температуры с достаточно высокой точностью. Результаты исследования подтверждают работоспособность модели и метода и их потенциал для применения в медицинской практике. Это может способствовать улучшению качества медицинского обслуживания и обеспечению более эффективного контроля за здоровьем пациентов. Научная новизна исследования заключается в предложенной модели, методе и СППР, работающей на их основе, позволяющей проводить удаленный мониторинг состояния показателей температуры тела в режиме реального времени и контроль общего состояния пациента.

Система зафиксировала все аномалии в показателях, получаемых с датчиков браслета, помимо тех, которые обзревались, и осведомляла об этом наблюдателей. Таким образом, эксперимент по оценке эффективности разработанной модели и метода, а также функциональности СППР "HelpMeTracker" для удаленного мониторинга состояния пациента, по проверке отклика приложения на различные отклонения в показателях температуры тела на реальных участниках можно считать успешным. Эксперимент показал, что предложенная модель и метод удаленного мониторинга пациентов являются полностью пригодным для контроля текущего состояния человека.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Farias F.A.C., Dagostini C.M., Bicca Y.A., Falavigna V.F., Falavigna A.* Remote Patient Monitoring: A Systematic Review. *Telemed J.E. Health.* – 2020 May. – 26 (5). – P. 576-583. – DOI: 10.1089/tmj.2019.0066. Epub 2019 Jul 17. PMID: 31314689.
2. *Stranieri Andrew & Balasubramanian Venki.* Remote Patient Monitoring for Healthcare. – 2019. – 10.4018/978-1-5225-7277-0.ch009.
3. Healthcare Mobile Application Market Size, Share & Trends Analysis Report By Type (Fitness Products Training, Appointment Booking & Construction), By Platform, By Technology, By End User, And Segment Forecasts, 2020-2027 // Grandviewresearch. – URL: <https://www.grandviewresearch.com/industry-analysis/healthcare-mobile-applications-market/methodology> (accessed: 01.02.2023).
4. *Гилка В.В., Орлова Ю.А., Хужахметова Д.Х., Кузнецова А.С.* Разработка метода и архитектуры мобильного приложения для удаленного мониторинга текущего состояния человека на основе данных жизненно важных показателей с датчиков // Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение. – 2021. – Т. 11, № 4. – С. 212-230.
5. *Gilka V.V., Kachanov Y.A., Kuznetsova A.S.* Architecture of the Android Application for Monitoring Person's Condition Based on Data Readings from Sensors of Smart Watches and Mobile Devices / Krouska A., Troussas C., Caro J. (eds) // *Novel & Intelligent Digital Systems: Proceedings of the 2nd International Conference (NiDS 2022): Lecture Notes in Networks and Systems.* – Vol 556. – Springer, Cham, 2023. – https://doi.org/10.1007/978-3-031-17601-2_22.
6. *Гилка В.В.* Общая архитектура взаимодействия компонентов Android приложения HelpMeTracker для удаленного отслеживания текущего состояния человека // России – творческую молодежь: Матер. XV Всерос. науч.-практ. студ. конф. (г. Камышин, 20-22 апреля 2022 г.). В 4 т. Т. 3 / под общ. ред. И.В. Степанченко; ВолгГТУ, КТИ (филиал) ВолгГТУ. – Волгоград, 2022. – С. 13-16.
7. Свидетельство о государственной регистрации программы для ЭВМ № 2023613348 Российской Федерация. HelpMeTracker: № 2023611626: заявл. 30.01.2023; опубл. 14.02.2023 / В.В. Гилка.
8. *Chougale Pankaj & Yadav Vaibhav & Gaikwad Anil & Student Bharati & Vidyapeeth.* Fire-base - overview and usage // *Journal of Engineering and Technology Management.* – 2022. – 2582-5208.
9. *Tanenbaum E., Weatherall D.* Computer networks. – 5th ed. – St. Petersburg: Piter, 2012. – 960 p.
10. *Mamdyal Miss & Sandupatla Miss & Saka I & Kothawale J & Shirashayad V & Kazi K.* GPS Tracking System. – 2022. – 2581-9429.
11. *Гилка В.В., Кузнецова А.С.* Тестирование работоспособности метода удаленного мониторинга, реализованного в HelpMeTracker на людях, и проверка реагирования приложения на отклонения в показателях здоровья // Вестник Дагестанского государственного технического университета. Технические науки. – 2023. – Т. 50, № 2. – С. 48-57. – DOI: 10.21822/2073-6185-2023-50-2-48-57.
12. *Покровский В.М.* Физиология человека: учебник / под ред. В.М. Покровского, Г.Ф. Коротько. – 3-е изд. – М.: Медицина, 2011. – 664 с. – ISBN 978-5-225-10008-7.
13. *Дудель Й., Рюэгг Й., Шмидт Р. [и др.].* Физиология человека. В 3 т. Т. 1. – М.: Мир, 1996. – 323 с. – ISBN 5-03-002544-8.
14. *Циммерман М., Ениг В., Вутке В. [и др.].* Физиология человека. В 3 т. Т. 2. – М.: Мир, 1996. – 313 с. – ISBN 5-03-002544-8.
15. *Ульмер Х.-Ф., Брюк К., Эве К. [и др.].* Физиология человека. В 3 т. Т. 3. – М.: Мир, 1996. – 198 с. – ISBN 5-03-002544-8.
16. *Гайворонский И.В., Ничипорук Г.И., Гайворонский А.И.* Анатомия и физиология человека: учебник для студ. учреждений сред. проф. образования. – 6-е изд., перераб. и доп. – М.: Издательский центр «Академия», 2011. – 496 с.
17. *Агаджанян Н.А., Смирнов В.М.* Нормальная физиология: учебник. – 3-е изд., испр. и доп. – М.: ООО «Изд-во «Медицинское информационное агентство», 2012. – 576 с.
18. *Зильбернагель С., Деспонулос А.* Наглядная физиология: пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2013. – 408 с.

19. Орлов П.С. Нормальная физиология: учебник. – 2-е изд., испр. и доп. – М.: ГЭОТАР-Медиа, 2010. – 832 с.
20. Агаджанян Н.А., Смирнов В.М. Нормальная физиология: учебник для студентов медицинских вузов. – М.: ООО «Изд-во «Медицинское информационное агентство», 2009. – 520 с.

REFERENCES

1. *Farias F.A.C., Dagostini C.M., Bicca Y.A., Falavigna V.F., Falavigna A.* Remote Patient Monitoring: A Systematic Review. *Telemed J.E. Health*, 2020 May, 26 (5), pp. 576-583. DOI: 10.1089/tmj.2019.0066. Epub 2019 Jul 17. PMID: 31314689.
2. *Stranieri Andrew & Balasubramanian Venki.* Remote Patient Monitoring for Healthcare. 2019. 10.4018/978-1-5225-7277-0.ch009.
3. Healthcare Mobile Application Market Size, Share & Trends Analysis Report By Type (Fitness Products Training, Appointment Booking & Construction), By Platform, By Technology, By End User, And Segment Forecasts, 2020-2027, Grandviewresearch. Available at: <https://www.grandviewresearch.com/industry-analysis/healthcare-mobile-applications-market/methodology> (accessed 01 February 2023).
4. *Gilka V.V., Orlova Yu.A., Khuzhakhmetova D.Kh., Kuznetsova A.S.* Razrabotka metoda i arkhitektury mobil'nogo prilozheniya dlya udalennogo monitoringa tekushchego sostoyaniya cheloveka na osnove dannykh zhiznenno vazhnykh pokazateley s datchikov [Development of a method and architecture of a mobile application for remote monitoring of a person's current state based on data from vital signs from sensors], *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta. Seriya: Upravlenie, vychislitel'naya tekhnika, informatika. Meditsinskoe priborostroenie* [News of the South-West State University. Series: Management, computer technology, computer science. Medical instrumentation], 2021, Vol. 11, No. 4, pp. 212-230.
5. *Gilka V.V., Kachanov Y.A., Kuznetsova A.S.* Architecture of the Android Application for Monitoring Person's Condition Based on Data Readings from Sensors of Smart Watches and Mobile Devices, Krouska A., Troussas C., Caro J. (eds), *Novel & Intelligent Digital Systems: Proceedings of the 2nd International Conference (NiDS 2022): Lecture Notes in Networks and Systems*, Vol 556. Springer, Cham, 2023. Available at: https://doi.org/10.1007/978-3-031-17601-2_22.
6. *Gilka V.V.* Obshchaya arkhitektura vzaimodeystviya komponentov Android prilozheniya HelpMeTracker dlya udalennogo otslezhivaniya tekushchego sostoyaniya cheloveka [General architecture of interaction between components of the Android application HelpMeTracker for remote tracking of a person's current state], *Rossii – tvorcheskuyu molodezh': Mater. XV Vseros. nauch.-prakt. stud. konf. (g. Kamyshin, 20-22 aprelya 2022 g.)* [Russia – creative youth: Proceedings of the XV All-Russian scientific and practical student conference (Kamyshin, April 20-22, 2022)]. In 4 vol. Vol. 3, under general. ed. I.V. Stepanchenko; VolgSTU, KTI (branch) VolgSTU. Volgograd, 2022, pp. 13-16.
7. *Gilka V.V.* Svidetel'stvo o gosudarstvennoy registratsii programmy dlya EVM № 2023613348 Rossiyskaya Federatsiya. HelpMeTracker: № 2023611626: zayavl. 30.01.2023: opubl. 14.02.2023 [Certificate of state registration of a computer program No. 2023613348 Russian Federation. HelpMeTracker: No. 2023611626: declared 01/30/2023: published 02/14/2023]
8. *Chougale Pankaj & Yadav Vaibhav & Gaikwad Anil & Student Bharati & Vidyapeeth.* Fire-base - overview and usage, *Journal of Engineering and Technology Management*, 2022, 2582-5208.
9. *Tanenbaum E., Weatherall D.* Computer networks. 5th ed. St. Petersburg: Piter, 2012, 960 p.
10. *Mamdyal Miss & Sandupatla Miss & Saka I & Kothawale J & Shirashayad V & Kazi K.* GPS Tracking System, 2022, 2581-9429.
11. *Gilka V.V., Kuznetsova A.S.* Testirovanie rabotosposobnosti metoda udalennogo monitoringa, realizovannogo v HelpMeTracker na lyudyakh, i proverka reagirovaniya prilozheniya na otkloneniya v pokazatelyakh zdorov'ya [Testing the performance of the remote monitoring method implemented in HelpMeTracker on people and checking the application's response to deviations in health indicators], *Vestnik Dagestanskogo gosudarstvennogo tekhnicheskogo universiteta. Tekhnicheskie nauki* [Bulletin of the Dagestan State Technical University. Technical science], 2023, Vol. 50, No. 2, pp. 48-57. DOI: 10.21822/2073-6185-2023-50-2-48-57.

12. *Pokrovskiy V.M.* Fiziologiya cheloveka: uchebnik [Human physiology: textbook], ed. by V.M. Pokrovskogo, G.F. Korot'ko. 3rd ed. Moscow: Meditsina, 2011, 664 p. ISBN 978-5-225-10008-7.
13. *Dudel' Y., Ryuegg Y., Shmidt R. [i dr.]*. Fiziologiya cheloveka [Human physiology]. In 3 vol. Vol. 1. Moscow: Mir, 1996, 323 p. ISBN 5-03-002544-8.
14. *Tsimmerman M., Enig V., Vutke V. [i dr.]*. Fiziologiya cheloveka [Human physiology]. In 3 vol. Vol. 2. Moscow: Mir, 1996, 313 s. ISBN 5-03-002544-8.
15. *Ul'mer Kh.-F., Bryuk K., Eve K. [i dr.]*. Fiziologiya cheloveka [Human physiology]. In 3 vol. Vol. 3. Moscow: Mir, 1996, 198 p. ISBN 5-03-002544-8.
16. *Gayvoronskiy I.V., Nichiporuk G.I., Gayvoronskiy A.I.* Anatomiya i fiziologiya cheloveka: uchebnik dlya stud. uchrezhdeniy sred. prof. obrazovaniya [Human anatomy and physiology: a textbook for students. institutions of secondary vocational education]. 6th ed., revised and expanded. Moscow: Izdatel'skiy tsentr «Akademiya», 2011, 496 p.
17. *Agadzhanian N.A., Smirnov V.M.* Normal'naya fiziologiya: uchebnik [Normal physiology: textbook]. 3rd ed., corrected and expanded. Moscow: ООО «Izd-vo «Meditsinskoe informatsionnoe agentstvo», 2012, 576 p.
18. *Zil'bernagl' S., Despopulos A.* Naglyadnaya fiziologiya [Visual physiology]: trans. from engl. Moscow: BINOM. Laboratoriya znaniy, 2013, 408 p.
19. *Orlov R.S.* Normal'naya fiziologiya: uchebnik [Normal physiology: textbook]. 2nd ed., corrected and expanded. Moscow: GEOTAR-Media, 2010, 832 p.
20. *Agadzhanian N.A., Smirnov V.M.* Normal'naya fiziologiya: uchebnik dlya studentov meditsinskikh vuzov [Normal physiology: a textbook for medical students]. Moscow: ООО «Izd-vo «Meditsinskoe informatsionnoe agentstvo», 2009, 520 p.

Статью рекомендовала к опубликованию д.т.н., профессор М.А. Бутакова.

Гилка Вадим Викторович – Волгоградский государственный технический университет; e-mail: gilka_vv@mail.ru; г. Волгоград, Россия; кафедра программного обеспечения автоматизированных систем; старший преподаватель.

Кузнецова Агнесса Сергеевна – e-mail: agnessakyz@yandex.ru; кафедра программного обеспечения автоматизированных систем; старший преподаватель.

Эль-Аит Джалалиддин Фирасович – e-mail: djalaliddin@yandex.ru; стоматологический факультет; студент.

Молдовская Ангелина Алексеевна – ГУ "Рыбницкое ЦРБ"; e-mail: amoldovskaia@gmail.com; г. Рыбница, Молдова; главный врач.

Gilka Vadim Viktorovich – Volgograd State Technical University; e-mail: gilka_vv@mail.ru; Volgograd, Russia; the department of software for automated systems; senior lecturer.

Kuznetsova Agnessa Sergeevna – e-mail: agnessakyz@yandex.ru; the department of software for automated systems; senior lecturer.

El-Ait Jalaliddin Firassovich – e-mail: djalaliddin@yandex.ru; faculty of dentistry; student.

Moldovskaya Angelina Alexeevna – Rybnitskoe Central City Hospital; e-mail: amoldovskaia@gmail.com; Rybnitsa, Moldova; head doctor.

Л.А. Гладков, Н.В. Гладкова, В.М. Курейчик

ПОДСИСТЕМА АВТОМАТИЧЕСКОГО АННОТИРОВАНИЯ ТЕКСТОВ НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ*

Рассматривается задача автоматического аннотирования текстов. Рассмотрена постановка задачи. Обоснована актуальность и важность разработки эффективных методов и программных систем для решения задачи автоматического реферирования текстов в современных информационных системах. Приведены определения понятий «данные» и «знания». Описан перечень задач, относящихся в направлении Data Mining. Подробно описана задача Text Mining и существующие методы ее решения. Рассмотрена задача реферирования текстов. Выделены основные этапы решения задачи суммаризации. Описаны основные методы автоматической обработки текста, выделены их достоинства и недостатки. Подробно рассмотрены методы реферирования и квазиреферирования. Проведен сравнительный анализ эффективности различных методов реферирования и квазиреферирования, выделены их ключевые достоинства и недостатки. Приведено краткое описание архитектуры encoder-decoder с точки зрения использования данной архитектуры в разрабатываемом алгоритме автоматического реферирования текстов. Приведено описание модели рекуррентных нейронных сетей, отмечены достоинства и недостатки подобных моделей. Рассмотрены архитектуры рекуррентной нейронной сети применительно к решению задачи автоматического реферирования текстов. Приведено описание модифицированной модели рекуррентной нейронной сети – нейронной сети долгой краткосрочной памятью. Приведено описание предложенного алгоритма автоматического реферирования и значения настроек его основных параметров. Приведено описание разработанной программной подсистемы автоматического реферирования. Выполнено компьютерное моделирование и приведены результаты, полученные в ходе вычислительных экспериментов. Выполнена оценка качества полученных решений. Определены оптимальные параметры разработанной программной системы. Сформулированы направления продолжения исследований.

Реферирование текстов; суммаризация; методы реферирования и квазиреферирования; рекуррентные нейронные сети; токенизация; стемминг; нейронные сет долгой краткосрочной памяти.

L.A. Gladkov, N.V. Gladkova, V.M. Kureichik

SUBSYSTEM FOR AUTOMATIC TEXT ANNOTATION BASED ON MACHINE LEARNING METHODS

This paper considers the problem of automatic text annotation. The formulation of the problem is considered. The relevance and importance of developing effective methods and software systems for solving the problem of automatic text summarization in modern information systems is substantiated. Definitions of the concepts “data” and “knowledge” are given. A list of tasks related to the Data Mining direction is described. The Text Mining problem and existing methods for solving it are described in detail. The problem of summarizing texts is considered. The main stages of solving the summation problem are highlighted. The main methods of automatic text processing are described, their advantages and disadvantages are highlighted. Abstractive summarization and extractive summarization methods are discussed in detail. A comparative analysis of the effectiveness of various abstracting and quasi-abstracting methods has been carried out, their key advantages and disadvantages have been highlighted. A brief description of the encoder-decoder architecture is given from the point of view of using this architecture in the developed algorithm

* Исследование выполнено за счет гранта Российского научного фонда № 22-21-00316, <https://rscf.ru/project/22-21-00316/>.

for automatic text summarization. A description of the model of recurrent neural networks is given, the advantages and disadvantages of such models are noted. The architecture of a recurrent neural network is considered in relation to solving the problem of automatic text summarization. A description of the modified model of a recurrent neural network – a neural network with long short-term memory – is given. A description of the proposed automatic abstracting algorithm and the settings of its main parameters are given. A description of the developed automatic abstracting software subsystem is given. Computer modeling is performed and the results obtained during computational experiments are presented. The quality of the solutions obtained was assessed. The optimal parameters of the developed software system are determined. Directions for continuing research are formulated.

Text summarization; text mining; abstractive summarization; extractive summarization methods; recurrent neural networks; tokenization; stemming; long short-term memory networks.

Введение. В современном мире информация играет ключевую роль во многих сферах. Количество информации растет в геометрической прогрессии, особенно в сети Internet. Мировая сеть содержит около 80-90% неструктурированной информации. Человечество старается оцифровать как можно большее количество такой информации в целях хранения, экономии места, удобства использования [1]. Для этого разрабатываются новые технологии хранения, создаются новые базы и хранилища данных. Сравнительно недавно появились и стали набирать популярность облачные хранилища ввиду удобства использования и хранения информации.

Далеко не вся подобная информация одинаково полезна и, более того, структурирована, что приводит к проблеме появления огромного количества данных, которые не поддаются ручной обработке. Для решения подобных проблем и получения профессиональных знаний используются методы data mining [2].

Стоит определить, что же такое «профессиональные знания». По сути – это сведения, которые характеризуют особенности деятельности и позволяют реализовать эту самую деятельность максимально эффективно. В связи с динамически развивающимися условиями рынка преимуществом современного специалиста является максимально широкий охват знаний в данной области. Например, для наиболее эффективного осуществления деятельности электронных магазинов необходимо обрабатывать большое количество неструктурированной информации (форумы, отзывы о продукции, информация о клиентах, пользовательский контент и т.д.). Поэтому актуальной является разработка предварительно обученной подсистемы, способной самостоятельно обрабатывающей пользовательские отзывы с целью получения короткого предложения, характеризующего эмоциональную окраску и смысл исходного экземпляра.

В зависимости от стоящих задач определяется подход к реферированию информации. В настоящее время существует большое количество алгоритмов для обработки неструктурированной информации. Одним из популярных подходов к решению данной задачи является использование нейронных сетей различных видов. Свою популярность они приобрели за счет эффективности работы в конкретной, узкоспециализированной сфере деятельности. Например, компания Google в 2018 представила новую технологию обработки естественного языка BERT (Bidirectional Encoder Representations from Transformers). Однако подобные системы имеют существенные недостатки, например, высокая требовательность к качеству информации, содержащейся в обучающей выборке. Это подтверждает вывод о востребованности подобных систем на основе нейронных сетей новой архитектуры и, во-вторых, о необходимости их совершенствования.

Методы решения задачи автоматического реферирования. Data Mining (с англ. «добыча данных») – это автоматизированный поиск полезной информации среди огромного количества массивов неструктурированной информации. Основной задачей Data Mining является определение паттернов и взаимосвязей, выявляе-

ния которых невозможно добиться при обычном анализе [2, 3]. Для решения подобных задач используют многочисленные математические и статистические алгоритмы. На данный момент существует множество алгоритмов (технологий), связанных с Data Mining – деревья решений, генетические алгоритмы, нейронные сети и производные.

Можно выделить три основных сферы применения Data Mining: Information Retrieval, Text Mining и Web Mining.

Information Retrieval («добыча данных») – группа методов, ориентированных на получение структурированных данных или же выборки данных меньшего размера [4–6]. Например, поисковая система, которая с помощью заданных алгоритмов позволяет получить необходимую информацию из большого массива документов. В качестве инструментов обычно выступают парсеры, методы индексации, фильтрации и сортировки данных.

Text Mining («добыча данных из текста») – это частный случай Data Mining, ориентированный на извлечение информации из текстов. В отличие от Information Retrieval, задачей в области Text Mining является анализ имеющихся данных с помощью математических методов, что позволяет получать новые знания [7]. Эта группа методов рассчитана на работу как частично обработанными, так и с «сырыми» данными.

Web Mining («добыча данных в Web») – набор методов и техник для получения данных из веб-источников [8]. Такие источники не являются текстовыми данными, соответственно и методы отличаются от Data Mining. Особенность в данном случае состоит в том, что в веб-ресурсах данные представлены в виде определенных форматов (Atom, SOAP, RSS, HTML), такие ресурсы располагают метаинформацией и информацией о структуре документа. Нужно также иметь ввиду наличие правил поисковой оптимизации (SEO).

Рассмотрим более подробно методы, относящиеся к группе Text Mining. Методы Text Mining направлены на превращение неструктурированных текстовых единиц в знания или же в подходящий для машинной обработки набор символов. Главными методами в Text Mining являются классификация и кластеризация, но далеко не всегда применяются именно они (зачастую они являются основой для различных надстроек и модификаций). Можно кратко выделить основные задачи, для решения которых используются методы Text Mining:

1. Извлечение понятий. Данная задача подразумевает повышение качества алгоритмов классификации, поиска и кластеризации с помощью получения новых «понятий» из текста. Например, необходимо создать базу данных из полезной информации, полученной в ходе анализа документов на естественном языке. Обычно в таких случаях осуществляется извлечение сущностей и связей, а также терминологии с последующим автореферированием [9].

2. Ответ на запросы (Question Answering). С помощью методов ответов на запросы мы можем дать возможность системе не только понять сам вопрос на естественном языке, но и предоставить пользователю понятный и вразумительный ответ. Примерный алгоритм состоит из поиска информации в частях документа, потенциально содержащих ответ, затем – фильтрация фраз, похожих на ответ, и, наконец, сам поиск правильного ответа.

3. Тематическое индексирование. Под «тематическим индексированием» изначально имелось ввиду определение для документов или запросов определенных индексов, которые могли бы отражать некие характерные особенности данных единиц. Но впоследствии это трансформировалось в задачу перевода текстовых единиц с естественного языка в формализованный вид, тогда как полученные описания представляют собой наборы ключевых слов и словосочетаний, отражающих тематическое содержание.

4. Поиск ключевых слов. В процессе поиска по ключевым словам ориентируются на результаты тематического индексирования для поиска документов, которые отвечают указанным требованиям (например, содержат пользовательские ключевые слова). Ключевыми словами в Text Mining можно назвать множество слов («терминов»), отражающих содержимое текста. Обычно метод используют для анализа частоты появления слов в тексте (основа для суммаризатора TF-IDF) [10].

5. Суммаризация (Text Summarization) [11]. С точки зрения человека реферирование и аннотирование текста – сложный вид интеллектуальной деятельности. Подобный вид деятельности требует большого количества временного ресурса. К тому же, в различных странах параллельно могут проходить исследования одного и того же материала, что связано с задержкой распространения информации в мире по разным причинам. Автоматизация подобных процессов – это выход из этой довольно сложной ситуации, который экономит время и ресурсы на обработку огромного количества текстов, многие из которых могут вообще не представлять никакой ценности для конкретного исследования.

Для эффективного извлечения скрытой информации из любого текста необходимо провести его предварительную обработку. В процессе такой обработки можно выделить три этапа [12]:

- ♦ очищение от знаков препинания – вариативный процесс, который зависит от поставленной задачи и используемого алгоритма;

- ♦ очищение от стоп-слов – комплексная задача. Состоит в удалении из текста лишних слов («шума»), которые негативно влияют на работу любого статистического метода. Как правило, это служебные части речи. В зависимости от используемого алгоритма может меняться и набор исключаемых частей речи;

- ♦ токенизация – самая сложная из всех этапов. Здесь, в первую очередь, необходимо разделить весь текст на составные части. Формат таких частей определяется алгоритмом. Как правило, текст разделяют на минимальные части – слова – и проводят дальнейшую обработку. В случае, когда необходимо получить список слов для каждого предложения текст разделяют на предложения с дальнейшим разложением на слова [13].

- ♦ стемминг – подразумевает приведение полученных слов к начальной форме. Более сложные системы включают алгоритмы, которые приводят к одной форме все синонимы.

Конечным результатом выполнения всех вышеперечисленных этапов является аннотация.

В зависимости от поставленной задачи методы автоматической обработки текста с целью получения какой-либо выдержки можно разделить на две группы – «extraction-based» и «abstraction-based» (извлечение информации и получение абстрактной информации). Более сложные структуры, к примеру - предложения, могут считаться ключевыми на основе совокупностей правил. Одна из причин считать предложение ключевым – наличие большого количества ключевых слов.

Для поиска ключевых единиц можно использовать несколько методов. Каждый из них по-своему эффективен в определенной ситуации и для определенных задач. Можно выделить три основных метода: статистический, позиционный, логико-семантический.

Статистический метод базируется на идее о том, что наиболее важные по смыслу слова встречаются в тексте чаще всего. Таким образом, предложение является ключевым, если в нем содержится определенное количество ключевых слов. Для частотного анализа используются различные статистические коэффициенты [14].

В позиционном методе ключевые единицы определяются в зависимости от расположения единицы в тексте. Как правило, это касается сложных лингвистических конструкций – предложения и абзацы. Например, предложение-заголовок вполне может являться ключевым элементом.

Последний метод – логико-семантический. Такой метод рассчитан на исследование структуры и семантики слов внутри текста. Задача состоит в выделении из текста предложения с наибольшим функциональным весом. Это зависит от множества различных факторов – наличия в предложении семантически значимых слов, взаимосвязей этого предложения с контекстом и т.д.

Итоговая генерация текста зависит от единиц, которые мы получим на выходе алгоритма: набор ключевых предложений или набор значимых слов (N-грамм). В любом случае главной задачей будет скрепление их в единую смысловую конструкцию. Смысловое разделение можно учесть непосредственно в ходе извлечения информации.

Реферирование и квазиреферирование. Можно выделить основные две задачи в процессе суммаризации: квазиреферирование (extractive summarization) и реферирование (abstractive summarization) [15].

1. Квазиреферирование. Цель квазиреферирования состоит в анализе входного текста (или выборки текстов) с последующим взвешиванием предложений или иных составных частей текста с целью определения наиболее значимых единиц. Получаемый в результате текст (summary) состоит из набора таких единиц. Методы квазиреферирования оперируют понятием «вес», который характеризует значимость лингвистической единицы для конкретного контекста [16]. Как правило, алгоритмы на основе такого подхода довольно просты для реализации и основаны на эвристиках. Основной недостаток данного метода состоит в возможной потере в результирующем тексте каких-либо семантических и синтаксических связей между выделенными единицами (например, предложениями).

Наиболее известным среди методов квазиреферирования является метод Latent Semantic Analysis (LSA). Он основан на построении матрицы с последующим ее сингулярным разложением. Поэтому каждое слово исходного текста подлежит взвешиванию. Способы определения веса предложения разнятся в зависимости от алгоритма. Получаемое в результате краткое содержание (аннотация) текста содержит наиболее значимые лингвистические единицы. Подобные графовые алгоритмы демонстрируют достаточно хорошую эффективность для решения конкретных задач.

К данной группе можно также отнести метод на основе определения веса TF-IDF [17]. TF-IDF – наиболее часто используемая мера определения значимости слова. Под мерой значимости мы понимаем вес единицы. TF-IDF – это произведение двух коэффициентов:

♦ TF (term frequency) – коэффициент, позволяющий измерить вес слова внутри одного текста (или частотность слова). Именно из-за этого показателя на этапе предварительной обработки текста необходимо избавиться от стоп-слов. Такие слова обычно обладают наибольшим показателем TF, но нулевой смысловой нагрузкой, что сказывается на эффективности взвешивания единиц с помощью этого коэффициента;

♦ IDF (inverted document frequency) – показатель, который позволяет оценить вес слова с учетом выборки текстов. Этот коэффициент хорошо себя проявляет при наличии задачи определения тематики текстов. Это инвертированный показатель частотности единицы внутри корпуса текстов. Количество совпадений в конкретном тексте роли не играет.

При перемножении данных коэффициентов можно получить общую характеристику слова (N-грамма). Полученный результирующий вес слова отражает смысловую важность слова внутри анализируемого текста с опорой на тексты данной тематики (или выборки) [18].

2. Реферирование. В основе данного подхода лежит определение основных смыслов текста и дальнейшая генерация краткого содержания на основе полученных знаний. В настоящее время в подобных системах используют различные виды нейронных сетей, для обучения которых используются наборы текстовой информации. Недостатком такого подхода является большой объем данных для обучения и необходимость их предварительной подготовки. Несмотря отмеченные недостатки, данный подход является наиболее перспективным из существующих. Как правило, для реализации метода реферирования используется архитектура encoder-decoder, а также transformer [19].

Методы квазиреферирования более просты в реализации по сравнению с методами реферирования. Они не нуждаются в обработанной и размеченной обучающей выборке. В большинстве случаев подобный подход предоставляет приемлемое качество полученного результирующего краткого содержания. Но для систем, работающих с большими массивами данных, требующих высокого качества аннотации и максимально исключают участие человека, используются именно методы реферирования.

Рекуррентные нейронные сети. Рекуррентные нейронные сети (RNNs) – это модели, широко используемые для обработки текстовых единиц на естественном языке, а наиболее эффективной архитектурой является – архитектура encoder-decoder LSTM (Long Short-term Memory Networks). Задачи, решаемые с помощью таких моделей – это частотный анализ и генерация нового текста [20]. Подобные решения используются для решения группы проблем последовательностей – seq2seq. По сути, проблема seq2seq заключается в определении следующего значения в последовательности или же выводе метки входной последовательности. Чаще всего это обозначается в виде отношений типа «один к одному» или «многие к одному».

Также могут существовать и задачи более сложного типа, например, определения последовательности входных данных и прогнозирования последовательности данных на выходе.

Цель использования рекуррентных нейронных сетей состоит в получении последовательности при получении или применении информации. Традиционно считается, что все входы и выходы в нейронных сетях независимы друг от друга, но, к сожалению, такая структура подходит для решения далеко не всех задач. Логично, например, что для предсказания последующего слова в последовательности необходимо учитывать и предыдущие слова. Рекуррентные нейронные сети именно поэтому и называются рекуррентными, что они выполняют определенной действие к элементу последовательности, учитывая всю предыдущую последовательность. Подобную идею можно интерпретировать и по-другому – такие RNN системы имеют «память». Теоретически можно предположить, что длина входных последовательностей может быть неограниченной, но на практике все обстоит несколько иначе (рис. 1), где

x_t – вход для временного шага t ;

s_t – представляет собой скрытое состояние для шага t и может интерпретироваться как «память». Это значение зависит от текущего входа и всех предыдущих состояний: $s_t = f(U x_t + W s_{t-1})$. В качестве функции f используется нелинейная \tanh или же $ReLU$. s_{-1} используется для вычисления первого скрытого состояния и при инициализации получает значение нуля (нулевого вектора).

o_t – представляет собой выход для шага t . К примеру, для предсказания слова в последовательности можно представить выход в качестве выхода вероятностей в словаре $o_t = \text{softmax}(Vs_t)$.

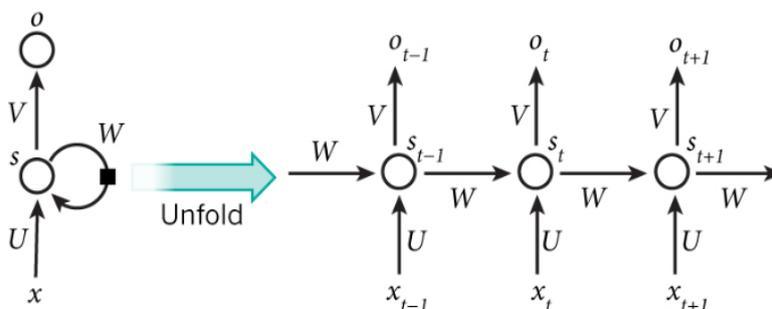


Рис. 1. Развертка рекуррентной нейронной сети

Таким образом, мы видим, что RNN можно развернуть в целую сеть. Допустим, что последовательность состоит из 10 слов. Тогда развертка будет включать 10 слоев.

Процесс обучения рекуррентной нейронной сети похож на обучение обычной нейронной сети. Обычно используется алгоритм обратного распространения ошибки (backpropagation), но с некоторыми оговорками. Исходя из того, что мы пользуемся одними и теми же параметрами на протяжении всех временных этапов сети, градиент зависит и от расчета текущего шага, и от предыдущих временных шагов. К примеру, если нам требуется вычислить градиент для $t = 3$, то мы «распространяем ошибку» на 2 шага, а затем суммируем полученные градиенты. Такая модификация backpropagation получила название «алгоритм обратного распространения ошибки сквозь время» или же «backpropagation through time» (BPTT). Недостаток такого обучения заключается в возможности анализировать только краткосрочные зависимости. Для обхода этих ограничений появились различные модификации, в том числе сети долгой краткосрочной памяти (Long Short-Term Memory, LSTM). От обычных RNN такие сети отличаются возможностью сохранять долгосрочные зависимости и способом вычисления скрытого состояния. Они рассчитаны на обучение долгосрочным зависимостям. Впервые LSTM описаны в работе Зеппа Хохрайтера и Юргена Шмихдхубера, опубликованной в 1997 году [21], после чего появилось множество модификаций данных сетей.

Если представить модуль обычной рекуррентной нейронной сети, то он будет выглядеть как один слой, содержащий функцию активации \tanh (гиперболический тангенс). Структура LSTM выглядит похожим образом за исключением одной детали. Внутри блока содержатся четыре слоя с определенным взаимодействием (рис. 2).

Красугольной идеей LSTM является наличие состояния ячейки (cell state). Информация попадает в блок, участвуя лишь в некоторых линейных преобразованиях. Так же LSTM может очищать состояние ячейки от информации, чему способствуют фильтры (gates).

В целом, LSTM – это огромный шаг в развитии нейронных сетей и рекуррентных нейронных сетей в частности. В настоящее время ведется работа над улучшением алгоритмов «attention», которые позволяют RNN брать данные из более крупного хранилища информации для более четкого определения контекста входящей последовательности.

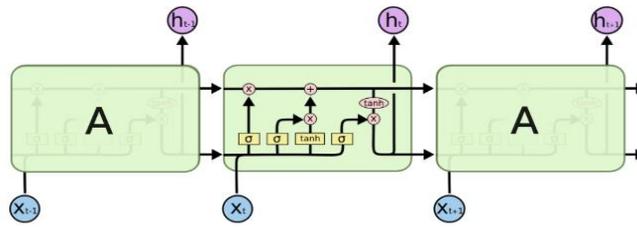


Рис. 2. Повторяющийся блок LSTM

Разработка алгоритма. Довольно часто отзывы клиентов могут быть громоздкими и описательными. Анализ таких отзывов вручную может отнимать значительное количество времени. В нашем случае целью является получение краткого оценочного реферата на основе входящего отзыва о продукте питания на английском языке. Для решения данной задачи была разработана система автоматического реферирования (суммаризатор), которая позволяет быстро определить основную идею отзыва и сформировать его краткое содержание. В качестве примера было решено использовать базу отзывов Amazon Fine Food с применением к нему модели seq2seq.

Использование модели «seq2seq» возможно для любой проблемы, которая связана с последовательной информацией. Нашей задачей является генерация реферата на основе отзывов пользователей. На вход будет поступать длинная последовательность слов, а на выходе мы должны получить краткое содержание. Соответственно, мы можем определить данную проблему seq2seq как «многие ко многим» [22].

При работе с нейронными сетями существует одна существенная проблема – невозможно сразу и заранее определить количество компонентов данной нейронной сети. В связи с этим были приняты примерные величины, используемые при обучении RNN. Например, было принято решение по общей структуре нейронной сети, а именно количество LSTM слоев равно 3. LSTM слои нам нужны для запоминания поступившей информации и взаимосвязи между ее элементами.

Для решения поставленных задач предложен следующий алгоритм обучения (рис. 3).

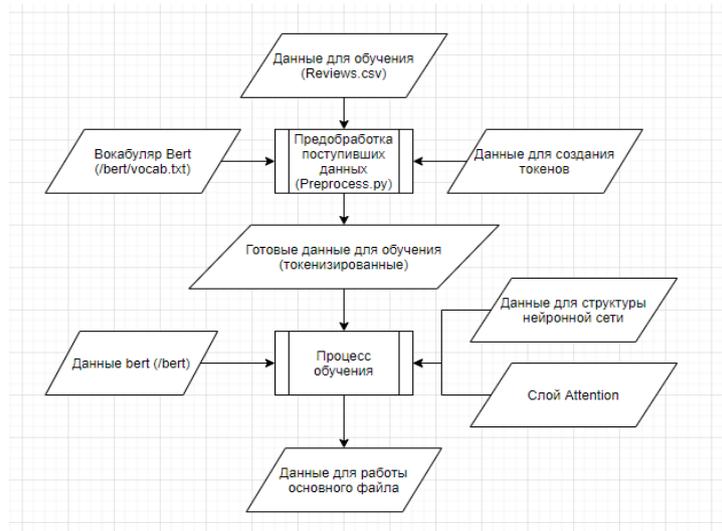


Рис. 3. Схема процесса обучения

Для выполнения предварительной обработки алгоритм использует следующие параметры:

- ◆ размеры обучающей и тестовой выборок - 0.9 и 0.9-1 соответственно;
- ◆ максимальная длина входной последовательности в токенах – 30.
- ◆ максимальная длина результирующей последовательности в токенах – 20.
- ◆ количество нейронов для слоев LSTM (влияет на качество и быстродействие) – 300.
- ◆ размерность эмбедингов для декодера – 100.
- ◆ максимальное количество итераций, после которого обучение остановится, если до этого не сработает условие окончания – 50.

После прохождения каждого цикла обучения производится оценка по следующим показателям: Loss; Validation Loss.

На основе разницы этих показателей может быть принято решение об остановке обучения и сохранении наиболее удачного результата.

Описание разработанной программной системы. Для реализации системы был выбран язык программирования Python в среде разработки PyCharm Community Edition 2019. Данная среда разработки наиболее эффективно использует возможности языка программирования Python, который, в свою очередь, хорошо справляется с задачами Data Mining, особенно Text Mining. Разработка проводилась на операционной системе Microsoft Windows.

В связи с ресурсоемкостью процесса обучения также была использована среда Google Colab, которая предоставляет необходимые серверные GPU мощности.

Объектно-ориентированный язык Python имеет в своем распоряжении огромное количество библиотек для работы с текстовыми и статистическими данными.

Программная система располагает удобным установщиком библиотек Python, а также удобными инструментами разработки на этом языке.

Программная оболочка для пользователя представляет собой исполняемый файл для запуска приложения. Результатом запуска является отображение главной страницы приложения (рис. 4). Текст (до 30 символов) вводится в окне ввода сверху, результирующее предложение выводится ниже.

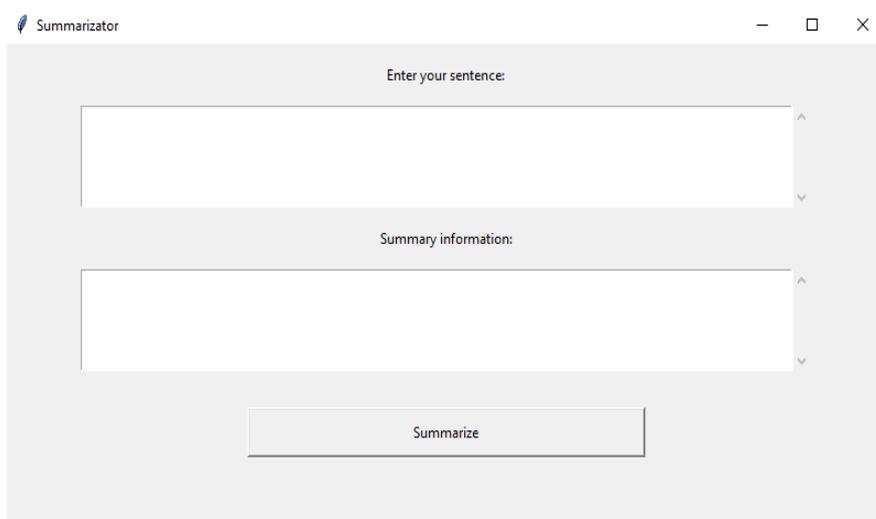


Рис. 4. Внешний вид оболочки

Результаты вычислительных экспериментов. Была проведена серия экспериментов для тестирования обучаемой модели. Изменялись различные параметры: количество последовательностей в выборке, количество нейронов для LSTM слоев, количество последовательностей, обрабатываемых одновременно и т.д. Количество единиц в выборке варьировалось от 10000 до 100000. Количество нейронов для LSTM слоев – от 50 до 200. Количество одновременно обрабатываемых последовательностей изменялось от 50 до 300.

В результате экспериментов выяснилось, что некоторые параметры невозможно повысить ввиду слабости машины для обработки. Было принято решение перенести обучение сети в Google Colab, где есть возможность обучать нейронную сеть на мощностях Google. Но для максимального качества работы алгоритма и этого оказалось недостаточно. На данный момент система обучена со следующими параметрами:

- ◆ размер выборки – 100000 отзывов;
- ◆ количество одновременно обрабатываемых отзывов – 200;
- ◆ количество нейронов для LSTM – 200;
- ◆ максимальное количество эпох – 50.

При данных параметрах был достигнут очень хороший результат, но стоит учесть, что обучение не прервалось автоматически ввиду большой разницы ошибок loss и val_loss. Результаты процесса обучения нейронной сети в Google Colab в виде диаграммы (рис. 5).

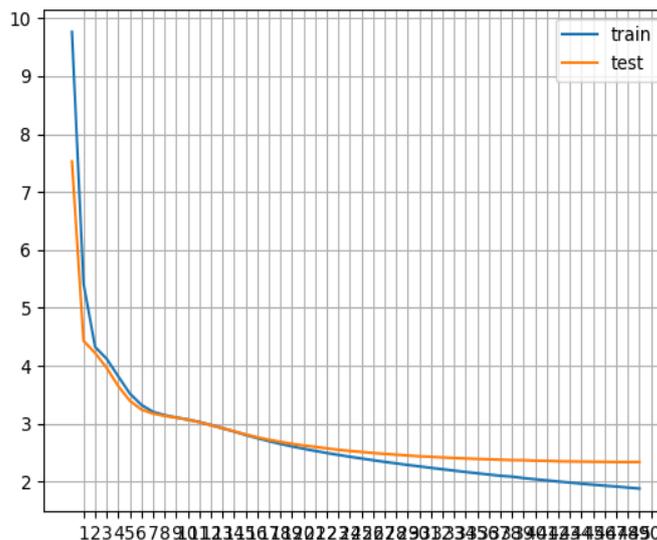


Рис. 5. Результаты обучения

На диаграмме можно увидеть зависимость loss и val_loss. При сильной разнице графиков происходит переобучение. Такого в нашем примере не наблюдается.

Оценить практическое качество работы алгоритма в данном конкретном случае довольно сложно. К примеру, очень популярная метрика ROUGE очень чувствительна к количеству слов, имеющих совпадение с ручным рефератом. Поэтому было принято решение использовать human evaluation.

Чтобы оценить результаты на практике было отобрано 10 предложений (отзывов) по адресу <https://www.foodnetwork.com/recipes/anne-thornton/fabulous-fudge-recipe2-1924263>.

Заключение. Задача разработки новых эффективных методов автоматического реферирования текстов становится все более востребованной по мере роста объемов обрабатываемой информации и необходимости скорости ее обработки. В зависимости от конкретной задачи применяются как алгоритмы для извлечения информации, так и для генерации совершенно новой. Все более активно для решения такого рода задач используются нейронные сети, но они требуют большого объема данных для обучения и точности калибровки. Так, например, компания Google в октябре 2019 года реализовала новый метод обработки естественного языка BERT (Bidirectional Encoder Representations from Transformers) на основе нейросетей новой архитектуры («трансформеры»).

В данной статье предложен подход к решению задачи автоматического реферирования отзывов о продуктах питания. Данный подход основан на использовании рекуррентной нейронной сети с определенной архитектурой и внедренным слоем BERT. Была разработана программная оболочка для демонстрации работы алгоритма.

Был проведен ряд экспериментов с параметрами сети и объемом обучающей выборки, что, в конечном итоге, позволило установить наиболее оптимальные настройки для эффективной работы алгоритма.

Определены возможности улучшения результатов работы системы за счет увеличения обучающей выборки, а также изменения метода поиска при построении результирующей последовательности.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Мордвинов А.В.* Разработка и исследование модели текста для его категоризации: автореф. дисс. ... канд. техн. наук: 05.13.01. – Н. Новгород, 2010. – 25 с.
2. *Треугода С.А.* Методы и алгоритмы автоматического реферирования текста на основе анализа функциональных отношений: автореф. дисс. ... канд. техн. наук: 05.13.01. – СПб., 2009. – 19 с.
3. *Лукашевич Н.В.* Модели и методы автоматической обработки неструктурированной информации на основе базы знаний онтологического типа: автореф. дисс. ... канд. техн. наук: 05.25.05. – М., 2014. – 32 с.
4. *Van Lierde H., Chow T.W.S.* Query-oriented text summarization based on hypergraph transversals // *Information Processing and Management*. – 2019. – Vol. 56, No. 4. – P. 1317-1338.
5. *Greengrass E.* *Information Retrieval: A Survey*: – University of Maryland. 2000. – 225 p.
6. *Manning D., Raghavan C., Schütze H.* *Introduction to Information Retrieval*: Cambridge. – England, 2008.
7. *Alguliev R.M., Isazade N.R., Abdi A., Idris N.* COSUM: Text summarization based on clustering and optimization // *Expert Systems*. – 2019. – Vol. 36, No. 1.
8. *Харламов А.* Технология автоматического смыслового анализа текстов TextAnalyst // *Вестник Московского государственного лингвистического университета*. – 2014. – С. 234-244.
9. *Хоай Л., Тузовский А.Ф.* Семантическое аннотирование документов в электронных библиотеках // *Известия Томского политехнического университета*. – 2013. – С. 157-164.
10. *Харламов А.* Когнитивный подход к смысловому анализу текстов // *Вестник Московского государственного лингвистического университета*. – 2013. – Т. 13, № 673. – С. 196-205.
11. *Gupta V., Bansal N., Sharma A.* Text summarization for big data: A comprehensive survey // *Lecture Notes in Networks and Systems*. – Delhi, 2019. – Vol. 56. – P. 503-516.
12. *Anam S.A., Muntasir Rahman A.M., Sleheen N.N., Arif H.* Automatic text summarization using fuzzy C-Means clustering // *2018 Joint 7th International Conference on Informatics, Electronics and Vision and 2nd International Conference on Imaging, Vision and Pattern Recognition*. – Kitakyushu, 2018. – P. 180-184.
13. *Chua S., Kulathuramaiyer N., Ranaivo-Malancon B., Iboi H.* A comparative Study of Sentiment-Based Graphs of Text Summaries // *2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences*. – Sarawak, 2018.

14. Siddiqui T. Generating abstractive summaries using sequence to sequence attention model // 2018 International Conference on Frontiers of Information Technology. Proceedings. – Karachi, 2018. – P. 212-217.
15. Sonawane S., Ghotkar A., Hinge S. Context-based multi-document summarization // Advances in Intelligent Systems and Computing. – 2018. – Vol. 812. – P. 153-165.
16. Alwis V. Intelligent E-news summarization // 18th International Conference on Advances in ICT for Emerging Regions. – Colombo, 2018. – P. 189-195.
17. Joshi A., Mehta K., Gupta N., Valloli V.K. Data generation using sequence-to-sequence // 2018 IEEE Recent Advances in Intelligent Computational Systems. – Pune, 2018. – P. 108-112.
18. Giglioli P., Sagar N., Rao A., Voyles J. Domain-Aware Abstractive Text Summarization for Medical Documents // Proceedings 2018 IEEE International Conference on Bioinformatics and Biomedicine. – New York, 2018. – P. 2338-2343.
19. Mahajani A., Pandya V., Maria I., Sharma D. Ranking-Based Sentence Retrieval for Text Summarization // 2018 2nd International Conference on Smart Innovations in Communications and Computational Sciences. – Mumbai, 2018. – P. 465-474.
20. Kirmani M., Manzoor Hakak N., Mohd M., Mohd M. Hybrid text summarization // 2nd International conference of the series Soft Computing: Theories and Applications. – 2017. – Kuruhsheetra, 2017. – P. 63-73.
21. Hochreiter S.; Schmidhuber J. Long short-term memory // Neural Computation: journal. – 1997. – Vol. 9, No. 8. – P. 1735-1780. – DOI: 10.1162/neco.1997.9.8.1735. – PMID 9377276.
22. Гладков Л.А., Гладкова Н.В., Бова В.В. Метод автоматического аннотирования текстов на основе гибридных интеллектуальных технологий // Информатизация и связь. – 2022. – № 2. – С. 54-60.

REFERENCES

1. Mordvinov A.V. Razrabotka i issledovanie modeli teksta dlya ego kategorizatsii: avtoref. dis. ... kand. tekhn. nauk [Development and research of a text model for its categorization: abstract of cand. of eng. sc. diss.]: 05.13.01. Nizhniy Novgorod, 2010, 25 p.
2. Trevgoda S.A. Metody i algoritmy avtomaticheskogo referirovaniya teksta na osnove analiza funktsional'nykh otноsheniy: avtoref. dis. ... kand. tekhn. nauk [Methods and algorithms for automatic text summarization based on the analysis of functional relationships: abstract of cand. of eng. sc. diss.]: 05.13.01. St. Petersburg, 2009, 19 p.
3. Lukashevich N.V. Modeli i metody avtomaticheskoy obrabotki nestrukturirovannoy informatsii na osnove bazy znaniy ontologicheskogo tipa: avtoref. dis. ... kand. tekhn. nauk [Models and methods for automatic processing of unstructured information based on an ontological knowledge base: abstract of cand. of eng. sc. diss.]: 05.25.05. Moscow, 2014, 32 p.
4. Van Lierde H., Chow T.W.S. Query-oriented text summarization based on hypergraph transversals, *Information Processing and Management*, 2019, Vol. 56, No. 4, pp. 1317-1338.
5. Greengrass E. Information Retrieval: A Survey: University of Maryland. 2000, 225 p.
6. Manning D., Raghavan C., Schütze H. Introduction to Information Retrieval: Cambridge. England. 2008.
7. Alguliev R.M., Isazade N.R., Abdi A., Idris N. COSUM: Text summarization based on clustering and optimization, *Expert Systems*, 2019, Vol. 36, No. 1.
8. Kharlamov A. Tekhnologiya avtomaticheskogo smyslovogo analiza tekstov TextAnalyst [Technology for automatic semantic analysis of texts TextAnalyst], *Vestnik Moskovskogo gosudarstvennogo lingvisticheskogo universiteta* [Bulletin of the Moscow State Linguistic University], 2014, pp. 234-244.
9. Khoay L., Tuzovskiy A.F. Semanticheskoe annotirovanie dokumentov v elektronnykh bibliotekakh [Semantic annotation of documents in electronic libraries], *Izvestiya Tomskogo politekhnicheskogo universiteta* [News of Tomsk Polytechnic University], 2013, pp. 157-164.
10. Kharlamov A. Kognitivnyy podkhod k smyslovomu analizu tekstov [Cognitive approach to semantic analysis of texts], *Vestnik Moskovskogo gosudarstvennogo lingvisticheskogo universiteta* [Bulletin of the Moscow State Linguistic University], 2013, Vol. 13, No. 673, pp. 196-205.
11. Gupta V., Bansal N., Sharma A. Text summarization for big data: A comprehensive survey, *Lecture Notes in Networks and Systems*. Delhi, 2019, Vol. 56, pp. 503-516.

12. Anam S.A., Muntasir Rahman A.M., Sleheen N.N., Arif H. Automatic text summarization using fuzzy C-Means clustering, *2018 Joint 7th International Conference on Informatics, Electronics and Vision and 2nd International Conference on Imaging, Vision and Pattern Recognition*. Kitakyushu, 2018, pp. 180-184.
13. Chua S., Kulathuramaiyer N., Ranaivo-Malancon B., Iboi H. A comparative Study of Sentiment-Based Graphs of Text Summaries, *2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences*. Sarawak, 2018.
14. Siddiqui T. Generating abstractive summaries using sequence to sequence attention model, *2018 International Conference on Frontiers of Information Technology. Proceedings*. Karachi, 2018, pp. 212-217.
15. Sonawane S., Ghotkar A., Hinge S. Context-based multi-document summarization, *Advances in Intelligent Systems and Computing*, 2018, Vol. 812, pp. 153-165.
16. Alwis V. Intelligent E-news summarization, *18th International Conference on Advances in ICT for Emerging Regions*. Colombo, 2018, pp. 189-195.
17. Joshi A., Mehta K., Gupta N., Valloli V.K. Data generation using sequence-to-sequence, *2018 IEEE Recent Advances in Intelligent Computational Systems*. Pune, 2018, pp. 108-112.
18. Giglioli P., Sagar N., Rao A., Voyles J. Domain-Aware Abstractive Text Summarization for Medical Documents, *Proceedings 2018 IEEE International Conference on Bioinformatics and Biomedicine*. New York. 2018, pp. 2338-2343.
19. Mahajani A., Pandya V., Maria I., Sharma D. Ranking-Based Sentence Retrieval for Text Summarization, *2018 2nd International Conference on Smart Innovations in Communications and Computational Sciences*. Mumbai, 2018, pp. 465-474.
20. Kirmani M., Manzoor Hakak N., Mohd M., Mohd M. Hybrid text summarization, *2nd International conference of the series Soft Computing: Theories and Applications*, 2017. Kuruhshehra, 2017, pp. 63-73.
21. Hochreiter S.; Schmidhuber J. Long short-term memory, *Neural Computation: journal*, 1997, Vol. 9, No. 8, pp. 1735-1780. DOI: 10.1162/neco.1997.9.8.1735. PMID 9377276.
22. Gladkov L.A., Gladkova N.V., Bova V.V. Metod avtomaticheskogo annotirovaniya tekstov na osnove gibridnykh intellektual'nykh tekhnologiy [Method for automatic annotation of texts based on hybrid intelligent technologies], *Informatizatsiya i svyaz'* [Informatization and communication], 2022, No. 2, pp. 54-60.

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

Гладков Леонид Анатольевич – Южный федеральный университет; e-mail: lagladkov@sfedu.ru; г. Таганрог, Россия; кафедра САПР; профессор.

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

Курейчик Виктор Михайлович – e-mail: vmkureychik@sfedu.ru; кафедра САПР; профессор.

Gladkov Leonid Anatol'evich – Southern Federal University; e-mail: lagladkov@sfedu.ru; Taganrog, Russia; the department of CAD; professor.

Gladkova Nadezhda Viktorovna – e-mail: nvgladkova@sfedu.ru; the department of CAD; senior teacher.

Kureichik Viktor Mikhaylovich – e-mail: vmkureychik@sfedu.ru; the department of CAD; professor.

Д.В. Тимошенко

ОСОБЕННОСТИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ЗАДАЧ ЭЛЕКТРОДИНАМИКИ ПРИЗЕМНОГО СЛОЯ

Исследуется проблема интегрирования системы уравнений электродинамики приземного слоя атмосферы для различных случаев электродного эффекта. Исходная система дифференциальных уравнений электродного эффекта приводится к так называемому уравнению полного тока, представляющему собой уравнение второго порядка параболического типа, рассматриваемое в двумерной области пространство-время. Уравнение полного тока позволяет связать совокупность основных факторов, влияющих на состояние электрического поля приземного слоя атмосферы: ток проводимости, турбулентный ток и ток, возникающий в результате конвективных процессов в атмосфере с так называемым полным током в приземном слое, отражающим изменение потенциала ионосферы. Описанный способ дает значительные преимущества в исследовании, поскольку в рамках одной модели позволяет осуществлять постановки различных задач электродинамики приземного слоя и проводить сравнительный анализ влияния на поведение электрического поля в приземном слое как отдельных факторов, так и их совокупностей. Целью работы является анализ математических аспектов рассматриваемых моделей с точки зрения подходов к интегрированию уравнения полного тока. В рамках исследования проводится сопоставление особенностей физических постановок и их влияние на структуру математических модели и их свойства с точки зрения математической физики. Рассмотренные различные физические постановки показывают, что даже в относительно простой математической модели, состоящей из одного уравнения, дополненного начальными-краевыми условиями, возникает достаточно широкий спектр как модельных формулировок, так и подходов к интегрированию рассмотренных моделей.

Математическое моделирование; электродинамика; приземный слой; атмосфера; электродный эффект; электрическое поле.

D.V. Timoshenko

FEATURES OF MATHEMATICAL MODELING IN THE GROUND LAYER ELECTRODYNAMICS

The paper examines the problem of integrating the system of electrodynamic equations of the atmosphere surface layer for various cases of the electrode effect. The original system of differential equations of the electrode effect is reduced to the so-called total current equation, which is a second-order equation of parabolic type, considered in a two-dimensional space-time region. The total current equation allows us to relate the set of main factors influencing the state of the electric field in the atmosphere surface layer: conduction current, turbulent current and current arising as a result of convective processes in the atmosphere with the so-called total current in the surface layer, reflecting changes in the potential of the ionosphere. The described method provides significant advantages in the study, since within the framework of one model it allows the formulation of various problems of the surface layer electrostatics and a comparative analysis of the influence of both individual factors and their combinations on the behavior of the electric field in the surface layer. The purpose of the work is to analyze the mathematical aspects of the models under consideration from the point of view of approaches to integrating the total current equation. The study compares the features of physical formulations and their influence on the structure of mathematical models and their properties from the point of view of mathematical physics. The various physical formulations considered show that even in a relatively simple mathematical model consisting of one equation supplemented with initial boundary conditions, a fairly wide range of both model formulations and approaches to integrating the considered models arises.

Mathematical modeling; electrostatics; surface layer; atmosphere; electrode effect; electric field.

Введение. Дифференциальные уравнения электродинамики атмосферы основываются на классических уравнениях Максвелла [1, 2], и в приближении электродного эффекта математическая модель горизонтально-однородного свободного от аэрозоля приземного слоя с учетом турбулентного и конвективного переносов [2–4] имеет вид:

$$\begin{cases} \frac{\partial n_{1,2}}{\partial t} \pm \frac{\partial}{\partial z} (b_{1,2} \cdot n_{1,2} \cdot E) - \frac{\partial}{\partial z} \left(D_T(z,t) \cdot \frac{\partial n_{1,2}}{\partial z} \right) + \frac{\partial}{\partial z} (v(z,t) \cdot n_{1,2}) = q(z,t) - \alpha n_1 n_2, & (1) \\ \frac{\partial E}{\partial z} = \frac{e}{\varepsilon_0} (n_1 - n_2) \end{cases}$$

Решениями уравнений в системе (1) являются функции распределения концентрации полярных аэроионов ($n_{1,2}$) и напряженность электрического поля (E) по времени (t) и высоте (z). Параметрами модели являются подвижности ($b_{1,2}$) полярных аэроионов, коэффициенты рекомбинации (α) и турбулентной диффузии (D_T), скорости конвективного переноса (v) и ионизации воздуха (q), а также элементарный заряд (e) и электрическая постоянная (ε_0).

Приведение системы уравнений электродного эффекта к уравнению полного тока. Положим в системе (1) $v = const$, $D_T = D_1 = const$ и исследуем модель методами теории подобия. Используя характерные значения и масштабы входящих в уравнения величин, введем безразмерные переменные:

$$t' = t/T, \quad z' = z/l_1, \quad n'_{1,2} = n_{1,2}/n_\infty, \quad E' = E/E_\infty, \quad n_\infty = \sqrt{q_\infty/\alpha}, \quad l_1 = D_1 \cdot \tau, \quad \tau = (q_\infty \alpha)^{-1/2}. \quad (2)$$

Индексом « ∞ » обозначим верхнюю границу электродного слоя. Тогда первые два уравнения системы (1) принимают вид:

$$\tau \frac{\partial n'_1}{\partial t'} + \tau b_1 E_\infty \frac{\partial (n'_1 E')}{\partial z'} - D_1 \tau \frac{\partial}{\partial z'} \left(z' \frac{\partial n'_1}{\partial z'} \right) + v \tau \frac{\partial n'_1}{\partial z'} = 1 - n'_1 n'_2, \quad (3)$$

$$\tau \frac{\partial n'_2}{\partial t'} - \tau b_2 E_\infty \frac{\partial (n'_2 E')}{\partial z'} - D_1 \tau \frac{\partial}{\partial z'} \left(z' \frac{\partial n'_2}{\partial z'} \right) + v \tau \frac{\partial n'_2}{\partial z'} = 1 - n'_1 n'_2. \quad (4)$$

Выделим из уравнений (3), (4) характерные пространственные масштабы: $L_{nE1} = b_1 E_\infty \tau$, $L_{nE2} = b_2 E_\infty \tau$, $L_D = D_1 \tau$, $L_v = v \tau$. Первые два определяют профили полярных аэроионов вследствие действия классического электродного эффекта [2,3], а последние соответствуют высоте конвективно-турбулентного электродного слоя [6]. Принимая значения параметров равными: $q = 7 \cdot 10^6 \text{ м}^{-3} \text{ с}^{-1}$, $E = -100 \text{ В/м}$, $D_1 = 0,1 \text{ м}^2/\text{с}$, $v = 0,1 \text{ м/с}$, получаем: $\tau \approx 300 \text{ с}$, $L_{nE1} = 3,6 \text{ м}$, $L_{nE2} = 4,2 \text{ м}$, $L_D = 30 \text{ м}$, $L_v = 30 \text{ м}$.

Далее запишем систему (1) в безразмерном виде:

$$\begin{cases} \frac{\tau}{T} \frac{\partial n'_{1,2}}{\partial t'} \pm \frac{b_{1,2} E_\infty}{D_1} \frac{\partial}{\partial z'} (n'_{1,2} \cdot E') - \frac{\partial}{\partial z'} \left(z' \cdot \frac{\partial n'_{1,2}}{\partial z'} \right) + \frac{v}{D_1} \frac{\partial n'_{1,2}}{\partial z'} = \frac{q}{q_\infty} - n'_1 \cdot n'_2, & (5) \\ \frac{\partial E'}{\partial z'} = \frac{e \cdot n_\infty l_1}{\varepsilon_0 E_\infty} (n'_1 - n'_2), \end{cases}$$

Системе (5) характеризуется следующие безразмерными параметрами (критериями подобия):

$$\xi_{1,2} = \frac{|b_{1,2}| \cdot E_{\infty} \cdot \tau}{l_1} = \frac{|b_{1,2}| E_{\infty}}{D_1}, \quad \chi = \frac{\nu}{D_1}, \quad \gamma = \frac{en_{\infty} l_1}{\varepsilon_0 E_{\infty}} = \frac{en_{\infty} D_1 \tau}{\varepsilon_0 E_{\infty}}. \quad (6)$$

Запишем систему (1) в безразмерной форме:

$$\begin{cases} \frac{\tau}{T} \frac{\partial n'_{1,2}}{\partial t'} \pm \xi_{1,2} \frac{\partial}{\partial z'} (n'_{1,2} \cdot E') - \frac{\partial}{\partial z'} \left(z' \cdot \frac{\partial n'_{1,2}}{\partial z'} \right) + \chi \frac{\partial n'_{1,2}}{\partial z'} = \frac{q}{q_{\infty}} - n_1 \cdot n_2, \\ \frac{\partial E'}{\partial z'} = \gamma (n'_1 - n'_2), \end{cases} \quad (7)$$

Физический смысл критериев подобия состоит в следующем [1, 2]: $\xi_{1,2} \geq 1$ характеризует собой случай классического электродного эффекта в приземном слое, $\xi_{1,2} < 1$ – турбулентный электродный эффект, при условии $\xi_{1,2} \ll 1$ уже имеет место, так называемое, приближение сильного турбулентного перемешивания, значение γ определяет влияние объёмного заряда на электрическое поле, параметр χ определяет соотношение конвективного и турбулентного переносов. Метеорологические процессы в атмосфере протекают в среднем $T \sim 1-3$ ч, т.е. ($\tau \ll T$). Последнее соотношение является основанием для дальнейшего рассмотрения стационарного варианта системы (1) и получения аналитических решений.

Из системы (1) следует уравнение для плотности полного электрического тока (j) в приземном слое атмосферы, который состоит из тока проводимости, турбулентного и конвективного токов [1, 3], [12–15]. В стационарном случае имеем:

$$\frac{\partial E}{\partial t} - \varepsilon_0 D_T(t, z) \frac{\partial^2 E}{\partial z^2} + \varepsilon_0 \nu(z) \frac{\partial E}{\partial z} + \lambda(z) E = j(t). \quad (8)$$

Здесь λ – электрическая проводимость воздуха.

Уравнение (8) представляет собой неоднородное дифференциальное уравнение второго порядка параболического типа с переменными коэффициентами. Данное уравнение применимо для описания широкого класса задач электродинамики приземного слоя атмосферы, поскольку позволяет учитывать одновременно воздействие локальных факторов на местности: проводимость атмосферы, температуру, метеорообстановку посредством соответствующих слагаемых в левой части, выражающих компоненты тока, и одновременно с этим модель отражает воздействие на электрическое поле приземного слоя глобальной вариации потенциала ионосферы, описываемого соответствующим законом изменения плотности полного тока в правой части уравнения (8). Решение уравнения (8) определяет зависимость напряженности электрического поля от времени и высоты в пределах приземного слоя [16–18].

Подходы к интегрированию уравнения полного тока для различных моделей электродного эффекта. Рассмотрим далее преобразования уравнения (8) для различных моделей электродного эффекта и соответствующие изменения подходов к его интегрированию. Можно выделить следующие основные модели электродного эффекта: турбулентный, конвективный, смешанный. Данные разновидности задач будут определять вид левой части дифференциального уравнения (8). Правая часть уравнения (плотность полного тока) также может иметь постоянный или переменный характер.

В дальнейшем в качестве общего для всех моделей электродного эффекта будем рассматривать одно из распространенных представлений глобального суточного изменения плотности тока [6, 15, 19, 20].

Таковым является представление плотности тока в виде гармонической функции, отражающей унитарную вариацию потенциала ионосферы [6]:

$$j(t) = j_0(1 - \alpha \sin(\omega t)), \quad \omega = 2\pi/T. \quad (9)$$

При значениях параметров $\alpha = 0,2$ и $T = 24$ ч функция суточного изменения электрического тока имеет вид волны с минимумом в 06^h UT и максимумом в 18^h UT и амплитудой 20% (отклонения от среднего значения), что достаточно хорошо согласуется с унитарной вариацией [3]. Поскольку зависимость плотности тока от координаты в рассматриваемых задачах не предполагается, не меньшая общности, будем считать, что суточные вариации описываются соотношением (9).

Модель турбулентного электродного слоя. Данная модель электродного эффекта предполагает учет в качестве основного фактора, влияющего на поведение напряженности поля в приземном слое, так называемый метеорологический фактор, определяющий скорость ветра в районе исследований. Данный фактор описывается коэффициентом турбулентной диффузии (D_T), который в общем случае может представлять собой функцию высоты и времени. В уравнении (8) величина (D_T) представляет собой коэффициент при старшей производной от напряженности поля по высоте и таким образом оказывает определяющее влияние на тип уравнения (8).

Рассмотрим следующие постановки задачи моделирования турбулентного электродного эффекта [16–18].

1. Коэффициент турбулентной диффузии – функция времени. В данной постановке предполагается, что величина (D_T) задается соотношением:

$$D_T(t) = D_0(2 - \cos \omega t), \quad (10)$$

повторяющим закон изменения плотности полного тока (9). Также предполагаем, что конвективные процессы отсутствуют, что соответствует условию $v(z) = 0$. Дополнительно будем считать, что электрическая проводимость воздуха также является постоянной: $\lambda = const$.

При описанных предположениях исследуемое уравнение (8) преобразуется к виду:

$$\frac{\partial E}{\partial t} - \varepsilon_0 D_0(2 - \cos \omega t) \frac{\partial^2 E}{\partial z^2} + \lambda E = j(t). \quad (11)$$

В качестве начальных и граничных условий для уравнения (11) принимаем следующие [3]:

$$E|_{t=0} = E_0 e^{-\frac{z}{L}}, \quad E|_{z=0} = E_0, \quad \frac{\partial E}{\partial z} \Big|_{z=z_\infty} = 0. \quad (12)$$

В результате получаем нестационарную начально-краевую задачу (11)–(12) для уравнения параболического типа (11) относительно неизвестной функции $E_1(t, z)$ с неоднородным начальным и граничными условиями (12). Данная задача представляет собой задачу с положительно определенным самосопряженным дифференциальным оператором относительно координаты в левой части уравнения (11), поэтому данная задача разрешима при помощи метода Фурье [6].

Решение задачи (11)–(12) подробно изложено в работе [8], здесь приведем только основные положения данного случая.

Поскольку граничные условия (12) являются неоднородными, необходимо выполнить следующую замену переменных: $E - E_0 = E_1(t, z)$. При указанной замене уравнение (11) сохраняет свою структуру, однако граничные условия становятся однородными:

Неизвестную функцию представим в виде ряда:

$$E_1 = \sum_{n=1}^{\infty} T_n(t) \Phi_n(z), \quad (13)$$

где функции $\{\Phi_n(z)\}_{n=1}^{\infty}$ образуют полную ортонормированную систему оператора Штурма-Лиувилля:

$$-D_0 \frac{d^2 \Phi}{dz^2} = \mu \Phi(z), \quad (14)$$

с граничными условиями:

$$\Phi(0) = \Phi(L) = 0. \quad (15)$$

Собственные числа и собственные функции задачи (14) имеют вид:

$$\mu_n = \left(\frac{\pi(2n-1)}{2L} \right)^2, \quad \Phi_n(z) = \sqrt{\frac{2}{L}} \sin \frac{\pi(2n-1)}{2L} z. \quad (16)$$

Таким образом, в рассматриваемом случае уравнение полного тока интегрируется по классической схеме метода Фурье.

2. Коэффициент турбулентной диффузии – функция координаты. Данный случай предполагает, что коэффициент турбулентной диффузии представляется в виде [3]:

$$D_T(z) = D_m z^m, \quad (17)$$

где $m = 0; 1; 4/3$.

При этом относительно выбранной закономерности поведения величины $D_T(z)$ можно рассматривать как нестационарную, так и стационарную задачу [19, 20].

В случае стационарной задачи уравнение полного тока преобразуется в обыкновенное дифференциальное уравнение относительно напряженности электрического поля приземного слоя. Такое уравнение решается как обыкновенное дифференциальное линейное неоднородное уравнение второго порядка с переменными коэффициентами. Для решения такого уравнения стандартным методом служит представление решения в виде степенного ряда. Соответствующее решение получено в [6].

Для нестационарной задачи уравнение (8) с учетом зависимости (17) преобразуется к виду [8].

$$\frac{\partial E}{\partial t} - \varepsilon_0 \frac{\partial}{\partial z} \left[D_T(z) \frac{\partial E}{\partial z} \right] + \lambda E = j(t). \quad (18)$$

Начальные и граничные условия для задачи (18) сохраняют форму (12).

Дифференциальный оператор в левой части уравнения (18) для коэффициента турбулентной диффузии, заданного зависимостью (17) не является положительно определенным [10]. В соответствии с теорией метода Фурье в данной ситуации допустимо существование так называемого «весового» множителя $g(z)$, при умножении на который обеих частей уравнения (8), соответствующий дифференциальный оператор вновь становится оператором Штурма-Лиувилля, а система собственных функций этого оператора является ортогональной с «весом» $g(z)$. Далее решение задачи (18) находится методом Фурье.

В случае, когда дифференциальный оператор в правой части уравнения (8) не удовлетворяет условиям положительной определенности и самосопряженности, и соответствующий весовой множитель отсутствует, это не позволяет построить полную ортонормированную систему собственных функций соответствующей краевой задачи, для интегрирования краевой задачи можно воспользоваться одним из приближенных методов. Одним из широко распространенных методов приближенного решения задачи (8) в этом случае является метод Галеркина [11]. Данный метод заключается в том, что выбирается конечная система базисных функций $\{u_i(x)\}_{i=1}^n$, составляющих часть некоторой полной системы, причем необходимо, чтобы функция $u_0(x)$ удовлетворяла неоднородным краевым условиям, в нашем случае: $u_0|_{z=0} = E_0$.

При этом функции $u_i(x)$ удовлетворяют однородному краевому условию:

$$\frac{\partial u_i}{\partial z} \Big|_{z=z_\infty} = 0.$$

В таком случае решение начально-краевой задачи можно искать в виде конечного соотношения:

$$E = u_0(z) + \sum_{i=1}^n C_i u_i(z). \quad (19)$$

Подстановка соотношения (18) в исходное дифференциальное уравнение (8) приводит к так называемой невязке:

$$R(z, C_1, C_2, \dots, C_n) = L[u_0] + \sum_{i=1}^n C_i L[u_i] - E(z). \quad (20)$$

В случае точного решения рассматриваемой краевой задачи величина невязки тождественно обращается в нуль ($R \equiv 0$); поэтому для получения приближенного решения, близкого к точному, необходимо подобрать коэффициенты C_i таким образом, чтобы невязка R была в каком-то смысле малой.

Для обеспечения этого условия в методе Галеркина формулируется требование, чтобы невязка была ортогональна к каждой из базисных функций $\{u_i(x)\}_{i=1}^n$, тогда при достаточно большом числе этих функций обеспечит малость невязки в силу следующей теоремы [11]:

Теорема. Пусть $\{u_i(x)\}_{i=1}^\infty$ – полная система функций с ненулевой нормой, ортогональных на отрезке $[a, b]$. Если непрерывная функция $f(x)$ ортогональна на отрезке $[a, b]$ ко всем функциям $\{u_i(x)\}_{i=1}^\infty$, то есть

$$E(z) = \frac{j}{L\lambda} e^{\lambda \varepsilon_0 V_0 (L-z)} \cdot z. \quad (27)$$

2. Далее рассмотрим общий случай зависимости скорости конвекции от высоты внутри электродного слоя: $V(z) = V_1 \frac{z}{l} \left(1 - \frac{z}{l}\right)$. В этом случае уравнение (26) принимает вид:

$$\varepsilon_0 V_1 \frac{z}{l} \left(1 - \frac{z}{l}\right) \frac{dE}{dz} + \lambda E = j_0. \quad (28)$$

Решение уравнения (28) имеет вид:

$$E(z) = E_0 \left(\frac{L-z_0}{z_0} \right)^{\frac{\lambda L}{\varepsilon_0 V_1}} \left(\frac{L-z}{z} \right)^{-\frac{\lambda L}{\varepsilon_0 V_1}} + \frac{j_0}{\lambda}. \quad (29)$$

3. Рассмотрим случай, когда коэффициент турбулентной диффузии ($D(z) = D_0 = const$) и скорость конвективного переноса ($V(z) = V_0 = const$) имеют соизмеримые значения.

Уравнение (24) примет вид:

$$-\varepsilon_0 D_0 \frac{d^2 E}{dz^2} + \varepsilon_0 V_0 \frac{dE}{dz} + \lambda E = j_0, \quad (30)$$

а граничные условия сохранятся в форме соотношений (2). Из уравнения (7) находим зависимость напряженности электрического поля от высоты:

$$E(z) = \frac{(E_0 - j_0/\lambda)}{e^{k_2 z_0} - e^{(-k_2)L + k_2 z_0}} e^{k_2 z} + j_0/\lambda, \quad k_2 = \frac{-\varepsilon_0 V_0 + \sqrt{(\varepsilon_0 V_0)^2 + 4\varepsilon_0 D_0 \lambda}}{-2\varepsilon_0 D_0}. \quad (31)$$

4. Рассмотрим случай, когда турбулентный перенос преобладает над конвективным, и при этом коэффициент турбулентной диффузии является функцией высоты: $D(z) = D_1 z$. В этом случае уравнение (24) принимает вид:

$$-\varepsilon_0 D_1 z \frac{d^2 E}{dz^2} + \lambda E = j_0, \quad (32)$$

а граничные условия сохраняют форму (25). Решение этого уравнения может быть найдено в виде степенного ряда относительно координаты, который следующим соотношением:

$$E(z) = \sum_{k=1}^{\infty} \frac{1 \cdot 3 \cdot 7 \cdots (3k-2)}{(3k!) \varepsilon_0 D_1} z^{3k} + \frac{j_0}{\lambda}. \quad (33)$$

5. Наконец, рассмотрим модель, описывающую одновременное влияние на поведение напряженности электрического поля турбулентных и конвективных процессов. В качестве законов, описывающих поведение соответствующих коэффициентов, примем следующие:

$$D(z) = D_1 z, V(z) = V_1 \frac{z}{l} \left(1 - \frac{z}{l}\right). \quad (34)$$

В этом случае уравнение (24) принимает вид:

$$-\varepsilon_0 D_1 z \frac{d^2 E}{dz^2} + \varepsilon_0 V_1 \frac{z}{l} \left(1 - \frac{z}{l}\right) \frac{dE}{dz} + \lambda E = j_0, \quad (35)$$

а граничные условия сохраняют форму (25).

Решение уравнения (35) также может быть представлено в виде степенного ряда:

$$E(z) = \sum_{k=1}^{\infty} \frac{(2k+3)(k-1)}{(k!)(k+2)! \varepsilon_0 (D_1 + V_1)} z^{3k+2} + \frac{j_0}{\lambda}. \quad (36)$$

Заключение. В работе рассмотрены особенности моделирования задач электродинамики приземного слоя на примере уравнения плотности полного электрического тока. Рассмотренные различные физические постановки показывают, что даже в относительно простой математической модели, состоящей из одного уравнения, дополненного начально-краевыми условиями, возникает достаточно широкий спектр как модельных формулировок, так и подходов к интегрированию рассмотренных моделей. Безусловно, наибольший интерес представляют случаи нестационарных задач второго порядка для параболических уравнений с дифференциальным оператором, не являющимся оператором Штурма-Лиувилля. Задачи данного вида не разрешимы с помощью классического метода Фурье и, вообще говоря, нуждаются в формировании самостоятельной теории построения системы собственных функций, что предполагается в качестве предмета для дальнейших исследований.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Морозов В.Н. Атмосферное электричество // Атмосфера. Справочник (справочные данные, модели). – Л.: Гидрометеоздат, 1991. – С. 394-408.
2. Куповых Г.В., Морозов В.Н., Шварц Я.М. Теория электродного эффекта в атмосфере. – Таганрог: Изд-во ТРТУ, 1998. – 123 с.
3. Куповых Г.В. Электродинамические процессы в приземном слое атмосферы. – Таганрог: Изд-во ТТИ ЮФУ, 2009. – 114 с.
4. Свидельский С.С., Литвинова В.С., Куповых Г.В., Клово А.Г. Формирование структуры атмосферного электродного слоя // Известия ЮФУ. Технические науки. – 2020. – № 5. – С. 130-141.
5. Шулейкин В.Н., Шукин Г.Г., Куповых Г.В. Развитие методов и средств прикладной геофизики – атмосферно-электрический мониторинг геологических неоднородностей и зон геодинамических процессов. – СПб.: РГГМУ, 2015. – 206 с.
6. Аджиев А.Х., Клово А.Г., Кудринская Т.В., Куповых Г.В., Тимошенко Д.В. Суточные вариации электрического поля в приземном слое атмосферы // Известия РАН. Физика атмосферы и океана. – 2021. – Т. 57, № 4. – С. 452-461.
7. Куповых Г.В., Белоусова О.В., Тимошенко Д.В., Клово А.Г., Кудринская Т.В. Электродинамическая модель турбулентно-конвективного приземного слоя: приближенные аналитические решения // Матер. IX Всероссийской научной конференции по атмосферному электричеству. – СПб.: ВКА им. А.Ф. Можайского, 2023. – С. 488-496.
8. Белоусова О.В., Тимошенко Д.В., Куповых Г.В. Математическое моделирование электрической структуры турбулентно-конвективного приземного слоя // Радиотехнические и телекоммуникационные системы. – 2023. – № 3 (51). – С. 12-19.

9. *Куповых Г.В., Тимошенко Д.В., Клово А.Г., Кудринская Т.В.* Влияние электродного эффекта на суточные вариации электрического поля атмосферы в приземном слое // Оптика атмосферы и океана. – 2023. – Т. 36, № 10. – С. 834-838.
10. *Тихонов А.Н., Самарский А.А.* Уравнения математической физики. – М.: Наука, 1972. – 736 с.
11. *Михлин С.Г.* Вариационные методы в математической физике. – М.: Гостехиздат, 1957. – 450 с.
12. *Williams E.R., Mareev E.A.* Recent progress on the global electrical circuit // Atmos. Res. – 2014. – Vol. 135–136. – P. 208-227.
13. *Liu C., Williams E.R., Zipser E.J., Burns G.* Diurnal variation of global thunderstorms and electrified shower clouds and their contribution to the global electrical circuit // J. Atmos. Sci. – 2010. – Vol. 67, No. 2. – P. 309-323.
14. *Mach D.M., Blakeslee R.J., Bateman M.G.* Global electric circuit implications of combined aircraft storm electric current measurements and satellite-based diurnal lightning statistics // J. Geophys. Res. – 2011. – Vol. 59, No. 1. – P. 183-204.
15. *Harrison R.G.* The Carnegie curve // Surveys in Geophysics. – 2013. – 34 (2). – P. 209-232.
16. *Kupovykh G.V., Timoshenko D.V., Klovo A.G., Kudrinskaya T.V.* Electrodynamical processes models in atmospheric surface layer // CATPID-2019. IOP Conf. Series: Materials Science and Engineering. – 2019. – Vol. 698 (4). – P. 44034.
17. *Kupovykh G., Klovo A., Timoshenko D.* The atmospheric electric field variations in the surface layer // Russian Open Conference on Radio Wave Propagation (RWP). 1-6 July 2019. – Publisher: IEEE, 2019. – P. 580-583.
18. *Kupovykh G.V., Timoshenko D.V., Kudrinskaya T.V., Klovo A.G.* Modeling of the atmospheric electric field local variations in the turbulent surface layer // Journal of Physics: IOP Conference Series. VIII All-Russian Conference on Atmospheric Electricity. – 2020. – Vol. 1604 (1). – P. 012003.
19. *Hoppel W.A.* Theory of the electrode effect // Journal Atmospheric and Terrestrial Physics. – 1967. – Vol. 29, No. 6. – P. 709-721.
20. *Evtushenko A.A., Mareev E.A.* On the generation of charge layers in MCS stratiform regions // Atmospheric Research. – 2009. – Vol. 91. – P. 272-280.
21. *Kudrinskaya T.V., Klovo A. G., Kupovykh G.V., Timoshenko D.V.* Reduction coefficient and electric field near plane electrode with geometric heterogeneity // Journal of Physics: IOP Conf. Series. VIII All-Russian Conference on Atmospheric Electricity. – 2020. – Vol. 1604 (1). – P. 012005.

REFERENCES

1. *Morozov V.N.* Atmosfernoe elektrichestvo [Atmospheric electricity], *Atmosfera. Spravochnik (spravochnye dannye, modeli)* [Atmosphere. Directory (reference data, models)]. Leningrad: Gidrometeoizdat, 1991, pp. 394-408.
2. *Kupovykh G.V., Morozov V.N., Shvarts Ya.M.* Teoriya elektrodnoogo efekta v atmosfere [Theory of the electrode effect in the atmosphere]. Taganrog: Izd-vo TRTU, 1998, 123 p.
3. *Kupovykh G.V.* Elektrodinamicheskie protsessy v prizemnom sloe atmosfery [Electrodynamical processes in the surface layer of the atmosphere]. Taganrog: Izd-vo TTI YuFU, 2009, 114 p.
4. *Svidel'skiy S.S., Litvinova V.S., Kupovykh G.V., Klovo A.G.* Formirovanie struktury atmosferного elektrodnoogo sloya [Formation of the structure of the atmospheric electrode layer], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2020, No. 5, pp. 130-141.
5. *Shuleykin V.N., Shchukin G.G., Kupovykh G.V.* Razvitie metodov i sredstv prikladnoy geofiziki – atmosferno-elektricheskiy monitoring geologicheskikh neodnorodnostey i zon geodinamicheskikh protsessov [Development of methods and means of applied geophysics - atmospheric-electric monitoring of geological heterogeneities and zones of geodynamic processes]. Saint Petersburg: RGGMU, 2015, 206 p.
6. *Adzhiev A.Kh., Klovo A.G., Kudrinskaya T.V., Kupovykh G.V., Timoshenko D.V.* Sutochnye variatsii elektricheskogo polya v prizemnom sloe atmosfery [Daily variations of the electric field in the surface layer of the atmosphere], *Izvestiya RAN. Fizika atmosfery i okeana* [Izvestia RAS. Physics of the atmosphere and ocean], 2021, Vol. 57, No. 4, pp. 452-461.

7. Kupovykh G.V., Belousova O.V., Timoshenko D.V., Klovo A.G., Kudrinskaya T.V. Elektrodynamiceskaya model' turbulentno-konvektivnogo prizemnogo sloya: priblizhennye analiticheskie resheniya [Electrodynamic model of the turbulent-convective surface layer: approximate analytical solutions], *Mater. IX Vserossiyskoy nauchnoy konferentsii po atmosferному elektrichestvu* [Materials of the IX All-Russian Scientific Conference on Atmospheric Electricity]. Saint Petersburg: VKA im. A.F. Mozhayskogo, 2023, pp. 488-496.
8. Belousova O.V., Timoshenko D.V., Kupovykh G.V. Matematicheskoe modelirovaniye elektricheskoy struktury turbulentno-konvektivnogo prizemnogo sloya [Mathematical modeling of the electrical structure of the turbulent-convective surface layer], *Radiotekhnicheskie i telekommunikatsionnye sistemy* [Radio engineering and telecommunication systems], 2023, No. 3 (51), pp. 12-19.
9. Kupovykh G.V., Timoshenko D.V., Klovo A.G., Kudrinskaya T.V. Vliyanie elektrodnoy efekta na sutochnye variatsii elektricheskogo polya atmosfery v prizemnom sloe [Influence of the electrode effect on daily variations of the atmospheric electric field in the surface layer], *Optika atmosfery i okeana* [Optics of the atmosphere and ocean], 2023, Vol. 36, No. 10, pp. 834-838.
10. Tikhonov A.N., Samarskiy A.A. Uravneniya matematicheskoy fiziki [Equations of mathematical physics]. Moscow: Nauka, 1972, 736 p.
11. Mikhlin S.G. Variatsionnye metody v matematicheskoy fizike [Variational methods in mathematical physics]. Moscow: Gostekhizdat, 1957, 450 p.
12. Williams E.R., Mareev E.A. Recent progress on the global electrical circuit, *Atmos. Res.*, 2014, Vol. 135-136, pp. 208-227.
13. Liu C., Williams E.R., Zipser E.J., Burns G. Diurnal variation of global thunderstorms and electrified shower clouds and their contribution to the global electrical circuit, *J. Atmos. Sci.*, 2010, Vol. 67, No. 2, pp. 309-323.
14. Mach D.M., Blakeslee R.J., Bateman M.G. Global electric circuit implications of combined aircraft storm electric current measurements and satellite-based diurnal lightning statistics, *J. Geophys. Res.*, 2011, Vol. 59, No. 1, pp. 183-204.
15. Harrison R.G. The Carnegie curve, *Surveys in Geophysics*, 2013, 34 (2), pp. 209-232.
16. Kupovykh G.V., Timoshenko D.V., Klovo A.G., Kudrinskaya T.V. Electrodynamic processes models in atmospheric surface layer, *CATPID-2019. IOP Conf. Series: Materials Science and Engineering*, 2019, Vol. 698 (4), pp. 44034.
17. Kupovykh G., Klovo A., Timoshenko D. The atmospheric electric field variations in the surface layer, *Russian Open Conference on Radio Wave Propagation (RWP). 1-6 July 2019*. Publisher: IEEE, 2019, pp. 580-583.
18. Kupovykh G.V., Timoshenko D.V., Kudrinskaya T.V., Klovo A.G. Modeling of the atmospheric electric field local variations in the turbulent surface layer, *Journal of Physics: IOP Conference Series. VIII All-Russian Conference on Atmospheric Electricity*, 2020, Vol. 1604 (1), pp. 012003.
19. Hoppel W.A. Theory of the electrode effect, *Journal Atmospheric and Terrestrial Physics*, 1967, Vol. 29, No. 6, pp. 709-721.
20. Evtushenko A.A., Mareev E.A. On the generation of charge layers in MCS stratiform regions, *Atmospheric Research*, 2009, Vol. 91, pp. 272-280.
21. Kudrinskaya T.V., Klovo A. G., Kupovykh G.V., Timoshenko D.V. Reduction coefficient and electric field near plane electrode with geometric heterogeneity, *Journal of Physics: IOP Conf. Series. VIII All-Russian Conference on Atmospheric Electricity*, 2020, Vol. 1604 (1), pp. 012005.

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

Тимошенко Дмитрий Владимирович – Южный федеральный университет; e-mail: dmitrytim@sfedu.ru; г. Таганрог, Россия; тел.: +78634371705; кафедра физико-математических основ инженерного образования; к.ф.-м.н.; доцент.

Timoshenko Dmitry Vladimirovich – Southern Federal University; e-mail: dmitrytim@sfedu.ru; Taganrog, Russia; phone: +78634371705; the department of physical and mathematical foundations of engineering education; cand. of phys. and math. sc.; associate professor.

Э.В. Мельник, И.Б. Сафроненкова, А.Ю. Таранов

**ОНТОЛОГИЧЕСКИЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ
ПЕРЕРАСПРЕДЕЛЕНИЯ ВЫЧИСЛИТЕЛЬНОЙ НАГРУЗКИ
В РАСПРЕДЕЛЕННОЙ СИСТЕМЕ МОНИТОРИНГА С МОБИЛЬНЫМИ
КОМПОНЕНТАМИ НА БАЗЕ РАСПРЕДЕЛЁННОГО РЕЕСТРА***

Рассмотрены проблемы, связанные с организацией вычислительного процесса в системах мониторинга с мобильными компонентами на базе распределенного реестра (РР), в том числе задача перераспределения вычислительной нагрузки. Требования к функционированию современных распределенных систем мониторинга включают согласованную работу узлов всей системы, принадлежащих различным слоям вычислительной среды, в т.ч. туманному и краевому слоям, которые отличаются высокой динамичностью. Совместное использование технологий РР и мобильных компонентов в составе распределенных систем мониторинга позволяет расширить круг задач, решаемых такими системами, в т.ч., по причине того, что снимает вопросы, связанные с синхронизацией географически распределённых копий данных. Однако при такой организации распределенной системы необходимо учитывать следующие особенности вычислительной среды: латентность, связанную с синхронизацией данных на узлах РР, изменение географического положения мобильных компонентов, ограниченность бортового энергоресурса и высокую динамичность туманного и краевого слоев. Проведенные ранее исследования показали, что в условиях высоко динамичных вычислительных сред эффективно применение метода сокращения поискового пространства на основе онтологического анализа. Для корректной работы данного метода необходимо разработать онтологическую модель, отражающую особенности рассматриваемой вычислительно-коммуникационной среды, включающей РР и мобильные компоненты. В статье представлена новая онтологическая модель функционирования распределенной системы мониторинга, учитывающая наличие мобильных компонентов и узлов РР. Разработаны продукционные правила размещения вычислительной нагрузки в туманных и краевых слоях и на их основе реализована программная модель, позволившая провести ряд вычислительных экспериментов. Результаты экспериментальных исследований продемонстрировали эффективность предложенного подхода и адекватность разработанной онтологической модели.

Распределенная система мониторинга; распределенный реестр; мобильный компонент; туманные вычисления; краевые вычисления; онтологическая модель; распределение вычислительной нагрузки.

E.V. Melnik, I.B. Safronenkova, A.Yu. Taranov

**ONTOLOGICAL APPROACH TO SOLVING THE WORKLOAD
RELOCATION PROBLEM IN A DISTRIBUTED MONITORING SYSTEM
WITH MOBILE COMPONENTS BASED ON A DISTRIBUTED LEDGER**

The paper considers the problems associated with the organization of the computing process in monitoring systems with mobile components based on a distributed ledger (DL), including the task of redistributing the computing load. The requirements for the functioning of modern distributed monitoring systems include the coordinated operation of nodes of the entire system belonging to various layers of the computing environment, including foggy and edge layers, which are highly dynamic. The joint use of DL technologies and mobile components as part of distributed monitoring systems makes it possible to expand the range of tasks solved by such systems, including due to the fact that it removes issues related to the synchronization of geographically distributed copies of data. However, with such an organization of a distributed system, it is necessary to take into

* Исследование выполнено в рамках реализации ГЗ ЮНЦ РАН 122020100270-3.

account the following features of the computing environment: latency associated with data synchronization at DL nodes, changes in the geographical location of mobile components, limited onboard energy resources and high dynamism of the fog and edge layers. Previous studies have shown that in highly dynamic computing environments, the use of the search space reduction method based on ontological analysis is effective. For the correct operation of this method, it is necessary to develop an ontological model reflecting the features of the considered computing and communication environment, including DL and mobile components. In this paper a new ontological model of the functioning of a distributed monitoring system has been developed, taking into account the presence of mobile components and DL nodes. Production rules for placing computational load in foggy and edge layers have been developed and a software model has been implemented based on them, which allowed a number of computational experiments to be carried out. The results of experimental studies have demonstrated the effectiveness of the proposed approach and the adequacy of the developed ontological model.

Distributed monitoring system; distributed ledger; mobile component; fog computing; edge computing; ontological model; workload relocation.

Введение. Концепции краевых и туманных вычислений подразумевает возможность смещения части вычислительной нагрузки ближе к месту генерации данных [1, 2]. Такое перераспределение позволяет более эффективно организовывать вычислительный процесс в географически распределенных системах [3, 4]. Распределенные системы мониторинга с мобильными компонентами, реализованные на базе концепции краевых и туманных вычислений, позволяют осуществлять сбор данных с географически распределенных территорий, чем значительно расширяют сферу своего применения. Однако актуализируется ряд вопросов, связанных с синхронизацией данных, осуществлением процедуры обращения к этим данным, а также ограниченностью энергоресурса мобильных компонентов. Внедрение в такие системы технологий распределенного реестра позволяет решить отмеченные проблемы, а именно: сократить время передачи данных и нагрузку на коммуникационную инфраструктуру за счет возможности работать с «ближайшей» копией РР. С другой стороны, процедура синхронизации данных на узлах РР вносит дополнительные временные издержки. Таким образом, комплексные временные издержки включают следующие компоненты:

- ◆ временные издержки, связанные с многократным решением задачи перераспределения вычислительной нагрузки в условиях динамичной среды и множества мобильных компонентов с ограниченным энергоресурсом;

- ◆ временные издержки, возникающие при синхронизации данных на узлах РР.

Поскольку временные издержки негативно сказываются на эффективности функционирования систем рассматриваемого класса, то необходимо принять меры, позволяющие их сократить.

В рамках данного исследования будем рассматривать методы сокращения временных издержек, связанных с решением задачи перераспределения вычислительной нагрузки в описанных условиях.

В настоящее время существуют различные подходы, методы и алгоритмы решения задачи назначения и перераспределения вычислительной нагрузки в туманной среде [5]. Возможность их эффективного использования применительно к данной предметной области будет рассмотрена в следующем разделе.

1. Анализ существующих методов решения задачи переноса вычислительной нагрузки. При организации вычислительного процесса одной из задач является задача назначения и перераспределения вычислительной нагрузки на вычислительные устройства. Анализ открытых источников показал, что в настоящее время существует достаточное количество методов/ подходов/ алгоритмов к ее решению [5]. Однако при решении данной задачи необходимо учитывать следующие особенности рассматриваемой предметной области, влияющие на выбор методов ее решения:

- ◆ большое число узлов-кандидатов для переноса вычислительной нагрузки;
- ◆ динамичность среды функционирования [6];
- ◆ ограниченный энергоресурс мобильных компонентов, входящих в состав системы;
- ◆ латентность синхронизации данных на узлах РР.

Перечисленные факторы накладывают существенные ограничения на известные методы (подходы, алгоритмы) решения задачи переназначения вычислительной нагрузки. Анализ применимости данных методов в имеющихся условиях приведен в табл. 1.

Таблица 1

Анализ применимости подходов, методов и алгоритмов для решения задачи перераспределения вычислительной нагрузки

Подход/Метод/ Алгоритм	Недостатки
Двухуровневый метод планирования ресурсов в туманных вычислениях [7]	Не ориентирован на сложные топологии
Механизм переноса вычислительной нагрузки для уменьшения задержек при выполнении задач в соответствии с характеристиками контейнеров [8]	Не учитывает влияния времени решения задачи переноса вычислительной нагрузки на эффективность решения пользовательской задачи
Алгоритм формирования расписания, основанный на эвристиках [9]	Высокая вычислительная сложность в случае большой размерности вычислительной среды
Четырехуровневая архитектура планирования с учетом задержек [10]	Не учитывает временные издержки, связанные с передачей данных
Алгоритм разгрузки для выполнения задач на свободных узлах [11]	Применим только в гомогенной среде
Подход к управлению ресурсами, основанный на динамическом планировании многопользовательских устройств с использованием методов классификации облачных серверов и разнородных устройств [12]	Применяется для задач малой размерности, ориентирован на взаимодействие с облачными ресурсами
Метод на основе локальных групп устройств (ЛГУ) [13]	Время получения решения может быть неприемлемо долгим за счет множественных информационных обменов
Модифицированный метод решения задач перераспределения рабочей нагрузки для систем мониторинга и прогнозирования на основе онтологии [14].	Не учитывает наличие мобильных компонентов и узлов распределенного реестра
Метод синтеза устройств реализации искусственных нейронных сетей, ориентированных на туманные вычисления. [15]	Ориентирован на ограниченное и заранее известное число устройств. Ориентирован на задачи, которые «удобно» разделить на подзадачи с заданными характеристиками (многослойные нейронные сети)

Из табл. 1 видно, что наиболее «подходящим» методом решения задачи назначения и перераспределения вычислительной нагрузки в имеющихся условиях является метод онтологического анализа, в основе которого лежит возможность усечения поискового пространства. Схема, отражающая суть данного метода, изображена на рис. 1.



Рис. 1. Схема метода онтологического анализа

Данный метод эффективно применялся с целью сокращения времени принятия решения в системах различного назначения, функционирующих в туманной среде. Так, в проекте РФФИ 20-04-60485, использование онтологического подхода позволило сократить время реконфигурации в автоматизированных системах раннего предупреждения социальных возмущений от 1,87 до 2,5 раз (в зависимости от конфигурации системы) в сравнении с аналогом.

Эксперименты, проведенные в ходе выполнения НИР «Методы и средства построения высоконадежных реконфигурируемых систем мониторинга и диагностики на базе технологий «цифровой экономики»», продемонстрировали сокращение временных издержек, возникающих в процессе организации вычислений в системах мониторинга и диагностики, реализованных на базе концепции туманных вычислений до 2,5 раз в сравнении с аналогом.

Таким образом, можно сделать вывод, что использование онтологического подхода с целью сокращения поискового пространства эффективно, о чем свидетельствуют проведенные ранее вычислительные эксперименты. Однако стоит отметить, что в использованных ранее для реализации данного метода онтологических моделях не были учтены особенности функционирования мобильных компонентов (в части энергопотребления) а также узлов РР (в части латентности синхронизации данных). Для эффективной реализации метода онтологического анализа применительно к рассматриваемой области, необходимо разработать новую онтологическую модель, которая лежит в основе метода.

2. Разработка онтологической модели распределенной системы мониторинга с мобильными компонентами на базе РР. Анализ предметной области. Для описания рассматриваемой предметной области необходимо провести ее всесторонний анализ, определив основные концепты, их свойства и связи между ними. Это потребует для последующего создания онтологической модели функционирования распределенной системы, которая является ядром онтологического подхода. Для этого подробно рассмотрим концепцию организации вычислительного процесса в туманной и краевом слоях в распределенной системе мониторинга с мобильными компонентами на базе РР (рис. 2).

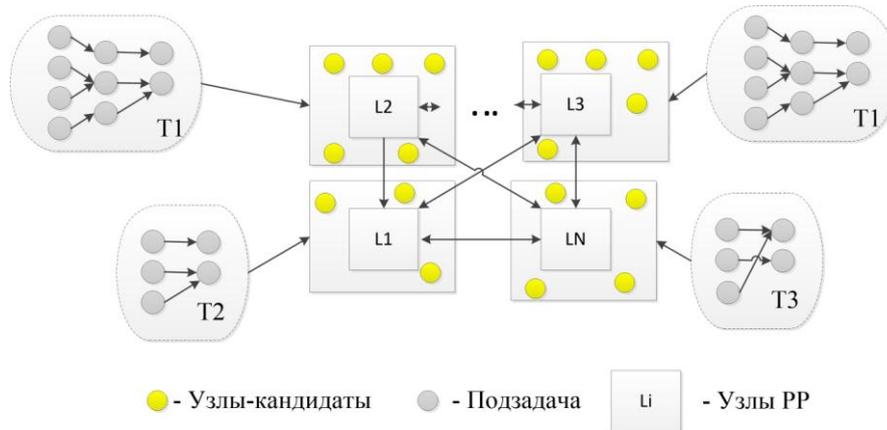


Рис. 2. Концепция организации вычислительного процесса в распределенной системе мониторинга с мобильными компонентами на базе РР

Распределенная система мониторинга функционирует в туманном слое и включает в себя следующие элементы:

- ◆ мобильные компоненты (БПЛА) – узлы краевого слоя;
- ◆ узлы РР;
- ◆ узлы туманного слоя (личные ПК, мобильные телефоны);
- ◆ узлы облачного слоя (ЦОД).

В процессе функционирования распределённая система мониторинга решает разнообразные задачи (T_i):

- ◆ сбор данных (преимущественно с использованием мобильных компонентов);
- ◆ передача данных;
- ◆ анализ данных (ресурсы мобильных компонентов – краевые устройства, узлы туманного слоя, узлы РР, облачные ресурсы);
- ◆ хранение данных;
- ◆ прогноз/ принятие решения (с использованием журнала записи данных, который реализуется за счет узлов РР).

Сформулируем основные проблемы, возникающие в процессе функционирования распределенной системы мониторинга с мобильными компонентами на базе РР:

- 1) ограниченный энергоресурс мобильных компонентов;
- 2) латентность при синхронизации данных в узлах РР.

Решение отмеченных проблем тесно связано с решением задачи назначения и перераспределения вычислительной нагрузки [16]. Ограниченность бортового ресурса приводит к необходимости организовывать вычислительный процесс таким образом, чтоб, имеющийся на борту энергоресурс, позволил выполнить назначенную на данное мобильное устройство функциональную задачу. Это возможно достичь за счет эффективного перераспределения задач на вычислительные узлы, которое позволяет высвободить ресурсы, необходимые для решения функциональных задач. Но, с другой стороны, занимает время, необходимое для решения задачи перераспределения вычислительной нагрузки, которая в имеющихся условиях является NP-сложной. Латентность, возникающая при реализации процедуры достижения консенсуса на узлах РР [17, 18], вносит дополнительные временные издержки, которые негативно сказываются на общем функционировании системы.

Таким образом, для решения задачи перераспределения вычислительной нагрузки необходимо использовать методы, позволяющие компенсировать временные издержки, вызванные перечисленными выше особенностями функционирования распределенных систем.

Разработка онтологической модели предметной области. На основании проведенного анализа предметной области, отражающей особенности функционирования распределенной системы мониторинга с мобильными компонентами на базе РР в части организации вычислительного процесса, получена следующая иерархия класс-подкласс выделенных концептов (рис. 3) [19, 20].

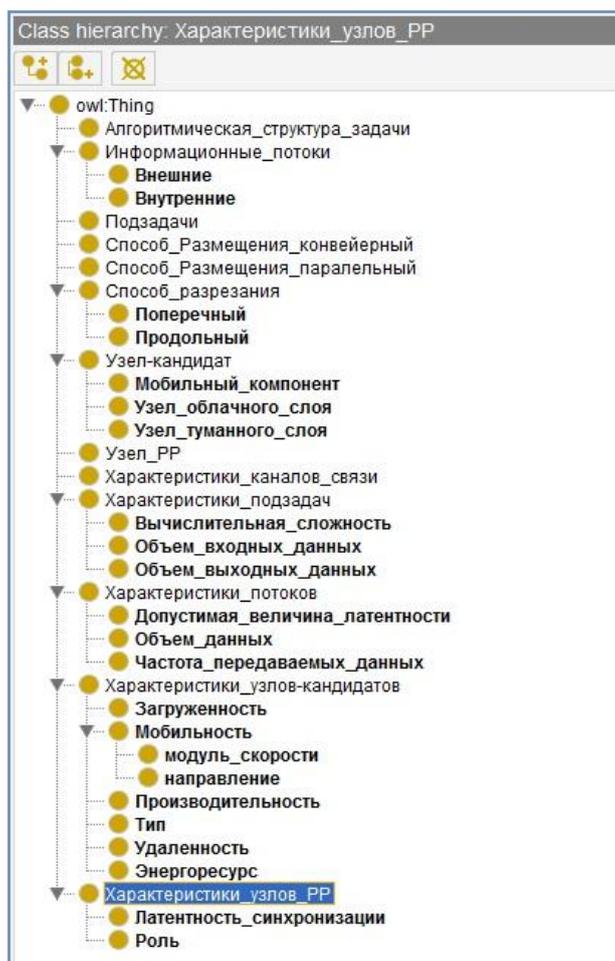


Рис. 3. Иерархическая структура концептов предметной области

На рис. 4 показано графовое отображение разработанной онтологической модели в неразвернутом виде, т.е. отображены основные концепты без «вложений». На рис. 5 показан общий вид онтологической модели.

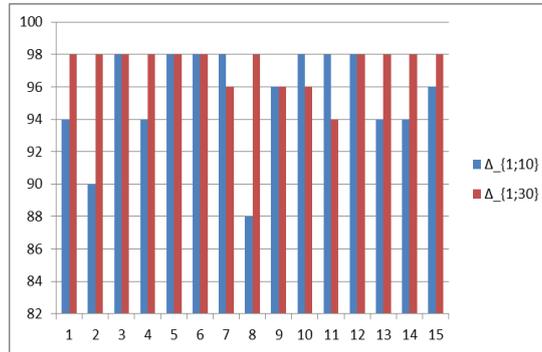


Рис. 6. Эффективность онтологического метода для $F_{нач} = 50$ для различной глубины туманного слоя

Эксперимент 2. Начальное количество узлов-кандидатов для размещения вычислительной нагрузки ($F_{нач}$) – 100, глубина туманного слоя – {1;10} хоп, глубина туманного слоя – {1;30} хоп (рис. 7).

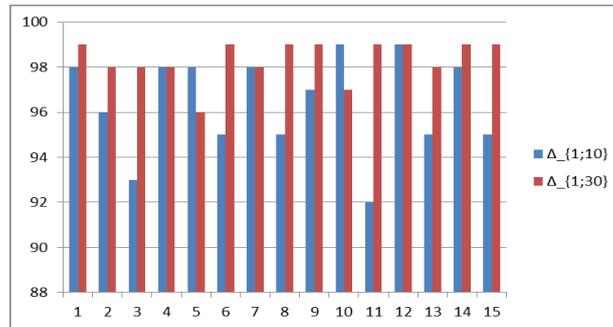


Рис. 7. Эффективность онтологического метода для $F_{нач} = 100$ для различной глубины туманного слоя

На основании проведенных экспериментов можно сделать вывод о том, что применение онтологического метода при решении задачи перераспределения вычислительной нагрузки позволяет сократить поисковое пространство в среднем на 93%. Это позволяет сократить время принятия решения при решении задачи перераспределения вычислительной нагрузки.

Заключение. В работе была рассмотрена проблема организации вычислительного процесса в распределённой системе мониторинга с мобильными компонентами на базе РР в части сокращения времени принятия решения. Был проведен анализ существующих методов (подходов, алгоритмов) к решению задачи перераспределения вычислительной нагрузки в туманной среде, на основе которого был сделан вывод о перспективности использования онтологического подхода. Однако имеющиеся разработанные онтологические модели и продукционные правила, применяемые к их классам, не учитывали наличие мобильных компонентов и узлов РР. В связи с этим, была разработана новая онтологическая модель, учитывающая особенности распределенной системы и среды ее функционирования. Также была разработана программная модель, позволяющая провести вычислительный эксперимент и подтверждающая эффективность применяемого метода для сокращения поискового пространства. Полученные результаты подтвердили адекватность разработанной онтологической модели.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Kaur K., Sachdeva M.* Fog Computing in IOT: An Overview of New Opportunities // Lecture Notes in Electrical Engineering. – 2020. – Vol. 605. – P. 113-119.
2. Cisco global cloud index: forecast and methodology (2014–2019). White paper.
3. *Клименко А.Б.* Метод ресурсосберегающего планирования распределенных вычислений в туманной вычислительной среде // Моделирование, оптимизация и информационные технологии. – 2022. – Т. 10 (3), № 3 (38). – С. 27-28.
4. *Naha R.K. et al.* Fog Computing: Survey of Trends, Architectures, Requirements, and Research Directions // IEEE Access. – 2018. – Vol. 6. – P. 47980-48009.
5. *Matrouk, K., Alatoun, K.* Scheduling Algorithms in Fog Computing: A Survey // International Journal of Networked and Distributed Computing. – 2021. – Vol. 9. – P. 59-74.
6. *Yi S., Qin Z., Li Q.* Security and Privacy Issues of Fog Computing: A Survey // Lecture Notes in Computer Science. – 2015. – Vol. 9204. – P. 685-695.
7. *Sun Y., Lin F., Xu, H.* Multi-objective Optimization of Resource Scheduling in Fog Computing Using an Improved NSGA-II // Wireless Personal Communications. – 2018. – Vol. 102. – P. 1369-1385.
8. *Yin L., Luo J., Luo H.* Tasks scheduling and resource allocation in fog computing based on containers for smart manufacturing // IEEE Trans. Ind. Inform. – 2018. – Vol. 14, No. 10. – P. 4712-4721.
9. *Wang J., Li D.* Task scheduling based on a hybrid heuristic algorithm for smart production line with fog computing // Sensors. – 2019. – Vol. 19, No. 5. – P. 1023.
10. *Wu C. -g., Li W., Wang L., Zomaya A.Y.* Hybrid Evolutionary Scheduling for Energy-Efficient Fog-Enhanced Internet of Things // IEEE Transactions on Cloud Computing. – 2021. – Vol. 9, No. 2. – P. 641-653.
11. *Li G., Yan J., Chen L., Wu J., Lin Q., Zhang Y.* Energy Consumption Optimization With a Delay Threshold in Cloud-Fog Cooperation Computing // IEEE Access. – 2019. – Vol. 7. – P. 159688-159697.
12. *Ghenai A., Kabouche Y., Dahmani W.* Multi-user dynamic scheduling-based resource management for Internet of Things applications // 2018 International Conference on Internet of Things, Embedded Systems and Communications (INTEC). – 2018. – P. 126-131.
13. *Мельник Э.В., Клименко А.Б., Иванов Д.Я.* Модель задачи формирования сообществ устройств информационно-управляющих систем в средах туманных вычислений // XIII Всероссийское совещание по проблемам управления ВСПУ-2019: Тр. – Тула: Тульский государственный университет. – 2018. – Вып. 2. – С. 1-6.
14. *Melnik E.V., Safronenkova I.B., Klimenko A.B.* A Modified Ontology-Based Method of Workload Relocation Problem Solving for Monitoring and Forecasting Systems // Lecture Notes in Networks and Systems. – 2021. – Vol. 229. – P. 145-152.
15. *Бахтин В.В., Тюрин С.Ф., Подлесных И.А.* Метод синтеза устройств нейросетевого распознавания на программируемой логике для реализации режима fog computing // Вестник Пермского национального исследовательского политехнического университета. Электротехника. Информационные технологии, системы управления. – 2022. – № 41. – С. 168-188.
16. *Brucker P.* Scheduling Algorithms. – Springer Berlin, Heidelberg, 2007. – 371 p.
17. *Fu X., Wang H., Shi P.* A survey of Blockchain consensus algorithms: mechanism, design and applications // Sci. China Inf. Sci. – 2021. – Vol. 64. – P. 121101.
18. *Niya S.R., Beckmann R., Stiller B.* DLIT: A Scalable Distributed Ledger for IoT Data // 2020 Second International Conference on Blockchain Computing and Applications (BCCA). – 2020. – P. 100-107.
19. *Водяхо А.И., Никифоров В.В.* Онтологические модели для систем реального времени // Онтология проектирования. – 2018. – Т. 8, № 2 (28). – С. 240-252.
20. *Ной Наталья Ф., Макгинесс Дебора Л.* Разработка онтологий 01: руководство по созданию вашей первой онтологии. – Режим доступа: http://www.labrate.ru/20181225/razrabotka_ontologiy_101_ruk.pdf (дата обращения: 03.10.2023).
21. Свид. о гос. регистрации прогр. для ЭВМ 2023665640 Российская Федерация. Модуль онтологического анализа для распределенной системы мониторинга с мобильными компонентами / Э.В. Мельник, И.Б. Сафроненкова, Таранов А.Ю. Заявл. 07.07.2023; зарегистр. 18.07.2023; опублик. 18.07.2023.

REFERENCES

1. Kaur K., Sachdeva M. Fog Computing in IOT: An Overview of New Opportunities, *Lecture Notes in Electrical Engineering*, 2020, Vol. 605, pp. 113-119.
2. Cisco global cloud index: forecast and methodology (2014–2019). White paper.
3. Klimenko A.B. Metod resursosberegayushchego planirovaniya raspredelennykh vychisleniy v tumannoy vychislitel'noy srede [A resource-saving method of distributed computation planning in fog-computing environment], *Modelirovanie, optimizatsiya i informatsionnye tekhnologii* [Modeling, optimization and information technology], 2022, Vol. 10 (3), No. 3 (38), pp. 27-28.
4. Naha R.K. et al. Fog Computing: Survey of Trends, Architectures, Requirements, and Research Directions, *IEEE Access*, 2018, Vol. 6, pp. 47980-48009.
5. Matrouk, K., Alatoun, K. Scheduling Algorithms in Fog Computing: A Survey, *International Journal of Networked and Distributed Computing*, 2021, Vol. 9, pp. 59-74.
6. Yi S., Qin Z., Li Q. Security and Privacy Issues of Fog Computing: A Survey, *Lecture Notes in Computer Science*, 2015, Vol. 9204, pp. 685-695.
7. Sun Y., Lin F., Xu, H. Multi-objective Optimization of Resource Scheduling in Fog Computing Using an Improved NSGA-II, *Wireless Personal Communications*, 2018, Vol. 102, pp. 1369-1385.
8. Yin L., Luo J., Luo H. Tasks scheduling and resource allocation in fog computing based on containers for smart manufacturing, *IEEE Trans. Ind. Inform.*, 2018, Vol. 14, No. 10, pp. 4712-4721.
9. Wang J., Li D. Task scheduling based on a hybrid heuristic algorithm for smart production line with fog computing, *Sensors*, 2019, Vol. 19, No. 5, pp. 1023.
10. Wu C. -g., Li W., Wang L., Zomaya A.Y. Hybrid Evolutionary Scheduling for Energy-Efficient Fog-Enhanced Internet of Things // *IEEE Transactions on Cloud Computing*. – 2021. – Vol. 9, No. 2. – P. 641-653.
11. Li G., Yan J., Chen L., Wu J., Lin Q., Zhang Y. Energy Consumption Optimization with a Delay Threshold in Cloud-Fog Cooperation Computing, *IEEE Access*, 2019, Vol. 7, pp. 159688-159697.
12. Ghenai A., Kabouche Y., Dahmani W. Multi-user dynamic scheduling-based resource management for Internet of Things applications, *2018 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC)*, 2018, pp. 126-131.
13. Mel'nik E.V., Klimenko A.B., Ivanov D.Ya. Model' zadachi formirovaniya soobshchestv ustroystv informatsionno-upravlyayushchikh sistem v sredakh tumannykh vychisleniy [Model of the problem of forming communities of information control systems devices in fog computing environments], *XIII Vserossiyskoe soveshchanie po problemam upravleniya VSPU-2019: Tr.* [All-Russian meeting on management problems of VSPU-2019: Proceedings]. Tula: Tul'skiy gosudarstvennyy universitet, 2018, Issue 2, pp. 1-6.
14. Melnik E.V., Safronenkova I.B., Klimenko A.B. A Modified Ontology-Based Method of Workload Relocation Problem Solving for Monitoring and Forecasting Systems, *Lecture Notes in Networks and Systems*, 2021, Vol. 229, pp. 145-152.
15. Bakhtin V.V., Tyurin S.F., Podlesnykh I.A. Metod sinteza ustroystv neyrosetevogo raspoznavaniya na programmiruемой logike dlya realizatsii rezhima fog computing [Method of synthesis of neural network recognition devices based on programmable logic for the implementation of fog computing mode], *Vestnik Permskogo natsional'nogo issledovatel'skogo politekhnicheskogo universiteta. Elektrotehnika. Informatsionnye tekhnologii, sistemy upravleniya* [PNRPU Bulletin. Electrotechnics, Informational Technologies, Control Systems], 2022, No. 41, pp. 168-188.
16. Brucker P. Scheduling Algorithms. Springer Berlin, Heidelberg, 2007, 371 p.
17. Fu X., Wang H., Shi P. A survey of Blockchain consensus algorithms: mechanism, design and applications, *Sci. China Inf. Sci.*, 2021, Vol. 64, pp. 121101.
18. Niya S.R., Beckmann R., Stiller B. DLIT: A Scalable Distributed Ledger for IoT Data, *2020 Second International Conference on Blockchain Computing and Applications (BCCA)*, 2020, pp. 100-107.
19. Vodyakho A.I., Nikiforov V.V. Ontologicheskie modeli dlya sistem real'nogo vremeni [Ontology models for real time systems], *Ontologiya proektirovaniya* [Design Ontology], 2018, Vol. 8, No. 2 (28), pp. 240-252.

20. *Noy Natal'ya F., Makginess Debora L.* Razrabotka ontologiy 01: rukovodstvo po sozdaniyu vashey pervoy ontologii [Ontology Development 101: A Guide to Creating Your First Ontology]. Available at: http://www.labrate.ru/20181225/razrabotka_ontologiy_101_ruk.pdf (accessed 03 October 2023).
21. *Mel'nik E.V., Safronenkova I.B., Taranov A.Yu.* Svid. o gos. registratsii progr. dlya EVM 2023665640 Rossiyskaya Federatsiya. Modul' ontologicheskogo analiza dlya raspredelennoy sistemy monitoringa s mobil'nymi komponentami [Certificate of state registration of a computer program 2023665640 Russian Federation. Ontological analysis module for a distributed monitoring system with mobile components]. Declared. 07.07.2023; registered. 18.07.2023; published. 18.07.2023.

Статью рекомендовал к опубликованию д.т.н. Ю.А. Кравченко.

Мельник Эдуард Всеволодович – Федеральный исследовательский центр Южный научный центр Российской академии наук; e-mail: evm17@mail.ru; г. Таганрог, Россия; тел.: 88634615459; г.н.с.

Сафроненкова Ирина Борисовна – e-mail: safronenkova050788@yandex.ru; тел.: +79604678753; м.н.с. лаб. ИТиПУ.

Таранов Антон Юрьевич – e-mail: mailfortexas@gmail.com; тел.: 88634615459; инженер лаб. ИТиПУ.

Melnik Eduard Vsevolodovich – Federal Research Center, The Southern Scientific Center of the Russian Academy of Sciences; e-mail: evm17@mail.ru; Taganrog, Russia; phone: +78634615459; chief researcher.

Safronenkova Irina Borisovna – e-mail: safronenkova050788@yandex.ru; phone: +79604678753; junior research fellow.

Taranov Anton Yurevich – e-mail: mailfortexas@gmail.com; phone: +78634615459; engineer.

Раздел III. Электроника, приборостроение и радиотехника

УДК 621.341.572

DOI 10.18522/2311-3103-2023-5-174-182

В.В. Бурлака, С.В. Гулаков, А.Ю. Головин, Д.С. Мироненко

СИСТЕМА БЕСПРОВОДНОЙ ПЕРЕДАЧИ ЭЛЕКТРИЧЕСКОЙ ЭНЕРГИИ

Широкое распространение получило оборудование, питающееся от встроенных аккумуляторов: беспилотные летательные аппараты, портативные радиостанции, тактические фонари, электрические транспортные средства и др. Зарядка аккумуляторов осуществляется зачастую контактным способом – подключением источника энергии посредством разъемного соединения. Это требует наличия технического персонала для производства коммутаций; организации защиты разъемного соединения от воздействия окружающей среды (влага, грязь и т.д.), а также защиты от поражения персонала электрическим током. Целью исследования является разработка технических средств беспроводной передачи электрической энергии, которые позволят исключить использование разъемных соединений, улучшить электробезопасность, а, главное – дадут возможность сделать процесс зарядки автоматическим. Результаты работы актуальны для создания автоматических систем доставки грузов с помощью беспилотных аппаратов; для реализации систем автоматической зарядки городского электротранспорта; для реализации автоматической зарядки беспилотных наземных, плавучих (в т.ч. подводных) и летательных аппаратов (разведывательных, патрульных и др.). Описана разработанная конструкция системы беспроводной передачи энергии мощностью до 250 Вт, пригодной для зарядки шестиземельных литиевых аккумуляторных батарей. Система работает с катушками диаметром 200 мм, полная работоспособность сохраняется при расстоянии между центрами катушек до 100 мм. КПД во всем диапазоне режимов не ниже 74 %, при измерении от сети 220 В до выхода на аккумуляторную батарею. При проектировании ставилась цель максимально снизить массу приемной части для облегчения ее установки на БПЛА и минимизации влияния на их тяговооруженность.

Беспроводная передача электрической энергии; резонансный трансформатор; синхронный выпрямитель; воздушный трансформатор; квазирезонансный инвертор напряжения.

V.V. Burlaka, S.V. Gulakov, A.Y. Golovin, D.S. Mironenko

THE WIRELESS ELECTRIC POWER TRANSFER SYSTEM

The equipment powered by built-in batteries has become widespread: unmanned aerial vehicles, portable radios, tactical flashlights, electric vehicles, etc. Charging of batteries is often carried out in a contact way – by connecting a power source by means of a detachable connection. This requires the presence of technical personnel for maintenance and replacement of batteries; requires the organization of protection of the battery connection from environmental influences (moisture, dirt, etc.), as well as protection against electric shock to personnel. The purpose of the research is to develop technical means of wireless transmission of electrical energy, which will eliminate the use of detachable connections, improve electrical safety, and, most importantly, will make it possible to make the charging process automatic. The results of the work are relevant for the implementation of automatic cargo delivery systems using unmanned vehicles; for the implementation of automatic charging systems for urban electric vehicles; for the implementation of

automatic charging of unmanned land, floating (including underwater) and aircraft (reconnaissance, patrol, etc.). The design of a wireless power transmission system with a power of up to 250 W, suitable for charging 6-cell lithium batteries, is described. The system works with coils with a diameter of 200 mm, full operability is maintained for a distance between the coils centers up to 100 mm. The efficiency in the entire range of operation modes is not lower than 74%, when measured from the 220 V mains to the output to the battery. When designing, the goal was to minimize the weight of the receiving part to facilitate its installation on the UAV and minimize the impact on its thrust-to-weight ratio.

Wireless energy transfer; resonant transformer; synchronous rectifier; air transformer; quasi-resonant voltage source inverter.

Введение. Широкое распространение мобильных электрических средств передвижения на земле, воде, в воздухе, различных гаджетов и другого оборудования, питающихся от аккумуляторных батарей (АКБ), ставят задачу организации надежной их зарядки в различных условиях (в т.ч. экстремальных), обеспечивающих автоматизацию процесса, малую зависимость от природных факторов (осадки, сырость, температурные параметры и др.), защиту от поражения электрическим током, скрытность зарядного устройства в случае необходимости. В наибольшей степени этим требованиям соответствует процесс беспроводной зарядки АКБ.

В настоящее время наиболее эффективным способом беспроводной передачи электрической энергии на небольшие расстояния является использование магнитно-связанных катушек [1–9]. Этот способ использован в стандарте беспроводных зарядок Qi [10–12] для мобильных устройств; прорабатываются варианты реализации беспроводной зарядки батарей электромобилей [13–15].

При создании таких систем основными критериями являются высокий КПД, небольшая себестоимость и массогабаритные показатели (особенно приемной части), простота пользования, соответствие санитарным нормам (индукция поля рассеяния), устойчивая работа в большом диапазоне изменения эквивалентных параметров воздушного трансформатора (расстояние между приемной и передающей катушками, смещение между их осями).

Основными путями повышения КПД являются использование высокодобротных катушек, намотанных литцендратом, установка синхронных выпрямителей на приемной стороне, а также оптимальное управление (выбор рабочей частоты и согласование сопротивления нагрузки). Здесь следует сразу отметить, что согласование сопротивления нагрузки требует применения дополнительного DC-DC преобразователя на приемной стороне [9, 16–19], что усложняет схему, ухудшает массогабаритные параметры приемной части и повышает ее себестоимость.

Соблюдение санитарных норм по индукции полей рассеяния достигается за счет проработки параметров воздушного трансформатора (рабочая индукция, конструкция магнитопровода) и применения экранирования. Причем экраны могут быть как из ферромагнитных материалов, так и комбинированные (ферромагнетик + проводник).

Задача обеспечения эффективной работы в большом диапазоне изменения эквивалентных параметров воздушного трансформатора вынуждает идти на компромиссы между эффективностью работы и удобством пользования. Для минимизации потерь энергии следует стремиться к максимальному коэффициенту связи передающей и приемной катушек.

Можно показать, что теоретический максимум КПД воздушного трансформатора является функцией произведения (kQ), где k – коэффициент связи, Q – добротность. Но требование обеспечения высокого коэффициента связи снижает удобство пользования системой, так как потребует точного и близкого распо-

ложения катушек. Так, зарядные устройства стандарта Qi работают при коэффициенте связи (0,3 ... 0,6). Беспроводные зарядные устройства для электромобилей рассчитываются на работу при меньшем коэффициенте связи, (0,15 ... 0,35) [13, 17].

При проектировании системы беспроводной передачи энергии, работающей в большом диапазоне изменения коэффициента связи и в большом диапазоне передаваемых мощностей (т.е. при изменении эквивалентного сопротивления нагрузки в широких пределах), усложняется задача синтеза регулятора, управляющего работой передающей стороны.

Подавляющее большинство систем строится с использованием принципа последовательной компенсации (series-series compensation) как в передатчике, так и в приемнике [3]. В результате имеется два магнитно-связанных резонансных контура. Но в такой системе, как известно, при коэффициенте связи, превышающем критический, происходит разделение полюсов и результирующая АЧХ становится «двугорбой», с двумя максимумами [16, 20]. Эта не монотонность АЧХ значительно усложняет синтез системы управления, разработчики предпочитают выбирать параметры воздушного трансформатора так, чтобы коэффициент связи во всем диапазоне режимов был меньше критического, и АЧХ имела один максимум.

Даже, несмотря на то, что такой компромисс (уменьшение коэффициента связи, а, следовательно, необходимость увеличения напряжения на передающей катушке и соответствующее увеличение тока намагничивания) ведет к увеличению потерь в системе.

Цель исследования состоит в разработке беспроводной системы передачи энергии, удовлетворяющей следующим критериям:

- ◆ простая приемная часть, желательно имеющая малую массу;
- ◆ отсутствие необходимости хорошего согласования резонансных частот передающего и приемного контуров;
- ◆ работа при минимальном коэффициенте связи не более 0,15 и при максимальном коэффициенте связи не менее 0,6;
- ◆ выходная мощность экспериментального образца не менее 200 Вт при выходном напряжении 22,2 ... 25,2 В (6S литиевая батарея);
- ◆ возможность масштабирования для увеличения мощности.

Вышеприведенный набор требований определил следующую структуру системы:

- ◆ приемник с последовательной компенсацией и мостовым синхронным выпрямителем, без преобразователя согласования сопротивления;
- ◆ приемная и передающая катушки без ферромагнитных сердечников, выполнены литцендратом, параметры контуров выбраны из условия минимизации числа витков;
- ◆ передатчик с отдельным управлением амплитудой и частотой выходного напряжения с выходным контуром, имеющим резонансную частоту значительно ниже, чем у приемного контура.

Передатчик системы беспроводной передачи энергии. Разнесение резонансных частот контуров снимает проблему разделения полюсов (АЧХ в рабочем диапазоне частот имеет один максимум), но приводит к тому, что преобразователь передатчика будет нагружен реактивным (индуктивным) током. Исходя из этого, для снижения потерь переключения применен квазирезонансный выходной каскад.

Кроме того, работа передающего контура «далеко» от резонансной частоты затрудняет использование принципа обратного рассеяния (back scattering) для организации передачи информации от приемника в передатчик. Поэтому в приемнике предусмотрена возможность отключения нагрузки при достижении максимального напряжения на выходе. Эта мера также повышает надежность работы прием-

ника (так, устройства стандарта QI выходят из строя при аварийном повышении напряжения на выходе – отсутствие элементов защиты объяснено в стандарте необходимостью уменьшения габаритов приемника [10]).

Принципиальная электрическая схема силовой части передатчика приведена на рис. 1. Входной выпрямитель передатчика состоит из диодного моста VD1 со сглаживающим конденсатором C1. Термистор NTC1 служит для ограничения тока заряда C1 при включении в сеть. Выпрямленное и сглаженное напряжение поступает на понижающий преобразователь, выполненный на элементах VT1, VT3, L1, C2. Он работает с широтно-импульсным управлением и обеспечивает регулируемое напряжение питания выходного полумоста VT2VT4. В качестве нижнего ключа (VT3) использован IGBT с антипараллельным диодом, хотя здесь можно оставить только диод. Выбранное решение преследует две цели: во-первых, позволяет обеспечить управление верхним транзистором (VT1) по схеме плавающего питания (bootstrap) – наличие IGBT гарантирует заряд «плавающего» конденсатора драйвера VT1; во-вторых, упрощает синтез регулятора напряжения на конденсаторе C2 за счет линеаризации передаточной характеристики преобразователя путем исключения бестоковых интервалов в токе L1.

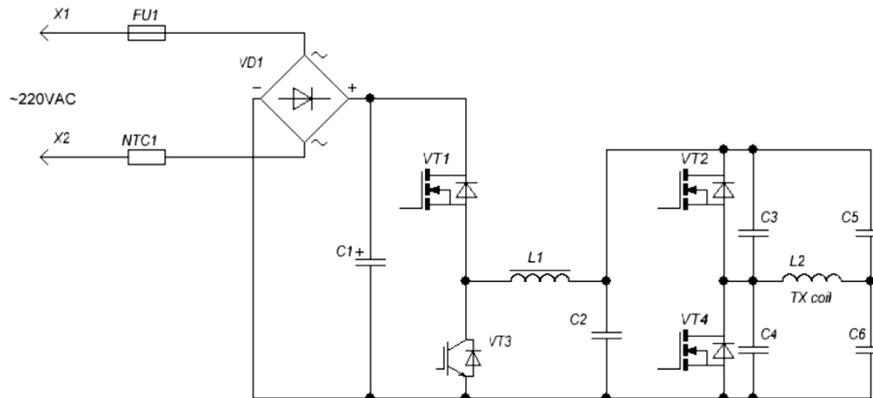


Рис. 1. Схема силовой части передатчика системы беспроводной передачи энергии

Применять в качестве VT3 полевой транзистор (MOSFET) не представляется рациональным из-за большого времени обратного восстановления диода, что характерно для высоковольтных MOSFET.

Выходной полумостовой преобразователь VT2VT4 работает с коэффициентом заполнения 50% и нагружен на передающий контур L2C5C6. Назначение конденсаторов C3 и C4 – обеспечение «мягкого» выключения силовых транзисторов VT2 и VT4. Так как выходной ток полумоста при работе системы индуктивный, то для снижения коммутационных потерь необходимо задержать рост напряжения на закрывающемся транзисторе. Это и обеспечивают конденсаторы C3 и C4, установленные максимально близко к соответствующим транзисторам – в противном случае индуктивности монтажа сведут на нет эффект от их установки. Открытие транзисторов при индуктивном характере выходного тока происходит мягко, при близком к нулю напряжении на канале, так как ток в это время протекает через обратный диод.

Таким образом, амплитуда напряжения на передающем контуре регулируется преобразователем VT1VT3L1C2, а частота – преобразователем VT2VT4L2C3C4C5C6.

Система управления передатчиком выполнена на однокристальном микроконтроллере с ядром ARM. На бортовой АЦП поступают сигналы: напряжение на C1, напряжение на C2, ток в L1, напряжение в средней точке соединения C5 и C6. Последний сигнал используется для оценки амплитуды и фазы тока в L2 и позволяет обойтись без установки дополнительного датчика тока. Контроллер формирует сигналы управления драйверами затворов VT1 – VT4.

Приёмник системы беспроводной передачи энергии. Принципиальная электрическая схема силовой части приемника приведена на рис. 2.

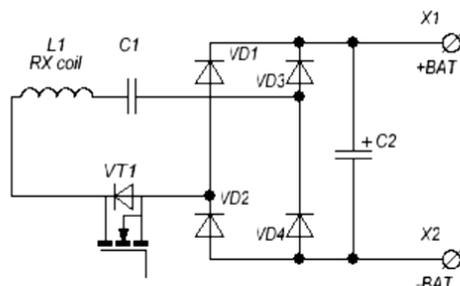


Рис. 2. Схема силовой части приемника системы беспроводной передачи энергии

Приемник состоит из контура L1C1, отключающего транзистора VT1, моста VD1 – VD4 (здесь в качестве диодов применены специализированные микросхемы синхронного выпрямления со встроенным силовым транзистором) и выходного сглаживающего конденсатора C2.

При выборе типов и номиналов элементов приемника учитывались максимальное напряжение на приемном контуре (при максимальном выходном токе) и рабочая частота. Эти две переменные однозначно определяют емкость C1 и индуктивность L1. Отключающий транзистор VT1 должен иметь допустимое напряжение «сток-исток» не менее удвоенной амплитуды ЭДС на L1 и минимальное сопротивление канала в открытом состоянии. Также следует учесть, что при закрытии VT1 энергия, запасенная в элементах контура L1C1, будет выделена в его канале в режиме лавинного пробоя. Поэтому паспортная энергия пробоя повторяющимися импульсами для VT1 должна быть не менее максимальной энергии, запасаемой в контуре L1C1.

Алгоритм работы беспроводной системы передачи энергии. Изготовлен экспериментальный образец беспроводной системы передачи энергии с применением описанных принципов построения.

Диаметры передающей и приемной катушек составляли 200 мм, обеспечивалась работа с дистанцией между катушками до 100 мм (при этом коэффициент связи составлял менее 0,1). Выходное напряжение приемника 25,2 В при максимальном токе до 10 А. Рабочая частота системы составляла 50 ... 70 кГц. В силовой части передатчика применены транзисторы AOT2500L, управляемые с применением драйвера IRS2184 с буферными повторителями на выходах; система управления передатчиком реализована на микроконтроллере STM32F030F4P6; в приемнике использованы микросхемы синхронного выпрямления DK5V45R10. КПД системы, измеренный от входа переменного напряжения 220 В 50 Гц до выхода на АКБ, составил не менее 74%.

Реализован следующий алгоритм работы.

В режиме ожидания производится периодический контроль появления приемника. Для этого в режиме опроса на передающую катушку подается напряжение с амплитудой 50 % максимальной и линейно уменьшающейся частотой (от 90 кГц до 30 кГц в эксперименте) и контролируется потребляемая полумостом активная мощность. Если она превышает предустановленное пороговое значение (порог определяется при калибровке – прямым измерением при отсутствии приемной катушки), считается, что приемник присутствует, и передатчик переходит в рабочий режим. Амплитуда напряжения в режиме опроса определяет, при каком коэффициенте связи (т.е. при каком расстоянии между приемной и передающей катушками) целесообразно начинать работу.

В рабочем режиме функционирует алгоритм поиска точки максимальной мощности типа РАО (Perturb And Observe). Рабочая частота изменяется на небольшую величину Δf , оценивается изменение потребляемой полумостом передатчика мощности ΔP . В экспериментальном образце время усреднения измерений составляло 20 мс (для уменьшения влияния помех с частотой сети). Если $\Delta P > 0$, изменение частоты продолжается в ту же сторону и Δf увеличивается по абсолютной величине, в противном случае знак Δf меняется на противоположный, а модуль Δf уменьшается до минимума.

Кроме этого, если мощность превышает предустановленный максимальный порог, изменение частоты блокируется, а напряжение питания полумоста уменьшается до снижения мощности до безопасной величины. Напротив, если точка максимальной мощности найдена, а величина мощности недостаточна, напряжение питания полумоста увеличивается.

При падении мощности ниже предустановленного порога передатчик переходит в режим опроса. Это произойдет либо при удалении приемника на значительное расстояние, либо при отключении приемника (закрытие VT1, рис. 2) при достижении конечного напряжения конца заряда АКБ, подключенной к приемнику.

В передатчике также реализован контроль температуры силовых компонентов и отключение при перегреве.

В приемнике наиболее тяжелый тепловой режим наблюдался у резонансного конденсатора (C1, рис. 2), собранного из 4-х параллельно включенных металлопленочных конденсаторов СВВ81 0,022 мкФ × 1600 В.

Микросборки синхронных выпрямителей VD1 – VD4 дополнительного теплоотвода не требовали, достаточно небольших полигонов на печатной плате. При выходном токе 10 А суммарная мощность, рассеиваемая на VD1 – VD4, не превышала 2 Вт.

Для реализации возможности регулирования передаваемой мощности следует обеспечить передачу данных между приемником и передатчиком. По мнению авторов, наиболее рационально использовать для этой цели маломощный радиоканал, поскольку использование методов параметрической передачи данных за счет изменения параметров приемного контура [21] при мощностях в сотни Вт ведут к существенному увеличению габаритов, массы и себестоимости приемной части. В экспериментальном образце реализована возможность радиосвязи с применением микросхем SYN115 в передатчике и SYN480R в приемнике. Они позволяют организовать передачу двоичных данных на частоте 433 МГц с амплитудной манипуляцией несущей. Использовалось программное кодирование ВМС (Biphase Mark Coding) с контролем целостности пакетов данных с помощью контрольной суммы CRC8.

При наличии канала передачи данных становится возможным также определение наличия посторонних предметов в поле передающей катушки, а также оптимизация режима работы по условию максимума КПД. Если же канал связи между приемником и передатчиком отсутствует, определение наличия инородных проводящих предметов в рабочей зоне невозможно – они неотличимы от приемника по влиянию на эквивалентные параметры передающей катушки.

Заключение. Предложен вариант построения простой системы беспроводной передачи электроэнергии, работающей в большом диапазоне изменения коэффициента связи приемной и передающей катушек, нетребовательной к параметрам компонентов, имеющей несложный алгоритм управления передатчиком и простое схемное решение приемной части.

Подобные системы беспроводной передачи энергии могут найти применение, например, для зарядки разведывательных БПЛА с вертикальным взлетом/посадкой, для зарядки подводных или наземных беспилотных аппаратов. Это позволит организовать их длительную автономную работу и многократную зарядку их батарей без необходимости привлечения обслуживающего технического персонала. Также подобные беспроводные системы меньшей мощности могут быть интегрированы в портативные изделия (рации, фонари и т.п.).

Работа выполняется при финансовой поддержке Минобрнауки России в рамках государственного задания FRRG-2023-0011.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Costanzo A., Dionigi M., Masotti D., Mongiardo M., Monti G., Tarricone L., and Sorrentino R. Electromagnetic Energy Harvesting and Wireless Power Transmission: A Unified Approach // Proceedings of the IEEE. – Nov. 2014. – Vol. 102, No. 11. – P. 1692-1711.
2. Covic G.A. and Boys J.T. Inductive Power Transfer // Proceedings of the IEEE. – June 2013. – Vol. 101, No. 6. – P. 1276-1289.
3. Bosshard R., Badstübner U., Kolar J.W., Stevanovic I. Comparative Evaluation of Control Methods for Inductive Power Transfer // Proceedings of the International Conference on Renewable Energy Research and Applications (ICRERA 2012), Nagasaki, Japan, November 11-14, 2012. – DOI: 10.1109/ICRERA.2012.6477400.
4. Garnica J., Chinga R.A. and Lin J. Wireless Power Transmission: From Far Field to Near Field // Proceedings of the IEEE. – June 2013. – Vol. 101, No. 6. – P. 1321-1331.
5. Kurs A., Karalis A., Moffatt R., Joannopoulos J.D., Fisher P., and Soljacic M. Wireless Power Transfer via Strongly Coupled Magnetic Resonances // Science. – June 2007. – Vol. 317, No. 5834. – P. 83-86.
6. Wei X., Wang Z., and Dai H. A Critical Review of Wireless Power Transfer via Strongly Coupled Magnetic Resonances // Energies. – July 2014. – Vol. 7, No. 7. – P. 4316-4341.
7. Nambiar S.C., Manteghi M. A simple wireless power transfer scheme for implanted devices // Radio Science Meeting (USNC-URSI NRS), United States National Committee of URSI National, 2014.
8. Ho S.L., Wang J., Fu W.N., and Sun M. A Comparative Study Between Novel Witricity and Traditional Inductive Magnetic Coupling in Wireless Charging // IEEE Transactions on Magnetics. – May 2011. – Vol. 47, No. 5. – P. 1522-1525.
9. Burlaka V.V., Podnebennaya S.K., Gulakov S.V. Analysis of Approaches to the Efficiency Improvement of Wireless Power Transmission Systems Using Low-Frequency Magnetic Fields // In proceedings of 2018 IEEE 38th International Conference on Electronics And Nanotechnology (ELNANO), Kyiv, 24-26 April, 2018. – Kyiv: Igor Sikorsky Kyiv Polytechnic Institute, 2018. – P. 572-575. – DOI: 10.1109/ELNANO.2018.8477481.
10. Qi Wireless Power Transfer System Description. Vol. I: Low Power Part 1: Interface Definition. Version 1.0.1. – Wireless Power Consortium, October 2010.
11. Qi Wireless Charging. – Режим доступа: www.qiwireless.com.
12. Wireless Power Consortium. – Режим доступа: www.wirelesspowerconsortium.com.

13. *Bosshard R., Kolar J.W.* Inductive Power Transfer for Electric Vehicle Charging – Technical Challenges and Tradeoffs // *IEEE Power Electronics Magazine*. – September 2016. – P. 22-30. – DOI: 10.1109/MPREL.2016.2583839.
14. *Qiu Chun, Chau K.T., Chunhua Liu, Chan C.C.* Overview of Wireless Power Transfer for Electric Vehicle Charging // *Proceedings of EVS27 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium (17-20 November 2013)*. – Barcelona, 2013. – P. 1-9.
15. *Villa J.L., Sanz J., Sallan J.* Inductive battery charging system for electric vehicles // *Proceedings of EVS27 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium (17-20 November 2013)*. – Barcelona, 2013. – P. 1-4.
16. *Bosshard R., Kolar J.W., Wunsch B.* Control Method for Inductive Power Transfer with High Partial-Load Efficiency and Resonance Tracking // *Proceedings of the International Power Electronics Conference – ECCE Asia (IPEC 2014), Hiroshima, Japan, May 18-21, 2014*. – DOI: 10.1109/IPEC.2014.6869889.
17. *Pinuela M., Yates D.C., Lucyszyn S., Mitcheson P.D.* Maximizing DC-to-Load Efficiency for Inductive Power Transfer // *IEEE Transactions on Power Electronics*. – 2013. – Vol. 28. – P. 2437-2447.
18. *Wang C.S., Covic G.A., and Stielau O.H.* Power Transfer Capability and Bifurcation Phenomena of Loosely Coupled Inductive Power Transfer Systems // *IEEE Trans. Ind. Electron.* – Feb. 2004. – Vol. 51, No. 1. – P. 148-157.
19. *Pantic Z., and Lukic S.M.* Framework and Topology for Active Tuning of Parallel Compensated Receivers in Power Transfer Systems // *IEEE Trans. Power Electron.* – Nov. 2012. – Vol. 27, No. 11. – P. 4503-4513.
20. *Бессонов Л.А.* Теоретические основы электротехники. В 2 т. Т. 1. Электрические цепи: учебник для ВУЗов. – 11-е изд., испр. и доп. – М.: Гардарики, 2007. – 701 с.
21. *MicroID ® 125 kHz RFID System Design Guide. Document No. DS51115F.* Microchip Technology Inc. – 2004. – 210 p. – Режим доступа: ww1.microchip.com/downloads/en/devicedoc/51115f.pdf.

REFERENCES

1. *Costanzo A., Dionigi M., Masotti D., Mongiardo M., Monti G., Tarricone L., and Sorrentino R.* Electromagnetic Energy Harvesting and Wireless Power Transmission: A Unified Approach, *Proceedings of the IEEE*, Nov. 2014, Vol. 102, No. 11, pp. 1692-1711.
2. *Covic G.A. and Boys J.T.* Inductive Power Transfer, *Proceedings of the IEEE*, June 2013, Vol. 101, No. 6, pp. 1276-1289.
3. *Bosshard R., Badstübner U., Kolar J.W., Stevanovic I.* Comparative Evaluation of Control Methods for Inductive Power Transfer, *Proceedings of the International Conference on Renewable Energy Research and Applications (ICRERA 2012), Nagasaki, Japan, November 11-14, 2012*. DOI: 10.1109/ICRERA.2012.6477400.
4. *Garnica J., Chinga R.A. and Lin J.* Wireless Power Transmission: From Far Field to Near Field, *Proceedings of the IEEE*, June 2013, Vol. 101, No. 6, pp. 1321-1331.
5. *Kurs A., Karalis A., Moffatt R., Joannopoulos J.D., Fisher P., and Soljacic M.* Wireless Power Transfer via Strongly Coupled Magnetic Resonances, *Science*, June 2007, Vol. 317, No. 5834, pp. 83-86.
6. *Wei X., Wang Z., and Dai H.* A Critical Review of Wireless Power Transfer via Strongly Coupled Magnetic Resonances, *Energies*, July 2014, Vol. 7, No. 7, pp. 4316-4341.
7. *Nambiar S.C., Manteghi M.* A simple wireless power transfer scheme for implanted devices, *Radio Science Meeting (USNC-URSI NRS), United States National Committee of URSI National, 2014*.
8. *Ho S.L., Wang J., Fu W.N., and Sun M.* A Comparative Study Between Novel Witricity and Traditional Inductive Magnetic Coupling in Wireless Charging, *IEEE Transactions on Magnetics*, May 2011, Vol. 47, No. 5, pp. 1522-1525.
9. *Burlaka V.V., Podnebennaya S.K., Gulakov S.V.* Analysis of Approaches to the Efficiency Improvement of Wireless Power Transmission Systems Using Low-Frequency Magnetic Fields, *In proceedings of 2018 IEEE 38th International Conference on Electronics And Nanotechnology (ELNANO), Kyiv, 24-26 April, 2018*. Kyiv: Igor Sikorsky Kyiv Polytechnic Institute, 2018, pp. 572-575. DOI: 10.1109/ELNANO.2018.8477481.

10. Qi Wireless Power Transfer System Description. Vol. I: Low Power Part 1: Interface Definition. Version 1.0.1. – Wireless Power Consortium, October 2010.
11. Qi Wireless Charging. Available at: www.qiwireless.com.
12. Wireless Power Consortium. Available at: www.wirelesspowerconsortium.com.
13. *Bosshard R., Kolar J.W.* Inductive Power Transfer for Electric Vehicle Charging – Technical Challenges and Tradeoffs, *IEEE Power Electronics Magazine*, September 2016, pp. 22-30. DOI: 10.1109/MPEL.2016.2583839.
14. *Qiu Chun, Chau K.T., Chunhua Liu, Chan C.C.* Overview of Wireless Power Transfer for Electric Vehicle Charging, *Proceedings of EVS27 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium (17-20 November 2013)*. Barcelona, 2013, pp. 1-9.
15. *Villa J.L., Sanz J., Sallan J.* Inductive battery charging system for electric vehicles, *Proceedings of EVS27 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium (17-20 November 2013)*. Barcelona, 2013, pp. 1-4.
16. *Bosshard R., Kolar J.W., Wunsch B.* Control Method for Inductive Power Transfer with High Partial-Load Efficiency and Resonance Tracking, *Proceedings of the International Power Electronics Conference – ECCE Asia (IPEC 2014), Hiroshima, Japan, May 18-21, 2014*. DOI: 10.1109/IPEC.2014.6869889.
17. *Pinuela M., Yates D.C., Lucyszyn S., Mitcheson P.D.* Maximizing DC-to-Load Efficiency for Inductive Power Transfer, *IEEE Transactions on Power Electronics*, 2013, Vol. 28, pp. 2437-2447.
18. *Wang C.S., Covic G.A., and Stielau O.H.* Power Transfer Capability and Bifurcation Phenomena of Loosely Coupled Inductive Power Transfer Systems, *IEEE Trans. Ind. Electron*, Feb. 2004, Vol. 51, No. 1, pp. 148-157.
19. *Pantic Z., and Lukic S.M.* Framework and Topology for Active Tuning of Parallel Compensated Receivers in Power Transfer Systems, *IEEE Trans. Power Electron*, Nov. 2012, Vol. 27, No. 11, pp. 4503-4513.
20. *Bessonov L.A.* Теоретические основы электротехники. В 2 т. Т. 1. Электрические цепи: учебник для ВУЗов [Theoretical foundations of electrical engineering. In 2 vol. Vol. 1. Electric circuits: a textbook for universities. 11th ed. Moscow: Gardariki, 2007, 701 p.
21. MicroID @ 125 kHz RFID System Design Guide. Document No. DS51115F. Microchip Technology Inc., 2004, 210 p. Available at: ww1.microchip.com/downloads/en/devicedoc/51115f.pdf.

Статью рекомендовал к опубликованию д.т.н., профессор К.Е. Румянцев.

Бурлака Владимир Владимирович – ФГБОУ ВО "ПГТУ"; e-mail: vburlaka@rambler.ru; г. Мариуполь, Россия; тел.: +79497048603; кафедра автоматизации электро- и теплоэнергетических комплексов; зав. кафедрой; д.т.н.; доцент.

Гулаков Сергей Владимирович – e-mail: gulakov.s.v@yandex.ru; тел.: +79497298638; кафедра автоматизации электро- и теплоэнергетических комплексов; д.т.н.; профессор.

Головин Андрей Юрьевич –e-mail: a.golovin@tms-2022.ru; тел.: +79497128054; кафедра автоматизации электро- и теплоэнергетических комплексов; м.н.с.

Мироненко Дмитрий Сергеевич – e-mail: mironenko.d.s@yandex.ru; тел.: +79496094732; кафедра информационных систем, технологий и автоматизации в строительстве; зав. кафедрой; к.т.н.; доцент.

Burlaka Vladimir Vladimirovich – FSBEI of HE "PSTU"; e-mail: vburlaka@rambler.ru; Mariupol, Russia; phone: +79497048603; the department of automation of electrical and thermal power complexes; head of the department; dr. of eng. sc.; associate professor.

Gulakov Sergey Vladimirovich – e-mail: gulakov.s.v@yandex.ru; phone: +79497298638; the department of automation of electrical and thermal power complexes; dr. of eng. sc.; professor.

Golovin Andrey Yurievich – e-mail: a.golovin@tms-2022.ru; phone: +79497128054; the department of automation of electrical and thermal power complexes; research associate.

Mironenko Dmitrii Sergeevich – e-mail: mironenko.d.s@yandex.ru; phone: +79496094732; the department of information systems, technologies and automation in construction; head of the department; cand. of eng. sc.; associate professor.

А.А. Жук

**БЫСТРОДЕЙСТВУЮЩИЕ ВЫХОДНЫЕ КАСКАДЫ ОПЕРАЦИОННЫХ
УСИЛИТЕЛЕЙ С ДИФФЕРЕНЦИРУЮЩЕЙ ЦЕПЬЮ КОРРЕКЦИИ
ПЕРЕХОДНОГО ПРОЦЕССА***

Впервые представлены схемотехнические решения с повышенной максимальной скоростью нарастания (спада) выходного напряжения семейства классических выходных каскадов класса АВ, которые являются основой многих интегральных микросхем операционных усилителей (544УД1, 153УД4, μ A741 и др.). Для этой цели в базовые схемы вводятся специальные элементы коррекции переходного процесса в режиме большого сигнала. Схемотехника данного класса реализуется как на биполярных (BJT), так и на КМОП транзисторах. Результаты компьютерного моделирования в среде LTspice XVII показывают, что в сравнении с классическими схемами за счет введения дополнительного дифференцирующего конденсатора и входного эмиттерного повторителя максимальная скорость спада выходного напряжения выходных каскадов увеличивается более чем в 500 раз. При этом рассмотренные схемы обеспечивают выходные напряжения с максимальной амплитудой от - 8.5 В до + 10 В при сравнительно низком сопротивлении нагрузки (до 2 кОм) и напряжениях питания ± 10 В. Показано, что при мелкосерийном производстве предложенных выходных каскадов рекомендовано их выполнение на базовых матричных кристаллах МН2ХА031 (ОАО «Интеграл», г. Минск, Беларусь), что позволит снизить себестоимость изготовления микросэлектронных изделий. Базовый матричный кристалл МН2ХА031 на основе техпроцесса 3СВiТ допускает работу предлагаемых схем в условиях низких температур (до - 197 °С), а также при воздействии проникающей радиации с поглощенной дозой гамма-квантов (до 1 Мрад) и потока нейтронов (до 10^{13} н/см²).

Аналоговая схемотехника; выходной каскад; операционный усилитель; дифференцирующая цепь коррекции; скорость нарастания выходного напряжения; полевые транзисторы; биполярные транзисторы.

А.А. Zhuk

**HIGH-SPEED OUTPUT STAGES OF OPERATIONAL AMPLIFIERS WITH
DIFFERENCING CIRCUIT CORRECTION OF TRANSITION PROCESS**

For the first time, a circuit design solution with an increased maximum output voltage rise (decay) rate of a family of classic AB class output stages, which are the basis of many integrated circuits of operational amplifiers (544UD1, 153UD4, μ A741, etc.), is presented. For this purpose, special elements of transient process in the large signal mode are introduced into the basic circuits of the output stages. The circuitry of this class is implemented on both bipolar (BJT) and CMOS transistors. The results of computer simulation in the LTspice XVII simulation software show that, in comparison with classical circuits, due to the introduction of an additional differentiating capacitor and an input emitter repeater, the maximum rate of decay in the output voltage of the output stages increases by more than 500 times. At the same time, the considered output stage circuits provide output voltages with maximum amplitude from - 8.5 V to + 10 V with a relatively low load resistance (up to 2 kOhm) and supply voltages of ± 10 V. It is shown that for small-scale production of the proposed output stages, their execution on basic matrix crystals MH2XA031 is recommended (OJSC "Integral", Minsk, Belarus), which will reduce the cost-effectiveness of manufacturing microelectronic products. The basic matrix crystal MH2XA031 based on the 3CBiT

* Исследование выполнено за счет гранта Российского научного фонда (проект № 23-79-10069).

process technology allows the proposed schemes to operate at low temperatures (up to $-197\text{ }^{\circ}\text{C}$), as well as under the influence of penetrating radiation with an absorbed dose of gamma quanta (up to 1 Mrad) and a neutron flux (up to 10^{13} n/cm^2).

Analogue circuitry; output stage; operational amplifier; differencing circuit correction; slew rate of output voltage; field-effect transistor; bipolar transistor.

Введение. Известно значительное количество схем выходных каскадов (ВК) аналоговых микроселектронных изделий, которые реализуются в виде различных модификаций эмиттерных (истоковых) повторителей [1–10] на биполярных (BJT) и полевых (JFet, КМОП, КНИ, КНС и др.) транзисторах, а также при их совместном включении. Известны также схемы эмиттерных повторителей [11–19], которые используются для согласования источника сигнала с низкоомной нагрузкой. Однако во многих ВК рассматриваемого класса не обеспечивается высокое быстродействие по одному из фронтов импульсных входных сигналов.

Основная цель настоящей статьи состоит в разработке и исследовании семейства классических выходных каскадов класса АВ с повышенной максимальной скоростью нарастания выходного напряжения при их реализации на основе радиационно-стойких базовых матричных кристаллов ОАО «Интеграл» МН2ХА030/031, в которых за счет введения дополнительного дифференцирующего конденсатора и дополнительного входного эмиттерного повторителя увеличивается максимальная скорость спада выходного напряжения.

1. Схема классического ВК. На рис. 1 изображена классическая схема ВК, которая используется в российских микросхемах операционных усилителей (ОУ) (544УД1 [1, с. 214], 153УД4 [1, стр. 224, рис. 4.12а], а также в ряде зарубежных микросхем, например, $\mu\text{A}741$ [3, стр. 323, рис. 15, стр. 331, рис. 29]). Рассматриваемая схема рекомендована для массового практического применения в аналоговых устройствах (см. [2], стр. 147, рис. 4.18, стр. 144, рис. 4.16, стр. 139, рис. 4.10). Конденсаторы C_{p1} и C_{p2} моделируют паразитные емкости в цепях базы выходных транзисторов VT2, VT3.

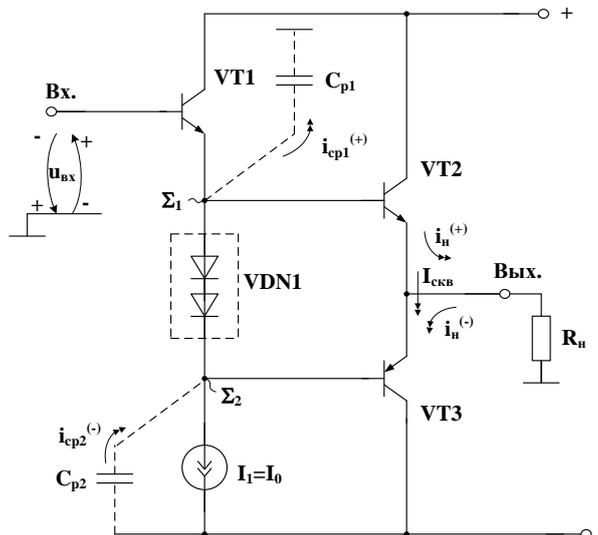


Рис. 1. Схема классического ВК [1]

Так, для схемы рис. 1 максимальная скорость нарастания выходного напряжения для отрицательной полярности импульсного входного сигнала ($SR^{(-)}$) существенно меньше, чем $SR^{(+)}$ для положительного входного сигнала. Это ограничивает ее области применения.

2. Схема предлагаемого быстродействующего ВК и результаты его компьютерного моделирования. На рис. 2 представлена схема предлагаемого ВК ОУ [20], который обеспечивает высокую скорость нарастания выходного напряжения, как для положительного, так и для отрицательного импульсного входного напряжения. Это обеспечивается за счет введения дополнительного корректирующего конденсатора C_k и входного эмиттерного повторителя на транзисторе VT4.

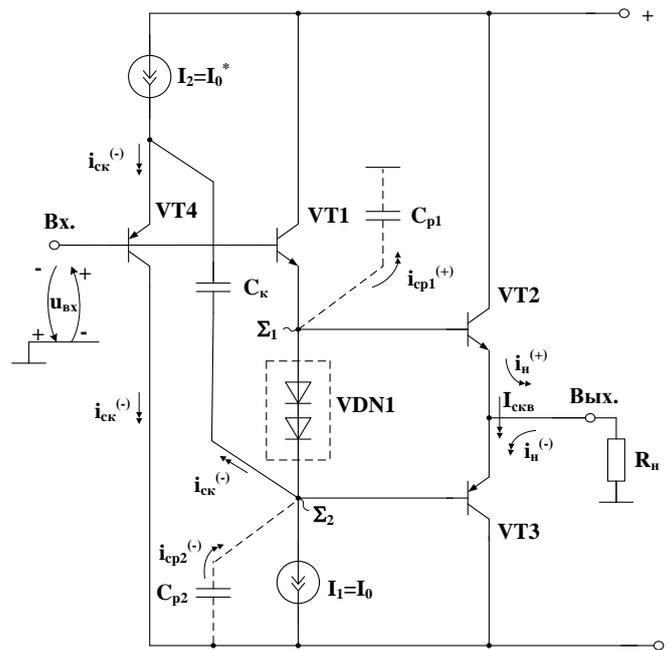


Рис. 2. Схема предлагаемого ВК ОУ [20]

2.1 Описание работы схемы ВК на рис. 2 при положительном входном импульсном сигнале большой амплитуды. Если на вход подается положительный импульсный сигнал большой амплитуды, то он с минимальной задержкой передается в цепь базы выходного транзистора VT2 и далее в цепь нагрузки R_n . При этом паразитная емкость C_{p1} перезаряжается большим импульсным током по цепи эмиттера входного транзистора VT1. Таким образом, в данном режиме рассматриваемый ВК имеет высокое быстродействие, что подтверждается результатами компьютерного моделирования на рис. 4 и данными из табл. 1, которые расположены ниже.

2.2 Работа схемы ВК (рис. 2) при отрицательном входном импульсном сигнале большой амплитуды. Если на вход подается отрицательный импульс большой амплитуды, то входной транзистор VT1 практически мгновенно запирается, а напряжение на базе выходного транзистора VT3 ($u_{\Sigma 2}$), определяющее закон изменения отрицательного выходного напряжения в нагрузке R_n , зависит от скорости перезаряда паразитной емкости C_{p2} :

$$\frac{du_{\Sigma 2}}{dt} = \frac{I_1 + i_{\text{CK}}^{(-)}}{C_{p2}}, \quad (1)$$

где I_1 – статический ток источника опорного тока I_1 , $i_{\text{CK}}^{(-)}$ – импульсный ток в конденсаторе C_K , зависящий от разности напряжений на эмиттере транзистора VT4 и базе выходного транзистора VT3.

Если выбрать емкость дополнительного корректирующего конденсатора C_K значительно больше, чем емкость паразитного конденсатора C_{p2} , то введение C_K существенно ускоряет процесс перезаряда паразитной емкости C_{p2} . Этот эффект положительно сказывается на скорости изменения напряжения на базе выходного транзистора VT3 ($u_{\Sigma 2}$), что в свою очередь повышает скорость изменения выходного напряжения ВК. Данный вывод подтверждается результатами компьютерного моделирования на рис. 5 и данными из табл. 1.

2.3 Результаты компьютерного моделирования ВК на рис.2. На рис. 3 показана схема для моделирования ВК на рис. 2 в среде LTspice при $t=27^\circ\text{C}$, $+V_{\text{CC}}=-V_{\text{EE}}=10\text{ В}$, $I_1\dots I_2=100\text{ мкА}$, $C_{p1}\dots C_{p2}=1\text{ пФ}$, $R_{\text{load}}=1\text{ МОм}$, $C_K=0$. При этом здесь и далее использовались модели транзисторов базового матричного кристалла МН2ХА031_25.01.21 Минского ОАО «Интеграл».

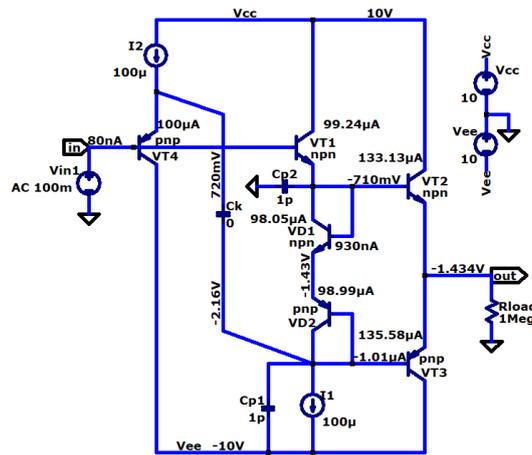


Рис. 3. Схема для моделирования ВК на рис. 2 в среде LTspice при $t=27^\circ\text{C}$

На рис. 4 представлена переходная характеристика переднего фронта выходного напряжения ВК на рис. 3 при $t=27^\circ\text{C}$, амплитуде входного импульса $u_{\text{вх}}=8\text{ В}$ и разных значениях емкости дополнительного корректирующего конденсатора C_K (при $C_K=0/10\text{ пФ}/20\text{ пФ}/30\text{ пФ}$).

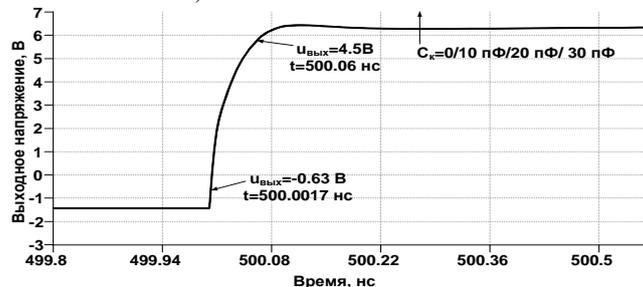


Рис. 4. Переходная характеристика переднего фронта выходного напряжения ВК на рис. 3

График на рис. 4 показывает, что дополнительный корректирующий конденсатор C_k и входной эмиттерный повторитель на транзисторе VT4 не оказывают влияние на скорость нарастания переднего фронта выходного напряжения.

На рис. 5 показана переходная характеристика заднего фронта выходного напряжения ВК на рис. 3 при $t=27^\circ\text{C}$, амплитуде входного импульса $u_{вх}=8\text{ В}$ и разных значениях емкости дополнительного корректирующего конденсатора C_k (при $C_k=0/10\text{ пФ}/20\text{ пФ}/30\text{ пФ}$).

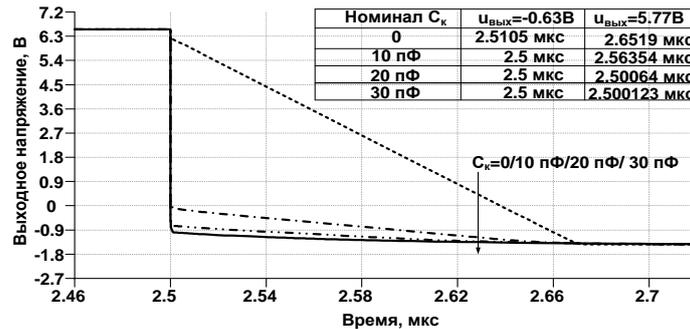


Рис. 5. Переходная характеристика заднего фронта выходного напряжения ВК на рис. 3

Таблица 1

Зависимость максимальной скорости нарастания выходного напряжения ВК на рис. 3 от емкости дополнительного корректирующего конденсатора (C_k)

Значение емкости C_k , пФ	Скорость нарастания переднего фронта ВК (рис. 4), $SR^{(+)}$	Скорость спада заднего фронта ВК (рис. 5), $SR^{(-)}$
0	109777 В/мкс	45,24 В/мкс
10	109777 В/мкс	100,72 В/мкс
20	109777 В/мкс	10000 В/мкс
30	109777 В/мкс	52032 В/мкс

Из табл. 1 следует, что введение транзистора VT4, источника опорного тока I_2 и конденсатора C_k повышает максимальную скорость нарастания выходного напряжения ВК для отрицательного фронта с 45,24 В/мкс до 52032 В/мкс, т.е. более чем в 1000 раз.

3. Модифицированная схема ВК и результаты его компьютерного моделирования. Применение специальной цепи смещения потенциалов на транзисторах VT5 и VT6, как показано на рис. 6, позволяет за счет изменения сопротивления резистора R1 управлять в широких пределах величиной сквозного тока выходных транзисторов VT2, VT3, а также регулировать зону нечувствительности на амплитудной характеристике ВК (рис. 7).

На рис. 7 приведена амплитудная характеристика ВК на рис. 3 в среде LTspice с разными сопротивлениями нагрузки при $t=27^\circ\text{C}$.

Компьютерное моделирование показывает, что максимальный сквозной ток выходных транзисторов VT2, VT3 устанавливается на уровне тока I_2 при нулевом сопротивлении резистора R1, а увеличение сопротивления этого резистора уменьшает величину сквозного тока до заданного, в т.ч. микроамперного, значения. Это

позволяет переводить выходные транзисторы VT2, VT3 в режим малых сквозных токов (10-20 мкА), что положительно сказывается на общем токопотреблении ВК при сохранении высоких значений $SR^{(-)}$ для отрицательного фронта (табл. 2).

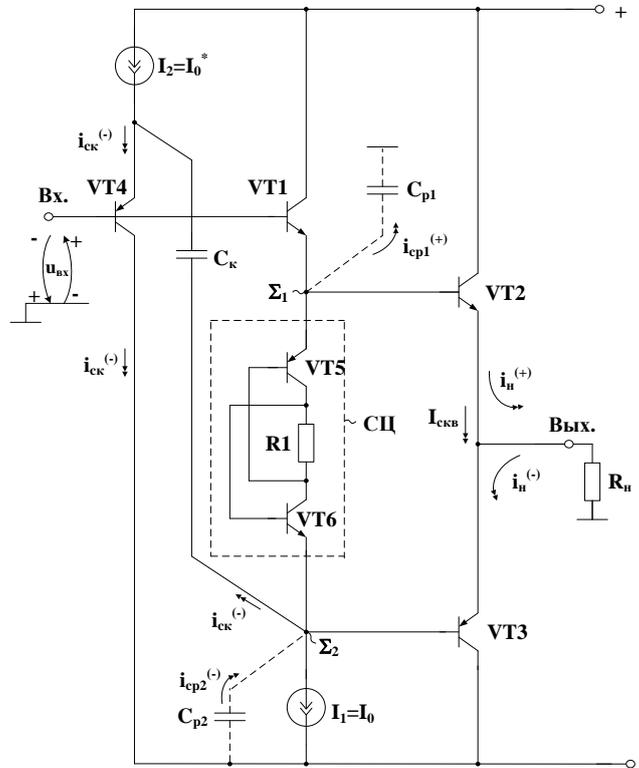


Рис. 6. Модифицированная схема ВК ОУ со специальной цепью установления сквозного тока выходных транзисторов [20]



Рис. 7. Амплитудная характеристика ВК на рис. 3

На рис. 8 представлена схема для моделирования быстродействующего ВК (рис. 6) в среде LTspice при $t=27^\circ\text{C}$, $+V_{\text{cc}}=-V_{\text{ee}}=10 \text{ В}$, $I_1 \dots I_2=100 \text{ мкА}$, $C_{p1} \dots C_{p2}=1 \text{ пФ}$, $R_{\text{load}}=1 \text{ МОм}$, $C_{\kappa}=0$, $R_1=1 \text{ кОм}$.

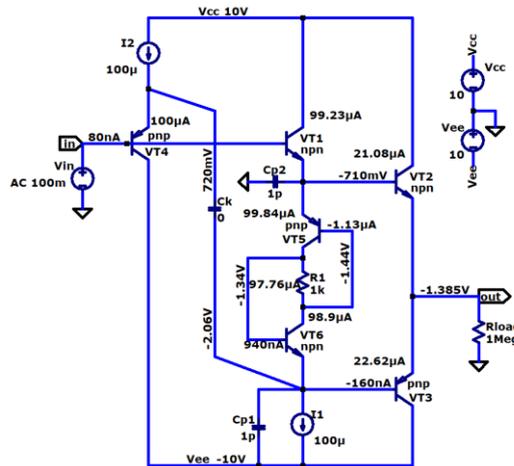


Рис. 8. Схема для моделирования ВК на рис. 6

На рис. 9 показана амплитудная характеристика ВК на рис. 8 в среде LTspice при $t=27^{\circ}\text{C}$.



Рис. 9. Амплитудная характеристика ВК на рис. 8 в среде LTspice

На рис. 10 приведена переходная характеристика заднего фронта выходного напряжения ВК на рис. 8 при $t=27^{\circ}\text{C}$, амплитуде входного импульса $u_{\text{вх}}=8\text{ В}$ и разных значениях емкости дополнительного корректирующего конденсатора $C_{\text{к}}$ (при $C_{\text{к}}=0/10\text{ пФ}/20\text{ пФ}/30\text{ пФ}$).

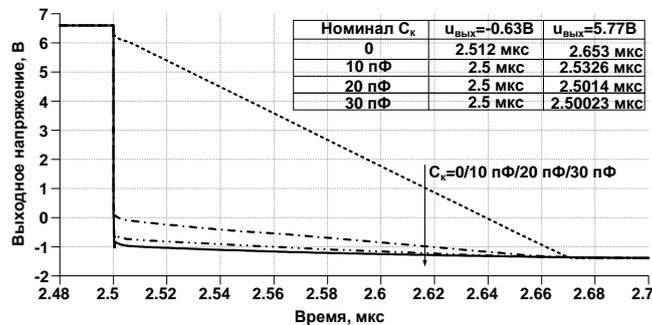


Рис. 10. Переходная характеристика заднего фронта выходного напряжения ВК на рис. 8

Таблица 2

**Зависимость максимальной скорости нарастания выходного напряжения
ВК на рис. 8 от емкости дополнительного корректирующего
конденсатора (C_k)**

Значение емкости C_k , пФ	Скорость нарастания переднего фронта ВК (рис. 10), $SR^{(+)}$	Скорость нарастания спада заднего фронта ВК (рис. 10), $SR^{(-)}$
0	93704 В/мкс	45,22 В/мкс
10	93704 В/мкс	196,31 В/мкс
20	93704 В/мкс	4571,42 В/мкс
30	93704 В/мкс	25600 В/мкс

4. Пример построения рассмотренных ВК на КМОП транзисторах. Рассмотренные ранее схемотехнические решения эффективны и при реализации предлагаемого ВК на КМОП транзисторах.

Данная КМОП схема (рис. 11) получена путем формальной замены биполярных транзисторов на КМОП транзисторы, которые имеют такие же полярности напряжений между выводами затвора, истока и стока, что и биполярные транзисторы между базой, эмиттером и коллектором.

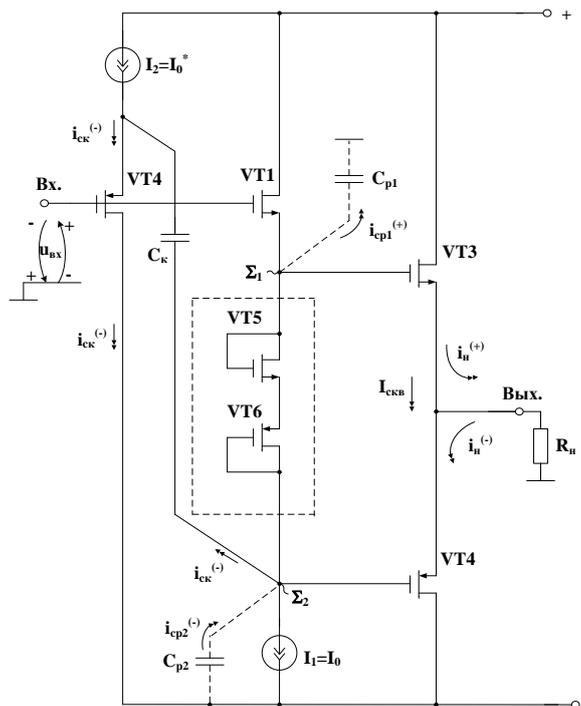


Рис. 11. Пример построения ВК ОУ при его реализации на КМОП транзисторах [20]

Заключение. Результаты компьютерного моделирования в среде LTspice показывают, что в сравнении с классическими схемами за счет введения дополнительного дифференцирующего конденсатора C_k и дополнительного входного эмиттерного повторителя на транзисторе VT4 максимальная скорость спада вы-

ходного напряжения выходных каскадов увеличивается более чем в 500 раз (до 25000 В/мкс) для схем на биполярных транзисторах. При этом рассмотренные схемы обеспечивают выходные напряжения с максимальной амплитудой от - 8.5 В до + 10 В при сравнительно низком сопротивлении нагрузки (до 2 кОм) и напряжениях питания ± 10 В. Схемотехника рассмотренных выходных каскадов может также быть адаптирована под технологические процессы, содержащие КМОП транзисторы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Якубовский С.В., Барканов Н.А., Ниссельсон Л.И. и др. Аналоговые и цифровые интегральные схемы. – Сов. радио, 1979.
2. Полонников Д.Е. Операционные усилители: принципы построения, теория, схемотехника. – М.: Энергоатомиздат, 1983.
3. Solomon J.E. The monolithic op amp: a tutorial study // IEEE Journal of Solid-State Circuits. – 1974. – Vol. 9, No. 6. – P. 314-332.
4. Floru F. Operational amplifier output stage. Patent for invention, No. US 6160451 of 2000, fig. 3.
5. Floru F. Improved operational amplifier output stage. Patent for invention, No. EP 1088391 of 2000, fig. 3.
6. Moraveji F. Output stage and method of enhancing output gain. Patent for invention, No. US 6586998 of 2003.
7. Dow R.N. Ultra high gain amplifier. Patent for invention, No. US 5614866 of 1997, fig. 2.
8. Bowers D.F. Rapid slewing unity gain buffer amplifier with boosted parasitic capacitance charging. Patent for invention, No. US 5323122 of 1994, fig. 2.
9. Damitio P., Alenin S. High output current wideband output stage/buffer amplifier. Patent application, No. US 2005/0264358 of 2005, fig. 1.
10. Жук А.А., Прокопенко Н.Н., Тутов А.Е. Двухтактный буферный усилитель на комплементарных биполярных транзисторах, № RU 2786191 от 2022.
11. Акопов В.П., Осипов В.Э. Выходной каскад операционного усилителя, № SU 1480094 от 1989.
12. Басий В.Т., Костырка В.С. и др. Эмиттерный повторитель, № SU 1300632 от 1987.
13. Исаков А.Б., Попов А.Э. и др. Эмиттерный повторитель, № RU 2003217 от 1993.
14. Акопов В.П., Осипов В.Э. Выходной каскад операционного усилителя, № SU 1569944 от 1990.
15. Шилов В.К., Вушкарник В.Н. и др. Повторитель напряжения, № SU 1012420 от 1983.
16. Мельник В.И., Сташків Ю.В. и др. Эмиттерный повторитель, № SU 1298853 от 1987.
17. Капитонов М.В., Прокопенко Н.Н., Ульянова В.И. Эмиттерный повторитель, № SU 1401566 от 1988.
18. Басий В.Т., Бородюк О.А. и др. Эмиттерный повторитель, № SU 1248028 от 1986.
19. Лучинский Л.И., Макеев Е.П. Повторитель напряжения, А.св. СССР № 425308 от 1974.
20. Жук А.А., Прокопенко Н.Н., Сергеенко М.А. Быстродействующий выходной каскад операционного усилителя, № RU 2802051 от 2023.

REFERENCES

1. Yakubovskiy S.V., Barkanov N.A., Nissel'son L.I. i dr. Analogovye i tsifrovye integral'nye skhemy [Analogue and digital integrated circuits]. Sov. radio, 1979.
2. Polonnikov D.E. Operatsionnye usiliteli: printsipy postroeniya, teoriya, skhemotekhnika [Operational amplifiers: principles of construction, theory, circuitry]. Moscow: Energoatomizdat, 1983.
3. Solomon J.E. The monolithic op amp: a tutorial study, *IEEE Journal of Solid-State Circuits*, 1974, Vol. 9, No. 6, pp. 314-332.
4. Floru F. Operational amplifier output stage. Patent for invention, No. US 6160451 of 2000, fig. 3.
5. Floru F. Improved operational amplifier output stage. Patent for invention, No. EP 1088391 of 2000, fig. 3.

6. *Moraveji F.* Output stage and method of enhancing output gain. Patent for invention, No. US 6586998 of 2003.
7. *Dow R.N.* Ultra high gain amplifier. Patent for invention, No. US 5614866 of 1997, fig. 2.
8. *Bowers D.F.* Rapid slewing unity gain buffer amplifier with boosted parasitic capacitance charging. Patent for invention, No. US 5323122 of 1994, fig. 2.
9. *Damitio P., Alenin S.* High output current wideband output stage/buffer amplifier. Patent application, No. US 2005/0264358 of 2005, fig. 1.
10. *Zhuk A.A., Prokopenko N.N., Titov A.E.* Dvukhtaktnyy bufernyy usilitel' na komplementarnykh bipolyarnykh tranzistorakh, № RU 2786191 ot 2022 [Two-cycle buffer amplifier on complementary bipolar transistors, No. RU 2786191 of 2022 (in Russian)].
11. *Akopov V.P., Osipov V.E.* Vykhodnoy kaskad operatsionnogo usilitelya, № SU 1480094 ot 1989 [Output stage of operational amplifier, No. SU 1480094 of 1989 (in Russian)].
12. *Basiy V.T., Kostyrka V.S. i dr.* Emitternyy povtoritel', № SU 1300632 ot 1987 [Emitter repeater, No. SU 1300632 of 1987 (in Russian)].
13. *Isakov A.B., Popov A.E. i dr.* Emitternyy povtoritel', № RU 2003217 ot 1993 [Emitter repeater, No. RU 2003217 of 1993 (in Russian)].
14. *Akopov V.P., Osipov V.E.* Vykhodnoy kaskad operatsionnogo usilitelya, № SU 1569944 ot 1990 [Output stage of operational amplifier, No. SU 1569944 of 1990 (in Russian)].
15. *Shipov V.K., Vushkarnik V.N. i dr.* Povtoritel' napryazheniya, № SU 1012420 ot 1983 [Voltage Repeater, No. SU 1012420 of 1983 (in Russian)].
16. *Mel'nik V.I., Stashkiv Yu.V. i dr.* Emitternyy povtoritel', № SU 1298853 ot 1987 [Emitter repeater, No. SU 1298853 of 1987 (in Russian)].
17. *Kapitonov M.V., Prokopenko N.N., Ul'yanova V.I.* Emitternyy povtoritel', № SU 1401566 ot 1988 [Emitter repeater, No. SU 1401566 of 1988 (in Russian)].
18. *Basiy V.T., Borodyuk O.A. i dr.* Emitternyy povtoritel', № SU 1248028 ot 1986 [Emitter repeater, No. SU 1248028 of 1986 (in Russian)].
19. *Luchinskiy L.I., Makeev E.P.* Povtoritel' napryazheniya, A.sv. SSSR № 425308 ot 1974 [Voltage Repeater, A.sv. USSR № 425308 of 1974 (in Russian)].
20. *Zhuk A.A., Prokopenko N.N., Sergeenko M.A.* Bystrodeystvuyushchiy vykhodnoy kaskad operatsionnogo usilitelya, № RU 2802051 ot 2023 [Fast output stage of operational amplifier, No. RU 2802051 of 2023 (in Russian)].

Статью рекомендовал к опубликованию д.т.н., профессор К.Е. Румянцев.

Жук Алексей Андреевич – Донской государственный технический университет; e-mail: alexey.zhuk96@mail.ru; г. Ростов-на-Дону, Россия; тел: +79185880301; младший научный сотрудник отдела «Управление научных исследований»; ассистент кафедры «Информационные системы и радиотехника».

Zhuk Alexey Andreevich – Don State Technical University; e-mail: alexey.zhuk96@mail.ru; Rostov-on-Don, Russia; phone: +79185880301; junior research fellow of the “Office of scientific research”; assistant of the department of information systems and radio engineering.

УДК 621.396.029

DOI 10.18522/2311-3103-2023-5-192-204

Б.М. Антипин, Е.М. Виноградов, О.А. Симонина, С.Л. Фёдоров

АНАЛИЗ ТРЕБОВАНИЙ К ПАРАМЕТРАМ ИЗЛУЧЕНИЙ РАДИОТЕХНИЧЕСКОГО ОБОРУДОВАНИЯ БЕСПИЛОТНЫХ АВИАЦИОННЫХ СИСТЕМ

На основе анализа отечественных и зарубежных нормативных документов рассматриваются особенности использования радиочастотного спектра в РФ для беспилотных авиационных систем (БАС). Рассмотрены типы радиоканалов для связи БАС и наземного пункта управления и определены частотные диапазоны, разрешённые для использования в целях связи с БАС согласно нормативно-правовой базе Российской Федерации. К лицензиро-

ванным относятся диапазоны 117,975–137 МГц; 5030–5091 МГц; 5850 – 6425 МГц, но есть возможность использования нелицензируемых диапазонов 137–174 МГц; 395–400 МГц; 430–460 МГц; 433,05–434,79 МГц; 860–880 МГц; 902–928 МГц; 1100–1400 МГц; 2400–2483,5 МГц. Показано, что существует проблема использования нелицензируемых диапазонов частот в связи с ограничениями на параметры радиоизлучения и диапазонов систем мобильной связи из-за отсутствия нормативной базы. В нормативно-правовой базе определены требования к уровням излучений радиоэлектронных средств (РЭС) БАС в лицензированных диапазонах, которые не должны превышать 100 мВт. В текущий момент в Российской Федерации нет нормативных требований к техническим параметрам РЭС БАС для их радиоканалов в нелицензируемых диапазонах, но при этом есть ограничения на устройства малого радиуса действия, использующие эти диапазоны. При этом существует ряд нормативных документов других стран, определяющих уровни сигналов в каналах управления полетов. Рассмотрены возможности применения БАС для целей радиоконтроля на примере опытной эксплуатации беспилотного летательного аппарата (БПЛА) «Орлан-10» с учётом существующих ограничений в области нормативно-правовой базы. Показано, что существует ряд ограничений для их эффективного использования в целях радиоконтроля: в области нормативно-правовой базы, технических характеристик антенн и зависимость результатов от высоты полёта. Также важным фактором являются ограничения по климатическим факторам. При этом именно для труднодоступных районов со сложными климатическими условиями такой метод мониторинга является наиболее перспективным.

БПЛА; БАС; выделенные частотные диапазоны; характеристики излучений; радиоконтроль.

B.M. Antipin, E.M. Vinogradov, O.A. Simonina, S.L. Fedorov

THE ANALYSIS OF REQUIREMENTS FOR RADIATION PARAMETERS OF RADIO EQUIPMENT IN UNMANNED AERIAL SYSTEMS

Based on the analysis of domestic and foreign regulatory documents, the features of the use of the radio frequency spectrum in the Russian Federation for unmanned aircraft systems (UAS) are considered. The types of radio channels for communication between the UAS and the ground control station are considered and the frequency ranges allowed for use for communication with the UAS according to the regulatory framework of the Russian Federation are determined. Licensed ranges include 117.975–137 MHz; 5030–5091 MHz; 5850–6425 MHz, but it is possible to use unlicensed bands 137–174 MHz; 395–400 MHz; 430–460 MHz; 433.05–434.79 MHz; 860–880 MHz; 902–928 MHz; 1100–1400 MHz; 2400–2483.5 MHz. It is shown that there is a problem of using unlicensed frequency ranges due to restrictions on the parameters of radio emission and ranges of mobile communication systems due to the lack of a regulatory framework. The regulatory framework defines the requirements for the radiation levels of radio-electronic means (RES) BASS in licensed ranges that should not exceed 100 MW. Currently, in the Russian Federation there are no regulatory requirements for the technical parameters of RES BAS for their radio channels in unlicensed ranges, but there are restrictions on short-range devices using these ranges. At the same time, there are a number of regulatory documents of other countries that determine the signal levels in flight control channels. The possibilities of using UAS for radio monitoring purposes are considered using the example of the trial operation of the Orlan-10 unmanned aerial vehicle (UAV), taking into account existing restrictions in the field of the regulatory framework. It is shown that there are a number of restrictions for their effective use for radio monitoring purposes: in the field of the regulatory framework, technical characteristics of antennas and the dependence of the results on the flight altitude. Limitations due to climatic factors are also an important factor. Moreover, it is precisely for hard-to-reach areas with difficult climatic conditions that this monitoring method is the most promising.

UAV; UAS; allocated frequency ranges; radiation characteristics; radio monitoring.

Введение. Высокая интенсивность использования БПЛА и востребованность результатов их применения в практической деятельности привели к росту числа компаний, занимающихся проектированием и изготовлением БПЛА по всему миру. 15 августа 2023 г. подписан Проект стратегии развития отрасли связи Российской Федерации на период до 2035 г. [1]. В документе указывается, что цифровизация различных секторов экономики приводит к росту потребностей абонентов в области качества связи, в частности, минимизации задержки. В сетях мобильной связи для каналов со скоростью передачи данных 1–10 Гбит/с время подключения должно быть не более 1 мс. Одним из активно развивающихся направлений, для которых важны такие параметры соединения, является управление беспилотными летательными аппаратами в режиме реального времени.

Также в этом документе отмечается, что необходимо создание и внедрение гибридных сетей связи, в том числе с использованием БАС. Предполагается, что такие сети связи будут представлять совокупность наземного и спутникового сегментов связи, что позволит эффективно использовать их для управления БАС. Особенностью России является наличие обширных труднодоступных территорий, а том числе Арктической зоны, что ставит задачи по развитию высокоскоростной связи с низкими задержками, а также развитие средств контроля таких территорий. В этой роли могут выступить БПЛА с возможностью управления в режиме реального времени. Легитимность и безопасность полётов БПЛА должны обеспечиваться выполнением требований нормативных документов, регламентирующих условия их применения и определяющих значения параметров излучений радиоэлектронных средств (РЭС) беспилотных авиационных систем, размещаемых на беспилотных воздушных судах (БВС – другое название БПЛА) и на пульте управления (ПУ). В связи с этим анализ требований к частотным диапазонам использования БПЛА, а также к значениям параметров их излучения является актуальной задачей.

1. Радиоканалы БАС. Состав радиоканалов для обеспечения использования БПЛА гражданского назначения должен предусматривать наличие нескольких специализированных радиоканалов, предназначенных для:

- ♦ передачи команд телеуправления полётом БПЛА (ПУ–БПЛА), при этом радиопередатчик входит в состав пульта управления. В каналах управления ПУ–БПЛА применяется псевдослучайная перестройка рабочей частоты (ППРЧ);

- ♦ приёма сигналов телеметрии от БПЛА (БПЛА–ПУ), когда передатчик находится на БПЛА. В канале телеметрии в режиме ППРЧ передаётся информация об уровне принимаемого БВС сигнала, напряжения аккумулятора, потребляемого тока, данные о координатах, высоте, направлении и скорости движения и др.;

- ♦ приёма сигналов систем глобального позиционирования GPS, ГЛОНАСС, Galileo, BeiDou: на борту БПЛА находится приёмник сигналов систем геопозиционирования, а информация о географических координатах БПЛА передаётся на ПУ в канале телеметрии;

- ♦ передачи команд управления полезной нагрузкой: включение и выключение видеокамеры, управление фокусным расстоянием оптической системы, сброс нагрузки и др. В этом случае радиопередатчик также входит в состав пульта управления;

- ♦ передачи фотографий и потокового видео для систем с видеоаппаратурой. В этом случае передатчик находится на БПЛА.

В современных системах управления БПЛА канал телеуправления и канал передачи команд управления полезной нагрузкой объединены в одной полосе частот передатчика на ПУ.

2. Выделенные полосы частот для РЭС БАС. История выделения полос частот для РЭС БАС содержит как Решения ГКРЧ, так и Распоряжения Правительства Российской Федерации и Указы Президента.

Распоряжением Правительства Российской Федерации от 3 апреля 2018 года №576-р в целях интенсивного развития российскими компаниями сферы беспилотных авиационных систем утверждён план мероприятий – дорожная карта – национальной технологической инициативы по направлению «Аэронет» [2].

Распоряжением Правительства Российской Федерации 16 сентября 2021 года №2587-р в план мероприятий национальной технологической инициативы по направлению «Аэронет» внесены изменения по совершенствованию законодательства и устранению административных барьеров с поэтапной реализацией в период с 2021 по 2035 год.

5 октября 2021 года Распоряжением Правительства Российской Федерации № 2806-р утверждена Концепция интеграции беспилотных воздушных судов в единое воздушное пространство Российской Федерации [3] и План реализации Концепции в части развития технологий.

Основные направления развития технологий интеграции беспилотных воздушных судов включают, в том числе, разработку и сертификацию цифровых радиолиний связи, контроля и управления в полосах частот, утверждённых регламентом радиосвязи МСЭ и Таблицей распределения полос радиочастот между радиослужбами РФ. Наземное и бортовое оборудование цифровых радиолиний связи, контроля и управления подлежит обязательной сертификации, для чего необходима разработка соответствующих сертификационных требований.

В Концепции обозначено требование об оснащении всех беспилотных воздушных судов, выполняющих полёты в едином воздушном пространстве, бортовыми ответчиками, обеспечивающими передачу идентификационного номера воздушного судна, его текущих координат и информации о параметрах движения воздушного судна. Основа навигационного обеспечения полётов БВС в РФ – навигационная спутниковая система ГЛОНАСС и GPS, а в перспективе Galileo и BeiDou.

На первом этапе в 2022 г. планировалось проведение научных и экспериментальных исследований возможности применения технологий и сетей связи для низколетящих БВС, использования полос радиочастот для цифровых радиолиний связи, контроля и управления, а также каналов связи внешних пилотов и разработка требований к цифровым радиолиниям связи, контроля и управления в полосах частот, утверждённых Всемирной конференцией радиосвязи МСЭ.

На втором этапе в срок 2025 года планируется разработка и сертификация цифровых радиолиний (включая спутниковые) для контроля и управления, связи с органами обслуживания воздушного движения (управления полётами) и экипажами других воздушных судов в полосах частот, утверждённых Всемирной конференцией радиосвязи МСЭ.

На заседании от 29 ноября 2021 года Решением ГКРЧ №21-60-05 [4] впервые выделила полосу радиочастот 5,850–6,425 ГГц с целью использования РЭС для организации связи с БВС, осуществляющих воздушное наблюдение и патрулирование с целью контроля состояния удалённых промышленных объектов.

В 2022 г. Решением ГКРЧ №22-65-03 [5] утверждён состав рабочей группы по подготовке предложений по выделению полос радиочастот РЭС для организации радиосвязи БВС и условиям их использования. Рабочей группой рассматривалась возможность вынесения на заседание ГКРЧ вопроса о выделении 15 новых полос радиочастот для БАС в РФ, в том числе:

- ◆ для организации каналов управления и телеметрии 117,975–137 МГц; 390–399,9 МГц; 860–876 МГц; 902–915 МГц; 921–925 МГц;
- ◆ для организации каналов, обеспечивающих использование полезной нагрузки в БАС 1427–1535 МГц; 2214–2226 МГц; 2328–2340 МГц; 2400–2500 МГц и 2580–2592 МГц;
- ◆ для применения РЭС стандарта LTE и действующих разрешений на использование радиочастот для РЭС стандарта LTE, установленными на БАС, в рамках действующих решений ГКРЧ об использовании полос радиочастот 1710–1785 МГц и 1805–1880 МГц;
- ◆ для проведения исследований о возможности использования полосы радиочастот 960–1200 МГц для организации командной линии БАС;
- ◆ для проведения исследований о возможности использования полосы радиочастот 5–5,2 и 5,85–6,425 ГГц для организации канала полезной нагрузки в БАС.

Однако против такого частотного «изобилия» выступили представители Минцифры, Росавиации, Роскосмоса и операторов мобильной связи. Одной из причин разногласий стала необходимость проведения предварительных исследований в области электромагнитной совместимости между радиоэлектронными средствами БАС и другими видами РЭС, работающими в каждой из рассматриваемых полос радиочастот, в первую очередь, в полосах 960–1200 МГц, 5–5,2 ГГц и 5,85–6,425 ГГц.

На заседании ГКРЧ 24 апреля 2023 года один из вопросов был посвящён рассмотрению рекомендаций Международной организации гражданской авиации (ИКАО) и сектора радиосвязи Международного союза электросвязи (МСЭ-Р) об использовании для линий управления и связи наземных систем дистанционного пилотирования воздушных судов ИКАО полос радиочастот 117,975–137 МГц, 960–1164 МГц и 5030–5091 МГц. В Решении №23-66-01 [6] отмечено, что в РФ для воздушной подвижной службы (ВПС) распределена только полоса 117,975–137 МГц, а полосы 960–1164 МГц и 5030–5091 МГц не распределены.

Относительно полосы 5030–5091 МГц принято решение, что её использование воздушной подвижной и воздушной подвижной спутниковой службой в РФ ограничивается системами воздушной связи, стандартизированными на международном уровне с внесением соответствующего изменения в Таблицу распределения полос радиочастот между радиослужбами РФ [7].

Стандартизация на международном уровне в этой полосе радиочастот подразумевает в том числе возможность использования бортовыми и стационарными станциями воздушной подвижной службы РЭС для радиоканалов управления и телеметрии БАС при условии исключения создания помех воздушной подвижной спутниковой службе и системам обеспечения посадки воздушных судов.

На текущий момент вопрос о необходимом частотном обеспечении благоприятных условий для развития рынка БАС во исполнение требований п.1 д)-3 поручения Президента России от 30.12.2022 № Пр-2548 [8] полностью не решён. Работа комиссии по подготовке предложений по использованию дополнительных полос радиочастот для линий управления и телеметрии, а также для линий полезной нагрузки в БАС продолжается. Таким образом, официальный частотный ресурс для РЭС БАС в РФ сегодня: 117,975 – 137 МГц; 5030 – 5091 МГц; 5850 – 6425 МГц. Кроме того, решением ГКРЧ от 07.05.2007 года № 07-20-03-001 «О выделении полос радиочастот устройствам малого радиуса действия» (в ред. от 10.03.2017) [9] для РЭС БПЛА, попадающих под категорию устройств малого радиуса действия, определены уровни излучений, полосы радиочастот и условия их использования.

На практике для каналов управления, телеметрии и полезной нагрузки при-меняют до 12 полос радиочастот в различных диапазонах, в том числе и в нелицензированных в РФ полосах радиочастот [10]. В РФ используются следующие нелицензированные полосы радиочастот для РЭС беспроводных авиационных систем:

- ◆ 137–174 МГц;
- ◆ 395–400 МГц для канала управления;
- ◆ 430–460 МГц: 433,05–434,79 МГц для канала управления, в полосе 438–448 МГц – для канала телеметрии;
- ◆ 860–880 и 902–928 МГц используются для канала управления БПЛА производства Autel и для канала управления FPV БПЛА других производителей, а также для передачи видеоизображения в потоковом режиме;
- ◆ 1100–1400 МГц для каналов управления FPV БПЛА, а также для каналов передачи видеоизображения;
- ◆ 2400–2483,5 МГц — используется в звене БПЛА-ПУ в том числе для потокового видео, и ПУ-БПЛА для канала управления С2.

Полосы радиочастот цифровых каналов систем подвижной радиосвязи, таких как GSM, DCS, UMTS и LTE, для управления БАС в РФ также не лицензированы.

3. Требования к параметрам излучений РЭС БАС. На текущий момент в РФ нет нормативных требований к техническим параметрам РЭС БАС (кроме устройств малого радиуса действия [9]) для радиоканалов управления, телеметрии, управления полезной нагрузкой, излучаемых мощностей РЭС в каждом канале, вида модуляции и т.д.

В зарубежных странах уровни сигналов в каналах управления полётом БВС, телеметрии и его нагрузкой определяются на основе требований на соответствие национальным сертификатам, к основным из которых следует отнести:

- ◆ сертификат федеральной комиссии связи (FCC – Federal Communication Commission) [11];
- ◆ сертификат Евросоюза (CE) [12];
- ◆ сертификат КНР (SRRC или SRMC – State Radio Regulation Committee) [13] для сертификации всех беспроводных продуктов, продаваемых в Китае.

Уровни эффективной изотропно излучаемой мощности (ЭИИМ) сигналов в радиоканалах БАС гражданского назначения, соответствующие принятым в разных странах национальных системах сертификации, приведены в табл. 1.

Таблица 1

Уровни ЭИИМ в радиоканалах БАС гражданского назначения

		CE dBm (мВт)	FCC dBm (мВт)	SRRC dBm (мВт)	SRMC dBm (мВт)	Для РФ dBm (мВт)
ПУ	2,4 ГГц	20 (100)	26 (400)	20 (100)	20 (100)	20 (100)
	5,8 ГГц	14 (25)	26 (400)	23 (200)	23 (200)	20 (100)
БВС	2,4 ГГц	20 (100)	33 (2000)	20 (100)	20 (100)	20 (100)
	5,8 ГГц	д/н	33 (2000)	д/н	д/н	20 (100)

В каналах управления применяется псевдослучайная перестройка рабочей частоты со скоростью перестройки – 350...1000 ск/с; видами модуляции FSK2, PSK2 (A/B) и OFDM; длительностями импульсов 500...2500 мкс и шириной спектра импульса – 0,3–2,0 МГц. Скорость передачи данных в таких каналах до 2-х Мбод. По данным отчёта МСЭ-Р М.2171 [21] возможная ширина полосы ППРЧ от 34 до 200 МГц, при этом стандартное значение составляет 80 МГц.

Число каналов ППРЧ составляет от 32-х до 100, при этом 100 каналов возможно получить только для полосы 200 МГц. Одной из особенностей БПЛА гражданского назначения производства DJI является автоматическая установка мощности сигналов передатчиков в зависимости от территории, на которой он используется. Эта информация содержится в памяти БПЛА, а местоположение определяется по принятым координатам от системы геопозиционирования. Запас мощности передатчиков, заложенный разработчиками, позволяет повысить мощность передатчиков до 400 мВт, а с установкой дополнительного усилителя (так называемого бустера) – до 2,0 Вт. Такие технические возможности являются фактором, который может провоцировать нарушение законодательства РФ в условиях необходимости достижения целей проводимых мероприятий.

4. Использование БАС для осуществления мероприятий радиоконтроля.

13 декабря 2011 года состоялось заседание Правительственной комиссии по федеральной связи и технологическим вопросам информатизации, на котором обсуждался вопрос использования беспилотных летательных аппаратов для радиоконтроля (РК) в труднодоступных районах, зонах чрезвычайных ситуаций и на местности с плотной застройкой. На заседании Правительственной комиссии принято решение рекомендовать Роскомнадзору организовать опытную эксплуатацию комплексов радиоконтроля на базе беспилотных летательных аппаратов [14].

Опытная эксплуатация комплекса радиоконтроля (КРК) на основе БПЛА «Орлан-10» самолётного типа производства ООО «СТЦ» [15] в подразделениях радиоконтроля предприятий радиочастотной службы РФ проводилась в 2012-2013 годах на производственной базе отдела мобильного радиоконтроля ФГУП «РЧЦ СЗФО». Ранее данные аппараты уже использовались для дистанционного зондирования Земли и аэрофотосъемки [16]. Основное предназначение КРК на БПЛА – осуществление мероприятий радиоконтроля в районах ограниченной доступности для мобильных комплексов радиоконтроля и не входящих в зону электромагнитной доступности стационарных средств радиоконтроля. Технические возможности КРК на базе БПЛА обеспечивают:

- ◆ выявление реально работающих РЭС в конкретном районе с целью определения местоположения и идентификации источников радиоизлучений, эксплуатирующихся без разрешительных документов, в том числе базовых станций сотовой связи различных стандартов, размещённых в удаленных и труднодоступных для наземных мобильных комплексов радиоконтроля районах;
- ◆ определение параметров излучений стационарных РЭС по полю, в том числе РЭС теле- и радиовещания, радиорелейных станций, базовых станций;
- ◆ определение местоположения источников радиоизлучений по данным радиопеленгования в контролируемом диапазоне частот.

Сектором радиосвязи международного союза электросвязи (МСЭ-R) в 2021 году подготовлен отчёт SM.2486-0 (06/21) «Применение коммерческих БПЛА для выполнения задач МСЭ-R по контролю за использованием спектра» [22]. В отчёте рассмотрен состав функциональных элементов системы радиоконтроля на основе коммерческого БПЛА, аспекты, касающиеся вопросов измерений и требования при работе с БПЛА, а также варианты их использования для измерения напряжённости поля радиосигналов и определения местоположения РЭС. Также оценена возможность применения БПЛА для решения задач, связанных с измерениями параметров сигналов, в том числе: измерение напряжённости электромагнитного поля, измерение зоны радиопокрытия, радиопеленгации источников радиоизлучений и др.

Однако опыт использования КРК на БПЛА показал как определенные достоинства данной технологии радиоконтроля, так и выявил ее существенные ограничения, приведенные в табл. 2.

Таблица 2

Результаты опытной эксплуатации БПЛА в интересах радиоконтроля

Возможности	Ограничения
Возможность проведения мероприятий РК в районах, где не обеспечивается транспортная доступность мобильных комплексов РК	Ограничения на полеты в районах, где расположены объекты Министерства обороны и атомные электростанции, в приграничных районах, над городами и пр. Ограничения по климатическим факторам внешней среды (дождь, снегопад, обледенение корпуса БПЛА)
Выявление работающих РЭС в контролируемом частотном диапазоне на большой территории в сжатые сроки	Антенны для пеленгования имеют ограничения по полосе радиочастот.
Повышение надежности КРК за счет резервирования комплекса двумя БПЛА, увеличение дальности при использовании одного БПЛА в качестве ретранслятора	Применение второго БПЛА удорожает КРК. В режиме ретранслятора передача результатов радиоконтроля на наземный ПУ в реальном масштабе времени не осуществляется, передаются только команды управления полетом. Вся информация о результатах РК записывается в бортовую память носителя
Выявление работающих РЭС в конкретном районе с целью определения местоположения и идентификации источников радиоизлучений, эксплуатирующихся без разрешительных документов, размещённых в удаленных, и труднодоступных для наземных мобильных комплексов радиоконтроля районах	Определение точного местоположения РЭС, действующих без разрешительных документов и не внесенных в БД ЧП требует применения мобильных и носимых средств пеленгования
Оценка уровней излучений РЭС	Оценка напряженности поля от РЭС на высоте полета не позволяет использовать эти данные для определения зон покрытия, где требуется проведение измерений на высотах 1,5 и 10 м [17, 18]
Наличие в комплекте поставки транспортного средства позволяет разместить в аппаратном отсеке автомобиля рабочие места операторов управления БПЛА и КРК на базе соответствующих АРМ	Высокая стоимость комплекта из 2-х КРК, транспортного средства и средств обеспечения. Большие эксплуатационные расходы на содержание

По результатам опытной эксплуатации КРК на базе БПЛА самолётного типа помимо приведённых в табл. 2 ограничений его применения также отмечена сложная процедура получения разрешения на полёты БПЛА и высокая себестоимость мероприятий РК с использованием КРК на БПЛА самолётного типа.

Отметим, что разработки в области совершенствования комплексов радиоконтроля для узких задач мониторинга, например, мониторинга сетей мобильной связи при условии разработки оптимального маршрута пролета БПЛА с учетом размещения базовых станций оператора [19], или с использованием БПЛА вертолётного типа [20], имеющих чаще всего относительно невысокую стоимость, по-прежнему ведутся. При этом возникает необходимость решения ряда оптимизационных задач, связанных с уменьшением стоимости мероприятий РК и выбора параметров КРК.

Заключение. Необходимость использования БАС различного назначения не вызывает сомнений. Однако проведённый анализ частотного обеспечения радиотехнического оборудования БАС показывает, что существуют проблемы с использованием такими системами радиочастотного спектра в РФ. Так, официально выделены для БАС частотные диапазоны 117,975–137 МГц, 5030–5091 МГц, 5850–6425 МГц, но можно использовать и нелегитимные полосы частот, на которые есть существенные ограничения на уровень излучения РЭС, что может сделать эксплуатацию некоторых БПЛА нелегитимным. При этом частотные диапазоны систем мобильной сотовой связи для управления БАС в РФ также не лицензированы, и есть необходимость проведения исследований в области электромагнитной совместимости радиосистем БАС и систем мобильной связи.

В нормативно-правовой базе РФ определены требования к уровням излучений радиотехнического оборудования БАС, которые не должны превышать 20 дБм (100 мВт), а также требования к модуляционным параметрам радиосигналов БАС.

Анализ опыта эксплуатации комплекса радиоконтроля для осуществления мероприятий радиоконтроля в районах ограниченной доступности для мобильных комплексов радиоконтроля и не входящих в зону электромагнитной доступности стационарных средств радиоконтроля показывает, что существует ряд ограничений для использования БПЛА. Так, выявлено, что кроме сложностей с получением разрешения на полеты, существует ряд ограничений по климатическим факторам внешней среды, затрудняющим эксплуатацию БПЛА в условиях Крайнего Севера. Также отмечены сложности с определением местоположения БПЛА и необходимость проведения дополнительных измерений другими методами для определения зон покрытия. Эти проблемы могут стать направлениями дальнейших исследований и новых инженерных решений в процессе совершенствования беспилотных систем в качестве средств мониторинга.

Исследование выполнено в рамках выполнения прикладных научных исследований по государственному заданию СПбГУТ на 2023 год, регистрационный номер 123060900012-6 в ЕГИСУ НИОКТР.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Проект стратегии развития отрасли связи Российской Федерации на период до 2035 года. – Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации. – URL: <https://digital.gov.ru/ru/documents/9120/> (дата обращения: 01.08.2023).
2. Об утверждении плана мероприятий («дорожной карты») по совершенствованию законодательства и устранению административных барьеров в целях обеспечения реализации плана мероприятий («дорожной карты») Национальной технологической инициативы по направлению «Аэронет»: Распоряжение Правительства РФ от 03.04.2018 г. N576-р (в ред. от 16.09.2021). Доступ из справ.-правовой системы «КонсультантПлюс».

3. Об утверждении Концепции интеграции беспилотных воздушных судов в единое воздушное пространство Российской Федерации и плана реализации Концепции в части развития технологий: Распоряжение Правительства РФ от 05.10.2021 г. N2806-р (в ред. от 05.12.2022). Доступ из справ.-правовой системы «КонсультантПлюс».
4. Решение ГКРЧ от 29 ноября 2021 года №21-60-05 «Об использовании полосы радиочастот 5850-6425 МГц радиоэлектронными средствами для организации связи с беспилотными воздушными судами (для служебного пользования). – Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации. – URL: <https://digital.gov.ru/ru/documents/8069/> (дата обращения: 23.08.2023).
5. Решение ГКРЧ от 23 декабря 2022 года №22-65-03 «Об использовании полос радиочастот радиоэлектронными средствами для организации связи в беспилотных авиационных системах». – Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации. – URL: <https://digital.gov.ru/ru/documents/8628/> (дата обращения: 01.08.2023).
6. Решение ГКРЧ от 24 апреля 2023 года №23-66-01 «Об использовании полос радиочастот радиоэлектронными средствами для организации связи в беспилотных авиационных системах гражданского назначения». – Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации. – URL: <https://digital.gov.ru/ru/documents/8969/> (дата обращения: 01.08.2023).
7. Постановление Правительства РФ от 18.09.2019 г. N1203-47 «Об утверждении Таблицы распределения полос радиочастот между радиослужбами Российской Федерации и признании утратившими силу некоторых постановлений Правительства Российской Федерации» (с изменениями и дополнениями). Доступ из справ.-правовой системы «Гарант».
8. Перечень поручений по вопросам развития беспилотных авиационных систем: Поручение Президента России от 30.12.2022 г. Пр-2548-р. Доступ из справ.-правовой системы «КонсультантПлюс».
9. Решение ГКРЧ от 07 мая 2007 года № 07-20-03-001 «О выделении полос радиочастот устройствам малого радиуса действия» (в ред. от 10.03.2017). – Электронный фонд нормативно-технической и нормативно-правовой информации Консорциума «Кодекс». – URL: <https://docs.cntd.ru/document/420342446> (дата обращения: 01.08.2023).
10. Макаренко С.И. Противодействие беспилотным летательным аппаратам: монография. – СПб.: Научное издание, 2020. – 204 с.
11. FCC Certification // NexPCB Knowledge Base. – URL: knowledgebase.nexpcb.com/articles/fcc/ (дата обращения: 01.08.2023).
12. Европейский сертификат качества – CE маркировка для Евросоюза. – Серконс. – URL: <https://www.serconsrus.ru/services/ce-marking/> (дата обращения: 01.08.2023).
13. SRRC Certification for telecommunication products in China. – MPR. – URL: <https://www.china-certification.com/en/china-srrc-type-approval/> (дата обращения: 01.08.2023).
14. Для радиоконтроля на труднодоступных территориях будут применяться беспилотные летательные аппараты. – Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации. – URL: <https://digital.gov.ru/ru/events/29464/> (дата обращения: 01.08.2023).
15. Захаров А.И., Иванов Н.В., Студеникин А.В. Возможности беспилотных летательных аппаратов при решении изыскательских задач // Инженерные изыскания. – 2012. – № 7. – С. 20-23.
16. Студеникин А.В. и др. Практика применения перспективных беспилотных летательных аппаратов для мониторинга и аэрофотосъемки // Современные проблемы дистанционного зондирования Земли из космоса. – 2012. – Т. 9, № 4. – С. 102.
17. Рекомендация МСЭ-R SM.378-7 «Измерение напряженности поля на станциях радиоконтроля». – Международный союз электросвязи. – URL: https://www.itu.int/dms_pubrec/itu-r/rec/sm/R-REC-SM.378-7-200702-I!!PDF-E.pdf (дата обращения: 01.08.2023).
18. Рекомендация МСЭ-R SM.1875-1 «Измерение покрытия DVB-T и проверка критериев планирования». – Международный союз электросвязи. – https://www.itu.int/dms_pubrec/itu-r/rec/sm/R-REC-SM.1875-1-201308-S!!PDF-R.pdf (дата обращения: 01.08.2023).

19. *Никитина А.В.* Оценка возможности применения беспилотных летательных аппаратов для задач радиоконтроля сетей мобильной связи // СПБНТОРЭС: Тр. ежегодной НТК. – 2021. – № 1. – С. 240-243.
20. *Ашихмин А.В. и др.* Способ однопозиционного местоопределения источников радиоизлучения с использованием бортового радиопеленгатора беспилотного летательного аппарата вертолетного типа // Системы управления, связи и безопасности. – 2021. – № 4. – С. 40-57.
21. Report ITU-R M.2171 «Characteristics of unmanned aircraft systems and spectrum requirements to support their safe operation in non-segregated airspace». – МСЭ. – URL: <https://www.itu.int/pub/R-REP-M.2171-2009> (дата обращения: 01.08.2023).
22. Report SM.2486-0 «Use of commercial drones for ITU-R spectrum monitoring tasks». – МСЭ. – URL: <https://www.itu.int/pub/R-REP-SM.2486-2021> (дата обращения: 01.08.2023).

REFERENCES

1. Proekt strategii razvitiya otrasli svyazi Rossiyskoy Federatsii na period do 2035 goda [Draft strategy for the development of the communications industry of the Russian Federation for the period up to 2035]. Ministry of Digital Development, Communications and Mass Communications of the Russian Federation (2023). Available at: <https://digital.gov.ru/ru/documents/9120/> (accessed 01 August 2023).
2. Ob utverzhdenii plana meropriyatiy («dorozhnoy karty») po sovershenstvovaniyu zakonodatel'stva i ustraneniyu administrativnykh bar'erov v tselyakh obespecheniya realizatsii plana meropriyatiy («dorozhnoy karty») Natsional'noy tekhnologicheskoy initsiativy po napravleniyu «Aeronet»: Rasporyazhenie Pravitel'stva RF ot 03.04.2018 g. N576-r (v red. ot 16.09.2021) [On the approval of the action plan («roadmap») to improve legislation and eliminate administrative barriers in order to ensure the implementation of the action plan («roadmap») The National Technological Initiative in the Aeronet direction: Order of the Government of the Russian Federation N576-p (2021)]. Available at: legal reference system «Konsul'tantPlyus».
3. Ob utverzhdenii Kontseptsii integratsii bespilotnykh vozdushnykh sudov v edinoe vozdushnoe prostranstvo Rossiyskoy Federatsii i plana realizatsii Kontseptsii v chasti razvitiya tekhnologii: Rasporyazhenie Pravitel'stva RF ot 05.10.2021 g. N2806-r (v red. ot 05.12.2022) [On approval of the Concept of Integration of Unmanned Aircraft into the Unified Airspace of the Russian Federation and the plan for the implementation of the Concept in terms of technology development: Order of the Government of the Russian Federation N2806-p (2021)]. Available at: legal reference system «Konsul'tantPlyus».
4. Reshenie GKRCCh ot 29 noyabrya 2021 goda №21-60-05 «Ob ispol'zovanii polosy radiochastot 5850-6425 MGts radioelektronnymi sredstvami dlya organizatsii svyazi s bespilotnymi vozdushnymi sudami (dlya sluzhebnogo pol'zovaniya) [Decision GRFC No. 21-60-05 (2021) “About the use of the radio frequency band 5850-6425 MHz by radio-electronic means for the organization of communication with unmanned aircraft (for official use)”. Ministry of Digital Development, Communications and Mass Communications of the Russian Federation. Available at: <https://digital.gov.ru/ru/documents/8069/> (accessed 01 August 2023).
5. Reshenie GKRCCh ot 23 dekabrya 2022 goda № 22-65-03 «Ob ispol'zovanii polos radiochastot radioelektronnymi sredstvami dlya organizatsii svyazi v bespilotnykh aviatsionnykh sistemakh» [Decision GRFC No. 22-65-03 (2022) “On the use of radio frequency bands by radio-electronic means for the organization of communication in unmanned aircraft systems”]. Ministry of Digital Development, Communications and Mass Communications of the Russian Federation. Available at: <https://digital.gov.ru/ru/documents/8628/> (accessed 01 August 2023).
6. Reshenie GKRCCh ot 24 aprelya 2023 goda №23-66-01 «Ob ispol'zovanii polos radiochastot radioelektronnymi sredstvami dlya organizatsii svyazi v bespilotnykh aviatsionnykh sistemakh grazhdanskogo naznacheniya» [Decision GRFC No. 23-66-01 (2023) “On the use of radio frequency bands by radio-electronic means for the organization of communication in unmanned civil aviation systems”. Ministry of Digital Development, Communications and Mass Communications of the Russian Federation. Available at: <https://digital.gov.ru/ru/documents/8969/> (accessed 01 August 2023).
7. Postanovlenie Pravitel'stva RF ot 18.09.2019 g. N1203-47 «Ob utverzhdenii Tablitsy raspredeleniya polos radiochastot mezhduradiosluzhbnami Rossiyskoy Federatsii i priznaniyu utrativshimi silu nekotorykh postanovleniy Pravitel'stva Rossiyskoy Federatsii» (s

- izmeneniyami i dopolneniyami) [Resolution of the Government of the Russian Federation N1203-47 (2019) “On the approval of the Table of Distribution of Radio Frequency Bands between the Radio Services of the Russian Federation and the invalidation of Certain resolutions of the Government of the Russian Federation (with amendments and additions)”. Available at: legal reference system «Garant».
8. Perechen' porucheniy po voprosam razvitiya bespilotnykh aviatsionnykh sistem: Poruchenie Prezidenta Rossii ot 30.12.2022 g. Pr-2548-r [List of instructions on the development of unmanned aircraft systems: Order of the President of Russia Пр-2548-p. (2022)]. Available at: legal reference system «KonsultantPlyus».
 9. Reshenie GKRCCh ot 07 maya 2007 goda № 07-20-03-001 «O vydelenii polos radiochastot ustroystvam malogo radiusa deystviya» (v red. ot 10.03.2017) [Decision GRFC № 07-20-03-001 (2007) “On the allocation of radio frequency bands to short-range devices”]. Electronic fund of normative, technical and regulatory information of the Code Consortium. Available at: <https://docs.cntd.ru/document/420342446> (accessed 01 August 2023).
 10. *Makarenko S.I.* Protivodeystvie bespilotnym letatel'nykh apparatam: monografiya [Counteraction by unmanned aerial vehicle: monograph]. Saint Petersburg: Naukoemkie tekhnologii, 2020, 204 p.
 11. FCC Certification, *NexPCB Knowledge Base*. Available at: knowledgebase.nexpcb.com/articles/fcc/ (accessed 01 August 2023).
 12. Evropeyskiy sertifikat kachestva – SE markirovka dlya Evrosoyuza [European Quality Certificate – CE marking for the European Union]. Serkons. Available at: <https://www.serconsrus.ru/services/ce-marking/> (accessed 01 August 2023).
 13. SRRC Certification for telecommunication products in China. MPR. Available at: <https://www.china-certification.com/en/china-srrc-type-approval/> (accessed 01 August 2023).
 14. Dlya radiokontrolya na trudnodostupnykh territoriyakh budut primenyat'sya bespilotnye letatel'nye apparaty [Unmanned aerial vehicles will be used for radio monitoring in remote areas]. Ministry of Digital Development, Communications and Mass Communications of the Russian Federation. Available at: <https://digital.gov.ru/ru/events/29464/> (accessed 01 August 2023).
 15. *Zakharov A.I., Ivanov N.V., Studenikin A.V.* Vozможности bespilotnykh letatel'nykh apparatov pri reshenii izyskatel'skikh zadach [The capabilities of unmanned aerial vehicles in solving survey tasks], *Inzhenernye izyskaniya* [Engineering surveys], 2012, No. 7, pp. 20-23.
 16. *Studenikin A.V. i dr.* Praktika primeneniya perspektivnykh bespilotnykh letatel'nykh apparatov dlya monitoringa i aerofotos"emki [The practice of using advanced unmanned aerial vehicles for monitoring and aerial photography], *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa* [Modern problems of remote sensing of the Earth from space], 2012, Vol. 9, No. 4, pp. 102.
 17. Rekomendatsiya MSE-R SM.378-7 «Izmerenie napryazhennosti polya na stantsiyakh radiokontrolya» [Recommendation ITU-R SM.378-7 “Field strength measurement at radio monitoring stations”]. International Telecommunication Union. Available at: https://www.itu.int/dms_pubrec/itu-r/rec/sm/R-REC-SM.378-7-200702-I!!PDF-E.pdf (accessed 01 August 2023).
 18. Rekomendatsiya MSE-R SM.1875-1 «Izmerenie pokrytiya DVB-T i proverka kriteriev planirovaniya» [Recommendation ITU-R SM.1875-1 “DVB-T coverage measurement and verification of planning criteria”]. International Telecommunication Union. Available at: https://www.itu.int/dms_pubrec/itu-r/rec/sm/R-REC-SM.1875-1-201308-S!!PDF-R.pdf (accessed 01 August 2023).
 19. *Nikitina A.V.* Otsenka vozможности primeneniya bespilotnykh letatel'nykh apparatov dlya zadach radiokontrolya setey mobil'noy svyazi [Assessment of the possibility of using unmanned aerial vehicles for the tasks of radio monitoring of mobile communication networks], *SPbNTORES: Tr. ezhegodnoy NTK* [SPbNTORES: Proceedings of the annual STC], 2021, No. 1, pp. 240-243.
 20. *Ashikhmin A.V. i dr.* Spособ odnopolozitsionnogo mestoopredeleniya istochnikov radioizlucheniya s ispol'zovaniem bortovogo radiopelengatora bespilotnogo letatel'nogo apparata vertoletnogo tipa [A method for single-position location of radio emission sources using an on-board radio direction finder of a helicopter-type unmanned aerial vehicle], *Sistemy upravleniya, svyazi i bezopasnosti* [Control, communication and security systems], 2021, No. 4, pp. 40-57.

23. Report ITU-R M.2171 “Characteristics of unmanned aircraft systems and spectrum requirements to support their safe operation in non-segregated airspace”. International Telecommunication Union. Available at: <https://www.itu.int/pub/R-REP-M.2171-2009> (accessed 01 August 2023).
21. Report SM.2486-0 “Use of commercial drones for ITU-R spectrum monitoring tasks”. International Telecommunication Union. Available at: <https://www.itu.int/pub/R-REP-SM.2486-2021> (accessed 01 August 2023).

Статью рекомендовал к опубликованию д.т.н., профессор К.Е. Румянцев.

Антипин Борис Маврович – Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича; e-mail: boris.antipin@sut.ru; г. Санкт-Петербург, Россия; к.т.н.; доцент кафедры телевидения и метрологии.

Виноградов Евгений Михайлович – e-mail: vinogradov_em@sut.ru; к.т.н.; доцент; старший научный сотрудник НИЛ «Радиоконтроля и ЭМС».

Симонина Ольга Александровна – e-mail: olga.simonina@sut.ru; к.т.н.; доцент кафедры радиосвязи и вещания.

Федоров Сергей Леонидович – e-mail: sergey.fedorov@sut.ru; к.т.н.; доцент; зав. кафедрой телевидения и метрологии.

Antipin Boris Mavrovich – Federal State Budget-Financed Educational Institution of Higher Education the Bonch-Bruevich Saint Petersburg State University of Telecommunications; e-mail: boris.antipin@sut.ru; Saint Petersburg, Russia; cand. of eng. sc.; associate professor of the department of television and metrology.

Vinogradov Evgeniy Mikhaylovich – e-mail: vinogradov_em@sut.ru; cand. of eng. sc.; associate professor senior researcher at the research laboratory «Radio control and electromagnetic compatibility».

Simonina Olga Alexandrovna – e-mail: olga.simonina@sut.ru; cand. of eng. sc.; associate professor of the department of radio communication and broadcasting.

Fedorov Sergey Leonidovich – e-mail: sergey.fedorov@sut.ru; cand. of eng. sc.; associate professor; head of the department of television and metrology.

УДК 621.341.572

DOI 10.18522/2311-3103-2023-5-204-213

В.В. Буллака, С.В. Гулаков, А.Ю. Головин, Д.С. Мироненко

УНИВЕРСАЛЬНЫЙ ДВУНАПРАВЛЕННЫЙ DC-AC ПРЕОБРАЗОВАТЕЛЬ

Представлено схемное решение и описание работы двунаправленного DC-AC преобразователя с трансформаторной изоляцией. В основе схемы лежат принципы, применяемые в резонансных LLC преобразователях и двойных активных мостах (DAB, Dual Active Bridge), но в предлагаемой конструкции используется биполярное построение инвертора со стороны переменного напряжения, а также совмещенное частотное и широтно-импульсное управление. Это дает возможность реализовать ряд функций в одном преобразователе, используя небольшое число коммутаций в силовой цепи. Так, возможна работа в режиме источника бесперебойного питания (преобразование постоянного напряжения АКБ в переменное 220 В 50 Гц); инвертора связи с сетью альтернативных источников (солнечных панелей); зарядного устройства (работающего как от сети, так и от солнечных панелей); сварочного инвертора для полуавтоматической сварки (причем сварку можно осуществлять как от сети, так и от АКБ); стабилизатора переменного напряжения; преобразователя постоянного напряжения (для «прикуривания» АКБ с разными напряжениями или степенью заряженности). Устройство с перечисленным набором функций может найти применение, например, в полевых условиях – для организации электропитания различных нагрузок, зарядки аккумуляторов, а также проведения мелких сварочных работ.

Схемное решение преобразователя позволяет при применении соответствующих алгоритмов управления работать с близким к единице коэффициентом мощности на стороне переменного напряжения. Это дает возможность обеспечить работу в «слабой» сети – от автономного генератора, при большой длине сетевого шнура и т.д.

Резонансный преобразователь; двойной активный мост; источник бесперебойного питания; сварочный инвертор; зарядное устройство; двунаправленный преобразователь.

V.V. Burlaka, S.V. Gulakov, A.Y. Golovin, D.S. Mironenko

UNIVERSAL BIDIRECTIONAL DC-AC CONVERTER

The paper presents a schematic solution and a description of the operation of a bidirectional DC-AC converter with transformer isolation. The topology is based on the principles used in resonant LLC converters and dual active bridges (DAB), but the proposed design uses bipolar inverter topology on the AC voltage side, as well as combined frequency and pulse width control. This makes it possible to implement a number of functions in a single converter using a small number of commutations in the power circuit. So, it is possible to implement the following operating modes: an uninterruptible power supply (conversion of a battery DC voltage to a 220 V 50 Hz AC); a grid-tie inverter for alternative sources (solar panels); a battery charger (operating both from the mains and from solar panels); a welding inverter for semi-automatic welding (and welding can be carried out both from mains, and from the battery); an AC voltage regulator; a DC/DC voltage converter (for “cranking up” batteries with different voltages or state-of-charge). A device with the listed set of functions can be used, for example, in the field – for organizing power supply of various loads, charging batteries, as well as performing minor welding operations. The circuit topology of the converter allows, when using appropriate control algorithms, to work with a power factor close to unity on the AC voltage side. This makes it possible to ensure operation in a “weak” grid – from an autonomous generator, with a large length of the power cord, etc.

Resonant converter; dual active bridge; uninterruptible power supply; welding inverter; battery charger; bidirectional converter.

Введение. В полевых условиях (сельскохозяйственные работы в поле, археологические и другие экспедиции, работа вдали от стационарных источников электрической энергии и т.п.), а также в аварийных ситуациях питание электрических потребителей осуществляется, как правило, от мобильных бензиновых или дизель-генераторов, источников бесперебойного питания (ИБП) с аккумуляторными батареями (АКБ), солнечных панелей, укомплектованных инвертором. Подключение к таким источникам энергии мощных электроприёмников (электроинструмент, сварочное оборудование), которые характеризуются резкими изменениями мощности во время работы, приводит к нарушению их работы.

К недостаткам генератора можно отнести производимый им акустический шум, а также необходимость пополнять запас топлива. ИБП имеет значительно меньший уровень шума, но требует периодической подзарядки АКБ. К тому же зачастую ИБП ограничен по мощности, что затрудняет обеспечение питания относительно мощных электроприёмников. Между тем, при решении задач ремонта, восстановления военной, сельскохозяйственной, другой техники в полевых условиях часто возникает необходимость проведения сварочных работ. При этом далеко не всегда есть возможность обеспечить качественное электрическое питание сварочного оборудования. Известные же технические решения реализации сварочного процесса от АКБ [1–3] отличаются высокой стоимостью и необходимостью применения специализированных АКБ с соответствующим зарядным устройством.

Цель исследования. Разработка и создание универсального DC-AC преобразователя, который позволит работать в различных режимах:

- ◆ ИБП – преобразование напряжения АКБ в 220 В 50 Гц;

- ◆ зарядка АКБ от сети или солнечных панелей; сварка от сети. При этом в режиме работы от сети переменного напряжения должна обеспечиваться гальваническая изоляция выходных низковольтных цепей;
- ◆ инвертор связи солнечных панелей с сетью;
- ◆ стабилизатор переменного напряжения;
- ◆ DC/DC режим – сварка от АКБ; «прикуривание».

Универсальный двунаправленный DC-AC преобразователь. Для реализации цели разработан универсальный преобразователь, с помощью которого можно реализовать сварочные работы с питанием как от электрической сети, так и от бортовых аккумуляторов техники; обеспечить питание электрических нагрузок переменным напряжением 220 В от аккумуляторов; выполнять функции пускозарядного устройства. Продолжительность работ при этом не ограничивается емкостью аккумуляторов – достаточно запустить двигатель, чтобы заработал генератор, компенсирующий их разряд. При этом в режиме работы от сети переменного напряжения обеспечивается гальваническая изоляция выходных низковольтных цепей. Известные решения двунаправленных DC/AC преобразователей [4–6] либо не имеют гальванической развязки, либо содержат два каскада преобразования с промежуточным звеном постоянного тока.

Функция ИБП (формирование напряжения 220 В 50 Гц от АКБ) предлагаемого преобразователя позволит обеспечить бесперебойное питание электроприемников, используя в качестве первичных источников энергии бортовые АКБ техники или другие автономные источники энергии (например, солнечные панели) [7, 8].

Универсальный преобразователь может выполнять функцию пуско-зарядного устройства, что позволит производить зарядку АКБ и облегчать запуск двигателей внутреннего сгорания техники. Причем зарядка АКБ может осуществляться как от переменного напряжения 220 В 50 Гц, так и от другой АКБ (т.е. в режиме «прикуривания», но с контролем тока заряда и напряжения конца заряда). За счет применения соответствующих алгоритмических решений в системе управления предлагаемый универсальный преобразователь может обеспечить близкий к единице коэффициент мощности со стороны входа переменного напряжения, что позволит эффективно работать в слабой электрической сети или от автономного генератора небольшой мощности. Схема силовой части преобразователя представляет собой доработанный вариант решения, опубликованного в работах [9, 10], и представлена на рис. 1.

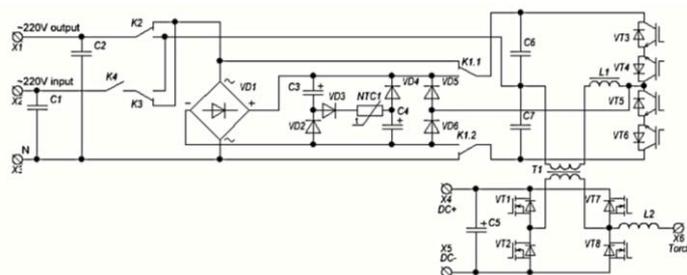


Рис. 1. Схема силовой части разработанного преобразователя

Устройство имеет вход и выход переменного напряжения 220 В с неразрывной нейтралью (клеммы X1 – X3), разъем для подключения АКБ (клеммы X4, X5) и разъем для подключения сварочной горелки в режиме сварки от АКБ (клемма X6).

Для выбора требуемого режима работы преобразователя осуществляют коммутацию элементов схемы с помощью реле K1 – K4, обеспечивающую нужную ее конфигурацию. Перечень режимов приведен в табл. 1.

Таблица 1

Состояние реле				Режим работы
K1	K2	K3	K4	
0	0	0	0	ИБП – преобразование напряжения АКБ в 220 В 50 Гц
1	0	0	1	Зарядка АКБ от сети или солнечных панелей; сварка от сети
0	0	0	1	Инвертор связи солнечных панелей с сетью
0	1	0	1	Стабилизатор переменного напряжения, понижение
0	0	1	1	Стабилизатор переменного напряжения, повышение
0	1	0	0	DC/DC режим – сварка от АКБ; «прикуривание»

Режимы работы устройства. Проанализируем перечисленные режимы работы устройства с приведением соответствующих схемных конфигураций силовой части в соответствии с состояниями реле K1 – K4.

Схема силовой части преобразователя в режиме ИБП приведена на рис. 2.

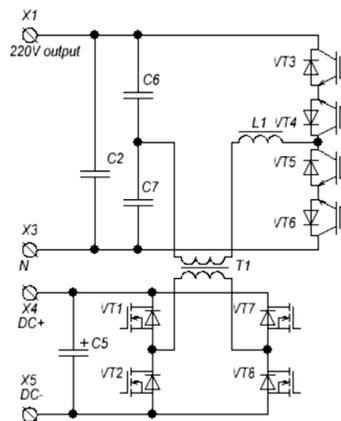


Рис. 2. Силовая часть в режиме ИБП

Следует сразу отметить, что такая же конфигурация используется и в режиме инвертора связи солнечных панелей с сетью – они подключаются к X4, X5 вместо АКБ, а вход 220 В (X2) соединяется с выходом 220 В (X1) за счет включения реле K4.

Напряжение АКБ поступает на мостовой преобразователь VT1VT2VT7VT8 и преобразуется в высокочастотное напряжение прямоугольной формы с изменяемым коэффициентом заполнения, которое подается на низковольтную обмотку силового высокочастотного трансформатора T1.

Высоковольтная часть преобразователя состоит из двуполярного полумоста VT3VT4VT5VT6, нагруженного на последовательно соединенные высоковольтную обмотку T1 и резонансный контур LC6C7. Конденсатор C2 является блокировочным по высокой частоте, снижает уровень высокочастотной составляющей в выходном напряжении и практически устраняет влияние нагрузки на резонансную частоту контура LC6C7 ($C2 \gg C6+C7$). Полумост VT3VT4VT5VT6 работает с фиксированным коэффициентом заполнения 50% и синхронизирован с низковольтным мостом VT1VT2VT7VT8. Эквивалентная схема замещения по высокой частоте может быть представлена так, как показано на рис. 3.

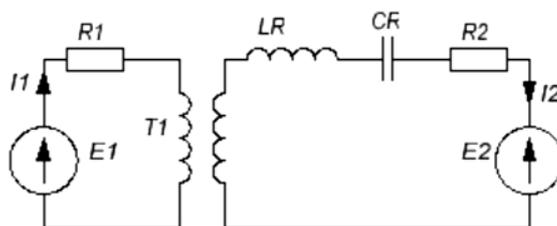


Рис. 3. Схема замещения силовой части преобразователя в режиме ИБП

Здесь E_1 – выходное напряжение моста VT1VT2VT7VT8, E_2 – выходное напряжение полумоста VT3VT4VT5VT6, LR – резонансная индуктивность (L_1 и индуктивность рассеяния T_1), CR – резонансная емкость ($C_6 + C_7$, учитывая, что $C_2 \gg C_6$ и $C_2 \gg C_7$), R_1, R_2 – суммарные активные сопротивления элементов. Формы напряжений на резонансной цепи показаны на рис. 4.

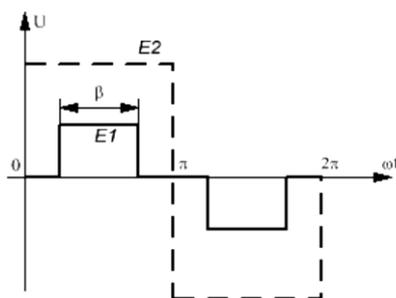


Рис. 4. Формы напряжений при работе в режиме ИБП

Для оценки поведения системы воспользуемся расчетом по основной гармонике, считая, что частота переключения равна резонансной частоте контура LRCR, а добротность контура велика, так что гармониками частоты переключения можно пренебречь.

Амплитуда основной гармоники сигнала E_2 составляет

$$E_{2m} = U_{C2} \cdot \frac{2}{\pi}$$

и зависит только от мгновенного напряжения на C_2 . Амплитуда основной гармоники E_1 зависит от напряжения АКБ U_{DC} и от коэффициента заполнения, представленного на рис. 4 углом β :

$$E_{1m} = U_{DC} \cdot \frac{4}{\pi} \cdot \sin \frac{\beta}{2}.$$

Фазы основной гармоники E_1 и E_2 совпадают – это неперемное условие корректной работы преобразователя.

Изменяя угол β за счет широтно-импульсного управления мостом VT1VT2VT7VT8, можно управлять амплитудой E_{1m} . Если пренебречь током намагничивания T_1 (что не внесет существенных погрешностей, поскольку он составляет 1 – 3 % рабочего тока), то, учитывая работу на резонансной частоте (полное сопротивление контура LRCR близко к 0), можно утверждать, что источники E_1 и E_2 соединены параллельно посредством T_1 и токи I_1 и I_2

совпадают по фазе с основными гармониками E1 и E2. Амплитуды же токов будут определяться разностью E1 и E2 с учетом коэффициента трансформации K_{T1} трансформатора T1.

Как было сказано выше, амплитуда E2 связана с напряжением на C2. Если $K_{T1} \cdot E_{1m} > E_{2m}$, то E1 будет работать с положительной мощностью (т.е. источником), а E2 – с отрицательной (т.е. приемником). Это приведет к росту напряжения на C2. И наоборот. Это дает возможность, изменяя угол β , получить на C2 переменное напряжение необходимой – синусоидальной формы. Для изменения знака напряжения на C2 меняется знак E1.

Таким образом, в режиме ИБП для получения выходного напряжения амплитудой U_M и угловой частотой Ω система управления изменяет угол β следующим образом:

$$\beta(t) = 2 \sin^{-1} \frac{|U_M \sin(\Omega t)|}{U_{DC} \cdot 2K_{T1}}$$

Для стабилизации выходного напряжения при изменении нагрузки имеется обратная связь по выходному напряжению и программная коррекция величины U_M .

Конфигурация силовой части преобразователя позволяет также обеспечить режим синхронной работы с сетью, т.е. выполнять функции инвертора связи с сетью альтернативных источников электроэнергии, например, солнечных панелей [11–16]. При этом меняется алгоритм управления – инвертор работает с обратной связью по мгновенному току сети и имитирует по отношению к сети отрицательное активное сопротивление. Это позволяет поддерживать высокий коэффициент мощности, близкую к нулю реактивную мощность и низкий уровень высших гармоник сетевого тока [17–21].

На высоковольтной части преобразователя – биполярном полумосте VT3VT4VT5VT6 и дросселе L1 – может быть построен бесступенчатый стабилизатор переменного напряжения. Схема силовой части в этом режиме показана на рис. 5.

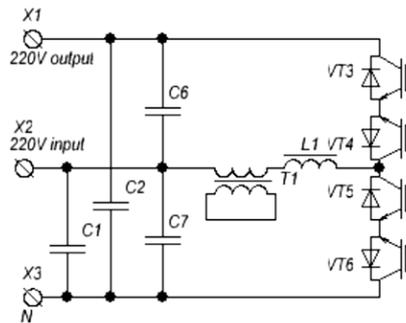


Рис. 5. Силовая часть в режиме стабилизатора переменного напряжения

Низковольтная часть в этом случае не используется, транзисторы VT2, VT8 (рис. 1) постоянно открыты и закорачивают низковольтную обмотку трансформатора T1. Управление коэффициентом передачи осуществляется путем изменения коэффициента заполнения сигналов управления биполярным полумостом. При этом напряжения на клеммах X1 и X2 относительно нейтрали сети X3 связаны соотношением $U_{X2} = U_{X1} \cdot D$, где D – коэффициент заполнения. Блокировочные конденсаторы C1 и C2 предотвращают проникновение компонент тока с частотой переключения в сеть и в цепь нагрузки. Следует отметить, что переключение режимов понижения или повышения напряжения осуществляется

электромеханическим способом – с помощью реле К2 и К3 (рис. 1). При этом функции клемм X1 и X2 меняются местами. Для исключения частых переключений реле при входном напряжении, близком к номинальному, программно введен гистерезис.

Предлагаемый универсальный преобразователь возможно использовать в качестве низковольтного DC/DC преобразователя, что может быть полезным, например, для решения задач сварки в полевых условиях от АКБ или заряда одной АКБ от другой (с разными напряжениями). В DC/DC режиме используется один полумост низковольтной части схемы, как показано на рис. 6.

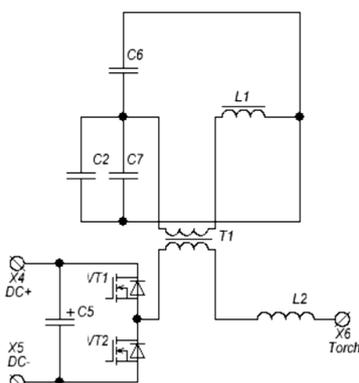


Рис. 6. Силовая часть в режиме сварки от АКБ

Все транзисторы высоковольтной части (VT3–VT6, рис. 1) открыты; трансформатор Т1 при работе находится в глубоком насыщении; АКБ-источник подключается к клеммам X4 и X5. Нагрузка подключается к клеммам X6 и X5. В ее качестве может выступать сварочная горелка для полуавтоматической сварки или другая АКБ (с меньшим напряжением). Следует также отметить, что в режиме «перекачки» заряда между АКБ ток может иметь оба знака, т.е. можно от низковольтной АКБ (подключенной к X6, X5) заряжать другую АКБ с большим напряжением (подключенную к X4, X5).

Как режим сварки, так и режим зарядки АКБ могут быть реализованы и при питании от сети переменного напряжения 220 В. Схема силовой части в этом режиме приведена на рис. 7.

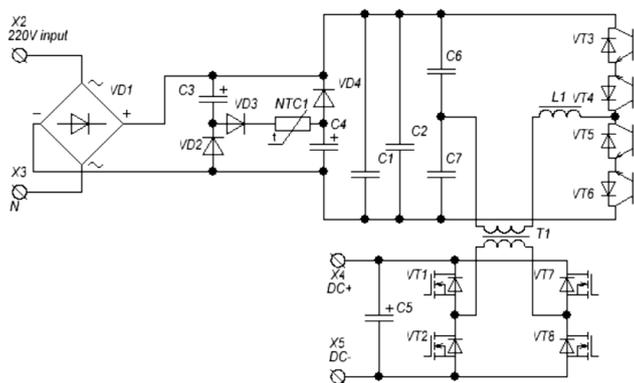


Рис. 7. Силовая часть в режиме сварки от сети, а также в режиме заряда АКБ от сети или солнечных панелей

Поскольку для обеспечения сварочного процесса необходимо постоянно поддерживать ионизацию дугового промежутка, в схеме необходимо предусмотреть накопитель, который обеспечит ток дуги в моменты прохождения сетевого напряжения через ноль. Исходя из этих соображений во входную цепь введен выпрямитель по схеме «valley-fill» [22, 23] на элементах VD1VD2C3VD3NTC1VD4C4. Особенностью этой схемы является то, что большую часть времени (когда модуль мгновенного напряжения сети больше половины амплитуды) нагрузка питается непосредственно от сети через мост VD1. Когда же модуль мгновенного напряжения сети меньше амплитуды, нагрузка питается от параллельно соединенных конденсаторов C3 и C4 через диоды VD2 и VD4. Заряжаются C3 и C4, будучи соединенными последовательно, через цепочку VD3NTC1, когда напряжение сети переходит через максимум. Термистор NTC1 установлен для ограничения тока заряда C3 и C4 при начальном включении преобразователя в сеть.

Остальная часть схемы представляет собой LLC преобразователь, принцип его работы и управления им были описаны выше при рассмотрении работы в режиме ИБП. Система управления обеспечивает контроль выходного тока и напряжения и управление ими в зависимости от задачи – сварка или заряд АКБ.

Заряд АКБ может быть также осуществлен от источника постоянного тока (например, солнечных панелей). При этом алгоритм управления предусматривает режим поиска точки максимальной мощности (MPPT – Maximum Power Point Tracking) [23] для обеспечения максимальной эффективности использования солнечной энергии. Реализован алгоритм типа РАО (Perturb And Observe), т.е. внесение возмущения и наблюдение с дальнейшим движением в сторону увеличения отбираемой мощности.

Заключение. Исходя из вышесказанного, наличие описанного универсального преобразователя в полевых условиях позволит обеспечить: питание электроприемников от бортовых АКБ транспортных средств; выполнение мелких ремонтных работ, связанных со сваркой; заряд АКБ от генератора или альтернативных источников (при их наличии); стабилизацию напряжения для питания устройств, требовательных к качеству напряжения (например, медицинское оборудование). Следует отметить возможность параллельного соединения нескольких преобразователей для увеличения мощности. При этом преобразователи соединяются между собой последовательной шиной передачи данных, один из них назначается ведущим, остальные – ведомыми.

Работа выполняется при финансовой поддержке Минобрнауки России в рамках государственного задания FRRG-2023-0008.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Welding with lithium-ion battery technology. Available at: <https://www.lorch.eu/en/solutions/innovations/battery-technology/>.
2. Welding inverter with built-in battery AccuMax 150. Available at: <https://www.castolin.com/ru-RU/product/accumax-150>.
3. Fronius AccuPocket 150 Battery-Powered Stick Welding Machine. Available at: <https://shopweldingsupplies.com/products/froniusaccupocket-battery-powered-150-stick-welding-machine#shopifyproduct-reviews>.
4. Aniket Dhere, Barkha Shete, Tanay Patil, Aranake V.M. Bidirectional AC-DC Power Converter, *VJER-Vishwakarma Journal of Engineering Research*, December 2018, Vol. 2, Issue 4.
5. Weilun Warren Chen. Bidirectional Three-Phase AC-DC Power Conversion Using DC-DC Converters And A Three-Phase Unfolder: A PhD. In *Electrical Engineering diss*. Logan, Utah State University, 2017. DOI: <https://doi.org/10.26076/64f2-457f>.

6. *Balaji Veerasamy*. Bidirectional AC/DC Converter with Improved Power Factor and Reduced DC Ripple for Battery and Photovoltaic Applications: Doctoral Thesis, Nagoya Institute of Technology. – 2015.
7. *Arau J., et al.* A Novel Uninterruptible Power Supply System With Active Power Factor Correction, in *IEEE Transactions on Power Electronics*, May 2002, Vol.17, No. 3, pp. 405-412. DOI: 10.1109/TPEL.2002.1004248.
8. *Burlaka V.V., Podnebennaya S.K., and Gulakov S.V.* Residential Uninterruptible Power Supply System with Renewable Energy Sources and Battery Pack, in *Proc. IEEE 3rd International Conference on Intelligent Energy and Power Systems (IEPS)*, 2018, pp. 125-128. DOI: 10.1109/IEPS.2018.8559498.
9. *Burlaka V., Gulakov S., Podnebennaya S., Kudinova E., Savenko O.* Bidirectional single stage isolated DC-AC converter, In *proceedings of 2020 IEEE KhPI Week on Advanced Technology (KhPI Week), Kharkiv, 05-10 October 2020*. Kharkiv: National Technical University “Kharkiv Polytechnic Institute”, 2020, pp. 343-346. DOI: 10.1109/KhPIWeek51551.2020.9250107
10. *Burlaka V., Gulakov S., Podnebennaya S., Kudinova E., Plakhtii O., Nerubatskyi V.* An Universal Bidirectional Three-Port DC/DC/AC Converter With Isolated AC Port, In *proceedings of 2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering (UKRCON), Lviv, 26-28 August 2021*. IEEE Catalog Number: CFP21K03-USB, pp. 367-372. DOI: 10.1109/UKRCON53503.2021.9575381.
11. *Abdel-Rahim O. and Abdelhameed E.* Ultimate transformerless boost DC-DC converter for renewable energy applications, *SVU-International Journal of Engineering Sciences and Applications*, 2021, Vol. 2, No. 2, pp. 63-69.
12. *Subramaniam U., Vavilapalli S., Padmanaban S., Blaabjerg F., Holm-Nielsen J.B., and Almakhles D.* A hybrid PV-battery system for ON-grid and OFF-grid applications-controller-in-loop simulation validation, *Energies*, 2020, Vol. 13, No. 3, p. 755. DOI: <https://doi.org/10.3390/en13030755>.
13. *Omar M.A. and Mahmoud M.M.* Design and simulation of a PV system operating in grid-connected and stand-alone modes for areas of daily grid blackouts, *International Journal of Photoenergy*, 2019, Vol. 2019, Article ID 5216583, 9 p. DOI: <https://doi.org/10.1155/2019/5216583>.
14. *Jadhav S., Devdas N., Nisar S., and Bajpai V.* Bidirectional DC-DC converter in solar PV system for battery charging application, in *Proceedings of the 2018 International Conference on Smart City and Emerging Technology (ICSCET)*, IEEE, Mumbai, India, 2018, January, pp. 1-4. DOI: 10.1109/ICSCET.2018.8537391.
15. *Baharudin N.H., Mansur T.M.N.T., Hamid F.A., Ali R., and Misrun M.I.* Topologies of DC-DC converter in solar PV applications, *Indonesian Journal of Electrical Engineering and Computer Science*, 2017, Vol. 8, No. 2, p. 36. DOI: <http://doi.org/10.11591/ijeecs.v8.i2.pp368-374>.
16. *Burlaka V., Gulakov S. and Podnebennaya S.* Low-Cost Transformerless Grid-Tie Inverter For Photovoltaic System, *2019 IEEE 6th International Conference on Energy Smart Systems (ESS)*. Kyiv, Ukraine, 2019, pp. 334-338. DOI: 10.1109/ESS.2019.8764200.
17. *Gowtham Raj T. and Saranya S.K.* Harmonic reduction using L-Z source inverter and voltage source inverter for distributed generation system, *Journal of Chemical and Pharmaceutical Sciences*, 2018, No. 1, pp. 280-283.
18. *Patel N., Kumar A. and Gupta N.* Three-Phase Single-Stage VSC Controlled Solar Photovoltaic System with Harmonic Filtering Capability Applied to DG, *2019 IEEE 2nd International Conference on Power and Energy Applications (ICPEA)*. Singapore, 2019, pp. 238-242. DOI: 10.1109/ICPEA.2019.8818509.
19. *Behera M.P., Ray P.K. and Beng G.H.* Single-phase grid-tied photovoltaic inverter to control active and reactive power with battery energy storage device, *2016 IEEE Region 10 Conference (TENCON)*. Singapore, 2016, pp. 1900-1904. DOI: 10.1109/TENCON.2016.7848352.
20. *Tao Peng, Peng Yang, Hanbing Dan, Hui Wang, Hua Han, Jian Yang, Hao Wang, Hui Dong, Patrick Wheeler.* A Single-Phase Bidirectional AC/DC Converter for V2G Applications, *Energies*, 2017, 10, 881. DOI: 10.3390/en10070881.
21. *Liven Pan, Chengning Zhang.* An Integrated Multifunctional Bidirectional AC/DC and DC/DC Converter for Electric Vehicles Applications, *Energies*, 2016, 9, 493. DOI: 10.3390/en9070493.

22. *Turchi J., Dalal D., Wang P., Jenck L.* Power Factor Correction (PFC). Handbook. Rev. 5. Semiconductor, Apr. 2014. Doc. HBD853/D.
23. *Rashid M.H.* Power Electronics. Handbook. 4th ed. Elsevier, 2018, 1510 p.

Статью рекомендовал к опубликованию д.т.н., профессор К.Е. Румянцев.

Бурлака Владимир Владимирович – ФГБОУ ВО "ПГТУ"; e-mail: vburlaka@rambler.ru; г. Мариуполь, Россия; тел.: +79497048603; кафедра автоматизации электро- и теплоэнергетических комплексов; зав. кафедрой; д.т.н.; доцент.

Гулаков Сергей Владимирович – e-mail: gulakov.s.v@yandex.ru; тел.: +79497298638; кафедра автоматизации электро- и теплоэнергетических комплексов; д.т.н.; профессор.

Головин Андрей Юрьевич – e-mail: a.golovin@tms-2022.ru; тел.: +79497128054; кафедра автоматизации электро- и теплоэнергетических комплексов; м.н.с.

Мироненко Дмитрий Сергеевич – e-mail: mironenko.d.s@yandex.ru; тел.: +79496094732; кафедра информационных систем, технологий и автоматизации в строительстве; зав. кафедрой; к.т.н.; доцент.

Burlaka Vladimir Vladimirovich – FSBEI of HE "PSTU"; e-mail: vburlaka@rambler.ru; Mariupol, Russia; phone: +79497048603; the department of automation of electrical and thermal power complexes; head of the department; dr. of eng. sc.; associate professor.

Gulakov Sergey Vladimirovich – e-mail: gulakov.s.v@yandex.ru; phone: +79497298638; the department of automation of electrical and thermal power complexes; dr. of eng. sc.; professor.

Golovin Andrey Yurievich – e-mail: a.golovin@tms-2022.ru; phone: +79497128054; the department of automation of electrical and thermal power complexes; research associate.

Mironenko Dmitrii Sergeevich – e-mail: mironenko.d.s@yandex.ru; phone: +79496094732; the department of information systems, technologies and automation in construction; head of the department; cand. of eng. sc.; associate 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 ссылку на номер публикации в списке литературы к статье. После аннотаций приводятся ключевые слова (словосочетания), несущие в тексте основную смысловую нагрузку (на русском и английском языках).

4. В тексте статьи следует использовать минимальное количество таблиц и иллюстраций. Рисунок должен иметь объяснения значений всех компонентов, порядковый номер, название, расположенное под рисунком. В тексте на рисунок дается ссылка. Таблица должна иметь порядковый номер, заголовок, расположенный над ней. Данные таблиц и рисунков не должны дублировать текст. Формулы должны быть набраны **в редакторе формул Word 7 for Windows**.

5. Цитаты тщательно сверяются с первоисточником и визируются автором на обратной стороне последней страницы: "Цитаты и фактический материал сверены". Подпись, дата.

6. Наличие пристатейного библиографического списка на русском и английском языках обязательно. **Ссылок должно быть не менее 20-ти**, из них на зарубежные источники – не менее 35 %. В тексте ссылки должны быть в квадратных скобках.

Примеры оформления литературы: а) для книг: фамилия, инициалы автора(ов), полное название книги, место, год издания, страницы; б) для статей: фамилия и инициалы автора(ов), полное название сборника, книги, газеты, журнала, где опубликована статья, место и год издания (сборника, книги), номер (для журнала), год и дата (для газеты), выпуск, часть (для сборника), страницы, на которых опубликована статья. Иностранная литература оформляется по тем же правилам.

Ссылки на неопубликованные работы не допускаются.

7. Рукопись должна быть тщательно вычитана. Редакционная коллегия оставляет за собой право при необходимости сокращать статьи, редактировать и отсылать авторам на доработку.

8. Статьи сопровождаются сведениями об авторе(ах) (фамилия, имя, отчество, ученое звание, должность, место работы, адрес, электронный адрес и номер телефона) на русском и английском языках.

9. Плата с аспирантов за публикацию рукописей не взимается.

Адрес журнала в Интернете: <http://izv-tn.tti.sfedu.ru/>.