



№2-2024

ISSN 1999-9429

ИЗВЕСТИЯ ЮФУ

ТЕХНИЧЕСКИЕ НАУКИ

- Системы управления и моделирование
- Электроника, нанотехнологии и приборостроение
- Алгоритмы обработки информации

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

№ 2 (238). 2024 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); Бородянский 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.

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

Технический редактор Ярошевич Н.В.

Оригинал-макет выполнен Ярошевич Н.В.

Дата выхода в свет 17.05. 2024 г. Формат 70×108 $\frac{1}{16}$. Бумага офсетная.

Офсетная печать. Усл. печ. л. – 25,4. Уч.-изд. л. – 19,7.

Заказ № 9515. 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, 2024

СОДЕРЖАНИЕ

РАЗДЕЛ I. СИСТЕМЫ УПРАВЛЕНИЯ И МОДЕЛИРОВАНИЕ

С.И. Родзин ВЫЧИСЛИТЕЛЬНАЯ МОДЕЛЬ КОЛЛЕКТИВНОГО ПОВЕДЕНИЯ ГРУППЫ ЖИВОТНЫХ: ЭФФЕКТИВНАЯ БИОЭВРИСТИКА ДЛЯ РЕШЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ	6
Ю.А. Брюхомицкий МОДЕЛЬ СИСТЕМЫ БИОМЕТРИЧЕСКОЙ ВЕРИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ ИНФОРМАЦИОННЫХ СИСТЕМ	17
А.Н. Бакуменко, В.А. Деркачев, В.В. Бахчевников, В.Т. Лобач МОДЕЛЬ АЛГОРИТМА ПОТОКОВОЙ МАРКИРОВКИ ШИРОКОФОРМАТНОГО ИЗОБРАЖЕНИЯ	25
С.Л. Беляков, А.В. Исаев МЕТОД ПОДДЕРЖКИ УСТОЙЧИВОСТИ СЕТИ ЭЛЕКТРОСНАБЖЕНИЯ НА ОСНОВЕ ГЕОИНФОРМАЦИОННОЙ МОДЕЛИ	31
Д.Ю. Кравченко МОДЕЛЬ ОНТОЛОГИИ ЗНАНИЙ ДЛЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ОБРАБОТКИ И АНАЛИЗА ТЕКСТОВ.....	38
П.П. Чернусь, П.П. Чернусь, А.А. Яковлев, Р.В. Сахабудинов, А.С. Голосий СИНТЕЗ СИСТЕМЫ УПРАВЛЕНИЯ КОМПЛЕКСА ПОДЪЕМА ПОДВОДНОГО ГРУЗА	50
И.Ю. Липко ПРОГРАММНЫЙ КОМПЛЕКС СТЕНДА ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ КАЧКИ КАТАМАРАНА В ВЕТРОВОЛНОВЫХ УСЛОВИЯХ	59
В.В. Соловьев, А.Я. Номерчук, Р.К. Филатов СИСТЕМНЫЙ АНАЛИЗ НАЗЕМНОЙ РОБОТИЗИРОВАННОЙ ПЛАТФОРМЫ СЕЛЬСКОХОЗЯЙСТВЕННОГО НАЗНАЧЕНИЯ.....	69
И.И. Князев ИССЛЕДОВАНИЕ МЕТОДОВ ПОСТРОЕНИЯ КАУЗАЛЬНЫХ ГРАФОВЫХ МОДЕЛЕЙ ДЛЯ СЛОЖНЫХ СОЦИОГУМАНИТАРНЫХ СИСТЕМ	83
Д.Г. Веселова, Н.Е. Сергеев ПРЕДИКТИВНАЯ АНАЛИТИКА ДЛЯ ПРИНЯТИЯ РЕШЕНИЙ В ДЕЦЕНТРАЛИЗОВАННЫХ СИСТЕМАХ	90
Л.А. Зинченко, В.В. Казаков, Р.Р. Моисеев, Е.С. Поляков ПРОГРАММНЫЙ МОДУЛЬ ДЛЯ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ МНОГОСЛОЙНЫХ ЭКРАНОВ ЗАЩИТЫ ЭЛЕКТРОННОЙ АППАРАТУРЫ ОТ ВОЗДЕЙСТВИЯ ТЯЖЕЛЫХ ЗАРЯЖЕННЫХ ЧАСТИЦ С ИСПОЛЬЗОВАНИЕМ GEANT4	100
Д.В. Семенихина, Я.А. Скотаренко ЧИСЛЕННЫЕ ИССЛЕДОВАНИЯ НЕЛИНЕЙНОЙ МЕТАПОВЕРХНОСТИ КЛИНОВИДНОЙ ФОРМЫ ДЛЯ РАЗРАБОТКИ РЕКОМЕНДАЦИЙ ПО УПРАВЛЕНИЮ ОТРАЖЕННЫМ ЭЛЕКТРОМАГНИТНЫМ ПОЛЕМ КРЫЛА ЛЕТАТЕЛЬНОГО АППАРАТА.....	109
С. Сингх, А.В. Прибыльский СИНТЕЗ СИСТЕМЫ СВЕРХБЫСТРОГО ОБНАРУЖЕНИЯ ПОЖАРООПАСНЫХ СИТУАЦИЙ НА ОСНОВЕ КОМПЛЕКСА ВЗАИМОСВЯЗАННЫХ ДАТЧИКОВ.....	121
А.В. Дьяков, К.Е. Румянцев АКТУАЛЬНЫЕ ПРОБЛЕМЫ РАДИОМОНИТОРИНГА В СИСТЕМЕ МЕРОПРИЯТИЙ ПО ОБЕСПЕЧЕНИЮ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ	132

РАЗДЕЛ II. ЭЛЕКТРОНИКА, НАНОТЕХНОЛОГИИ И ПРИБОРОСТРОЕНИЕ

И.И. Дементьев, А.О. Костина ОБ АНАЛИТИЧЕСКОЙ НЕОПРЕДЕЛЕННОСТИ КОМПОНЕНТ ТЕНЗОРА ЭНЕРГЕТИЧЕСКИХ КОЭФФИЦИЕНТОВ ЭЛЕКТРОМЕХАНИЧЕСКОЙ СВЯЗИ ДЛЯ ПЛЕНОЧНЫХ ПОЛИМЕРНЫХ ПЬЕЗОМАТЕРИАЛОВ	142
А.В. Лабынцев, Н.Е. Пономарев, А.Н. Саржанов, Д.В. Харланов ХАРАКТЕРИСТИКИ ДВУХРЕЗОНАТОРНОГО МИКРОПОЛОСКОВОГО ЗВЕНА ГРЕБЕНЧАТОГО ТИПА С ОПТИМАЛЬНОЙ ДЛИНОЙ ОБЛАСТИ СВЯЗИ МЕЖДУ РЕЗОНАТОРАМИ.....	153
Л.К. Самойлов, Д.Ю. Денисенко, А.Е. Титов МАКСИМАЛЬНЫЕ ДИНАМИЧЕСКИЕ ПОГРЕШНОСТИ ФИЛЬТРОВ ЛЕЖАНДРА В СИСТЕМАХ КОНТРОЛЯ И УПРАВЛЕНИЯ	164
Б.Г. Коноплев АНАЛИЗ ВЫПРЯМИТЕЛЬНЫХ СВОЙСТВ НАНОМЕТРОВЫХ МОП-ТРАНЗИСТОРОВ В ДИОДНОМ ВКЛЮЧЕНИИ ПРИ СВЕРХНИЗКИХ НАПРЯЖЕНИЯХ	175
А.Н. Зикий, А.С. Кочубей ТРАНЗИСТОРНЫЙ ГЕНЕРАТОР С ДИЭЛЕКТРИЧЕСКИМ РЕЗОНАТОРОМ	183
Д.В. Харланов, А.В. Лабынцев ПОТЕНЦИАЛЬНЫЕ ВОЗМОЖНОСТИ ФИЛЬТРА НА ШПИЛЕЧНЫХ РЕЗОНАТОРАХ С МЕТАЛЛИЗИРОВАННЫМИ ОТВЕРСТИЯМИ.....	191
А.В. Бугакова, О.В. Дворников, Н.Н. Прокопенко, В.А. Чеховский, Д.В. Клейменкин СХЕМОТЕХНИЧЕСКИЕ ОСОБЕННОСТИ ПРОЕКТИРОВАНИЯ ВЫСОКОТЕМПЕРАТУРНЫХ АНАЛОГОВЫХ МИКРОСХЕМ НА GaN И GaAs ТРАНЗИСТОРАХ	202

РАЗДЕЛ III. АЛГОРИТМЫ ОБРАБОТКИ ИНФОРМАЦИИ

М.В. Пикалов, А.М. Письмеров НАСТРОЙКА ПАРАМЕТРОВ ГЕНЕТИЧЕСКОГО АЛГОРИТМА ПРИ ПОМОЩИ АНАЛИЗА ЛАНДШАФТА ФУНКЦИИ ПРИСПОСОБЛЕННОСТИ И МАШИННОГО ОБУЧЕНИЯ.....	221
П.А. Гессен, А.И. Лизин, В.А. Павлова, В.Д. Саенко, В.А. Тупиков РЕКУРСИВНЫЙ АЛГОРИТМ АНАЛИЗА И ВОССТАНОВЛЕНИЯ КОНТУРОВ В СИСТЕМАХ НАВИГАЦИИ И НАВЕДЕНИЯ.....	228
И.О. Шепель МЕТОД ГЕНЕРАЦИИ КАРТЫ ЗАНЯТОСТИ МОБИЛЬНОГО РОБОТА ПО ДАННЫМ МУЛЬТИСПЕКТРАЛЬНОЙ СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ	236
И.А. Шипов, Е.В. Ветошкин ПОДХОДЫ К МОДУЛЬНОМУ ПРОЕКТИРОВАНИЮ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ НАВИГАЦИОННЫХ СИСТЕМ НАЗЕМНЫХ ПОДВИЖНЫХ ОБЪЕКТОВ.....	246
Е.А. Титенко, И.Е. Чернецкая, М.А. Титенко, Э.В. Мельник, Д.А. Трокоз РАСШИРЕННАЯ ПРОДУКЦИОННАЯ МАШИНА ВЫВОДА ДЛЯ РЕАЛИЗАЦИИ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ	254
А.А. Белевцев, А.М. Белевцев, В.А. Балыбердин МЕТОДИКА ПОСТРОЕНИЯ ДОРОЖНЫХ КАРТ РАЗВИТИЯ ТЕХНОЛОГИЧЕСКИХ ТРЕНДОВ И ТЕХНОЛОГИЙ В УСЛОВИЯХ ОГРАНИЧЕНИЙ	268
А.В. Руденко, М.А. Руденко ПОДДЕРЖКА ПРИНЯТИЯ ВРАЧЕБНЫХ РЕШЕНИЙ ПРИ ПЛАНИРОВАНИИ ПРОВЕДЕНИЯ ПРОЦЕДУРЫ ЛАЗЕРНОЙ ЛИТОТРИПСИИ.....	278
РАЗДЕЛ IV. СООБЩЕНИЕ ОБ ОТЗЫВЕ ПУБЛИКАЦИИ	290

CONTENT

SECTION I. CONTROL SYSTEMS AND MODELING

S.I. Rodzin A COMPUTATIONAL MODEL OF THE COLLECTIVE BEHAVIOR OF A GROUP OF ANIMALS: EFFECTIVE BIO HEURISTICS FOR SOLVING APPLIED GLOBAL OPTIMIZATION PROBLEMS	6
Yu.A. Bryuhomitsky MODEL OF THE SYSTEM OF BIOMETRIC VERIFICATION OF INFORMATION SYSTEMS USERS	17
A.N. Bakumenko, V.A. Derkachev, V.V. Bakhchevnikov, V.T. Lobach MODEL OF ALGORITHM FOR STREAMING LABELING OF WIDE FORMAT IMAGES	26
S.L. Belyakov, A.V. Isaev METHOD OF SUPPORTING THE STABILITY OF THE POWER SUPPLY NETWORK BASED ON A GEOINFORMATION MODEL	32
D.Yu. Kravchenko KNOWLEDGE ONTOLOGY MODEL FOR INTELLIGENT TEXT PROCESSING AND ANALYSIS SYSTEMS.....	39
P.P. Chernus, P.P. Chernus, A.A. Yakovlev, R.V. Sakhabudinov, A.S. Golosiy SYNTHESIS OF THE UNDERWATER CARGO LIFTING COMPLEX CONTROL SYSTEM	51
I.Y. Lipko THE SOFTWARE PACKAGE FOR THE CATAMARAN ROLL STAND IN WIND WAVE CONDITIONS SIMULATION	60
V.V. Soloviev, A.Y. Nomerchuk, R.K. Filatov SYSTEM ANALYSIS OF A GROUND ROBOTIC PLATFORM FOR AGRICULTURAL PURPOSE	70
I.I. Knyazev RESEARCH OF METHODS FOR CONSTRUCTING CAUSAL GRAPH MODELS FOR COMPLEX SOCIO-HUMANITARIAN SYSTEMS.....	83
D.G. Veselova, N.E. Sergeyev PREDICTIVE ANALYTICS FOR DECISION-MAKING IN DECENTRALIZED SYSTEMS.....	90
L.A. Zinchenko, V.V. Kazakov, R.R. Moiseev, E.S. Polyakov A SOFTWARE FOR AUTOMATED DESIGN OF MULTILAYER SHIELDING FOR ELECTRONIC EQUIPMENT PROTECTION FROM HEAVY CHARGED PARTICLES BASED ON GEANT4.....	101
D.V. Semenikhina, Ya.A. Skotarenko NUMERICAL STUDIES OF A NONLINEAR WEDGE-SHAPED METASURFACE FOR THE DEVELOPMENT OF RECOMMENDATIONS FOR CONTROL OF THE REFLECTED ELECTROMAGNETIC FIELD OF AN AIRCRAFT WING.....	109
S. Singh, A.V. Pribylskiy SYNTHESIS OF A SYSTEM FOR ULTRA-FAST DETECTION OF FIRE-HAZARDOUS SITUATIONS BASED ON A COMPLEX OF INTERCONNECTED SENSORS	122
A.V. Dyakov, K.E. Rumyantsev CURRENT PROBLEMS OF RADIOMONITORING IN THE SYSTEM OF ACTIONS TO ENSURE INFORMATION SECURITY	133

SECTION II. ELECTRONICS, NANOTECHNOLOGY AND INSTRUMENTATION

I.I. Dementev, A.O. Kostina	
ABOUT ANALYTICAL UNCERTAINTY THE COMPONENT OF THE TENSOR OF POWER ELECTROMECHANICAL FORCE FACTORS FOR FILM POLYMERIC PIEZO MATERIAL	142
A.V. Labyntsev, N.E. Ponomarev, A.N. Sarzhanov, D.V. Kharlanov	
CHARACTERISTICS OF A TWO-RESONATOR COMB-TYPE MICROSTRIP ELEMENT WITH AN OPTIMAL LENGTH OF THE COUPLING REGION BETWEEN THE RESONATORS.....	153
L.K. Samoilov, D.Yu. Denisenko, A.E. Titov	
MAXIMUM DYNAMIC ERRORS OF LEGENDRE FILTERS IN CONTROL AND CONTROL SYSTEMS.....	165
B.G. Konoplev	
ANALYSIS OF THE RECTIFYING PROPERTIES OF NANOMETER MOS TRANSISTORS IN A DIODE CONNECTION AT ULTRALOW VOLTAGE	175
A.N. Zikiy, A.S. Kochubey	
TRANSISTOR OSCILLATOR WITH DIELECTRIC RESONATOR.....	184
D.V. Kharlanov, A.V. Labyntsev	
POTENTIAL CAPABILITIES OF THE FILTER ON HAIRPIN RESONATORS WITH METALLIZED HOLES.....	192
A.V. Bugakova, O.V. Dvornikov, N.N. Prokopenko, V.A. Tchekhovski, D.V. Kleimenkin	
CIRCUIT FEATURES OF HIGH-TEMPERATURE ANALOG MICROCIRCUITS ON GaN AND GaAs TRANSISTORS.....	203

SECTION III. INFORMATION PROCESSING ALGORITHMS

M.V. Pikalov, A.M. Pismerov	
GENETIC ALGORITHM PARAMETER TUNING USING EXPLORATORY LANDSCAPE ANALYSIS AND MACHINE LEARNING	221
P.A. Gessen, A.I. Lizin, V.A. Pavlova, V.D. Saenko, V.A. Tupikov	
RECURSIVE ANALYSIS ALGORITHM AND RESTORATION OF CONTOURS IN NAVIGATION AND GUIDANCE SYSTEMS.....	229
I.O. Shepel	
METHOD FOR GENERATING A MOBILE ROBOT OCCUPANCY MAP FROM MULTISPECTRAL VISION SYSTEM DATA	237
I.A. Shipov, E.V. Vetoshkin	
APPROACHES TO MODULAR DESIGN OF SOFTWARE FOR LANDINGS NAVIGATION SYSTEMS.....	247
E.A. Titenko, I.E. Chernetskaya, M.A. Titenko, E.V. Melnik, D.A. Trokoz,	
ADVANCED PRODUCTION OUTPUT ENGINE FOR IMPLEMENTING PARALLEL COMPUTING.....	254
A.A. Belevtsev, A.M. Belevtsev, V.A. Balyberdin	
METHODOLOGY FOR CONSTRUCTING ROADMAPS FOR THE DEVELOPMENT OF TECHNOLOGICAL TRENDS AND TECHNOLOGIES UNDER CONDITIONS OF RESTRICTIONS.....	269
A.V. Rudenko, M.A. Rudenko	
SUPPORT FOR MEDICAL DECISION-MAKING WHEN PLANNING THE LASER LITHOTRIpsy PROCEDURE	279
SECTION IV. REPORT OF RETRACTION.....	290

Раздел I. Системы управления и моделирование

УДК 004.832.23

DOI 10.18522/2311-3103-2024-2-6-16

С.И. Родзин

ВЫЧИСЛИТЕЛЬНАЯ МОДЕЛЬ КОЛЛЕКТИВНОГО ПОВЕДЕНИЯ ГРУППЫ ЖИВОТНЫХ: ЭФФЕКТИВНАЯ БИОЭВРИСТИКА ДЛЯ РЕШЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ*

Перспективным решением задач глобальной оптимизации являются метаэвристики, инспирированные природой, представляющие собой недетерминированные алгоритмы, исследующие пространство поиска, решений, обучающиеся в процессе поиска, не привязанные к конкретной задаче, хотя и не гарантирующие точное решение. Целью данного исследования является разработка эффективного алгоритма для решения прикладных проблем глобальной оптимизации многомерных одномеральных и мультимодальных функций, встречающихся в задачах инженерного проектирования, обработки изображений и компьютерного зрения, энергетики и энергоменеджмента, анализа данных и машинного обучения, робототехники. Для достижения этой цели в статье предлагается вычислительная модель коллективного поведения группы животных и эффективный алгоритм дифференциально-векторного движения. Модель включает разнообразные паттерны поведения в группе животных: удерживать текущую позицию; двигаться в направлении к ближайшим соседям или, наоборот, от ближайших соседей; двигаться случайным образом; конкурировать за позицию. В коллективной памяти хранится информация о местоположении доминирующих особей группы и направлении движения группы, лучшие позиции агентов с учетом механизмов конкуренции и доминирования в группе. Алгоритм был экспериментально протестирован на семи известных многомерных одномеральных и мультимодальных функциях. Результаты были сопоставлены с генетическим алгоритмом, алгоритмом роя частиц, гравитационного поиска дифференциальной эволюции. Предлагаемый алгоритм показал лучшие результаты, нежели конкурирующие алгоритмы, на всех тестовых функциях. Это объясняется лучшим балансом нового алгоритма между скоростью сходимости и диверсификацией пространства поиска решений. Проверка полученных результатов с использованием T-критерия суммы рангов Уилкоксона для независимых выборок показала, что результаты по алгоритму являются статистически значимыми. Также проводилось сравнение с одним из наиболее эффективных алгоритмов непрерывной оптимизации BFGS - квазиньютоновским итерационным алгоритмом численной оптимизации, предназначенным для нахождения локального экстремума одномеральных функций. Результаты оказались сопоставимы для многомерных функций. Алгоритм также сравнивался с методом мультистарта в задаче глобальной оптимизации мультимодальных функций и доказал свое преимущество по времени и точности найденных решений.

Биоэвристика; глобальная оптимизация; дифференциально-векторное движение; память; многомерная функция; агент; оператор; популяция; критерий Уилкоксона.

S.I. Rodzin

A COMPUTATIONAL MODEL OF THE COLLECTIVE BEHAVIOR OF A GROUP OF ANIMALS: EFFECTIVE BIO HEURISTICS FOR SOLVING APPLIED GLOBAL OPTIMIZATION PROBLEMS

A promising solution to global optimization problems are metaheuristics inspired by nature, which are non-deterministic algorithms that explore the search space, solutions, learning in the search process, not tied to a specific task, although they do not guarantee accurate solutions. The purpose of this study is

* Исследование выполнено за счет гранта Российского научного фонда № 23-21-00089, <https://rscf.ru/project/23-21-00089/> в Южном федеральном университете.

to develop an effective algorithm for solving applied problems of global optimization of multidimensional single-modal and multimodal functions found in engineering design, image processing and computer vision, energy and energy management, data analysis and machine learning, robotics. To achieve this goal, the article proposes a computational model of the collective behavior of a group of animals and an effective algorithm for differential vector motion. The model includes various patterns of behavior in a group of animals: to hold the current position; to move towards the nearest neighbors or, conversely, from the nearest neighbors; to move randomly; to compete for a position. The collective memory stores information about the location of the dominant individuals of the group and the direction of movement of the group, the best positions of agents, taking into account the mechanisms of competition and dominance in the group. The algorithm was experimentally tested on seven known multidimensional single-modal and multimodal functions. The results were compared with a genetic algorithm, a particle swarm algorithm, and a gravitational search for differential evolution. The proposed algorithm showed better results than competing algorithms on all test functions. This is due to the better balance of the new algorithm between the rate of convergence and the diversification of the solution search space. Verification of the results obtained using the Wilcoxon sum of ranks *T*-test for independent samples showed that the results of the algorithm are statistically significant. A comparison was also made with one of the most effective continuous optimization algorithms of BFGS - a quasi-Newtonian iterative numerical optimization algorithm designed to find the local extremum of single-modal functions. The results were comparable for multidimensional functions. The algorithm was also compared with the multistart method in the problem of global optimization of multi-extreme functions and proved its advantage in terms of time and accuracy of the solutions found.

Bioheuristics; global optimization; differential vector motion; memory; multidimensional function; agent; operator; population; Wilcoxon criterion.

Введение. Найти произвольный локальный оптимум относительно просто, используя классические методы оптимизации. Найти глобальный оптимум функции гораздо труднее. Детерминированные методы такие как ветвей и границ, численные методы, методы алгебраической геометрии часто неприменимы, дают лишь теоретическую гарантию нахождения локального оптимума. Перспективным решением задачи глобальной оптимизации являются метаэвристики, представляющие собой недетерминированные алгоритмы, исследующие пространство поиска, решений, обучающиеся в процессе поиска, не привязанные к конкретной задаче, хотя и не гарантирующие точное решение. Активной областью исследований является разработка метаэвристик, инспирированных природой. Эта область включает эволюционные и роевые биоэвристики [1], биоэвристики, основанные на физических и химических процессах, на особенностях многоклеточных организмов, способных к фотосинтезу, а также биоэвристики, инспирированные когнитивными процессами и деятельностью человека [2]. Здесь природа выступает источником концепций, механизмов и принципов проектирования искусственных вычислительных систем для решения сложных вычислительных задач глобальной оптимизации, является примером адаптивного решения задачи оптимизации.

В настоящее время известно свыше 500 метаэвристик [3]. Наиболее цитируемым является роевой алгоритм (*Particle Swarm Optimization, PSO*) [4]. По результатам решения трех инженерных тестовых задач (растяжение/сжатие пружины), проектирование сосуда высокого давления, сварка балки) были определены четыре лучших биоэвристики: алгоритм императорских пингвинов (*Emperor Penguin Optimization, EPO*), алгоритм охоты орлов (*Aquila Optimizer, AO*), алгоритм хамелеона (*Chameleon Swarm Algorithm, ChSA*) и алгоритм африканских стервятников (*African Vulture Optimization Algorithm, AVOA*) [5].

Известными примерами приложений глобальной оптимизации является вычислительная филогенетика, проблема коммивояжера и проектирование электрических схем, инженерная безопасность зданий, гипотеза Кеплера, задача упаковки, калибровка моделей распространения радиосигналов [6].

Около половины существующих метаэвристик инспирированы феноменами поведения животных [3, 7]:

1. Птиц:
 - ◆ гнездовой паразитизм кукушек, *Cuckoo Search, CS*;
 - ◆ стая кур в поисках пищи, *Chicken Swarm Optimization, CSO*;
 - ◆ поиск пищи, ее хранение и воровство вороной, *Crow Search Algorithm, CSA*;

- ◆ поиск пищи совой, *Owl Search Algorithm, OSA*;
- ◆ поиск пищи колибри, *Hummingbird's Optimization Algorithm, HOA*;
- ◆ кооперативное охотничье поведение ястребов, *Harris Hawks Optimization, HHO*;
- ◆ охота беркута, *Golden Eagle Optimizer, GEO*;
- ◆ полет голубей, *Pigeon Optimization Algorithm, POA*;
- ◆ миграция чаек, *Seagull Optimization Algorithm, SeOA*;
- ◆ выживание в суровых условиях полярных регионов императорских пингвинов, *Emperor Penguin Optimizer, EPO*;
- ◆ полет африканских стервятников в поисках пищи, *AVOA*;
- 2. Млекопитающих:
 - ◆ летучих мышей, *Bat Algorithm, BA*;
 - ◆ дельфинов, *Dolphin Echolocation, DEO*;
 - ◆ кашалотов, *Sperm Whale Algorithm, SWA*;
 - ◆ горбатых китов, *Whale Optimization Algorithm, WOA*;
 - ◆ серых волков, *Grey Wolf Optimizer, GWO*;
 - ◆ паукообразных обезьян, *Spider Monkey Optimization, SMO*;
 - ◆ горилл, *Gorilla Troops Optimizer, GTO*;
 - ◆ пятнистых гиен, *Spotted Hyena Optimizer, SHO*;
 - ◆ белок-летяг, *Squirrel Search Algorithm, SSA*;
 - ◆ львов, *Lion's Algorithm, LA*;
 - ◆ слонов, *Elephant Search Algorithm, ESA*;
- 3. Насекомых:
 - ◆ муравьиной колонии, *Ant Colony Optimization, ACO*;
 - ◆ пчелиной колонии, *Artificial Bee Colony, ABC*;
 - ◆ мухи дрозофилы, *Drosophila Food Search Optimization, DFO*;
 - ◆ самок комаров, *Mosquito Host Seeking Algorithm, MHSA*;
 - ◆ бабочек-монархов, *Monarch Butterfly Optimization, MBO*;
 - ◆ мотыльков, *Moth Flame Optimization, MFO*;
 - ◆ саранчи, *Locust Swarm, LS*;
 - ◆ кузнечиков, *Grasshopper Optimization Algorithm, GOA*;
 - ◆ стрекоз, *Dragonfly Algorithm, DA* [8];
 - ◆ жуков короедов, *Pity Beetle Algorithm, PBA* [9].

В большинстве из упомянутых алгоритмов для моделирования коллективного поведения животных используется концепция их индивидуального поведения. Ключевой принцип индивидуального поведения состоит в том, что простые повторяющиеся взаимодействия между агентами порождают сложные поведенческие паттерны на уровне колонии или стаи. Примерами являются сеть муравьиных феромонных следов, танцы пчел-разведчиков, миграция рыбных косяков.

Однако исследования в [10] показали существование коллективной памяти в группах животных. Наличие такой памяти свидетельствует о том, что предыстория групповой структуры влияет на коллективное поведение на последующих этапах. С учетом этого появляется возможность моделировать коллективное поведение группы с помощью простых индивидуальных правил и настройки общей памяти.

В статье рассматривается биологическая и вычислительная модель коллективной памяти стаи животных в поисках пищи, а также предлагается эффективная биоэвристика для решения прикладных задач глобальной оптимизации.

В алгоритме поисковые агенты моделируют группу животных, взаимодействующих друг с другом на основе простых поведенческих правил, реализуемых в виде математических операторов. Операторы применяются к каждому агенту с учетом того, что вся группа имеет общую память, в которой хранятся их собственные лучшие позиции. Подход сопоставляется с известными метаэвристиками, а получаемые результаты свидетельствуют об эффективности алгоритма.

Биологическая модель. Паттерны коллективного поведения таких организмов, как муравьи, рыбы, птицы уже давно привлекали внимание натуралистов и ученых. Однако несмотря на долгую историю научных исследований, связь между индивидуальными и групповыми паттернами поведения начала устанавливаться не так давно [11].

Например, агенты в группе часто вынуждены быстро принимать решение о направлении движения в неопределенной и опасной среде. Группа может состоять из агентов, имеющих различный статус и не осведомленных о состоянии других индивидов или об угрозе. В соответствии с принципом конкуренции и доминирования группа может иметь иерархическую структуру. В [12] установлено, что такая структура приводит к более стабильным группам с лучшими свойствами сплоченности среди агентов. Повторяющиеся взаимодействия в группе масштабируются до коллективного поведения и принятия решений в широком диапазоне типов групп животных от насекомых до птиц. Даже среди людей в определенных обстоятельствах имеют место быть сходные паттерны поведения. При этом коллективное принятие решений имеет существенную черту – общую память.

Несмотря на разнообразие передвижений группы животных, многие из коллективных паттернов порождаются простыми правилами, которым следуют отдельные агенты группы: (1) удерживать текущую позицию; (2) двигаться в направлении к ближайшим соседям или, наоборот, от ближайших соседей; (3) двигаться случайным образом; (4) конкурировать за определенную позицию. Выбор правила определяется в соответствии с внутренней мотивацией индивида.

Однако существование коллективной памяти в группах животных влияет на коллективное поведение. Так в коллективной памяти может храниться информация о местоположении доминирующих особей группы или направлении движения группы. Это открывает возможности для моделирования сложного коллективного поведения группы, используя простые индивидуальные правила и общую память. В этой работе при построении алгоритма коллективного поведения животных используется поведенческая модель группы животных, что позволяет определить новые операторы поиска оптимального решения. Также используется коллективная память для хранения лучших позиций агентов (лучших решений) с учетом механизмов конкуренции и доминирования в группе.

Алгоритм коллективного поведения животных (АКПЖ). Алгоритм АКПЖ предполагает наличие набора операторов, реализующих правила взаимодействия и коллективного поведения животных. Каждое решение в пространстве поиска оптимума представляет собой позицию животного. В общей памяти хранятся наилучшие решения в каждом поколении (M_g) и наилучшие решения в течение всех предыдущих поколений (M_h).

Алгоритм АКПЖ является итерационным, начинается со случайной инициализации популяции решений и определения наилучшей позиции в группе. Затем применяются операторы алгоритма пока не будет выполнен критерий останова. Критерием завершения алгоритма является заранее заданное число итераций NI . АКПЖ относится к классу алгоритмов дифференциально-векторного движения, характерной чертой которых является использование механизмов рекомбинации имеющихся репрезентативных решений для создания новых решений [5]. Операторы алгоритма реализуют передвижения относительно ближайших соседей, случайные передвижения, а также обновление общей памяти.

Рассмотрим шаги алгоритма подробнее.

Инициализация популяции.

Алгоритм начинается с инициализации позиций популяции животных (агентов) $A = \{a_1, a_2, \dots, a_{N_p}\}$, где N_p – размер популяции. Каждая позиция агента a_i представляет собой D -мерный вектор, содержащий значения параметров, подлежащих оптимизации. Эти значения случайным образом и равномерно распределяются между заранее заданной начальной нижней и верхней границей j -го параметра a_j^{low} и a_j^{high} соответственно:

$$a_{i,j} = a_j^{low} + rand(0,1) \cdot (a_j^{high} - a_j^{low}), j = 1, 2, \dots, D; i = 1, 2, \dots, N_p, \quad (1)$$

Здесь $a_{i,j}$ – это j -й параметр i -го агента.

Все начальные позиции A сортируются в соответствии с фитнес функцией приспособленности (доминирования) для формирования новой популяции $X = \{x_1, x_2, \dots, x_{N_p}\}$ с последующим сохранением в памяти наилучших решений в каждом поколении (M_g) и наилучших решений в течение всех предыдущих поколений (M_h).

Определение наилучшей позиции в популяции.

В процессе определения наилучшей позиции генерируются первые $B \{a_1, a_2, \dots, a_B\}$ новых позиций агентов популяции A . Эти позиции вычисляются с учетом значений, содержащихся в памяти всех предыдущих поколений M_h , с небольшим случайным возмущением вокруг них:

$$a_i = m_h^l + v, \quad (2)$$

где $l \in \{1, 2, \dots, B\}$, m_h^l – l -й элемент памяти M_h , v – случайный вектор малой длины.

Оператор передвижения относительно ближайших соседей.

Согласно биологической метафоре, животные в соответствии с внутренней мотивацией случайно сближаются или отдаляются от ближайших соседей. Необходимы новые операторы, моделирующие такой биологический паттерн. С этой целью генерируется случайное число r в диапазоне $[0, 1]$. Если r не больше некоторого порогового значения H , то индивидуальная позиция агента перемещается (сближается или отдаляется) относительно ближайшей наилучшей позиции из M_h , иначе – перемещается к ближайшей наилучшей позиции из M_g :

$$a_i = \begin{cases} x_i \pm r \cdot (m_h^{\text{ближ}} - x_i), & \text{если } r \leq H \\ x_i \pm r \cdot (m_g^{\text{ближ}} - x_i), & \text{если } r > H \end{cases} \quad (3)$$

где $i \in \{B+1, B+2, \dots, N_p\}$, $m_h^{\text{ближ}}$ и $m_g^{\text{ближ}}$ – ближайшие наилучшие позиции из M_h и M_g для x_i .

Оператор случайного передвижения.

Следуя биологической модели, один агент с некоторой вероятностью P случайным образом изменяет свою позицию. Это поведенческое правило реализуется с учетом следующего выражения:

$$a_i = \begin{cases} r & \text{с вероятностью } P \\ x_i & \text{с вероятностью } (1 - P) \end{cases} \quad (4)$$

где $i \in \{B+1, B+2, \dots, N_p\}$, r – случайный вектор, определяемый в пространстве поиска. Этот оператор аналогичен повторной инициализации агента в случайной позиции, как это указано в (1).

Обновление общей памяти.

После выполнения операторов определения наилучшей позиции в популяции, передвижения относительно ближайших соседей и случайного передвижения для всех N_p агентов, необходимо обновить память M_h .

Чтобы актуализировать M_h , используется механизм доминирования. Животные, взаимодействующие внутри группы, поддерживают между собой минимальную дистанцию ρ . Эта дистанция зависит от того, насколько агрессивно ведет себя животное. Следовательно, когда два животных противостоят друг другу на таком расстоянии, преобладает наиболее доминирующая особь тем временем другие удаляются.

В алгоритме АКПЖ память M_h обновляется согласно следующей процедуры:

1. Элементы M_h и M_g объединяются: $M_U = M_h \cup M_g$.
2. Каждый элемент $m_U^i \in M_U$ сравнивается попарно с оставшимися элементами памяти $\{m_U^1, m_U^2, \dots, m_U^{2B-1}\}$. Если расстояние между сравниваемыми элементами меньше ρ , то преимущество получает элемент с лучшим значением фитнес-функции, а другой удаляется.
3. Из полученных на предыдущем шаге элементов M_U выбирается наилучшее значение B для построения нового M_h .

Значение ρ влияет на скорость сходимости алгоритма и время вычислений, поэтому параметр ρ вычисляется с учетом следующего уравнения:

$$\rho = \frac{\prod_{j=1}^D (a_j^{high} - a_j^{low})}{10 \cdot D}, \quad (5)$$

где a_j^{high} и a_j^{low} представляют собой предварительно заданные нижнюю и верхнюю границы j -го параметра соответственно в D -мерном пространстве.

Алгоритм АКПЖ включает следующие шаги.

Шаг 1. Установка параметров N_p, B, H, P и NI .

Шаг 2. Случайная инициализация позиций популяции агентов $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{N_p}\}$ согласно (1).

Шаг 3. Сортировка позиций \mathbf{A} в соответствии с фитнес функцией для формирования новой популяции $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N_p}\}$ с последующим сохранением в памяти наилучших решений в каждом поколении \mathbf{M}_g и наилучших решений в течение всех предыдущих поколений \mathbf{M}_h .

Шаг 4. Выбор первых B позиций из \mathbf{A} и сохранение их в памяти \mathbf{M}_g .

Шаг 5. Обновление \mathbf{M}_h (на первой итерации $\mathbf{M}_h = \mathbf{M}_g$).

Шаг 6. Вычисление первых позиций B новой популяции решений $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_B\}$ согласно (2).

Шаг 7. Вычисление остальных элементов, используя операторы передвижения относительно ближайших соседей и случайного передвижения:

```

for  $i=B+1: N_p$ 
  if ( $r_1 < P$ ) then
    оператор передвижения относительно ближайших соседей
      {if ( $r_2 < H$ ) then
         $\mathbf{a}_i = \mathbf{x}_i \pm r \cdot (\mathbf{m}_{h}^{ближайший} - \mathbf{x}_i)$ 
      else if
         $\mathbf{a}_i = \mathbf{x}_i \pm r \cdot (\mathbf{m}_{g}^{ближайший} - \mathbf{x}_i)$ 
      }
    else if
      оператор случайного передвижения
      {
         $\mathbf{a}_i = \mathbf{r}$ 
      }
  end for ( $r_1, r_2 \in rand(0,1)$ )
    
```

Шаг 8. Если условие останова NI выполнено, то алгоритм завершается; иначе – возврат к шагу 3.

Наилучшее значение в \mathbf{M}_h представляет глобальное решение задачи оптимизации.

Результаты экспериментов. Для оценки эффективности представленного алгоритма использовался набор из 7 тестовых функций, представленных в [13], где приводится описание нескольких десятков унимодальных и мультиэкстремальных функций, а также результаты их тестирования различными метаэвристиками для многомерных задач глобальной оптимизации:

- ◆ $f_1(\mathbf{X}) = \sum_{i=1}^n x_i^2, x_i \in [-100, 100]^n, f_{opt} = 0;$
- ◆ $f_2(\mathbf{X}) = \sum_{i=1}^n |x_i| + \prod_{i=1}^n |x_i|, x_i \in [-10, 10]^n, f_{opt} = 0;$
- ◆ $f_3(\mathbf{X}) = \sum_{i=1}^n (\sum_{j=1}^i x_j)^2, x_i \in [-100, 100]^n, f_{opt} = 0;$
- ◆ $f_4(\mathbf{X}) = \sum_{i=1}^{n-1} [100(x_{i-1} - x_i^2)^2 + (x_i - 1)^2], x_i \in [-30, 30]^n, f_{opt} = 0;$
- ◆ $f_5(\mathbf{X}) = \sum_{i=1}^n -x_i \sin(\sqrt{|x_i|}), x_i \in [-500, 500]^n, f_{opt} = -418,98 * n;$
- ◆ $f_6(\mathbf{X}) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10], x_i \in [-5, 12; 5, 12]^n, f_{opt} = 0;$
- ◆ $f_7(\mathbf{X}) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right), x_i \in [-600, 600]^n, f_{opt} = 0.$

Здесь n – размерность функции (n устанавливалась равной 30), f_{opt} – минимальное значение функции.

Чтобы продемонстрировать эффективность, а также вычислительные характеристики предложенного алгоритма создана программная среда на языке программирования *C#*. Тестирование проводилось на персональном компьютере с процессором *Intel Core i7* с ОЗУ-8 Гб в среде *Windows 10*.

Вначале для исследования влияния параметров P и H на результативность алгоритма были протестированы четыре функции: f_1, f_3, f_4, f_7 . При этом максимальное число итераций $NI = 1000$, размер популяции $N_p = 50$, $B = 10$. Поиск глобального минимума для каждой из функций проводился 30 раз.

В табл. 1 представлены результаты экспериментов по алгоритму АКПЖ: наилучшие значения для каждой функции μ и дисперсия σ^2 , усредненные за 30 прогонов при фиксированном значении параметра $H = 0,8$ и изменении P от 0,5 до 0,9 с шагом 0,1.

Таблица 1

Результаты АКПЖ с различными значениями параметра P для функций f_1, f_3, f_4, f_7 при $H = 0,8$

Функция	f_1	f_3	f_4	f_7
$P=0,5; \mu (\sigma^2)$	$2,63 \cdot 10^{-11} (2,13 \cdot 10^{-12})$	$5,71 \cdot 10^{-13} (1,11 \cdot 10^{-14})$	$5,68 \cdot 10^{-11} (2,21 \cdot 10^{-12})$	$1,57 \cdot 10^{-2} (1,25 \cdot 10^{-3})$
$P=0,6; \mu (\sigma^2)$	$1,98 \cdot 10^{-17} (6,51 \cdot 10^{-18})$	$7,78 \cdot 10^{-19} (1,52 \cdot 10^{-20})$	$1,54 \cdot 10^{-17} (1,68 \cdot 10^{-18})$	$1,14 \cdot 10^{-6} (3,71 \cdot 10^{-7})$
$P=0,7; \mu (\sigma^2)$	$1,28 \cdot 10^{-23} (3,54 \cdot 10^{-24})$	$4,47 \cdot 10^{-27} (3,60 \cdot 10^{-28})$	$5,11 \cdot 10^{-22} (4,42 \cdot 10^{-23})$	$2,81 \cdot 10^{-8} (5,21 \cdot 10^{-9})$
$P=0,8; \mu (\sigma^2)$	$2,33 \cdot 10^{-29} (4,41 \cdot 10^{-30})$	$7,62 \cdot 10^{-31} (4,23 \cdot 10^{-32})$	$9,02 \cdot 10^{-28} (6,77 \cdot 10^{-29})$	$4,21 \cdot 10^{-10} (4,87 \cdot 10^{-11})$
$P=0,9; \mu (\sigma^2)$	$4,53 \cdot 10^{-23} (5,12 \cdot 10^{-24})$	$3,42 \cdot 10^{-26} (3,54 \cdot 10^{-27})$	$4,77 \cdot 10^{-20} (1,94 \cdot 10^{-21})$	$4,58 \cdot 10^{-4} (6,92 \cdot 10^{-5})$

В табл. 2 представлены результаты экспериментов по алгоритму АКПЖ: наилучшие значения для каждой функции μ и дисперсия σ^2 , усредненные за 30 прогонов при фиксированном значении параметра $P = 0,8$ и изменении H от 0,5 до 0,9 с шагом 0,1.

Для наглядности наилучшие результаты выделены шрифтом. Результаты свидетельствуют о том, что правильное сочетание различных значений параметров повышает точность решения.

Сравнение производительности АКПЖ на 7 тестовых функциях производилось с результатами, полученными с помощью генетического алгоритма (*GA*) [14], алгоритма роя частиц (*PSO*) [15], алгоритм гравитационного поиска (*GSA*) [16] и алгоритма дифференциальной эволюции (*DE*) [17]. Для корректного сравнения устанавливались одни и те же общие параметры управления: размер популяции $N_p = 50$, максимальное число итераций $NI = 1000$.

Таблица 2

Результаты АКПЖ с различными значениями параметра H для функций f_1, f_3, f_4, f_7 при $P = 0,8$

Функция	f_1	f_3	f_4	f_7
$H=0,5; \mu (\sigma^2)$	$2,23 \cdot 10^{-10} (8,92 \cdot 10^{-11})$	$5,71 \cdot 10^{-10} (5,12 \cdot 10^{-11})$	$8,80 \cdot 10^{-9} (5,55 \cdot 10^{-10})$	$1,81 \cdot 10^{-4} (2,16 \cdot 10^{-5})$
$H=0,6; \mu (\sigma^2)$	$3,35 \cdot 10^{-18} (3,21 \cdot 10^{-19})$	$3,24 \cdot 10^{-18} (1,32 \cdot 10^{-19})$	$6,72 \cdot 10^{-21} (1,11 \cdot 10^{-22})$	$2,89 \cdot 10^{-6} (6,43 \cdot 10^{-7})$
$H=0,7; \mu (\sigma^2)$	$3,85 \cdot 10^{-22} (6,78 \cdot 10^{-23})$	$6,29 \cdot 10^{-27} (8,26 \cdot 10^{-28})$	$1,69 \cdot 10^{-23} (1,34 \cdot 10^{-24})$	$2,36 \cdot 10^{-7} (3,75 \cdot 10^{-8})$
$H=0,8; \mu (\sigma^2)$	$2,33 \cdot 10^{-29} (4,41 \cdot 10^{-30})$	$7,62 \cdot 10^{-31} (4,23 \cdot 10^{-32})$	$9,02 \cdot 10^{-28} (6,77 \cdot 10^{-29})$	$4,21 \cdot 10^{-10} (4,87 \cdot 10^{-11})$
$H=0,9; \mu (\sigma^2)$	$4,72 \cdot 10^{-21} (6,29 \cdot 10^{-22})$	$5,41 \cdot 10^{-22} (5,28 \cdot 10^{-23})$	$7,39 \cdot 10^{-21} (4,41 \cdot 10^{-22})$	$3,02 \cdot 10^{-4} (4,37 \cdot 10^{-6})$

Настройки параметров для каждого из конкурирующих алгоритмов соответствовали оригинальным работам:

- ♦ согласно [14] *GA* использует арифметический кроссовер, гауссовскую мутацию и селекцию колеса рулетки. Вероятности операторов кроссинговера и мутации были установлены на 0,3 и 0,1 соответственно;
- ♦ в алгоритме *PSO* устанавливались параметры $c_1 = c_2 = 2$, а коэффициент инерции (ω) линейно уменьшается от 0,9 до 0,2;
- ♦ в алгоритме *GSA* параметры $G_0 = 100$, $\alpha = 20$; $T = 1000$. В [16] эти значения были установлены в качестве наилучших;
- ♦ в алгоритме *DE* согласно [17] вероятность кроссинговера равна 0,9, а весовой коэффициент F равен 0,8.

Результаты сравнения алгоритма АКПЖ с конкурирующими алгоритмами *GA*, *PSO*, *GSA*, *DE* на тестовых функциях $f_1, f_2, f_3, f_4, f_5, f_6, f_7$ приведены в табл. 3.

Сравнение проводилось по показателям среднее наилучшее значение функции (*av*) и медиана наилучшего значения функции на последней итерации (*med*). Наилучший результат для каждой функции выделен шрифтом.

Согласно табл. 3 АКПЖ показал лучшие результаты, нежели конкурирующие алгоритмы *GA*, *PSO*, *GSA* и *DE* на всех семи тестовых функциях. Это объясняется лучшим балансом АКПЖ между скоростью сходимости и диверсификацией пространства поиска решений.

Таблица 3

Результаты АКПЖ в сравнении с алгоритмами *GA*, *PSO*, *GSA*, *DE*

Функция		f_1	f_2	f_3	f_4	f_5	f_6	f_7
<i>GA</i>	<i>av</i>	23,13	1,07	$5,6 \cdot 10^3$	$1,1 \cdot 10^3$	$-1,2 \cdot 10^4$	5,90	1,16
	<i>med</i>	21,87	1,13	$5,6 \cdot 10^3$	$1,0 \cdot 10^3$	$-1,2 \cdot 10^4$	5,71	1,14
<i>PSO</i>	<i>av</i>	$1,8 \cdot 10^{-3}$	2,0	$4,1 \cdot 10^{-3}$	$3,6 \cdot 10^4$	$-9,8 \cdot 10^3$	55,1	0,01
	<i>med</i>	$1,2 \cdot 10^{-3}$	$1,9 \cdot 10^{-3}$	$2,2 \cdot 10^{-3}$	$1,7 \cdot 10^3$	$-9,8 \cdot 10^3$	56,6	0,008
<i>GSA</i>	<i>av</i>	$7 \cdot 10^{-11}$	$4 \cdot 10^{-3}$	$0,16 \cdot 10^3$	25,16	$-2,8 \cdot 10^3$	15,32	0,29
	<i>med</i>	$7 \cdot 10^{-11}$	$4 \cdot 10^{-3}$	$0,15 \cdot 10^3$	25,18	$-2,6 \cdot 10^3$	14,42	0,04
<i>DE</i>	<i>av</i>	11,21	0,95	0,12	0,25	$-4,1 \cdot 10^3$	30,12	$1 \cdot 10^{-3}$
	<i>med</i>	13,21	1,05	0,09	0,31	$-4,1 \cdot 10^3$	31,43	$1 \cdot 10^{-3}$
АКПЖ	<i>av</i>	$2 \cdot 10^{-29}$	$5 \cdot 10^{-20}$	$8 \cdot 10^{-31}$	$9 \cdot 10^{-28}$	$-1,2 \cdot 10^4$	$1 \cdot 10^{-3}$	$1 \cdot 10^{-13}$
	<i>med</i>	$1 \cdot 10^{-20}$	$3 \cdot 10^{-11}$	$1 \cdot 10^{-19}$	$3 \cdot 10^{-18}$	$-1,2 \cdot 10^4$	$7 \cdot 10^{-4}$	$1 \cdot 10^{-13}$

Также оценивалась статистическая значимость полученных результатов. С этой целью применялся *T*-критерий суммы рангов Уилкоксона [18] для независимых выборок, найденных каждым из сравниваемых алгоритмов на 30 тестовых запусках, при уровне значимости 5%. Значение $T < 0,05$ рассматривалось как адекватное доказательство против нулевой гипотезы, которая отвергается. Предложенный алгоритм АКПЖ превосходит конкурирующие, а экспериментальные результаты по алгоритму являются статистически значимыми.

Также проводилось сравнение с одним из наиболее эффективных алгоритмов непрерывной оптимизации *BFGS* – квазиньютоновским итерационным алгоритмом численной оптимизации, предназначенным для нахождения локального экстремума нелинейного функционала без ограничений [19]. Эксперименты проводились в *MatLab*.

В первом эксперименте сравнивалась производительность алгоритмов *BFGS* и АКПЖ на унимодальной функции. В унимодальных функциях глобальный минимум совпадает с локальным минимумом. *BFGS* имеет высокую скорость локальной сходимости, которая зависит от размерности задачи. В качестве эталона для сравнения была выбрана гладкая и дифференцируемая функция Розенброка (f_5).

В эксперименте оба алгоритма *BFGS* и АКПЖ решали задачу минимизации функции f_5 при различной размерности функции. Для реализации *BFGS* в качестве исходной матрицы принималась $B_0 = I$ (первый шаг эквивалентен градиентному спуску). В качестве показателей эффективности сравнения рассматривались время решения и число итераций для достижения минимума. Для *BFGS* в качестве условия останова алгоритма принималось $\|g_5(X)\| \leq 1 \times 10^{-6}$, где $g_5(X)$ является градиентом $f_5(X)$. Для АКПЖ критерием останова является условие, когда больше не происходят изменения в памяти M_n .

В табл. 4 представлены результаты обоих алгоритмов с учетом следующих размерностей функции f_5 : $n \in \{2, 10, 30, 50, 70, 100, 120\}$.

Для обеспечения согласованности результатов учитывались два показателя: среднее затраченное время (t_{cp}) и среднее число итераций (IN_{cp}) после 30 прогонов алгоритмов.

Таблица 4

Результаты сравнения алгоритма *BFGS* и алгоритма АКПЖ на многомерной функции Розенброка $f_5(X)$

n	f_5	t_{cp} (с)		IN_{cp}	
		<i>BFGS</i>	АКПЖ	<i>BFGS</i>	АКПЖ
2		0,15	4,21	6	89
10		0,55	5,28	22	98
30		1,35	5,44	41	108
50		2,77	5,88	68	112
70		4,23	6,11	93	115
100		5,55	6,22	105	121
120		6,64	6,71	125	129

Из табл. 4 видно, что алгоритм *BFGS* превосходит АКПЖ по показателям t_{cp} и IN_{cp} при небольших значениях размерности. Однако при размерности функции $f_5(X)$, начиная с $n \geq 70$, результаты по алгоритму АКПЖ сопоставимы с результатами *BFGS*. Тот факт, что алгоритм *BFGS* превосходит предлагаемый гибридный алгоритм, нельзя рассматривать как недостаток АКПЖ, учитывая ограничения, накладываемые на функции методом *BFGS*.

В случае решения задачи глобальной оптимизации мультиэкстремальной функции алгоритм *BFGS* попадает в ловушку локальных оптимумов. Это ограничивает его применение для решения задач глобальной оптимизации. Одним из наиболее широко используемых для решения задач глобальной оптимизации является метод мултистарта (*MS*) [20]. В *MS* в качестве исходного решения случайным образом выбирается точка из допустимой области, а затем из нее запускается алгоритм непрерывной оптимизации. Затем процесс повторяется до тех пор, пока не будет достигнуто значение, близкое к глобальному оптимуму.

Во втором эксперименте сравнивалась производительность *MS* и АКПЖ на мульти-модальных функциях $f_6(X)$ и $f_7(X)$. Сравнение проводилось по показателям среднее затраченное время (t_{cp}), среднее число итераций (IN_{cp}) и среднее наилучшее значение функции (av) после 30 прогонов алгоритмов. Результаты сравнения приведены в табл. 5.

Таблица 5

Результаты сравнения алгоритма *MS* и алгоритма АКПЖ на multimodalных функциях $f_6(X)$ и $f_7(X)$ при $n = 30$

Функция	<i>MS</i>			АКПЖ		
	t_{cp}	IN_{cp}	av	t_{cp}	IN_{cp}	av
$f_6(X)$	45,4	23,3	$1,2 \cdot 10^{-2}$	10,2	633	$1,0 \cdot 10^{-3}$
$f_7(X)$	72,1	102,3	$4,51 \cdot 10^{-10}$	15,8	884	$1,14 \cdot 10^{-13}$

Из табл. 5 видно преимущество АКПЖ по времени и точности найденных решений.

Заключение. Представлена биоэвристика для решения прикладных задач глобальной оптимизации, основанная на вычислительной модели коллективного поведения группы животных, взаимодействующих между собой на основе простых поведенческих правил, реализуемых в виде математических операторов и имеющих общую память.

Алгоритм был экспериментально протестирован на множестве из семи известных одно-модальных и multimodalных функций. Его производительность сравнивалась с генетическим алгоритмом, алгоритмом роя частиц, алгоритмом гравитационного поиска и алгоритмом дифференциальной эволюции. Эксперименты показали, что предлагаемый алгоритм в целом превосходит конкурирующие алгоритмы по всем тестовым функциям,

а полученные результаты являются статистически значимыми. Сравнение предлагаемой биоэвристики с одним из наиболее эффективных алгоритмов непрерывной оптимизации *BFGS* на унимодальной функции с размерностью от $n = 2$ до $n = 120$ показало, что *BFGS* превосходит АКПЖ при небольших значениях размерности, однако, начиная с $n = 70$, их результаты становятся сопоставимыми при том, что на функции, оптимизируемые алгоритмом АКПЖ, не накладывается ограничений, в отличие от *BFGS*. Сравнение АКПЖ с методом мултистарта на мултимодальных функциях также показывает его преимущество по времени и точности найденных решений.

Результаты, в основном, обусловлены двумя причинами: (1) операторы, моделирующие биологические паттерны поведения позволяют лучше исследовать пространство поиска оптимума; (2) мултимодальной оптимизации не только сохраняется разнообразие решений, содержащихся в памяти M_h , но даже увеличивается за счет реализации принципа конкуренции.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Курейчик В.В., Родзин С.И. Вычислительные модели эволюционных и роевых биоэвристик (обзор) // Информационные технологии. – 2021. – Т. 27, № 10. – С. 507-520.
2. Курейчик В.В., Родзин С.И. Вычислительные модели биоэвристик, основанных на физических и когнитивных процессах (обзор) // Информационные технологии. – 2021. – Т. 27, № 11. – С. 563-574.
3. Dragoi E.N., Dafinescu V. Review of Metaheuristics Inspired from the Animal Kingdom // Mathematics. – 2021. – No. 9. – P. 2335.
4. Bonyadi M., Michalewicz Z. Particle swarm optimization for single objective continuous space problems: a review // Evolutionary Computation. – 2017. – Vol. 25 (1). – P. 1-54.
5. Родзин С.И. Современное состояние биоэвристик: классификация, бенчмаркинг, области применения // Известия ЮФУ. Технические науки. – 2023. – № 2. – С. 280–298.
6. Глобальная оптимизация: Википедия. Свободная энциклопедия. – URL: [https://ru.wikipedia.org/wiki/Глобальная оптимизация](https://ru.wikipedia.org/wiki/Глобальная_оптимизация) (дата обращения: 20.01.2024).
7. Курейчик В.В., Родзин С.И. Биоэвристики, инспирированные фауной (обзор) // Информационные технологии. – 2023. – Т. 29, № 11. – С. 559-573.
8. Mirjalili S. Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi objective problems // Neural Comput. Appl. – 2015. – Vol. 27. – P. 1053-1073.
9. Kallioras N.A., et al. Pity beetle algorithm – A new metaheuristic inspired by the behavior of bark beetles // Adv. Eng. Softw. – 2018. – Vol. 121. – P. 147-166.
10. Couzin I.D. Collective minds // Nature. – 2007. – Vol. 445. – P. 715-728.
11. Bode N., Wood A., Franks D. The impact of social networks on animal collective motion // Anim. Behav. – 2011. – Vol. 82 (1). – P. 29-38.
12. Broom M., Koenig A., Borries C. Variation in dominance hierarchies among group-living animals: modeling stability and the likelihood of coalitions // Behav. Ecol. – 2009. – Vol. 20. – P. 844-855.
13. Long W., et al. Solving high-dimensional global optimization problems using an improved sine cosine algorithm // Expert systems with applications. – 2019. – Vol. 123. – P. 108-126.
14. Hamzaçebi C. Improving genetic algorithms' performance by local search for optimization // Appl. Math. Comput. – 2008. – Vol. 196 (1). – P. 309-317.
15. Kennedy J., Eberhart R.C. Particle swarm optimization // Proc. IEEE Inter. Conf. on Neural Networks. – 1995. – Vol. 4. – P. 1942-1948.
16. Rashedi E., Nezamabadi-pour H., Saryazdi S. GSA: a gravitational search algorithm // Inf. Sci. – 2009. – Vol. 179. – P. 2232-2248.
17. Storn R., Price K. Differential evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces // Technical report TR-95–012. ICSI. Berkeley. CA. – 1995.
18. Wilcoxon F. Individual comparisons by ranking methods Frank Wilcoxon // Bio-metrics Bulletin. – 2006. – Vol. 1 (6). – P. 80-83.
19. Al-Baali M. On the behavior of bombined extra-updating/self-scaling BFGS method // Jour. Comput. Appl. Math. – 2001. – Vol. 134. – P. 269-281.
20. *art i* , et al. Intelligent multi-start methods // Handbook of Metaheuristics. – Cham: Springer, 2019. – P. 221-243.

REFERENCES

1. 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 Technology], 2021, Vol. 27, No. 10, pp. 507-520.
2. Kureychik V.V., Rodzin S.I. Vychislitel'nye modeli bioevristik, osnovannykh na fizicheskikh i kognitivnykh protsessakh (obzor) [Computational models and bio heuristics based on physical and cognitive processes (review)], *Informatsionnye tekhnologii* [Information Technology], 2021, Vol. 27, No. 11, pp. 563-574.
3. Dragoi E.N., Dafinescu V. Review of Metaheuristics Inspired from the Animal Kingdom, *Mathematics*, 2021, No. 9, pp. 2335.
4. Dragoi E.N., Dafinescu V. Review of Metaheuristics Inspired from the Animal Kingdom, *Mathematics*, 2021, No. 9, pp. 2335.
5. Rodzin S.I. Sovremennoe sostoyanie bioevristik: klassifikatsiya, benchmarking, oblasti primeneniya [Computational models of bioheuristics based on physical and cognitive processes (review)], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2023, No. 2, pp. 280-298.
6. Global'naya optimizatsiya: Vikipediya. Svobodnaya entsiklopediya [Global optimization: Wikipedia. The free encyclopedia]. Available at: [https://ru.wikipedia.org/wiki/Global'naya optimizatsiya](https://ru.wikipedia.org/wiki/Global'naya_optimizatsiya) (accessed 20 January 2024).
7. Kureychik V.V., Rodzin S.I. Bioevristiki, inspirirovannye faunoy (obzor) [Bio heuristics inspired by fauna (review)], *Informatsionnye tekhnologii* [Information Technology], 2023, Vol. 29, No. 11, pp. 559-573.
8. Mirjalili S. Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi objective problems, *Neural Comput. Appl.*, 2015, Vol. 27, pp. 1053-1073.
9. Kallioras N.A., et al. Pity beetle algorithm – A new metaheuristic inspired by the behavior of bark beetles, *Adv. Eng. Softw.*, 2018, Vol. 121, pp. 147-166.
10. Couzin I.D. Collective minds, *Nature*, 2007, Vol. 445, pp. 715-728.
11. Bode N., Wood A., Franks D. The impact of social networks on animal collective motion, *Anim. Behav.*, 2011, Vol. 82 (1), pp. 29-38.
12. Broom M., Koenig A., Borries C. Variation in dominance hierarchies among group-living animals: modeling stability and the likelihood of coalitions, *Behav. Ecol.*, 2009, Vol. 20, pp. 844-855.
13. Long W., et al. Solving high-dimensional global optimization problems using an improved sine cosine algorithm, *Expert systems with applications*, 2019, Vol. 123, pp. 108-126.
14. Hamzaçebi C. Improving genetic algorithms' performance by local search for optimization, *Appl. Math. Comput.*, 2008, Vol. 196 (1), pp. 309-317.
15. Kennedy J., Eberhart R.C. Particle swarm optimization, *Proc. IEEE Inter. Conf. on Neural Networks*, 1995, Vol. 4, pp. 1942-1948.
16. Rashedi E., Nezamabadi-pour H., Saryazdi S. GSA: a gravitational search algorithm, *Inf. Sci.*, 2009, Vol. 179, pp. 2232-2248.
17. Storn R., Price K. Differential evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces, Technical report TR-95–012. ICSI. Berkeley, CA, 1995.
18. Wilcoxon F. Individual comparisons by ranking methods Frank Wilcoxon, *Bio-metrics Bulletin*, 2006, Vol. 1 (6), pp. 80-83.
19. Al-Baali M. On the behavior of combined extra-updating/self-scaling BFGS method, *Jour. Comput. Appl. Math.*, 2001, Vol. 134, pp. 269-281.
20. *art i ., et al.* Intelligent multi-start methods, *Handbook of Metaheuristics*. Cham: Springer, 2019, pp. 221-243.

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

Родзин Сергей Иванович – Южный федеральный университет; e-mail: srodzin@sfedu.ru; г. Таганрог, Россия; кафедра математического обеспечения и применения ЭВМ; профессор.

Rodzin Sergey Ivanovich – Southern Federal University; e-mail: srodzin@sfedu.ru; Taganrog, Russia; phone: +78634371673; the department of software engineering; professor.

Ю.А. Брюхомицкий

МОДЕЛЬ СИСТЕМЫ БИОМЕТРИЧЕСКОЙ ВЕРИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ ИНФОРМАЦИОННЫХ СИСТЕМ

Предлагается гибридная модель системы текстонезависимой динамической верификации пользователей информационных систем, которая основана на комплексном использовании искусственных иммунных систем и искусственных нейронных сетей. Подлежащие верификации данные текстонезависимой динамической биометрии пользователей представлены, двумя последовательностями информационных единиц фиксированного размера векторов признаков, соответствующих образам двух классов – «свой» и «чужой». Такое представление ориентировано на массово-параллельную децентрализованную обработку данных, принятую в искусственных иммунных системах. Последующая верификация пользователей обоих классов реализуется с помощью вероятностной искусственной нейронной сети, которая в признаковом пространстве вычисляет плотности вероятности концентрации информационных единиц обоих классов. В дополнение к характеристикам плотности вероятности информационных единиц используются допустимые цены ошибок 1-го и 2-го рода для образов каждого класса. Итоговый результат биометрической верификации работающего пользователя контролируется на основании текущего сравнения совокупных статистических оценок плотности вероятности и допустимой цены ошибок образов каждого из двух классов. Предлагаемый подход к верификации личности работающего пользователя позволяет предложить общую схему этой процедуры для существенно различных модальностей динамической биометрии: голоса, рукописи и клавиатурного набора. Реализация такого подхода для биометрии конкретной модальности будет несколько отличаться, но общая схема верификации может быть сохранена. Преимуществами предлагаемого подхода являются: возможность текстонезависимого анализа динамической биометрии различной модальности, произвольного объема, содержания и языка; возможность принятия верификационного решения в непрерывном режиме в темпе поступления работы пользователя; в перспективе повышать точность работы системы верификации путем увеличения размерности нейронной сети; возможность использования истории анализа результатов верификации реальных пользователей для последующей более точной настройки системы. Относительным недостатком работы является необходимость программной реализации нейронной сети большой размерности. Однако в перспективе этот недостаток быстро нивелируется с повышением производительности средств вычислительной техники.

Текстонезависимая биометрическая верификация личности по динамическим биометрическим параметрам; искусственная иммунная система; вероятностная нейронная сеть; статистическая оценка плотности вероятности; цена ошибки классификации.

Yu.A. Bryuhomitsky

MODEL OF THE SYSTEM OF BIOMETRIC VERIFICATION OF INFORMATION SYSTEMS USERS

A hybrid model of the system of text-independent dynamic verification of users of information systems, which is based on the integrated use of artificial immune systems and artificial neural networks, is proposed. The verifiable data of text-independent dynamic user biometrics are represented by two sequences of information units of fixed-size feature vectors corresponding to the images of two classes – 'friend' and 'stranger'. This representation is oriented towards the massively parallel decentralized data processing adopted in artificial immune systems. The subsequent verification of the users of both classes is realized by a probabilistic artificial neural network, which computes the probability densities of the concentration of information units of both classes in the feature space. In addition to the probability density characteristics of the information units, the allowable 1st and 2nd kind error prices for images of each class are used. The final result of biometric verification of the working user is controlled based on the current comparison of the aggregate statistical estimates of the probability density and the acceptable price of errors of the images of each of the two classes. The proposed approach to verifying the identity of a working user allows to propose a general scheme of this procedure for significantly different modalities of dynamic biometrics: voice, handwriting, and keyboard typing. The implementation of such an approach for specific modality biometrics will be slightly different, but the general verification scheme can be maintained. The advantages of the proposed approach are: the possibility of text-independent analysis of dynamic biometry of different modality, arbitrary volume, content and language; possibility of making a

verification decision in continuous mode at the rate of user's work arrival; in the future to increase the accuracy of the verification system by increasing the dimensionality of the neural network; the possibility of using the history of analysis of verification results of real users for further more accurate tuning of the system. A relative disadvantage of the work is the necessity of program realization of a neural network of large dimensionality. However, in the future, this disadvantage will be quickly leveled with the increase of computing performance.

Text-independent biometric verification of identity by dynamic biometric parameters; artificial immune system; probabilistic neural network; statistical estimation of probability density of information units of two classes; price of classification errors.

Введение. Биометрическая идентификация личности в настоящее время получила широкое распространение. При этом большинство систем основано на распознавании личности по статическим биометрическим параметрам. При наличии многих преимуществ статической биометрии (удобство использования, компактность биометрических эталонов, относительная простота процедур регистрации и идентификации, возможность идентификации больших потоков людей), она имеет и ряд недостатков (открытость биометрических параметров, допускающая использование муляжей, высокая стоимость, негативное отношение некоторых слоев общества к сбору биометрических данных).

Применяются также динамические системы биометрической идентификации личности, основанные на анализе индивидуальных особенностей хорошо заученных подсознательных движений человека. Такие системы используются преимущественно как средство аутентификации личности при входе в информационные системы (ИС). Практическое применение в настоящее время получили системы анализа голоса [1–3], рукописи [4–8] и клавиатурного почерка [9–13]. Преимущества этих систем: низкая стоимость, возможность сохранения образа личности в тайне и быстрой смены этого образа в случае его компрометации. Недостатки этих систем – сравнительно меньшая точность и зависимость результатов идентификации от психофизического состояния человека (испуг, стресс и т.п.). Вместе с тем этот недостаток в некоторых приложениях может эффективно использоваться для контроля психофизического состояния личности (допуск к работе, характеризующейся высокой ценой ошибки, выявление лжи и правонарушений со стороны пользователей (аналог полиграфа) и т.п.).

Особое место в динамических биометрических системах идентификации занимают т.н. текстонезависимые технологии, в которых вместо парольных слов и фраз используется личностные особенности воспроизведения произвольных текстов: голосом, рукописью, клавиатурным набором. В таких системах эталоны личностей строятся на основе достаточно больших образцов текста соответствующей модальности. При этом возникает ряд проблем, связанных с трудностью формирования, анализа и сопоставления эталонов с предъявляемыми образцами. Вместе с тем, эти проблемы уже возможно удовлетворительно решать в задачах информационной безопасности, связанных с текущим контролем работы пользователей в ИС. Примерами таких задач являются: непрерывная скрытая верификация работающих пользователей ИС; скрытое выявление инсайдеров, осуществляющих злонамеренные и правонарушающие действия в ИС; выявление отклонений в психофизическом состоянии личности; аудит безопасности ИС на основе интерактивного взаимодействия администратора ИС с пользователями; выявление лжи (аналог полиграфа) и другие подобные задачи.

Постановка задачи. Биометрическая аутентификация личности может реализовываться в двух вариантах:

- ♦ путем предъявления личностью только биометрических признаков, которые последовательно сопоставляются со биометрическими эталонами всех пользователей ИС, зарегистрированных в биометрической базе данных ИС, с целью выявления признаков, схожих к предъявленным (аутентификация);

- ♦ путем предварительного предъявления личностью не биометрического идентификатора, позволяющего извлечь из биометрической базы данных ИС соответствующий ему биометрический эталон (если таковой имеется) с последующим сопоставлением его с предъявленным биометрическим признаком (верификация).

Данная работа ориентирована на систему верификации личности, получившую преимущественное распространение при использовании биометрических идентификаторов. Для этого предлагается использовать гибридную модель, сочетающую в себе две технологии искусственного интеллекта: искусственные иммунные системы (ИИС) – для представления образов текстонезависимой динамической биометрии) [14–20] и искусственные нейронные сети (ИНС) – для сопоставления и распознавания этих образов [21].

В текстонезависимой динамической биометрии данные представлены, по существу, сигналами (функциями времени), структура которых отражает индивидуальные особенности текущего воспроизведения личностью произвольных текстов различной модальности. Для распознавания таких сигналов предложено множество методов, которые обычно сводятся сначала к переводу их в статическое представление в частотной, или в частотно-временной областях. После чего задача распознавания решается уже в формате статического представления исходных сигналов.

В данной работе для верификации работающей в ИС личности предлагается альтернативный подход к распознаванию сигналов, который сводится к представлению их последовательностями информационных единиц фиксированного размера, в виде векторов признаков двух классов пользователей – «свой» и «чужой». Такое представление хорошо согласуется с массово-параллельной децентрализованной обработкой данных, принятой в ИИС.

На этапе верификации пользователей двух классов «свой» – «чужой» используется вероятностная нейронная сеть (PNN – Probabilistic Neural Network), являющаяся модификацией ИНС радиально-базисных функций (RBF) [22-24]. Применение сети PNN позволяет фиксировать плотность концентрации последовательностей информационных единиц обоих классов в признаковом пространстве. Результат верификации «свой» – «чужой» определяется путем сравнения статистических оценок плотности вероятности распределения образов каждого из двух классов. В дополнение к характеристикам плотности вероятности информационных единиц каждого класса предлагается также использовать цену ошибок, как допустимых ошибок 1-го и 2-го рода процедуры верификации.

Решение поставленной задачи. Воспроизведение произвольного текста средствами динамической биометрии любой модальности реализуется совокупностью заученных подсознательных движений, которые преобразуются в электрические сигналы (функции времени). В динамической биометрии сигналы $\mathbf{x}(t)$ разной модальности имеют и различную мерность. В общем случае их следует рассматривать как многомерные:

$$\mathbf{x}(t) = x_1(t), x_2(t), \dots, x_n(t).$$

На этапе предварительной обработки сигналы $\mathbf{x}(t)$ оцифровываются $\mathbf{x}(t) \rightarrow \mathbf{x}(t_i)$, $i = 1, 2, \dots$, масштабируются. Далее из них исключаются длительные паузы, не обусловленные особенностями воспроизведения текста. В голосовой биометрии исключаются также неинформативные с точки зрения распознавания голоса фонемы шипящих звуков.

Отсчеты сигнала $\mathbf{x}(t_i)$, $i = 1, 2, \dots$ рассматриваются как точки метрического пространства E^n , представленные векторами признаков $\mathbf{x}(t_i) = x_1(t_i), x_2(t_i), \dots, x_n(t_i)$, а сам сигнал $\mathbf{x}(t_i)$, – как последовательность $\{\mathbf{x}(t_i)\}_{i=1}^{\infty}$ элементов, представленных векторами признаков: $\mathbf{x}(t_i)$.

Проведенные ранее исследования [11–13], показывают, что индивидуальные особенности динамической биометрии личности в большей степени проявляются не характером воспроизведения одиночных символов и фонем, а морфемно обусловленных последовательно идущими фрагментами текста. Для использования этого феномена последовательность $\{\mathbf{x}(t_i)\}_{i=1}^{\infty}$ расчленяется на фрагменты $\{\mathbf{x}(t_i)\}_{i=1}^r$ одинакового размера по r отсчетов в каждом фрагменте.

Результатом будет новая последовательность $\{\mathbf{y}(t_j)\}_{j=1}^{\infty}$, каждый элемент которой содержит r векторов $\mathbf{x}(t_i)$ исходной последовательности $\{\mathbf{x}(t_i)\}_{i=1}^{\infty}$:

$$\mathbf{y}(t_j)_{j=1}^{\infty} = \{\mathbf{x}(t_i)\}_{i=1}^r.$$

В реальной системе верификации последовательность $\{\mathbf{y}(t_j)\}_{j=1}^{\infty}$ ограничивается n элементами – $\{\mathbf{y}(t_j)\}_{j=1}^n$. При этом каждый вектор $\mathbf{y}(t_j)$ последовательности $\{\mathbf{y}(t_j)\}_{j=1}^n$ будет представлен s -мерным вектором, содержащим $s = n \times r$ компонент:

$$\mathbf{y}(t_j) = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1r} \\ y_{21} & y_{22} & \dots & y_{2r} \\ \dots & \dots & \dots & \dots \\ y_{n1} & y_{n2} & \dots & y_{nr} \end{bmatrix}.$$

В итоге текстонезависимая динамическая биометрия личности будет представлена последовательностью $\{\mathbf{y}(t_j)\}_{j=1}^n$ s -мерных векторов признаков в метрическом пространстве E^n .

В связи с применением в данной работе комплексного использования аппарата ИИС и ИНС PNN представляется целесообразным использовать модель отрицательного отбора (МОО) ИИС [14–20].

Применение МОО ИИС для решения задачи верификации пользователей ИС по принципу «свой»-«чужой» предполагает вначале создание эталона \mathbf{P} «своего» пользователя (в данном случае, в виде конечной последовательности $\mathbf{P} = \{\mathbf{y}(t_j)\}_{j=1}^n$), а затем совокупности m детекторов \mathbf{D} в том же формате $\mathbf{D} = \{\mathbf{d}(t_j)\}_{j=1}^m$, которые должны отличаться от элементов последовательности $\mathbf{P} = \{\mathbf{y}(t_j)\}_{j=1}^n$ на некоторую заданную величину δ_0 .

Соотношение числа элементов эталона \mathbf{P} и числа детекторов \mathbf{D} в общем случае зависит от задачи и вида применяемых детекторов. В данной работе, в связи с использованием ИНС PNN, для классификации «свой» – «чужой» предлагается использовать равное число элементов эталона \mathbf{P} и числа детекторов \mathbf{D} , т. е. $n = m$.

Простейший способ создания детекторов \mathbf{D} состоит из двух фаз. В первой фазе осуществляется случайная генерация кандидатов в детекторы $\hat{\mathbf{D}}$, равномерно распределенных в пространстве признаков E^n . Во второй фазе кандидаты в детекторы $\hat{\mathbf{D}}$ по координатам сопоставляются с векторами эталона \mathbf{P} на основе меры близости Евклида:

$$\delta_{yd} = \sqrt{\sum_{l=1}^s (y_{jl} - d_{jl})^2}.$$

Если $\delta_{yd} > \delta_0$, то кандидат $\hat{\mathbf{d}}_j$ в детекторы $\hat{\mathbf{D}}$ приобретает статус детектора \mathbf{d}_j , в противном случае он уничтожается. Таким образом формируется популяция заданного числа N_d детекторов \mathbf{D} .

В рабочем режиме производится верификация текстонезависимых биометрических данных работающей в ИС неизвестной личности, представленных двумя совокупностями данных (элементов эталона \mathbf{P} и детекторов \mathbf{D}). Верификация указанных данных осуществляется на два класса «свой» – «чужой» с использованием модифицированной PNN-сети [21–24].

ИНС PNN представляет собой параллельную реализацию статистических методов Байеса и ориентирована исключительно на задачи классификации.

Верификацию биометрических данных работающей в ИС личности на два класса «свой» – «чужой» предлагается реализовать на основе формального правила:

- ◆ класс с более плотным распределением в области неизвестного образца, будет иметь преимущество перед другим классом;
- ◆ класс с более высокой ценой ошибки классификации будет иметь преимущество перед другим классом.

В ИНС PNN оценка плотности распределения образцов осуществляется методом Парцена (Parzen). Исходными данными для классификации являются две последовательности равной длины:

- ◆ $\{\mathbf{y}(t_j)\}_{j=1}^n$ – эталонная последовательность «своего»;

♦ $\{\mathbf{d}(t_j)\}_{j=1}^n$ – последовательность детекторов «чужого».

Для каждого элемента последовательностей $\{\mathbf{y}(t_j)\}_{j=1}^n$ и $\{\mathbf{d}(t_j)\}_{j=1}^m$ формируются упрощенные функция Гаусса, которые отличаются от классической отсутствием коэффициента $1/\sigma\sqrt{2\pi}$ перед экспонентой, что позволяет получить максимальное значение функции плотности вероятностей, равное единице:

$$\varphi(\mathbf{y}_j) = \exp\left(-\frac{\|\mathbf{y} - \mathbf{y}_j\|^2}{2\sigma^2}\right); \quad \varphi(\mathbf{d}_j) = \exp\left(-\frac{\|\mathbf{d} - \mathbf{d}_j\|^2}{2\sigma^2}\right),$$

где σ – параметр, задающий ширину функции.

Функции плотностей распределения вероятностей элементов для каждого класса образуются путем суммирования функций Гаусса элементов последовательностей $\{\mathbf{y}(t_j)\}_{j=1}^n$ и $\{\mathbf{d}(t_j)\}_{j=1}^m$:

$$\varphi(\mathbf{Y}) = \sum_{i=1}^{L_k} \exp\left(-\frac{\|\mathbf{y} - \mathbf{y}_i\|^2}{2\sigma^2}\right); \quad \varphi(\mathbf{D}) = \sum_{i=1}^{L_k} \exp\left(-\frac{\|\mathbf{d} - \mathbf{d}_i\|^2}{2\sigma^2}\right),$$

где L_k – объем обучающей выборки каждого класса.

Структура ИНС PNN для решения задачи текстонезависимой верификации личности по динамическим биометрическим параметрам приведена на рис. 1.

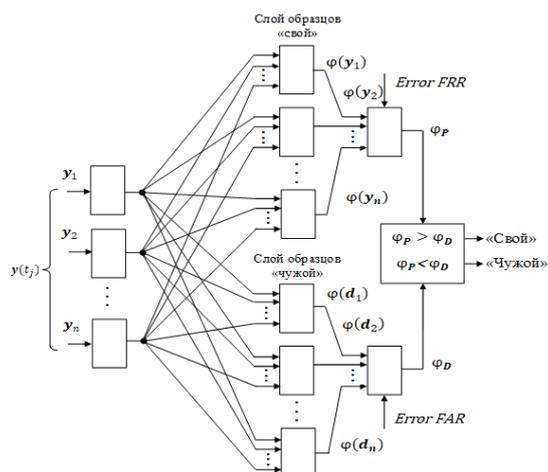


Рис. 1. PNN-сеть для решения задачи верификации личности

Входной слой и слой образов образуют полно связную структуру. На входы сети последовательно поступают s -мерные векторы \mathbf{y}_j последовательности $\mathbf{P} = \{\mathbf{y}_j\}_{j=1}^n$, биометрических данных неизвестной личности. Слой образов представлен двумя группами нейронов «свой» и «чужой» по n нейронов в каждой группе. Веса обеих матриц связей слоя образов определяются значениями компонент соответствующих s -мерных образов входных векторов \mathbf{y}_j и \mathbf{d}_j :

$$\mathbf{W}_y = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1n} \\ y_{21} & y_{22} & \dots & y_{2n} \\ \dots & \dots & \dots & \dots \\ y_{s1} & y_{s2} & \dots & y_{sn} \end{bmatrix}; \quad \mathbf{W}_d = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ d_{s1} & d_{s2} & \dots & d_{sn} \end{bmatrix}.$$

Слой суммирования представлен двумя нейронами, каждый из которых суммирует плотность вероятности своей группы:

$$\varphi^P = \varphi(\mathbf{y}_1), \varphi(\mathbf{y}_2), \dots, \varphi(\mathbf{y}_n); \quad \varphi^D = \varphi(\mathbf{d}_1), \varphi(\mathbf{d}_2), \dots, \varphi(\mathbf{d}_n).$$

Выходной нейрон выполняет функцию дискриминатора величины суммарной плотности вероятности $\varphi = \varphi^P + \varphi^D$ с учетом цены ошибок 1-го рода (FRR) и 2-го рода (FAR) и в итоге производит классификацию работающей личности:

$$\varphi = \begin{cases} \text{при } (\varphi_P + FRR) > (\varphi_D + FAR) - \text{«свой»}; \\ \text{при } (\varphi_P + FAR) < (\varphi_D + FAR) - \text{«чужой»}. \end{cases}$$

(мерности величин $\varphi_P, \varphi_D, FRR, FAR$) для конкретных задач верификации должны быть согласованы).

В ИНС PNN необходимо провести предварительную нормализацию входных векторов \mathbf{y}_j и \mathbf{d}_j :

$$y_j^H = y_j / \sqrt{\sum_{j=1}^n y_j^2}; \quad d_j^H = d_j / \sqrt{\sum_{j=1}^n d_j^2}.$$

Такая операция превращает входные векторы \mathbf{y}_j и \mathbf{d}_j в векторы единичной длины \mathbf{y}_j^H и \mathbf{d}_j^H в пространстве E^n .

Исходя из соответствия метрик входных векторов \mathbf{y}_j , \mathbf{d}_j и весов \mathbf{W}_y , нормализация производится и для весов \mathbf{W}_y

$$w_{ij}^H = w_{ij} / \sqrt{\sum_{j=1}^n w_{ij}^2}.$$

С учетом нормализации и последующих преобразований плотности вероятностей для обеих последовательностей будут выглядеть так:

$$\varphi(\mathbf{y}_j) = \exp\left(\frac{1}{\sigma^2} \sum_{j=1}^n y_j^H \cdot w_{ij}^H - 1\right); \quad \varphi(\mathbf{d}_j) = \exp\left(\frac{1}{\sigma^2} \sum_{j=1}^n d_j^H \cdot w_{ij}^H - 1\right)$$

Функции активности каждого нейрона слоя суммирования определяет значение плотностей вероятностей обоих классов φ_P и φ_D :

$$\begin{aligned} \varphi_P &= \sum_{j=1}^n \exp\left(\frac{1}{\sigma^2} \sum_{j=1}^n y_j^H \cdot w_{ij}^H - 1\right); \\ \varphi_D &= \sum_{j=1}^n \exp\left(\frac{1}{\sigma^2} \sum_{j=1}^n d_j^H \cdot w_{ij}^H - 1\right). \end{aligned}$$

В режиме верификации через обученную сеть PNN непрерывно по n текущих отсчетов $\mathbf{y}(t_j)$ пропускается входная биометрическая последовательность $\{\mathbf{y}(t_j)\}_{j=1}^n$ априори неизвестной личности.

ИНС PNN по итоговым значениям плотности вероятности и ошибок первого и второго рода φ_P и φ_D обоих классов определяет принадлежность входной последовательности $\mathbf{y}(t_j)$ к одному из двух возможных классов «свой» или «чужой».

Принятие верификационного решения, осуществляется, исходя из соотношения совокупной оценки плотностей вероятности φ_P и φ_D с учетом ошибок 1-го и 2-го рода:

$$\mathbf{y}(t_j) \equiv \begin{cases} \text{«свой»}, \text{ если } \varphi_P > \varphi_D; \\ \text{«чужой»}, \text{ если } \varphi_P \leq \varphi_D. \end{cases}$$

Принятие решения «свой» – «чужой» осуществляется в непрерывном режиме в темпе поступлении входной последовательности текстонезависимой биометрии $\{\mathbf{y}(t_j)\}_{j=1}^n$ априори неизвестной личности.

Заключение. Предлагаемый подход к верификации личности работающего оператора ИС позволяет предложить общую схему этой процедуры для существенно различных модальностей динамической биометрии: голоса, рукописи и клавиатурного набора. Реализация такого подхода для биометрии конкретной модальности может отличаться, но общая схема верификации будет сохранена.

Преимуществами предлагаемого подхода являются:

- ◆ возможность текстонезависимого анализа динамической биометрии различной модальности, произвольного объема, содержания и языка;
- ◆ возможность принятия верификационного решения в непрерывном режиме в темпе поступления входной последовательности текстонезависимой биометрии различной модальности априори неизвестной личности;
- ◆ перспектива, вследствие развития информационных технологий, повышения точности работы системы верификации путем увеличения размерности слоя образцов PNN-сети.
- ◆ возможность использования истории анализа результатов верификации реальных пользователей для последующей более точной настройки ИНС PNN в части оптимального соотношения примеров «своего», «чужого» и выбора цены ошибок 1-го и 2-го рода.

Относительным недостатком предлагаемого подхода в текущее время является необходимость реализации нейронной сети PNN большой размерности. В перспективе же развития электронных и программных технологий ИС этот недостаток будет быстро нивелироваться.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Матвеев Ю.Н.* Технологии биометрической идентификации личности по голосу и другим модальностям // Вестник МГТУ им. Н.Э. Баумана. Серия «Приборостроение». – 2012. – № 2. – С. 46-61.
2. *Campbell W., Assaleh K., Broun C.* Speaker recognition with polynomial classifiers // IEEE Trans. Speech Audio Process. – 2002. – Vol. 10, No. 4. – P. 205-212.
3. *Брюхомицкий Ю.А.* Иммунологическая модель текстонезависимой голосовой идентификации личности // Известия ЮФУ. Технические науки. – 2022. – № 2 (226). – С. 6-13.
4. *Анисимова Э.С.* Идентификация онлайн-подписи с помощью оконного преобразования Фурье и радиального базиса // Компьютерные исследования и моделирование. – 2014. – Т. 6, № 3. – С. 357-364.
5. *Jain A.K., Friederike D.G., Connel S.D.* On-line signature verification // Pattern Recognition. – 2002. – Vol. 35 (12). – P. 2963-2972.
6. *Plamondon R., Srihari S.* On-line and Off-line Handwriting Recognition: A Comprehensive Survey // IEEE Trans. PAMI. – 2000. – Vol. 22 (1). – P. 63-84.
7. *Брюхомицкий Ю.А., Абрамов Е.С.* Верификация рукописных текстов с использованием иммунологических и нейросетевых технологий // Вопросы защиты информации. – М.: Научные и информационные издания ФГУП «НТИ оборонного комплекса «КОМПАС». – С. 31-37.
8. *Брюхомицкий Ю.А.* Иммунологический метод верификации рукописи с использованием векторного представления данных // Известия ЮФУ. Технические науки. – 2016. – № 9 (182). – С. 50-57.
9. *Мазниченко Н.И., Гвозденко М.В.* Анализ возможностей систем автоматической идентификации клавиатурного почерка // Вестник Национального технического университета «Харьковский политехнический институт». Серия «Информатика и моделирование». – 2008. – Вып. № 24. – С. 77-82.
10. *Скубицкий А.В.* Анализ применимости метода реконструкции динамических систем в системах биометрической идентификации по клавиатурному почерку // Инфокоммуникационные технологии. – 2008. – Т. 6, № 1. – С. 51-53.
11. *Брюхомицкий Ю.А.* Клавиатурный мониторинг на основе иммунологического клонирования // Безопасность информационных технологий. – 2016. – № 4 (40). – С. 5-11.
12. *Брюхомицкий Ю.А.* Иммунологический метод клавиатурного мониторинга // Вестник Брестского государственного технического университета. Физика, математика, информатика. – 2016. – № 5 (101). – С. 28-32.
13. *Брюхомицкий Ю.А.* Цепочный метод клавиатурного мониторинга // Известия ЮФУ. Технические науки. – 2009. – № 11. – С. 135-145.
14. *Dasgupta D.* Artificial Immune Systems and Their Applications, Ed. – Springer-Verlag, 1999.
15. *De Castro L.N., Timmis J.I.* Artificial Immune Systems: A New Computational Intelligence Approach, London: Springer-Verlag, 2000. – 357 p.
16. *Hofmeyr S. and Forrest S.* Architecture for an Artificial Immune System // Evolutionary Computation. – 2000. – 8 (4). – P. 443-473.
17. *Specht D.F.* Probabilistic neural networks // Neural Networks. – 1990. – No. 3. – P. 109-118.
18. *Чернышев Ю.О., Венцов Н.Н., Григорьев Г.В.* Искусственные иммунные системы: обзор и современное состояние // Программные продукты и системы. – 2014. – № 4. – С. 136-142.

19. Зайцев С.А., Субботин С.А. Обобщенная модель искусственной иммунной системы // Proceedings. – Berlin–Heidelberg: Springer-Verlag, 2003. – Ser. LNCS 2723. – P. 195-206.
20. Литвиненко В.И., Дидык А.А., Захарченко Ю.А. Компьютерная система для решения задач классификации на основе модифицированных иммунных алгоритмов // Автоматика. Автоматизация. Электротехнические комплексы и системы. – 2008. – № 2 (22).
21. Нейронные сети: полный курс. – 2-е изд.: пер. с англ. – М.: Издательский дом «Вильямс», 2006. – 1104 с.
22. Spech D.F. Probabilistic neural networks // Neural Networks. – 1990. – No. 3. – P. 109-118.
23. Каллан Р. Основные концепции нейронных сетей. – М.: Вильямс, 2001. – 291 с.
24. Брехомитский Ю.А. Верификация динамических биометрических параметров личности на основе вероятностной нейронной сети // Известия ЮФУ. Технические науки. – 2020. – № 5 (215). – С. 52-60.

REFERENCES

1. Matveev Yu.N. Tekhnologii biometricheskoy identifikatsii lichnosti po golosu i drugim modal'nostyam [Technologies for biometric personal identification by voice and other modalities], *Vestnik MGTU im. N.E. Baumana. Seriya «Priborostroenie»* [Bulletin of MSTU im. N.E. Bauman. Series "Instrument making"], 2012, No. 2, pp. 46-61.
2. Campbell W., Assaleh K., Broun C. Speaker recognition with polynomial classifiers, *IEEE Trans. Speech Audio Process*, 2002, Vol. 10, No. 4, pp. 205-212.
3. Bryukhomitskiy Yu.A. Immunologicheskaya model' tekstonezavisimoy golosovoy identifikatsii lichnosti [Immunological model of text-independent voice identification of a person], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2022, No. 2 (226), pp. 6-13.
4. Anisimova E.S. Identifikatsiya onlayn-podpisi s pomoshch'yu okonnogo preobrazovaniya Fur'e i radial'nogo bazisa [Identification of an online signature using a windowed Fourier transform and a radial basis], *Komp'yuternye issledovaniya i modelirovanie* [Computer Research and Modeling], 2014, Vol. 6, No. 3, pp. 357-364.
5. Jain A.K., Friederike D.G., Connel S.D. On-line signature verification, *Pattern Recognition*, 2002, Vol. 35 (12), pp. 2963-2972.
6. Plamondon R., Srihari S. On-line and Off-line Handwriting Recognition: A Comprehensive Survey, *IEEE Trans. PAMI* 2000, Vol. 22 (1), pp. 63-84.
7. Bryukhomitskiy Yu.A., Abramov E.S. Verifikatsiya rukopisnykh tekstov s ispol'zovaniem immunologicheskikh i neyrosetevykh tekhnologiy [Verification of handwritten texts using immunological and neural network technologies], *Voprosy zashchity informatsii* [Issues of information protection]. Moscow: Nauchnye i informatsionnye izdaniya FGUP «NTTS oboronnoogo kompleksa «KOMPAS», pp. 31-37.
8. Bryukhomitskiy Yu.A. Immunologicheskiy metod verifikatsii rukopisi s ispol'zovaniem vektornogo predstavleniya dannykh [Immunological method of manuscript verification using vector representation of data], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2016, No. 9 (182), pp. 50-57.
9. Maznichenko N.I. Gvozdenko M.V. Analiz vozmozhnostey sistem avtomaticheskoy identifikatsii klaviaturnogo pocherka [Analysis of the capabilities of systems for automatic identification of keyboard handwriting], *Vestnik Natsional'nogo tekhnicheskogo universiteta «Khar'kovskiy politekhnicheskii institut». Seriya «Informatika i modelirovanie»* [Bulletin of the National Technical University "Kharkiv Polytechnic Institute". Series "Informatics and Modeling"], 2008, Issue № 24, pp. 77-82.
10. Skubitskiy A.V. Analiz primenimosti metoda rekonstruktsii dinamicheskikh sistem v sistemakh biometricheskoy identifikatsii po klaviaturnomu pocherku [Analysis of the applicability of the method for reconstructing dynamic systems in biometric identification systems based on keyboard handwriting], *Infokommunikatsionnye tekhnologii* [Infocommunication technologies], 2008, Vol. 6, No. 1, pp. 51-53.
11. Bryukhomitskiy Yu.A. Klaviaturnyy monitoring na osnove immunologicheskogo klonirovaniya [Keyboard monitoring based on immunological cloning], *Bezopasnost' informatsionnykh tekhnologiy* [Information technology security], 2016, No. 4 (40), pp. 5-11.
12. Bryukhomitskiy Yu.A. Immunologicheskiy metod klaviaturnogo monitoringa [Immunological method of keyboard monitoring], *Vestnik Brestskogo gosudarstvennogo tekhnicheskogo universiteta. Fizika, matematika, informatika* [Bulletin of Brest State Technical University. Physics, mathematics, computer science], 2016, No. 5 (101), pp. 28-32.

13. Bryukhomitskiy Yu.A. Tsepochnyy metod klaviaturnogo monitoringa [Chain method of keyboard monitoring], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2009, No. 11, pp. 135-145.
14. Dasgupta D. Artificial Immune Systems and Their Applications, Ed. Springer-Verlag, 1999.
15. De Castro L.N., Timmis J.I. Artificial Immune Systems: A New Computational Intelligence Approach, London: Springer-Verlag, 2000, 357 p.
16. Hofmeyr S. and Forrest S. Architecture for an Artificial Immune System, *Evolutionary Computation*, 2000, 8 (4), pp. 443-473.
17. Specht D.F. Probabilistic neural networks, *Neural Networks*, 1990, No. 3, pp. 109-118.
18. Chernyshev Yu.O., Ventsov N.N., Grigor'ev G.V. Iskusstvennyye immunnye sistemy: obzor i sovremennoe sostoyanie [Artificial immune systems: review and current state], *Programmnye produkty i sistemy* [Software products and systems], 2014, No. 4, pp. 136-142.
19. Zaytsev S.A., Subbotin S.A. Obobshchennaya model' iskusstvennoy immunnoy sistemy [Generalized model of an artificial immune system], *Proceedings*. Berlin–Heidelberg: Springer-Verlag, 2003. Ser. LNCS 2723, pp. 195-206.
20. Litvinenko V.I., Didyk A.A., Zakharchenko Yu.A. Komp'yuternaya sistema dlya resheniya zadach klassifikatsii na osnove modifitsirovannykh immunnykh algoritmov [Computer system for solving classification problems based on modified immune algorithms], *Avtomatika. Avtomatizatsiya. Elektrotekhnicheskie komplekсы i sistemy* [Automation. Automation. Electrical complexes and systems], 2008, No. 2 (22).
21. Neyronnye seti: polnyy kurs [Neural networks: a complete course]. 2nd ed.: transl. from engl. Moscow: Izdatel'skiy dom «Vil'yams», 2006, 1104 p.
22. Spech D.F. Probabilistic neural networks, *Neural Networks*, 1990, No. 3, pp. 109-118.
23. Kallan R. Osnovnyye kontseptsii neyronnykh setey [Basic concepts of neural networks]. Moscow: Vil'yams, 2001, 291 p.
24. Bryukhomitskiy Yu.A. Verifikatsiya dinamicheskikh biometricheskikh parametrov lichnosti na osnove veroyatnostnoy neyronnoy seti [Verification of dynamic biometric personality parameters based on a probabilistic neural network], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2020, No. 5 (215), pp. 52-60.

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

Брюхомицкий Юрий Анатольевич – Южный федеральный университет; e-mail: bryukhomitskiy@sfedu.ru; г. Таганрог, Россия; тел.: 88634371905; кафедра безопасности информационных технологий; с.н.с.; доцент.

Bryukhomitskiy Yuriy Anatoly – Southern Federal University; e-mail: bryukhomitskiy@sfedu.ru; Taganrog, Russia; phone: +78634371905; the department of security in data processing technologies; senior researcher; associate professor.

УДК 004.93

DOI 10.18522/2311-3103-2024-2-25-31

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

МОДЕЛЬ АЛГОРИТМА ПОТОКОВОЙ МАРКИРОВКИ ШИРОКОФОРМАТНОГО ИЗОБРАЖЕНИЯ

В настоящей статье предложен алгоритм обработки широкоформатного изображения для применения в системах, работающих в режиме реального масштаба времени с высокоскоростным потоком видеоданных. Вопрос предварительной обработки изображения, его кластеризации, сегментации и маркировки имеет особую важность для систем обработки видеопотока высокого разрешения в режиме реального времени. Кроме того, при реализации таких алгоритмов остро стоит вопрос минимизации затрат вычислительных ресурсов программируемых логических интегральных схем (ПЛИС), на которых происходит непосредственное развертывание алгоритмов потоковой обработки изображений. Минимальное потребление ресурсов обеспечивают однопроходные алгоритмы маркировки, в которых отсутствует необходимость буферизации изображения, что имеет особую важность при обработке широкоформатного изображения высокого разрешения. Однако, при реализации одиночного прохода изображения через систему обработки может происходить создание множества дополнительных маркеров подлежащих дальнейшему объединению, особенно при анализе изображения с большим разрешением. Созданные дополнитель-

ные маркеры требуют увеличения требований количеству используемых ячеек памяти на ПЛИС. Описываемый в статье алгоритм потоковой маркировки широкоформатного изображения высокого разрешения позволяет производить маркировку потокового видеоизображения высокого разрешения снижая вероятность создания дополнительных меток подлежащих дальнейшему объединению. Суть улучшения алгоритма относительно стандартного однопроходного состоит в добавлении к сканирующей маске дополнительных элементов, которые позволяют избежать ситуации появления различных меток, соответствующих одному объекту, что позволяет при минимальном увеличении объема используемой памяти на ПЛИС избежать дублирования меток и перерасхода памяти устройства. Проведено моделирование алгоритма для реализации на ПЛИС с помощью инструмента Xilinx System Generator for DSP в связке со средой для модельно-ориентированного проектирования (МОП) Matlab Simulink. Представлены результаты работы алгоритма на изображениях, полученных с высокоскоростной линейной камеры TELEDYNE DALSA LA-CC-04K05B-00- с использованием мезонина Integre Technologies LLC FMC-200-A, а также отладочной платы Xilinx ZYNQ Ultrascale+ MPSoC ZCU106.

Маркировка; сегментация; связанные области; ПЛИС; потоковая обработка.

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

MODEL OF ALGORITHM FOR STREAMING LABELING OF WIDE FORMAT IMAGES

This article proposes a wide-format image processing algorithm for use in systems operating in real time with a high-speed video data stream. The issue of image preprocessing, its clustering, segmentation and labeling is of particular importance for systems for processing high-resolution video streams in real time. In addition, when implementing such algorithms, there is an urgent issue of minimizing the cost of computational resources of programmable logic integrated circuits (FPGAs), on which the direct deployment of streaming image processing algorithms takes place. Minimal resource consumption is ensured by single-pass marking algorithms, which eliminate the need for image buffering, which is especially important when processing high-resolution wide-format images. However, when implementing a single pass of an image through the processing system, many additional markers may be created that are subject to further combining, especially when analyzing images with high resolution. The additional markers created require an increase in the requirements for the number of usable memory cells on the FPGA. The algorithm for streaming high-resolution wide-format images described in the article makes it possible to label high-resolution streaming video images, reducing the likelihood of creating additional tags that need to be further combined. The essence of improving the algorithm relative to the standard one-pass one is to add additional elements to the scanning mask, which avoid the appearance of different labels corresponding to the same object, which allows, with a minimal increase in the amount of memory used on the FPGA, to avoid duplication of labels and overuse of device memory. The algorithm was simulated for implementation on an FPGA using the Xilinx System Generator for DSP tool in conjunction with the Matlab Simulink environment for model-based design (MBD). The results of the algorithm are presented on images obtained from a high-speed linear camera TELEDYNE DALSA LA-CC-04K05B-00-R using the Integre Technologies LLC FMC-200-A mezzanine, as well as the Xilinx ZYNQ Ultrascale+ MPSoC ZCU106 development board.

Labeling; segmentation; connected areas; FPGA; stream processing.

Введение. При разработке систем компьютерного зрения фундаментальной задачей является выбор и реализация алгоритма предобработки и маркировки изображения. Данный алгоритм не должен вносить излишних временных задержек, для обеспечения работы в режиме реального времени и должен обеспечивать корректную маркировку объектов на изображении исключая создание «лишних» маркеров или объединения различных объектов. При работе с изображениями высокого разрешения обостряется проблема создания дополнительных маркеров, последующее объединение которых влечет увеличение времени обработки изображения и количества используемых вычислительных ресурсов ПЛИС. Множество существующих различных алгоритмов обработки изображений можно сгруппировать следующим образом: двухпроходный (классический) алгоритм [1], алгоритм многократного сканирования [2], алгоритм параллельной обработки, алгоритм трассировки контура [3–4] и алгоритм однократного прохода [5].

Согласно сравнительному исследованию алгоритмов маркировки связанных областей [6–8] следует, что однократный алгоритм маркировки связанных областей имеет самый низкий относительно прочих алгоритмов уровень использования ресурсов. Что позволяет развёртывать системы на меньших, более дешёвых ПЛИС.

Постановка задачи. Решаемая задача заключается в модификации однократного алгоритма маркировки изображения для обработки высокоскоростного потока видеоданных высокого разрешения. В качестве входных данных использовалось изображение семян пшеницы, полученное с высокоскоростной линейной камеры TELEDYNE DALSA LA-CC-04K05B-00-R при пролете семян мимо объектива камеры. В результате работы алгоритма должно быть получено маркированное изображение с количеством меток равным количеству объектов на изображении. Решаемую задачу можно разбить на два этапа. Первый этап - предварительная обработка изображения [9–10]. Данный этап включает в себя разделение полноцветного изображения на цветовые каналы, бинаризация изображения, применение процедур фильтрации, эрозии и сглаживания. Второй этап заключается в осуществлении маркировки подготовленного изображения путем однократного прохода с минимизацией возможности создания меток подлежащих дальнейшему объединению.

Модель алгоритма потоковой маркировки изображения. Модель алгоритма обработки изображения показана на рис. 1. На вход алгоритма поступает поток пикселей изображения и сигнал разрешения

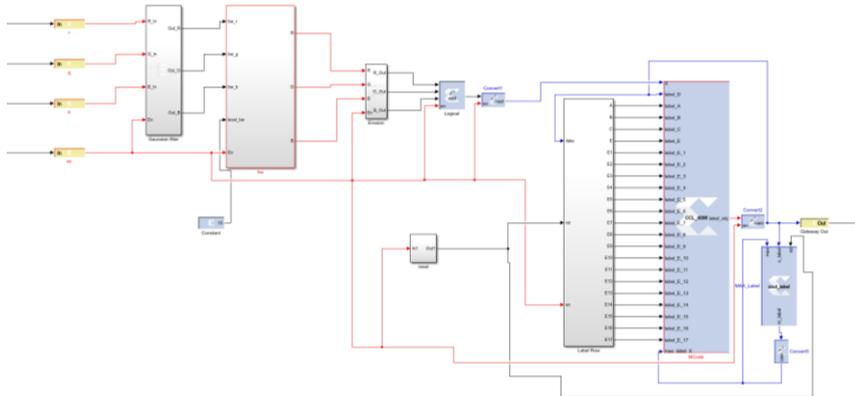


Рис. 1. Модель алгоритма обработки изображения

Первым шагом производится размытие изображения путем применения фильтра Гаусса [11–12]. Результат работы фильтра передается в блок бинаризации, где, согласно задаваемому порогу, происходит перевод изображения в бинарное представление. Бинарное изображение передается в следующий блок, реализующий процедуру эрозии. Все описанные процедуры выполняются в каждом цветовом канале, после процедуры эрозии из трех бинарных изображений красного, зеленого и синего каналов соответственно, создается суммарное бинарное изображение. Описанная процедура позволяет сгладить границы объектов и убрать шум на изображении.

Имея подготовленное бинарное изображение можно производить непосредственную маркировку объектов. Процедура маркировки производится путем сканирования изображения маской показанной на рис. 2.

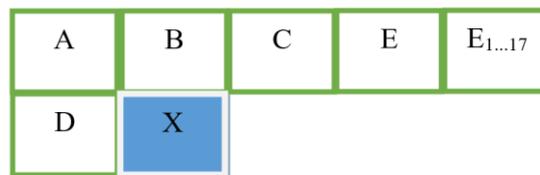


Рис. 2. Сканирующая маска

В классическом однопроходном алгоритме сканирующая маска содержит и использует 4 элемента(пикселя) A,B,C,D для маркировки пикселя X. Однако в виду высокого разрешения обрабатываемого изображения такая маска порождает ситуацию когда одному объекту присваивается несколько маркеров. Для обработки широкоформатного изображения в предлагаемом алгоритме добавлены элементы E-E₁₇ которые позволяют минимизировать вероятность создание различных меток, принадлежащих одному объекту, и подлежащих объединению. Это не приводит к увеличению числа используемых ячеек памяти поскольку для работы алгоритма буферизируется одна строка изображения целиком. Процедура хранения меток предыдущей строки изображения организована с помощью сдвигового регистра, а извлечение необходимых меток производится путем снятия информации с соответствующих триггеров регистра.

На выходе процедуры маркировки получается размеченное бинарное изображение. Пикселям, принадлежащим фону, присваивается нулевая метка, а пиксели, принадлежащие объектам, имеют метку с номером объектов в порядке их обнаружения. Таким образом, максимальное значение метки равно количеству обнаруженных объектов на изображении.

Результат. В качестве входных данных для модели использовался поток пикселей с линейной камеры TELEDYNE DALSA LA-CC-04K05B-00-R с длиной строки 4096 пикселей. Результаты работы вышеописанной модели показан на рис. 3.

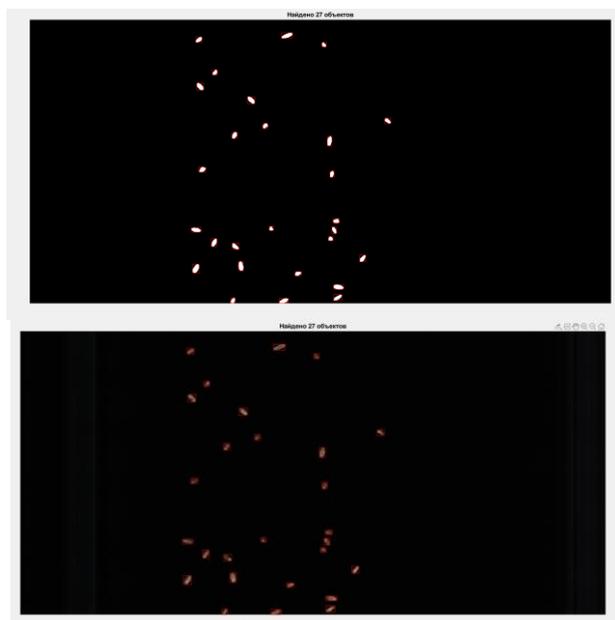


Рис. 3. Результаты работы модели

Полученное маркированное бинарное изображение позволяет производить дальнейшую обработку изображения в зависимости от назначения системы компьютерного зрения.

Результаты оценки использования ресурсов ПЛИС для развертывания описанного алгоритма представлен табл. 1

Таблица 1

BRAMs	DSPs	LUTs	Registers
0	0	599 (0.2%)	33468 (6%)

Выводы. Предлагаемый алгоритм потоковой маркировки широкоформатного изображения позволяет производить разметку высокоскоростного потока пикселей с минимизацией возможности создания меток подлежащих дальнейшему объединению. Показа-

на минимизация объема дополнительной памяти ПЛИС, требуемой для реализации представленного алгоритма, относительно стандартного однопроходного. Применение показанной оптимизации приводит к уменьшению количества дополнительных меток, принадлежащих одному объекту, что может существенно повлиять на объем ресурсов, требуемых для реализации алгоритма маркировки на ПЛИС. Показана возможность реализации алгоритма на ПЛИС и произведена оценка требуемых для реализации ресурсов. Для Xilinx ZYNQ Ultrascale+ MPSoC ZCU106 количество требуемых ресурсов для реализации представленного алгоритма следующее: таблиц LUT 599 (0,2% от всего количества), регистров 33468 (6: от всего количества). Результаты оценки подтверждают выгоду использования однопроходных алгоритмов, а модернизированное сканирующее окно и алгоритм выбора метки обеспечивают сохранение качества маркировки с увеличением размера и разрешения обрабатываемого изображения.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Rosenfeld A., Pfaltz J.L.* Sequential operations in digital picture processing // Journal of the ACM (JACM). – 1966. – 13.4. – P. 471-494.
2. *Haralick R.* Some neighborhood operations // Real Time/Parallel Computing Image Analysis. – 1981. – P. 11-35.
3. *Chang F. and Chen C.-J.* A Component-Labeling Algorithm Using Contour Tracing Technique // Proceedings of the 7th International Conference on Document Analysis and Recognition (ICIDAR 2003). – 2003. – 0-7695-1960-1/03.
4. *Chang F., Chen C.-J., Lu C.* A linear-time component-labeling algorithm using contour tracing technique // Computer Vision and Image Understanding. – 2004. – Vol. 93, No. 2. – P. 206-220.
5. *Bailey D., Johnston C.* Single pass connected components analysis // Image and Vision Computing New Zealand. – 2008. – P. 282-287.
6. *Johnston C., Bailey D.* FPGA implementation of a single pass connected components algorithm // Electronic Design, Test and Applications. – 2008. – P. 228-231.
7. *Trein J., Schwarzbacher A., Hoppe B., Noffz K. H., Trenchel T.* Development of a FPGA Based Real-Time Blob Analysis Circuit // Irish Systems and Signals Conference. – 2007. – P. 121-126.
8. *Walczyk R., Armitage A., Binnie T.D.* Comparative study on connected component labeling algorithms for embedded video processing systems // IPCV. – 2010. – Vol. 10. – URL: researchgate.net/publication/220808599_Comparative_Study_on_Connected_Component_Labeling_Algorithms_for_Embedded_Video_Processing_Systems.
9. *Yaman S. et al.* Image and Video Processing Applications Using Xilinx System Generator // 2019 7th International Symposium on Digital Forensics and Security (ISDFS). – IEEE, 2019. – P. 1-5.
10. *Raut N.P., Gokhale A.V.* FPGA implementation for image processing algorithms using xilinx system generator // IOSR Journal of VLSI and Signal Processing (IOSR-JVSP). – 2013. – Vol. 2, No. 4. – P. 26-36.
11. *Ляхов П.А., Валуева М.В.* Применение сглаживающих фильтров для очистки от шума изображений в оттенках серого // Наука. Инновации. Технологии. – 2015. – № 3. – С. 37-50.
12. *Sghaier A. et al.* FPGA implementation of filtered image using 2D Gaussian filter // International Journal of Advanced Computer Science and Applications. – 2016. – Vol. 7, No. 7.
13. *Bakumenko A. et al.* Connected component labeling algorithm in streaming image processing with FPGAs // SPIE Future Sensing Technologies 2021. – SPIE, 2021. – Vol. 11914. – P. 328-340.
14. System Generator for DSP User Guide, UG640, 14.1. April 2012.
15. Zynq UltraScale+ Devices Register Reference UG1087 (v1.8) July 30, 2021.
16. *Bakumenko A. et al.* Crop seed classification based on a real-time convolutional neural network // SPIE Future Sensing Technologies. – SPIE, 2020. – Vol. 11525. – P. 654-667.
17. *Ковалев А.В., Бакуменко А.Н.* Алгоритм маркировки связанных областей при потоковой обработке изображения // Инженерный вестник Дона. – 2022. – № 4 (88). – С. 191-201.
18. *Бахчевников В.В., Деркачев В.А., Бакуменко А.Н.* Способ использования средств быстрого прототипирования для реализации сверточной нейронной сети на ПЛИС // Известия ЮФУ. Технические науки. – 2020. – № 3 (213). – С. 146-156.
19. *Деркачев В.А., Бахчевников В.В., Бакуменко А.Н.* Классификатор изображений семян сельскохозяйственных культур с использованием сверточной нейронной сети // Известия ЮФУ. Технические науки. – 2020. – № 4 (214). – С. 32-39.
20. *Бахчевников В.В., Деркачев В.А., Бакуменко А.Н.* Метод разработки искусственных нейронных сетей для прототипирования на ПЛИС с помощью средств Xilinx System Generator и Matlab // Компьютерные и информационные технологии в науке, инженерии и управлении" КомТех-2020". – 2020. – С. 462-468.

REFERENCES

1. Rosenfeld A., Pfaltz J.L. Sequential operations in digital picture processing, *Journal of the ACM (JACM)*, 1966, 13.4, pp. 471-494.
2. Haralick R. Some neighborhood operations, *Real Time/Parallel Computing Image Analysis*, 1981, pp. 11-35.
3. Chang F. and Chen C.-J. A Component-Labeling Algorithm Using Contour Tracing Technique, *Proceedings of the 7th International Conference on Document Analysis and Recognition (ICIDAR 2003)*, 2003. 0-7695-1960-1/03.
4. Chang F., Chen C.-J., Lu C. A linear-time component-labeling algorithm using contour tracing technique, *Computer Vision and Image Understanding*, 2004, Vol. 93, No. 2, pp. 206-220.
5. Bailey D., Johnston C. Single pass connected components analysis, *Image and Vision Computing New Zealand*, 2008, pp. 282-287.
6. Johnston C., Bailey D. FPGA implementation of a single pass connected components algorithm, *Electronic Design, Test and Applications*, 2008, pp. 228-231.
7. Trein J., Schwarzbacher A., Hoppe B., Noffz K.H., Trenchel T. Development of a FPGA Based Real-Time Blob Analysis Circuit, *Irish Systems and Signals Conference*, 2007, pp. 121-126.
8. Walczyk R., Armitage A., Binnie T.D. Comparative study on connected component labeling algorithms for embedded video processing systems, *IPCV*, 2010, Vol. 10. Available at: researchgate.net/publication/220808599_Comparative_Study_on_Connected_Component_Labeling_Algorithms_for_Embedded_Video_Processing_Systems.
9. Yaman S. et al. Image and Video Processing Applications Using Xilinx System Generator, *2019 7th International Symposium on Digital Forensics and Security (ISDFS)*. IEEE, 2019, pp. 1-5.
10. Raut N.P., Gokhale A.V. FPGA implementation for image processing algorithms using xilinx system generator, *IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)*, 2013, Vol. 2, No. 4, pp. 26-36.
11. Lyakhov P.A., Valueva M.V. Primenenie sglazhivayushchikh fil'trov dlya ochistki ot shuma izobrazheniy v ottenkakh serogo [Application of anti-aliasing filters for removing noise from images in various grayscales], *Nauka. Innovatsii. Tekhnologii* [Science. Innovation. Technologies], 2015, No. 3, pp. 37-50.
12. Sghaier A. et al. FPGA implementation of filtered image using 2D Gaussian filter, *International Journal of Advanced Computer Science and Applications*, 2016, Vol. 7, No. 7.
13. Bakumenko A. et al. Connected component labeling algorithm in streaming image processing with FPGAs, *SPIE Future Sensing Technologies 2021*. SPIE, 2021, Vol. 11914, pp. 328-340.
14. System Generator for DSP User Guide, UG640, 14.1. April 2012.
15. Zynq UltraScale+ Devices Register Reference UG1087 (v1.8) July 30, 2021.
16. Bakumenko A. et al. Crop seed classification based on a real-time convolutional neural network, *SPIE Future Sensing Technologies*. SPIE, 2020, Vol. 11525, pp. 654-667.
17. Kovalev A.V., Bakumenko A.N. Algoritm markirovki svyazannykh oblastey pri potokovoy obrabotke izobrazheniya [Algorithm for marking connected areas during streaming image processing], *Inzhenernyy vestnik Dona* [Engineering Bulletin of the Don], 2022, No. 4 (88), pp. 191-201.
18. Bakhchevnikov V.V., Derkachev V.A., Bakumenko A.N. Sposob ispol'zovaniya sredstv bystrogo prototipirovaniya dlya realizatsii svertochnoy neyronnoy seti na PLIS [Method of using fast prototyping tools for implementing a convolutional neural network on FPGA], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2020, No. 3 (213), pp. 146-156.
19. Derkachev V.A., Bakhchevnikov V.V., Bakumenko A.N. Klassifikator izobrazheniy semyan sel'skokhozyaystvennykh kul'tur s ispol'zovaniem svertochnoy neyronnoy seti [Classifier of images of agricultural seeds using a convolutional neural network], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2020, No. 4 (214), pp. 32-39.
20. Bakhchevnikov V.V., Derkachev V.A., Bakumenko A.N. Metod razrabotki iskusstvennykh neyronnykh setey dlya prototipirovaniya na PLIS s pomoshch'yu sredstv Xilinx System Generator i Matlab [Method for developing artificial neural networks for prototyping on FPGAs using Xilinx System Generator and Matlab], *Kompyuternye i informatsionnye tekhnologii v nauke, inzhenerii i upravlenii* "KomTekh-2020" [Computer and information technologies in science, engineering and management" KomTech - 2020"], 2020, pp. 462-468.

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

Бакуменко Алексей Николаевич – Южный федеральный университет; e-mail: baku@sfedu.ru; г. Таганрог, Россия; тел.: +79886031853; кафедра радиотехнических и телекоммуникационных систем; ассистент.

Деркачев Владимир Александрович – e-mail: vderkachev@sfedu.ru; тел.: +79614154733; кафедра радиотехнических и телекоммуникационных систем; ассистент.

Бахчевников Валентин Владимирович – e-mail: bahchevnikov@sfedu.ru; тел.: +79518289271; кафедра радиотехнических и телекоммуникационных систем; старший преподаватель.

Лобач Владимир Тихонович – e-mail: vtlobach@sfedu.ru; тел.: +79185233270; кафедра радиотехнических и телекоммуникационных систем; доцент.

Bakumenko Alexey Nikolaevich – Southern Federal University; e-mail: baku@sfedu.ru; Taganrog, Russia; phone: +79886031853; the department of radio engineering & telecommunication systems; assistant professor.

Derkachev Vladimir Aleksandrovich – e-mail: vderkachev@sfedu.ru; phone +79614154733; the department of radio engineering & telecommunication systems; assistant professor.

Bakhchevnikov Valentin Vladimirovich – e-mail: bahchevnikov@sfedu.ru; phone: +79518289271; the department of radio engineering & telecommunication; senior lecturer.

Lobach Vladimir Tikhonovich – e-mail: vtlobach@sfedu.ru; phone: +79185233270; the department of radio engineering & telecommunication systems; associate professor.

УДК 519.688

DOI 10.18522/2311-3103-2024-2-31-38

С.Л. Беляков, А.В. Исаев

МЕТОД ПОДДЕРЖКИ УСТОЙЧИВОСТИ СЕТИ ЭЛЕКТРОСНАБЖЕНИЯ НА ОСНОВЕ ГЕОИНФОРМАЦИОННОЙ МОДЕЛИ

Рассматривается задача управления распределением энергетической мощности на территории, покрытой интеллектуальной энергосетью. Целью управления является стабилизация потока энергии при наличии внешних воздействий, вызванных изменением состояния окружающей среды. Присутствие уязвимостей обусловлено природой энергетических сетей и при определенном стечении обстоятельств провоцирует аномалии энергоснабжения. Воздействие факторов внешней среды разнообразно по содержанию, не позволяя уверенно прогнозировать актуальные угрозы. Описываются геоинформационные модели, использующие представление знаний образами. Ее использование дает возможность оценивать актуальность известных угроз. Концептуально образ включает в себя центр и допустимые преобразования центра в некотором контексте. Рассматривается случай, когда угроза нарушения работы интеллектуальной сети оценивается трансформированием образа в заданную область пространства, в которой размещается интеллектуальная сеть. Принципиальной особенностью предлагаемого подхода является оценка реализуемость явления или события в заданной области пространства. Операция переноса ситуации заставляет учитывать особенности топологии заданной области. Основную роль начинают играть не столько атрибуты ситуации, как параметры порождающей ее инфраструктуры. Особенностью предлагаемого подхода следует считать перенос смыслового контекста, который представлен допустимыми преобразованиями образа. Программная функция трансформирования связывается со слоем картографического изображения. Для заданного объекта, входящего в прецедент-оригинал, определяется область размещения. Ее граница определяется свойствами объекта. Если размер области размещения позволяет сконструировать объект соответствующего класса, этот объект создается. Оценивается достоверность результата путем применения экспертного знания о качестве объектов рассматриваемого класса. Перечисленные действия выполняются не только над геометрией пространственного объекта. Аналогично концепции образного представления геометрии используются временные и семантические атрибуты. Формирование списка угроз при заданном состоянии внешней среды составляет суть управления устойчивости. Анализируются особенности алгоритмизации процедуры трансформации образов. Приводится способ оценки достоверности трансформирования. Применение предложенного подхода перспективно для интеллектуальных систем энергоснабжения, поведение которых сложным образом зависит от факторов внешней среды.

Геоинформационное моделирование; интеллектуальная сеть; управление энергопотреблением; интеллектуальные ГИС; представление знаний.

S.L. Belyakov, A.V. Isaev

METHOD OF SUPPORTING THE STABILITY OF THE POWER SUPPLY NETWORK BASED ON A GEOINFORMATION MODEL

The article considers the problem of controlling the distribution of energy power in an area covered by an intelligent energy network. The management objective is to stabilize the energy flow in the presence of external influences caused by changes in the surrounding environment. Vulnerabilities in the system are inherent due to the nature of energy networks and under certain circumstances can lead to anomalies in energy supply. External environmental factors vary in content, making it difficult to confidently predict current threats. Geoinformation models utilizing image-based knowledge representation are described. Their use enables the assessment of the relevance of known threats. Conceptually, an image comprises a center and permissible transformations of that center within a certain context. The case is considered where the threat to the functioning of the intelligent network is assessed by transforming the image into a specified area of space where the intelligent network is located. The key feature of the proposed approach is the evaluation of the feasibility of an event occurring in a given space. The operation of transferring the situation requires consideration of the topology of the specified area. The attributes of the generating infrastructure become more significant than the attributes of the situation itself in this approach. A distinctive feature of the proposed approach is the transfer of semantic context represented by permissible transformations of the image. The software transformation function is linked to a layer of cartographic representation. For a given object in the original precedent, its placement area is determined, with the boundary being defined by the object's properties. If the size of the placement area allows for the construction of an object of the corresponding class, that object is created. The credibility of the result is evaluated by applying expert knowledge about the quality of objects of the considered class. The listed actions are performed not only on the geometry of spatial objects but also on their temporal and semantic attributes, akin to the concept of image-based representation of geometry. Forming a list of threats given a specific state of the external environment constitutes the essence of stability management. The features of algorithmizing the image transformation procedure are analyzed, and a method for assessing the credibility of transformation is provided. The application of the proposed approach holds promise for intelligent energy supply systems, whose behavior is intricately linked to external environmental factors.

Geoinformation modelling; smart grid; energy management; intelligent GIS; knowledge representation, transfer learning.

Введение. Устойчивое функционирование интеллектуальных сетей представляет интерес как в практическом, так и в теоретическом отношениях [1, 2]. Реально функционирующее оборудование в системах поставки энергии способно лишь частично противостоять внешним отрицательным воздействиям. К ним относят воздействия окружающей природной среды, техносферы, сбои и отказы оборудования и программ самих интеллектуальных сетей. Проблема устойчивости должна решаться, в частности, развитием системы управления в отношении ее интеллектуализации. «Умные» устройства, приборы, транспорт и города рассматриваются в настоящее время как бесспорная перспектива. С теоретической точки зрения, достижение устойчивости требует адекватных моделей объектов управления. Особый интерес представляют геоинформационные модели, отражающие явления и процессы реального мира в картографическом виде.

Интеллектуальные геоинформационные модели представляют собой особый класс образно-знаковых моделей, включающих в себя карты, планы, схемы систем реального мира и знания об их поведении [3]. Часть знаний закладывается в географические электронные карты уже на этапе их изготовления. Богатое содержание карт повышает эффективность принимаемых решений. Однако, этих знаний недостаточно для работы в условиях неопределенности. В частности, обеспечение устойчивости требует оценки текущего состояния сети, его классификацию и выработку управляющего воздействия [4]. Такая парадигма управления требует особой концептуальной модели для описания и оценки пространственных ситуаций. Специфика в том, чтобы эффективно использовать богатый опыт принятия решений в условиях неполноты, неопределенности и неоднозначности данных о состоянии интеллектуальной среды. Программный инструментарий геоинформационных систем (ГИС) позволяет пользователю фиксировать большой объем детальных данных о прецедентах принятия решения. Тем не менее, данные не несут идеи о том, по какому пути должна пойти выработка решений. Этот вопрос исследуется в настоящей работе.

Особое внимание при использовании геоинформационных моделей, с нашей точки зрения, следует уделять переносу содержащихся в них знаний (transfer learning). В отличие от моделей машинного обучения предварительная обученность моделей ГИС не связывается с предварительной обработкой наборов данных (data sets) и установлением некоторых разумных начальных значений параметров модели. Поведение геоинформационных моделей сильно зависит от топологии пространства. Например, представляется невозможным перенести опыт устранения аварийных ситуаций систем энергоснабжения из заболоченной области на зону песчаной пустыни. Однако, эксперты-аналитики способны «вообразить» то, как могла бы выглядеть известная ситуация в совершенно другой местности. Эта особенность чрезвычайно полезна и привлекает к исследованию того, как следует представлять опыт, какова логика переноса опыта и как можно аргументировать то или иное решение.

Целью настоящей работы является описание геоинформационной модели переноса знаний об аномальных ситуациях, влияющих на качество работы интеллектуальной сети.

Постановка задачи. Под свойством устойчивости интеллектуальной сети будем понимать ее способность противостоять угрозам нарушения нормального функционирования. Природа угроз различна, они обладают широким набором варьируемых параметров. Как следствие, объективно существует неопределенность того, какие из угроз актуальны на текущий момент времени. Существенно важна также пространственная локализация угроз: интеллектуальная сеть являются территориально рассредоточенными объектами. В задаче прогнозирования угроз и выработки защитных мер всегда будет важен опыт наблюдения событий и явлений, ранее приводивших к нарушениям работы интеллектуальной сети. Возможность переноса опыта применения защитных мер с одних ситуаций на другие имеет особую значимость для пространственных объектов. Возможность с помощью ГИС решать задачу «как ЭТО будет выглядеть ТАМ» с этой точки зрения является перспективным инструментом обеспечения устойчивости. Геоинформационной моделью для поддержания устойчивости интеллектуальных сетей будем представлять как множество объектов:

$$M = \{ C, B, G, P \}, \quad (1)$$

где C – множество контекстов, в которых могут решаться задачи пространственного анализа; B – множество пространственных областей, в которых размещается интеллектуальная сеть, G – набор процедур переноса опыта изучения прецедентов опасных ситуаций; P – множество прецедентов, наблюдавшихся ранее при работе интеллектуальной сети. С помощью геоинформационной модели должна решаться задача:

$$\begin{cases} W(G(B_w, P_a, C_w)) \rightarrow \max, \\ B_w \subseteq B, P_a \subseteq P, C_w \subseteq C. \end{cases} \quad (2)$$

Здесь W – критерий достоверности переноса опыта, B_w – рабочая область пространства, в которой выполняется анализ угроз, P_a – список актуальных угроз, C_w – контекст анализа. Чтобы обеспечить решение задачи (2), необходимо:

- ◆ определить форму представления знаний о прецеденте, которые позволят переносить их в заданную пространственную область;
- ◆ представить процедуру переноса знания;
- ◆ обосновать критерий оценки результата переноса.

Если решены перечисленные вопросы, система управления интеллектуальной сети выполняет следующее:

1. Принимает в картографическом виде информацию об аномальном событии или процессе. На соответствующие слои электронной карты наносятся необходимые картографические объекты.
2. Формируется область анализа на карте, которая обновляется соответственно имеющимся оперативным данным о внешней среде.
3. Выбираются прецеденты известных ситуаций, которые потенциально могут произойти в анализируемой области.

4. Прецеденты, возможность повторения которых достаточно высока, считаются угрозами в ближайшем будущем.

5. Вырабатываются управляющие сигналы, направленные на снижение вероятности реализации угроз и минимизации ущерба в случае наступления.

Обзор публикаций. Построение и использование геоинформационных моделей для управления интеллектуальными сетями применяется уже много лет [5, 6]. Данный подход рассматривается как выработка управляющих воздействий на основе многокритериальной оценки состояния объекта [7]. Особенностью данного механизма является комплексная оценка территорий и накопление прецедентов аварийных ситуаций. Управление строится на результатах статистических обобщений непрерывно накапливаемых данных о состоянии среды, в которой функционирует интеллектуальная сеть [8]. Обращает на себя внимание структурирование прецедентов, минимально отличающее их от общей классификации объектов ГИС [9]. С одной стороны, это удешевляет обслуживание информационной базы ГИС, с другой – понижает достоверность пространственного анализа воздействий на интеллектуальная сеть. Имеется настоятельная необходимость изучать более сложные информационные структуры прецедентов [10].

Case based reasoning (CBR) относится к давно применяемым методам повторного использования опыта [11]. В его основе лежит применение метрики близости прецедентов и гипотезы компактности. В соответствии с этим, решения в сходных ситуациях должны быть близки по смыслу. Данный принцип эффективен для анализа пространственных ситуаций, однако сложность подбора выразительной и адекватной меры сходства заставляет обращаться к другим принципам [12]. В частности, требует исследования вопрос о том, какие модификации CBR должны использоваться для переноса опыта из одной области пространства в другую.

Учитывая то, что ГИС ориентированы на непрерывное накопление геопространственных данных и относятся к системам больших данных (big data), много публикаций посвящено применению методов анализа данных в системах [13]. Значительный интерес проявляется к моделям машинного обучения [14, 15]. Данный подход целесообразен тогда, когда большие объемы данных характеризуют непосредственно процесс выработки решений. Практика обеспечения устойчивости для интеллектуальной сети такими данными не обладает. По содержанию это данные о состоянии внешней среды и единичных логически обоснованных решениях. Здесь необходимо исследовать особый путь обучения интеллектуальной системы, использующий обширные данные о топологии пространства, в котором размещается интеллектуальная сеть [16].

Передача знаний является элементов исследований глобальной проблемы добычи и использования знаний. Работы этого направления исследуют модели представления имеющегося знания для его передачи заинтересованной стороне. Примером является работа [17]. В качестве основы используется модель пропускной способности (Throughput Model), включающая в себя восприятие, информацию, выработку суждений и принятие решений. Перечисленные элементы рассматриваются как интеллектуальные операции, основанные на знаниях. Передача знаний рассматривается как поддержка алгоритмического процесса модели пропускной способности. Уровень концептуального представления знаний в работах данного направления соответствует организационному, когда цикл управления интеллектуальной сетью реализуется экспертами. Таким образом, исследования в этом направлении концентрируются на системе понятий, которые частично охватывают интеллектуальность автоматизированной системы управления интеллектуальная сеть.

Исследования интеллектуальной сети с точки зрения кибербезопасности проводят анализ угроз и уязвимостей, которые способны породить инциденты сбоев и отказов [18, 19]. Это помогает систематизировать знания об опасностях и выработке мер защиты. При этом в работах по данному вопросу не рассматривается применение геоинформационных моделей.

Представление прецедентов пространственных ситуаций. Известным методом применения опыта является CBR [11]. Если задано пространственное положение двух прецедентов p_a и p_b , их сходство оценивается принятой метрикой расстояния $d(p_b, p_a)$. Минимальное расстояние соответствует максимальному сходству. Процедура CBR в геоинформационной модели трудно применима по двум причинам:

- ◆ расстояние $d(p_b, p_a)$ требует наличия на карте двух прецедентов. В случае, когда точное расположение одного из них не известно, усложняет ситуацию;
- ◆ для картографического пространства метрика $d(p_b, p_a)$ нелинейна. Для любого объекта карты должен выполняться набор топологических соотношений, определяемых целью создания карты.

Анализируя возможные пути решения задачи (2), будем учитывать, что всякий прецедент является подмножеством пространственных объектов (K) и отношений (R), представленных в геоинформационной модели:

$$p_i \in P \Rightarrow p_i \subseteq K \cup R, p_i \neq \emptyset . \quad (3)$$

Каждый прецедент размещен в области на карте B_{p_a} . Перенос прецедента означает определение новой области \tilde{B}_{p_a} , в которой будут построены картографические объекты и отношения нового прецедента \tilde{p}_i . В общем случае все элементы пары:

$$(\tilde{p}_i, K \cup R) \quad (4)$$

должны быть топологически согласованы. Содержательно это означает, что новое положение любого элемента прецедента \tilde{p}_i должно в точности повторять пространственное положение аналогичного элемента исходного прецедента p_i . Из-за того, что области B_{p_a} и \tilde{B}_{p_a} покрывают различные множества картографических объектов, точная топологическая согласованность становится чрезвычайно маловероятной. При этом практика пространственного анализа показывает, что расположение элементов прецедента \tilde{p}_i не является жестко фиксированным относительно друг друга и окружающих объектов. Возможные вариации положения определяются смыслом используемого контекста $C_w \subseteq C$. Однако, для этого должны быть известны границы таких изменений положения. Необходимость использования знаний о допустимых преобразованиях приводит к следующей концептуальной модели [20]. Любой прецедент описывается как образ:

$$\langle c, h_1(c), h_2(c), \dots, h_M(c) \rangle, \quad (5)$$

в котором c – это центр образа, а $h_i(c)$ – допустимые преобразования образа. Центр включает в себя картографическое представление прецедента в заданном контексте. Любое допустимое преобразование является набором картографических объектов, отражающих изменение значений свойств центра, но сохраняющих смысл прецедента. В качестве примера на рис. 1 показаны два экземпляра допустимых преобразований для прецедента «повреждение электрического кабеля», заданного точечным объектом. На рис. 1,а изображен образ прецедента как точечный объект на карте (условно обозначен крестом) в его области B_{p_a} . Показаны все значимые объекты, характеризующие аварийную ситуацию и принятое впоследствии решение (здесь не показано). Допустимым преобразованием h_1 является положение места размещения обрыва провода для свободного пространства относительно ближайших зданий (рис. 1,б). Этот фактор существенно повлиял на принятое впоследствии решение. Допустимым преобразованием h_2 считается возможное положение обрыва на двух соседних линиях (рис. 1,с). Картографическое изображение каждого допустимого преобразования включает ссылки на объекты, определяющие их геометрию (показаны треугольниками). Такие ссылки используются для объяснения геометрических характеристик допустимого преобразования.

Концептуальная модель (5) определяет представление знаний, необходимое для последующего переноса прецедентов. Существенно то, что топологическая согласованность в отличие от (4) должна контролироваться для пары:

$$(\tilde{p}_i, \tilde{B}_{p_i}), \quad (6)$$



Рис. 1. Пример допустимых преобразований

Это не только снижает трудоемкость задания и проверки ограничений, но и делает перенос прецедента более интеллектуальным, т.е. адаптированным к наперед не известным особенностям местности.

Подчеркнем, что модель (5) представляется картографическими объектами, не включает привязку к ним знаний известных моделей: продукционной, семантической, фреймовой и других.

Перенос прецедентов. Процедура переноса прецедента в виде (5) заключается в том, чтобы в заданной области \tilde{B}_{p_a} при наличии областей допустимых преобразований разместить соответствующие элементы прецедента \tilde{p}_a . Алгоритмические особенности этой процедуры следующие:

- ◆ каждое допустимое преобразование размещается на отдельном слое как набор картографических объектов. Их типы и значения атрибутов должны быть аргументами функции отображения каждого элемента прецедента в заданную локацию;
- ◆ каждому слою ставится в соответствие программная функция, интерпретирующая объекты данного слоя и возвращающая как результат область для размещения элемента прецедента. Если область пуста, прецедент переместить не удастся;
- ◆ каждому слою ставится в соответствие программная функция, конструирующая элемент нового прецедента в непустой области размещения;
- ◆ завершающим этапом является вызов функции оценки достоверности результата.

Ключевым для функции трансформирования является оценка достоверности. Результат может включать в себя только часть элементов, соответствующих исходному прецеденту. Например, преобразование h_2 на рис. 1 может не трансформироваться, если два кабеля расположены слишком близко друг к другу. Однако, эта особенность, по мнению эксперта, может сохранять возможность повторного применения известного решения, быть допустимым преобразованием результата трансформирования. Следовательно, возможную оценку результата трансформирования можно получить также соответственно модели (5). Образ может быть представлен, как

$$\langle F_{TR}(p_a, \tilde{B}_{p_a}), h_1(F_{TR}(p_a, \tilde{B}_{p_a})), \dots, h_M(F_{TR}(p_a, \tilde{B}_{p_a})) \rangle,$$

где $F_{TR}(p_a, \tilde{B}_{p_a})$ – функция трансформирования прецедента в заданную область. От эксперта потребуется указать значение оценки достоверности для каждого допустимого преобразования:

$$(\nu_1, h_1(F_{TR}(p_a, \tilde{B}_{p_a})), \dots, (\nu_M, h_M(F_{TR}(p_a, \tilde{B}_{p_a})))) \\ \nu_r \in [0, \nu_{max}].$$

Алгоритм оценки достоверности заключается в последовательном переборе слоев с изображениями h_k до тех пор, пока не будет найдено наиболее близкое допустимое преобразование.

Заключение. Предложенный в данной работе подход к управлению устойчивостью отличается использованием особого представления опыта эксплуатации интеллектуальной сети. Образы пространственных ситуаций, порожденных аномальными природными явлениями и событиями, используются для воспроизведения этих ситуаций в указанной пространственной области. Управление интеллектуальной сетью приобретает свойство

интеллектуальности, поскольку получает механизм адаптации к ранее не известным условиям. В данном случае это топология среды функционирования интеллектуальная сеть. База знаний о прецедентах аномальных ситуаций, заложенная в геоинформационную модель, используется для построения списка угроз в указанной пространственной локации. Дальнейшим направлением исследований по данной теме мы считаем поиск новых форм представления прецедентов для трансформирования опыта.

Проведенное исследование позволило сделать следующие выводы. Во-первых, устойчивость интеллектуальной сети существенно зависит от факторов природной среды. Их динамика и непредсказуемость делает полезными интеллектуальные геоинформационные модели. Они включают в себя процедуры и знания для воспроизведения процессов и явлений в разных локациях. Во-вторых, модель представления знаний обладает спецификой, обусловленной необходимостью трансформирования прецедентов. Концептуально это описание допустимых преобразований объекта, события или явления, сохраняющих его смысловую целостность. Дальнейшим направлением исследований по данной теме мы считаем поиск новых форм представления прецедентов для трансформирования опыта.

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. Fotopoulou Maria, Rakopoulos Dimitrios, Petridis Stefanos, Drosatos Panagiotis. Assessment of smart grid operation under emergency situations, *Energy*, 2024, Vol. 287, 129661.
2. Tan Zhiwei, Li Zhuo. Digital twins for sustainable design and management of smart city buildings and municipal infrastructure, *Sustainable Energy Technologies and Assessments*, 2024, Vol. 64, 103682.
3. Veenendaal Bert, Brovelli Maria Antonia, Wu Lixin. Cloud/web mapping and geoprocessing services – Intelligently linking geoinformation, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, Vol. 114, pp. 243-244.
4. Renugadevi N., Saravanan S., Naga Sudha C.M. IoT based smart energy grid for sustainable cities, *Materials Today: Proceedings*, 2023, Vol. 81, Part 2, pp. 98-104.
5. Alhamwi A., Medjroubi W., Medjroubi T., and Agert C. GIS-based urban energy systems models and tools: Introducing a model for the optimisation of flexibilisation technologies in urban areas, *Applied Energy*, 2017, Vol. 191, pp. 1-9.
6. Zhao Fang, Fashola Olushola I., Olarewaju Tolulope I., Onwumere Ijeoma. Smart city research: A holistic and state-of-the-art literature review, *Cities*, 2021, Vol. 119, 103406.
7. Mardania A., Zavadskas E., and Zakuana K. A review of multi-criteria decision-making applications to solve energy management problems: Two decades from 1995 to 2015, *Renewable and Sustainable Energy Reviews*, 2017, Vol. 71, pp. 216-256.
8. Bhatt J., Shah V., and Jan O. An instrument at engineer's review on smart grid: Critical applications and parameters, *Renewable and Sustainable Energy Reviews*, 2014, Vol. 40, pp. 2017-20139.
9. rówczyńska ., Skiba ., Sztubecka ., Bazan-Krzywoszańska A., Kazak J.K., Gajownik P. Scenarios as a tool supporting decisions in urban energy policy: The analysis using fuzzy logic, multi-criteria analysis and GIS tools, *Renewable and Sustainable Energy Reviews*, 2021, Vol. 137, 110598.
10. Sylla Abdourahim. Thierry Coudert, Laurent Geneste, A Case-Based Reasoning (CBR) approach for Engineer-To-Order systems performance evaluation, *IFAC-PapersOnLine*, 2021, Vol. 54, Issue 1, pp. 717-722.
11. Kolodner J. Case-Based Reasoning, Morgan Kaufmann, San Mateo, 1993.
12. Zhuoa H.H., Kambhampati Z. Model-lite planning: Case-based vs. model-based approaches, *Artificial Intelligence*, 2017, 246, pp. 1-21.
13. Mostafa Noha, Haitham Saad Mohamed Ramadan, Elfarouk Omar. Renewable energy management in smart grids by using big data analytics and machine learning, *Machine Learning with Applications*, 2022, Vol. 9, 100363.
14. Jing Qianzhen, Yan Jing, Wang Yanxin, He Ruixin, Lu Lei. A novel differentiable neural network architecture automatic search method for GIS partial discharge pattern recognition, *Measurement*, 2022, Vol. 195, 111154.
15. Li Yunqin, Yabuki Nobuyoshi. Tomohiro Fukuda, Integrating GIS, deep learning, and environmental sensors for multicriteria evaluation of urban street walkability, *Landscape and Urban Planning*, 2023, Vol. 230, 104603.

16. Salenbien R., Wack Y., Baelmans M., Blommaert M. Geographically informed automated non-linear topology optimization of district heating networks, *Energy*, 2023, Vol. 283, 128898.
17. Rodgers Waymond, Cardenas Jesus A., Gemoets Leopoldo A., Sarfi Robert J. A smart grids knowledge transfer paradigm supported by experts' throughput modeling artificial intelligence algorithmic processes, *Technological Forecasting and Social Change*, 2023, Vol. 190, 122373.
18. Jean Paul A. Yaacoub, Javier Hernandez Fernandez, Hassan N. Noura, Ali Chehab. Security of Power Line Communication systems: Issues, limitations and existing solutions, *Computer Science Review*, 2021, Vol. 39, 100331.
19. Dennis Amelia, Weston Dale, Amlôt i chard, Arnold Andreas, Carbon Danielle, Carter Holly. The role of pre-incident information and responder communication in effective management of casualties, including members of vulnerable groups, during a decontamination field exercise, *International Journal of Disaster Risk Reduction*, 2023, Vol. 94, 103806.
20. Belyakov S., Belyakova M., Bozhenyuk A., Rozenberg I. The features of generations of solutions by intellectual information systems, In: *Advances in Intelligent Systems and Computing*. Springer International Publishing Switzerland, 2016, Vol. 451, pp. 221-229.

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

Беляков Станислав Леонидович – Южный федеральный университет; e-mail: sbelyakov@sfedu.ru, beliacov@yandex.ru; г. Таганрог, Россия; тел.: +79185109607; д.т.н.; профессор кафедры информационно-аналитических систем безопасности имени профессора Берштейна Леонида Самойловича.

Исаев Александр Викторович – e-mail: alis@sfedu.ru, shura.isaev.99@mail.ru; тел.: +79518352784; аспирант кафедры информационно-аналитических систем безопасности имени профессора Берштейна Леонида Самойловича.

Belyakov Stanislav Leonidovich – Southern Federal University; e-mail: sbelyakov@sfedu.ru, beliacov@yandex.ru; Taganrog, Russia; phone: +79185109607; dr of eng. sc.; professor of the department of information and analytical security systems named after professor Leonid Samoilovich Bershtein.

Isaev Alexander Viktorovich – Southern Federal University; e-mail: alis@sfedu.ru, shura.isaev.99@mail.ru; phone: +79518352784; postgraduate student of the department of information and analytical security systems named after professor Leonid Samoilovich Bershtein.

УДК 004.89

DOI 10.18522/2311-3103-2024-2-38-50

Д.Ю. Кравченко

МОДЕЛЬ ОНТОЛОГИИ ЗНАНИЙ ДЛЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ОБРАБОТКИ И АНАЛИЗА ТЕКСТОВ*

Статья посвящена решению научной проблемы создания верхнеуровневого описания модели онтологии знаний для интеллектуальных систем обработки и анализа текстов на естественном языке, построенной на основе оригинальной компонентной архитектуры, обеспечивающей необходимый уровень детализации спецификаций анализируемой текстовой информации. Актуальность данной задачи обусловлена необходимостью развития теоретических основ построения информационных моделей семантических зависимостей внутри текстов на естественном языке. Автором даны определения основным терминам исследуемой предметной области. Представлена формализованная постановка решаемой задачи. Проблема «информационного взрыва», причиной возникновения которой стал экспоненциальный рост объемов цифровой информации, привела к ситуации, когда до 95% информационного потока содержит неструктурированные данные. В подобных условиях, крайне актуальной становится задача создания эффективных интеллектуальных систем поиска и приобретения знаний, в том числе, интеллектуальных систем обработки и анализа текстов на естественном языке. Научным направлением решения этой частной задачи является Text Mining (ТМ) – раскопка знаний в текстовой информации. В качестве примера прикладной задачи использования приобретенных знаний, в данном исследовании, рассматривается

* Исследование выполнено за счет гранта Российского научного фонда № 22-71-10121, <https://rscf.ru/project/22-71-10121/> в Южном федеральном университете.

значимая проблема информационной поддержки процессов предупреждения и/или ликвидации последствий чрезвычайных ситуаций. В данной задаче исходными данными являются потоки текстовых сообщений (новостной информации, отчетов о техническом состоянии техногенных объектов, информации о природных явлениях и т.п.), поступающих в центры принятия решений, а на выходе формируются прогностические оценки и/или конкретные инструкции относительно оценки ситуации и предпринимаемых действий определенными специалистами. Одной из причин, сдерживающих развитие интеллектуальных систем обработки и анализа текста для решения задач поиска, приобретения и использования знаний, является недостаточно высокий уровень эффективности моделей и алгоритмов, обеспечивающих комплексное решение описанных выше задач искусственного интеллекта с учетом особенностей семантики и контекста.

Онтология знаний; обработка и анализ текстов; семантика; информационная поддержка; чрезвычайные ситуации; поддержка принятия решений; структурирование информации.

D.Yu. Kravchenko

KNOWLEDGE ONTOLOGY MODEL FOR INTELLIGENT TEXT PROCESSING AND ANALYSIS SYSTEMS

The article is devoted to solving the scientific problem of creating a top-level description of a knowledge ontology model for intelligent systems for processing and analyzing texts in natural language, built on the basis of an original component architecture that provides the necessary level of detail in the specifications of the analyzed text information. The relevance of this task is due to the need to develop the theoretical foundations for constructing information models of semantic dependencies within texts in natural language. The author gives definitions to the main terms of the subject area under study. A formalized definition of the problem being solved is presented. The problem of the "information explosion," which was caused by the exponential growth in the volume of digital information, has led to a situation where up to 95% of the information flow contains unstructured data. In such conditions, the task of creating effective intelligent systems for searching and acquiring knowledge, including intelligent systems for processing and analyzing texts in natural language, becomes extremely urgent. The scientific direction for solving this particular problem is Text Mining (TM) – the excavation of knowledge in text information. As an example of the applied task of using acquired knowledge, this study examines the significant problem of information support for the processes of preventing and/or eliminating the consequences of emergency situations. In this task, the initial data are streams of text messages (news information, reports on the technical condition of man-made objects, information about natural phenomena, etc.) arriving at decision-making centers, and the output is formed by predictive assessments and/or specific instructions regarding the assessment situations and actions taken by certain specialists. One of the reasons hindering the development of intelligent text processing and analysis systems for solving problems of searching, acquiring and using knowledge is the insufficiently high level of models and algorithms efficiency that provide a comprehensive solution to the above-described problems of artificial intelligence, taking into account the peculiarities of semantics and context.

Knowledge ontology; text processing and analysis; semantics; information support; emergency situations; decision support; information structuring.

Введение. Повышенный интерес исследователей к проблеме поиска и приобретения знаний при обработке и анализе текстовой информации привёл к появлению значительного числа определений основных терминов в данной предметной области. Следствием такой терминологической рассогласованности стала проблема неопределенности и нечеткости при описании базовых понятий: информация, данные и знания [1]. В контексте Text Mining (TM) под априорной информацией будем понимать входные текстовые документы, включающие в себя наборы неструктурированных или слабо структурированных данных. Текстовый документ является основной лингвистической единицей естественного языка. Построение модели представления текста позволяет закодировать семантические характеристики информации в виде вектора, который в дальнейшем применяется для решения различных прикладных задач [2, 3].

Задача TM – поиск и приобретение, а затем использование знаний. Знания (англ. knowledge) – это сложная иерархия элементов ценной информации с выявленными зависимостями и закономерностями между фактами, событиями, явлениями и процессами [4]. Ценность информации определяется на основе расчета вероятностных оценок достижения

цели решаемой прикладной задачи до и после получения определенной информации [5]. *Поиск знаний* (англ. *knowledge retrieval*) – информационный процесс возврата информации к структурированной форме. Под *приобретением знаний* (англ. *knowledge acquisition*) для текстовой информации, после приведения ее к некоторой структуре на этапе поиска, будем понимать систематизацию полученных структурированных знаний [2–5].

Результат последовательных процессов поиска и приобретения знаний определяется достоверностью, надежностью, релевантностью, интерпретируемостью нового знания. Достижимость данных характеристик знаний напрямую связана с проблемой снижения уровня *информационной неопределенности* – нехватки информации для решения поставленных прикладных задач [6].

В качестве примера прикладной задачи использования приобретенных знаний, в данном исследовании, рассматривается значимая проблема информационной поддержки процессов предупреждения и/или ликвидации последствий чрезвычайных ситуаций. В данной задаче исходными данными являются потоки текстовых сообщений (новостной информации, отчетов о техническом состоянии техногенных объектов, информации о природных явлениях и т.п.), поступающих в центры принятия решений, а на выходе формируются прогностические оценки и/или конкретные инструкции относительно оценки ситуации и предпринимаемых действий определенными специалистами [7].

Технологии анализа и обработки текстов на естественном языке (NLP) предъявляют ряд требований к разработке интеллектуальных систем, обеспечивая тем самым конфигурируемость данных инструментов. Особенности каждой конфигурации интеллектуальной системы анализа и обработки текста для решения задач поиска, приобретения и использования знаний зависят от имеющихся пересечений определенных разделов лингвистики [8], а именно: технологии построения и анализа слов (морфология); технологии построения и анализа предложений (синтаксис); технологии содержательного анализа и смысловой оценки (семантика).

1. Задачи поиска, приобретения и использования знаний при анализе и обработке текстов. Представим указанные в названии пункта задачи поиска, приобретения и использования знаний в виде процесса извлечения знаний (*knowledge extraction*). Данный процесс в самой общей форме представляет собой последовательность следующих этапов, показанных на рис. 1: поступление корпуса текста на вход интеллектуальной системы; решение задач предобработки текста; применение оператора преобразования полученной после предобработки входных данных текстовой информации.



Рис. 1. Процесс извлечения знаний в общем виде

Формально данное отображение запишем в следующем виде:

$$F_{\text{extraction}}: [Doc_x, x = 1 \dots T] \rightarrow \text{knowledge}, \quad (1)$$

где $F_{\text{extraction}}$ – оператор преобразования полученной после предобработки входных данных текстовой информации; Doc_x – множество текстовых документов; knowledge – знания, необходимые для решения последующих, иерархически вышестоящих задач, например, для решения задачи информационной поддержки процессов предупреждения и/или ликвидации последствий чрезвычайных ситуаций.

Одним из *способов представления* полученных знаний является построение модели онтологии, описывающей предметную область в виде множества концептов (понятий) $(K_i, i = 1 \dots N$, где N – количество концептов) с заданной на нём системой связей $(R_q, q = 1 \dots M$, где M – количество связей), являющихся по своей сути несимметричными семантическими отношениями. Таким образом, элементом подобной модели является кортеж длины три: $\langle k_i, r_q, k_j \rangle, i \neq j$. Система понятий, подходящая для реализации машинного вывода в контексте решения прикладных задач использования приобретенных знаний, подразумевает применение таких категорий семантических отношений, как, например, «эквивалентность», «класс – подкласс», «часть – целое» и другие [9, 10].

Отметим, что тексты на естественном языке в явном или неявном виде содержат в себе компоненты смысла, имеющие в своей основе структуру описанных выше кортежей, имитирующих знания человека и позволяющих строить информационные модели с практически неограниченными возможностями масштабирования. Именно процессы построения и обработки подобных информационных моделей (онтологий, графов знаний, моделей мира) лежат в основе методов поиска, приобретения и использования знаний при анализе и обработке текста.

Описанный выше способ представления данных делит процессы поиска и приобретения знаний на этапы *извлечения концептов* (concept extraction) и *извлечения отношений* между ними (relation extraction). В интеллектуальных информационных системах происходит комбинирование данных этапов. Способ комбинирования определяет тип конкретной интеллектуальной информационной системы (ИИС). Так, например, *открытые* ИИС извлекают все возможные отношения между концептами в корпусе текста, а *закрытые* – только отношения из ранее заданной выборки концептов.

Задача *поиска знаний* в интеллектуальных системах обработки и анализа текстов на естественном языке в большей степени связана с качественным построением синтаксической схемы текста на основе применения парсера. С этой точки зрения наиболее важными будем считать процедуры структурирования и фильтрации текстовой информации, которые позволят определить основные смысловые элементы текста, такие как, например, *ключевые слова* (key words).

2. Постановка задачи. Извлечение ключевых слов является фундаментальной задачей в *обработке естественного языка* (Natural Language Processing, NLP) и включает в себя выявление и извлечение наиболее релевантных и значимых слов или фраз из заданного текста. Парсеры играют ключевую роль в этом процессе, анализируя синтаксическую структуру текста и помогая выявлять ключевые компоненты, представляющие важные понятия или темы.

1. *Синтаксический анализ структуры для извлечения ключевых слов.* Для выполнения извлечения ключевых слов с использованием синтаксического анализа структуры, можно сосредоточиться на конкретных синтаксических единицах, таких как именные фразы (NPs) [11] и глагольные фразы (VPs). Эти фразы часто являются хорошими кандидатами на роль ключевых слов, так как обычно содержат важную информацию о субъекте, объекте или действии в предложении. Например, рассмотрим предложение: "The swift fox jumps over the lazy dog." С помощью синтаксического анализа структуры будут выделены следующие фразы:

- ◆ именные фразы (NPs): "The swift fox", "the lazy dog";
- ◆ глагольная фраза (VP): "jumps over".

Из извлеченных фраз можно выделить следующие ключевые слова: "swift fox" (быстрая лиса); "lazy dog" (ленивая собака); "jumps over" (перепрыгивает через), как существенные компоненты данного предложения.

2. *Синтаксический анализ зависимостей для извлечения ключевых слов.* Синтаксический анализ зависимостей помогает определить отношения между субъектом, глаголом и объектом, а также другие существенные синтаксические зависимости, которые вносят вклад в общий смысл предложения. Ключевые слова могут быть извлечены из этих зависимостей, причем предпочтение отдается словам, несущим значительный семантический вес и играющим важные роли в структуре предложения [12].

В контексте извлечения ключевых слов текст преобразуется в граф, где вершины представляют собой возможные ключевые слова, а ребра – их отношения. Взаимосвязь между ключевыми фразами-кандидатами может быть определена по тому, как часто они встречаются вместе или насколько семантически близки.

Предположим, что строится ориентированный граф $G = (V, E)$, где V – множество вершин, а E – множество ребер. Оценка или важность вершины определяется как [11, 12]:

$$S(V_i) = (1 - d) + d * \sum_{j \in In(V_i)} \frac{1}{|Out(V_j)|} S(V_j), \quad (2)$$

где $In(V_i)$ – это набор вершин, которые указывают на V_i , а $Out(V_i)$ – это набор вершин, на которые указывает V_i . При этом d – это коэффициент затухания, который устанавливается в диапазоне от 0 до 1.

В рамках решения задачи поиска знаний в интеллектуальных системах обработки и анализа текстов на естественном языке данное исследование предполагает создание дополнительного фильтра на выходе парсера, что позволит извлечь смысловую часть предложения из полученной синтаксической схемы текстовой информации для дальнейшего использования в процессах приобретения знаний.

При решении задач приобретения знаний необходимо сформулировать постановку и условия задачи, равно как и само понятие решения. Пусть существует определенный подход S , позволяющий на множестве K анализируемых объектов знаний из множества P разных предметных областей атрибутам (характеристикам) X ставить в соответствие значения (величины) Y , которые приводят к оптимальной обработке разнородных распределенных информационных ресурсов R на основе применения методов M . При этом обработка информационных ресурсов должна проводиться с учетом контекста H , что является необходимым условием повышения эффективности информационного процесса приобретения знаний (функции) F . Таким образом, Y является достоверным выводом, а S – решающим подходом, построенным на основе комплекса методов и правдоподобных рассуждений. Очевидно, что в данном случае эффективность вывода напрямую связана с интеллектуальностью метода анализа и обработки имеющихся информационных ресурсов. Полученный вывод представляет собой систему проанализированных отношений и соответствий, приводящих к обнаружению неявных зависимостей и закономерностей между атрибутами X на множестве объектов знаний K . Идентифицированные новые зависимости и закономерности между объектами знаний являются основой для приобретения нового знания [4].

Реализация моделей и методов приобретения знаний на основе использования интеллектуальных систем существенно изменило аппарат формальных рассуждений. Основной особенностью задач построения рассуждений в контексте данного исследования является наличие информационной неопределенности и большой размерности, что требует получения и анализа значительного числа альтернативных вариантов структуры системно значимых отношений между объектами знаний исследуемых предметных областей.

Задача *приобретения знаний* сводится к оценке семантической близости понятий (концептов), позволяющей распределить их по классам для дальнейшего построения онтологии предметной области.

3. Построение модели онтологии знаний для интеллектуальных систем обработки и анализа текстов. Построение верхнеуровневой модели онтологии знаний для интеллектуальных систем обработки и анализа текстов проведено с учетом результатов анализа проблем поиска, приобретения и использования знаний, представленных в данной работе ранее.

Несмотря на тот факт, что подход к анализу и обработке текстов на основе извлечения ключевых слов не является новым, в определенных условиях он может дать значительный эффект при определении смысла предложения. Приходя на вход, корпус текста после предварительной обработки приобретает вид входного набора данных, содержащего множество предложений. Для повышения эффективности процесса извлечения ключевых слов онтология имеет список множества имеющихся ключевых слов (*Key Words, KWs*), что

позволяет в дальнейшем на основе процедур нормализации и определения их синонимии находить прецеденты, ускоряющие работу интеллектуальной системы обработки и анализа текстов (рис. 2).



Рис. 2. Компонент онтологии для извлечения и обработки ключевых слов

Таким образом, представленный на рис. 2 компонент верхнеуровневого описания онтологии знаний включает в себя множество предложений, поступивших на вход интеллектуальной системы (3), и множество ключевых слов [13–17], извлеченных и сохраненных в онтологии (4):

$$Sent = \{sent_1, \dots, sent_n\}, \tag{3}$$

$$KW_s = \{kw_1, \dots, kw_m\}. \tag{4}$$

На данных множествах построено бинарное отношение инцидентности (5) между *Sent* и *KW_s*:

$$Rel_{SentKW_s} \subseteq Sent \times KW_s. \tag{5}$$

В онтологии обязательно должно храниться предложение, из которого было выбрано ключевое слово (*KW*). Также необходимо хранить информацию о том, является *KW* объектом или субъектом, так как это сильно влияет на векторную репрезентацию *KW*. Вектор *KW* так же хранится в онтологии. Каждое вновь извлеченное ключевое слово попадает в онтологию либо новым элементом множества *KW_s*, либо как синоним уже существующего *KW*. При этом, необходимо учитывать лингвистические особенности нового элемента онтологической модели (часть речи, род, число, лицо и т.п.), так как эта информация значительно влияет на смысл предложения (рис. 3) [18–21]. Например, в следующих предложениях: «Я системный администратор»; «Я выполняю функции системного администратора»; «Я отвечаю за системное администрирование»; «Он хотел бы стать системным администратором», – при одинаковых (или почти одинаковых) ключевых словах определяется разный смысл. Нужен каталог *KW_s* для каждой доменной (предметной) области, так как *KW_s* нужны до начала разбора предложения. Все слова в онтологии надо делить на категории «значимые» (*Meaningful Words, MWs*) и «уточняющие» (*Clarifying Words, CWs*). Для «значимых» слов должны быть сформированы списки синонимов, а также учтены лингвистические особенности.

Онтология имеет эффективный механизм управления ключевыми словами, что позволяет организовать поиск синонимов (их «склеивание») и, при необходимости, проводить модификацию онтологии (изменение системы внутренних связей между понятиями (концептами)) [20, 21]. Для этого построена иерархия ключевых слов, в которой на верхнем уровне находится поисковое ключевое слово (*Search Key Word, SKW*), включенное в запрос пользователя на входе интеллектуальной системы, например, «транспортное предприятие». Затем, после обработки предложений, на следующем уровне иерархии появляется набор актуальных ключевых слов (*Actual Key Word, AKW*), например, более широкое понятие – «транспортно-логистический кластер» или «транспортный холдинг» и т.п. Данная группа синонимичных *AKW* на следующем уровне иерархии объединяется в единый класс, характерным понятием которого является нормализованное ключевое слово (*Normalized Key Word, NKW*) – часто встречающееся ключевое слово в данном контексте, с наибольшим значением семантической близости, полученной на основе оценки косинусной меры сходства с вектором признаков данного класса (рис. 3).

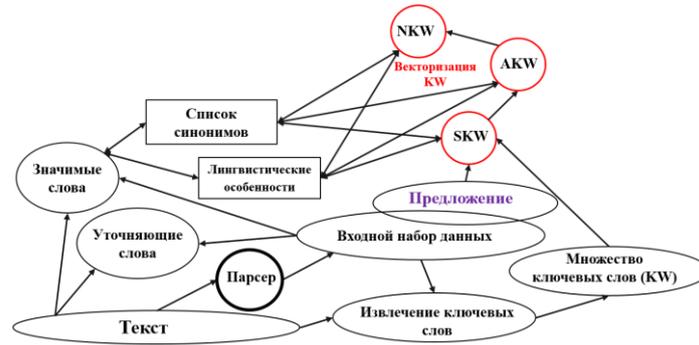


Рис. 3. Расширенный компонент онтологии для извлечения и обработки ключевых слов

Таким образом, представленный на рис. 3 расширенный компонент верхнеуровневого описания онтологии знаний включает в себя множество «значимых» слов с точки зрения извлекаемого смысла MWs (6), множество «уточняющих» слов CWs (7), множество поисковых ключевых слов $SKWs$ (включает в себя извлекаемое из предложений множество KWs , а также ключевые слова, поступающие в интеллектуальную систему через запросы пользователей) (8), множество актуальных ключевых слов $AKWs$ (9), множество нормализованных ключевых слов $NKWs$ (10):

$$MWs = \{mw_1, \dots, mw_e\}, \quad (6)$$

$$CWs = \{cw_1, \dots, cw_j\}, \quad (7)$$

$$SKWs = \{skw_1, \dots, skw_x\}, \quad (8)$$

$$AKWs = \{akw_1, \dots, akw_y\}, \quad (9)$$

$$NKWs = \{nkw_1, \dots, nkw_z\}, \quad (10)$$

На данных множествах построены следующие бинарные отношения инцидентности (11) и (12):

$$Rel_{SentMWs} \subseteq Sent \times MWs, \quad (11)$$

$$Rel_{SentCWs} \subseteq Sent \times CWs, \quad (12)$$

а также следующие бинарные отношения инцидентности (13)-(17) между синонимами:

$$Syn_{SKWsAKWs} \subseteq SKWs \times AKWs, \quad (13)$$

$$Syn_{SKWsNKWs} \subseteq SKWs \times NKWs, \quad (14)$$

$$Syn_{AKWsNKWs} \subseteq AKWs \times NKWs, \quad (15)$$

$$Syn_{AKWsAKWs} \subseteq AKWs \times AKWs, \quad (16)$$

$$Syn_{MWsMWs} \subseteq MWs \times MWs, \quad (17)$$

Глаголы (*Verbs*) тоже хранятся в онтологии, а также имеют нормализованный вид (*NVerbs*) и взаимосвязи с определенными KWs на основе выявленных закономерностей и зависимостей между ними (иногда один глагол может быть связан с несколькими объектами). В онтологии описаны критерии принадлежности к определенной группе контекстов для кластеризации нормализованных смыслов. Предложения с определенными в них KWs распределены по контекстам (предметным областям). Предусмотрена фильтрация слов (наиболее часто встречающихся), определяющих конкретный контекст. Сам контекст определяется на основе «реального смысла» *Real Meaning* (триплета «*субъект – глагол – объект*» («*Sbj – Verb – Obj*»)), построенного на основе слов взятых непосредственно из предложения) (рис. 4). Наиболее проблемной задачей является построение кластеров определенных контекстов.

Таким образом, полное верхнеуровневое описание онтологии знаний для интеллектуальных систем обработки и анализа текстов (рис. 4) включает в себя множества глаголов ($Verbs$) (18) и нормализованных глаголов ($NVerbs$) (19):

$$Verbs = \{verb_1, \dots, verb_s\}, \quad (18)$$

$$NVerbs = \{nverb_1, \dots, nverb_h\}. \quad (19)$$

Отношения синонимии между элементами данных множеств представлено следующим бинарным отношением инцидентности (20):

$$Syn_{VerbsNVerbs} \subseteq Verbs \times NVerbs. \quad (20)$$

В построенной онтологии определены ограничения и правила. Например, можно установить правило, что каждое предложение должно содержать как минимум одно существительное и один глагол. Это поможет проверять полноту и состоятельность онтологической структуры при решении задач извлечения знаний. Необходимо предусмотреть возможность масштабируемости онтологии, что позволит добавлять и изменять ее компоненты.

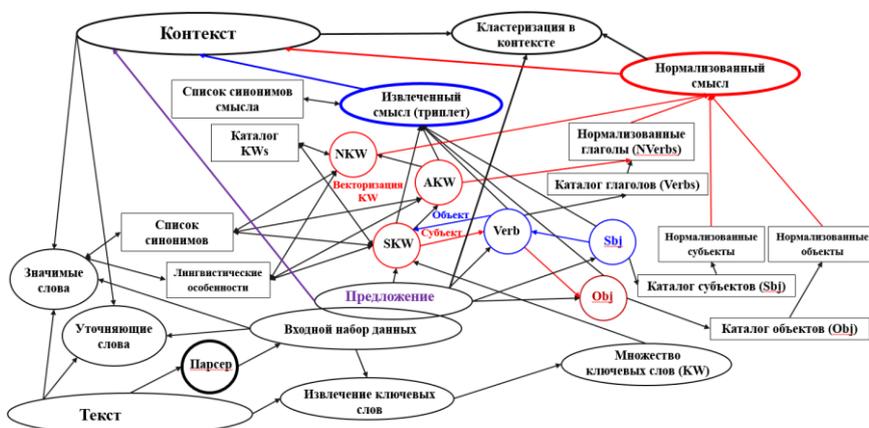


Рис. 4. Верхнеуровневое описание модели онтологии знаний для интеллектуальных систем обработки и анализа текстов

Имеющиеся в онтологии субъекты (Sbj) и объекты (Obj) наполняют множество концептов (понятий) ($K_i, i = 1 \dots N$, где N – количество концептов) с заданной на нём системой связей (бинарных отношений) ($R_q, q = 1 \dots M$, где M – количество связей), являющихся по своей сути несимметричными семантическими отношениями, как это было показано в главе 1.

Конечное множество лексических меток (словарь онтологии) представлено множеством L (21):

$$L = \{L_1, \dots, L_u\}. \quad (21)$$

Выражением (22) представлено антисимметричное, транзитивное, нереклексивное бинарное отношение, являющееся отношением частичного порядка на множестве понятий (концептов) K :

$$Rel_K \subseteq K \times K, Rel_K \in K. \quad (22)$$

Бинарное отношение инцидентности между множествами L и K задано следующим выражением (23):

$$Rel_{LK} \subseteq L \times K. \quad (23)$$

Бинарное отношение инцидентности между множествами L и R задано следующим выражением (24):

$$Rel_{LR} \subseteq L \times R. \quad (24)$$

Тогда под формальной моделью онтологии верхнего уровня описания будем понимать следующий кортеж (25):

$$O = \langle K, R, L, I, Sent, KW, Word, Verb, Rel, Syn, A \rangle, \quad (25)$$

где I – множество экземпляров понятий; $KW = \{KW_s, SKW_s, AKW_s, NKW_s\}$ – интегрированное множество ключевых слов; $Word = \{MW_s, CW_s\}$ – интегрированное множество слов в онтологии; $Verb = \{Verbs, NVerbs\}$ – интегрированное множество глаголов в онтологии; $Rel = \{Rel_k, Rel_{LK}, Rel_{LR}, Rel_{SentKW_s}, Rel_{SentMW_s}, Rel_{SentCW_s}\}$ – интегрированное множество бинарных отношений между элементами онтологии; $Syn = \{Syn_{SKW_sAKW_s}, Syn_{SKW_sNKW_s}, Syn_{AKW_sNKW_s}, Syn_{AKW_sAKW_s}, Syn_{MW_sMW_s}, Syn_{VerbsNVerbs}\}$ – интегрированное множество бинарных отношений синонимии между элементами онтологии; A – аксиомы онтологии.

Таким образом, в данном пункте работы построена верхнеуровневая модель онтологии знаний, которая отличается от известных аналогов применением оригинальной компонентной архитектуры, позволяющей обеспечить высокий уровень детализации спецификаций анализируемой текстовой информации. Представленная модель даёт понимание термина «онтология» в зависимости от контекста текстовой информации и целей его анализа и обработки. В этом заключается гибкость интеллектуальных систем обработки и анализа текстов на естественном языке, повышающая эффективность процессов извлечения знаний.

4. Результаты вычислительного эксперимента. Предложенная верхнеуровневая модель онтологии протестирована при решении задачи структурирования знаний, основным этапом которой является извлечение ключевых слов. Онтология построена в специально созданном программном приложении на основе декомпозиции информации о различных вариантах возникновения предпосылок чрезвычайных ситуаций, а также действий по ликвидации их возможных последствий.

Исследователь задает произвольное количество уровней декомпозиции. Такая организация информационной модели позволяет выделить уровни детализации анализируемой информации. Каждый узел онтологии имеет привязку к определенному элементу информации.

Первому уровню онтологии принадлежит только один понятийный узел. Он является корнем модели и обобщает все ее содержимое, являясь метаинформацией (контекстом). Каждый новый уровень раскрывает содержание концептов (понятий) предыдущего уровня, тем самым раскрывая для исследователя всю предметную область и обнаруживая наличие междисциплинарных связей с другими узлами информационной модели.

Онтология включает в себя множество предметных областей $F = \{F_1, F_2, F_3 \dots F_z\}$, где $z = \overline{1, m}$, m – integer constant (количество предметных областей), причем, $F_z = \{O_{z1}, O_{z2}, O_{z3}, \dots, O_{zk_z}\}$, где k_z – integer constant (количество элементов (концептов) знаний в F_z), причем, в случае задания отношений между концептами одной предметной области получим, что $\forall i, j \rightarrow i = \overline{1, k_z}; j = \overline{1, k_z}; i \neq j \exists C_z[i][j] = n$, где n – количество связей между O_{zi} и O_{zj} , z – integer constant (номер исследуемой предметной области). В случае задания всех, в т.ч. междисциплинарных, отношений между концептами предметных областей получим, что $\forall a, b \rightarrow a = \overline{1, \sum_{z=1}^m k_z}; b = \overline{1, \sum_{z=1}^m k_z}; a \neq b \exists C_{interdisc}[a][b] = p$, где p – количество всех, в т.ч. междисциплинарных, связей между концептами.

Алгоритм структурирования знаний на основе извлечения ключевых слов, необходимый для анализа и обработки предложенной информационной модели, описан в работе [7]. В результате проведенных исследований получена следующая зависимость временной сложности работы алгоритма от числа вершин онтологии (рис. 5).

Временная сложность алгоритма в представленном примере составила $O(n^2)$, где n – количество анализируемых алгоритмом входных данных. По мнению автора в большинстве реальных случаев информационной поддержки процессов предупреждения и/или ликвидации последствий чрезвычайных ситуаций не понадобится онтология большой размерности. Более вероятным диапазоном используемого количества вершин в он-

тологии будет $n \in [5000: 50000]$, в этом случае ВСА алгоритма структурирования информации при поддержке принятия решений по предупреждению и ликвидации последствий чрезвычайных ситуаций составит $O(n)$.



Рис. 5. Зависимость времени работы метода от числа вершин семантической сети

Заключение. В данной статье представлена разработка верхнеуровневого описания модели онтологии знаний для интеллектуальных систем обработки и анализа текстов на естественном языке, построенной на основе оригинальной компонентной архитектуры, обеспечивающей необходимый уровень детализации спецификаций анализируемой текстовой информации. Актуальность данной задачи обусловлена необходимостью развития теоретических основ построения информационных моделей семантических зависимостей внутри текстов на естественном языке.

В статье представлен аналитический обзор задач поиска, приобретения и использования знаний при анализе и обработке текстов на естественном языке. В качестве примера прикладной задачи использования приобретенных знаний, в данном исследовании, рассматривается значимая проблема информационной поддержки процессов предупреждения и/или ликвидации последствий чрезвычайных ситуаций.

Основными результатами проведенного исследования являются следующие:

1. Даны определения основным терминам исследуемой предметной области.
2. Представлена формализованная постановка решаемой задачи.
3. Построена верхнеуровневая модель онтологии знаний, которая отличается от известных аналогов применением оригинальной компонентной архитектуры, позволяющей обеспечить высокий уровень детализации спецификаций анализируемой текстовой информации. Представленная модель даёт понимание термина «онтология» в зависимости от контекста текстовой информации и целей его анализа и обработки. В этом заключается гибкость интеллектуальных систем обработки и анализа текстов на естественном языке, повышающая эффективность процессов извлечения знаний.

Для оценки эффективности применения данной информационной модели разработано программное приложение и проведен вычислительный эксперимент. Полученные результаты проведенных экспериментальных исследований подтверждают эффективность предложенной верхнеуровневой модели онтологии знаний для интеллектуальных систем обработки и анализа текстов на естественном языке.

В развитие данного исследования автор планирует в целях детализации системы отношений между элементами предложенной онтологической структуры провести разработку низкоуровневой модели онтологии знаний, отличающейся применением оригинальной структуры связей между концептами, которая позволит строить множество смысловых паттернов и проводить оценку их семантической близости.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Барсегян А., Куприянов М., Степаненко В., Холод И. Технологии анализа данных: Data Mining, Text Mining, Visual Mining, OLAP. – 2-е изд. – СПб.: БХВ-Петербург, 2008.
2. Кравченко Д.Ю., Кравченко Ю.А., Курейчик В.В., Марков В.В. Математическое описание процесса поддержки принятия решений при оценке семантической близости знаний в конкретизированной модели онтологии // Современные компьютерные технологии: Матер. II научно-методической конференции НПП. – Таганрог: Изд-во ЮФУ, 2021. – С. 25-28.

3. *Хорошевский, В.Ф.* Семантические технологии: ожидания и тренды // Открытые семантические технологии проектирования интеллектуальных систем. – 2012. – № 2. – С. 143-158.
4. *Кравченко Ю.А.* Концептуальные основы рефлексивно-адаптивного подхода к построению интеллектуальных информационных систем // Известия ЮФУ. Технические науки. – 2011. – № 7 (120). – С. 167-171.
5. *Корогодина В.И., Корогодина В.Л.* Информация как основа жизни. – Дубна: Издательский центр «Феникс», 2000. – 208 с.
6. *Кравченко Ю.А.* Информационные модели приобретения знаний и методы их классификации, структурирования, интеграции и семантического поиска: дис. ... д-ра техн. наук: 05.13.17. – Таганрог: 2021. – 314 с.
7. *Кравченко Д.Ю., Герасименко Е.М., Кравченко Ю.А., Кулиев Э.В.* Поддержка принятия решений по предупреждению и ликвидации последствий чрезвычайных ситуаций на основе нечеткого метода структурирования информации // Известия ЮФУ. Технические науки. – 2023. – № 2 (232). – С. 201-212.
8. *Agarwal M.* An Overview of Natural Language Processing // International Journal for Research in Applied Science and Engineering Technology. – 2019. – No. 7. – P. 2811-2813.
9. *Наумов В.Н.* Анализ данных и машинное обучение. Методы и инструментальные средства. – СПб: ИПЦ СЗИУ РАНХиГС, 2020. – 260 с.
10. *Пимешков В.К., Шишаев М.Г.* Методы извлечения знаний из естественно-языковых текстов // Тр. Кольского научного центра РАН. Серия: Технические науки. – 2022. – Т. 13, № 2. – С. 31-45.
11. *Mao X., Huang S., Li R., Shen L.* Automatic keywords extraction based on co-occurrence and semantic relationships between words // IEEE Access. – 2020. – Vol. 8. – P. 117528-117538.
12. *Kravchenko Yury, Mansour Ali, Mohammad Juman, Silega Nemury, Kravchenko Daniil.* Harnessing key phrases in constructing a concept-based semantic representation of text using clustering techniques // Lecture Notes in Computer Science. – Vol. 14335. – P. 190-201.
13. *Kravchenko D.Y., Kravchenko Y.A., Kursitys I.O.* Architecture and Method of Integrating Information and Knowledge on the Basis of the Ontological Structure // Advances in Intelligent Systems and Computing. – 2018. – Vol. 658. – P. 93-103.
14. *Kravchenko D.Y., Kravchenko Yu.A., Kursitys I.O.* Ontological Approach for Designing a Multi-agent Behavior Model in the Internet Environment // Journal of Physics: Conference Series. – 2019. – Vol. 1333. – No. paper 032043.
15. *Ефремова О.А., Павлов С.В.* Онтологическая модель интеграции разнородных по структуре и тематике пространственных баз данных в единую региональную базу данных // Онтология проектирования. – 2017. – Т. 7, № 3 (25). – С. 323-333.
16. *Kozierkiewicz-Hetmanska A., Pietranik M.* The Knowledge Increase Estimation Framework for Ontology Integration on the Concept Level // Journal of Intelligent and Fuzzy Systems. – 2017. – Vol. 32. – P. 1161-1172.
17. *Кравченко Ю.А., Бранченко Д.С.* Исследование преимуществ онтологического подхода в семантическом анализе данных // Информационные технологии, системный анализ и управление. – 2016. – Т. 1. – С. 146-151.
18. *Brandt S., Kontchakov R., Ryzhikov V. et al.* Ontology-Based Data Access with a Horn Fragment of Metric Temporal Logic // 31st Conference on Artificial Intelligence. – 2017. – Vol. 31. – P. 1-17.
19. *Eiter T., Parreira J.X., Schneider P.* Spatial Ontology Mediated Query Answering over Mobility Streams // The Semantic Web - 14th International Conference. – 2017. – Part I. – Vol. 10249. – P. 219-237.
20. *Кравченко Д.Ю., Кравченко Ю.А., Мансур А.М., Мохаммад Ж.Х.* Метод автоматического извлечения ключевых слов // Тр. международного научно-технического конгресса «Интеллектуальные системы и информационные технологии – 2022» («ИС & ИТ-2022», «IS&IT'22»). Т. 1. Научное издание. – Таганрог: Изд-во Ступина С.А., 2022. – С. 90-97.
21. *Кравченко Д.Ю., Кравченко Ю.А., Мансур А.М., Мохаммад Ж.Х.* Модифицированный метод построения семантического представления текста на основе методов кластеризации и взвешивания терминов // Тр. XII международной научно-технической конференции «Технологии разработки информационных систем (ТРИС-2022)». – Таганрог: 2022. – С. 94-100.

REFERENCES

1. *Barsegyan A., Kupriyanoa M., Stepanenko V., Kholod I.* Tekhnologii analiza dannykh: Data Mining, Text Mining, Visual Mining, OLAP [Data analysis technologies: Data Mining, Text Mining, Visual Mining, OLAP]. 2nd ed. St. Petersburg: BKhV-Peterburg, 2008.

2. *Kravchenko D.Yu., Kravchenko Yu.A., Kureychik V.V., Markov V.V.* Matematicheskoe opisaniye protsessy podderzhki prinyatiya resheniy pri otsenke semanticheskoy blizosti znaniy v konkretizirovannoy modeli ontologii [Mathematical description of the decision support process when assessing the semantic proximity of knowledge in a specified ontology model], *Sovremennyye komp'yuternyye tekhnologii: Mater. II nauchno-metodicheskoy konferentsii NPR* [Modern computer technologies: Proceedings of the II scientific and methodological conference of scientific and scientific research]. Taganrog: Izd-vo YuFU, 2021, pp. 25-28.
3. *Khoroshevskiy, V.F.* Semanticheskie tekhnologii: ozhidaniya i trendy [Semantic technologies: expectations and trends], *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem* [Open semantic technologies for designing intelligent systems], 2012, No. 2, pp. 143-158.
4. *Kravchenko Yu.A.* Kontseptual'nyye osnovy refleksivno-adaptivnogo podkhoda k postroeniyu intellektual'nykh informatsionnykh sistem [Conceptual foundations of a reflexive-adaptive approach to the construction of intelligent information systems], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2011, No. 7 (120), pp. 167-171.
5. *Korogodin V.I., Korogodina V.L.* Informatsiya kak osnova zhizni [Information as the basis of life]. Dubna: Izdatel'skiy tsentr «Feniks», 2000, 208 s.
6. *Kravchenko Yu.A.* Informatsionnyye modeli priobreteniya znaniy i metody ikh klassifikatsii, strukturirovaniya, integratsii i semanticheskogo poiska: dis. ... d-ra tekhn. nauk [Information models of knowledge acquisition and methods of their classification, structuring, integration and semantic search: dr. of eng. sc. diss.]: 05.13.17. Taganrog: 2021, 314 p.
7. *Kravchenko D.Yu., Gerasimenko E.M., Kravchenko Yu.A., Kuliev E.V.* Podderzhka prinyatiya resheniy po preduprezhdeniyu i likvidatsii posledstviy chrezvychaynykh situatsiy na osnove nechetkogo metoda strukturirovaniya informatsii [Support for decision-making to prevent and eliminate the consequences of emergency situations based on the fuzzy method of information structuring], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2023, No. 2 (232), pp. 201-212.
8. *Agarwal M.* An Overview of Natural Language Processing, *International Journal for Research in Applied Science and Engineering Technology*, 2019, No. 7, pp. 2811-2813.
9. *Naumov V.N.* Analiz dannykh i mashinnoe obucheniye. Metody i instrumental'nyye sredstva [Data analysis and machine learning. Methods and tools]. St. Petersburg: IPTS SZIU RANKHiGS, 2020, 260 p.
10. *Pimeshkov V.K., Shishaev M.G.* Metody izvlecheniya znaniy iz estestvenno-yazykovykh tekstov [Methods for extracting knowledge from natural language texts], *Tr. Kol'skogo nauchnogo tsentra RAN. Seriya: Tekhnicheskie nauki* [Proceedings of the Kola Scientific Center of the Russian Academy of Sciences. Series: Technical Sciences], 2022, Vol. 13, No. 2, pp. 31-45.
11. *Mao X., Huang S., Li R., Shen L.* Automatic keywords extraction based on co-occurrence and semantic relationships between words, *IEEE Access*, 2020, Vol. 8, pp. 117528-117538.
12. *Kravchenko Yury, Mansour Ali, Mohammad Juman, Silega Nemury, Kravchenko Daniil.* Harnessing key phrases in constructing a concept-based semantic representation of text using clustering techniques, *Lecture Notes in Computer Science*, Vol. 14335, pp. 190-201.
13. *Kravchenko D.Y., Kravchenko Y.A., Kursitya I.O.* Architecture and Method of Integrating Information and Knowledge on the Basis of the Ontological Structure, *Advances in Intelligent Systems and Computing*, 2018, Vol. 658, pp. 93-103.
14. *Kravchenko D.Y., Kravchenko Yu.A., Kursitya I.O.* Ontological Approach for Designing a Multi-agent Behavior Model in the Internet Environment, *Journal of Physics: Conference Series*, 2019, Vol. 1333, No. paper 032043.
15. *Efremova O.A., Pavlov S.V.* Ontologicheskaya model' integratsii raznorodnykh po strukture i tematike prostranstvennykh baz dannykh v edinuyu regional'nuyu bazu dannykh [Ontological model for integrating spatial databases of heterogeneous structure and subject matter into a single regional database], *Ontologiya proektirovaniya* [Design Ontology], 2017, Vol. 7, No. 3 (25), pp. 323-333.
16. *Kozierkiewicz-Hetmanska A., Pietranik M.* The Knowledge Increase Estimation Framework for Ontology Integration on the Concept Level, *Journal of Intelligent and Fuzzy Systems*, 2017, Vol. 32, pp. 1161-1172.
17. *Kravchenko Yu.A., Branchenko D.S.* Issledovanie preimushchestv ontologicheskogo podkhoda v semanticheskoy analize dannykh [Study of the advantages of the ontological approach in semantic data analysis], *Informatsionnyye tekhnologii, sistemnyy analiz i upravleniye* [Information technologies, system analysis and management], 2016, Vol. 1, pp. 146-151.
18. *Brandt S., Kontchakov R., Ryzhikov V. et al.* Ontology-Based Data Access with a Horn Fragment of Metric Temporal Logic, *31st Conference on Artificial Intelligence*, 2017, Vol. 31, pp. 1-17.
19. *Eiter T., Parreira J.X., Schneider P.* Spatial Ontology Mediated Query Answering over Mobility Streams, *The Semantic Web - 14th International Conference*, 2017, Part I, Vol. 10249, pp. 219-237.

20. Kravchenko D.Yu., Kravchenko Yu.A., Mansur A.M., Mokhammad Zh.Kh. Metod avtomaticheskogo izvlecheniya klyuchevykh slov [Method for automatic extraction of keywords], *Tr. mezhdunarodnogo nauchno-tekhnicheskogo kongressa «Intellectual'nye sistemy i informatsionnye tekhnologii – 2022» («IS & IT-2022», «IS&IT'22»)* [Proceedings of the international scientific and technical congress “Intelligent systems and information technologies - 2022” (“IS & IT-2022”, “IS&IT'22”)]. Vol. 1. Scientific publication. Taganrog: Izd-vo Stupina S.A., 2022, pp. 90-97.
21. Kravchenko D.Yu., Kravchenko Yu.A., Mansur A.M., Mokhammad Zh.Kh. Modifitsirovannyi metod postroeniya semanticheskogo predstavleniya teksta na osnove metodov klasterizatsii i vzhivaniya terminov [A modified method for constructing a semantic representation of text based on clustering and term weighting methods], *Tr. XII mezhdunarodnoy nauchno-tekhnicheskoy konferentsii «Tekhnologii razrabotki informatsionnykh sistem (TIS -2022)»* [Proceedings of the XII International Scientific and Technical Conference “Information Systems Development Technologies (TRIS-2022)”]. Taganrog: 2022, pp. 94-100.

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

Кравченко Даниил Юрьевич – Южный федеральный университет; e-mail: dkravchenko@sfedu.ru; г. Таганрог, Россия; тел.: 88634371651; кафедра систем автоматизированного проектирования; аспирант.

Kravchenko Daniil Yurievich – Southern Federal University; e-mail: dkravchenko@sfedu.ru; Taganrog, Russia; phone: +78634371651; the department of computer aided design, postgraduate.

УДК 621.337.11:004.942:519.876.5

DOI 10.18522/2311-3103-2024-2-50-59

П.П. Чернусь, П.П. Чернусь, А.А. Яковлев, Р.В. Сахабудинов, А.С. Голосий

СИНТЕЗ СИСТЕМЫ УПРАВЛЕНИЯ КОМПЛЕКСА ПОДЪЕМА ПОДВОДНОГО ГРУЗА

Для транспортировки подводного груза применяется судно-носитель (СН). Судно носитель оснащается комплексом подъема подводного груза, предназначенным для захвата груза, его подъема, закрепления на СН и транспортировки к пункту базирования. Комплекс подъема подводного груза включает: спускаемый модуль, механизм подъема, механизм фиксации, механизм демпфирования и систему управления. В статье представлены результаты разработки математической модели подводного груза, который описан на основании теорем об изменении количества движения и кинетического момента механической системы. В уравнениях линейного и углового перемещений груза присутствует присоединенная к нему масса жидкости. С учётом относительно малых линейных и угловых скоростей, а также малых углов вращения выполнена линеаризация уравнения динамики в скалярном виде. Модель груза реализована в стандартных блоках системы имитационного моделирования. Для синтеза системы управления груз представлен передаточной функцией в виде апериодического звена второго порядка. В работе синтезирован трехконтурный подчиненный регулятор груза с обратными связями по положению, по скорости и по току. Получены выражения для расчёта обобщённых динамических характеристик замкнутой систем второго порядка в зависимости от относительного коэффициента усиления регулятора, рассчитаны параметры системы. Исследования, проведенные на математических моделях системы, позволили получить начальные сведения о линейном и угловом перемещении спускаемого модуля в установившемся режиме, перемещении точек внешнего и внутреннего подвесов, величине силы на тросах, о моменте и скорости, развиваемых электродвигателями лебедок. Моделирование режимов спуска, стабилизации и подъема позволили скорректировать параметры оборудования и добиться удовлетворительных результатов функционирования комплекса.

Моделирование; подводный груз; масса; инерция; кинетический момент; качка; линеаризация; система управления; передаточная функция; замкнутый контур; Пи-регулятор.

P.P. Chernus, P.P. Chernus, A.A. Yakovlev, R.V. Sakhabudinov, A.S. Golosiy

SYNTHESIS OF THE UNDERWATER CARGO LIFTING COMPLEX CONTROL SYSTEM

A carrier vessel is used to transport an underwater cargo. The given vessel is equipped with a unit for lifting underwater cargoes designed to capture it, lift it, secure it on the vessel and transport it to the base point. The given unit includes the following components: the descent module, the lifting mechanism, the locking mechanism, the damping mechanism and the control system. The paper presents the findings of the mathematical model development of the main components for the unit to make a crucial contribution to obtaining reliable results – an underwater cargo, an asynchronous motor, a cable suspension and a compensation mechanism. The underwater cargo is described based on the theorems about the change in the amount of motion and kinetic moment of the mechanical system. In the equations of linear and angular displacement of a cargo, there is a mass of liquid in it. The cable suspension model takes into account its deformation in motion while operating. The model of an asynchronous motor with a short-circuited rotor is obtained from a generalized circuit. A vector control method is provided, the rotor flow coupling vector is taken as the base vector. The compensation mechanism model is based on an adiabatic process in a macroscopic system, where there is no heat released into. For the control system synthesis, the cargo is represented by a transfer function in the form of an aperiodic link of the second order. There is a three-loop PID controller synthesized with feedback in position, speed and current. Equations are obtained for calculating the generalized dynamic characteristics of a closed second-order system, and the controller parameters are calculated. The findings carried out on mathematical models of the system help us to obtain initial information about the linear and angular displacement of the descent module in steady state, the movement of points of external and internal suspensions, the magnitude of the force on the cables, the torque and speed developed by electric motors of winches. Modeling of descent, stabilization and ascent modes made it possible to adjust the parameters of the equipment and achieve satisfactory results of the complex's operation.

Modeling; underwater cargo; mass; inertia; kinetic moment; pitching; linearization; control system; transfer attribute; closed loop; proportional integral (PI) controller.

Введение. Схема механизма подъема основана на применении тросовых лебедок; четыре точки подвеса тросовых систем размещены в диаметральной плоскости судна-носителя (рис. 1). Здесь T1, T2, T3, T4 – точки подвеса груза.

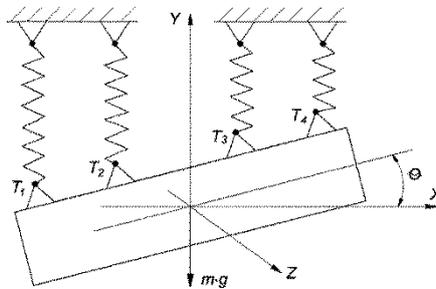


Рис. 1. Расположение точек подвеса и осей координат груза

Математическая модель груза. Динамика твердого тела может быть описана следующими уравнениями

$$\frac{\delta Q}{\delta t} + \omega \times Q = \sum_i F_i, \tag{1}$$

$$\frac{\delta K_O}{\delta t} + V \times Q + \omega \times K_O = \sum_i \rho_i \times F_i, \tag{2}$$

где Q – вектор количества движения механической системы; K_O – вектор кинетического момента механической системы относительно подвижного центра O ; $\frac{\delta Q}{\delta t}, \frac{\delta K_O}{\delta t}$ – производные векторов Q и K_O относительно подвижной системы координат $OXYZ$; V – вектор скорости подвижного центра O ; ω – вектор угловой скорости системы $OXYZ$; $\sum_i F_i, \sum_i \rho_i \times F_i$ – векторы суммарных силы и момента относительно центра O от всех внешних сил, действующих на систему [1].

Векторы Q количества движения и K_O кинетического момента твёрдого тела относительно подвижного центра O определяются в соответствии с выражениями

$$Q = mV + m\omega \times \rho_c, \quad (3)$$

$$K_O = m\rho_c \times V + J\omega, \quad (4)$$

где J – тензор инерции твёрдого тела.

Сила, приводящая груз в движение, должна перемещать не только сам груз, но и некоторое количество присоединенной жидкости вслед за ним [2, 3]. Аналогичная ситуация возникает и при вращении подводного груза [4]. Присоединенные массы и моменты инерции зависят от геометрии подводного груза [5]. Объединенные в симметрическую матрицу они имеют вид:

$$\lambda = \{\lambda_{ij}; i, j = 1, \dots, 6\}. \quad (5)$$

Тогда уравнения для количества движения и кинетического момента примут вид:

$$Q = mV + m\omega \times \rho_c + \sum_{i=1}^3 (\lambda_i, x) e_i, \quad (6)$$

$$K_O = m\rho_c \times V + J\omega + \sum_{i=1}^3 (\lambda_{i+3}, x) e_i, \quad (7)$$

где $x = (V_x, V_y, V_z, \omega_x, \omega_y, \omega_z)^T$ – вектор кинематических параметров движения, спроецированных на оси связанной системы координат; λ_i – i -я строка матрицы λ ; (λ_i, x) – скалярное произведение векторов λ_i и x ; e_i – орты связанной системы координат. Следует отметить, что векторы Q и K_O являются линейными формами вектора X

$$Q = A_Q \cdot x, \quad (8)$$

$$K_O = A_K \cdot x, \quad (9)$$

где

$$A_Q = \begin{pmatrix} m + \lambda_{11} & \lambda_{12} & \lambda_{13} & \lambda_{14} & mz_c + \lambda_{15} & -m\gamma_c + \lambda_{16} \\ \lambda_{21} & m + \lambda_{22} & \lambda_{23} & -mz_c + \lambda_{24} & \lambda_{25} & mx_c + \lambda_{26} \\ \lambda_{31} & \lambda_{32} & m + \lambda_{33} & m\gamma_c + \lambda_{34} & -mx_c + \lambda_{25} & \lambda_{26} \end{pmatrix}, \quad (10)$$

$$A_K = \begin{pmatrix} \lambda_{41} & -mz_c + \lambda_{42} & m\gamma_c + \lambda_{43} & J_{xx} + \lambda_{44} & -J_{xy} + \lambda_{45} & -J_{xz} + \lambda_{46} \\ mz_c + \lambda_{51} & \lambda_{52} & -mx_c + \lambda_{53} & -J_{yx} + \lambda_{54} & J_{yy} + \lambda_{55} & -J_{yz} + \lambda_{56} \\ -m\gamma_c + \lambda_{61} & mx_c + \lambda_{62} & \lambda_{63} & -J_{zx} + \lambda_{64} & -J_{zy} + \lambda_{65} & J_{zz} + \lambda_{66} \end{pmatrix}. \quad (11)$$

В итоге имеет место следующая система уравнений

$$A_Q \frac{dx}{dt} = \sum_i F_i - \omega \times (A_Q x), \quad (12)$$

$$A_K \frac{dx}{dt} = \sum_i \rho_i \times F_i - V \times (A_Q x) - \omega \times (A_K x). \quad (13)$$

После преобразований, а также с учетом только перемещений вертикально вверх и вращения вокруг оси Z уравнения динамики в скалярном виде имеют вид

$$(m + \lambda_{22})\dot{v}_y + \lambda_{26}\dot{\omega}_z = \frac{c\gamma\rho v_y^2}{2} S + A - G + F_{By} + m\gamma\omega_z^2, \quad (14)$$

$$(J_{zz} + \lambda_{66})\dot{\omega}_z + \lambda_{26}\dot{v}_y = \frac{c\gamma\rho v_y^2}{2} SL + \gamma_c G \sin \theta - M_{Bz} - m\gamma\omega_z v_y. \quad (15)$$

Проводя линеаризацию с учётом малых линейных и угловых скоростей, малых углов, получена следующая модель груза, представленная в стандартных блоках системы имитационного блочного моделирования MATLAB Simulink (рис. 2).

Синтез системы управления. Основным преимуществом подчиненной структуры является возможность компенсации возмущений на груз регулирования, которые могут быть вызваны, как внешними воздействиями (температура, напряжение питания, момент нагрузки и т.д.), так и внутренними перекрестными связями в многомерных системах. В дальнейшем рассматривается обобщенный ПИД-регулятор с реальным дифференцирующим звеном, передаточная функция которого определяется выражением

$$W_P(s) = k_{\Pi} + \frac{k_{\text{И}}}{s} + \frac{k_{D S}}{T_{D S} s + 1}. \quad (16)$$

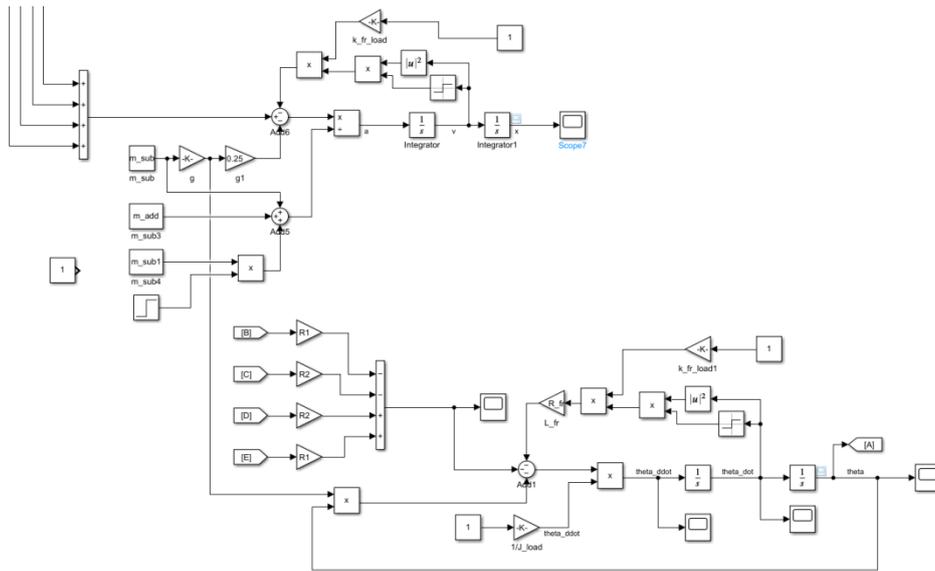


Рис. 2. Модель груза

Синтез осуществлялся по так называемой стандартной методике [13, 17]. Суть данного подхода заключается в следующем. Синтез регуляторов проводится по контурам, в которых передаточная функция груза представляется звеном второго порядка, цель которого – обеспечить замкнутому контуру требуемые динамические характеристики. Например, если передаточная функция замкнутого контура имеет вид

$$W(s) = \frac{b_0}{s^2 + 2\xi\omega_0s + \omega_0^2}, \quad (17)$$

где $\xi = 0,707$, то она называется оптимумом по модулю и доставляет минимум интегральному критерию от ошибки регулирования и дает перерегулирование $\sigma = 4,3\%$:

$$I_{ITAE} = \int_0^{\infty} t \cdot e(t) dt, \quad (18)$$

Рассмотрим подробнее упрощённую модель асинхронной машины с короткозамкнутым ротором [6–11], и определимся с выбором типов регуляторов [14–16, 18, 19] (рис. 3).

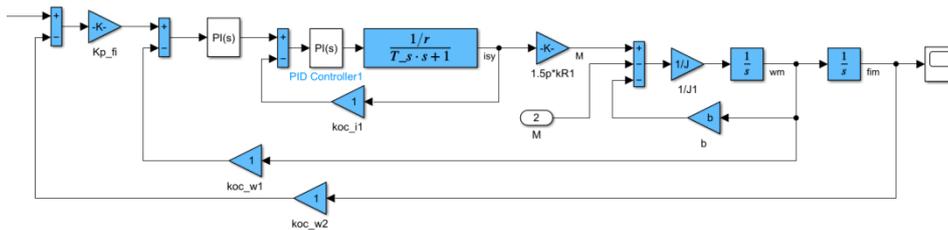


Рис. 3. Передаточная функция эквивалентной схемы

Токовый контур имеет передаточную функцию

$$W_{\text{рази}} = \frac{1/r}{T_s \cdot s + 1} = \frac{29}{0,032s + 1}. \quad (19)$$

Для коррекции применим ПИ-регулятор [20], рассчитанный по стандартной методике [13], в целях компенсации большей постоянной времени. Соотношение параметров

$$\frac{k_p}{k_i} = T_s \quad (20)$$

Примем $k_i = 5$, $k_p = 0,16$. Тогда передаточная функция токового контура

$$W_i = \frac{5(0,032s+1)}{s} \cdot \frac{29}{0,032s+1} = \frac{145}{s}. \quad (21)$$

Замкнутая передаточная функция токового контура

$$W_{i\text{зам}} = \frac{145}{s+145} = \frac{1}{0,007s+1}. \quad (22)$$

В свою очередь, скоростной контур имеет следующую передаточную функцию

$$W_{\text{раз}\omega} = \frac{1}{0,007 \cdot s+1} \cdot \frac{\frac{1}{b}}{J \cdot s+1} = \frac{1}{0,007 \cdot s+1} \cdot \frac{79}{80 \cdot s+1}. \quad (23)$$

Выберем $\xi = 1$. Тогда переходной процесс контура будет без перерегулирования. Также применим ПИ-регулятор.

Рассчитаем систему: $\omega_0 = 101$; $k_{p\omega} = 71,4$; $k_{i\omega} = 0,89$.

$$W_{\text{зам}\omega} = \frac{\frac{1}{k_{oc}}}{\frac{1}{\omega_0^2} \cdot s^2 + \frac{2 \cdot \xi}{\omega_0} \cdot s + 1} \approx \frac{1}{0,014 \cdot s+1} \quad (24)$$

Контур по положению имеет передаточную функцию

$$W_{\text{раз}\varphi} = \frac{1}{s \cdot (0,014 \cdot s+1)}. \quad (25)$$

Для контура по положению выберем $\xi = 0,707$. Тогда переходной процесс контура будет с перерегулированием в 4 %.

Рассчитаем систему: $\omega_0 = 50,5$; $k_{p\omega} = 35,7$.

Моделирование режимов функционирования. Математическая модель комплекса подъема подводного груза разработана в пакете MATLAB Simulink [12]. В общей модели присутствуют блоки, являющиеся частными моделями: асинхронного двигателя (motor), тросового подвеса, механизма компенсации, груза (cargo) – спускаемого модуля с грузом или без него, системы управления. Результаты моделирования показали следующее:

1. Выбранные параметры механизма подъёма в части мощности электродвигателя, передаточных отношений редуктора лебёдок и синтезированная система управления обеспечивают режим стабилизации при заданных параметрах качки СН и параметрах спускаемого модуля. Размах колебаний крайних точек подвеса не превосходит (0,06–0,07) м, а угловые колебания минус $0,1^\circ$. Силы в точках подвеса лежат в пределах от 120 до 160 кН (рис. 4).

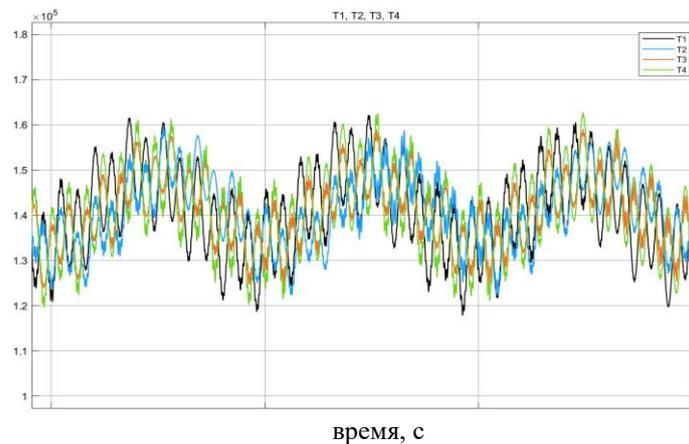


Рис.4. Силы на тросах в установившемся режиме стабилизации

2. Режим подъема в зависимости от положения судна по отношению к вертикали волны имеет три специфические фазы (рис. 5). На спокойной воде (рис. 5,а) каретка упирается в наружный шток компенсатора и неподвижна. Натяжение тросов равно неуровненности груза. Скорость подъема от 0,05 до 0,10 м/с обеспечивается лебедкой.

При волнении СН сначала взбирается на гребень волны (рис. 5,б); перемещения штока и ускорения положительны и максимальны. При этом каретка утапливает наружный шток компенсатора, полиспаст «растягивается». Усилия на тросах максимальны, груз перемещается со средней скоростью от 0,05 до 0,10 м/с.

В фазе спуска во впадину волны (рис. 5,в) перемещения штока и ускорения также максимальны, но с отрицательным знаком. Внутренний шток компенсатора выдвигается, каретка отодвигается, полиспаст «сжимается». Усилия на тросах минимальны, груз перемещается с практически постоянной скоростью.

В процессе моделирования режима подъема были уточнены параметры механизма компенсации – снижено начальное давление (до 6 МПа) и уменьшен диаметр наружного поршня.

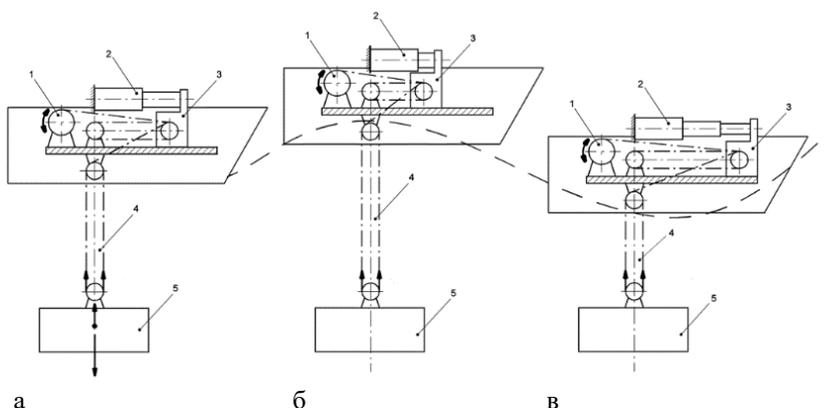


Рис. 5. Схема работы в режиме подъема: 1 – лебедка; 2 – гидроцилиндр; 3 – каретка; 4 – полиспаст; 5 – груз

В результате, в режиме свободного подъема наблюдается незначительное колебание средней скорости от 0,05 до 0,15 м/с (рис. 6) при заданных ходах кареток механизмов компенсации $\pm 1,5$ м. При подъеме наблюдается увеличение угловых колебаний спускаемого модуля, но их величина не превосходит 1° , что можно считать допустимым. Провисания тросов не происходит, усилия на узлах подвеса не превышают 270 кН (рис. 7).

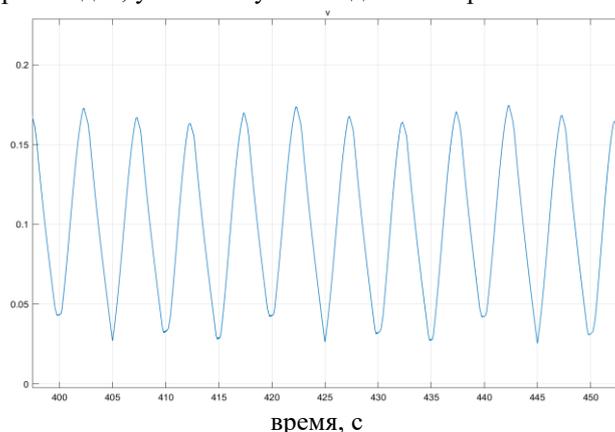


Рис. 6. Колебания скорости центра тяжести СМ в режиме подъема

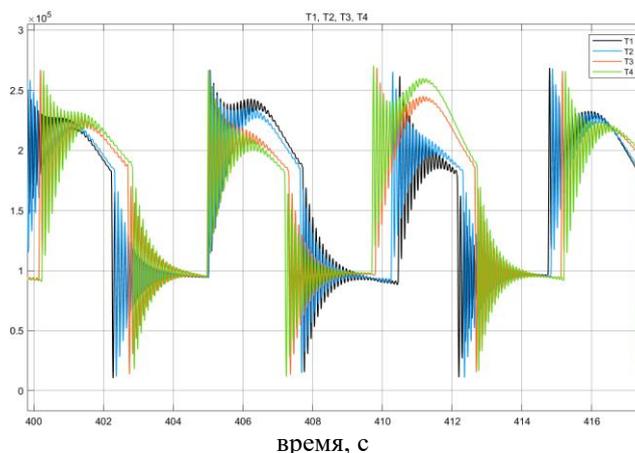


Рис. 7. Силы, действующие в точках подвеса СМ в режиме подъёма

3. Режим спуска является наиболее сложным для системы. Это связано с тем, что в режиме спуска груз обладает повышенными динамическими характеристиками по сравнению с режимом подъёма, а приведённая масса компенсатора достаточно велика. Поэтому наблюдается сдвиг по фазе реакции груза и компенсатора, что приводит к увеличению амплитуды колебаний. Кроме этого отсутствие стабилизации на спуске приводит к выходу на ограничение в 2 м компенсатора и, как следствие, наблюдаются удары гидrocилиндров компенсатора в упоры, что требует введения демпфирующих устройств.

Для исключения рывков и провисов тросов введена стабилизация (режим управляемого спуска). При управляемом спуске механизм компенсации не работает, максимальные обороты электродвигателей лебёдок не превосходят 1250 об/мин (рис. 8).

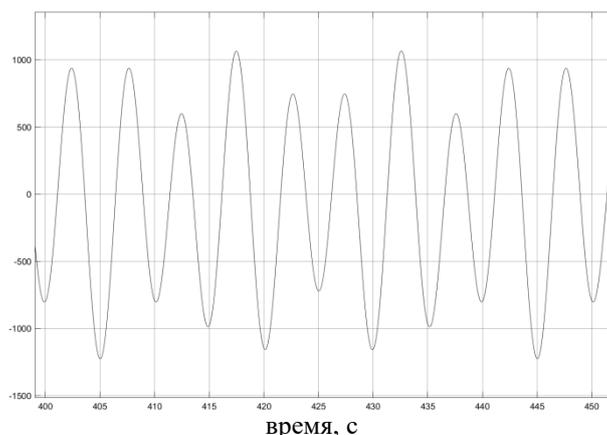


Рис. 8. Обороты ротора ЭД крайней лебёдки в режиме управляемого спуска

4. Выбранные параметры механизма подъёма имеют определенные резервы по увеличению массы спускаемого модуля или увеличению допустимой амплитуды качки на 20–25%.

Заключение. Для комплекса подъема подводного груза синтезирована система управления, в контурах обратной связи которой применены корректирующие ПИ-регуляторы. В модели объекта управления учтены силы и моменты, действующие на груз, в том числе масса присоединенной к грузу жидкости. В выражениях обобщённых динамических характеристик замкнутой системы второго порядка в зависимости от относительного коэффициента усиления регулятора, рассчитаны его параметры.

Имитационное моделирование позволило оценить динамические характеристики замкнутой системы при ее функционировании во всех режимах работы. Поскольку гидравлическая система подъема включает тросовую систему, проведены исследования на модели предельных усилий в наиболее напряженных режимах и анализ колебательных процессов спускаемого модуля. Для электропривода на базе асинхронного двигателя оценены предельные обороты вращения ротора.

В целом контур управления объектом с механизмом подъема подтвердил работоспособность, реализуемость и приемлемые характеристики как на основных, так и на дополнительных режимах функционирования.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Никущенко Д.В. К вопросу о выборе системы координат при исследовании динамики подводных грузов // Морские интеллектуальные технологии. – 2015. – Т. 1, № 2 (28). – С. 19.
2. Huang H., Tang Q., Li H. et al. Vehicle-manipulator system dynamic modeling and control for underwater autonomous manipulation // Multibody Syst Dyn. – 2017. – 41. – P. 125-147. – <https://doi.org/10.1007/s11044-016-9538-3>.
3. Sharma A.K., Saha S.K. Simplified drag modeling for the dynamics of an underwater manipulator // IEEE Journal of Oceanic Engineering. – 2019. – Vol. 46, No. 1. – P. 40-55.
4. Куришин А.В. Комплексирование на подводном аппарате данных инерциальной навигационной системы, магнитометра и глобальной навигационной спутниковой системы ГЛОНАСС: дис. ... канд. техн. наук: 05.13.01. – М., 2016. – 144 с.
5. Chang Z. et al. Dynamics Simulation of Grasping Process of Underwater Vehicle-Manipulator System // Journal of Marine Science and Engineering. – 2021. – Vol. 9, No. 10. – P. 1131.
6. Афоничев Д.Н., Пиляев С.Н., Степин М.А. Математическая модель управляемого асинхронного электродвигателя // Повышение эффективности использования мобильных энергетических средств в различных режимах движения. – 2017. – С. 10-15.
7. Омельченко Е.Я. Динамические математические модели асинхронных двигателей. – 2012.
8. Кыюнг Н.С. и др. Моделирование электропривода с частотным управлением асинхронного двигателя // Известия Тульского государственного университета. Технические науки. – 2014. – № 3. – С. 221-228.
9. Мецержяков В.Н. и др. Система управления частотным асинхронным синхронизированным электроприводом // Известия высших учебных заведений. Проблемы энергетики. – 2021. – Т. 23, № 3. – С. 116-126.
10. Семёнов А.С. Моделирование режимов работы асинхронного двигателя в пакете программ MATLAB // Вестник Северо-Восточного федерального университета им. М.К. Аммосова. – 2014. – Т. 11, № 1. – С. 51-59.
11. Ahmed A., Kaur K. Mathematical model for an asynchronous motor implemented in MATLAB // International journal of scientific & technical development. – 2023.
12. Moler C. Numerical Computing with MATLAB. Electronic ed. The MathWorks, Inc, Natick, MA, 2014.
13. Герман-Галкин С.Г. Matlab & Simulink. Проектирование мехатронных систем на ПК. – СПб.: Корона-Век, 2008. – 386 с.
14. Dogruer T., Tan N. Design of PI controller using optimization method in fractional order control systems // IFAC-PapersOnLine. – 2018. – Vol. 51, No. 4. – P. 841-846.
15. Черный Г.Г. Газовая динамика. – М.: Наука, 1988. – 424 с.
16. Haddrell A.E. et al. Accounting for changes in particle charge, dry mass and composition occurring during studies of single levitated particles // The Journal of Physical Chemistry A. – 2012. – Vol. 116, No. 40. – P. 9941-9953.
17. Бесекерский В.А., Попов Е.П. Теория систем автоматического управления. – СПб.: Профессия, 2003. – 752 с.
18. Bequette B.W. Process control: modeling, design, and simulation. – Prentice Hall Professional, 2003.
19. Bar-Kana I. Adaptive control: A simplified approach // Control and Dynamic Systems. – 2012. – Vol. 25. – P. 187-235.
20. Гайворонский С.А., Езангина Т.А. Методика выбора параметров ПИ-регулятора для интервальной системы автоматического управления // Векторы благополучия: экономика и социум. – 2012. – № 3 (4). – С. 143-147.

REFERENCES

1. *Nikushchenko D.V.* К вопросу о выборе системы координат при исследовании динамики подводных грузов [On the issue of choosing a coordinate system when studying the dynamics of underwater cargo], *Morskie intellektual'nye tekhnologii* [Marine intelligent technologies], 2015, Vol. 1, No. 2 (28), pp. 19.
2. *Huang H., Tang Q., Li H. et al.* Vehicle-manipulator system dynamic modeling and control for underwater autonomous manipulation, *Multibody Syst Dyn*, 2017, 41, pp. 125-147. Available at: <https://doi.org/10.1007/s11044-016-9538-3>.
3. *Sharma A.K., Saha S.K.* Simplified drag modeling for the dynamics of an underwater manipulator, *IEEE Journal of Oceanic Engineering*, 2019, Vol. 46, No. 1, pp. 40-55.
4. *Kurshin A.V.* Kompleksirovanie na podvodnom apparate dannykh inertsiyal'noy navigatsionnoy sistemy, magnitometra i global'noy navigatsionnoy sputnikovoy sistemy GLONASS: dis. ... kand. tekhn. nauk [Integration of data from an inertial navigation system, magnetometer and global navigation satellite system GLONASS on an underwater vehicle: cand. of eng. diss.]: 05.13.01. Moscow, 2016, 144 p.
5. *Chang Z. et al.* Dynamics Simulation of Grasping Process of Underwater Vehicle-Manipulator System, *Journal of Marine Science and Engineering*, 2021, Vol. 9, No. 10, pp. 1131.
6. *Afonichev D.N., Pilyaev S.N., Stepin M.A.* Matematicheskaya model' upravlyаемого asinkhronnogo elektrodvigatelya [Mathematical model of a controlled asynchronous electric motor], *Povyshenie effektivnosti ispol'zovaniya mobil'nykh energeticheskikh sredstv v razlichnykh rezhimakh dvizheniya* [Increasing the efficiency of using mobile energy vehicles in various driving modes], 2017, pp. 10-15.
7. *Omel'chenko E.Ya.* Dinamicheskie matematicheskie modeli asinkhronnykh dvigateley [Dynamic mathematical models of asynchronous motors], 2012.
8. *Kyong N.S. i dr.* Modelirovanie elektroprivoda s chastotnym upravleniem asinkhronnogo dvigatelya [Modeling of an electric drive with frequency control of an asynchronous motor], *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki* [News of Tula State University. Technical science], 2014, No. 3, pp. 221-228.
9. *Meshcheryakov V.N. i dr.* Sistema upravleniya chastotnym asinkhronnym sinkhronizirovannym elektroprivodom [Control system for frequency asynchronous synchronized electric drive], *Izvestiya vysshikh uchebnykh zavedeniy. Problemy energetiki* [News of higher educational institutions. Energy problems], 2021, Vol. 23, No. 3, pp. 116-126.
10. *Semenov A.S.* Modelirovanie rezhimov raboty asinkhronnogo dvigatelya v pakete programm MATLAB [Modeling of operating modes of an asynchronous motor in the MATLAB software package], *Vestnik Severo-Vostochnogo federal'nogo universiteta im. MK Ammosova* [Bulletin of the North-Eastern Federal University named after. MK Ammosova], 2014, Vol. 11, No. 1, pp. 51-59.
11. *Ahmed A., Kaur K.* Mathematical model for an asynchronous motor implemented in MATLAB, *International journal of scientific & technical development*, 2023.
12. *Moler C.* Numerical Computing with MATLAB. Electronic ed. The MathWorks, Inc, Natick, MA, 2014.
13. *German-Galkin S.G.* Matlab & Simulink. Proektirovanie mekhatronnykh sistem na PK [Matlab & Simulink. Design of mechatronic systems on a PC]. St. Petersburg: Korona-Vek, 2008, 386 p.
14. *Dogruer T., Tan N.* Design of PI controller using optimization method in fractional order control systems, *IFAC-PapersOnLine*, 2018, Vol. 51, No. 4, pp. 841-846.
15. *Chernyy G.G.* Gazovaya dinamika [Gas dynamics]. Moscow: Nauka, 1988, 424 p.
16. *Haddrell A.E. et al.* Accounting for changes in particle charge, dry mass and composition occurring during studies of single levitated particles, *The Journal of Physical Chemistry A*, 2012, Vol. 116, No. 40, pp. 9941-9953.
17. *Besekerskiy V.A., Popov E.P.* Teoriya sistem avtomaticheskogo upravleniya [Theory of automatic control systems]. St. Petersburg: Professiya, 2003, 752 p.
18. *Bequette B.W.* Process control: modeling, design, and simulation. Prentice Hall Professional, 2003.
19. *Bar-Kana I.* Adaptive control: A simplified approach, *Control and Dynamic Systems*, 2012, Vol. 25, pp. 187-235.
20. *Gayvoronskiy S.A., Ezangina T.A.* Metodika vybora parametrov PI-regulyatora dlya interval'noy sistemy avtomaticheskogo upravleniya [Methodology for selecting PI controller parameters for an interval automatic control system], *Vektory blagopoluchiya: ekonomika i sotsium* [Vectors of well-being: economics and society], 2012, No. 3 (4), pp. 143-147.

Статью рекомендовал к опубликованию д.т.н., профессор Ю.Л. Вященко.

Чернусь Петр Павлович – БГТУ «ВОЕНМЕХ» им. Д.Ф. Устинова; e-mail: petr.chernus@yandex.ru; г. Санкт-Петербург, Россия; тел.: +78124900518; к.т.н.; доцент.

Чернусь Павел Павлович – e-mail: pavel.chernus@yandex.ru; тел.: +78124900518; к.т.н.; доцент.

Яковлев Александр Анатольевич – АО «Конструкторское бюро «Арсенал» им. М.В. Фрунзе; e-mail: aa.yakovlev@kbarsenal.ru; г. Санкт-Петербург, Россия; тел.: +79052500726; к.т.н.; начальник управления.

Сахабудинов Роман Владиславович – Санкт-Петербургский государственный университет аэрокосмического приборостроения; e-mail: stapler1@yandex.ru; г. Санкт-Петербург, Россия; тел.: +79211852881; к.т.н.; доцент.

Голосий Александр Сергеевич – e-mail: a.s.golosiy@mail.ru; тел.: +79211852881; к.т.н.; доцент.

Chernus Peter Pavlovich – BSTU "VOENMEH" D.F. Ustinova; e-mail: petr.chernus@yandex.ru; St. Petersburg, Russia; phone: +78124900518; cand. of eng. sc.; associate professor.

Chernus Pavel Pavlovich – e-mail: pavel.chernus@yandex.ru; phone: +78124900518; cand. of eng. sc.; associate professor.

Yakovlev Aleksandr Anatolievich – ARSENAL DESIGN BUREAU JSK; e-mail: aa.yakovlev@kbarsenal.ru; Saint Petersburg, Russia; phone: +79052500726; cand. of eng. sc.; head of department.

Sakhabudinov Roman Vladislavovich – Saint Petersburg State University of Aerospace Instrumentation; e-mail: stapler1@yandex.ru; Saint Petersburg, Russia; phone: +79211852881; cand. of eng. sc.; associate professor.

Golosiy Aleksandr Sergeevich – e-mail: a.s.golosiy@mail.ru; phone: +79213025903; cand. of eng. sc.; associate professor.

УДК 681.5, 517.935

DOI 10.18522/2311-3103-2024-2-59-69

И.Ю. Липко

ПРОГРАММНЫЙ КОМПЛЕКС СТЕНДА ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ КАЧКИ КАТАМАРАНА В ВЕТРОВОЛНОВЫХ УСЛОВИЯХ

Статья посвящена актуальной теме имитации качки катамарана и программному обеспечению, алгоритмам позволяющим реализовать эту имитацию. Актуальность вытекает из необходимости тестировать навигационные устройства и их алгоритмы, строить большие тренажёрные комплексы имитации объектов надводного, воздушного и наземного базирования. Данная статья будет сосредоточена на описании программного и алгоритмического обеспечения для создания имитатора качки катамарана, находящегося на поверхности воды, при воздействии на него ветроволновых воздействий. Имитационный стенд для воспроизведения качки катамарана включает в себя платформу Стюарта, датчики, микроконтроллеры и персональный компьютер оператора. Верхняя часть платформы перемещается подобно палубе катамарана, а нижняя часть жёстко закреплена. Результаты экспериментов учитываются датчиками, закреплёнными на подвижной части, и обрабатываются микроконтроллером-прототипом навигационной системы. Считается, что навигационная система должна получать данные об объекте качки и возвращать оператору вектор состояния, содержащий угол и скорость крена, угол и скорость дифферента, высоту и скорость вертикальной качки. Применяемые математические модели описывают динамику катамарана и системы управления, уменьшающую амплитуду качки, внешние возмущения морской волны, формируемые по спектру Пирсона-Московица. Для получения траекторий качки катамарана реализован алгоритм Рунге-Кутты 4-го порядка с фиксированным шагом, а для фильтрации и восстановления полного вектора состояний катамарана реализован фильтр Калмана по схеме с «с задержкой». Разработано программное обеспечение для воспроизведения крена катамарана на имитационном стенде, фильтрации и восстановления вектора состояния, графического отображения результатов эксперимента в виде графиков, сохранения результатов в файлах. Для лучшей интерпретации результатов входные параметры графического интерфейса и алгоритмов имеют простые наглядные параметры: скорость и направление ветра, начальное состояние катамарана и служебные, а результаты представляются в виде графиков. В статье представлено подробное описание взаимосвязи модулей программ, применяемых математических алгоритмов, указаны входные и выходные параметры. Результаты имитации показывают достаточное качество воспроизведения качки катамарана. Однако были выявлены незначительные погрешности, связанные с механическими ограничениями применённой кинематической модели.

Имитационный стенд; качка катамарана; фильтр Калмана; программный комплекс; платформа Стюарта.

I.Y. Lipko

THE SOFTWARE PACKAGE FOR THE CATAMARAN ROLL STAND IN WIND WAVE CONDITIONS SIMULATION

The article is devoted to the actual problem of catamaran roll imitation, software and algorithms that allow to implement this simulation. The relevance arises from the need to test navigation devices and their algorithms, to build large simulator complexes to reproduce roll and impulses of surface, air and ground-based vehicles. This article will focus on the description of software and algorithms for developing a roll simulator of a catamaran located on the sea surface when it disturbed by windwave influences. The simulation test bench for reproducing the catamaran roll includes a Stewart platform, sensors, micro-controllers and an operator's personal computer. The upper part of the platform moves like the deck of a catamaran, and the lower part is rigidly fixed. The experimental results are taken into account by sensors mounted on a movable part and processed by a microcontroller-navigation system prototype. It is believed that the navigation system should receive data about the pitching object and return to the operator a state vector containing the angle and speed of roll, angle and speed of pitch, height and speed of heave. The applied mathematical models describe the dynamics of a catamaran and a control system that reduces the amplitude of pitching, external disturbances of the sea wave formed by the Pearson-Moskowitz spectrum. To obtain the catamaran's roll trajectories, the Runge-Kutta 4th order algorithm with a fixed step is implemented, and for filtering and restoring the full vector of catamaran states the Kalman filter is implemented according to the "delayed" scheme. Software has been developed to reproduce the roll of a catamaran on a simulation test bench, filter and restore the state vector, graphically display the results of the experiment in the form of graphs, and save the results in files. For a better interpretation of the results, the input parameters of the graphical interface and algorithms have simple visual parameters: the wind speed and direction, the initial state of the catamaran and the service ones, and the results are presented in the form of graphs. The article provides a detailed description of the relationship between program modules, applied mathematical algorithms, and input and output parameters. The simulation results show a sufficient quality of reproduction of the rolling of the catamaran. However, minor errors were identified due to the mechanical limitations of the applied kinematic model.

Test bench; catamaran roll; Kalman filter; software package; Stewart platform.

Введение. Данная статья посвящена разработке программного обеспечения и реализации алгоритмов для имитационного стенда качки катамарана. Имитационные стенды – это сложные устройства, состоящие из нескольких частей и предназначенные для тестирования электронных устройств или программного обеспечения бортовых систем, тренировки людей в условиях, максимально приближенных к естественным. По сравнению с численным моделированием у исследователя появляются условия, когда его теоретические модели не ограничены математическим представлением реального мира.

Рассматриваемый нами имитационный стенд предназначен для моделирования крена катамарана, а также для тестирования и отладки бортовых измерительных систем в реальных условиях и сбора данных для проверки теоретических результатов.

Имитация является одним из важных этапов проверки и тестирования работоспособности электронных устройств и систем [1, 2]. При разработке устройств имитационное тестирование проводится после моделирования и перед эксплуатацией. Моделирование необходимо для синтеза элементов управления и проверки алгоритмов на применимость. А цель имитационного моделирования состоит в воспроизведении поведения исследуемой системы на основе статистических и других представлений о поведении системы, с учётом наиболее существенных взаимосвязей между её элементами.

Известно, что при исследовании и моделировании судов исследователи используют бассейн для генерации волн [3, 4]. В этом бассейне в воде качается небольшой прототип реального судна. Специальный генератор волн создает волны, похожие на настоящие морские. Исследователь следит за положением судна, углами крена и т.д. Этот метод очень хорош для исследования мореходных качеств судна, но имеет ряд ограничений, например, сложность повторяемости, стоимость одного эксперимента.

Обычно необходимые алгоритмы для стенда включают в себя методы подготовки, проведения экспериментов и сбора результатов. Подготовка эксперимента включает генерацию траектории. Проведение и сбор результатов включают в себя контроль всех

действий, выполняемых стендом, регистрацию и сохранение данных. В нашем случае это соответственно: генерация траектории крена катамарана, фильтрация данных, сбор данных в файлы.

Стенд состоит из платформы Стюарта с нижней неподвижной частью и верхней подвижной частью, соответствующей палубе судна, а также рабочего места оператора на базе персонального компьютера.

Известно несколько примеров использования платформы Стюарта [5] для моделирования крена судов и морских сооружений при волнении моря [6–9]. Некоторые из них используют две объединенные платформы для моделирования морских волн (нижняя платформа) и палубы судна (верхняя платформа).

Аналогичная ситуация складывается и в других областях: автомобилестроении, самолетостроении и вертолетостроении [10–14], где тренажеры используются для обучения операторов машин или для тестирования устройств. Для перемещения подвижной части платформы в большинстве из них используются линейные приводы, такие как гидравлические, электромеханические или пневматические. Эти приводы имеют множество преимуществ, но очень дороги. В противовес им мы используем вращающиеся шаговые двигатели, которые менее дороги.

У нас есть несколько концептуальных требований к имитационному стенду и рабочему пространству оператора:

1. Простые интуитивно понятные параметры проведения экспериментов.
2. Необходимо повторять эксперимент много раз. Идеальная последовательность действий представляет собой цикл: задать входные данные, запустить эксперимент, собрать, показать и изучить данные, экспортировать данные в другие приложения. Например, мы разработали несколько прогнозирующих алгоритмов [15], которые должны быть протестированы в условиях, приближенных к реальным. Повторение экспериментов позволяет проверить качество алгоритмов и программ.
3. Имитационный стенд должен поддерживать установку большинства доступных датчиков.

Дальнейшее содержание статьи состоит из разделов, описывающих каждую часть стенда. В начале будет дано общее описание, тракты данных, конструкция и оснащение стенда, а затем подробнее каждый из блоков стенда. В конце показаны и обсуждены результаты воспроизведения траекторий стендом и заключение.

Конструкция и аппаратное оснащение стенда. В данном разделе описывается общая структура имитационного стенда, программное обеспечение, связь блоков, позволяющих воспроизводить качку катамарана, приём данных с датчиков и их обработку.

Разработанный стенд предназначен для имитации качки судна под воздействием внешних ветроволновых воздействий и отладки алгоритмов навигационных систем, алгоритмов прогноза и оценки рисков ситуаций качки катамарана. Физически стенд состоит из следующих связанных частей: компьютер для вычислений, микроконтроллер и платформа Стюарта, непосредственно реализующая качку. В созданной платформе Стюарта используются вращающиеся шаговые двигатели и ножки, соединяющие нижнюю и верхнюю части платформы (рис. 1), что позволило снизить стоимость конструкции. Все моторы, силовая электроника, микроконтроллер и провода закрыты жёстким металлическим каркасом чтобы защитить от несанкционированного доступа (рис. 2).

Компьютер выполняет программные блоки (рис. 3), которые реализованы в виде программных модулей, предназначенных для генерации качки катамарана и преобразования траектории качки в углы поворота моторов платформы Стюарта. Эти блоки по сути «порождают» качку для имитации.

Блок «Генерация траекторий» – это программа, предназначенная для создания траекторий качки катамарана. Программа запускается с использованием параметров командной строки или специального файла конфигурации. Параметры содержат в себе вектор начального состояния катамарана, «зерно» генератора случайных чисел, служебные

флаги. В результате программа создает CSV-файл, содержащий в каждой строке обратный отсчет времени и соответствующее состояние катамарана: угол и скорость крена, угол и скорость дифферента, высоту вертикальной качки. Для генерации траекторий применяется метод интегрирования Рунге-Кутты 4-го порядка с дискретным шагом.

«Блок пересчёта» решает обратную задачу кинематики и переводит движение по траектории крена в соответствующие сигналы управления двигателями платформы Стюарта. CSV-файл с траекторией из предыдущего блока является входными данными программы. Результатом этого блока является CSV-файл со строками, которые содержат углы поворота двигателей платформы. Специальная программа реализует выполнение перемещений платформы.

Блок «Моторы» – это физическая система связанных моторов, реализующих движение подвижной части платформы. Блок «Датчик углов» – это устройство, используемое для получения данных о текущем наклоне платформы и скорости этого наклона. Блок «Фильтрация и восстановление» – это программные модули, выполняющие операции обработки сырых данных от датчика, восстановлению полного вектора состояния и передающие результат в блок обработки «Расчёт оценок», в котором могут быть реализованы любые алгоритмы, необходимые для отладки навигационного оборудования.

По сути, группа из блоков «Датчик углов», «Фильтрация и восстановление», «Расчёт оценок» представляет собой прототип навигационного устройства для считывания данных о крене и его последующей обработки. Данные о крене принимаются датчиком, установленным на подвижной поверхности платформы, и передаются в микрокомпьютер, где производятся расчеты и отображаются на экране компьютера.

В нашем случае, в качестве прототипа, используются: 3DOF-сенсор MPU6050 для получения углов и ускорений качки и микроконтроллер NUCLEO-F401RE, на котором реализованы алгоритмы обработки данных от датчика, фильтрации и восстановления вектора состояния. Описанные выше блоки являются модулями разработанных программ и написаны на C++ в Qt Creator.

Оператор стенда через графический интерфейс программы имеет возможность задавать погодные условия, в которых будет функционировать судно, наблюдать за текущим положением и графиками качки.

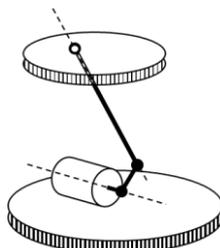


Рис. 1. Схема соединения одной опоры, мотора и оснований платформы Стюарта



Рис. 2. Стенд во время отладки устройства

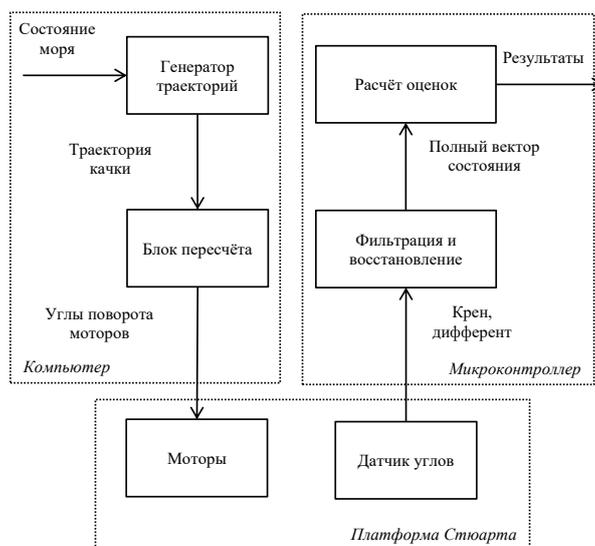


Рис. 3. Тракт данных и параметров стенда

Модель судна и возмущений. В данном разделе даётся описание математических моделей катамарана с системой управления которая используется для моделирования и воспроизведения траекторий качки стендом, внешних ветроволновых возмущений.

Катамаран представляет собой судно с двумя одинаковыми симметричными корпусами, соединёнными палубой. Рассматриваемый катамаран помимо пера руля снабжён подводным крылом и подруливающими клапанами. Подводное крыло катамарана обеспечивает плавность движения на высоких скоростях, путём создания силы F_T и момента M_T на корпус судна, а подруливающие клапана создают силу F_F и момент M_F для изменения курса (рис. 4). В совокупности эти исполнительные механизмы вместе с пером руля позволяют уменьшить качку катамарана на волнении. Оперирование этими устройствами реализуется с помощью системы управления по обратной связи.

Используемая модель катамарана имеет следующие размеры: длина 90 м, ширина 25.96 м, осадка 2.6 м, водоизмещение 734.54 м.

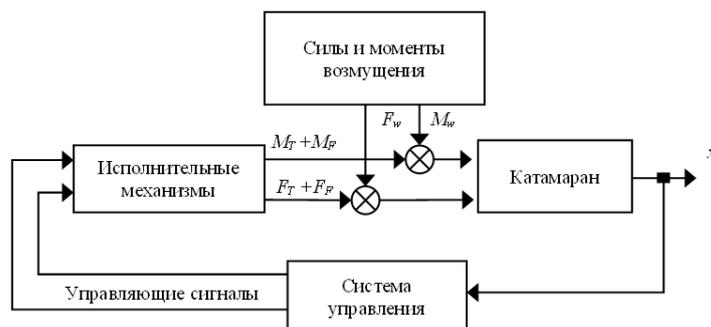


Рис. 4. Функциональная схема катамарана с системой управления

В общем случае в модели движения судна существует связь между всеми видами качки судна [16, 17]. Но в статье мы будем рассматривать только крен, угол дифферента и вертикальную качку катамарана, потому что именно они играют наибольшую роль при реализации различных опасных режимов [15, 18]. Уравнение пространства-состояния для катамарана:

$$\begin{cases} \dot{x} = Ax + B(u + w) \\ y = Cx \end{cases}, \quad (3)$$

где $x = [\dot{\eta}, \zeta, \eta, \zeta, \dot{\theta}, \theta]$ – вектор состояния, включающий скорость вертикальной качки, скорость дифферента, вертикальная качка, угол дифферента, скорость крена, крен; $u = \begin{bmatrix} F_T + F_F \\ M_T + M_F \end{bmatrix}$ – вектор управления, включающий силу и момент, возникающие на исполнительных механизмах, $w = \begin{bmatrix} F_w \\ M_w \end{bmatrix}$ – вектор ветроволновых возмущений, включающий силу и момент, действующие на корпус судна; матрицы объекта и управления

$$A = \begin{bmatrix} -0.9073 & -25.1097 & -14.1503 & -17.4945 & 0.001 & 0 \\ 0.0514 & -0.503 & 0.2442 & -12.4 & 0.001 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0.001 & 0.001 & -0.01 & -0.01 & -5 & -15 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.0082 & 0.0000083 \\ -0.00016 & 0.000017 \\ 0 & 0 \\ 0 & 0 \\ 0.002 & 0.0000083 \\ 0 & 0 \end{bmatrix},$$

и матрица наблюдений

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

поскольку вертикальную качку реального судна измерить довольно трудно. Но для генерации траекторий используется единичная матрица C .

Объект-катамаран (тройка матриц $\{A, B, C\}$ невырождена) замыкается обратной связью с линейно-квадратичным регулятором, обеспечивающим уменьшение амплитуды крена и дифферента (рис. 2), с матрицей

$$K = \begin{bmatrix} 59.3026 & 26.6427 & 18336.7 & 2006.9 & -1807.6 & -2950.1 \\ -1.2133 & 2.5576 & 14.2237 & -2.8566 & 2.1435 & 93.5241 \end{bmatrix}$$

в законе управления $u = -Kx$.

Морская волна, образуясь под влиянием ветра, создаёт силы и моменты, которые действуют на корпус катамарана. Энергия и статистика её распределения по частотам описывается спектром Пирсона-Московица.

Для симуляции действия морской волны используется формирующий фильтр, на выходе которого получаем силы и моменты (рис. 5). Сначала белый шум подаётся на вход блока линейной аппроксимации спектра Пирсона-Московица. Блок линейной аппроксимации спектра в качестве параметра принимает скорость ветра, влияющий на генерируемые погодные условия и желаемое состояние моря, а на выходе возвращает возвышение волны. Затем этот сигнал подаётся на блок линейной аппроксимации сил и моментов, в результате чего получаются силы и моменты, подаваемые на модель катамарана.

Спектр Пирсона-Московица $S(\omega) = 8.1 \cdot 10^{-3} \cdot 9.8^2 \omega^{-5} e^{-0.74 \cdot g/v \cdot \omega^{-4}}$ линеаризован передаточной функцией второго порядка (рис. 6)

$$H_{\text{ПМ}}(s) = \frac{K_{\omega} s}{s^2 + 2\lambda\omega_0 s + \omega_0^2},$$

где $K_{\omega} = 2\lambda\omega_0\sigma$ – коэффициент усиления, $\sigma^2 = \max_{0 < \omega < 2\pi} S(\omega)$ – константа, описывающая интенсивность волнения, λ – демпфирующий коэффициент, ω_0 – доминирующая частота волнения, $g = 9.8 \text{ м/с}^2$ – ускорение свободного падения, v – скорость ветра м/с.

Передаточная функция блока линейной аппроксимации силы и момента представляется константой K_f . Она подбирается так, чтобы амплитуда результирующего сигнала удовлетворяла значащей высоте волны из статистических данных.

Адекватность полученной модели внешних возмущений на этапе создания контролировалась путём сопоставления статистики выхода блоков и значащих высот волн с имеющимися справочными данными статистики ветра и волнения морей [19].

В результате симулирования комплексной модели «внешние возмущения – катамаран» генерируется траектория движения судна, включающая в себя углы крена, дифферента и курса, высоту возвышения палубы на волне.



Рис. 5. Схема формирования сил и моментов внешнего возмущения

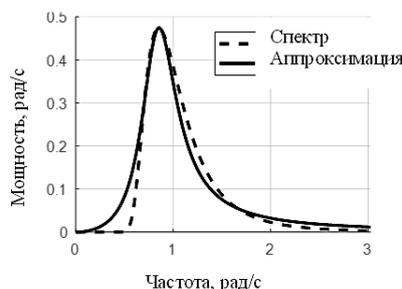


Рис. 6. Спектр Пирсона-Московица и его линейное приближение

Фильтрация данных. В данном разделе приводится описание результатов синтеза фильтра Калмана, используемого для фильтрации данных и восстановления полного вектора состояния катамарана.

Для получения фильтра Калмана комплексная система «катамаран с управлением и внешние возмущения» дискретизуется с тактом 0.01 с, поскольку этот фильтр будет функционировать на микроконтроллере. В результате синтеза с параметрами

$$Q = E(\gamma\gamma^T) = I_2 \cdot 10^7, \quad R = E(\zeta\zeta^T) = I_4 \cdot 10^{-1},$$

$$N = E(\gamma\zeta^T) = O_{2,4},$$

где I_n – единичная матрица размера n , $O_{n,m}$ – нулевая матрица с n строками и m столбцами, γ – шум в системе, ζ – шум при измерении; имеем дискретный наблюдатель состояния, реализованный по схеме «с задержкой» [20], учитывающий только значения прошлого ($n - 1$) шага

$$\hat{x}_{n+1,n} = (A_e - L_e C_e) \hat{x}_{n,n+1} + [B_e - L_e D_e \quad L_e] \cdot [O_{4,1} \quad y_{raw}]^T,$$

$$\begin{bmatrix} \hat{y}_{n,n-1} \\ \hat{x}_{n,n-1} \end{bmatrix} = \begin{bmatrix} C_e \\ I_6 \end{bmatrix} \hat{x}_{n,n-1} + \begin{bmatrix} D_e & O_{10,4} \\ O_{6,4} & O_{6,4} \end{bmatrix} \cdot \begin{bmatrix} O_{4,1} \\ y_{raw} \end{bmatrix},$$

где $\hat{x}_{n,n-1}$ – оценка состояния катамарана, $\hat{y}_{n,n-1}$ – оценка выхода катамарана, y_{raw} – 4x1 вектор считанных с датчика значений,

$$A_e = \begin{bmatrix} 0.8770 & -0.0091 & -0.0762 & -0.0119 & 0.0053 & 0.0244 \\ -1.0854 & -0.0195 & -0.7805 & -0.1685 & -0.0077 & 0.4085 \\ 0.0067 & 0.0975 & 0.5590 & 0.0360 & -0.0055 & 0.0040 \\ -0.1307 & -0.7889 & -5.9193 & -0.0658 & -0.0661 & -0.0626 \\ 0.0004 & -0.0007 & 0.0075 & 0.0013 & 0.9390 & 0.0930 \\ 0.0013 & 0.0231 & 0.0866 & 0.0200 & -1.1861 & 0.8465 \end{bmatrix},$$

$$B_e = \begin{bmatrix} 0.0597 & 0.0857 & 0.0001 & 0.0003 \\ -0.0674 & 0.5716 & 0.0026 & 0.0359 \\ -0.0077 & -0.0980 & -0.0004 & -0.0067 \\ 0.1209 & 0.7826 & 0.0026 & 0.0231 \\ -0.0004 & 0.0007 & 0.0001 & 0.0003 \\ -0.0013 & -0.0229 & -0.0002 & 0.0004 \end{bmatrix},$$

$$C_e = \begin{bmatrix} I_2 & 0_{2,4} \\ 0_{2,4} & I_2 \\ I_6 \end{bmatrix}, D_e = O_{10,4},$$

и матрица усиления, полученная в результате решения алгебраического уравнения Риккати:

$$L_e = \begin{bmatrix} 0.0597 & 0.0857 & 0.0001 & 0.0003 \\ -0.0674 & 0.5716 & 0.0026 & 0.0359 \\ -0.0077 & -0.0980 & -0.0004 & -0.0067 \\ 0.1209 & 0.7826 & 0.0026 & 0.0231 \\ -0.0004 & 0.0007 & 0.0001 & 0.0003 \\ -0.0013 & -0.0229 & -0.0002 & 0.0004 \end{bmatrix}.$$

Учитывая очевидные оптимизации (перемножение нулевых и единичных матриц) в вышеприведённом алгоритме, написан модуль, который используется в программе фильтрации и оценки вероятности.

Результаты моделирования. Рассмотрим несколько примеров работы имитационного стенда.

На рис. 7 показаны сгенерированная численным методом траектория (синим) и показания датчиков отработанного движения стендом (красным). Для рис. 7,а заданы параметры: скорость ветра 13.2 м/с, «зерно»-число 11622, вектор начального состояния нулевой, а для рис. 7,б заданы параметры: скорость ветра 12.9 м/с, «зерно»-число 25, вектор начального состояния нулевой

Результаты воспроизведения траектории достаточно хорошо повторяют основные движения, ошибка воспроизведения (рис. 8) находится на приемлемом уровне и в среднем не превышает 0.5° за редким исключением в 1° для крена и в среднем не превышает 0.03° за редким исключением в 0.05° для дифферента. Однако при более пристальном наблюдении можно увидеть, что в некоторых случаях не происходит воспроизведения граничных значений крена, как, например на рис., 50-60 с, 110-120 с, 160-170 с.

Основная причина этого была выявлена в ходе нескольких экспериментов и наблюдений. Она заключается в ограничении кинематических конфигураций опоры вала мотора и ножек, соединённых с подвижной платформой. Шаровые шарниры в подошве ножек являются слабым местом этой конструкции, что приводит к небольшому дребезжанию во время движения. Шум не виден глазами, но заметный датчиком. Чтобы уменьшить этот шум, мы используем балластный груз на верхней поверхности, который существенно снижает этот дребезг.

За время эксплуатации стенда проблем с программным обеспечением не возникало.

Заключение. Разработанное программное обеспечение успешно генерирует траекторию движения катамарана, качественно фильтрует и восстанавливает выходной сигнал. Механически стенд воспроизводит сигнал также качественно с редкими нюансами, заключающимися в особенностях кинематических связей платформы Стюарта.

Программное обеспечение прототипа навигационной системы имеет небольшой размер и оптимизированно, что позволяет выполнять его на микроконтроллере: фильтрация и восстановление вектора состояния с помощью фильтра Калмана.

Дальнейшая работа над улучшением стенда заключается в программной реализации других алгоритмов создания траекторий для других типов судов, алгоритмов фильтрации данных, улучшении механических характеристик стенда.

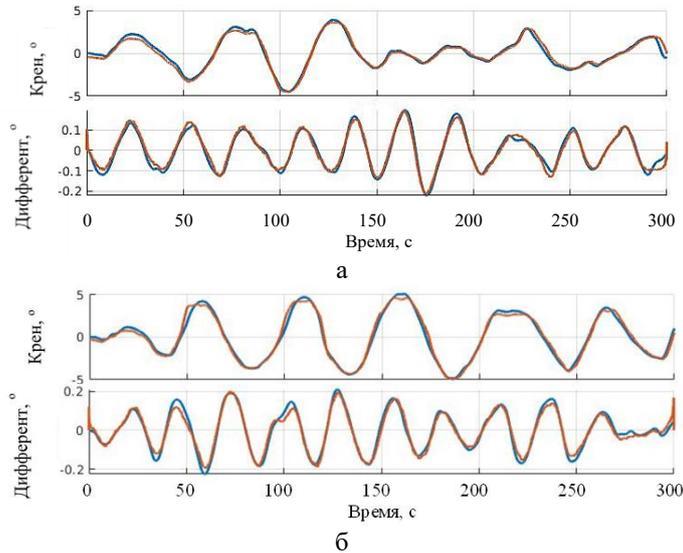


Рис. 7. Сравнение реализаций качки симуляции (синим) с воспроизведением имитатором (красным) для крена и дифферента

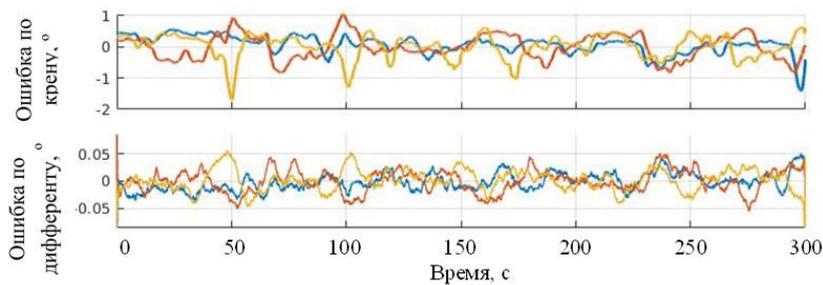


Рис. 8. Ошибки воспроизведения траекторий в трёх реализациях по крену и дифференту

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Вагуценко Л.Л., Вагуценко А.Л., Заичко С.И. Бортовые автоматизированные системы контроля мореходности. – Одесса: Феникс, 2005. – 272 с.
2. Бородай И.К., Живица С.Г., Платонов В.Г. Мореходность корабля и пути ее изучения (к 50-летию лаборатории мореходности) // Тр. Крыловского государственного научного центра. – 2018. – № 2 (384). – С. 147-152.
3. Akhmediev N., Chabchoub A., Hoffmann N., Onorato M. Super Rogue Waves: Observation of a Higher-Order Breather in Water Waves // Phys. Rev. – 2012. – P. 2-7.
4. Dematteis G., Grafke T., Vanden-Eijnden E. Rogue Waves and Large Deviations in Deep Sea // Proceedings of the National Academy of Sciences. – 2018. – Vol. 115, No. 5. – P. 855-860.
5. Stewart D. A Platform with Six Degrees of Freedom // Proceedings of the Institution of Mechanical Engineers. – 1965. – Vol. 180. – P. 371-378.
6. Andrievsky B. et al. Control of pneumatically actuated 6-DOF Stewart platform for driving simulator // 19th International Conference on Methods and Models in Automation and Robotics, Miedzyzdroje, 2014. – P. 663-668.
7. Krasnov E.I., Mikhaylov V.V., Sergeev S.L., Stuchenkov A.B. Development of a motion system for a training simulator based on Stewart platform // International Conference “Stability and Control Processes” in Memory of V.I. Zubov, St. Petersburg, 2015. – P. 99-101.

8. *Silva- uñiz Diego, Campos Julio, Fernández Enrique*. Stewart Platform Motion Control Automation with Industrial Resources to Perform Cycloidal and Oceanic Wave Trajectories // *Machines*. – 2022. – DOI: 10.711.10.3390/machines10080711.
9. *Lebrón ., Valente V., Sobczyk ., Perondi E*. Control of an Electrohydraulic Stewart Platform Manipulator as Vessels Motion Simulator // 9th FPNI Ph.D. Symposium on Fluid Power, 2016. – V001T01A033. – DOI: 10.1115/FPNI2016-1553.
10. *Villacís C. et al*. Real-time flight simulator construction with a network for training pilots using mechatronics and cyber-physical system approaches // 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering, Chennai. – P. 238-247.
11. *Advani Sunjoo, Hosman Ruud, Potter Mario*. Objective Motion Fidelity Qualification in Flight Training Simulators. – 2007. – DOI: 10.2514/6.2007-6802.
12. *Affan Muhammad, Ahmed Syed, Manek Ali, Uddin Riaz*. Design and Implementation of the Washout Filter for the Stewart-Gough Motion Platform. – 2019. – P. 415-419.
13. *Zhu Daoyang, Yan Jun, Duan Shaoli*. A Design of Nonlinear Scaling and Nonlinear Optimal Motion Cueing Algorithm for Pilot's Station // *Wuhan University Journal of Natural Sciences*. – 2021. – Vol. 26. – P. 513-520.
14. *Lukianov Petro, Kabanyachyi Volodymyr*. Mathematical model of stable equilibrium operation of the flight simulator based on the Stewart platform // *Aviation*. – 2023. – Vol. 27. – P. 119-128.
15. *Лунко И.Ю.* Моделирование редких событий качки беспилотного катамарана // *Морские интеллектуальные технологии*. – 2021. – № 4-1 (54). – С. 219-226.
16. *Fossen T.I*. Handbook of marine craft hydrodynamics and motion control. – Jhon Wiley & Sons, 2011. – ISBN 978-1-119-99149-6.
17. *Kramar V*. Development of the particular vessel mathematical models // *Annals of DAAAM and Proceedings of the International DAAAM Symposium*. – 2016. – Vol. 27, Issue 1. – P. 34-39.
18. *Galeazzi R., Christian H., Mogens B., Fossen T. I*. Stabilisation of parametric roll resonance by combined speed and fin stabilizer control // *Proceedings of the European Control Conference, Budapest, Hungary, 2009*. – P. 4895-4900.
19. *Лопатухин Л.И., Бухановский А.В., Чернышева Е.С.* Справочные данные по режиму ветра и волнения Берингова и Белого морей. Российский морской регистр судоходства, 2010. – 566 с. – ISBN 978-5-89331-135-8.
20. *Lewis F.L., Xie L., Popa D*. Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory. – 2nd ed. – CRC Press, 2017.

REFERENCES

1. *Vagushchenko L.L., Vagushchenko A.L., Zaichko S.I*. Bortovye avtomatizirovannye sistemy kontrolya morekhodnosti [Onboard automation systems of seakeeping]. Odessa: Feniks, 2005, 272 p.
2. *Boroday I.K., Zhivitsa S.G., Platonov V.G*. Morekhodnost' korablya i puti ee izucheniya (k 50-letiyu laboratorii morekhodnosti) [Ship seakeeping and the ways of studying it (to the 50th anniversary of Seakeeping Laboratory)], *Tr. Krylovskogo gosudarstvennogo nauchnogo tsentra* [Transactions of the Krylov State Research Centre], 2018, No. 2 (384), pp. 147-152.
3. *Akhmediev N., Chabchoub A., Hoffmann N., Onorato M*. Super Rogue Waves: Observation of a Higher-Order Breather in Water Waves // *Phys. Rev.* – 2012. – P. 2-7.
4. *Dematteis G., Grafke T., Vanden-Eijnden E*. Rogue Waves and Large Deviations in Deep Sea, *Proceedings of the National Academy of Sciences*, 2018, Vol. 115, No. 5, pp. 855-860.
5. *Stewart D*. A Platform with Six Degrees of Freedom, *Proceedings of the Institution of Mechanical Engineers*, 1965, Vol. 180, pp. 371-378.
6. *Andrievsky B. et al*. Control of pneumatically actuated 6-DOF Stewart platform for driving simulator, *19th International Conference on Methods and Models in Automation and Robotics, Miedzyzdroje, 2014*, pp. 663-668.
7. *Krasnov E.I., Mikhaylov V.V., Sergeev S.L., Stuchenkov A.B*. Development of a motion system for a training simulator based on Stewart platform, *International Conference "Stability and Control Processes" in memory of V.I. Zubov, St. Petersburg, 2015*, pp. 99-101.
8. *Silva- uñiz Diego, Campos Julio, Fernández Enrique*. Stewart Platform Motion Control Automation with Industrial Resources to Perform Cycloidal and Oceanic Wave Trajectories, *Machines*, 2022. DOI: 10.711.10.3390/machines10080711.
9. *Lebrón ., Valente V., Sobczyk ., Perondi E*. Control of an Electrohydraulic Stewart Platform Manipulator as Vessels Motion Simulator, *9th FPNI Ph.D. Symposium on Fluid Power, 2016*. V001T01A033. DOI: 10.1115/FPNI2016-1553.

10. Villacís C. et al. Real-time flight simulator construction with a network for training pilots using mechatronics and cyber-physical system approaches, *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering, Chennai*, pp. 238-247.
11. Advani Sunjoo, Hosman Ruud, Potter Mario. Objective Motion Fidelity Qualification in Flight Training Simulators, 2007. DOI: 10.2514/6.2007-6802.
12. Affan Muhammad, Ahmed Syed, Manek Ali, Uddin Riaz. Design and Implementation of the Washout Filter for the Stewart-Gough Motion Platform, 2019, pp. 415-419.
13. Zhu Daoyang, Yan Jun, Duan Shaoli. A Design of Nonlinear Scaling and Nonlinear Optimal Motion Cueing Algorithm for Pilot's Station, *Wuhan University Journal of Natural Sciences*, 2021, Vol. 26, pp. 513-520.
14. Lukianov Petro, Kabanyachyi Volodymyr. Mathematical model of stable equilibrium operation of the flight simulator based on the Stewart platform, *Aviation*, 2023, Vol. 27, pp. 119-128.
15. Lipko I.Yu. Modelirovanie redkikh sobyitij kachki bespilnogo katamarana [Rare roll event modeling of unmanned catamaran], *Morskie intellektual'nye tekhnologii* [Marine intellectual technologies], 2021, No. 4-1 (54), pp. 219-226.
16. Fossen T.I. Handbook of marine craft hydrodynamics and motion control. Jhon Wiley & Sons, 2011. ISBN 978-1-119-99149-6.
17. Kramar V. Development of the particular vessel mathematical models, *Annals of DAAAM and Proceedings of the International DAAAM Symposium*, 2016, Vol. 27, Issue 1, pp. 34-39.
18. Galeazzi R., Christian H., Mogens B., Fossen T. I. Stabilisation of parametric roll resonance by combined speed and fin stabilizer control, *Proceedings of the European Control Conference, Budapest, Hungary, 2009*, pp. 4895-4900.
19. Lopatukhin L.I., Bukhanovskiy A.V., Chernysheva E.S. Spravochnye dannye po rezhimu vetra i volneniya Beringova i Belogo morey [Reference data on the wind and wave regime of the Bering and White Seas]. Rossiyskiy morskoy registr sudokhodstva [Russian Maritime Register of Shipping], 2010, 566 p. ISBN 978-5-89331-135-8.
20. Lewis F.L., Xie L., Popa D. Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory. 2nd ed. CRC Press, 2017.

Статью рекомендовал к опубликованию д.т.н., профессор В.Х. Пшихопов.

Липко Иван Юрьевич – Севастопольский государственный университет; e-mail: ivanlipko@yandex.ru; г. Севастополь, Россия; тел. +79787391279; старший преподаватель; научный сотрудник.

Lipko Ivan Urievich – Sevastopol State University; e-mail: ivanlipko@yandex.ru; Sevastopol, Russia; phone +79787391279; senior lecturer; research officer.

УДК 631.171:303.732.4

DOI 10.18522/2311-3103-2024-2-69-82

В.В. Соловьев, А.Я. Номерчук, Р.К. Филатов

СИСТЕМНЫЙ АНАЛИЗ НАЗЕМНОЙ РОБОТИЗИРОВАННОЙ ПЛАТФОРМЫ СЕЛЬСКОХОЗЯЙСТВЕННОГО НАЗНАЧЕНИЯ

Целью данной работы является проведение системного анализа мобильных роботизированных платформ, которые могут использоваться в сельском хозяйстве для транспортировки грузов и прополки сорняков. Данное исследование является актуальным из-за увеличения численности населения, уменьшения пахотных площадей, естественного оттока населения из сельской местности и снижения количества сельскохозяйственной техники. Для достижения поставленной цели в работе сформировано «дерево» целей системного анализа конструктивного исполнения платформ, которые предваряют и декомпозируют этапы проектирования и разработки роботов сельскохозяйственного назначения. Из-за наличия нечетких и вербальных показателей оценки экспертами роботизированных платформ, авторы предлагают определять их в виде нечетких интервалов, которые с применением аддитивной свертки позволяют получить комплексный показатель, который может представлять либо в нечетком виде, либо в виде пессимистической, оптимистической или нейтральной оценок. При этом весовые коэффициенты аддитивной свертки также можно представить в нечетком виде. Для этого предложены операции умножения и сложения нечетких интервалов. Для проведения имитационного моделирования представлена струк-

тура программного обеспечения с применением объектно-ориентированного подхода. За счет перегрузки классических операций сложения и умножения удалось реализовать алгебраические операции с нечеткими интервалами без усложнения вычислений. Результаты моделирования подтвердили работоспособность подхода и позволили определить конструктивное исполнение, компоновку, двигатели и движители для сельскохозяйственной платформы. Предложенные методы можно использовать перед этапами проектирования и разработки роботов различного назначения, а использование показателей в нечетком виде позволяет снизить нагрузку на экспертов.

Роботизированная платформа; системный анализ; аддитивная свертка; нечеткий интервал; оценка конструкции.

V.V. Soloviev, A.Y. Nomerchuk, R.K. Filatov

SYSTEM ANALYSIS OF A GROUND ROBOTIC PLATFORM FOR AGRICULTURAL PURPOSE

The aim of this work is to conduct a systemic analysis of mobile robotic platforms that can be used in agriculture for cargo transportation and weed control. This study is relevant due to the increasing population, decreasing arable land, natural population outflow from rural areas, and reduction in agricultural machinery. To achieve the set goal, a "tree" of objectives for the systemic analysis of the constructive implementation of platforms has been formed, which precedes and decomposes the stages of designing and developing agricultural robots. Due to the presence of fuzzy and verbal evaluation indicators by experts of robotic platforms, the authors suggest defining them in the form of fuzzy intervals, which, with the use of additive convolution, allow obtaining a composite indicator that can be presented either in fuzzy form or in the form of pessimistic, optimistic, or neutral assessments. At the same time, the weighting coefficients of additive convolution can also be presented in fuzzy form. For this purpose, operations of multiplication and addition of fuzzy intervals are proposed. To conduct simulation modeling, the structure of software is presented using an object-oriented approach. By overloading classical addition and multiplication operations, it was possible to implement algebraic operations with fuzzy intervals without complicating calculations. The modeling results confirmed the feasibility of the approach and allowed determining the constructive implementation, layout, engines, and actuators for the agricultural platform. The proposed methods can be used before the stages of designing and developing robots for various purposes, and the use of indicators in fuzzy form allows reducing the burden on experts.

Robotic platform; systemic analysis; additive convolution; fuzzy interval; construction assessment.

Введение. Потребность в увеличении количества продуктов питания, при ограниченности пахотных земель, естественном оттоке населения из сельской местности и снижении количества сельскохозяйственной техники требуют интенсивной роботизации сельского хозяйства [1]. Сельскохозяйственная робототехника может послужить толчком к цифровизации сельского хозяйства за счет интеграции различных технологий, т.к. роботы могут использоваться в полном цикле производства продукции АПК от предпосевной подготовки земель до упаковки готовой продукции [2]. Отечественные и зарубежные исследователи публикуют статьи по разработке конструкции, систем управления и систем распознавания образов для наземных мобильных платформ сельскохозяйственного назначения.

В общем случае сельскохозяйственная платформа содержит либо гусеничный, либо колесный движитель [3] и оснащается сенсорами, приемником сигнала GPS/ГЛОНАСС, бортовым вычислителем, средствами передачи данных пользователю и навесным оборудованием в виде манипуляторов, лазерных пропольщиков и т.п.

В работе [4] описана структура системы интеллектуального управления в садоводстве, которая рассчитывает нормы внесения удобрений и сроки проведения защитных мероприятий. Указано, что такая система может быть развернута на роботизированной платформе, которая оснащена системой технического зрения с алгоритмами интеллектуальной обработки информации, системой адаптивного вождения в полевых условиях, навесным оборудованием в виде распылителей удобрений.

Автор работы [5] представил роботов для обработки почвы, посадки растений, борьбы с сорняками, сбора урожая. Указано, что роботы должны иметь полный привод для работы на неподготовленном грунте.

В статье [6] в качестве объекта исследования представлен робот, который может перемещаться в междурядьях высокорослых растений, на примере кукурузы.

В публикации [7] представлен автономный робот на базе трактора, способный обработать 99,5% листьев растений по результатам тестовых испытаний. В работе [8] авторы также модифицировали серийный мини-трактор и интегрировали систему автономного управления движением и навесным оборудованием.

В работах [9–12] представлены четырехколесные роботы сельскохозяйственного назначения с поворотными и неповоротными колесами. Также в статье [13] описана концепция платформы с маломощным дизельным двигателем, который приводит в движение четыре гидромотора колес. Платформа имеет полный привод и поворотные колеса.

В работе [14], в отличие от других работ, рассматривается модульная роботизированная платформа сельскохозяйственного назначения. Авторы отмечают, что уровень развития технологий способствует внедрению робототехники в АПК, но отсутствие правил, стандартов и протоколов затрудняют развитие данной отрасли. Предлагается полноприводный робот с электромоторами и редукторами на каждом колесе.

Публикация [15] посвящена разработке модульного робота с гусеничным движителем с электроприводами и редукторами.

По результатам обзора публикаций можно отметить, что авторы предлагают различные варианты роботизированных платформ для сельского хозяйства без системного осмысления конструктивного и компоновочного исполнения, с привязкой к спектру решаемых задач.

Постановка задачи. Предметом исследования является роботизированная платформа сельскохозяйственного назначения для транспортировки грузов. Также на платформе может быть установлен дельта-манипулятор для прополки сорняков на кукурузных полях с шириной междурядий 45 см и предельной высотой культуры 50 см. При большей высоте кукурузы необходимость в прополке отпадает.

Требуется:

- ◆ предложить метод оценки конструктивного исполнения роботизированной платформы в зависимости от кинематической схемы и размещения полезной нагрузки;
- ◆ предложить метод оценки двигателей и движителей для выбора технических средств, ориентированных на сельскохозяйственное применение;
- ◆ разработать программное обеспечение, позволяющее провести моделирование и определить оптимальную конструкцию платформы.

Этапы системного анализа для решения задачи разработки роботизированного комплекса. При разработке роботизированного комплекса будем руководствоваться этапами системного анализа по Н.П. Федоренко [16], как наиболее приближенного к решению технических задач и предполагающего выполнение следующих этапов.

Этап 1 – Формулирование проблемы. Для замены ручного труда и исключения применения гербицидов необходима роботизированная платформа сельскохозяйственного назначения, на которую можно установить дельта-манипулятор для прополки сорняков в автономном режиме в полевых условиях, обеспечивая безопасное перемещение в среде и сохранение культурных растений.

Этап 2 – Определение целей. Совокупность целей можно представить в виде дерева целей, как представлено на рис. 1. Цели системного анализа в данном случае – разработка роботизированной платформы (0), разработка привода (01), разработка корпуса (02), анализ кинематических схем (011), анализ движителей (0111), анализ двигателей (0112), формирование требований к двигателю (01121), формирование требований движителю (01111), анализ вариантов исполнения (021), анализ источников питания (0211), анализ полезной нагрузки (0212), формирование требований к автономности (02111), формирование требований к полезной нагрузке (02121).

В соответствии с разработанной иерархией целей далее будет предложена двухэтапная процедура определения оптимальной конструкции платформы:

- ◆ на первом этапе вычисляются частные оценки различных платформ и определяется конфигурация с соответствующими кинематическими характеристиками и расположением нагрузки (движение по иерархии вверх);

♦ на втором этапе формируются требования к двигателю, движителю, нагрузке и автономности, как к оптимальной совокупности для различных конфигураций (движение по иерархии вниз).

По результатам двух этапов будет определена кинематическая схема платформы, расположение полезной нагрузки, требуемые двигатели, движители и параметры автономности.

Этап 3 – Систематизация информации. Согласно дереву целей для роботизированной платформы можно выделить следующие факторы:

01111 – формирование требований к движителю: эффективность на грунте, надежность, ремонтпригодность, серийное производство, удельная нагрузка на грунт и т.п.;

01121 – формирование требований к двигателю: КПД, энергозатраты, диапазон скоростей, серийное производство, надежность, компактность, обслуживание и т.п.;

0111 – анализ движителей: колесный, гусеничный, комбинированный;

0112 – анализ двигателей: электрический коллекторный, электрический бесколлекторный, гидравлический, пневматический и т.п.;

011 – анализ кинематических схем: трех- (S1), четырех- (S2), шестиколесная (S3), гусеничная (S4), полноприводная (S5), с поворотными колесами (S6), с пассивными поворотными колесами (S7), с пассивными неповоротными колесами (S8);

02121 – формирование требований к полезной нагрузке: емкость с жидкостью, грузы в коробках, грузы в мешках, дельта-манипулятор;

02111 – формирование требований к автономности: работоспособность не менее 8 часов в сутки;

0212 – анализ полезной нагрузки: простота погрузки/выгрузки, простота установки/демонтажа;

0211 – анализ источников питания: аккумуляторы, дизель-генераторы, солнечные панели, комбинированные схемы;

021 – анализ вариантов исполнения: размещение оборудования и нагрузки над уровнем поверхности, боковое размещение, размещение спереди, размещение сзади, захват не менее одного междуружья.

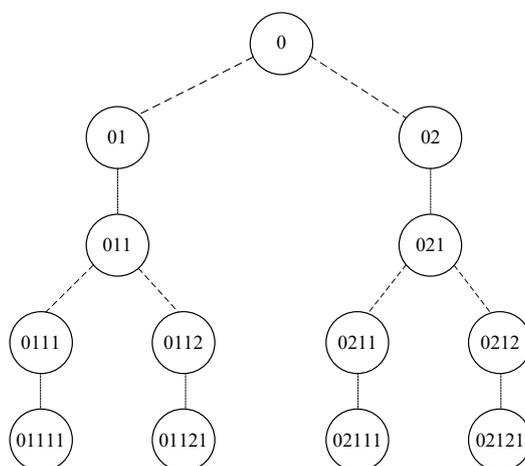


Рис. 1. Дерево целей для разработки роботизированной платформы

Этап 4 – Проработка альтернатив. Одним из вариантов создания альтернативных конструкций роботизированной платформы является их автоматическая генерация с оценкой на каждом уровне иерархии дерева и комплексной оценкой в целом, что позволит исключить неэффективные решения.

В соответствии с деревом целей возможны разные кинематические схемы платформ с различными сочетаниями движителей, как представлено в табл. 1.

Таблица 1

Варианты кинематических схем роботизированных платформ

	S1	S2	S3	S4	S5	S6	S7	S8
S1	–	–	–	–	–	+++	+	+
S2	–	–	–	–	+	+	+	+
S3	–	–	–	–	+	+	+	+
S4	–	–	–	+	–	–	–	–

Согласно таблице, при заданных условиях можно реализовать 12 кинематических схем платформ, представленных на рис. 2.

Для каждой кинематической схемы возможны различные варианты размещения полезной нагрузки и оборудования, что приводит к 48 вариантам компоновки платформ.

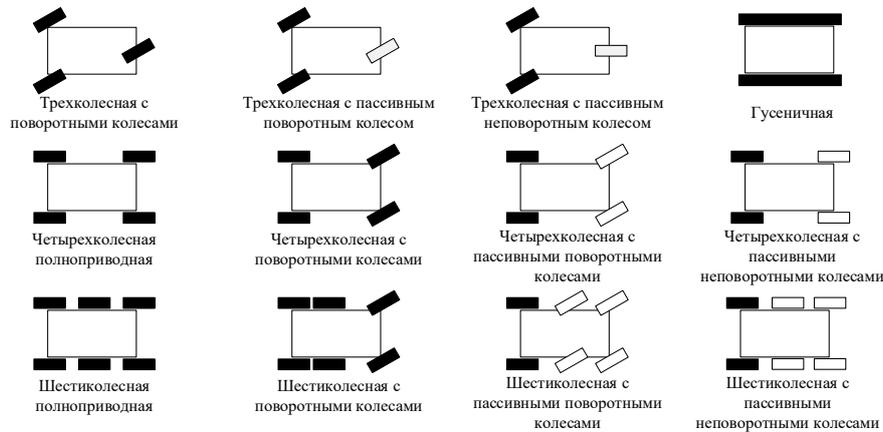


Рис. 2. Базовые кинематические схемы роботизированных платформ

Метод оценки конструктивного исполнения роботизированной платформы.

Для отбора альтернатив необходимо получить численные оценки качества каждого варианта компоновки платформы и сформировать критерий оценки.

При разработке метода формирования критерия качества будем пользоваться индуктивными рассуждениями, т.е. покажем выкладки для формирования критерия оценки для платформы в целом, а затем «спустимся» по дереву вниз и получим оценки, тех компонентов, которые могут дать требуемую оценку платформы.

Роботизированную платформу будем оценивать по следующим показателям: маневренность (M), удельная нагрузка на грунт (L), грузоподъемность (C), сложность конструкции (S), функциональность в прополке сорняков (P).

На маневренность оказывает влияние тип двигателей, тип движителей, их количество и компоновка. Удельная нагрузка определяется типом движителей и их количеством. Грузоподъемность определяется количеством движителей и компоновкой. Сложность конструкции определяется типом движителей, двигателей и компоновкой. Функциональность в прополке сорняков определяется компоновкой. Взаимосвязь показателей с составляющими можно отразить с использованием логики отношений на основе теоретико-множественного подхода [17].

Например, введем бинарное отношение R из множества V в множество M :

$$R \subset V \times M, \tag{1}$$

которое определяет вклад в оценку маневренности платформы отдельных составляющих.

Таким образом получается, что при заданных исходных данных оценки отдельных показателей определяются соответствующими отношениями на множествах исходных данных и обобщающий критерий можно представить в следующем виде:

$$J=F(M, L, C, S, P), \quad (2)$$

где F – функция свертки, связывающая показатели M, L, C, S, P .

Для сложных технических систем при выборе функции F необходимо [18]: обосновать допустимость свертки, с учетом однородности критериев; выполнить нормализацию критериев; учесть приоритет критериев; построить функцию свертки.

Традиционным методом проверки однородности является критерий Стьюдента. Для его применения необходима выборка значимых параметров за некоторый промежуток времени, на основе которой рассчитывается уровень значимости и проверяется гипотеза однородности.

Нормализация показателей в (2) необходима из-за разных мощностей множеств M, L, C, S, P для унификации их размерности и диапазона изменения. Таким образом, возникает необходимость сведения критериев к единой мере. Обозначим частный критерий эффективности k , а его максимальное и минимальное значение как k_{\max} и k_{\min} соответственно. Тогда согласно табл. 2, можно выделить несколько подходов к нормализации.

Таблица 2

Подходы к нормализации критериев

№	Выражение	Диапазон изменения
1	$\bar{k} = \frac{k_{\min}}{k}$	$\left[\frac{k_{\min}}{k_{\max}}, 1 \right]$
2	$\bar{k} = 1 - \frac{k}{k_{\max}}$	$\left[0, 1 - \frac{k_{\min}}{k_{\max}} \right]$
3	$\bar{k} = \frac{k_{\max} - k}{k_{\max} - k_{\min}}$	$[0, 1]$

Относительно процесса оценки конструктивного исполнения платформы критерий (2) является максимизируемым. Любое из указанных выражений подходит для унификации размерности критериев. Однако, следует обратить внимание, что выражения 1 и 2 используются со значениями k_{\max} и k_{\min} взятыми среди всех показателей из (2). В выражении 3 значения k_{\max} и k_{\min} берутся для каждого из рассматриваемых критериев. Таким образом, если априорно достоверно известны точные значения k_{\max} и k_{\min} всех критериев, для нормализации (2) можно использовать выражения 1 и 2. В противном случае необходимо использовать выражение 3.

Существует несколько подходов к учету приоритета критериев. Наибольшее распространение получил учет приоритета за счет введения весовых коэффициентов критериев, каждый из которых определяет вклад частного критерия в общую оценку качества. В нашем случае аддитивная свертка реализует принцип справедливой компенсации абсолютных значений нормированных критериев и записывается в виде [19]:

$$J = \sum_{i=1}^n \alpha_i \bar{k}_i, \quad (3)$$

где α_i – весовой коэффициент критерия \bar{k}_i ; J – значение обобщенного критерия; n – количество частных критериев.

Также следует отметить, что затруднительно получить номинальные значения элементов M, L, C, S, P как значимых параметров в виде четких численных величин. Это связано с тем, что данные параметры изменяются в некоторых интервалах. Поэтому можно их определять треугольными нечеткими числами или нечеткими интервалами. В данном случае под нечетким интервалом будем понимать выпуклую нечеткую величину, функция принадлежности которой задается следующим образом:

$$\forall u, v, \exists \omega \in [u, v], \mu_Q(\omega) \geq \min(\mu_Q(u), \mu_Q(v)). \quad (4)$$

Интервалы задаются четверкой параметров $\tilde{M}=(\underline{m}, \bar{m}, \alpha, \beta)$, где \underline{m} и \bar{m} – соответственно нижнее и верхнее модальное значение интервала, а α и β представляют собой левый и правый коэффициент нечеткости. Кроме того, треугольное нечеткое число будем рассматривать как частный случай нечеткого интервала с $\underline{m} = \bar{m}$. Также при $\underline{m} = \bar{m}$ и $\alpha = \beta$ получим четкое описание параметра.

Удобством нечеткого представления параметров является то, что их значения автоматически нормированы степенями принадлежности $\mu(p)$, но нормализация частных критериев в любом случае необходима.

В (3) весовые коэффициенты критериев обычно определяются с использованием экспертных суждений по оценке значимости частных критериев [20]. При этом экспертные оценки также можно представить в виде нечетких интервалов. Таким образом, выражение (3) можно переписать в виде:

$$\tilde{J} = \sum_{i=1}^n \tilde{\alpha}_i \tilde{k}_i, \quad (5)$$

где $\tilde{\alpha}_i$ – оценка значимости критерия \tilde{k}_i ; \tilde{J} – обобщенный критерий в виде нечеткого интервала.

Из (5) следует, что необходимо определить операции умножения и сложения нечетких интервалов.

Сумма двух интервалов $\tilde{M}_i=(\underline{m}_i, \bar{m}_i, \alpha_i, \beta_i)$ и $\tilde{M}_j=(\underline{m}_j, \bar{m}_j, \alpha_j, \beta_j)$ – есть также интервал $\tilde{M}=(\underline{m}, \bar{m}, \alpha, \beta)$, где

$$\alpha = \alpha_i + \alpha_j; \quad \beta = \beta_i + \beta_j; \quad \underline{m} = \underline{m}_i + \underline{m}_j, \quad \bar{m} = \bar{m}_i + \bar{m}_j. \quad (6)$$

Обобщим полученные формулы на k интервалов:

$$\alpha = \sum_{i=1}^k \alpha_i; \quad \beta = \sum_{i=1}^k \beta_i; \quad \underline{m} = \sum_{i=1}^k \underline{m}_i, \quad \bar{m} = \sum_{i=1}^k \bar{m}_i. \quad (7)$$

Произведение двух интервалов $\tilde{M}_i=(\underline{m}_i, \bar{m}_i, \alpha_i, \beta_i)$ и $\tilde{M}_j=(\underline{m}_j, \bar{m}_j, \alpha_j, \beta_j)$ – есть также интервал $\tilde{M}=(\underline{m}, \bar{m}, \alpha, \beta)$, где

$$\begin{aligned} \alpha &= \alpha_i \alpha_j, \quad \beta = \beta_i \beta_j, \\ \underline{m} &= \min \{ \underline{m}_i \cdot \underline{m}_j, \bar{m}_i \cdot \underline{m}_j, \underline{m}_i \cdot \bar{m}_j, \bar{m}_i \cdot \bar{m}_j \}, \\ \bar{m} &= \max \{ \underline{m}_i \cdot \underline{m}_j, \bar{m}_i \cdot \underline{m}_j, \underline{m}_i \cdot \bar{m}_j, \bar{m}_i \cdot \bar{m}_j \}. \end{aligned} \quad (8)$$

С использованием (6)-(8) вычисляется обобщенный критерий (5) в виде нечеткого интервала. Здесь следует отметить, что если необходимо оперировать четкими значениями обобщенного критерия, то необходимо определить тип оценок. Например, пессимистическую, оптимистическую и нейтральную оценки, которые можно вычислить по формулам:

$$J_p = \max(\tilde{J}), \quad (9)$$

$$J_o = \min(\tilde{J}), \quad (10)$$

$$J_N = \frac{\sum_{u_j \in \tilde{J}} u_j}{|\tilde{J}|}. \quad (11)$$

На основании вышеизложенного предлагается следующий метод оценки конструктивного исполнения роботизированной платформы:

Этап 1. Определение множеств оценочных $P = \{P_1, P_2, \dots, P_n\}$ параметров роботизированной платформы.

Этап 2. Введение бинарных отношений R на множествах P .

Этап 3. Выявление множеств значимых параметров $P^* \subseteq P$ технологического процесса.

Этап 4. Определение номинальных значений значимых параметров \tilde{P}_0^* в виде нечетких интервалов.

Этап 5. Определение параметров нечетких интервалов $\tilde{P}^* = (\underline{p}, \bar{p}, \alpha, \beta)$ для измеряемых параметров.

Этап 6. Определение экспертных оценок значимости критериев $\tilde{\alpha}$.

Этап 7. Введение операций на нечетких интервалах.

Этап 8. Введение частных критериев эффективности $\tilde{K} = F(\tilde{P}_0^*, \tilde{P}^*, \tilde{\alpha})$.

Этап 9. Вычисление обобщенного критерия \tilde{J} в виде нечеткого интервала.

Этап 10. Определение типа оценок и их вычисление.

Метод оценки двигателей и движителей. На основе оценки конструктивного исполнения роботизированной платформы необходимо выбрать конкретные двигатели и движители с учетом требований к автономности.

Согласно дереву целей, построенному при системном анализе, в методе необходимо учесть:

- ◆ оценки движителя: эффективность на грунте, надежность, ремонтпригодность, серийное производство, удельная нагрузка на грунт и т.п.;
- ◆ оценки двигателя: КПД, энергозатраты, диапазон скоростей, серийное производство, надежность, компактность, обслуживание и т.п.;
- ◆ типы движителей: колесный, гусеничный, комбинированный;
- ◆ типы двигателей: электрический коллекторный, электрический бесколлекторный, гидравлический, пневматический и т.п.;
- ◆ источники питания: аккумулятор, дизель-генератор, солнечная панель, комбинированные схемы.

Тогда для оценки двигателей и движителей предлагается следующий метод:

Этап 1. Формирование базы двигателей на основе информации из справочников и каталогов производителей.

Этап 2. Оценка параметров двигателей: КПД, энергозатраты, диапазон скоростей, серийное производство, надежность, компактность, обслуживание, в том числе в нечетком представлении.

Этап 3. Свертка оценок и вычисление оценок двигателей по формуле (3) или (5).

Этап 4. Формирование базы движителей на основе информации из справочников, каталогов производителей.

Этап 5. Оценка параметров движителей: эффективность на грунте, надежность, ремонтпригодность, серийное производство, удельная нагрузка на грунт, в том числе в нечетком представлении.

Этап 6. Свертка оценок и вычисление оценки движителей по формуле (3) или (5).

Этап 7. Формирование базы источников питания на основе информации из справочников, каталогов производителей.

Этап 8. Оценка источников питания для выбранной конфигурации платформ.

Этап 9. Поиск всех комбинаций из п. 3, п. 6 и п. 8.

Этап 10. Свертка оценок комбинаций из п. 9 по формуле (3) или (5).

Этап 11. Выбор комбинации с максимальной оценкой.

Модульная структура программного обеспечения. Для оценки функциональности роботизированной платформы разработано программное обеспечение на языке Python со следующей модульной структурой:

- ◆ модуль чтения исходных данных: содержит информацию о параметрах нечетких переменных, весовых коэффициентах нечетких чисел, частных оценках платформ, которые загружаются из текстового файла;

◆ библиотека классов: содержит описание класса нечеткая переменная (FuzzyVar), описание класса лингвистическая переменная (LinguisticVar), описание класса платформы (Robot);

◆ основной программный модуль: содержит описание главной формы приложения, виджетов для вывода списка возможных конфигураций платформ, текста с описанием платформ и их изображений из графических файлов формата *.png, а также вывод графиков лингвистических переменных, при создании экземпляров класса Robot.

Класс FuzzyVar описывает нечеткие переменные в виде нечетких интервалов и операции над нечеткими числами. Класс LinguisticVar описывает лингвистические переменные, которые создаются на базе нечетких переменных. Класс Robot описывает параметры роботизированных платформ, включая лингвистические переменные параметров: маневренность (*M*), удельная нагрузка на грунт (*L*), грузоподъемность (*C*), сложность конструкции (*S*), функциональность в прополке сорняков (*P*).

Классы LinguisticVar и Robot не наследуют свойства и методы, а получают их напрямую в виде списка в методе инициализации, чем обеспечивается связь между классами.

Для реализации операций с нечеткими интервалами потребовалось перегрузить методы add, mul и реализовать выражения (6)-(8) в программном коде. Этот подход позволил выполнять операции с переменными, заданными в виде нечетких интервалов, как с обычными числами. При этом, возвращаемый результат перегруженных функций “_add_”, “_mul_” также оформлен как новый экземпляр класса FuzzyVar, что позволяет выполнять операции над двумя и более нечеткими интервалами в едином математическом выражении и не разбивать его на отдельные операции.

В табл. 3 представлены значимые параметры для оценки роботизированной платформы, количество и значения термов для их описания в виде лингвистических переменных. В общем случае количество терм-множеств может быть различной и корректироваться в зависимости от задачи, в которой будет использован метод.

Таблица 3

Лингвистические переменные параметров платформы

Параметр	Номер терма				
	1	2	3	4	5
Маневренность	Низкая	Ниже среднего	Средняя	Выше среднего	Высокая
Нагрузка на грунт	Высокая	Средняя	Низкая	-	-
Грузоподъемность	Низкая	Средняя	Высокая	-	-
Сложность конструкции	Высокая	Средняя	Низкая	-	-
Функциональность	Низкая	Ниже среднего	Средняя	Выше среднего	Высокая

Следует отметить, что терм-множества лингвистических переменных параметров платформы введены таким образом, что термы с минимальными номерами, соответствуют худшим параметрам, а максимальные термы – лучшим, что позволило исключить операцию вычитания при расчете нечеткой свертки оценки конструктивного исполнения платформ.

Для оценки роботизированных платформ предлагается система кодирования параметров, которая в общем случае описывается следующей конструкцией:

$$w^* _r^* _p^* _pr^* _&,$$

где w^* – количество колес; r^* - количество поворотных колес; p^* – количество пассивных колес; pr^* – количество пассивных поворотных колес; $\&$ – место расположения нагрузки на платформе (s – боковая, r – сзади, f – впереди, t – вверху над колесной базой).

Здесь символ “*” определяет цифру. Согласно табл. 3 эксперты оценили каждую платформу и сформировали вектор номеров термов для каждой конфигурации, как представлено в табл. 4.

Таблица 4

Кодирование и оценка частных параметров конфигураций роботизированных платформ

Трехколесная с поворотными колесами			
w3 r3 p0 pr0 s	w3 r3 p0 pr0 r	w3 r3 p0 pr0 f	w3 r3 p0 pr0 t
M5 L2 C1 S3 P2	M5 L2 C1 S3 P1	M5 L1 C1 S3 P1	M5 L2 C1 S2 P2
...			
Шестиколесная с пассивными неповоротными колесами			
w6 r0 p4 pr0 s	w6 r0 p4 pr0 r	w6 r0 p4 pr0 f	w6 r0 p4 pr0 t
M1 L3 C3 S1 P4	M1 L3 C3 S1 P1	M1 L3 C3 S1 P1	M1 L3 C3 S1 P4

Результаты экспериментов. При использовании программного обеспечения были сформированы частные оценки для всех 48 конфигураций в виде нечетких интервалов, которые автоматически объединены в нечеткие переменные.

Интерфейс приложения для оценки функциональности роботизированных платформ представлен на рис. 3.

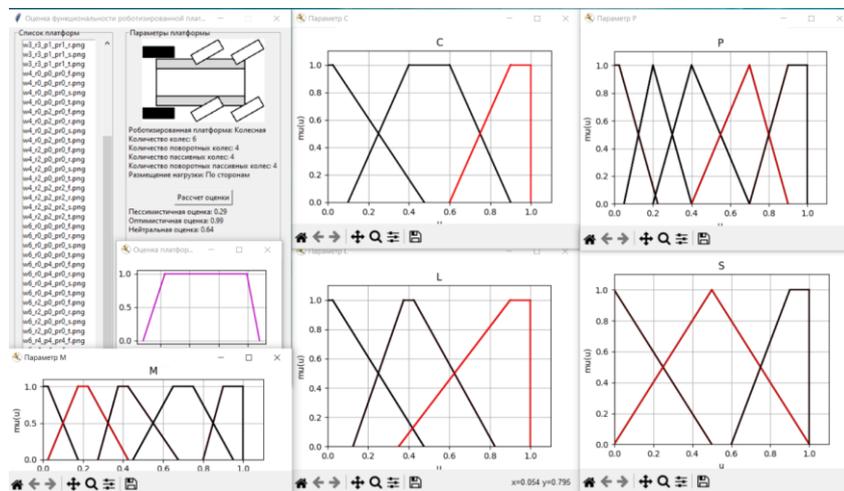


Рис. 3. Интерфейс приложения для оценки платформ

Пользователю доступен список возможных конфигураций платформ. При выборе конфигурации подгружается ее изображение и приводится развернутое описание, в соответствии с закодированной последовательностью. При нажатии кнопки «Расчет оценки» открываются графические формы с лингвистическими переменными параметров. Термы, соответствующие оценкам экспертов, выделяются цветом. Оценка платформы в виде нечеткого интервала выстраивается в графическом окне. Также выводятся значения песимистичной, оптимистичной и нейтральной оценок. Результаты моделирования представлены в табл. 5.

Таблица 5

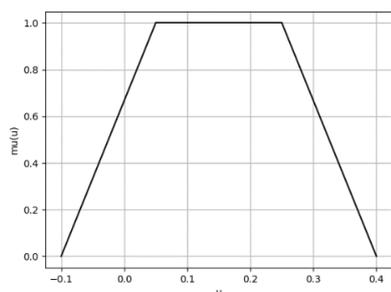
Результаты оценок конфигураций платформ

№	Конфигурация	Пессимистичная оценка	Оптимистичная оценка	Нейтральная оценка
...
5	w3 r2 p1 pr0 f	0,06625	0,37375	0,22
...
37	w6 r0 p4 pr0 f	0,0225	0,4725	0,2475
38	w6 r0 p4 pr0 r	0,0225	0,4725	0,2475
...
44	w6 r2 p0 pr0 t	0,495	1,35	0,9225
...

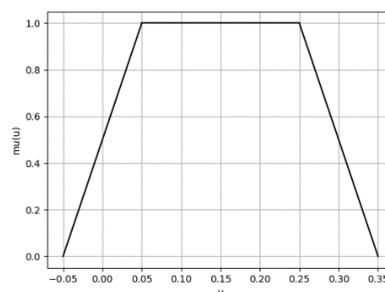
По результатам моделирования установлено, что при заданных частных оценках параметров:

- ◆ шестиколесная платформа с полным приводом и двумя поворотными колесами и расположением нагрузки над колесной базой является лучшим вариантом по совокупности характеристик;
- ◆ трехколесная платформа с передним расположением нагрузки и шестиколесная платформа с четырьмя неповоротными пассивными колесами является худшим вариантом по совокупности характеристик.

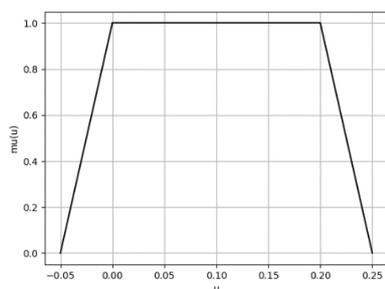
При этом, важно то, что весовые коэффициенты для расчета оценки по формуле (5) также могут задаваться в виде нечетких интервалов как представлено на рис. 5. Они определяются экспертами на основе своей уверенности о степени важности того или иного параметра, которые дают вклад в общую оценку конфигурации роботизированной платформы. Очевидно, ширина плато интервала характеризует величину степени уверенности эксперта относительно его среднего значения, а наклон боковых сторон определяет граничные значения его мнения относительно того или иного параметра. Для определения ширины интервала, можно при анкетировании эксперта разрешить использование модификаторов: около, сильно больше, немного больше, меньше и т.п., на основе которых, вычислять значение.



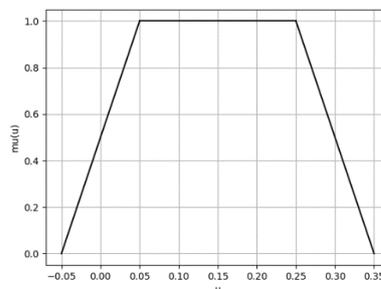
$a - k_1$



$b - k_2$



$v - k_3$



$z - k_4$

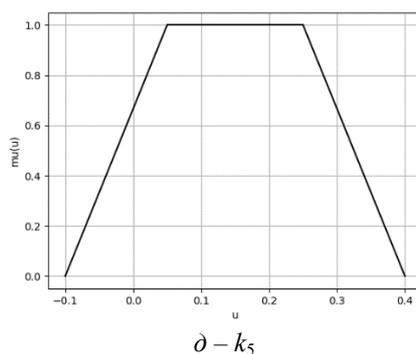


Рис. 5. Весовые коэффициенты критериев в виде нечетких интервалов

Следует отметить, что при такой системе оценки, требование $\sum k_i = 1$ не выдерживается, так как выбирается максимальная из возможных оценок для выбранной конфигурации платформ.

Заключение. В статье проведен системный анализ роботизированных платформ сельскохозяйственного назначения и сформировано дерево целей, позволяющее применить метод оценки оптимального конструктивного исполнения и метод оценки двигателей и движителей. Метод оценки конструктивного исполнения роботизированной платформы отличается простотой реализации, допускающей компьютерное моделирование и возможность учета экспертных знаний, в том числе заданных нечетко. Применение нечетких интервалов и математических операций над ними, позволяет учитывать трудноформализуемые показатели, например «сложность конструкции» или «функциональность в прополке сорняков» при вычислении комплексной оценки платформ с учетом двигателей и движителей. Применение аддитивной свертки критериев качества, допускающей использование весовых коэффициентов, заданных нечетко, позволяет применять метод оценки конструктивного исполнения платформ при небольшом количестве экспертов.

Предложенные методы могут использоваться для оценки конструктивного исполнения платформ в различных областях промышленности и сельского хозяйства за счет изменения или расширения показателей качества.

В дальнейшем авторы планируют рассмотреть эффективность свертки критериев с использованием нечеткой меры и нечетко-интегральной свертки [21].

Авторы благодарны руководству ООО «Алькема-Элитное» (г. Краснодар) за консультативную и экспертную поддержку при подготовке данной работы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Загазежева О.З., Бербекова М.М. Основные тренды развития роботизированных технологий в сельском хозяйстве // Известия КБНЦ РАН. – 2021. – № 5 (103). – С. 11-20.
2. Соловьев В.В., Шадрина В.В., Номерчук А.Я., Филатов Р.К. Перспективы развития сельскохозяйственной робототехники в условиях импортозамещения // Сб. трудов XIII Всероссийской Школы-семинара, молодых ученых, аспирантов, студентов и школьников. – Ростов-на-Дону – Таганрог: ЮФУ, 2022. – С. 27-34.
3. Santos Valle, S. and Kienzle. Agriculture 4.0 - Agricultural robotics and automated equipment for sustainable crop production // Integrated Crop Management. – Rome, 2020. – Vol. 24.
4. Измайлов А.Ю., Смирнов И.Г., Хорт Д.О. Цифровые агротехнологии в системе «Умный сад» // Садоводство и виноградарство. – 2018. – 6. – С. 33-39. – DOI: 10.31676/0235-2591-2018-6-33-39.
5. Oliveira L.F., Moreira A.P., & Silva M.S. Advances in Agriculture Robotics: A State-of-the-Art Review and Challenges Ahead // Robotics. – 2021. – 10. – 52 p.
6. Higuti Vitor & Velasquez Andres & agalhães Daniel & Becker arcelo & Chowdhary Girish. Under canopy light detection and ranging-based autonomous navigation // Journal of Field Robotics. – 2018. – No. 36. – 10.1002/rob.21852.
7. Adamides George. Agricultural Robots in Targeted Spraying: A mini State-of-the-Art review // Robotics & Automation Engineering Journal. – 2017. – No. 2. – 10.19080/RAEJ.2018.02.555581.

8. *Emmi Luis & Gonzalez-de-Soto Mariano & Pajares Gonzalo & Gonzalez-de-Santos Pablo*. New Trends in Robotics for Agriculture: Integration and Assessment of a Real Fleet of Robots // *The ScientificWorld Journal*. – 2014. – 404059. – 10.1155/2014/404059.
9. *Corpe Samuel & Tang Liqiong & Abplanalp Phillip*. GPS-guided modular design mobile robot platform for agricultural applications // *Proceedings of the International Conference on Sensing Technology, ICST, 2013*. – P. 806-810. – 10.1109/ICSensT.2013.6727763.
10. *Chebroly Nived & Lottes Philipp & Schaefer Alexander & Winterhalter Wera & Burgard Wolfram & Stachniss Cyrill*. Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields // *The International Journal of Robotics Research*. – 2017. – No. 36. – 10.1177/0278364917720510.
11. *Tabile Rubens & Godoy Eduardo & Pereira Robson & Tangerino Giovana & Porto Arthur & Inamasu Ricardo*. Design and development of the architecture of an agricultural mobile robot // *Engenharia Agricola*. – 2011. – No. 31. – P. 130-142. – 10.1590/S0100-69162011000100013.
12. *Bawden Owen, Ball David, Kulk Jason, Perez Tristan, & Russell Raymond*. A lightweight, modular robotic vehicle for the sustainable intensification of agriculture / In Chen, C (Ed.) // *Proceedings of the 16th Australasian Conference on Robotics and Automation 2014*. Australian Robotics and Automation Association (ARAA), Australia. – P. 1-9.
13. *Abhishesh P. et al*. Multipurpose Agricultural Robot Platform: Conceptual Design of Control System Software for Autonomous Driving and Agricultural Operations Using Programmable Logic Controller // *World Academy of Science, Engineering and Technology, International Journal of Agricultural and Biosystems Engineering*. – 2017. – 4.
14. *Pecka Aldis & Osadcuks Vitalijs*. Conceptual design of modular multi functional agricultural mobile robot. – 2018. – P. 202-206. – 10.22616/trd.24.2018.031.
15. *Nielsen S.H., Jensen K., Bøgild A., Jørgensen O.J., Jacobsen N.J., Jæger C.L.D., Jørgensen . N.* A Low Cost, Modular Robotics Tool Carrier For Precision Agriculture Research // *11th International Conference on Precision Agriculture International Society of Precision Agriculture*. – 2012.
16. *Черняк Ю.И.* Системный анализ в управлении экономикой. – М.: Экономика, 1975. – 193 с.
17. *Финаев В.И.* Модели систем принятия решений: учеб. пособие. – Таганрог: Изд-во ТРТУ, 2005. – 118 с.
18. *Штойер Р.* Многокритериальная оптимизация: пер. с англ. – М.: Радио и связь, 1992. – 504 с.
19. *Лотов А.В., Поспелова И.И.* Многокритериальные задачи принятия решений: учеб. пособие. – М.: МАКС Пресс, 2008. – 197 с.
20. *Дмитриев М.Г., Ломазов В.А.* Оценка чувствительности линейной свертки частных критериев при экспертном определении весовых коэффициентов // *Искусственный интеллект и принятие решений*. – 2014. – № 1. – С. 52-56. – ISSN: 2071-8594.
21. *Соловьев В.В.* Алгоритм оценки эффективности функционирования сложных технических систем // *Матер. Всероссийской научной конференции: Перспективы развития гуманитарных и технических систем: в 3 частях, Ч. 2*. – Таганрог: Изд-во ТТИ ЮФУ, 2011. – С. 61-62.

REFERENCES

1. *Zagazezheva O.Z., Berbekova M.M.* Osnovnye trendy razvitiya robotizirovannykh tekhnologiy v sel'skom khozyaystve [Main trends in the development of robotic technologies in agriculture], *Izvestiya KBNTS RAN* [News of the KBSC RAS], 2021, No. 5 (103) pp. 11-20.
2. *Solov'ev V.V., Shadrina V.V., Nomerchuk A.Ya., Filatov R.K.* Perspektivy razvitiya sel'skokhozyaystvennoy robototekhniki v usloviyakh importozameshcheniya [Prospects for the development of agricultural robotics in the conditions of import substitution], *Sb. trudov XIII Vserossiyskoy SHkoly-seminara, molodykh uchenykh, aspirantov, studentov i shkol'nikov* [Collection of works of the XIII All-Russian School-Seminar, young scientists, graduate students, students and schoolchildren]. Rostov-on-Don – Taganrog: YuFU, 2022, pp. 27-34.
3. *Santos Valle, S. and Kienzle*. Agriculture 4.0 - Agricultural robotics and automated equipment for sustainable crop production, *Integrated Crop Management*. Rome, 2020, Vol. 24.
4. *Izmaylov A.Yu., Smirnov I.G., Khort D.O.* Tsifrovye agrotekhnologii v sisteme «Umnyy sad» [Digital agricultural technologies in the “Smart Garden” system], *Sadovodstvo i vinogradarstvo* [Gardening and viticulture], 2018, 6, pp. 33-39. DOI: 10.31676/0235-2591-2018-6-33-39.
5. *Oliveira L.F., Moreira A.P., & Silva M.S.* Advances in Agriculture Robotics: A State-of-the-Art Review and Challenges Ahead, *Robotics*, 2021, 10, 52 p.
6. *Higuti Vitor & Velasquez Andres & agalhães Daniel & Becker arcelo & Chowdhary Girish*. Under canopy light detection and ranging-based autonomous navigation, *Journal of Field Robotics*, 2018, No. 36, 10.1002/rob.21852.
7. *Adamides George*. Agricultural Robots in Targeted Spraying: A mini State-of-the-Art review, *Robotics & Automation Engineering Journal*, 2017, No. 2. 10.19080/RAEJ.2018.02.555581.

8. Emmi Luis & Gonzalez-de-Soto Mariano & Pajares Gonzalo & Gonzalez-de-Santos Pablo. New Trends in Robotics for Agriculture: Integration and Assessment of a Real Fleet of Robots, *The ScientificWorld Journal*, 2014, 404059. 10.1155/2014/404059.
9. Corpe Samuel & Tang Liqiong & Abplanalp Phillip. GPS-guided modular design mobile robot platform for agricultural applications, *Proceedings of the International Conference on Sensing Technology, ICST, 2013*, pp. 806-810. 10.1109/ICSensT.2013.6727763.
10. Chebrolu Nived & Lottes Philipp & Schaefer Alexander & Winterhalter Wera & Burgard Wolfram & Stachniss Cyrill. Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields, *The International Journal of Robotics Research*, 2017, No. 36. 10.1177/0278364917720510.
11. Tabile Rubens & Godoy Eduardo & Pereira Robson & Tangerino Giovana & Porto Arthur & Inamasa Ricardo. Design and development of the architecture of an agricultural mobile robot, *Engenharia Agricola*, 2011, No. 31, pp. 130-142. 10.1590/S0100-69162011000100013.
12. Bawden Owen, Ball David, Kulk Jason, Perez Tristan, & Russell Raymond. A lightweight, modular robotic vehicle for the sustainable intensification of agriculture, In Chen, C (Ed.), *Proceedings of the 16th Australasian Conference on Robotics and Automation 2014. Australian Robotics and Automation Association (ARAA), Australia*, pp. 1-9.
13. Abhishesh P. et al. Multipurpose Agricultural Robot Platform: Conceptual Design of Control System Software for Autonomous Driving and Agricultural Operations Using Programmable Logic Controller, *World Academy of Science, Engineering and Technology, International Journal of Agricultural and Biosystems Engineering*, 2017, 4.
14. Pecka Aldis & Osadcuks Vitalijs. Conceptual design of modular multi functional agricultural mobile robot, 2018, pp. 202-206. 10.22616/rrd.24.2018.031.
15. Nielsen S.H., Jensen K., Bøgild A., Jørgensen O.J., Jacobsen N.J., Jæger C.L.D., Jørgensen . N. A Low Cost, Modular Robotics Tool Carrier For Precision Agriculture Research, *11th International Conference on Precision Agriculture International Society of Precision Agriculture*, 2012.
16. Chernyak Yu.I. Sistemnyy analiz v upravlenii ekonomikoy [System analysis in economic management], Moscow: Ekonomika, 1975, 193 p.
17. Finaev V.I. Modeli sistem prinyatiya resheniy: ucheb. posobie [Models of decision-making systems: textbook]. Taganrog: Izd-vo TRTU, 2005, 118 p.
18. Shtoyer R. Mnogokriterial'naya optimizatsiya [Multicriteria optimization]: trans. from engl. Moscow: Radio i svyaz', 1992, 504 p.
19. Lotov A.V., Pospelova I.I. Mnogokriterial'nye zadachi prinyatiya resheniy: ucheb. posobie [Multicriteria decision-making problems: textbook]. Moscow: MAKS Press, 2008, 197 p.
20. Dmitriev M.G., Lomazov V.A. Otsenka chuvstvitel'nosti lineynoy svertki chastnykh kriteriev pri ekspertnom opredelenii vesovykh koeffitsientov [Assessing the sensitivity of linear convolution of partial criteria in the expert determination of weighting coefficients], *Iskusstvennyy intellekt i prinyatie resheniy* [Artificial intelligence and decision making], 2014, No. 1, pp. 52-56. ISSN: 2071-8594.
21. Solov'ev V.V. Algoritm otsenki effektivnosti funktsionirovaniya slozhnykh tekhnicheskikh sistem [Algorithm for assessing the effectiveness of the functioning of complex technical systems], *Mater. Vserossiyskoy nauchnoy konferentsii: Perspektivy razvitiya gumanitarnykh i tekhnicheskikh sistem* [Materials of the All-Russian Scientific Conference: Prospects for the development of humanitarian and technical systems]: in 3 part. Part 2. Taganrog: Izd-vo TTI YuFU, 2011, pp. 61-62.

Статью рекомендовал к опубликованию д.т.н., профессор В.Х. Пшихопов.

Соловьев Виктор Владимирович – Южный федеральный университет; e-mail: vvsolovev@sfedu.ru; г. Таганрог, Россия; тел.: +79043438844; кафедра систем автоматического управления; старший преподаватель.

Номерчук Александр Яковлевич – e-mail: aynomerchuk@sfedu.ru; тел.: +79185945894; кафедра систем автоматического управления; старший преподаватель.

Филатов Роман Константинович – КБ «АРК» ДГТУ; e-mail: r.k.filatov@gmail.com; г. Таганрог, Россия; тел.: +79996965820; к.т.н.; руководитель.

Soloviev Viktor Vladimirovich – Southern Federal University; e-mail: vvsolovev@sfedu.ru; Taganrog, Russia; phone: +79043438844; the department of automatic control systems; senior lecturer.

Nomerchuk Alexandr Yakovlevich – e-mail: aynomerchuk@sfedu.ru; phone: +79185945894; the department of automatic control systems; senior lecturer.

Filatov Roman Konstantinovich – СВ ARK DSTU; e-mail: r.k.filatov@gmail.com; Taganrog, Russia; phone: +79996965820; cand. of eng. sc.; supervisor.

И.И. Князев**ИССЛЕДОВАНИЕ МЕТОДОВ ПОСТРОЕНИЯ КАУЗАЛЬНЫХ ГРАФОВЫХ МОДЕЛЕЙ ДЛЯ СЛОЖНЫХ СОЦИОГУМАНИТАРНЫХ СИСТЕМ**

Сложные социогуманитарные системы - это разновидность систем, которые изучаются в социологии, антропологии, экономике, политологии, и других гуманитарных науках. Эти системы характеризуются сложностью взаимодействий между их составляющими элементами, которые могут быть как людьми (индивидами, группами), так и культурными, социальными, экономическими и политическими аспектами. Например, общество как социогуманитарная система состоит из различных элементов, таких как люди, культура, институты, ценности и так далее. Они взаимодействуют между собой, образуя сложную сеть связей и влияний, которая определяет поведение и развитие общества. Чтобы лучше понять такие системы, используются различные подходы, включая системный анализ, социальную сетевую теорию, теорию сложности и другие методы. Эти подходы помогают выявить основные закономерности в функционировании сложных социогуманитарных систем и предсказать их развитие в будущем. В данной статье рассматриваются подходы к выявлению причинно-следственных связей, выделяются основные требования к построению этих связей в контексте сложных социогуманитарных систем, имеющих дело, в основном, со слабоструктурированной информацией, часто в виде естественного языка и текстов. Были определены слабые и сильные стороны выявленных подходов, а также рассмотрены примеры использования современных методов построения графов на разных задачах: выявление рисков в бизнесе, анализ социальных явлений, выявление наличия причинности в текстах. Исследование показало, что наиболее продуктивными являются методы машинного обучения, например языковые модели для извлечения знаний из текста в совокупности с нейросетевыми технологиями и графовым представлением знаний. Они требуют уверенных знаний математики, статистики и программирования, как минимум на языке Python, имеющих самую внушительную инструментальную поддержку для решения задач машинного обучения. Также, выявление причинности основывается не только на корреляции, но и на других методах, таких как тест Грейнджера, используемый для анализа временных рядов.

Каузальные графы; построение причинно-следственных связей; представление знаний; машинное обучение; социогуманитарные системы; выявление причинности.

I.I. Knyazev**RESEARCH OF METHODS FOR CONSTRUCTING CAUSAL GRAPH MODELS FOR COMPLEX SOCIO-HUMANITARIAN SYSTEMS**

Complex socio-humanitarian systems are a type of systems that are studied in sociology, anthropology, economics, political science, and other humanities. These systems are characterized by the complexity of interactions between their constituent elements, which can be both human (individuals, groups) and cultural, social, economic and political aspects. For example, society as a socio-humanitarian system consists of various elements, such as people, culture, institutions, values, and so on. They interact with each other, forming a complex network of connections and influences that determines the behavior and development of society. To better understand such systems, various approaches are used, including systems analysis, social network theory, complexity theory, and other methods. These approaches help to identify the main patterns in the functioning of complex socio-humanitarian systems and predict their development in the future. This article discusses approaches to identifying cause-and-effect relationships and highlights the basic requirements for building these connections in the context of complex socio-humanitarian systems that deal mainly with semi-structured information, often in the form of natural language and texts. The strengths and weaknesses of the identified approaches were identified, and examples of the use of modern methods of constructing graphs on various tasks were considered: identifying risks in business, analyzing social phenomena, identifying the presence of causality in texts. The study showed that the most productive methods are machine learning, for example, language models for extracting knowledge from text in combination with neural network technologies and graph representations of knowledge. They require solid knowledge of mathematics, statistics and programming, at least in Python, with the most impressive tool support for solving machine learning problems. Also, identifying causality is based not only on correlation but also on other methods such as the Granger test used for time series analysis.

Causal graphs; building cause-and-effect relationships; knowledge representation; machine learning; socio-humanitarian systems; identification of causality.

Введение. Каузальная графовая модель - это математическая модель, которая используется для представления причинно-следственных отношений между различными переменными в системе. В отличие от простых статистических моделей, каузальные графовые модели позволяют не только описывать статистические связи между переменными, но и делать выводы о причинно-следственных взаимосвязях.

В основе каузальных графовых моделей лежит концепция направленных графов, где узлы представляют переменные, а направленные ребра указывают направление влияния между переменными. Такие модели помогают исследователям выявить не только статистические связи между данными, но и понять, как одна переменная влияет на другую.

Важным преимуществом каузальных графовых моделей является возможность проведения анализа причинно-следственных отношений на основе данных, что помогает исследователям делать более обоснованные выводы о влиянии переменных друг на друга в сложных системах. Эти модели активно применяются в различных областях, таких как медицина, экономика, социология и машинное обучение. Они помогают выявить факторы, влияющие на определенные явления, и улучшить принятие решений на основе понимания причинно-следственных связей.

Особенности каузальных графовых моделей для социогуманитарных систем.

При исследовании были выделены следующие особенности таких моделей:

1. Многоуровневость переменных: в социогуманитарных системах часто есть многоуровневая природа переменных, где факторы на разных уровнях взаимосвязаны. Каузальные графовые модели позволяют учитывать эту сложность и взаимосвязи между разными уровнями.

2. Обратные связи и циклы: часто встречаются обратные связи и циклические взаимодействия между переменными. При построении каузальных графов нужно учитывать возможность таких взаимосвязей, чтобы корректно моделировать причинно-следственные отношения.

3. Неопределенность и контекст: социогуманитарные системы могут быть подвержены различным воздействиям и иметь неопределенные результаты. При построении каузальных графов нужно учитывать контекст и возможную неопределенность, чтобы точно интерпретировать результаты моделирования.

4. Интерпретируемость и вовлеченность экспертов: важно, чтобы каузальные графовые модели были интерпретируемыми для специалистов в социогуманитарных областях. Построение моделей с участием экспертов из социогуманитарных областей поможет создать более достоверные и полезные модели.

В соответствии с потребностями социогуманитарных систем были выделены следующие аспекты, необходимые для построения графовых моделей:

1. Тщательный анализ данных: выявить ключевые переменные и их взаимосвязи перед построением модели.

2. Экспертное мнение и контекст: участие экспертов из социогуманитарных областей поможет лучше понять специфику системы и правильно интерпретировать результаты моделирования.

3. Учет неопределенности и динамичности: необходимо учитывать неопределенность и динамичность социогуманитарных систем при построении моделей, чтобы они были более адаптивными к изменениям в системе.

4. Тестирование и валидация.

Для сбора данных, используемых потом для построения каузальных графовых моделей используются методы выявления причинности.

Методы выявления причинности — это инструменты и подходы, которые позволяют исследователям определять причинно-следственные взаимосвязи между переменными в данных и системах. Эти методы помогают не только выявить связи между переменными, но и сделать выводы о том, как одна переменная влияет на другую.

Основные методы выявления причинности. Когда мы говорим о построении каузальных графовых моделей, методы выявления причинности становятся ключевым этапом в процессе. В контексте построения каузальных графовых моделей методы выявления при-

чинности помогают исследователям определить направление влияния между переменными и построить корректные и информативные модели. На основе результатов изучения причинности строятся направленные графы, отражающие причинно-следственные отношения между переменными.

Проанализировав литературу из открытых ресурсов ScienceDirect, были выделены основные методы, применяемый для выявления причинности, сравнение приведено в табл. 1.

Таблица 1

Основные методы выявления причинности

Метод	Ключевые особенности	Недостатки	Преимущества
Экспериментальные исследования	Проведение контролируемых тестов с измерением показателей переменных	- серия испытаний в разных условиях более информативна, чем серия испытаний в одинаковых условиях. Повторение в теоретически эквивалентных, но конкретно различных обстоятельствах обеспечивает лучший тест, чем даже самое тщательно спланированное одиночное исследование - ограниченные возможности - несамостоятельный метод	В рамках узкой задачи гуманитарной направленности метод может быть достаточно эффективен [1]
Наблюдательные исследования	Происходит наблюдение за поведением системы и объектов без вмешательства в нее. На основе полученных наблюдений производится анализ и выявление закономерностей	- необходимо использовать дополнительные статистические методы для анализа и расчёта ключевых показателей и их связей	Эффективен в рамках поставленной задачи при наличии большого количества исходных данных для анализа
Методы машинного обучения	Использование алгоритмов машинного обучения для выявления связей	- результат сильно зависит от качества данных - поиск наборов данных для обучения - требуются знания ЯП (Python, R)	- Возможность не только выявить взаимосвязи в рамках гипотезы, но и выявить аномалии и скрытые связи - обработка огромных массивов данных - прогнозирование эффектов

Экспериментальные исследования. В экспериментах проводятся контролируемые тесты, где изменяются определенные переменные, чтобы определить их влияние на другие переменные. Эксперименты позволяют устанавливать причинно-следственные связи. Работа [1] посвящена исследованию использования экспериментов в социальных науках, представлен обзор существующего состояния и предложения для будущих исследований.

В работе авторами был зафиксирован впечатляющий рост использования экспериментов в социальных науках. Как и любой инструмент, эксперименты для одних целей лучше, чем для других, и имеют соответствующие сильные и слабые стороны. Разные области и субдисциплины, которые ограничиваются использованием одного метода, ограничивают типы исследовательских вопросов, которые они могут решать, а также связывают свои знания с ограничениями метода. Авторы считают, что методологический плюрализм социальных наук является одной из их сильных сторон.

Наблюдательные исследования. В наблюдательных исследованиях изучаются данные из натуральной среды без вмешательства исследователя. На основе анализа данных пытаются найти возможные причинно-следственные отношения. В работе [2] задачей было оценить, в какой степени индивидуальный ресурс (медицинская грамотность) и социальные ресурсы (поддержка друзей и семьи) смягчают связь между рядом индивидуальных характеристик (пол, возраст, семейный достаток и депрессивные чувства) и проблему использования социальных сетей (ПИСС); также связь между ПИСС и показателями здоровья (самооценка здоровья, удовлетворенность жизнью и трудности со сном), как на межнациональном, так и на национальном уровне.

Исследование проведено на 22 226 подростках из шести европейских стран с использованием данных опроса "Поведение детей школьного возраста в отношении здоровья" за 2017/2018 год. Использовались модели случайных эффектов и анализ модераторов (используется для проверки гипотезы о различиях во вкладе независимой переменной в зависимую при разных уровнях переменной-модератора).

Методы машинного обучения. Машинное обучение позволяет строить модели, которые могут выявлять причинно-следственные связи на основе больших объемов данных и сложных паттернов.

1. Регрессионный анализ: данный метод используется для определения статистической связи между одной зависимой переменной и одной или более независимых переменных. Это позволяет моделировать и предсказывать значения зависимой переменной на основе независимых факторов.

2. Кластерный анализ: данный метод позволяет группировать события на основе их сходства и выявлять паттерны и взаимосвязи внутри каждой группы. Он может быть полезен, когда в данных отсутствуют явно заданные зависимые переменные.

3. Методы временных рядов: эти методы используются для анализа последовательных данных во времени, таких как данные о временных рядах событий. Примерами методов временных рядов являются ARIMA [3,4] (авторегрессионное интегрированное скользящее среднее), SARIMA (сезонная модель ARIMA) [5, 6] и LSTM [7–9] (долгая краткосрочная память).

Отдельно стоит отметить тест причинно-следственной связи по Грейнджеру [10, 11], который используется в эконометрике для выявления каузальных связей во временных рядах. Реализация этого метода доступна в MATLAB, Python и других современных инструментах. Корреляции не всегда достаточно для подтверждения причинно-следственных связей, потому используют в дополнение этот тест, однако, причинность по Грейнджеру является необходимым, но не достаточным условием причинно-следственной связи.

Работа [12] демонстрирует применение данного подхода к выявлению взаимосвязей. Целью исследования было выявление эффектов, связанных с самоубийствами военнослужащих, сосредоточившись на конкретном историческом случае самоубийства кронпринца Рудольфа в 1889 году.

Для достижения цели исследования были проведены анализ временных рядов самоубийств среди военнослужащих в период с 1873 по 1910 год. Подход авторов основан на методах прерванного анализа временных рядов, позволяющих выявить изменения в тенденциях после конкретного события, такого как самоубийство кронпринца.

4. Байесовская сеть: данная модель представляет события в виде вершин графа, где ребра между вершинами указывают на вероятностные зависимости между событиями. Байесовские сети могут использоваться для моделирования взаимосвязей и предсказания вероятности возникновения событий.

5. Ассоциативные правила: этот метод используется для поиска скрытых взаимосвязей между событиями на основе частоты их совместной встречаемости. Например, анализ ассоциативных правил может помочь выявить, какие события часто происходят вместе и как одно событие может быть следствием другого.

Конечный выбор метода зависит от типа данных, доступной информации и цели анализа взаимосвязей в событиях. Важно учитывать, что эти методы машинного обучения могут использоваться в комбинации или совмещаться с другими методами для более точного анализа.

Таким образом, методы выявления причинности являются важным инструментом при построении каузальных графовых моделей, позволяя более точно представить причинно-следственные связи в данных и системах.

Применение каузальных графов в исследованиях. В работе [13] выявление причинно-следственной связи между событиями является важнейшей исследовательской задачей в области обработки естественного языка. Однако, существующие методы либо игнорируют фоновые знания о событиях, либо не учитывают влияние шума графа знаний на представления событий. В этой статье предложена новая сеть расширенных знаний на основе контраста графов (GCKAN) для задачи идентификации причинно-следственных связей событий, которая объединяет комплексные фоновые знания о событиях из графов знаний и устраняет проблему шума графа знаний. Во-первых, предлагается модуль расширения описательных знаний для агрегирования информации о соседях с одним переходом в графах знаний и получения значимых описательных знаний о событиях. Затем модуль расширения реляционных знаний предназначен для кодирования информации о многопрыжковом пути и изучения скрытых знаний о рассуждениях между парами событий. Кроме того, в двух модулях разработаны схемы контрастного обучения на основе достоверности и степени, которые подавляют шум графа знаний во время агрегации информации и создают более надежные представления событий с учетом знаний. Обширные эксперименты с тремя общедоступными наборами данных демонстрируют последовательное превосходство GCKAN над современными методами, основанными на знаниях. Эксперименты с шумовыми помехами и эксперименты по межтематической адаптации дополнительно подтверждают надежность и обобщение GCKAN.

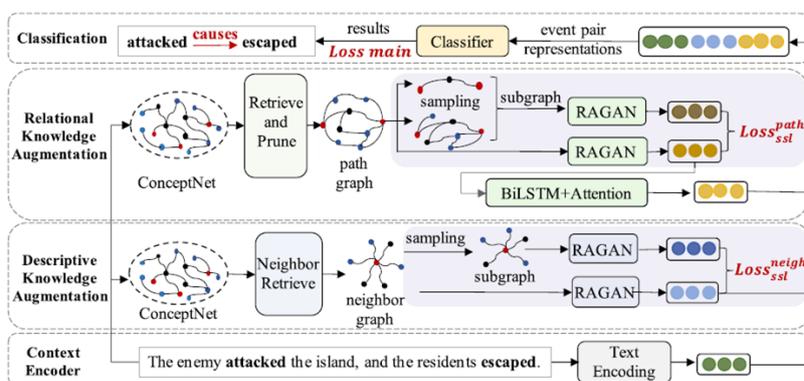


Рис. 1. Архитектура модели преобразования знаний

На рис. 1 представлена архитектура от авторов которая состоит из модуля кодирования контекста, модуля описательного расширения знаний, модуля реляционного расширения знаний и модуля классификации. Фиолетовые части обозначают контрастное обучение графа.

Выделим инструменты и методы, задействованные в этой модели и рассмотрим подробнее.

ConceptNet – открытая и бесплатная база знаний, структурно реализованная в виде семантической сети, включающей знания о мире для понимания слов. Предыдущие версии ConceptNet – краудсорсинговый проект, в рамках которого создатели управляли веб-сайтом (Open Mind Common Sense), собирая факты от людей, которые заходили на этот сайт [14].

Данные поступают из множества различных источников, содержит реляционные знания, предоставленные Open Mind Common Sense и родственными ему проектами на других языках. В источники входит Википедия, Викисловарь, Open Multilingual WordNet [15, 16], онтология высокого уровня из OpenCyc, ряд знаний об интуитивных словесных ассоциациях людей извлекаются при помощи «целенаправленных игр». Например, факты на английском языке из Verbosity, словесной игры, ранее запущенной проектом GWAP, а на японском языке – из игры «nadya.jp» от Nihon Unisys и Dentsu.

Преобразование текста делается с помощью модели BERT [17].

Расширение описательных знаний, которое использует сеть внимания графов с учетом отношений (RAGAN) для агрегирования описательных знаний о событиях и разрабатывает основанную на надежности схему контрастного обучения графов, чтобы сделать представление знаний менее чувствительным к связям с шумным соседом;

BiLSTM [18] для расширения реляционных знаний, которое сначала организует пути между событиями в графы, а затем изучает многошаговые реляционные знания с помощью сетей внимания графов, учитывающих отношения.

Далее все сводится к задаче бинарной классификации, определяя, есть ли причинно-следственная связь между событиями.

На рис. 2 представлен пример описания ситуации «Враг атаковал остров, и жители бежали». Граф отражает связи с близкими по значению словами, на ребрах указан вид связей, на вершинах слова.

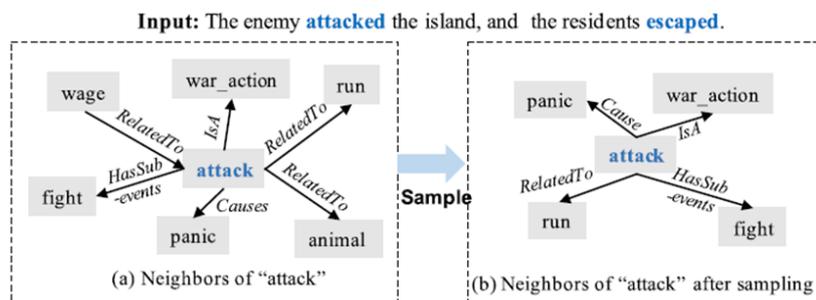


Рис. 2. Фрагмент визуализации каузального графа

В работе [19] применяется графовая модель выявления неявных знаний в области выявления рисков бизнеса. Целью данного исследования было проанализировать цепочки рисков и извлечь неявную информацию из раскрытых документов. Во-первых, была предложена классификация ребер и предложены подходящие метки для ребер графа цепочки рисков. Кроме того, разработана классификация типов ребер в гетерогенных графах с использованием нейронных сетей графов (GNN[20]). Это было достигнуто путем определения шести рисков и построения графиков цепочек рисков. Результаты показали, что классификация по типу ребер оказалась эффективным подходом по сравнению с существующим методом. Этот метод потенциально может помочь инвесторам увеличить свою прибыль и принять более обоснованные решения.

Заключение. В данной работе было проведено исследование каузальных графов в задачах социогуманитарных систем. Были выделены основные достоинства и недостатки применения этого подхода в рамках социогуманитарных аспектов, а также требования к их построению с учетом особенностей предметной области. Среди них были выделены и описаны следующие: многоуровневость переменных, обратные связи и циклы, неопределенность и контекст, интерпретируемость и вовлеченность экспертов.

Исследуя вышеописанный подход, были определены основные методы выявления причинности, в таблице указаны их сильные и слабые стороны, выделенные из современных исследований. Как показало исследование, на данном этапе развития мощнейшим инструментом являются методы машинного обучения, но они требуют качественной подготовки и знаний в области математики, статистики и программирования.

Помимо теоретических выводов, было продемонстрировано применение каузальных графов на примере анализа текстов, в основе которых, зачастую лежит именно рассматриваемый в этой работе подход.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Mize T.D., Manago B. The past, present, and future of experimental methods in the social sciences, *Social Science Research*, 2022, Vol. 108, 102799. ISSN 0049-089X.
2. Lahti H., Kulmala M., Hietajärvi L., Lyyra N., Kleszczewska D., Boniel-Nissim M., Furstova J., van den Eijnden R., Sudeck G., Paakkari L. What Counteracts Problematic Social Media Use in Adolescence? A Cross-National Observational Study, *Journal of Adolescent Health*, 2024, Vol. 74, Issue 1, pp. 98-112. ISSN 1054-139X.
3. Asteriou Dimitros, Hall Stephen G. ARIMA Models and the Box–Jenkins Methodology, *Applied Econometrics*. 2nd ed. Palgrave MacMillan, 2011, pp. 265-286. ISBN 978-0-230-27182-1.
4. Terence C. Mills. Time Series Techniques for Economists. Cambridge University Press, 1990. ISBN 978-0-521-34339-8.
5. Percival Donald B., Walden Andrew T. Spectral Analysis for Physical Applications. Cambridge University Press. 1993. ISBN 978-0-521-35532-2.
6. Shumway R.H. and Stoffer D.S. Time Series Analysis and Its Applications: With R Examples. Springer, 2017. DOI: 10.1007/978-3-319-52452-8.
7. Gers F., Schraudolph N., Schmidhuber J. Learning precise timing with LSTM recurrent networks (engl.), *Journal of Machine Learning Research: journal*, 2002, Vol. 3, pp. 115-143.
8. Greff K., Srivastava K., Koutnik Jan, Steunebrink Bas ., Schmidhuber Jürgen. LSTM: A Search Space Odyssey, *Neural and Evolutionary Computing*, 2015
9. Gers F.A., Schmidhuber J. LSTM Recurrent Networks Learn Simple Context Free and Context Sensitive Languages (engl.), *IEEE Transactions on Neural Networks (engl.)* russ.: journal, 2001, Vol. 12, No. 6, pp. 1333-1340. DOI: 10.1109/72.963769.
10. Kaminski M., Ding M., Truccolo W.A., Bressler S.L Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance, *Biol Cybern*, 2001, 85, pp. 145-57.
11. Ancona N., Marinazzo D., Stramaglia S. Radial basis function approaches to nonlinear Granger causality of time series, *Physical Review E*, 2004, Vol. 70, 056221.
12. Arendt F., Mestas M. Suicide among soldiers and social contagion effects: An interrupted time-series analysis, *Social Science & Medicine*, 2023, Vol. 320, 115747. ISSN 0277-9536.
13. Ding L., Chen J., Du P., Xiang Ya. Event causality identification via graph contrast-based knowledge augmented networks, *Information Sciences*, 2024, Vol. 656, 119905.
14. ConceptNet. Режим доступа: Available at: <https://conceptnet.io/> (accessed 28 March 2024).
15. Bond F., Paik K. A survey of wordnets and their licenses, *In Proceedings of the 6th Global WordNet Conference (GWC 2012)*. Matsue, 2012, pp. 64-71.
16. Bond F., Vossen P., McCrae John. CILI: the Collaborative Interlingual Index./ F. Bond, P. Vossen, John McCrae, Christiane Fellbaum, *In Proceedings of the 8th Global WordNet Conference (GWC2016)*. Bucharest, 2016, pp 50-57.
17. Pre-training of deep bidirectional transformers for language understanding, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, (Vol. 1: Long and Short Papers), 2019, pp. 4171-4186.
18. Cheng F., Miyao Y. Classifying temporal relations by bidirectional lstm over dependency paths, *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, 2017, Vol. 2: Short Papers, pp. 1-6.
19. Hajime Sasaki, Motomasa Fujii, Hiroki Sakaji, Shigeru Masuyama. Enhancing Risk Identification with GNN: Edge Classification in Risk Causality from Securities Reports, *International Journal of Information Management Data Insights*, 2024, Vol. 4. Issue 1, 100217. ISSN 2667-0968.
20. Wu Lingfei, Cui Peng, Pei Jian; Zhao Liang. Graph Neural Networks: Foundations, Frontiers, and Applications. Springer Singapore, 2022, 725 p.

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

Князев Иван Игоревич – Южный федеральный университет; e-mail: ikniazev@sfedu.ru; г. Таганрог, Россия; тел.: +79265347104; аспирант.

Knyazev Ivan Igorevich – Southern Federal University; e-mail: ikniazev@sfedu.ru; Taganrog, Russia; phone: +79265347104; postgraduate student.

Д.Г. Веселова, Н.Е. Сергеев

ПРЕДИКТИВНАЯ АНАЛИТИКА ДЛЯ ПРИНЯТИЯ РЕШЕНИЙ В ДЕЦЕНТРАЛИЗОВАННЫХ СИСТЕМАХ

В настоящее время актуальность использования криптоактивов стремительно растет. В последние годы торговля криптовалютой стала одной из наиболее обсуждаемых тем в мире финансов и инвестиций. Криптовалюты, такие как Биткоин, Эфириум, привлекают внимание миллионов людей благодаря своей инновационности, потенциалу высокой прибыли и возможности децентрализации. Технология блокчейн, на которой основаны криптовалюты, является одной из самых инновационных и перспективных технологий на рынке. Изучение торговли криптовалютами позволяет понять, как частные инвесторы и компании могут использовать блокчейн-технологии для инвестирования и развития своих бизнесов. Одной из главных причин популярности торговли криптовалютой является ее высокий уровень волатильности. Курс криптовалют может быстро меняться, что предоставляет возможности для получения прибыли. Данная статья направлена на исследование использования предиктивной аналитики для принятия решений в децентрализованных системах на примере торговли криптовалютами на централизованных и децентрализованных биржах. В настоящей работе проводится исследование децентрализованных и централизованных систем с целью дальнейшего создания систем поддержки принятия решений. Приведено общее описание и схемы работы децентрализованных и централизованных динамических систем на примере исследования криптобирж. Данная научная статья исследует типовую структуру централизованной и децентрализованной криптобиржи, анализируя основные компоненты и принципы ее функционирования. В статье рассматривается внутренняя организация биржи, включая систему хранения цифровых активов, механизмы выполнения сделок, обеспечение безопасности и управление рисками. Также обсуждается взаимодействие биржи с участниками рынка, а также регулируемыми органами. Помимо этого, данная научная статья исследует правила и принципы работы трейдеров и маркет мейкеров на централизованных и децентрализованных криптобиржах. Рассматриваются основные стратегии и тактики, которые используют участники рынка для обеспечения ликвидности и оптимизации операций торговли. В статье проводится сравнение подходов к торговле на различных типах криптобирж с учетом их особенностей и влияния на ценовую динамику криптовалют. Представленные результаты могут способствовать более глубокому пониманию процессов в торговле криптовалютами и оптимизации стратегий принятия решений для инвесторов и торговцев на рынке криптоактивов.

Децентрализация; динамические системы; криптовалюта; алгоритмическая торговля; предиктивная аналитика; блокчейн.

D.G. Veselova, N.E. Sergeyev

PREDICTIVE ANALYTICS FOR DECISION-MAKING IN DECENTRALIZED SYSTEMS

Currently, the relevance of using crypto assets is growing rapidly. In recent years, cryptocurrency trading has become one of the most discussed topics in the world of finance and investment. Cryptocurrencies such as Bitcoin, Ethereum, attract the attention of millions of people due to their innovativeness, high profit potential, and decentralization possibilities. The blockchain technology, on which cryptocurrencies are based, is one of the most innovative and promising technologies in the market. Studying cryptocurrency trading helps understand how private investors and companies can use blockchain technologies for investment and business development. One of the main reasons for the popularity of cryptocurrency trading is its high level of volatility. The cryptocurrency exchange rate can change quickly, providing opportunities for profit. This article focuses on exploring the use of predictive analytics for decision-making in decentralized systems using cryptocurrency trading on centralized and decentralized exchanges as an example. The research conducted in this work aims to investigate decentralized and centralized systems to further develop decision support systems. A general description and operation schemes of decentralized and centralized dynamic systems are provided using cryptocurrency exchanges as a research example. This scientific article examines the typical structure of centralized and decentralized cryptocurrency exchanges, analyzing the fundamental components and principles of their functioning. The article discusses the internal organization of the exchange, including the system for storing digital

assets, transaction execution mechanisms, security provisions, and risk management. It also examines the interaction between the exchange and market participants, as well as regulatory bodies. Furthermore, this scientific article explores the rules and principles of operation for traders and market makers on centralized and decentralized cryptocurrency exchanges. It covers the main strategies and tactics used by market participants to ensure liquidity and optimize trading operations. The article compares the trading approaches on different types of cryptocurrency exchanges considering their specific features and impact on cryptocurrency price dynamics. The presented results can contribute to a deeper understanding of cryptocurrency trading processes and optimize decision-making strategies for investors and traders in the crypto asset market.

Decentralization; dynamic systems; cryptocurrency; algorithmic trading; predictive analytics; blockchain.

Введение. Исследование децентрализованных и централизованных систем является актуальным по многим причинам. Во-первых, децентрализация предоставляет уникальные возможности для создания более устойчивых, прозрачных и безопасных систем.

Во-вторых, децентрализация имеет большое значение для многих отраслей и областей деятельности. Например, в сфере финансов и банковского дела децентрализованные системы позволяют проводить безопасные и прозрачные финансовые операции, минуя посредников и устраняя необходимость в доверии к институтам.

В-третьих, централизованные системы по-прежнему широко используются и важны для функционирования многих организаций и государств. Исследование централизованных систем помогает понять их преимущества и ограничения, а также идентифицировать возможности улучшения эффективности и безопасности.

Определение 1. Предиктивная аналитика — это процесс анализа данных, который использует различные методы и модели для прогнозирования будущих событий или результатов. Данный вид аналитики позволяет предсказывать вероятные будущие сценарии на основе данных о прошлых тенденциях и поведении [1].

Предиктивная аналитика является важным инструментом для трейдеров и инвесторов на криптобиржах, поскольку позволяет прогнозировать будущие тенденции и изменения цен на криптовалюты, используя следующие факторы:

1. Волатильность рынка.
2. Большое количество данных.
3. Повышение эффективности торговли.
4. Быстрота реакции на изменения рынка.
5. Уменьшение ошибок и повышение надежности.

В целом, предиктивная аналитика является ценным инструментом для трейдеров на криптобиржах, позволяющим прогнозировать тенденции и изменения на рынке криптовалют. Она помогает принимать осознанные решения, повышает эффективность торговли, снижает риски и повышает надежность в прогнозировании [2].

Далее приведем самые популярные и известные подходы к анализу и аналитике рынков [3, 4]:

1. Линейная регрессия: Данный алгоритм часто применяется для прогнозирования цен на криптовалюты. Он позволяет анализировать исторические данные о ценах для предсказания будущих изменений цен.

2. ARIMA (Авторегрессионное Интегрированное Скользящее Среднее): Модели ARIMA широко используются для анализа временных рядов при прогнозировании криптовалют. Они учитывают как прошлые значения, так и скользящие средние для предсказания будущих цен.

3. LSTM (Долгая Краткосрочная Память): LSTM – это тип рекуррентной нейронной сети (RNN), способной предсказывать будущие цены на основе исторических данных. Она часто используется для прогнозирования временных рядов в торговле криптовалютами.

4. Случайный лес: Алгоритм случайного леса использует ансамбль деревьев решений для анализа различных предикторов и прогнозирования. Он может применяться для предсказания цен и тенденций на рынке криптовалют.

5. Градиентный бустинг (GBM): GBM – это техника машинного обучения, которая комбинирует несколько слабых моделей прогнозирования для создания более сильной общей модели прогнозирования. Она часто применяется для предсказания цен и паттернов на рынке криптовалют.

Для сравнительного анализа приведем список продуктов, которые используются на данный момент для предсказания и аналитики поведения рынка [4, 5]:

1. TradingView: Эта платформа предоставляет инструменты для расширенного построения графиков, технического анализа и получения данных в режиме реального времени для различных криптовалют.

2. Coinigy: Coinigy предлагает комплексный набор инструментов для технического анализа, построения графиков и управления портфелем на нескольких биржах криптовалют.

3. CryptoCompare: CryptoCompare предлагает широкий спектр продуктов, включая данные рынка в режиме реального времени, исторические данные, агрегацию новостей и различные инструменты аналитики для криптовалютной индустрии.

4. CoinMetrics: CoinMetrics предоставляет всесторонний анализ данных on-chain и off-chain для криптовалют. Она предлагает исследования тенденций рынка, объемы транзакций, активность сети и многое другое.

Типовая структура централизованной криптобиржи. На схеме ниже (рис. 1) представлена типовая структура работы системы централизованной криптобиржи.

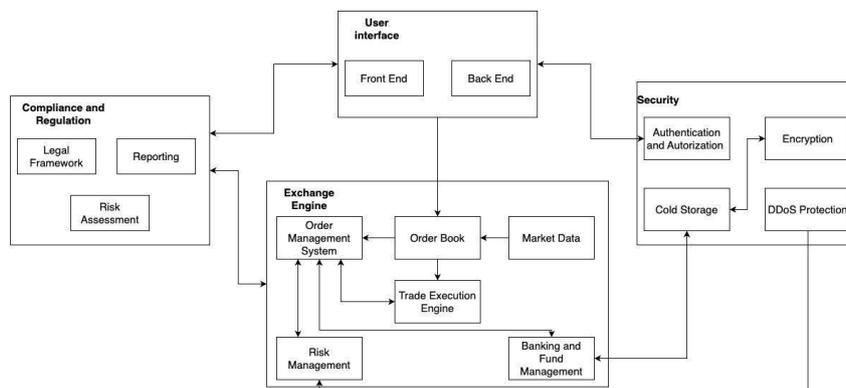


Рис. 1. Схема типовой структуры централизованной криптобиржи

Определение 2. Ордер – это заявка от трейдера на покупку или продажу определенного количества криптовалюты по определенной цене. Ордер можно разделить на два типа: ордер на покупку (buy order) и ордер на продажу (sell order). Когда ордер исполняется (то есть, когда его цена соответствует текущей цене на бирже), совершается сделка между покупателем и продавцом.

Разберем на примере структуру на примере типовой централизованной криптобиржи Binance [5, 6]. В состав инфраструктуры входят следующие компоненты:

1. Пользовательский интерфейс (User Interface):
 - a. Интерфейс (Frontend).
 - b. Бэкэнд (Backend): данный компонент является связующим звеном между фронтендом и ядром биржи.
2. Движок биржи (Exchange Engine):
 - a. Система управления ордерами (Order Management System): данная часть системы отвечает за жизненный цикл ордеров, на вход система получает заявки на покупку и продажу крипто активов, сопоставляет их между собой, следуя правилам матчинга ордеров на рынке и исполняет ордера.
 - b. Книга ордеров (Order Book): данная часть системы отвечает за трекинг всех активных заявок на покупку и продажу крипто активов, их цен, эмаунтов и времени создания и жизни.

- c. Маркет дата (Market Data): предоставляет реалтайм информацию о ценах различных крипто активов, торговых объемов, глубины стакана. данный компонент используется для иллюстрации графиков движения рынка, работы книги ордеров и системы менеджмента ордеров.
 - d. Движок исполнения ордеров (Trade Execution Engine): этот компонент отвечает за изменение балансов пользователей, исходя из исполненных ими торговых заявок.
 - e. Управление рисками (Risk Management) внедряет меры по контролю рисков и отслеживает торговые операции, чтобы предотвратить незаконное или манипулятивное поведение.
 - f. Управление средствами (Banking and Fund Management) обрабатывает процессы депозита и вывода средств, управляет балансами пользователей и обеспечивает безопасность средств, хранимых на бирже.
3. Система безопасности (Security):
- a. Аутентификация и авторизация (Authentication and Authorization): реализует безопасные механизмы входа, двухфакторную аутентификацию и управление доступом на основе ролей для защиты учетных записей пользователей.
 - b. Шифрование (Encryption): шифрует чувствительные данные, такие как учетные данные пользователей, закрытые ключи и финансовые транзакции, для обеспечения конфиденциальности.
 - c. Холодное хранение (Cold Storage): хранит значительную часть средств в автономных "холодных" кошельках, которые меньше подвержены возможным нарушениям безопасности.
 - d. Защита от DDoS-атак (DDoS Protection).
4. Система регулирования (Compliance and Regulation):
- a. Правовая система (Legal Framework): Соблюдение соответствующих законов и регуляций, связанных с криптовалютами, финансовыми услугами и защитой потребителей.
 - b. Отчетность (Reporting): Предоставляет необходимые отчеты регулирующим органам в соответствии с законом.
 - c. Оценка рисков (Risk Assessment): Периодически проводит оценку рисков и проверки, чтобы выявить и устранить проблемы, связанные с соблюдением и безопасностью.

В централизованной криптовалютной бирже обычно доступны следующие типы ордеров [7, 8]:

1. Ордер рыночной цены (Market order). Ордер на покупку или продажу определенного количества криптовалюты по лучшей доступной рыночной цене.
2. Ордер лимитированной цены (Limit order). Ордер на покупку или продажу определенного количества криптовалюты по указанной цене (лимит). Сделка будет выполнена только в случае достижения рынком указанной цены или лучше.
3. Ордер стоп-лимит (Stop order) также известный как стоп-лосс ордер, это ордер на покупку или продажу определенного количества криптовалюты, когда рыночная цена достигает заданной уровня-триггера. Это используется для ограничения потенциальных убытков или фиксации прибыли.
4. Ордер стоп-лимит с лимитированной ценой (Stop-Limit Order) похож на ордер стоп-лимит, этот ордер срабатывает, когда рыночная цена достигает указанного уровня-триггера. Однако, он выполняется как ордер с лимитированной ценой по указанной цене или лучше.

Типовая структура децентрализованной криптобиржи. Далее предлагается рассмотреть типовую структуру децентрализованной биржи на примере Uniswap v3 [9, 10].

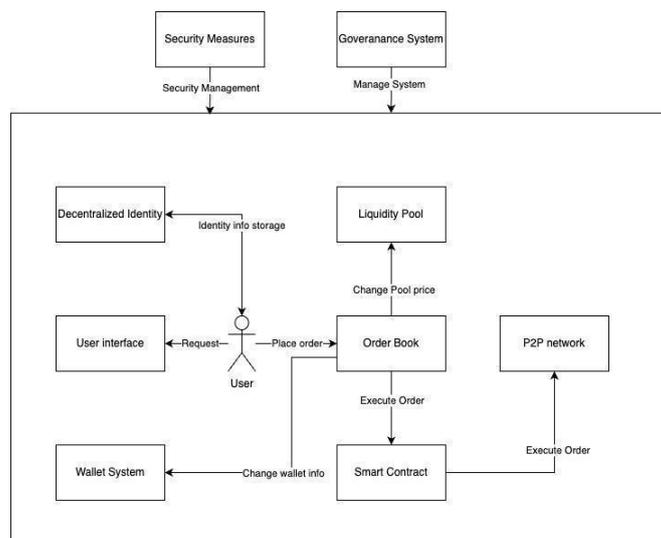


Рис. 2. Схема типовой структуры децентрализованной криптобиржи

1. Пользовательский интерфейс (User interface):
 - a. Веб-сайт или мобильное приложение, предоставляющее доступ к функциям биржи.
 - b. Позволяет пользователям создавать учетные записи, управлять кошельками, размещать ордера и просматривать историю транзакций.
2. Система кошельков (Wallet System):
 - a. У каждого пользователя есть личный кошелек, связанный с его учетной записью.
 - b. Хранит приватные ключи пользователя, необходимые для подписи транзакций.
 - c. Предоставляет безопасный способ хранения, отправки и получения криптовалют.
3. Книга ордеров (Order Book):
 - a. Управляет и сопоставляет покупку и продажу ордеров от пользователей.
 - b. Хранит и отображает текущие заявки покупки и продажи.
 - c. Позволяет пользователям размещать рыночные, лимитные и стоп-ордера.
4. Смарт контракты (Smart Contract):
 - a. Реализует логику и правила для выполнения сделок.
 - b. Обеспечивает прозрачность, безопасность и доверие с помощью технологии блокчейн.
 - c. Используется для получения он-чейн данных сторонними децентрализованными приложениями.
5. Пул ликвидности (Liquidity Pool):
 - a. Децентрализованный фонд, используемый для торговли.
 - b. Пользователи могут вносить средства в пул и получать вознаграждение.
 - c. Обеспечивает достаточную ликвидность для торговых операций.
6. Децентрализованная система хранения данных пользователей (Decentralized Identity):
 - a. Предоставляет пользователям контроль над их персональной информацией.
 - b. Обеспечивает безопасное и конфиденциальное создание учетных записей.
 - c. Позволяет пользователям сохранять владение и контроль над своими данными.
7. Peer-to-Peer сеть (P2P network):
 - a. Использует децентрализованную сеть для сопоставления ордеров и осуществления сделок.
 - b. Пользователи взаимодействуют друг с другом напрямую без посредников.
 - c. Снижает зависимость от центрального агентства и способствует доверительным сделкам.

8. Система безопасности (Security Measures):
 - a. Внедряет надежные меры безопасности для защиты средств пользователей и данных.
 - b. Использует шифрование для коммуникаций пользователей и хранения данных.
 - c. Проводит регулярные аудиты безопасности и программы поиска уязвимостей.
 - d. Предлагает аутентификацию с несколькими факторами для повышения безопасности.
9. Система управления (Governance System):
 - a. Позволяет держателям токенов участвовать в принятии решений.
 - b. Предоставляет механизмы голосования для обновлений платформы, структур комиссий и других предложений.
 - c. Способствует демократическому и общественному участию.
10. Интеграция с блокчейн-сетями (Other Blockchain Integration):
 - a. Обеспечивает совместимость с различными блокчейн-сетями.
 - b. Пользователи могут торговать различными поддерживаемыми биржей криптовалютами.

В децентрализованных крипто валютных биржах (DEX), таких как Uniswap, доступные типы ордеров отличаются из-за специфики децентрализованной платформы. Основной тип ордера, используемый на DEX, это Ордер Автоматизированного маркет мейкера (АММ). Вместо традиционных типов ордеров, DEX полагаются на пулы ликвидности и АММ, где трейдеры предоставляют ликвидность в пулы, и активы торгуются напрямую с ними. В этом случае взаимодействие осуществляется со смарт контрактами, и цены определяются алгоритмически на основе соотношений ликвидности каждой монеты в пуле [11].

Для сравнения приведем несколько примеров самых популярных децентрализованных криптобиржи [12]. Биржи были выбраны по количеству «залоченых» средств (TVL) и активности пользователей и проторгованному объему (24h Volume) [13].

1. Uniswap – децентрализованная биржа, построенная на блокчейне Ethereum. Она позволяет пользователям торговать токенами ERC-20 непосредственно из своих кошельков.

2. PancakeSwap – децентрализованная биржа, построенная на Binance Smart Chain (BSC). Она предоставляет возможности низкой стоимостной и быстрой торговли для токенов в этой сети.

3. SushiSwap – децентрализованная биржа, которая предоставляет расширенные возможности по сравнению с Uniswap, такие как возможности фарминга¹ и стейкинга.

4. 1inch – децентрализованный агрегатор обмена, который находит лучшие пути для торговли через несколько децентрализованных бирж, предлагая пользователям лучшие цены.

5. Curve Finance – децентрализованная биржа, оптимизированная для торговли стейбл монетами². Она предлагает низкую волатильность и низкие комиссии для обмена стейбл монет.

6. Balancer – децентрализованная биржа и автоматизированный портфельный менеджер. Она позволяет пользователям создавать и управлять пулами ликвидности с несколькими токенами и настраиваемыми весами.

¹ Фарминг – это процесс участия в децентрализованных финансовых (DeFi) проектах, где пользователи предоставляют свои криптовалютные активы для получения вознаграждения в виде процентов или новых токенов.

² Стейблкоины – это криптовалюты, стоимость которых привязана к стабильным активам, таким как фиатные валюты (например, доллар США), драгоценные металлы или другие стабильные активы.

7. Bancor – это децентрализованный протокол ликвидности, который позволяет пользователям торговать токенами непосредственно из своих кошельков с непрерывной ликвидностью.

8. dYdX – это децентрализованная платформа маржинальной торговли, построенная на блокчейне Ethereum. Она позволяет пользователям торговать с использованием плеча, при этом сохраняя контроль над своими средствами.

Правила и принципы работы трейдеров и маркет мейкеров. Одной из важных информационных составляющих, способных существенно влиять на поведение трейдеров криптовалют, является новости о регулирующих мерах и правительственной политике в отношении криптовалют. Основные принципы торговли для автоматизированных маркет-мейкеров на крипто биржах, как децентрализованных, так и централизованных, могут включать следующие аспекты [14, 15]:

1. Ликвидность. Маркетмейкеры стремятся обеспечить адекватную ликвидность торговых пар на бирже. Они предоставляют постоянную двустороннюю котировку, готовы купить и продать активы по различным ценам.

2. Спреды. Хорошие маркет-мейкеры стараются поддерживать низкие спреды между ценами покупки и продажи, чтобы сделки были более выгодными для трейдеров. Это обычно привлекает больше участников на биржу и способствует большей активности торговли.

3. Автоматизация. Автоматизированные маркет-мейкеры базируются на технологии, которая позволяет им быстро и эффективно реагировать на изменения цен и выставлять соответствующие ордера. Это обеспечивает непрерывность котировок и более точное отражение текущих рыночных условий.

4. Управление рисками. Маркет-мейкеры должны быть осведомлены о рисках, связанных с рыночной волатильностью и прочими потенциальными угрозами. Они должны разрабатывать стратегии управления рисками, чтобы минимизировать потери и обеспечить стабильность своей деятельности.

5. Технический анализ. Маркет-мейкеры основывают свои решения на техническом анализе рынка. Они используют инструменты анализа данных и моделирование ценовых движений для принятия решений о котировках и исполнении ордеров.

6. Глубина рынка. Маркет-мейкеры стремятся поддерживать глубину рынка для активов, чтобы облегчить трейдерам покупку и продажу значительных объемов. Это также позволяет снизить влияние крупных ордеров на рыночные цены.

7. Мониторинг и обновление. Маркет-мейкеры постоянно мониторят рынок и обновляют свои котировки в соответствии с изменениями спроса и предложения. Это помогает им адаптироваться к рыночным условиям и поддерживать конкурентоспособность.

Помимо этого, необходимо принимать в учет, что имеется некоторое количество факторов, которые влияют на поведение рынка. В ходе анализа рынка, было выделено несколько факторов, которые имеют непосредственное влияние на котировки на децентрализованных и централизованных биржах:

1. Предложение и спрос. Основной экономический принцип предложения и спроса играет значительную роль в определении цен на платформах DEX и CEX. Если спрос на конкретную криптовалюту превышает предложение, цена обычно возрастает, и наоборот.

2. Ликвидность рынка. Большая ликвидность обычно приводит к более плавным движениям цен и узким разрывам между ценой продажи и покупки. Платформы CEX часто имеют более высокую ликвидность по сравнению с DEX, так как обычно имеют более крупную пользовательскую базу, больше торговых пар и профессиональных маркетмейкеров.

3. Торговый объем. Торговый объем представляет собой общий объем торговой активности криптовалютой за определенный период. Большой объем торгов часто указывает на большое участие на рынке и может способствовать увеличению волатильности цен.

4. Манипуляция рынком и спекуляция. Рынки криптовалют подвержены манипуляции рынком, схемам "качалка" и спекулятивным торговым операциям. Эти факторы могут влиять на цены как на DEX, так и на CEX.

5. Регулирование. Государственные регулирования и политика, связанная с криптовалютами и биржами, могут серьезно повлиять на их цены. Положительные регуляторные меры, такие как признание или принятие криптовалюты, могут привести к ее росту, тогда как негативные действия могут вызвать снижение цен.

6. Технологические обновления и платформы. Обновления в блокчейн-технологии, протоколах или конкретных платформах DEX/CEX могут повлиять на воспринимаемую стоимость и функциональность криптовалют, торгуемых на этих платформах. Положительные новости и обновления могут привести к увеличению интереса к покупке и повышению цен.

7. Безопасность и взломы. Случаи нарушения безопасности, взломы или уязвимости, связанные с определенными платформами DEX или CEX, могут негативно сказаться на ценах. Пользователи могут потерять доверие к безопасности платформы, что может привести к потенциальному снижению цен.

8. Инвесторское поведение. Настроения и психология рынка также влияют на цены криптовалют. Положительные новости и оптимизм рынка или негативные новости и страх могут поднять или опустить цены, так как инвесторы реагируют на изменяющиеся рыночные условия.

Таким образом, для прогнозирования поведения рынка в той или иной ситуации, необходимо руководствоваться не только стандартными входными данными, но и учитывать социальные факторы [16, 17].

Помимо основных внешних факторов, необходимо рассмотреть виды стратегий, которые используют трейдеры и брокеры в своей деятельности. В рамках данной работы были выделены следующие стратегии, которые позволяют трейдерам реализовывать торговые операции:

1. Перемещающееся среднее пересечение (Moving Average Crossover). Данная стратегия основана на сравнении коротких (например, 50-дневное) и длинных (например, 200-дневное) скользящих средних с целью выявления сигналов к покупке и продаже при пересечении краткосрочной скользящей средней сверху или снизу от долгосрочной скользящей средней.

2. Полосы Боллинджера (Bollinger Bands). Этот алгоритм использует стандартные отклонения для создания верхней и нижней торговых полос вокруг скользящей средней. Когда цена касается верхней полосы, генерируется сигнал к продаже, а когда она касается нижней полосы, генерируется сигнал к покупке.

3. Возврат к среднему (Mean Reversion). Данная стратегия предполагает, что цены активов с течением времени имеют тенденцию возвращаться к своему среднему значению. Трейдеры идентифицируют перекупленные и перепроданные условия, открывая сделки в противоположном направлении и ожидая, что цена вернется к своему среднему уровню.

4. Агрессивная торговля (Breakout Trading). Этот алгоритм стремится выявить значимые уровни цен, такие как уровни поддержки и сопротивления, и осуществляет сделки, когда цена выходит за пределы этого диапазона. Это позволяет извлечь прибыль из увеличенной волатильности и импульса после прорыва.

5. Подход следования за трендом (Trend Following). Как следует из названия, эта стратегия стремится захватить тренды в ценах активов. Она включает в себя определение направления тренда и открытие сделок в этом направлении, с надеждой на продолжение тренда и получение прибыли от его продолжения.

6. Арбитраж (Arbitrage). Этот алгоритм стремится извлечь прибыль из различий в ценах между различными рынками или активами с низким или нулевым риском. Трейдеры одновременно покупают и продают связанные активы по разным ценам, извлекая прибыль из разницы в ценах [18].

7. Парная торговля (Pairs Trading). Данная стратегия предполагает идентификацию двух взаимосвязанных активов и одновременное открытие длинных и коротких позиций по ним. Трейдер ожидает, что разница между активами вернется к своему среднему значению, извлекая прибыль из относительных изменений цен.

8. Торговля с высокой частотой (HFT): Стратегии HFT используют мощные компьютеры и алгоритмы для выполнения большого количества сделок в течение секунд или миллисекунд. Эти алгоритмы основаны на небольших различиях в ценах или торговых моделях для извлечения прибыли из краткосрочных неэффективностей рынка [19].

9. Торговля на волатильности (Volatility Trading). Данная стратегия основана на извлечении прибыли из изменений в неявной волатильности опционов или других производных инструментов. Трейдеры могут открывать позиции, когда ожидается рост или снижение волатильности, используя различные индикаторы и модели волатильности.

10. Торговля на основе новостей (News-Based Trading). Трейдеры, использующие эту стратегию, анализируют новостные события и экономические показатели для предсказания реакции рынка. Позиции открываются перед или сразу после важных новостных объявлений, с целью извлечения прибыли из результирующих изменений цен [20].

Заключение. Целью последующего исследования является разработка метода предиктивной аналитики для децентрализованных систем. В рамках исследования предполагается на примере процессов рынка криптовалют на децентрализованных биржах разработать математическую модель функционирования бирж для предсказания поведения сложных децентрализованных систем. Предполагается, что разработанная математическая модель будет представлять процесс в виде иерархической структуры централизованной системы, а при помощи предложенных методов интерполяции появится возможность приведения модели децентрализованных систем к модели централизованных систем. Перечень последующих задач исследования:

1. Разработать методологию приведения модели работы децентрализованных систем к модели централизованных систем на примере блокчейн технологий.
2. Разработать метод предиктивной аналитики для предсказания состояния децентрализованной системы на примере блокчейн технологий.
3. Используя разработанную методологию, предложить инструмент для принятия решений в работе с децентрализованными системами при помощи искусственного интеллекта.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Зыбин Е.Ю., Косьянчук В.В.* Алгебраический критерий обнаружения факта и времени возникновения отказов в системах управления динамическими объектами // Известия РАН. Теория и системы управления. – 2016. – № 4. – С. 50-61.
2. *Мироновский Л.А.* Функциональное диагностирование динамических систем. – СПб.: Изд-во МГУ-ГРИФ, 1998.
3. *Jiang Y, Yin S, Kaynak O.* Data-driven monitoring and safety control of industrial cyber-physical systems: Basis and beyond // IEEE Access. – 2018. – Vol. 6. – P. 47374-47384.
4. *Angeris G., Chitra T.* Improved Price Oracles: Constant Function Market Makers. In Proceedings of the 2nd ACM Conference on Advances in Financial Technologies (AFT '20) // Association for Computing Machinery, – New York, 2020. – P. 80-91.
5. A Deep Dive into Decentralized Exchanges. – URL: <https://coinmarketcap.com/alexandria/article/a-deep-dive-into-decentralizedexchanges> (accessed: 21/01/2021).
6. *Alexander C., Heck D.F.* Price discovery in bitcoin: The impact of unregulated markets // Journal of Financial Stability. – 2020. – No. 50. – P. 1-18.
7. *Angeris G., Chitra T.* Improved price oracles: Constant function market makers // Conference: AFT '20: 2nd ACM Conference on Advances in Financial Technologies. – 2020. – P. 80-91.
8. *Chen M.A., Wu Q., Yang B.* How valuable is FinTech innovation? // Review of Financial Studies. – 2019. – Vol. 32 (5). – P. 2062-2106.
9. *Cohen J.E.* Between Truth and Power: The Legal Constructions of Informational Capitalism. – New York: Oxford University Press, 2019. – 376 p.
10. *Cong L.W., He Z.* Blockchain disruption and smart contracts // Review of Financial Studies. – 2019. – Vol. 32 (5). – P. 1754-1797.
11. *Egorov M.* StableSwap - Efficient Mechanism for Stablecoin Liquidity. – URL: <https://www.curve.fi/stableswap-paper.pdf>.
12. *Lin L.X.* Deconstructing decentralized exchanges // Stanford Journal of Blockchain Law & Policy. – URL: <https://stanfordjblp.pubpub.org/pub/deconstructing-dex> (accessed: 26/01/2021).

13. Murray A., Kuban S., Josefy M., Anderson J. Contracting in the smart era: the implications of blockchain and decentralized autonomous organizations for contracting and corporate governance // *Academy of Management Perspectives*. – 2019. – P. 51-55.
14. Buterin V. Ethereum whitepaper. URL: <https://ethereum.org/en/whitepaper/>.
15. Adams H., Zinsmeister N., Salem M., Keefer R., Robinson D. Uniswap v3 Core. – URL: <https://uniswap.org/whitepaper-v3.pdf>.
16. Li X., Wang R., Li Y., Qin X., Liu S. Predicting Bitcoin Prices from Twitter with LSTM Network // *IEEE Access*. – 2021. – P. 17253-17265.
17. Kim Y., Kim J., Kim H., Rhim J., & Ryu K.H. Bitcoin price prediction using machine learning: An approach to predict the price of Bitcoin based on Twitter sentiment analysis // *International Conference on Platform Technology and Service (PlatCon)*. – 2018. – P. 1-5.
18. Yoon J.H., & Chang H. Predicting cryptocurrency prices using machine learning techniques // *Applied Intelligence*. – 2020. – P. 1258-1272.
19. Verkouteren J.M., & van Vliet B. Machine learning for financial prediction: Hype or hope // *Expert Systems with Applications*. – 2019. – P. 1-22.
20. Turel O., Ul Ibrahim M., & Altun G.B. Predicting Cryptocurrency Prices Using Machine Learning Techniques: A Comparative Analysis // *Journal of Organizational and End User Computing (JOEUC)*. – 2021. – P. 79-94.

REFERENCES

1. Zybin E.Yu., Kos'yanchuk V.V. Algebraicheskiy kriteriy obnaruzheniya fakta i vremeni vozniknoveniya otkazov v sistemakh upravleniya dinamicheskimi ob'ektami [Algebraic criteria for detecting the occurrence and timing of failures in control systems of dynamic objects], *Izvestiya RAN. Teoriya i sistemy upravleniya* [Proceedings of the Russian Academy of Sciences. Theory and control systems], 2016, No. 4, pp. 50-61.
2. Mironovskiy L.A. Funktsional'noe diagnostirovanie dinamicheskikh system [Functional diagnostics of dynamic systems]. St. Petersburg: Izd-vo MGU-GRIF, 1998.
3. Jiang Y, Yin S, Kaynak O. Data-driven monitoring and safety control of industrial cyber-physical systems: Basis and beyond, *IEEE Access*, 2018, Vol. 6, pp. 47374-47384.
4. Angeris G., Chitra T. Improved Price Oracles: Constant Function Market Makers. In Proceedings of the 2nd ACM Conference on Advances in Financial Technologies (AFT '20), *Association for Computing Machinery*, New York, 2020, pp. 80-91.
5. A Deep Dive into Decentralized Exchanges. Available at: <https://coinmarketcap.com/alexandria/article/a-deep-dive-into-decentralized-exchanges> (accessed: 21/01/2021).
6. Alexander C., Heck D.F. Price discovery in bitcoin: The impact of unregulated markets, *Journal of Financial Stability*, 2020, No. 50, pp. 1-18.
7. Angeris G., Chitra T. Improved price oracles: Constant function market makers, *Conference: AFT '20: 2nd ACM Conference on Advances in Financial Technologies*, 2020, pp. 80-91.
8. Chen M.A., Wu Q., Yang B. How valuable is FinTech innovation?, *Review of Financial Studies*, 2019, Vol. 32 (5), pp. 2062-2106.
9. Cohen J.E. Between Truth and Power: The Legal Constructions of Informational Capitalism. New York: Oxford University Press, 2019, 376 p.
10. Cong L.W., He Z. Blockchain disruption and smart contracts, *Review of Financial Studies*, 2019, Vol. 32 (5), pp. 1754-1797.
11. Egorov M. StableSwap - Efficient Mechanism for Stablecoin Liquidity. Available at: <https://www.curve.fi/stableswap-paper.pdf>.
12. Lin L.X. Deconstructing decentralized exchanges // *Stanford Journal of Blockchain Law & Policy*. Available at: <https://stanfordjblp.pubpub.org/pub/deconstructing-dex> (accessed: 26/01/2021).
13. Murray A., Kuban S., Josefy M., Anderson J. Contracting in the smart era: the implications of blockchain and decentralized autonomous organizations for contracting and corporate governance, *Academy of Management Perspectives*, 2019, pp. 51-55.
14. Buterin V. Ethereum whitepaper. Available at: <https://ethereum.org/en/whitepaper/>.
15. Adams H., Zinsmeister N., Salem M., Keefer R., Robinson D. Uniswap v3 Core. Available at: <https://uniswap.org/whitepaper-v3.pdf>.
16. Li X., Wang R., Li Y., Qin X., Liu S. Predicting Bitcoin Prices from Twitter with LSTM Network, *IEEE Access*, 2021, pp. 17253-17265.
17. Kim Y., Kim J., Kim H., Rhim J., & Ryu K.H. Bitcoin price prediction using machine learning: An approach to predict the price of Bitcoin based on Twitter sentiment analysis, *International Conference on Platform Technology and Service (PlatCon)*, 2018, pp. 1-5.

18. Yoon J.H., & Chang H. Predicting cryptocurrency prices using machine learning techniques, *Applied Intelligence*, 2020, pp. 1258-1272.
19. Verkouteren J.M., & van Vliet B. Machine learning for financial prediction: Hype or hope, *Expert Systems with Applications*, 2019, pp. 1-22.
20. Turel O., Ul Ibrahim M., & Altun G.B. Predicting Cryptocurrency Prices Using Machine Learning Techniques: A Comparative Analysis, *Journal of Organizational and End User Computing (JOEUC)*, 2021, pp. 79-94.

Статью рекомендовал к опубликованию, д.т.н., профессор А.Н. Целых.

Сергеев Николай Евгеньевич – Южный федеральный университет; e-mail: nesergeev@sfedu.ru; г. Таганрог, Россия; тел.: +79281742585; доцент; д.т.н., профессор.

Веселова Диана Геннадьевна – e-mail: diaveselova@sfedu.ru; тел.: +79818640317; аспирант.

Sergeev Nikolay Yevgenyevich – Southern Federal University; e-mail: nesergeev@sfedu.ru; Taganrog, Russia; phone: +79281742585; associate professor; dr. of eng. sc.; professor.

Veselova Diana Gennadievna – e-mail: diaveselova@sfedu.ru; phone: +79818640317; graduate student.

УДК 004.021

DOI 10.18522/2311-3103-2024-2-100-109

Л.А. Зинченко, В.В. Казаков, Р.Р. Моисеев, Е.С. Поляков

**ПРОГРАММНЫЙ МОДУЛЬ ДЛЯ АВТОМАТИЗИРОВАННОГО
ПРОЕКТИРОВАНИЯ МНОГОСЛОЙНЫХ ЭКРАНОВ ЗАЩИТЫ ЭЛЕКТРОННОЙ
АППАРАТУРЫ ОТ ВОЗДЕЙСТВИЯ ТЯЖЕЛЫХ ЗАРЯЖЕННЫХ ЧАСТИЦ
С ИСПОЛЬЗОВАНИЕМ GEANT4**

Представлено открытое программное обеспечение для автоматизации проектирования экранов радиационной защиты с использованием Geant4 для защиты электронного оборудования от тяжелых заряженных частиц. В статье представлена выбранная архитектура для реализации предложенного подхода, а также формулируются требуемые входные данные и результирующие выходные данные. Описан маршрут проектирования экрана на основе входных данных о материале как последовательность создания соответствующих классов-наследников. Данная статья представляет собой глубокое исследование, посвященное разработке открытого программного обеспечения на базе фреймворка Geant4, который использует метод Монте-Карло, нацеленного на автоматизацию процесса проектирования экранов радиационной защиты с целью обеспечения эффективной защиты электронного оборудования от воздействия тяжелых заряженных частиц. В статье подробно рассматривается архитектура разработанного программного обеспечения, включая описание основных компонентов и технологий, применяемых в его создании, а также определение необходимых входных данных и формулирование требований к программному продукту. Представленный маршрут проектирования экрана описывается как последовательность создания соответствующих классов-наследников и их взаимодействия в рамках разработанной архитектуры, что обеспечивает эффективность и точность расчетов радиационной защиты. Результаты данной работы представляют собой новый инновационный подход к проектированию экранов радиационной защиты, который имеет потенциал значительно повысить надежность и безопасность электронных систем в условиях воздействия тяжелых заряженных частиц. Они оказывают важное практическое значение для специалистов в области радиационной защиты и разработки электроники, предоставляя им эффективный инструмент для проведения анализа и оптимизации экранов радиационной защиты. Кроме того, результаты исследования представляют интерес для исследователей, работающих в области моделирования радиационных эффектов и разработки новых методов защиты электроники от воздействия радиации. В целом, статья представляет собой значимый вклад в область радиационной защиты и электроники, а также является основой для дальнейших исследований и разработок в этом направлении.

Радиация; экранирование; Geant4; фреймворк; заряженные частицы.

L.A. Zinchenko, V.V. Kazakov, R.R. Moiseev, E.S. Polyakov

A SOFTWARE FOR AUTOMATED DESIGN OF MULTILAYER SHIELDING FOR ELECTRONIC EQUIPMENT PROTECTION FROM HEAVY CHARGED PARTICLES BASED ON GEANT4

The article presents open source software for automating the design of radiation protection screens using Geant4 to protect electronic equipment from heavy charged particles. The article presents the chosen architecture for the implementation of the proposed approach, and also formulates the required input data and the resulting output data. The route of designing a screen based on input data about the material is described as a sequence of creating appropriate successor classes. This article is an in-depth study devoted to the development of open source software based on the Geant4 framework, which uses the Monte Carlo method aimed at automating the design process of radiation protection screens in order to ensure effective protection of electronic equipment from the effects of heavy charged particles. The article examines in detail the architecture of the developed software, including a description of the main components and technologies used in its creation, as well as determining the necessary input data and formulating requirements for the software product. The presented screen design route is described as a sequence of creating appropriate successor classes and their interaction within the framework of the developed architecture, which ensures the efficiency and accuracy of radiation protection calculations. The results of this work represent a new innovative approach to the design of radiation protection screens, which has the potential to significantly increase the reliability and safety of electronic systems under the influence of heavy charged particles. They are of great practical importance for specialists in the field of radiation protection and electronics development, providing them with an effective tool for analyzing and optimizing radiation protection screens. In addition, the results of the study are of interest to researchers working in the field of modeling radiation effects and developing new methods for protecting electronics from radiation exposure. In general, the article represents a significant contribution to the field of radiation protection and electronics, and is also the basis for further research and development in this direction.

Radiation; shielding; Geant4; framework; charged particles.

Введение. В последние годы электронная аппаратура широко используется в различных космических программах. Однако бортовое электронное оборудование, предназначенное для работы в экстремальных условиях, должно быть защищено от радиационного воздействия [1–4].

При прохождении поясов Ван Аллена бортовое электронное оборудование подвергается воздействию тяжелых заряженных частиц (ТЗЧ) [3], электронов и прочих частиц и может накапливать заряд, полученный при воздействии ТЗЧ [5–8]. В частности, биты данных в памяти микросхемы могут изменяться под воздействием ТЗЧ, что может привести к сбоям и некорректной работе бортовых электронных устройств. В связи с вышеописанными проблемами требуется защита бортовых электронных устройств от ТЗЧ.

Одним из методов защиты электронного оборудования является экранирование. В работе [9] рассматривается программное обеспечение для проектирования экранов. Однако в данном случае было использовано программное обеспечение SRIM для оценки воздействия ТЗЧ.

Программное обеспечение Geant4 [10–14], используя в своей основе более сложные по сравнению с программой SRIM методики расчета и модели, позволяет оценивать эффекты и процессы в различных материалах. Тем не менее, одним из серьезных недостатков использования программного обеспечения Geant4 является сложность настройки для проведения вычислений и сложность анализа полученных данных, что требует значительного времени инженера-проектировщика для анализа различных альтернативных проектных решений. Еще одним недостатком Geant4 является необходимость вручную вводить данные: для расчета каждого альтернативного проектного решения: генератор должен подготовить файл с входными параметрами или использовать генератор входных файлов. Если инженер хочет разработать собственный сценарий для расчета, он сталкивается с трудностями из-за структуры API Geant4 и необходимости глубоких познаний в языке программирования Python или C++.

Растущий объем данных, который требуется обработать, требуют внедрения новых подходов к работе с информацией, например, веб-ориентированных мобильных инструментов для обработки данных.

В статье представлен модуль разработанного программного обеспечения для автоматизации проектирования экранов с целью защиты электронного оборудования от ТЗЧ. Разработанный модуль автоматизирует процесс ввода данных для расчета при помощи интерфейса и позволяет хранить результаты расчета.

Обзор программного обеспечения Geant4. Geant4 [10] представляет собой программный пакет, разработанный для моделирования прохождения частиц через материал. Это программное обеспечение может быть использовано для моделирования детекторов и систем в высокоэнергетической физике, проведения экспериментов в ядерных установках, а также при проектировании электронной аппаратуры.

Geant4 основан на широком наборе физических моделей для моделирования взаимодействиями частиц и вещества в широком диапазоне энергий. Данные были получены из многих источников по всему миру и в настоящее время Geant4 является хранилищем, включающим в себя большую часть известной информации о взаимодействиях частиц.

Пакет программного обеспечения Geant4 включает в себя следующие инструменты:

- ◆ Настройка геометрии системы.
- ◆ Настройка используемых материалов.
- ◆ Генерация первичных событий: задание энергии частиц, угла, положения и т.д.
- ◆ Отслеживание прохождения частиц через материалы и электромагнитные поля.
- ◆ Физические процессы, определяющие взаимодействие частиц.
- ◆ Реакция чувствительных компонентов детектора.
- ◆ Генерация данных о событиях.
- ◆ Хранение событий и треков.
- ◆ Визуализация детектора и траекторий частиц.

Организация процесса моделирования с использованием Geant4 основана на объектно-ориентированной методологии. Geant4 предоставляет следующие возможности взаимодействия с программным пакетом:

- ◆ Пользовательский интерфейс (UI).
- ◆ Интерпретаторы команд.
- ◆ Управление набором инструментов.

Особенностью пользовательского интерфейса является взаимодействие с Geant4 в интерактивном формате с использованием стороннего фреймворка, такого как OpenGL, Qt.

У Geant4 есть встроенный набор команд, которые управляют процессом моделирования. С помощью встроенных команд пользователи могут устанавливать геометрию, изменять список физических процессов, которые должны учитываться и создавать первичные события. Управление моделированием с помощью команд может осуществляться непосредственно в консоли в интерактивном режиме или с использованием скрипта.

Geant4 предоставляет широкий выбор создания геометрии: базовые формы, такие как параллелепипед (G4Box), трубы (G4Tube), сферы и т. д., а также, используя булевы функции, пользователи могут создавать сложные формы с использованием вырезов, объединений, исключений и т.д.

Geant4 написан на C++. Это позволяет пользователю подходить к процессу моделирования с использованием объектно-ориентированной методологии [17]. Пользователь может взаимодействовать более гибко с программным пакетом Geant4, используя встроенные классы и функции и создавая свои собственные скрипты. Объектно-ориентированные методы помогают управлять сложностью и ограничивать зависимости, определяя единый интерфейс и общие организационные принципы для всех физических моделей. В этой структуре функционирование моделей понятнее, а создание и добавление новых моделей представляет собой хорошо определенную процедуру, практически не влекущую изменений в существующем коде.

Существует вариант реализации на языке C++ [16]. Особенностью реализации приложений на C++ является прямой доступ к Geant4. Для Geant4 написана подробная документация на C++, которая описывает процесс моделирования и использования инструмента. Однако для этого требуются продвинутое знание и навыки программирования на C++, необходимо понимание принципов объектно-ориентированного программирования, таких как наследование, полиморфизм и т. д., что существенно повышает требования к пользователю. Более того, для компиляции проекта, написанного с использованием Geant4, пользователи должны использовать CMake [19]. Тем не менее, все трудности моделирования на C++ компенсируются производительностью проекта - вычисления и моделирование будут более продуктивными по сравнению с другими вариантами.

Преимущества:

- 1) Высокая производительность: написание скрипта на C++ обеспечивает высокую производительность, что особенно важно для сложных вычислительных задач.
- 2) Полный доступ к Geant4: полный доступ ко всей функциональности Geant4, что позволяет создавать сложные симуляции и настраивать их до мельчайших деталей.
- 3) Полное соответствие документации на C++ существенно снижает риск ошибок, связанных с компиляцией проекта и с ядром Geant4.

Недостатки:

- 1) Сложность и продолжительность разработки: разработка моделей на C++ может быть сложной и занимать много времени.
- 2) Управление памятью: пользователям придется самостоятельно управлять памятью, что может привести к такой ошибке, как утечка памяти.
- 3) Требуется более глубокое понимание C++ и Geant4.

Также существует вариант реализации с использованием Python библиотеки Geant4Py [17]. Этот вариант основан на языке программирования Python и представляет собой высокоуровневую оболочку над Geant4 с более простым и понятным интерфейсом. Это абстрактный уровень, который делает разработку более доступной, но также может ограничивать некоторые возможности Geant4.

Преимущества:

- 1) Удобство и скорость разработки: простой и читаемый синтаксис Python упрощает создание моделей и прототипов.
- 2) Управление памятью: Python автоматически обрабатывает управление памятью, снижая риск утечек памяти и облегчая разработку.

Недостатки:

- 1) Низкая производительность: Python обычно медленнее по сравнению с C++, поэтому может быть не подходящим для сложных вычислительных задач.
- 2) Ограниченный доступ к Geant4: библиотека Geant4Py для Python может иметь ограничения в доступе к функциональности Geant4 по сравнению с версией на C++.

Существует также вариант реализации с использованием Python библиотеки Geant4-rubind [18]. Этот метод также основан на языке программирования Python, но представляет собой альтернативное решение для взаимодействия с Geant4 по сравнению с rubind11 (Geant4Py). Он имеет слабую связь с модулем g4py, на котором основан Geant4Py, но сохраняет API, близкое к стандарту C++. Более того, с использованием Python становится возможным использование сторонних модулей для расширения возможностей моделирования (визуализация, хранение данных, аналитика и т. д.). Однако на данный момент этот модуль не позволяет полностью перенести API Geant4 на Python и проводить моделирование в полном объеме, как в версии на C++.

Преимущества:

- 1) Прямой доступ к Geant4 на Python: этот метод обеспечивает прямой доступ к функциональности API Geant4 на Python, обеспечивая баланс между производительностью и удобством разработки.
- 2) Относительно высокая производительность: по сравнению с Geant4Py, geant4-rubind может обеспечивать лучшую производительность с использованием библиотеки rubind11.

Недостатки:

1) Более сложное использование: настройка и использование geant4-pybind может потребовать более глубоких знаний как C++, так и Python.

2) Как и предыдущее решение на Python, этот модуль не позволяет реализовать слишком сложные моделирования.

На основании вышеизложенного могут быть сделаны следующие выводы.

◆ Если проект требует максимальной производительности и полного доступа к Geant4, то выбор C++ может быть предпочтительным.

◆ Если основное внимание уделяется скорости разработки и нет требований к высокой производительности, то использование библиотек Python может быть более удобным.

◆ Если необходим баланс между производительностью и удобством разработки, и требуется доступ к Geant4, то geant4-pybind может быть подходящим выбором.

В табл. 1 систематизированы особенности каждого из вариантов реализации.

Таблица 1

Сравнение фреймворков

Фреймворк	Сравниваемые параметры			
	Производительность	Удобство разработки	Доступ к Geant4	Управление памятью
C++	Очень высокая	Низкая	Полный	Да
Geant4Py	Низкая	Высокая	Ограниченный	Да (автоматическое)
Geant4-pybind	Средняя	Средняя	Средний	Да (частично автоматическое)

В данной работе для реализации программного модуля был выбран Geant4-pybind, так как он обеспечивает баланс между производительностью и удобством разработки, а также предоставляет доступ к Geant4. Это означает, что разрабатываемый модуль удобно было разрабатывать и в то же время имелся доступ к различным возможностям Geant4.

Модуль автоматизации проектирования экранов защиты от ТЗЧ. Для реализации был разработан веб-интерфейс для ввода входных данных, который позволяет пользователю определить все необходимые условия и данные.

На рис. 1 показан состав разработанного программного обеспечения.

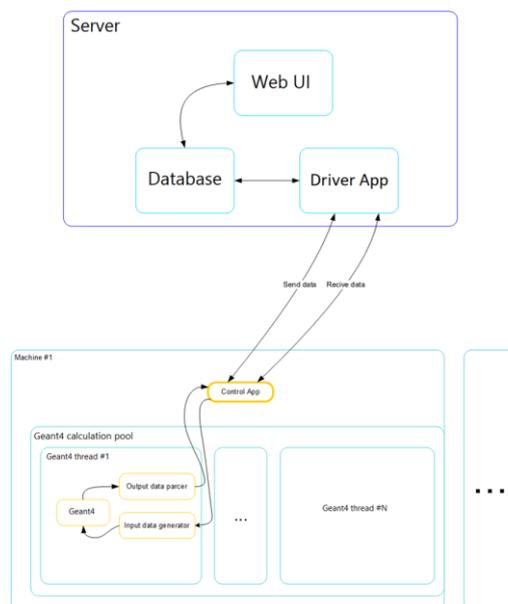


Рис. 1. Состав разработанного программного модуля, основанного на Geant4

Задача моделирования была разделена на отдельные подзадачи следующим образом. Серверу были назначены следующие функции:

- ◆ Хранение входных данных для расчета.
- ◆ Хранение результатов расчета.

Клиенту были поручены следующие задачи:

- ◆ Получение данных для расчета из различных альтернативных проектных решений.
- ◆ Генерация входного файла для расчета.
- ◆ Расчет многослойных экранов с использованием программы Geant4.
- ◆ Анализ результатов расчета.
- ◆ Вывод данных в формате JSON.

Такой подход позволяет использовать распределенную систему вычислений. Это позволило ускорить расчеты, разделяя устройства пользователей и узлы моделирования. С веб-интерфейсом пользователи могут использовать смартфоны или ноутбуки и быстро переключаться между ними в режиме реального времени.

Входные данные включают в себя:

- ◆ экран: количество слоев в экране, их толщину, используемые материалы;
- ◆ частицы: тип частицы, ее энергию, количество запущенных частиц.

Вся необходимая информация сохраняется в удобном формате JSON.

Входной JSON-файл содержит следующие поля:

- ◆ Название и описание экрана.
- ◆ Количество слоев в экране.
- ◆ Материалы экрана.
- ◆ Химические элементы для каждого материала.
- ◆ Атомные номера.
- ◆ Стандартная атомная масса.
- ◆ Плотность материала.
- ◆ Первичные частицы.
- ◆ Тип частицы.
- ◆ Атомная масса.
- ◆ Энергия первичных частиц.

После ввода данных в веб-интерфейс скрипт получает запрос от сервера входных данных в формате JSON. На следующем этапе согласно входным параметрам выполняется моделирование.

Для описания геометрии и материала экрана использован класс *G4UserDetectorConstruction*. Данный класс требует переопределения двух методов: *Construct*, который создает и возвращает физический объем (геометрия, материалы и расположение объекта), и *ConstructSDandField*, который задает «чувствительность» для экрана, чтобы обнаруживать проходящие через него частицы и определять ее параметры (кинетическая и потенциальная энергии, координаты, тип частицы и т.д.) [20].

Далее скрипт создает класс-наследник от *G4UserDetectorConstruction* для создания экранов.

В методе *Construct* класса определяется материал экранов. Для создания экрана это может быть как металл, так и сплав. В результате получается объект *G4Material*, который описывает все свойства материала: название, плотность, и т.д. Затем этот объект передается в *G4LogicalVolume* для создания логического объема.

Далее указываются габариты экрана. В нашей работе предполагается, что все экраны являются параллелепипедами, и в JSON-файле пользователь может определить их толщину. Для этого в Geant4 используется класс *G4Box*, который дает возможность создать параллелепипед, указывая в качестве аргументов его размеры. По аналогии с *G4Material*, объект класса *G4Box* передается в логический объем.

Логический объем описывает материал, электромагнитные поля, геометрию и детектор для данного объема, что необходимо для последующего моделирования. Следует отметить, что в логическом объеме, который является объектом класса *G4LogicalVolume*

[12], детектируются и генерируются вторичные частицы, и определяется изменение их энергии вследствие взаимодействия с экраном. Завершающим этапом описания экрана является их физическое расположение в пространстве.

Объект класса *G4PhysicalVolume* определяет координаты центра тела и его вращение относительно системы координат, связанной с центром родительского объема, внутри которого располагаются экраны, называемый *World*. В терминологии Geant4 объем *World* [15] является пространством, который окружает экраны. Он создается также по вышеописанному принципу и нужен для размещения всех дочерних объемов. В нашем случае для *World* выбран материал воздух *G4_Air*.

На рис. 2 показаны маршрут создания описания альтернативного проектного решения.

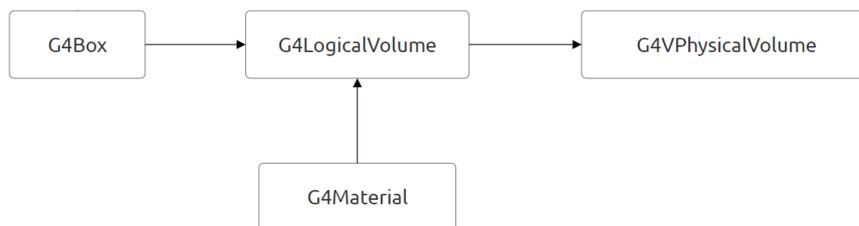


Рис. 2. Классы Geant4 для описания экранов защиты электронной аппаратуры

На следующем шаге необходимо указать, с какими частицами взаимодействует экран.

Для этого был использован класс *G4VUserPrimaryGeneratorAction*, который позволяет пользователям указать количество частиц, энергию, направление, а также расстояние, на котором они генерируются.

На этапе моделирования определяется класс, который является наследником класса *G4VUserPrimaryGeneratorAction* для указания взаимодействия частиц с экраном. В этом дочернем классе устанавливается число первичных частиц, энергия и направление.

Программа извлекает всю информацию о частицах из входного JSON-файла, предоставленного веб-сервером.

Для моделирования воздействия частиц программное обеспечение устанавливает каждый экран как чувствительный. Создается объект-потомок класса *G4VSensitiveDetector*. Таким образом, на каждом шаге прохождения частицы через экран сценарий анализирует текущие параметры частицы и также сохраняет данные для создания выходных данных.

Программное обеспечение сохраняет выходные данные в формате JSON. Выходной JSON-файл содержит следующие поля:

- ◆ Материалы экрана.
- ◆ Название материала.
- ◆ Количество задержанных первичных частиц.
- ◆ Количество задержанных вторичных частиц.
- ◆ Общее количество запущенных частиц.
- ◆ Общее количество излученных первичных частиц.
- ◆ Общее количество излученных вторичных частиц.

Выходные данные описывают созданный пользователем экран и результаты моделирования. JSON-файл содержит материал каждого слоя экрана и количество частиц, застрявших в соответствующем слое. Помимо первично сгенерированных частиц, выходные данные содержат информацию о вторично сгенерированных частицах: это частицы, которые образовались в результате взаимодействия других частиц с экраном. Эта информация важна, поскольку даже если все первичные частицы были задержаны экраном, это не гарантирует, что вторичные частицы не пройдут через экран и не повредят электронную аппаратуру.

На основе полученных результатов моделирования можно судить о степени защиты электронной аппаратуры от заряженных частиц.

Заключение. В статье рассмотрены особенности реализации разработанного программного обеспечения для автоматизации проектирования экранов для защиты электронного оборудования от тяжелых заряженных частиц. Выбранная клиент-серверная архитектура позволяет получать удаленный доступ к параметрам уже рассчитанных экранов с различных устройств и в различных операционных системах, а также выполнять расчет альтернативных проектных решений с использованием распределенных вычислений. Таким образом, программное обеспечение позволяет проводить автоматизированное проектирование экранов защиты электронной аппаратуры от воздействия ТЗЧ.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Shakhnov V., Zinchenko L., Kosolapov I., Filippov I.* Modeling and Optimization of Radiation Tolerant Microsystems // EMS '14 Proceedings of the 2014 European Modelling Symposium. – 2014. – P. 484-489.
2. *Glushko A.A., Zinchenko L.A., Shakhnov V.A.* Simulation of the impact of heavy charged particles on the characteristics of field-effect silicon-on-insulator transistors // Journal of Communications Technology and Electronics. – 2015. – No. 10.
3. *Кузнецов Н.В.* Радиационная опасность на околоземных орбитах и межпланетных траекториях космических аппаратов. – URL: <http://nuclphys.sinp.msu.ru/crd/index.html> (дата обращения: 01.03.2024).
4. *Glushko A.A., Morozov S.A., Chistyakov M.G.* Investigation of the sensitive region of the MOSFET to the effects of secondary particles arising from ionizing radiation // Microelectronics. – 2023. – Vol. 52, No. 4. – P. 282-289.
5. *Терехов В.В.* Выбор модели аппроксимации для исследования электрических характеристик микросистем под влиянием тяжелых ионов // Будущее машиностроения России. Всероссийская конференция молодых учёных и специалистов (с международным участием), 14-я, Москва, 21-24 сентября 2021 г.: Сб. докладов: в 2 т. МГТУ им. Н.Э. Баумана (национальный исследовательский ун-т), Союз машиностроителей России. – 2021. – Т. 2. – С. 247-253.
6. *Novikov L.S., Mileev V.N. et al.* Sinotic Radiation effects on materials in space technology // Surface. X-ray, synchrotron and neutron studies. – 2009. – No. 3. – P. 32-48.
7. *Хаффнер Дж.* Ядерное излучение и защита в космосе. – М.: Атомиздат, 1971.
8. *Александров П.А., Жук В.И., Литвинов В.Л.* Способы построения отказоустойчивых цифровых микросхем и оценки вероятности их отказа, вызванного облучением. – М., 2019.
9. *Shakhnov V.A., Zinchenko L.A., Rezchikova E.V., et al.* Visual Analytics and Its Applications in Electronic Engineering Education : BMSTU Case Study // 2022 6th International Conference on Information Technologies in Engineering Education, Inforino 2022 – Proceedings. – DOI: 10.1109/Inforino53888.2022.9782907.
10. *Allison J. et al.* Recent Developments in Geant4 // Nucl. Instrum. Meth. A. – 2016. – 835. – P. 186-225.
11. *Allison J. et al.* Geant4 Developments and Applications // IEEE Trans. Nucl. Sci. – 2006. – 53. – P. 270-278.
12. *Christopher M Poole, Iwan Cornelius, Jamie V Trapp, Christian M Langton.* A CAD Interface for GEANT4 // Australasian physical & engineering sciences in medicine.
13. *Agostinelli S. et al.* Geant4 - A Simulation Toolkit // Nucl. Instrum. Meth. A. – 2003. – 506. – P. 250-303.
14. *Geant4 User's Guide for Toolkit Developers: Geant4 Docs.* – URL: <https://geant4-userdoc.web.cern.ch/UsersGuides/ForToolkitDeveloper/html/index.html> (дата обращения: 22.10.2023).
15. *Geant4 Guide for Physics Lists: Geant4 Docs.* – URL: <https://geant4-userdoc.web.cern.ch/UsersGuides/PhysicsListGuide/html/index.html> (дата обращения: 22.10.2023).
16. *Geant4 - API.* – URL: <https://simoncblyth.github.io/env/notes/geant4/api/> (дата обращения: 22.10.2023).
17. *Geant4Py: Geant4-Python interface.* – URL: <https://github.com/koichi-murakami/g4python> (дата обращения: 22.10.2023)
18. *Geant4_pybind: Alternative Python bindings for Geant4.* – URL: https://github.com/HaarigerHarald/geant4_pybind (дата обращения: 22.10.2023).
19. *Tutorial and examples of Geant4.* – URL: <https://dev.asifmoda.com/en/geant4> (дата обращения: 22.10.2023).
20. *Sensitive Detector in Geant4: Seminar on Software for Nuclear, Subnuclear and Applied Physics.* – URL: https://agenda.infn.it/event/8781/contributions/75434/attachments/54976/64833/Cirrone_Sensitive_Detector.pdf (дата обращения: 22.10.2023).

REFERENCES

1. *Shakhnov V., Zinchenko L., Kosolapov I., Filippov I.* Modeling and Optimization of Radiation Tolerant Microsystems, *EMS '14 Proceedings of the 2014 European Modelling Symposium*, 2014, pp. 484-489.
2. *Glushko A.A., Zinchenko L.A., Shakhnov V.A.* Simulation of the impact of heavy charged particles on the characteristics of field-effect silicon-on-insulator transistors, *Journal of Communications Technology and Electronics*, 2015, No. 10.
3. *Kuznetsov N.V.* Radiatsionnaya opasnost' na okolozemnykh orbitakh i mezhplanetnykh traektoriyakh kosmicheskikh apparatov [Radiation danger in near-earth orbits and interplanetary trajectories of spacecraft]. Available at: <http://nuclphys.sinp.msu.ru/crd/index.html> (accessed 01 March 2024).
4. *Glushko A.A., Morozov S.A., Chistyakov M.G.* Investigation of the sensitive region of the MOSFET to the effects of secondary particles arising from ionizing radiation, *Microelectronics*, 2023, Vol. 52, No. 4, pp. 282-289.
5. *Terekhov V.V.* Vybora modeli approksimatsii dlya issledovaniya elektricheskikh kharakteristik mikrosistem pod vliyaniem tyazhelykh ionov [The choice of an approximation model for studying the electrical characteristics of microsystems under the influence of heavy ions], *Budushchee mashinostroeniya Rossii. Vserossiyskaya konferentsiya molodykh uchennykh i spetsialistov (s mezhdunarodnym uchastiem), 14-ya, Moskva, 21-24 sentyabrya 2021 g.: Sb. dokladov: v 2 t. MGTU im. N.E. Baumana (natsional'nyy issledovatel'skiy un-t), Soyuz mashinostroyteley Rossii* [The future of machine building in Russia. All-Russian Conference of Young Scientists and Specialists (with International participation), 14th, Moscow, September 21-24, 2021: Collection of reports: in 2 vol. Bauman Moscow State Technical University (National Research University), Union of Machine Builders of Russia], 2021, Vol. 2, pp. 247-253.
6. *Novikov L.S., Mileev V.N. et al.* Sinotic Radiation effects on materials in space technology, *Surface. X-ray, synchrotron and neutron studies*, 2009, No. 3, pp. 32-48.
7. *Khaffner Dzh.* Yadernoe izluchenie i zashchita v kosmose [Radiation and shielding in space]. Moscow: Atomizdat, 1971.
8. *Aleksandrov P.A., Zhuk V.I., Litvinov V.L.* Sposoby postroeniya otkazoustoychivyykh tsifrovyykh mikroskhem i otsenki veroyatnosti ikh otkaza, vyzvannogo oblucheniem [Methods of constructing fault-tolerant digital chips and assessing the probability of their failure caused by radiation]. Moscow, 2019.
9. *Shakhnov V.A., Zinchenko L.A., Rezchikova E.V., et al.* Visual Analytics and Its Applications in Electronic Engineering Education : BMSTU Case Study, *2022 6th International Conference on Information Technologies in Engineering Education, Inforino 2022 – Proceedings.* – DOI: 10.1109/Inforino53888.2022.9782907.
10. *Allison J. et al.* Recent Developments in Geant4, *Nucl. Instrum. Meth. A*, 2016, 835, pp. 186-225.
11. *Allison J. et al.* Geant4 Developments and Applications, *IEEE Trans. Nucl. Sci.*, 2006, 53, pp. 270-278.
12. *Christopher M Poole, Iwan Cornelius, Jamie V Trapp, Christian M Langton.* A CAD Interface for GEANT4, *Australasian physical & engineering sciences in medicine*.
13. *Agostinelli S. et al.* Geant4 - A Simulation Toolkit, *Nucl. Instrum. Meth. A*, 2003, 506, pp. 250-303.
14. Geant4 User's Guide for Toolkit Developers: Geant4 Docs. Available at: <https://geant4-userdoc.web.cern.ch/UsersGuides/ForToolkitDeveloper/html/index.html> (accessed 22 October 2023).
15. Geant4 Guide for Physics Lists: Geant4 Docs. Available at: <https://geant4-userdoc.web.cern.ch/UsersGuides/PhysicsListGuide/html/index.html> (accessed 22 October 2023).
16. Geant4 - API. Available at: <https://simoncbyth.github.io/env/notes/geant4/api/> (accessed 22 October 2023).
17. Geant4Py: Geant4-Python interface. Available at: <https://github.com/koichi-murakami/g4python> (accessed 22 October 2023)ю
18. Geant4_pybind: Alternative Python bindings for Geant4. Available at: https://github.com/HaarigerHarald/geant4_pybind (accessed 22 October 2023).
19. Tutorial and examples of Geant4. Available at: <https://dev.asifmoda.com/en/geant4> (accessed 22 October 2023).
20. Sensitive Detector in Geant4: Seminar on Software for Nuclear, Subnuclear and Applied Physics. Available at: https://agenda.infn.it/event/8781/contributions/75434/attachments/54976/64833/Cirrone_SensitiveDetector.pdf (accessed 22 October 2023).

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

Зинченко Людмила Анатольевна – МГТУ им. Н.Э. Баумана; г. Москва, Россия; e-mail: lyudmillaa@mail.ru; тел.: +79057299365; д.т.н.; профессор; профессор кафедры ИУ4.

Казakov Вадим Вячеславович – e-mail: kazakov-vadim-v@yandex.ru; тел.: +79689369731; к.т.н.; старший преподаватель кафедры ИУ4.

Моисеев Роман Романович – e-mail: r.roman.moiseev.r@gmail.com; тел.: +79163710048; студент.

Поляков Евгений Сергеевич – e-mail: ysvidetel@gmail.com; тел.: +79163710048; студент.

Zinchenko Lyudmila Anatolyevna – Bauman Moscow State Technical University; e-mail: lyudmillaa@mail.ru; Moscow, Russia; phone: +79057299365; dr. of eng. sc.; professor.

Kazakov Vadim Vyacheslavovich – e-mail: kazakov-vadim-v@yandex.ru; phone: +79689369731; cand. of eng. sc.; senior lecturer.

Moiseev Roman Romanovich – e-mail: r.roman.moiseev.r@gmail.com; phone: +79163710048; student.

Polyakov Evgeny Sergeevich – e-mail: ysvidetel@gmail.com; phone: +79163710048; student.

УДК 621.371

DOI 10.18522/2311-3103-2024-2-109-121

Д.В. Семенихина, Я.А. Скотаренко

**ЧИСЛЕННЫЕ ИССЛЕДОВАНИЯ НЕЛИНЕЙНОЙ МЕТАПОВЕРХНОСТИ
КЛИНОВИДНОЙ ФОРМЫ ДЛЯ РАЗРАБОТКИ РЕКОМЕНДАЦИЙ
ПО УПРАВЛЕНИЮ ОТРАЖЕННЫМ ЭЛЕКТРОМАГНИТНЫМ ПОЛЕМ КРЫЛА
ЛЕТАТЕЛЬНОГО АППАРАТА**

Целью исследования является оценка эффективности применения нелинейной метаповерхности на объекте клиновидной формы для управления отраженным от него электромагнитным полем. В исследовании нелинейная метаповерхность представляет собой двумерный упорядоченный массив, элементы которого, взаимодействуя с электромагнитным полем падающей волны, способны управлять амплитудой и фазой прошедшей и отраженной волн, позволяя таким образом управлять волновым фронтом. Задачей исследования является поиск и проверка метода анализа клиновидных структур с нелинейными нагрузками различной конфигурации. При решении задачи исследования, рассматривается модель задней кромки крыла летательного аппарата с продольными нелинейными нагрузками на обеих гранях. Исследуемые нелинейные элементы моделируются сосредоточенными нагрузками в зонах на поверхности задней кромки крыла летательного аппарата. Подбором параметров сосредоточенных нагрузок задается вид их нелинейности. Описывается метод численного анализа модели, основанный на решении задачи рассеяния и задачи излучения электромагнитного поля. Представлена пошаговая последовательность действий от создания объемной модели задней кромки крыла летательного аппарата, до задания параметров расчета и получения результатов. Приведены результаты моделирования задней кромки крыла летательного аппарата с различным числом нелинейных нагрузок, различным положением нагрузок относительно друг друга и от кромки, различными параметрами сосредоточенных нагрузок. Полученные результаты электродинамического моделирования позволяют сделать выводы относительно параметров нагрузок, их количества и расположения. Были исследованы различные комбинации нагрузок, а именно две, три и четыре нелинейных нагрузки на каждой грани. Результаты для идеально проводящего клина с двумя нагрузками, показали низкую эффективность в части возможности увеличения уровней кратных гармоник относительно уровня основной в рассеянном от структуры электромагнитном поле. Анализ клина с тремя нагрузками, показал возможность увеличения уровней кратных гармоник на 13 дБ в секторе углов $\pm 5^\circ$. Стабильное, в широкой полосе частот, увеличение уровня кратных гармоник на 13-23 дБ в секторе углов $\pm 80^\circ$ достигается при размещении на задней кромке крыла летательного аппарата четырех нагрузок.

Нелинейная метаповерхность; нелинейные нагрузки; возбуждение клина; задняя кромка крыла.

D.V. Semenikhina, Ya.A. Skotarenko

**NUMERICAL STUDIES OF A NONLINEAR WEDGE-SHAPED METASURFACE
FOR THE DEVELOPMENT OF RECOMMENDATIONS FOR CONTROL
OF THE REFLECTED ELECTROMAGNETIC FIELD OF AN AIRCRAFT WING**

The aim of the study is to evaluate the effectiveness of using a nonlinear metasurface on a wedge-shaped object to control the electromagnetic field reflected from it. In the study, a nonlinear meta-surface is a two-dimensional ordered array, the elements of which, interacting with the electromagnetic field of an

incident wave, are able to control the amplitude and phase of the transmitted and reflected waves, thus allowing the wavefront to be controlled. The task of the research is to find and verify a method for analyzing wedge-shaped structures with nonlinear loads of various configurations. When solving the researching task, a model of the trailing edge of a wing of an aircraft with longitudinal nonlinear loads on both sides is considered. The studied nonlinear elements are modeled by concentrated loads in zones on the surface of the trailing edge of the aircraft wing. By selecting the parameters of concentrated loads, the type of their nonlinearity is set. A method of numerical analysis of the model based on solving the scattering problem and the problem of electromagnetic field radiation is described. A step-by-step sequence of actions is presented from creating a three-dimensional model of the trailing edge of an aircraft wing, to setting calculation parameters and obtaining results. The results of modeling the trailing edge of an aircraft wing with a different number of nonlinear loads, different positions of loads relative to each other and from the edge, and different parameters of concentrated loads are presented. The obtained results of electrodynamic modeling allow us to do conclusions regarding the parameters of loads, their number and location. Various combinations of loads were investigated, notably two, three and four nonlinear loads on each face. The results for an ideally conducting wedge with two loads showed low efficiency in terms of the possibility of increasing the levels of multiple harmonics relative to the main level in an electromagnetic field scattered from the structure. The analysis of a wedge with three loads showed the possibility of increasing the levels of multiple harmonics by 13 dB in the $\pm 5^\circ$ angle sector. Stable, in a wide frequency band, an increase in the level of multiple harmonics by 13-23 dB in the $\pm 80^\circ$ angle sector is achieved when four loads are placed on the trailing edge of the aircraft wing.

Nonlinear metasurface; nonlinear loads; wedge excitation; wing trailing edge.

Введение. В исследовании рассматривается нелинейная метаповерхность (МП). Метаповерхность – это электромагнитная структура, представляющая собой двумерный упорядоченный массив из элементов субволновых размеров, которые, взаимодействуя с электромагнитным полем (ЭМП) падающей волны, способны управлять амплитудой, фазой и поляризацией прошедшей и отраженной волн, осуществляя таким образом полный контроль над волновым фронтом [1]. Метаповерхности обладают такими свойствами, как электромагнитная прозрачность и маскировка [2], поглощение [3, 4] и др.

В исследовании возможность управления ЭМП основывается на эффекте нелинейного рассеяния (ЭНР), который был открыт в середине прошлого века [5]. Согласно исследованиям ЭНР [6] установили, что электромагнитная энергия мощных передатчиков при падении на объекты искусственного происхождения порождает в спектре рассеянного поля комбинационные и гармонические частотные составляющие. Основными естественными источниками ЭНР являются оксиды, например, оксид железа (Fe_2O_3).

Следует отметить, что в большей мере, чем естественные источники, нелинейное рассеяние порождают полупроводниковые элементы, такие как диоды, транзисторы, смесители и т.п. Полупроводниковые элементы отличаются видом нелинейности [7], это обусловлено разной технологией производства.

Многие физические явления в природе описываются линейными математическими моделями, однако протекающие в них процессы, как правило сложнее и должны содержать нелинейные члены [8].

Потребность в построении нелинейных моделей возникает в связи с тем, что процессы, протекающие в определённых условиях, имеют нелинейный характер. Их не получается описать линейными уравнениями, в тоже время они представляют научный интерес, так как возможно позволят найти полезные применения [9, 10].

Целью исследования является разработка метода оценки применения нелинейных нагрузок (НН) на обеих гранях задней кромки крыла летательного аппарата (ЛА), для решения задачи рассеяния электромагнитных волн (ЭМВ) подобными [11–13] структурами.

Обоснование выбора модели задней кромки крыла летательного аппарата. Большой научно-прикладной интерес представляет решение задачи дифракции ЭМВ на объектах сложной формы в так называемых центрах рассеяния («блестящих точках») [14], которые вносят основной вклад в отраженное объектом ЭМП. Частным случаем таких точек у ЛА выступают, например, клиновидные грани задних кромок крыльев (ЗКК), кромок рулей курсовой устойчивости (киля) и т.д. [15].

Для управления ЭМП ЛА в задней полусфере используется метод внедрения нелинейной метаповерхности в области ЗКК. Суть метода заключается в использовании эффекта нелинейного рассеяния [16], для переноса энергии рассеянного объектом ЭМП с основной частоты на частоты гармоник. За счет нелинейных свойств метаповерхности, должен обеспечиваться минимальный уровень сигнала на частоте основной гармоники и максимальный на кратных частотах. Нелинейная МП представляет собой конформный двумерный материал периодической структуры с включениями нелинейных элементов (диоды, варикапы и т.д.) [17]. Конформными называют метаповерхности, которые повторяют геометрию тела, на котором должны быть установлены.

Размеры модели задней кромки крыла ЛА были выбраны следующие [18, 19] (рис. 1):

- ◆ длина торца клина составляет 100 мм;
- ◆ длина клина до сопряжения с цилиндром 200 мм;
- ◆ угол раскрыва клина равен $\pi/6$.

Рабочий диапазон частот выбирается исходя из стандартных диапазонов средств радиолокации. Исходя из этих соображений, был выбран диапазон частот 3-9 ГГц (для основной частоты).

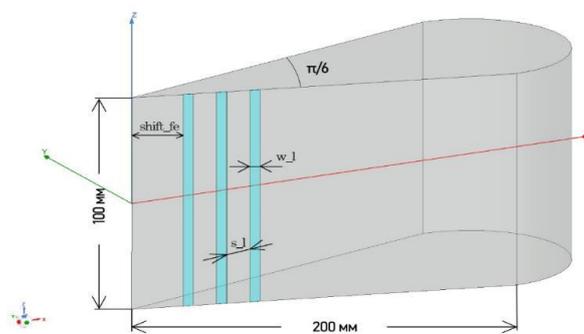


Рис. 1. Исследуемая модель задней кромки крыла ЛА

Постановка задачи. Конструкция задней кромки крыла ЛА. Задачу численного исследования задней кромки крыла ЛА можно разделить на два этапа. Первый этап – это решение задачи рассеяния, второй этап – решение задачи излучения.

На первом этапе выполняется оценка влияния плоской падающей электромагнитной волны на возбуждаемые токи, поиск минимального и максимального значения поля. На втором этапе задаются токи в виде портов возбуждения в области сосредоточенной нагрузки, например, на первой нагрузке, остальные нагрузки при этом задаются как сосредоточенная нагрузка (нагрузка с R, L, C параметрами, где R – сопротивление, L – индуктивность и C – ёмкость).

Была разработана линейная модель ЗКК, показанная на рисунке 1. Модель представляет собой электродинамическую структуру в виде бесконечного идеально проводящего клина длиной $length_w = 200$ мм, шириной $width_w = 100$ мм, углом раскрыва 30° . Задача решалась методом конечных элементов в среде численного электродинамического моделирования Ansys HFSS, которая позволяет решать только линейные задачи, поэтому нелинейные нагрузки будем аппроксимировать нагрузками с сосредоточенными параметрами R, L, C. Сосредоточенная нагрузка представляет собой колебательный RLC – контур с импедансными граничными условиями [20]:

$$E_{tan} = Z_s(\hat{n} \times H_{tan}), \quad (1)$$

где \hat{n} – единичный вектор, перпендикулярный поверхности; E_{tan} – составляющая электрического поля, касательная к поверхности; H_{tan} – составляющая магнитного поля, касательная к поверхности;

$Z_s = R_s + jX_s$ – поверхностный импеданс, где R_s – активное сопротивление в Омах на квадрат, а X_s – реактивное сопротивление.

Параметры сосредоточенной нагрузки подбираются таким образом, чтобы обеспечить максимум и минимум поля на исследуемой частоте.

Нагрузки расположены на *shift_fe* расстоянии от кромки клина до первой нагрузки, а расстояние между нагрузками *s_l*, при этом ширина нагрузок *w_l*. Модель возбуждается плоской падающей волной с частотой ω . Необходимо исследовать характеристики ЭМП заданной структуры в зависимости от параметров нагрузок. Целью исследования является выбор параметров сосредоточенной нагрузки и её положения на клине, обеспечивающих за счет пространственного резонанса наибольшее отношение компонент векторов напряженностей полей в точке наблюдения, определяемых на кратных частотах 2ω , 3ω .

Считаем, что клин является идеально проводящим. Условия периодичности (граничные условия Master-Slave) на боковых гранях границы задачи устанавливаем для указания бесконечности клина (рис. 2,а). На остальной части модели устанавливаем граничные условия излучения на бесконечности (Radiation Boundary). Возбуждение клина происходит плоской падающей волной (Plane Wave), перпендикулярной ребру клина, (рис. 2,б).

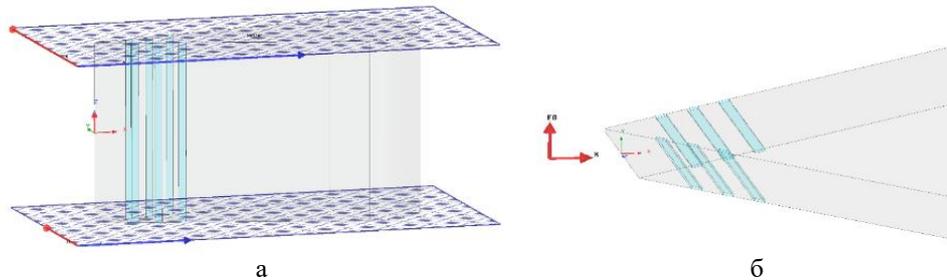


Рис. 2. Исследуемая модель задней кромки крыла ЛА указание: а – граничных условий и б – условий возбуждения плоской падающей волной (Plane Wave)

За источник возбуждения ЭМП на частоте второй гармоники 2ω примем сосредоточенный порт (Lumped Port), который будет моделировать излучение на частотах гармоник. Решая параметрическую задачу, будем искать такие параметры L и C , при которых происходит возбуждение и резонанс. Для предварительной оценки значений L и C выполним расчет резонанса для параллельного колебательного контура (КК), используем известную формулу [21] для нахождения резонансной частоты:

$$\omega_{\text{рез}} = \frac{1}{\sqrt{L \cdot C}}, \quad (2)$$

где L – ёмкость, C – индуктивность КК.

Полученные по формуле 2 значения ёмкости и индуктивности будем задавать в параметрах (L , C) сосредоточенной нагрузки.

Анализ характеристик ЭМП идеально проводящего клина с нагрузками. На первом этапе решается задача рассеяния, структура облучается плоской падающей волной на частоте 3 ГГц. Параметры, описывающие положения нагрузок, варьируются, и выполняется оценка отраженного поля. На рис. 3 и 4 представлены зависимости поля от положения нагрузок на клине и расстоянии между нагрузками.

После нахождения параметров расположения нагрузок, при которых получается минимум поля, выполняем анализ влияния параметров R , L и C сосредоточенных нагрузок. Результаты изменения поля в зависимости от значений ёмкости и индуктивности представлены на рис. 5 и 6.

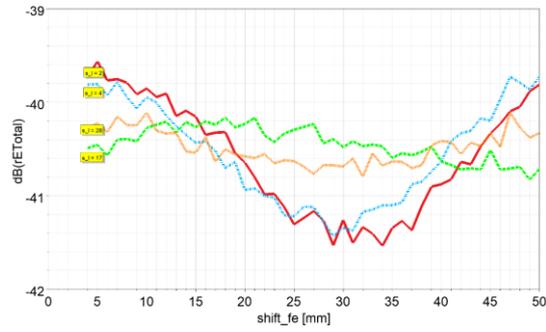


Рис. 3. Полное поле (дБ) в зависимости от положения нагрузок на клине $shift_fe$ (мм) и расстояния между нагрузками s_l (мм)

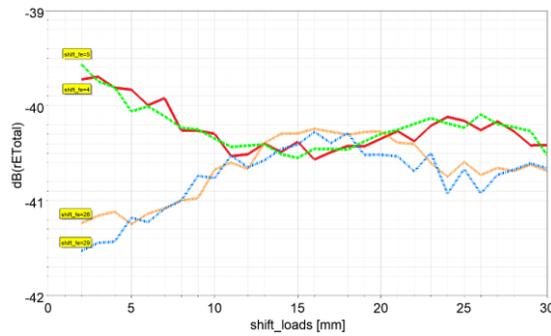


Рис. 4. Полное поле (дБ) в зависимости от расстояния между нагрузками s_l (мм) и положения нагрузок на клине $shift_fe$ (мм)

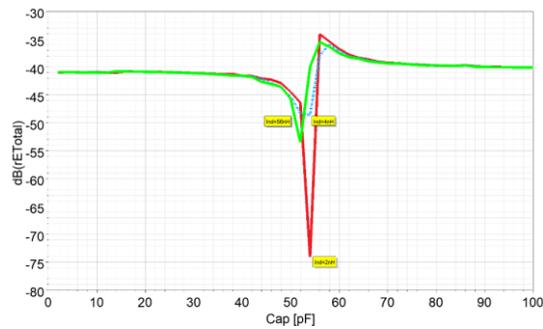


Рис. 5. Полное поле (дБ) в зависимости от ёмкости нагрузки Cap (пФ)

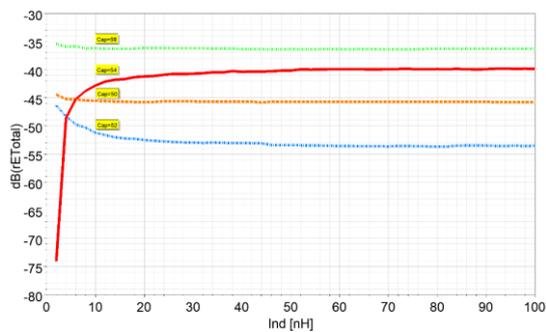


Рис. 6. Полное поле (дБ) в зависимости от индуктивности нагрузки Ind (нГн)

Находим резонансные значения емкости и индуктивности для частоты 3 ГГц: $C_{ap}=54$ пФ, $Ind=2$ нГн, при которых наблюдаем уменьшение поля в обратном падающей волне направлении относительно эталона на 33,5 дБ (рис. 7).

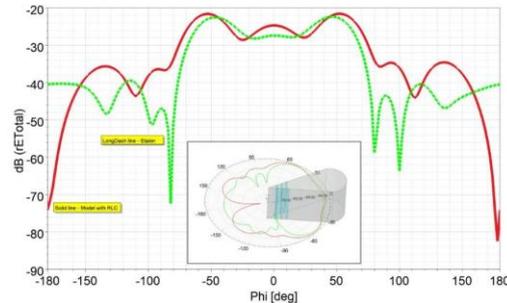


Рис. 7. Полное поле (дБ) в зависимости от параметров нагрузки (C_{ap} , Ind):
красная сплошная линия – клин с нагрузками;
зеленая пунктирная линия – эталон (только клин)

На втором этапе решается задача возбуждения на частоте второй гармоники 2ω (6 ГГц). За основу взята модель из первого этапа. Вместо первой от кромки клина нагрузки, на каждой грани устанавливается порт возбуждения (Lumped Port) (рис. 8).

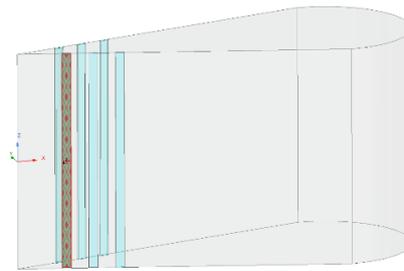


Рис. 8. Модель задачи возбуждения, порт возбуждения выделен красным цветом

Результаты расчета полного поля для частоты 6 ГГц представлены на рис. 9. Наблюдаем увеличение поля на 12 дБ, что подтверждает переход энергии с основной частоты на вторую гармонику. Полученные результаты позволяют оценить эффективность управления характеристиками рассеяния исследуемого объекта для конкретных конфигураций нагрузок.

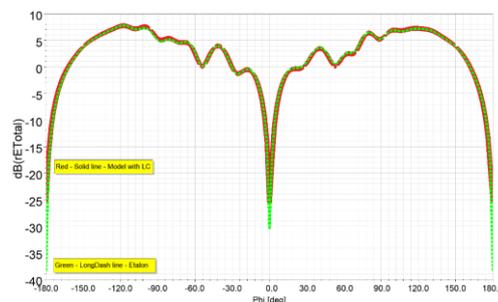


Рис. 9. Полное поле (дБ) от угла Φ на удвоенной частоте (6 ГГц):
красная сплошная линия - клин с нагрузками и портами возбуждения;
зеленая пунктирная линия – эталон (только порты возбуждения, без нагрузок)

Далее повторяем проделанные действия для разного числа нагрузок на клине.

Анализ модели клина с двумя нагрузками. Исследуем модель с двумя нагрузками (рис. 10), решим сначала задачу рассеяния.

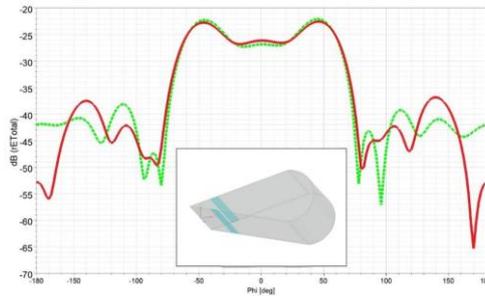


Рис. 10. Полное поле (дБ) в зависимости от параметров нагрузки (C_{ap} , I_{nd}):
красная сплошная линия - клин с нагрузками;
зеленая пунктирная линия – эталон (только клин)

На рис. 10 представлены результаты снижения поля для случая, когда на кромке клина установлено две нагрузки. Разница уровня полного поля между эталоном и моделью с нагрузками составляет 10 дБ.

Затем исследуем задачу возбуждения для модели с двумя нагрузками.

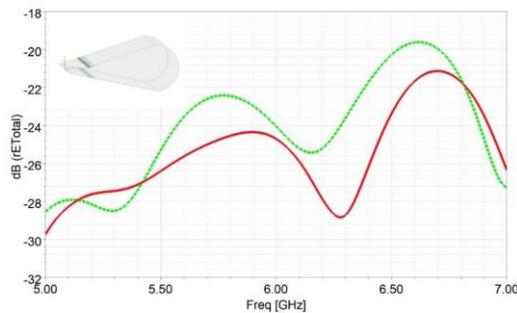


Рис. 11. Полное поле (дБ) в зависимости от частоты анализа для:
красная кривая – клин с нагрузками и источником на первой нагрузке,
зеленая кривая – эталон (только клин с источником на первой нагрузке)

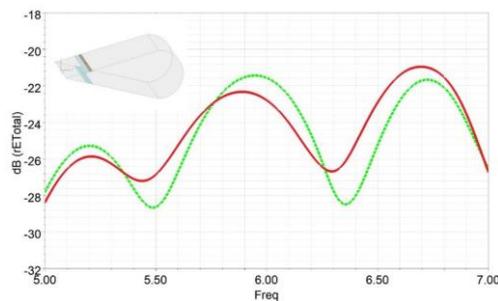


Рис. 12. Полное поле (дБ) в зависимости от частоты анализа для:
красная кривая – клин с нагрузками и источником на второй нагрузке,
зеленая кривая – эталон (только клин с источником на второй нагрузке)

Исследование модели с двумя нагрузками (см. рис. 11, 12), показали невысокую эффективность в части возможности увеличения уровней кратных гармоник относительно уровня основной в рассеянном от структуры ЭМП. В основном вся энергия на кратных частотах плавно распределяется по частотам и не даёт существенного увеличения уровня ЭМП.

Анализ моделей клина с тремя и четырьмя нагрузками. Аналогичным образом исследуем модели с тремя и четырьмя нагрузками.

В случае трех нагрузок (рис. 13) максимального снижения уровня поля на основной частоте удалось добиться при следующих параметрах нагрузки: при $shift_fe=20$ мм, $shift_loads=8$ мм, $Cap=54$ пФ, $Ind=2$ нГн.

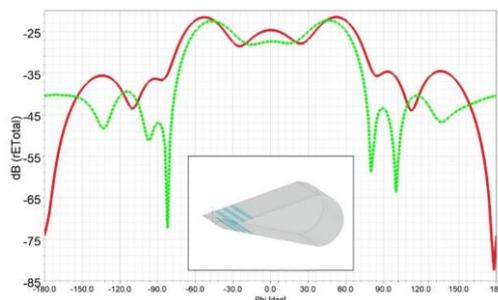


Рис. 13. Полное поле (дБ) от угла Φ на частоте 3 ГГц, при $shift_fe=20$ мм, $shift_loads=8$ мм, $Cap=54$ пФ, $Ind=2$ нГн: красная сплошная линия – клин с нагрузками; зеленая пунктирная линия – эталон (только клин)

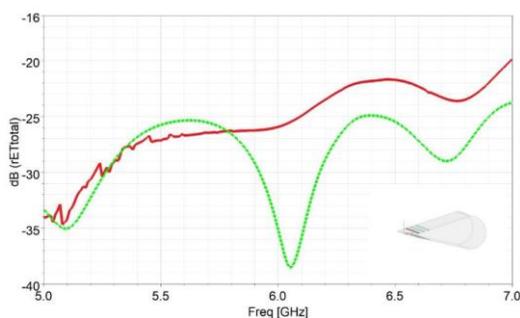


Рис. 14. Полное поле (дБ) в зависимости от частоты анализа для: красная кривая – клин с нагрузками и источником на первой нагрузке, зеленая кривая – эталон (только клин с источником на первой нагрузке)

Исследовались три модели с тремя нагрузками, одна из которых замещалась на источник (см. рис. 14-16). Расчеты показали эффективность в конфигурации, когда источник возбуждения находится на первой нагрузке. Для этого случая максимальное увеличение поля составляет 13 дБ на частоте 6,06 ГГц и 9 дБ на частоте 6 ГГц (см. рис. 14).

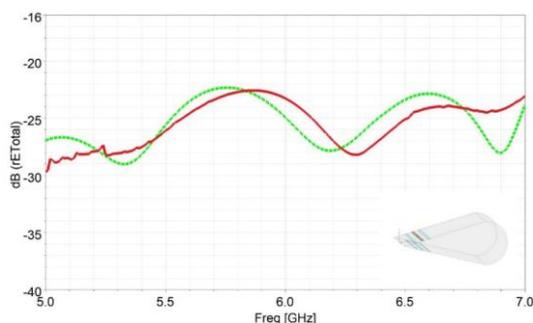


Рис. 15. Полное поле в зависимости от частоты анализа для: красная кривая – клин с нагрузками и источником на второй нагрузке, зеленая кривая – эталон (только клин с источником на второй нагрузке)

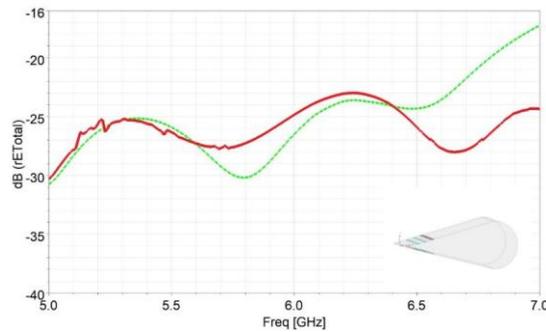


Рис. 16. Полное поле (дБ) в зависимости от частоты анализа для: красная кривая – клин с нагрузками и источником на третьей нагрузке, зеленая кривая – эталон (только клин с источником на третьей нагрузке)

Анализ модели с четырьмя нагрузками показал возможность снижения уровня поля на основной частоте на 52 дБ (рис. 17). Этого удалось добиться при следующих параметрах нагрузки: при $shift_fe=27$ мм, $shift_loads=2$ мм, $Cap=48$ пФ, $Ind=8$ нГн.

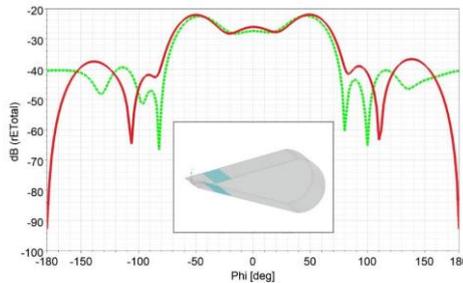


Рис. 17. Полное поле (дБ) в зависимости от угла Φ : для четырёх нагрузок на частоте 3 ГГц для $shift_fe=27$ мм, $shift_loads=2$ мм: красная сплошная линия – клин с нагрузками; зеленая пунктирная линия – эталон (только клин)

Исследование моделей с четырьмя нагрузками показали эффективность в конфигурации, когда источник возбуждения находится на второй нагрузке, как видно из сравнения результатов решения задач возбуждения (рис. 18-21). В данной конфигурации наблюдается увеличение уровня поля на кратных гармониках в широкой полосе частот.

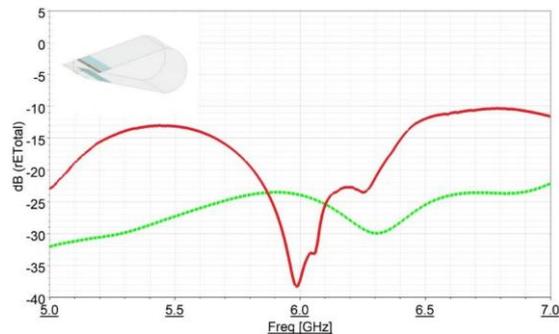


Рис. 18. Полное поле (дБ) в зависимости от частоты анализа для: красная кривая – клин с нагрузками и источником на первой нагрузке, зеленая кривая – эталон (только клин с источником на первой нагрузке)

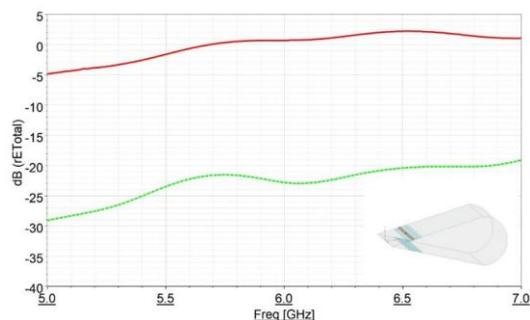


Рис. 19. Полное поле (дБ) в зависимости от частоты анализа для: красная кривая – клин с нагрузками и источником на второй нагрузке, зеленая кривая – эталон (только клин с источником на второй нагрузке)

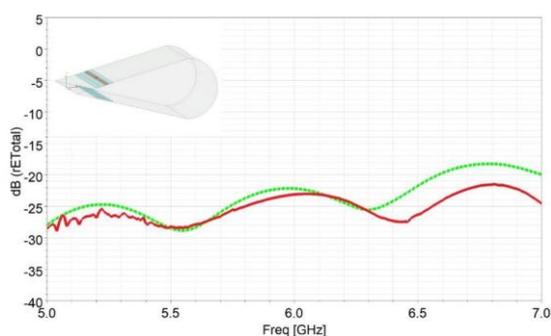


Рис. 20. Полное поле (дБ) в зависимости от частоты анализа для: красная кривая – клин с нагрузками и источником на третьей нагрузке, зеленая кривая – эталон (только клин с источником на третьей нагрузке)

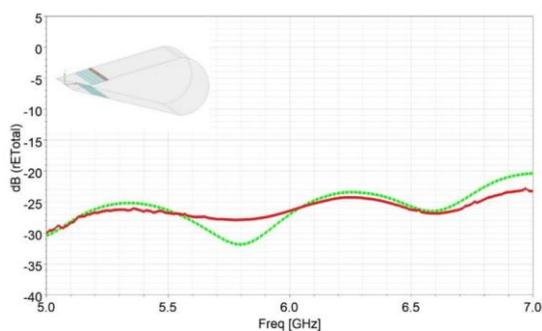


Рис. 21. Полное поле (дБ) в зависимости от частоты анализа для: красная кривая – клин с нагрузками и источником на четвертой нагрузке, зеленая кривая – эталон (только клин с источником на четвертой нагрузке)

Выводы:

- ◆ в ходе электродинамического моделирования задачи найдены параметры нагрузок, при которых возникает пространственный резонанс;
- ◆ определены расположение нагрузок на клине и их количество, при которых обеспечивается увеличение относительного уровня гармонических составляющих;
- ◆ численные исследования идеально проводящего клина с двумя нагрузками, показали низкую эффективность в части возможности увеличения уровней кратных гармоник относительно уровня основной в рассеянном от структуры ЭМП. Так для задачи рас-

сеяния удалось добиться снижения уровня полного поля на основной частоте (3 ГГц), по сравнению с эталоном, в секторе углов $\pm 20^\circ$ всего на 10 дБ. Для задачи возбуждения на удвоенной частоте (6 ГГц), когда источник расположен на первой от кромки клина нагрузке зафиксировано снижение уровня полного поля на 2 дБ, а когда источник расположен на второй нагрузке уровень полного поля относительно эталона не изменился;

◆ численные исследования идеально проводящего клина с тремя нагрузками, показали возможность увеличения уровней кратных гармоник на 13 дБ в секторе углов $\pm 5^\circ$ относительно уровня основной в рассеянном от структуры ЭМП, при условии, что источник возбуждения расположен на первой нагрузке от кромки клина. Увеличение относительного уровня обеспечивается путём снижения уровня рассеяния на основной гармонике;

◆ численные исследования идеально проводящего клина с четырьмя нагрузками, показали возможность увеличения уровней кратных гармоник от 13 дБ до 23 дБ в секторе углов $\pm 80^\circ$ относительно уровня основной гармоники в рассеянном от структуры ЭМП, при условии, что источник возбуждения расположен на второй нагрузке от кромки клина. Увеличение относительного уровня обеспечивается путём снижения уровня рассеяния на основной гармонике;

◆ численный анализ модели клина с разным количеством нагрузок показал, что оптимальным вариантом, в исследуемом диапазоне частот, является использование четырех нагрузок, т.к. при этом обеспечивается стабильное увеличение уровня кратных гармоник на 13–23 дБ в широкой полосе частот.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Максимова А.Г., Одит М.А. Широкополосная поглощающая метаповерхность на основе воды. – СПб.: Санкт-Петербургский государственный электротехнический университет «ЛЭТИ».
2. Bergmair I., [at al.]. Single and multiplayer metamaterials fabricated by nanoimprint lithography // *Nanotechnology*. – 2011. – 22. – P. 325301.
3. Watts C.M., [at al.]. Metamaterial, Electromagnetic Wave Absorbers // *Advanced materials*. – 2012. – Vol. 24.
4. Wen Q.-Y., [at al.]. Perfect Metamaterial Absorbers in Microwave and Terahertz Bands // *Metamaterial*. – 2012.
5. Eastman A., Horle L. The Generation of Spurious Signal by Nonlinearity of the Transmission Path // *Proc. IRE*. – 1940. – Vol. 28. – P. 438.
6. Штейншленгер В.Б., Мисежников Г.С. Исследование эффекта нелинейного рассеяния радиоволн металлическими объектами // *Радиотехника и электроника*. – 1984. – Т. 39. – Вып. 6. – С. 902-906.
7. Величко Д.В., Рубанов В.Г. Полупроводниковые приборы и устройства: учеб. пособие. – Белгород: Политекра, 2006. – 184 с.
8. Oleksandr Malysukin and Vincent Fusco. Combined Frequency-Time Domain Analysis of FSS with Nonlinear Loads Using Volterra Integral Equations, ECIT Institute, Queens University of Belfast Queens Rd, Queens Island, Belfast BT3 9DT, N. Ireland.
9. Liao D. Scattering and Imaging of Nonlinearly Loaded Antenna Structures in Half-Space Environments // *IEEE Transactions on Antennas and Propagation*. – 2014. – 62 (8). – P. 4230-4240. – DOI: 10.1109/tap.2014.2322875.
10. Chien-Chang Huang, & Tah-Hsiung Chu. Analysis of wire scatterers with nonlinear or time-harmonic loads in the frequency domain // *IEEE Transactions on Antennas and Propagation*. – 1993. – 41 (1). – P. 25-30. – DOI: 10.1109/8.210111.
11. Zhu Z.-W.*, Chen H.-Y., Zhang H.-B., Zhou P.-H., Deng L.-J., and Xie J.-L. Patterned resistive strip loading for edge scattering suppression of a finite wedge // *Progress in Electromagnetics Research M*. – 2012. – Vol. 25. – P. 27-38.
12. Jacques Periaux, Mourad Sefrioui, Bertrand Mantel. RCS multi-objective optimization of scattered waves by active control elements using Gas, Dassault Aviation. 78 Quai Marcel Dassault, 92214 SaintCloud, France.
13. Knott E.F. Suppression of edge scattering with impedance strings // *IEEE Transactions on Antennas and Propagation*. – 1997. – 45 (12). – P. 1768-1773. – DOI: 10.1109/8.650194.
14. Семенихина Д.В., Юханов Ю.В., Привалова Т.Ю. Теоретические основы радиоэлектронной борьбы: учеб. пособие. – Таганрог, 2009. – 266 с.
15. Геворкян А.В. Гашение полей рассеяния объектов системой антенн Вивальди: дис. ... канд. техн. наук: 05.12.07. – Таганрог, 2016. – 226 с.

16. Семенихина Д.В. Анализ эффекта нелинейного рассеяния электромагнитных волн // Электродинамика и техника СВЧ и КВЧ. – 1998. – Т. 6, № 1-2 (21). – С. 119-126.
17. Sievenpiper D.F. Nonlinear Grounded Metasurfaces for Suppression of High-Power Pulsed RF Currents // IEEE Antennas and Wireless Propagation Letters. – 2011. – 10. – P. 1516-1519. – DOI: 10.1109/lawp.2011.2182593.
18. Zhu Z-W, Chen H-Y, Zhang H-B, Zhou P-H, Deng L-J, Xie J-L. Patterned resistive strip loading for edge scattering suppression of a finite wedge // Prog. Electromagn. Res. M. – 2012. – 25. – P. 27-38.
19. Цихош Э. Сверхзвуковые самолеты: Справочное руководство: пер. с польск. – М.: Мир, 1983. – 432 с.
20. Марков Г.Т., Чаплин А.Ф. Возбуждение электромагнитных волн. – М.: Изд-во Радио и связь, 1983. – 296 с.
21. Баскаков С.И. Радиотехнические цепи и сигналы. – М.: Изд-во «Высшая школа», 2000. – 459 с.

REFERENCES

1. Maksimova A.G., Odit M.A. Shirokopolosnaya pogloshchayushchaya metapoverkhnost' na osnove vody [Broadband observable metasurface based on water]. St. Petersburg: Sankt-Peterburgskiy gosudarstvennyy elektrotekhnicheskii universitet «LETI».
2. Bergmair I., [at al.]. Single and multiplayer metamaterials fabricated by nanoimprint lithography, *Nanotechnology*, 2011, 22, pp. 325301.
3. Watts C.M., [at al.]. Metamaterial, Electromagnetic Wave Absorbers, *Advanced materials*, 2012, Vol. 24.
4. Wen Q.-Y., [at al.]. Perfect Metamaterial Absorbers in Microwave and Terahertz Bands, *Metamaterial*, 2012.
5. Eastman A., Horle L. The Generation of Spurious Signal by Nonlinearity of the Transmission Path, *Proc. IRE*, 1940, Vol. 28, pp. 438.
6. Shteynshlenger V.B., Mizezhnikov G.S. Issledovanie effekta nelineynogo rasseyaniya radiovoln metallicheskimi ob'ektami [Study of the effect of nonlinear scattering of radio waves by metal objects], *Radiotekhnika i elektronika* [Radio engineering and electronics], 1984, Vol. 39, Issue 6, pp. 902-906.
7. Velichko D.V., Rubanov V.G. Poluprovodnikovye pribory i ustroystva: ucheb. posobie [Semiconductor devices and devices: textbook]. Belgorod: Politerra, 2006, 184 p.
8. Oleksandr Malyskin and Vincent Fusco. Combined Frequency-Time Domain Analysis of FSS with Nonlinear Loads Using Volterra Integral Equations, ECIT Institute, Queens University of Belfast Queens Rd, Queens Island, Belfast BT3 9DT, N. Ireland.
9. Liao D. Scattering and Imaging of Nonlinearly Loaded Antenna Structures in Half-Space Environments, *IEEE Transactions on Antennas and Propagation*, 2014, 62 (8), pp. 4230-4240. DOI: 10.1109/tap.2014.2322875.
10. Chien-Chang Huang, & Tah-Hsiung Chu. Analysis of wire scatterers with nonlinear or time-harmonic loads in the frequency domain, *IEEE Transactions on Antennas and Propagation*, 1993, 41 (1), pp. 25-30. DOI: 10.1109/8.210111.
11. Zhu Z.-W.*, Chen H.-Y., Zhang H.-B., Zhou P.-H., Deng L.-J., and Xie J.-L. Patterned resistive strip loading for edge scattering suppression of a finite wedge, *Progress in Electromagnetics Research M*, 2012, Vol. 25, pp. 27-38.
12. Jacques Periaux, Mourad Sefrioui, Bertrand Mantel. RCS multi-objective optimization of scattered waves by active control elements using Gas, Dassault Aviation. 78 Quai Marcel Dassault, 92214 SaintCloud, France.
13. Knott E.F. Suppression of edge scattering with impedance strings, *IEEE Transactions on Antennas and Propagation*, 1997, 45 (12), pp. 1768-1773. DOI: 10.1109/8.650194.
14. Semikhina D.V., Yukhanov Yu.V., Privalova T.Yu. Teoreticheskie osnovy radioelektronnoy bor'by: ucheb. posobie [Theoretical foundations of electronic warfare. learning materials]. Taganrog, 2009, 266 p.
15. Gevorkyan A.V. Gashenie poley rasseyaniya ob'ektov sistemoy antenn Vival'di: dis. ... kand. tekhn. nauk [Damping of scattering fields of objects by Vivaldi antenna system: cand. of eng. sc. diss.]: 05.12.07. Taganrog, 2016, 226 p.
16. Semikhina D.V. Analiz effekta nelineynogo rasseyaniya elektromagnitnykh voln [Analysis of the effect of non-linear scattering of electromagnetic waves], *Elektrodinamika i tekhnika SVCh i KVCh* [Electrodynamics and technology VHF and EHF], 1998, Vol. 6, No. 1-2 (21), pp. 119-126.
17. Sievenpiper D.F. Nonlinear Grounded Metasurfaces for Suppression of High-Power Pulsed RF Currents, *IEEE Antennas and Wireless Propagation Letters*, 2011, 10, pp. 1516-1519. DOI: 10.1109/lawp.2011.2182593.

18. Zhu Z-W, Chen H-Y, Zhang H-B, Zhou P-H, Deng L-J, Xie J-L. Patterned resistive strip loading for edge scattering suppression of a finite wedge, *Prog. Electromagn. Res. M*, 2012, 25, pp. 27-38.
19. Tsikhosh E. Sverkhzvukovye samolety: Spravochnoe rukovodstvo [Supersonic Aircraft: Reference Manual]: transl. from polish. Moscow: Mir, 1983, 432 p.
20. Markov G.T., Chaplin A.F. Vozbuzhdenie elektromagnitnykh voln [Excitation of electromagnetic waves]. Moscow: Izd-vo Radio i svyaz', 1983, 296 p.
21. Baskakov S.I. Radiotekhnicheskie tsepi i signaly [Radio engineering circuits and signals]. Moscow: Izd-vo «Vysshaya shkola», 2000, 459 p.

Статью рекомендовала к опубликованию к.т.н. Н.Н. Кисель.

Семенихина Диана Викторовна – Южный федеральный университет; e-mail: d_semenikhina@mail.ru; г. Таганрог, Россия; тел.: +78634371634; д.т.н.; профессор.

Скотаренко Ярослав Александрович – e-mail: yaroslav.cko@mail.ru; тел.: +79044460124; аспирант.

Semenikhina Diana Viktorovna – Southern Federal University; e-mail: d_semenikhina@mail.ru; Taganrog, Russia; phone: +78634371733; dr. of eng. sc.; professor.

Skotarenko Yaroslav Aleksandrovich – e-mail: yaroslav.cko@mail.ru; phone: +79044460124; post-graduate student.

УДК 004.942, 654.924.56

DOI 10.18522/2311-3103-2024-2-121-132

С. Сингх, А.В. Прибыльский

СИНТЕЗ СИСТЕМЫ СВЕРХБЫСТРОГО ОБНАРУЖЕНИЯ ПОЖАРООПАСНЫХ СИТУАЦИЙ НА ОСНОВЕ КОМПЛЕКСА ВЗАИМОСВЯЗАННЫХ ДАТЧИКОВ

Современные технологии и городская инфраструктура требуют инновационных подходов к обнаружению пожароопасных ситуаций. Эффективное и сверхбыстрое обнаружение возгораний становится неотъемлемой частью обеспечения безопасности. С этой целью синтезируются и реализуются системы способные обнаруживать и информировать об пожароопасной ситуации за считанные секунды, в статье синтезируется одна из таких систем. Исследование и синтез математической модели цифрового универсального пожарного датчика, который в свою очередь является комплексом взаимосвязанных датчиков, актуально в связи с постоянным развитием инфраструктуры систем, возрастающей сложностью электрооборудования и необходимость сокращению ущерба, возникающего при возникновении и распространении пожаров. Предиктивная диагностика работоспособности электрооборудования, позволяет своевременно выявлять и устранять потенциальные угрозы пожарной безопасности. В рамках данного исследования приводится теоретическая математическая модель реального цифрового универсального пожарного датчика, сперва в упрощенном варианте, затем в усложненном с учетом конструкции и статистического подхода к задаче нахождения порогов срабатывания датчика, приведено описание параметров математической модели и последовательного принципа работы. Данный датчик представляет собой инновационное решение в области пожарной безопасности, которое обеспечивает высокий уровень контроля и эффективности в реальном времени. На основе теоретических моделей, представленных в статье, разработана математическая модель датчика, которая смоделирована с использованием программного средства Simulink на реальных данных, полученных от производителя датчика. Результаты моделирования показали, что модель корректно описывает поведение реального датчика на всех каналах и может быть использована в дальнейших исследованиях, таких как прогнозирование и обнаружения пожароопасных ситуаций с использованием нейронных сетей. Синтез предложенной системы необходим для дальнейших исследований в области прогнозирования и обнаружения пожароопасных ситуаций на основе полученной математического модели.

Пожарный датчик; simulink; сигнал детекции; стандартное отклонение; функция активации; математическое моделирование.

S. Singh, A.V. Pribylskiy

SYNTHESIS OF A SYSTEM FOR ULTRA-FAST DETECTION OF FIRE- HAZARDOUS SITUATIONS BASED ON A COMPLEX OF INTERCONNECTED SENSORS

Modern technologies and urban infrastructure require innovative approaches to detecting fire hazards. Effective and ultra-fast fire detection is becoming an integral part of safety. For this purpose, systems capable of detecting and informing about a fire hazard situation in a matter of seconds are synthesized and implemented; one of such systems is synthesized in the article. The research and synthesis of a mathematical model of a digital universal fire sensor, which in turn is a complex of interconnected sensors, is relevant due to the constant development of system infrastructure, the increasing complexity of electrical equipment and the need to reduce damage arising from the outbreak and spread of fires. Predictive diagnostics of electrical equipment performance allows timely identification and elimination of potential fire safety threats. Within the framework of this research, a theoretical mathematical model of a real digital universal fire sensor is presented, first in a simplified version, then in a more complicated version, taking into account the design and statistical approach to the problem of finding the sensor response thresholds, a description of the parameters of the mathematical model and the sequential principle of operation is given. This sensor is an innovative fire safety solution that provides a high level of control and efficiency in real time. Based on the theoretical models presented in the article, a mathematical model of the sensor has been developed, which is simulated using the Simulink software tool on real data obtained from the sensor manufacturer. The simulation results showed that the model correctly describes the behavior of a real sensor on all channels and can be used in further research, such as predicting and detecting fire situations using neural networks. The synthesis of the proposed system is necessary for further research in the field of forecasting and detection of fire hazardous situations based on the obtained mathematical model.

Fire sensor; simulink; detection signal; standard deviation; activation function; mathematical modeling.

Введение. В условиях постоянного технологического развития и растущей сложности инфраструктурных систем вопросы пожарной безопасности приобретают особую актуальность. В этом контексте внедрение инновационных решений, способных обеспечивать высокий уровень контроля и эффективности в реальном времени, становится приоритетной задачей. Одним из таких ключевых элементов становится универсальный цифровой пожарный датчик, обладающий рядом выдающихся особенностей, спроектированных для обеспечения безопасности в самых сложных условиях.

Цифровой универсальный пожарный датчик – это комплекс взаимосвязанных датчиков, выполняющих функцию обнаружения пожароопасных ситуаций. Этот комплекс датчиков включает в себя датчики температуры, датчика дыма, измеряющих задымленность в ультрафиолетовом и инфракрасных спектрах. Подробнее его структура и конструкция рассмотрена в статьях [1-6].

Важным аспектом функциональности универсального пожарного датчика является его способность контролировать работоспособность пожарно-технического оборудования [1–8] в режиме реального времени. Эта характеристика открывает новые возможности для оперативного реагирования на пожарные угрозы, особенно в условиях движения поездов метро под землей, где эффективность системы безопасности является фактором первостепенной важности.

Максимальная эффективность противопожарной защиты электрооборудования становится важным аспектом в современных комплексах инфраструктурных систем. Цифровой универсальный пожарный датчик предлагает интегрированный подход к обеспечению безопасности, гарантируя не только работоспособность всех средств первичного пожаротушения, но и предоставляя цифровую историю жизненного цикла основных элементов пожарно-технического оборудования и текущего состояния охраняемых зон.

Кроме того, предиктивная диагностика работоспособности электрооборудования сторонних производителей в охраняемых зонах является важным моментом в предотвращении пожароопасных ситуаций. Пожарный датчик выходит за пределы традиционных методов и предоставляет интеллектуальные решения для предупреждения возможных неполадок, способных привести к пожарам.

Цель исследования заключается в представлении математической модели датчика, тем самым открывая новые перспективы для применения в различных областях, а также демонстрации практического потенциала для дальнейших исследований в области сенсорных технологий.

Математическая модель цифрового универсального датчика. Конструктивно датчик состоит из 14 элементов [1], 8 из которых датчики температуры, 3 ультрафиолетовых датчика дыма, 3 инфракрасных датчика дыма исполненных в металлическом корпусе. Упомянутые датчики контролируется и настраивается дистанционно в режиме реального времени в том числе чувствительность и пороговые значения. Датчик выполнен таким образом, чтобы воспринимать дым вокруг себя со всех возможных направлений [1, 2], что позволяет обнаруживать дым значительно раньше стандартных пожарных датчиков по изменению плотности дыма в разных точках пространства [3].

Предположим, что у нас есть функции чувствительности и пороговых значений для каждого датчика [9,10]. Обозначим чувствительность датчика как S_i и пороговое значение как T_i , где i – индекс датчика.

Тогда математическая модель ультрафиолетовых датчиков B_1, B_2, B_3 :

$$D_U = \frac{1}{3} \sum_{i=1}^3 \max \left(0, \frac{B_i}{S_{U_i}} - T_{U_i} \right), \quad (1)$$

где S_{U_i} – чувствительность инфракрасного датчика i ; T_{U_i} – пороговое значение ультрафиолетового датчика i ; $\max(0, x)$ – функция активации ReLU.

Математическая модель инфракрасных датчиков R_1, R_2, R_3 :

$$D_I = \frac{1}{3} \sum_{i=1}^3 \max \left(0, \frac{R_i}{S_{I_i}} - T_{I_i} \right), \quad (2)$$

где S_{I_i} – чувствительность инфракрасного датчика i ; T_{I_i} – пороговое значение инфракрасного датчика i

Математическая модель температурных датчиков R_1, R_2, R_3

$$D_T = \frac{1}{8} \sum_{i=1}^8 \max \left(0, \frac{T_i}{S_{T_i}} - T_{T_i} \right), \quad (3)$$

где S_{T_i} – чувствительность температурного датчика i ; T_{T_i} – пороговое значение температурного датчика i

Общий сигнал детекции D [6, 7]:

$$D = \max(D_U, D_I, D_T). \quad (4)$$

Модель учитывает чувствительность и пороговые значения для каждого датчика и суммирует их с учетом каждого их типа. Модель также использует функцию активации ReLU [11–13] для учета только положительных значений, т. к. негативные значения не имеют физического смысла в данном контексте.

Эта модель предполагает, что датчик обнаруживает дым, если хотя бы один из ультрафиолетовых, инфракрасных или температурных датчиков выдаёт сигнал. Максимизация используется для учета всех трех типов датчиков.

Конструктивно датчик выполнен в металлическом корпусе [1,14], на сегодняшний день основная часть этих датчиков устанавливается в вагоны метрополитена. Работа датчика в метрополитене может подвергнуться различным механическим воздействиям и условиям окружающей среды [15], которые могут повлиять на его работоспособность:

- ◆ Механическая прочность: В вагоне метро могут возникнуть вибрации, удары и другие механические воздействия в результате движения по рельсам, торможения и стоянок. Это может потенциально повлиять на механическую целостность датчика.

- ◆ Устойчивость: Устойчивость датчика к вибрациям и переменам температуры также может быть важной характеристикой, особенно если вагон метро подвергается значительным температурным колебаниям.

- ◆ Экранирование: Экранирование может быть важным для предотвращения воздействия электромагнитных полей, которые могут присутствовать в вагоне метро.

Из всего это следует вывод, что математической модели стоит учесть и коэффициенты механической прочности $K_{\text{мех}}$, устойчивости $K_{\text{уст}}$ и экранирования $K_{\text{экр}}$.

Пороговое значение датчиков устанавливается с помощью стандартных статистических методов. Одним из простых, но эффективных методов установки порогового значения является метод двойного стандартного отклонения от среднего, общий вид которого имеет вид:

Пусть X – случайная величина, а σ – стандартное отклонение. Тогда двойное стандартное отклонение от среднего σ_d будет выглядеть как:

$$\sigma_d = 2\sigma. \quad (5)$$

Если среднее значение обозначено как μ , то формула двойного стандартного отклонения от среднего:

$$\sigma_d = 2\sqrt{\text{Var}(X)}, \quad (6)$$

где $\text{Var}(X)$ – дисперсия случайной величины X .

С учетом новой пороговой функции и коэффициентов механической прочности, устойчивости и экранирования математическая модель универсального датчика примет вид:

Математическая модель ультрафиолетовых датчиков B_1, B_2, B_3 :

$$D_{U_n} = \frac{1}{3} \sum_{i=1}^3 K_{\text{мех}} * K_{\text{уст}} * K_{\text{экр}} * \max\left(0, \frac{X_i - M_{U_i}}{SD_{U_i}} - 2SD_{U_i}\right), \quad (7)$$

где SD_{U_i} – стандартное отклонение сигнала ультрафиолетового датчика i ; X_i – значение сигнала от датчика i ; M_{U_i} – среднее значение сигнала от ультрафиолетового датчика; $\max(0, x)$ – функция активации ReLU.

Математическая модель инфракрасных датчиков R_1, R_2, R_3 :

$$D_{I_n} = \frac{1}{3} \sum_{i=1}^3 K_{\text{мех}} * K_{\text{уст}} * K_{\text{экр}} * \max\left(0, \frac{X_i - M_{I_i}}{SD_{I_i}} - 2SD_{I_i}\right), \quad (8)$$

где SD_{I_i} – стандартное отклонение сигнала инфракрасного датчика i ; X_i – значение сигнала от датчика i ; M_{I_i} – среднее значение сигнала от ультрафиолетового датчика.

Математическая модель температурных датчиков T_1, T_2, T_3 :

$$D_{T_n} = \frac{1}{8} \sum_{i=1}^8 K_{\text{мех}} * K_{\text{уст}} * K_{\text{экр}} * \max\left(0, \frac{X_i - M_{T_i}}{SD_{T_i}} - 2SD_{T_i}\right), \quad (9)$$

где SD_{T_i} – стандартное отклонение сигнала температурного датчика i ; X_i – значение сигнала от датчика i ; M_{T_i} – среднее значение сигнала от температурного датчика

Новый общий сигнал детекции D_n :

$$D_n = \max(D_{U_n}, D_{I_n}, D_{T_n}) \quad (10)$$

Опишем работы математической модели универсального датчика с помощью последовательного описания.

1. Сбор данных – на это этапе ультрафиолетовые, инфракрасные, температурные датчики собирают данные о сигналах в реальном времени.

2. Вычисление отклонений – для каждого датчика вычисляются отклонения сигналов от средних значений ($M_{U_i}, M_{I_i}, M_{T_i}$) в единицах стандартного отклонения ($SD_{U_i}, SD_{I_i}, SD_{T_i}$).

3. Активация датчиков – для каждого датчика применяется функция активация, которая моделирует активацию датчика при превышении порогового значения.

4. Механические коэффициенты – являются статическими параметрами, считаются неизменными во времени.

5. Общий сигнал детекции – определяется как максимум взвешенных сигналов от всех датчиков, что является обобщенной оценкой обнаружения возможной пожароопасной ситуации.

6. Решение об активации датчика – активация датчика зависит от вложенных алгоритмов работы системы, но как правило система считается активированной, когда D_n превышает заданный порог, что будет служить сигналом обнаружения пожароопасной ситуации.

7. Динамическая регулировка параметров: динамические параметры модели, такие как среднее значений, стандартного отклонение, пороговые значения, являются регулируемыми и настраиваемыми в режиме реального времени для оптимизации работы системы в различных условиях окружающей среды.

Данная математическая модель основана на статистическом подходе представления системы.

Математическая модель в Simulink. Математическая модель [16–19] ультрафиолетового датчика представлена на рис. 1:

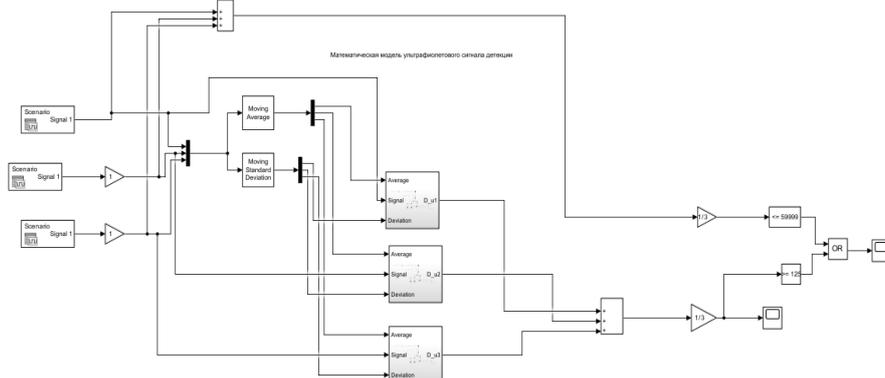


Рис. 1. Математическая модель ультрафиолетового сигнала детекции

Срабатывания датчика происходит, если среднее значения входных данных меньше 59999 [1, 2, 20] или сигнал детекции больше 125, последнее следует из эмпирического порога, т.к. считается что диапазон 60125-60000, не является пожаром [3, 21].

На рис. 1 имеется блок изображенный на рис. 2

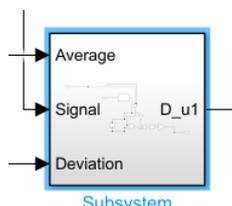


Рис. 2. Подсистема расчета сигнала с датчика

Её развернутый вид представлен на рис. 3:

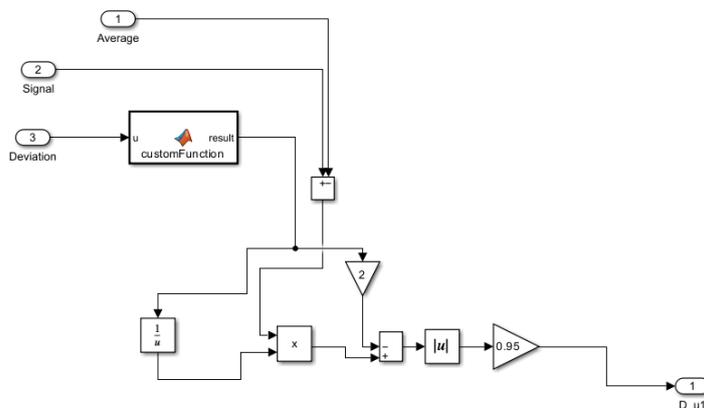


Рис. 3. Развернутая подсистема, изображенная на рис. 2

Математическая модель инфракрасного датчика совпадает с моделью ультрафиолетового датчика.

Математическая модель температурных датчиков представлена на рис. 4.

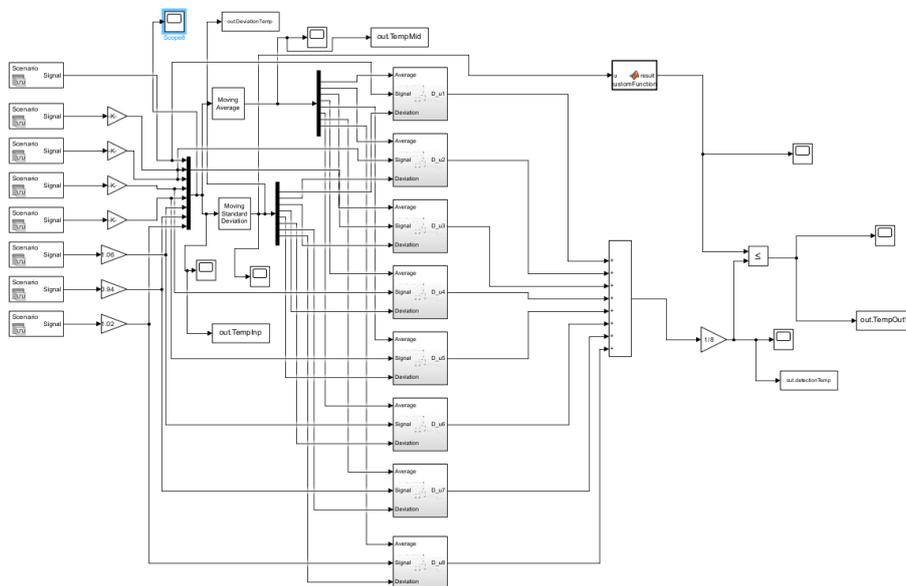


Рис. 4. Математическая модель сигнала детекции температурных датчиков

Моделирование математической модели. Про моделируем полученные модели, для этого используем реальные данные, измеренные с помощью цифрового пожарного датчика, они представляют из себя отметку во времени и соответствующее ему значение с датчиков температуры и дыма. Всего используется 37764 отметки, что соответствует примерно данным за 10,5 часов. Пример данных представлен на рис. 5.

TIME	DATA	TIME	DATA
0	23.49	0	60121
1	23.49	1	60121
2	23.49	2	60121
3	23.49	3	60121
4	23.47	4	60122
5	23.47	5	60122
6	23.47	6	60122

Рис. 5. Пример входных данных

Температура датчиками измеряется в градусах Цельсия, пороговое значение рассчитывается с помощью двойного стандартного отклонения от среднего рассмотренного ранее, соответственно на основании рассчитанного отклонения выставляется пороговое значение температуры, при выходе за пределы порога срабатывает сигнал о пожароопасной ситуации.

Задымленность датчиками дыма измеряется в пределах от 0 до 60125 [1], кроме порогового значения, также учитывается эмпирически найденный минимальный порог отсутствия пожароопасных ситуаций, который равен 59999, значения меньше этого порога являются сигналом о наличии дыма.

Выполним моделирование участка с температурными датчиками.

Инициализируем для каждого из температурных датчиков входной сигнал, рис. 6.

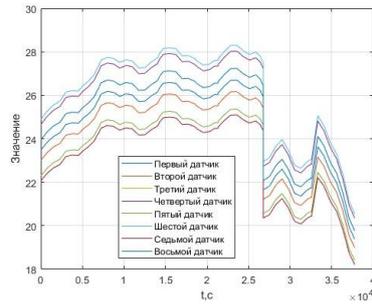


Рис. 6. Входные данные из температурных датчиков

Стандартное отклонение температурных датчиков представлено на рис. 7.

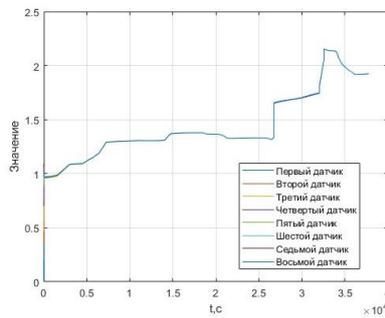


Рис. 7. График изменения стандартного отклонения датчиков температуры

Среднее значение температурного датчика рис. 8

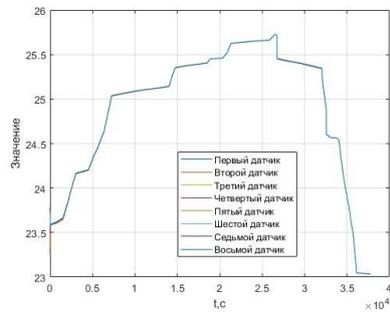


Рис. 8. График изменений среднего значения температуры

Общий сигнал детекции изображен на рис. 9.

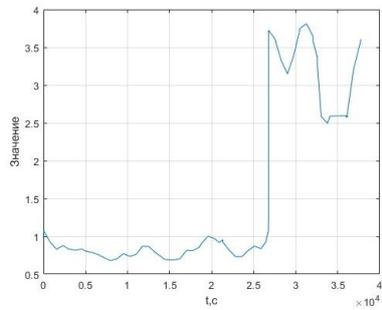


Рис. 9. Общий сигнал детекции температурных датчиков

Тогда выходной сигнал, изменяющийся от 0 до 1 (рис. 10), где 0 – пожароопасная ситуация отсутствует, 1 – возникла пожароопасная ситуация. Изменение порога приведено на рис. 11.

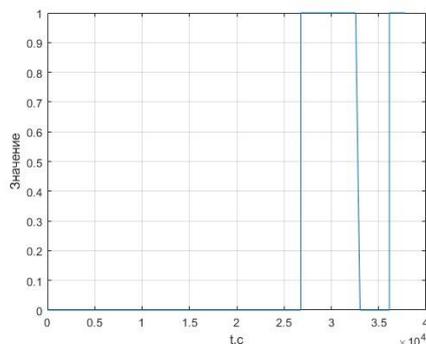


Рис. 10. График состояния срабатывания пожарного датчика по каналу температуры

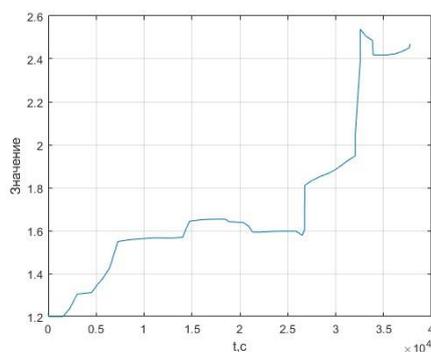


Рис. 11. Изменение порогового значения датчика температуры

Промоделируем работу датчика дыма в ультрафиолетовом спектре. Входные данные, поступающие от датчика в модель представлены на рис. 12.

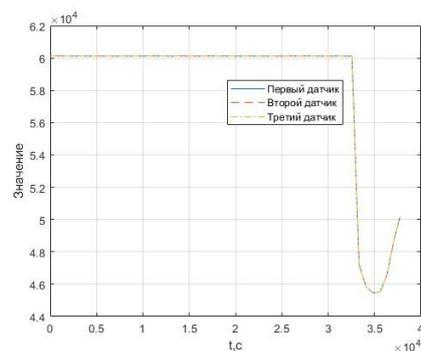


Рис. 12. График изменение входных данных ультрафиолетового спектра

Из рис. 12 видно, что входные данные по всем трем датчикам совпадают, соответственно среднее значение и стандартное отклонение по каждому из датчиков будет совпадать.

Изменение стандартного отклонения представлено на рис. 13.

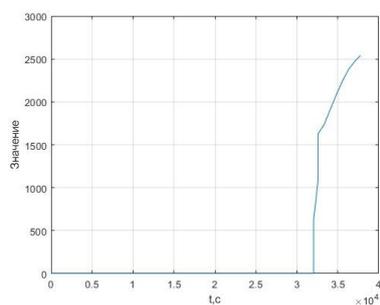


Рис. 13. График изменения стандартного отклонения ультрафиолетовых датчиков

Среднее значение сигнала ультрафиолетовых датчика приведено на рис. 14.

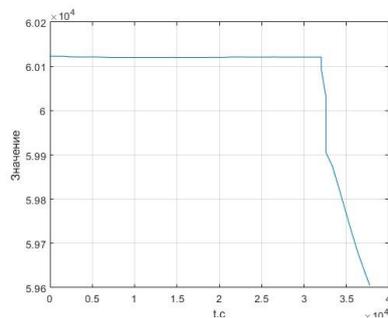


Рис. 14. График изменений среднего значения задымленности

Сигнал детекции ультрафиолетовых датчиков изображен на рис. 15.

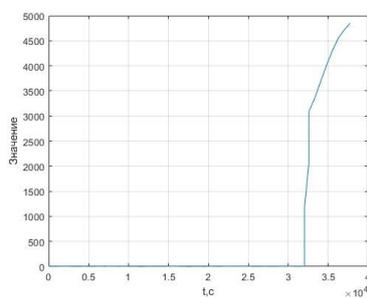


Рис. 15. Общий сигнал детекции ультрафиолетовых датчиков

Тогда выходной сигнал, изменяющийся от 0 до 1 (рис. 16), где 0 – пожароопасная ситуация отсутствует, 1 – возникла пожароопасная ситуация.

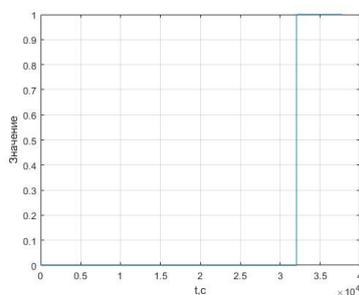


Рис. 16. График состояния срабатывания пожарного датчика по каналу ультрафиолетовых датчиков

Моделированием инфракрасного датчика можно пренебречь, т.к. его модель совпадает с моделью ультрафиолетового датчика, а также входные данные взятые за 10,5 часов также совпали. Это означает что пожароопасная ситуация в тот день затрагивала оба спектра: инфракрасный и ультрафиолетовый.

Обсуждение. Сравнивая результаты полученные из рисунков 6-16, можно сделать вывод, что моделируемая математическая модель температурного датчика работает корректно, в этом можно убедиться сравним графики детекции (рис. 9), выходного сигнала (рис. 10) и порогового значения (рис. 11), сопоставив порог с сигналом детекции и сравнив его с выходным графиком. Аналогичным образом можно подтвердить корректность моделирования инфракрасного и ультрафиолетовых датчиков.

Возникшие участки пожароопасных ситуации при моделировании совпадают с реальными зафиксированными компанией производителя цифрового универсального пожарного датчика. Принимая факт, что математическая модель в Simulink составлялись на основании теоретических математических моделях, представленных в формулах 7, 8, 9, 10, можно сделать вывод, что приведенные математические модели корректно описывают поведение реального цифрового пожарного датчика.

Заключение. Решения в области пожарной безопасности позволяют уменьшить ущерб причиняемый живым существам и материальным ценностям

В данной статье описана математическая модель реального цифрового пожарного датчика, выполнена его моделирование с использованием программного средства Simulink. Полученная математическая модель позволяет не только описать цифровой датчик, но и вносит вклад в будущие исследования по прогнозированию и обнаружению пожароопасных ситуаций с использованием пожарного датчика.

Дальнейшие исследования предполагают построение нейронной модели пожарного датчика, а также решения задачи прогнозирования пожароопасных ситуаций на основании данных из цифрового пожарного датчика с использованием нейронных сетей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Саутин И.Г. Концепция построения безопасной противопожарной автоматики // Алгоритм безопасности. – 2015. – № 4.
2. Саутин И.Г. Сверхраннее обнаружение дыма. Новые возможности // Алгоритм безопасности. – 2016. – № 5.
3. Саутин И.Г. Безопасность зданий и сооружений // Безопасность. Средства размещения. – 2017. – № 1.
4. Саутин И.Г. Противопожарная защита: технологии и решения // Транспорт. Противопожарная защита. Пожарная автоматика. Средства спасения. – 2018.
5. Саутин И.Г. Особое мнение. Можно ли доверить свою жизнь дымовому пожарному извещателю? // Алгоритм безопасности. – 2019. – № 6.
6. Зайцев А.В. Достоверность и своевременность обнаружения пожара, и как их учесть в нормах на СПС // Алгоритм безопасности. – 2016. – № 2.
7. Иванов А.Н. К вопросу об оценке эффективности пожарной автоматики // Пожарная безопасность: современные вызовы. Проблемы и пути решения: Матер. Всерос. науч.-практ. конф. – СПб.: С.-Петербург. ун-т ГПС МЧС России, 2021.
8. Преснов А.И. [и др.]. Пожарная техника: учеб.: в 2-х ч. Ч. 2. – СПб.: С.-Петербург. Ун-т ГПС МЧС России, 2016. – 404 с.
9. Qureshi W.S., Ekpanyapong M., Dailey M.N., Rinsurongkawong S., Malenichev A., Krasotkina O. QuickBlaze: Early fire detection using a combined video processing approach // Fire Technol. – 2016. – 52. – P. 1293-1317. – DOI: 10.1007/s10694-015-0489-7.
10. Mulholland, G.W. How well are we measuring smoke? // Fire and Materials. – June 1982. – Vol. 6, No. 2. – P. 65-67.
11. Bywater D. Detection of Real Fires by Carbon Monoxide Detectors—Foreign Experience. The Results of 10 Years of Research Lead to a Leap in Fire Detection Technology. Available online: https://www.aktivsb.ru/statii/obnaruzhenie_realnykh_pozharov_detektorami_ugarnogo_gaza_zarubezhnyy_opyt.html (accessed on 20 December 2023).
12. Luck H. and Hase K.R. Signal Detection Aspects in Automatic Fire Detection // Fire Safety Journal. – 1983. – 6. – P. 233-240.

13. Fedorov A., Bytcinskaya T., Lukyanchenko A., Hung T.D. Trends in the development of automatic fire detectors // *Technol. Technosphere Saf.* – 2009. – 23. – P. 111-114. Available online: <https://cyberleninka.ru/article/n/tendentsii-razvitiya-avtomaticheskikh-pozharnyh-izveschateley-1> (accessed on 20 December 2023).
14. Petrov A.E., Fedorov A.V., Kochegarov A.V., Lomaev E.N., Preobrazhenskiy A.P. The Analysis of Network Models for the Design of Industrial and Fire Safety Systems for Oil Refineries // *IOP Conf. Ser. Earth Environ. Sci.* – 2021. – 808. – 012024.
15. Bogdan L., Cristina B. The design of temperature control system using PIC18f4620; 2010.
16. MathWorks. «MATLAB & Simulink Help Center» MathWorks. – 2023. – <https://www.mathworks.com/help/index.html>. Accessed 20 December 2023.
17. Moler Cleve B. Numerical Computing with Matlab. The MathWorks, Inc., Natick, 2004.
18. Chapra Steven C. Applied Numerical Methods with MATLAB for Engineers and Scientists. McGraw Hill Companies, Inc. – 2nd ed. – New York, 2008.
19. Trench William F. Elementary Differential Equations. – Trench, 2013.
20. Ukil A., Braendle H., Krippner P. Distributed temperature sensing: Review of technology and applications // *IEEE Sens. J.* – 2011. – 12. – P. 885-892. – DOI: 10.1109/JSEN.2011.2162060.
21. Fawad Khan, Zhiguang Xu, Recent Advances in Sensors for Fire Detection, PMID: 35590999, 2022 Apr 26, DOI: 10.3390/s22093310.

REFERENCES

1. Sautin I.G. Kontsepsiya postroeniya bezopasnoy protivopozharnoy avtomatiki [The concept of constructing safe fire-fighting automation], *Algoritm bezopasnosti* [Safety Algorithm], 2015, No. 4.
2. Sautin I.G. Sverkhrannee obnaruzhenie dyma. Novye vozmozhnosti [Ultra-early smoke detection. New features], *Algoritm bezopasnosti* [Safety Algorithm], 2016, No. 5.
3. Sautin I.G. Bezopasnost' zdaniy i sooruzheniy [Safety of buildings and structures], *Bezopasnost'. Sredstva razmeshcheniya* [Safety. Accommodation facilities], 2017, No. 1.
4. Sautin I.G. Protivopozharnaya zashchita: tekhnologii i resheniya [Fire protection: technologies and solutions], *Transport. Protivopozharnaya zashchita. Pozharnaya avtomatika. Sredstva spaseniya* [Transport. Fire protection. Fire automatics. Rescue means], 2018.
5. Sautin I.G. Osoboe mnenie. Mozhno li doverit' svoyu zhizn' dymovomu pozharnomu izveshchatel'yu? [Special opinion. Can you trust your life to a smoke alarm?], *Algoritm bezopasnosti* [Security algorithm], 2019, No. 6.
6. Zaytsev A.V. Dostovernost' i svoevremennost' obnaruzheniya pozhara, i kak ikh uchest' v normakh na SPS [Reliability and timeliness of fire detection, and how to take them into account in ATP standards], *Algoritm bezopasnosti* [Safety Algorithm], 2016, No. 2.
7. Ivanov A.N. K voprosu ob otsenke effektivnosti pozharnoy avtomatiki [On the issue of assessing the effectiveness of fire automatics], *Pozharnaya bezopasnost': sovremennye vyzovy. Problemy i puti resheniya: Mater. Vseros. nauch.-prakt. konf.* [Fire safety: modern challenges. Problems and solutions: Materials of Vseros. scientific-practical conf.]. St. Petersburg: S.-Peterb. un-t GPS MChS Rossii, 2021.
8. Presnov A.I. [i dr.]. Pozharnaya tekhnika: ucheb. [Fire equipment: textbook]: in 2 part. Part 2. St. Petersburg: S.-Peterb. Un-t GPS MChS Rossii, 2016, 404 p.
9. Qureshi W.S., Ekpanyapong M., Dailey M.N., Rinsurongkawong S., Malenichev A., Krasotkina O. QuickBlaze: Early fire detection using a combined video processing approach, *Fire Technol.*, 2016, 52, pp. 1293-1317. DOI: 10.1007/s10694-015-0489-7.
10. Mulholland, G.W. How well are we measuring smoke?, *Fire and Materials*, June 1982, Vol. 6, No. 2, pp. 65-67.
11. Bywater D. Detection of Real Fires by Carbon Monoxide Detectors—Foreign Experience. The Results of 10 Years of Research Lead to a Leap in Fire Detection Technology. Available online: https://www.aktivsb.ru/statii/obnaruzhenie_realnykh_pozharov_detektorami_ugarnogo_gaza_zarubezhnyy_opyt.html (accessed on 20 December 2023).
12. Luck H. and Hase K.R. Signal Detection Aspects in Automatic Fire Detection, *Fire Safety Journal*, 1983, 6, pp. 233-240.
13. Fedorov A., Bytcinskaya T., Lukyanchenko A., Hung T.D. Trends in the development of automatic fire detectors, *Technol. Technosphere Saf.*, 2009, 23, pp. 111-114. Available online: <https://cyberleninka.ru/article/n/tendentsii-razvitiya-avtomaticheskikh-pozharnyh-izveschateley-1> (accessed on 20 December 2023).
14. Petrov A.E., Fedorov A.V., Kochegarov A.V., Lomaev E.N., Preobrazhenskiy A.P. The Analysis of Network Models for the Design of Industrial and Fire Safety Systems for Oil Refineries, *IOP Conf. Ser. Earth Environ. Sci.*, 2021, 808, 012024.
15. Bogdan L., Cristina B. The design of temperature control system using PIC18f4620; 2010.

16. MathWorks. «MATLAB & Simulink Help Center» MathWorks, 2023. Available at: <https://www.mathworks.com/help/index.html>. Accessed 20 December 2023.
17. Moler Cleve B. Numerical Computing with Matlab. The MathWorks, Inc., Natick, 2004.
18. Chapra Steven C. Applied Numerical Methods with MATLAB for Engineers and Scientists. McGraw Hill Companies, Inc. 2nd ed. New York, 2008.
19. Trench William F. Elementary Differential Equations. Trench, 2013.
20. Ukil A., Braendle H., Krippner P. Distributed temperature sensing: Review of technology and applications, *IEEE Sens. J.*, 2011, 12, pp. 885-892. DOI: 10.1109/JSEN.2011.2162060.
21. Fawad Khan, Zhiguang Xu, Recent Advances in Sensors for Fire Detection, PMID: 35590999, 2022 Apr 26, DOI: 10.3390/s22093310.

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

Сингх Санны – Южный федеральный университет; e-mail: singkh@sfedu.ru; г. Таганрог, Россия; тел.: +79885751350; кафедра систем автоматического управления; аспирант.

Прибыльский Алексей Васильевич – e-mail: apribylsky@sfedu.ru; тел.: +79885619718; кафедра систем автоматического управления; к.т.н.; доцент.

Singh Sanni – Southern Federal University; e-mail: singkh@sfedu.ru; Taganrog, Russia; phone: +79885751350; the department of automatic control systems; graduate student.

Pribylskiy Alexey Vasilievich – e-mail: apribylsky@sfedu.ru; phone: +79885619718; the department of automatic control systems; cand. of eng. sc.; associate professor.

УДК 004.056.53+347.837+654.16

DOI 10.18522/2311-3103-2024-2-132-141

А.В. Дьяков, К.Е. Румянцев

АКТУАЛЬНЫЕ ПРОБЛЕМЫ РАДИОМОНИТОРИНГА В СИСТЕМЕ МЕРОПРИЯТИЙ ПО ОБЕСПЕЧЕНИЮ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Беспроводные сети передачи данных порождают угрозы, от которых невозможно защититься традиционными для проводных сетей средствами, поскольку в таком случае невозможно обеспечить эквивалент безопасности проводных сетей в силу физических свойств канала связи. Целью статьи является определение актуальных проблем, существующих при обеспечении информационной безопасности (ИБ) в беспроводных сегментах сетей передачи данных. Для достижения поставленной цели из банка угроз безопасности информации ФСТЭК России осуществлена выборка угроз, потенциально реализуемых в беспроводных сетях. Установлено, что реализация таких угроз может приводить к полному набору нарушений состояния ИБ, а именно: к нарушению конфиденциальности, целостности и доступности информации. Рассмотрены существующие практические способы обеспечения ИБ в беспроводных сегментах сетей. Анализ этих способов указал на присутствующую при радиомониторинге техническую возможность создания дополнительного рубежа в системе эшелонированной защиты информации. В свою очередь, это обеспечивает потенциал обнаружения уязвимостей и вторжений на канальном уровне сетевого взаимодействия как в локальных сетях предприятий, так и в крупномасштабных сетях общего пользования. В соответствии с целью сгруппированы аспекты построения такого рубежа защиты, связанные с контролем канального уровня сетевого взаимодействия беспроводных устройств, уменьшением размеров частотно-территориальных кластеров и правовым обеспечением. Обзором публикаций выявлен разрыв между существующими подходами к радиомониторингу и обеспечению ИБ, также обнаружена слабая развитость направления, связанного с исследованиями в области обнаружения и предотвращения беспроводных вторжений. Полученный результат указывает на необходимость пересмотра сложившейся концепции радиомониторинга и разработки соответствующих организационно-технических мер для его интеграции в систему мероприятий по обеспечению ИБ, что должно помочь решить проблему своевременного обнаружения и предотвращения вторжений в беспроводные сегменты сетей передачи данных, а также выявления уязвимых элементов инфраструктуры этих сетей.

Информационная безопасность; радиоконтроль; радиомониторинг; системы обнаружения вторжений; СОВ; беспроводные сети; Wi-Fi; WIDS; WIPS.

A.V. Dyakov, K.E. Rumyantsev

CURRENT PROBLEMS OF RADIOMONITORING IN THE SYSTEM OF ACTIONS TO ENSURE INFORMATION SECURITY

Wireless data transmission networks generate threats that cannot be protected against by means traditional for wired networks, because in this case it is impossible to provide equivalent security of wired networks due to the physical properties of the communication channel. The purpose of the article is to determine the actual problems that exist in ensuring information security (IS) in wireless segments of data networks. To achieve this goal, a selection of threats potentially realizable in wireless networks has been made from the information security threat bank of FSTEC of Russia. It is established that the realization of such threats can lead to a full set of violations of the state of IS, namely: violation of confidentiality, integrity and availability of information. The existing practical ways of providing IS in wireless segments of networks are considered. The analysis of these methods pointed out the technical possibility of creating an additional boundary in the system of echeloned information protection. In turn, this provides the potential to detect vulnerabilities and intrusions at the link layer of network communication both in local networks of enterprises and in large-scale public networks. In accordance with the goal, aspects of building such a defense frontier are grouped, related to control of the link layer of network interaction of wireless devices, reduction of frequency-territorial clusters and legal support. The review of publications reveals a gap between the existing approaches to radio monitoring and IS provision, and also reveals poor development of the direction related to research in the field of detection and prevention of wireless intrusions. The obtained result indicates the need to revise the existing concept of radio monitoring and develop appropriate organizational and technical measures for its integration into the system of measures to ensure IS, which should help to solve the problem of timely detection and prevention of intrusions into wireless segments of data networks, as well as the identification of vulnerable elements of the infrastructure of these networks.

Information security; radiocontrol; radiomonitoring; intrusion detection systems; wireless networks; Wi-Fi; WIDS; WIPS.

Введение. Беспроводная передача информации, не ограничивающая абонентов определенной точкой пространства, всегда вызывала большой интерес. Сегодня с уверенностью можно сказать о том, что беспроводная техника буквально завоевывает мир. Среди причин такой популярности можно отметить мобильность и эстетическую привлекательность: помимо свободы передвижения отсутствие проводов делает интерьер опрятным, исчезают трудности, связанные с распутыванием кабелей и механической надёжностью соединений. Рост популярности беспроводных технологий объясняется также развитием самой инфраструктуры: появляются новые стандарты передачи данных, обеспечивающие высокую скорость передачи информации, увеличивается количество базовых станций и, как следствие, расширяется зона покрытия. Встроенное программное обеспечение устройств существенно сокращает время на настройку соединения: пользователю достаточно всего один раз подключиться к защищённой сети, после чего повторные подключения к ней будут осуществляться устройством автоматически. В случае мобильной связи с использованием SIM-карт, средств бесконтактной оплаты с технологией NFC, а также устройств радиочастотной идентификации (RFID) у пользователя принципиально отсутствует необходимость в каких-либо настройках и получении для этого специальных знаний.

Вместе с тем каждый новый этап распространения беспроводных технологий создаёт такие угрозы, которые зачастую трудно спрогнозировать, не говоря уже о разработке и внедрении контрмер для предотвращения реализации этих угроз. С позиции возможного нарушения состояния информационной безопасности (ИБ) беспроводной сегмент телекоммуникационной сети является самым уязвимым звеном по ряду причин. Во-первых, в реальных условиях радиоэлектронные средства (РЭС), действующие в беспроводной сети, всегда будут подвержены непреднамеренному интерференционному воздействию других РЭС и излучений иных источников электромагнитных помех. Нарушение электромагнитной совместимости приводит не только к снижению скорости передачи данных по радиоканалу, но и вовсе к потере связи и приведению информационной системы в состояние «отказ в обслуживании» (англ. Denial of Service, DoS). Во-вторых, между конечными точками беспроводного сегмента сети находится неограниченная область – ра-

диоэфир, в связи с чем у злоумышленника появляется возможность не преодолевать физическую защиту и не контактировать с аппаратными средствами информационной системы. Анонимность беспроводных вторжений [1] делает для нарушителя эту возможность ещё более привлекательной. Воздействия, при которых беспроводные сети становятся уязвимыми для несанкционированного доступа, с точки зрения обеспечения ИБ представляют гораздо большую опасность по сравнению с DoS-атаками.

Для оценки негативных последствий реализации угроз в беспроводных сетях можно воспользоваться банком данных угроз безопасности информации ФСТЭК России [2], выбрав из него угрозы, потенциально реализуемые в этих сетях. Полученные таким образом данные консолидированы в табл. 1 и визуализированы на диаграмме (рис. 1).

Таблица 1

Последствия реализации угроз в беспроводных сетях

Идентификатор угрозы	Последствия реализации угрозы		
	нарушение конфиденциальности	нарушение целостности	нарушение доступности
УБИ.011			+
УБИ.030	+	+	+
УБИ.069	+	+	
УБИ.083	+	+	+
УБИ.116		+	
УБИ.125	+	+	+
УБИ.126	+		+
УБИ.133	+		
УБИ.139	+	+	+
УБИ.140			+

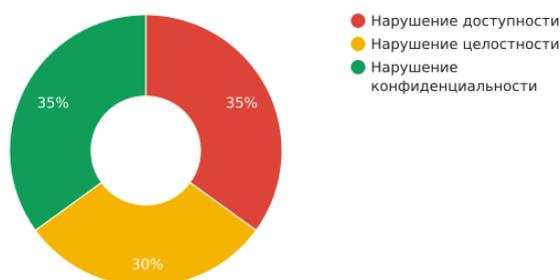


Рис. 1. Диаграмма распределения последствий реализации угроз в беспроводных сетях

Распределение последствий реализации угроз в беспроводных сетях указывает на то, что нарушения доступности, целостности и конфиденциальности информации возможны и находятся приблизительно в одинаковых пропорциях относительно друг друга.

Для всестороннего рассмотрения проблемы вторжений в системы беспроводной передачи данных помимо определения угроз и последствий их реализации необходимо также установить категории нарушителей. Модель нарушителя, построенная согласно нормативным документам [3], будет носить весьма общий характер. Основными факторами, которые окажутся неучтенными, являются цели и мотивация нарушителя. В некоторых случаях эти факторы являются решающими [4]. В оценке потенциала нарушителя, даже при учете его компетенции, основополагающую роль играет мотивация [5]: так, в отсутствие необходимой компетенции именно мотивация толкает злоумышленника на совершение нарушения.

Проблема контроля канального уровня сетевого взаимодействия. Арсенал средств защиты информации в телекоммуникационных каналах традиционно представляет собой межсетевые экраны, а также может включать в себя средства криптографиче-

ского шифрования и системы обнаружения и предотвращения вторжений (англ. Intrusion Detection and Prevention System, IDS/IPS). Последние в качестве входной информации используют поток пакетов сетевого трафика [6–8]. Рассматривая перечисленные средства в иерархии семиуровневой модели сетевого взаимодействия OSI [9], можно сделать вывод от том, что действуют они в подавляющем большинстве на сетевом уровне модели OSI. Работа с канальным уровнем модели OSI зачастую сводится лишь к ограничительным мерам, правила которых задаются вручную и являются по сути статическими, например: фильтрация адресов устройств в управляемых коммутаторах и настройка паролей беспроводных интерфейсов. При этом, в соответствии с принципом декапсуляции, подробности взаимодействия, происходящего на канальном уровне, при движении к более высшим уровням утрачиваются и уже на сетевом уровне отсутствуют. Таким образом, доступ к канальному уровню сетевого взаимодействия возможен только лишь с низшего слоя модели OSI – физического уровня, что для беспроводных сетей является радиоэфиром.

Идея контроля активности на канальном уровне посредством специализированных радиоприёмных устройств не нова: беспроводные IDS/IPS (англ. Wireless IDS/IPS, WIDS/WIPS) существуют и представляют собой системы, которые осуществляют мониторинг радиоэфира, анализируют полученную информацию об источниках радиосигнала, их взаимодействиях и аномальных активностях, а также предотвращают действия, противоречащие настроенной политике предотвращения вторжений.

Преимущества построения IDS/IPS на более низких уровнях сетевой модели OSI очевидны: эшелонированная защита обеспечивает собственные меры безопасности на каждом уровне, если нарушитель сумеет преодолеть один рубеж безопасности, то он столкнется со следующим, что значительно затруднит проникновение в информационную систему через периметр сети.

Необходимо отметить, что научное направление построения WIDS/WIPS развито слабо. В русскоязычных публикациях упоминания таких систем встречаются нечасто и преимущественно в рекламных брошюрах ведущих иностранных производителей сетевого оборудования. В зарубежных публикациях научный интерес представляет эксплуатационная документация к оборудованию WIDS/WIPS, из которой можно сложить некоторое представление об алгоритмах действия этих устройств. По существу, упомянутые WIDS/WIPS являются дополнительным функционалом сетевого оборудования семейства стандартов IEEE 802.11 и позволяют организовывать ещё один рубеж безопасности беспроводного сегмента телекоммуникационной корпоративной сети, состоящей из одной или нескольких точек доступа Wi-Fi, находящихся в ограниченном пространстве. Такие WIDS и WIPS предназначены для защиты интересов именно владельца базовой станции (точки доступа) и находящейся за ней информационной системы, при этом у абонентов нет возможности удостовериться в надёжности сети, к которой они производят подключение. Кроме того, контроль телекоммуникационных сетей других стандартов и масштабирование таких WIDS/WIPS до уровня сетей общего пользования не представляется возможным.

Доступ к канальному уровню модели OSI также может быть получен через физический уровень посредством техники, предназначенной для контроля за излучениями РЭС и стоящей на вооружении у специальных служб, осуществляющих радиоконтроль и радиомониторинг. Специфика данных мероприятий, заключающаяся в закрытости указанных служб, накладывает существенный отпечаток на направленность публикаций: подавляющее большинство из них посвящены решению широко известных научных и технических проблем, совершенствованию существующих методов и их программно-аппаратных реализаций. Так, наиболее полно типовые задачи радиомониторинга, аспекты построения и функционирования систем радиоконтроля, а также опыт осуществления мероприятий радиоконтроля в сложной помеховой обстановке в промышленных центрах, внутри зданий и на открытой местности освещены в [10, 11]. Благодаря высокой концентрации теоретической и практической информации книги по существу являются настольными справочниками и высоко востребованы специалистами в области радиомониторинга, руководителями радиоконтрольных служб, сотрудниками силовых ведомств и служб безопасности государственных и коммерческих структур.

Вместе с тем, несмотря на глубокое и многогранное изложение материала в указанных источниках, а также немалое количество обзорных и рационализаторских публикаций, основанных на данных источниках, о существовании острых и актуальных практических трудностей применения существующего радиоконтрольного оборудования можно судить лишь по единичным публикациям. Так, наличие серьёзных пробелов, вызванных формальным подходом к радиоконтролю, открыто заявлено лишь в [12]. Авторы недвусмысленно и эмоционально указывают на то, что современный (по состоянию на 2017 год) радиоконтроль невозможен без анализа трафика цифровых каналов связи, причём извлекаемая информация носит не второстепенный характер, а должна являться приоритетным фактором принятия решений. Также авторами показаны простые и одновременно яркие примеры абсурдности формального подхода в интересах обеспечения защиты информации, заключающегося в анализе амплитудно-частотных характеристик излучений «классическим радиоконтролем». Повсеместное внедрение беспроводной цифровой связи привело к тому, что эта проблема за считанные годы распространилась из отдельных защищаемых помещений на масштабные территории мегаполисов и государств в целом.

Невозможность осуществления полноценного радиоконтроля без вскрытия передаваемой информации подтверждается также и в [13], где коллектив авторов (сотрудников одной из ведущих организаций-производителей российского радиоконтрольного оборудования) предлагает способ адресного пеленгования базовых станций GSM, UMTS, LTE сетей сотовой связи. При этом авторы также заявляют о том, что «в большинстве сотовых систем связи, в частности, UMTS и LTE, множество базовых станций, характеризующихся своими идентификационными параметрами, осуществляют одновременную передачу в одном частотном диапазоне. Это делает принципиально невозможным применение фазоразностного метода пеленгования с непосредственным вычислением пеленга и требует выделения сигнала каждой обнаруженной БС из суммарного группового сигнала». Вводится понятие «тонкой структуры» сигналов базовых станций.

Проблема уменьшения размеров частотно-территориальных кластеров. Поскольку радиочастотный спектр как ресурс ограничен, увеличение пропускной способности каналов (и, как следствие, их широкополосности) становится возможным путём увеличения плотности размещения РЭС с одновременным уменьшением зон обслуживания каждого средства. Таким образом, доля аппаратуры стандартов и технологий беспроводной связи на коротких расстояниях с каждым годом будет увеличиваться вместе с ростом широкополосной такой аппаратуры. Приведённая тенденция развития беспроводных сетей связи свидетельствует о том, что всё большее количество РЭС будет одновременно использовать одни и те же радиочастотные каналы. Начиная с технологий связи третьего поколения (3G, UMTS), стало возможным развёртывать масштабные одночастотные беспроводные сети. Такая особенность делает традиционные амплитудно-частотные методы радиоконтроля неэффективными, поскольку стационарные комплексы радиоконтроля в таком случае смогут обнаруживать лишь один ближайший передатчик, под мощным излучением которого будут маскированы излучения от других передатчиков, находящихся на большем удалении и работающих в том же частотном канале. Уменьшение размеров частотно-территориальных кластеров приводит к тому, что развитие сети дорогостоящих стационарных комплексов радиоконтроля в перспективе не приведёт к сколь-нибудь значимому увеличению охвата территории радиоконтролем. С этой проблемой уже сталкиваются службы ИБ, обеспечивающие защиту информации в рамках даже относительно небольшого ограниченного пространства предприятий, применяя пространственно-распределённые системы радиоконтроля (TORNADO-RxMTCA, RS1000, АРК-АБС и прочие).

В случае контроля сетей общего пользования, развёрнутых на масштабных территориях, стратегия увеличения количества стационарных мониторинговых комплексов невозможна, поскольку количество последних должно будет расти соразмерно с количеством базовых станций (точек беспроводного доступа). В настоящее время единственной «панацеей» для решения подобных задач являются мобильные комплексы радиоконтроля. Двигаясь по установленному маршруту, они позволяют обнаруживать большее количество РЭС по сравнению со стационарными изделиями. Вместе с тем возникает пробле-

ма идентификации РЭС, поскольку помимо обнаружения самого факта излучения необходимо вскрыть содержимое передаваемой информации и выделить уникальные идентификаторы-позывные, позволяющие указать на присутствие излучения от конкретных РЭС в электромагнитном групповом «смоге». Так, поиски незаконно-действующих РЭС превращаются в «поиск иголки в стоге сена»: априори неизвестно, существуют ли такие средства, а обнаружить их становится возможным лишь оказавшись в непосредственной близости, при этом размер зоны обнаружения неуклонно сокращается из-за ввода в эксплуатацию новых средств. В результате, вместе с постоянным ростом трудозатрат на осуществление мероприятий радиомониторинга, также растёт и нагрузка на специалистов, быстрее вырабатывается ресурс дорогостоящего радиоконтрольного оборудования, увеличиваются расходы на горюче-смазочные материалы и на поддержание в надлежащем состоянии автомобильного парка. Кроме того, подобная тактика поиска вторжений в сети и уязвимых элементов её инфраструктуры не обеспечивает непрерывности контроля.

Проблема правового обеспечения. Действующий в настоящее время на территории Российской Федерации порядок радиоконтроля РЭС гражданского назначения основан на том, что в ходе мероприятий радиоконтроля его объектами главным образом являются параметры излучений РЭС с привязкой к местам их установки. Сами РЭС и их взаимодействие между собой представляют интерес лишь в экстраординарных случаях [14], например при установлении причин нарушений связи в цифровых системах и локализации источников DoS-атак, при этом специалисты радиоконтроля де-факто вынуждены действовать за пределами руководящих документов [15], то есть заниматься «самостоятельностью» со всеми вытекающими отсюда последствиями. В таких случаях успех мероприятия не гарантирован даже при наличии у специалиста соответствующего оборудования, опыта и некоторых других сопутствующих факторов, включая аналитические способности, физическую силу, артистизм, способность пойти на риск, и даже интуицию, поэтому проведение подобных операций является своего рода искусством.

Таким образом, при радиоконтроле анализу подлежит именно физический уровень модели сетевого взаимодействия OSI; требования к обязательному контролю канального уровня и тем более руководящие документы отсутствуют, несмотря на наличие соответствующих технических возможностей. Кроме того, подавляющее большинство РЭС, являющихся абонентскими станциями (стационарными, носимыми и возимыми), не представляют интерес у надзорных органов и оказываются неохваченными радиоконтролем.

Ранее упомянутая проблема, заключающаяся в уменьшении размеров частотно-территориальных кластеров, хорошо известна специалистам радиоконтроля и вынуждает их применять совместно со средствами измерения так-называемые «индикаторы» [16]. Данное вспомогательное оборудование позволяет получать с канального уровня модели OSI не только сведения о передаваемой в эфир идентификационной информации (MAC-адреса устройств, CID, LAC, MCC, MNC базовых станций сотовой связи и т.п.), но и о подробностях сетевого взаимодействия, что даёт возможность строить топологии сетей. В качестве примера таких индикаторов можно привести носимые изделия отечественного производства RAD-001 (ООО НПФ «Радиян-М»), Барс-GSM (ООО «ГАИП»), Барс-У-мини (ООО «СТЦ»). Широчайшим спектром возможностей подобного рода обладают и мобильные комплексы на основе радиоприёмного устройства Аргамак-ИС (АО «ИРКОС»). Однако, нормативно-правовой статус у таких средств в настоящее время отсутствует, что стало возможным после череды законодательных новшеств. В действующем законодательстве [17] определено, что в качестве доказательств правонарушений могут рассматриваться показания специальных технических средств, под которыми понимаются приборы, утверждённые в установленном порядке в качестве средств измерения, имеющие соответствующие сертификаты и метрологическую поверку. Индикаторы (вспомогательное оборудование) не являются средствами измерения и не могут обладать метрологическими характеристиками в принципе. Кроме того, формируемые индикаторными средствами отчёты, могут быть свободно отредактированы, то есть сфальсифицированы. Таким образом, полученные при помощи индикаторов сведения невозможно использовать в качестве доказательства факта нарушения. По этим причинам, при от-

сутствии у владельца РЭС действий, направленных на устранение нарушения в добровольном порядке, возникают прецеденты невозможности принуждения владельца РЭС к устранению нарушений.

Вместе с тем, сама процедура установления юридического владельца РЭС зачастую сопряжена со значительными трудностями процессуального характера: нередки случаи, когда, несмотря на совокупность прямых и косвенных признаков, нарушитель отрицает факт владения неразрешённым РЭС [18]. При контроле соблюдения требований об обязательной идентификации абонентов при подключении к сети «Интернет» [19] Роскомнадзор, в чьём ведении находится осуществление соответствующего контроля, не имеет полномочий на проведение проверок в отношении физических и юридических лиц, не являющихся операторами связи. Данные о таких случаях передаются в правоохранительные органы (МВД, ФСБ, прокуратуру РФ), однако дальнейшее движение материалов нередко останавливается из-за ряда иных вполне очевидных причин, включающих мало-значительность гипотетического правонарушения и несоразмерность сил и средств, необходимых для установления всех обстоятельств, возможному размеру вреда и тяжести последствий. Также отсутствует механизм установления факта организации физическим лицам публичных сетей с использованием точек доступа, закрытых паролем [20].

Кроме того, в соответствии с постановлением [21] у органов радиочастотной службы существует ряд задач, которые невозможно выполнить качественно без применения индикаторных средств, позволяющих анализировать каналный уровень сетевого взаимодействия [16]. Наиболее ответственными среди этих задач можно отметить радиоконтроль базовых станций сотовой связи в пограничной зоне (вблизи государственной границы) при международной правовой защите частотных присвоений, работы по оценке выполнения операторами связи требований к защите сетей связи от несанкционированного доступа к ним и передаваемой по ним информации, а также проведение мониторинга в целях выявления угроз устойчивости, безопасности и целостности функционирования на территории Российской Федерации информационно-телекоммуникационной сети «Интернет» и сети связи общего пользования.

Сложившаяся ситуация, связанная с несовершенством правового обеспечения, нередко приводит к абсурдности радиомониторинга в системе мероприятий по обеспечению ИБ вследствие того, что получаемые данные не имеют юридической силы и не могут служить основанием для нейтрализации возникающих угроз. В свою очередь, количественное накопление проблем приводит к качественным изменениям в обществе: наблюдая за неспособностью государства дать ответы на возникающие вызовы и угрозы, граждане утрачивают доверие к власти, что в конечном итоге вынуждает их решать возникающие проблемы способами, лежащими вне правового поля.

Выводы. В настоящее время сложилась парадоксальная ситуация, когда между защитой информации и радиомониторингом образовалась своеобразная техническая, методологическая и нормативно-правовая пустота. Такое положение является существенным препятствием как для эффективного осуществления мероприятий по обеспечению ИБ, так и надзорной деятельности за излучениями РЭС и регулирования использования радиочастотного спектра. Разработка и внедрение организационно-технических мер, позволяющих преодолеть проблему контроля беспроводных сетей на канальном уровне сетевого взаимодействия, проблему уменьшения размеров частотно-территориальных кластеров и проблему правового обеспечения, должны создать благоприятные условия для качественного социально-экономического развития общества.

Данные, получаемые в ходе радиомониторинга беспроводных сетей с подробностями взаимодействия на канальном уровне, могут быть востребованы не только в целях предотвращения вторжений в такие сети, но и в иных задачах, например:

- ◆ получение дополнительного фактора идентификации личности (владельца устройства) в системах распознавания клиентов (в банках, клубах, библиотеках и иных общественных местах);
- ◆ решение широкого спектра оперативно-разыскных задач, включая поиск похищенного имущества и пропавших без вести лиц;

- ◆ установление личности граждан, поступающих в лечебные учреждения в состоянии нарушенного сознания;
- ◆ контроль за соблюдением ИБ в части ограничений на внос техники в режимные зоны;
- ◆ выявление базовых станций, позволяющих подключаться к сети «Интернет» без идентификации (в нарушение требований [19]);
- ◆ выявление беспроводного сегмента сетей, не ограничивающих доступ к запрещённой информации (нарушение требований [22]).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Меррит М., Поллино Д.* Безопасность беспроводных сетей / пер. с англ. Семенова А.В. – М.: Компания АйТи; ДМК Пресс, 2004. – 288 с.
2. Банк данных угроз безопасности информации // ФСТЭК России. – URL: <https://bdu.fstec.ru/threat> (дата обращения: 27.01.2024).
3. Методический документ «Методика оценки угроз безопасности информации» (утв. Федеральной службой по техническому и экспортному контролю 5 февраля 2021 г.).
4. *Голембиовская О.М., Рытов М.Ю., Шинаков К.Е., Горлов А.П., Губсков Ю.А., Голембиовский М.М., Кондрашова Е.В.* Этапы формирования модели угроз и модели нарушителя информационной безопасности с учетом изменений законодательства Российской Федерации. – 2-е изд. – Саратов: Вузовское образование, 2024. – 265 с.
5. *Голембиовская О.М., Рытов М.Ю., Шинаков К.Е., Голембиовский М.М., Кондрашова Е.В.* Формализация подхода к определению степени ущерба и потенциала нарушителя. – 2-е изд. – Саратов: Вузовское образование, 2024. – 75 с.
6. *Полтавцева М.А., Лаврова Д.С.* Высокопроизводительные системы обнаружения вторжений: учеб. пособие. – 2-е изд. – М.; Вологда: Инфра-Инженерия, 2023. – 152 с. – ISBN 978-5-9729-1213-1.
7. *Шелухин О.И., Руднев А.Н., Савелов А.В.* Системы обнаружения вторжений в компьютерные сети: учеб. пособие. – М.: Московский технический университет связи и информатики, 2013. – 88 с.
8. *Басыня Е.А.* Сетевая информационная безопасность: учебник. – М.: Национальный исследовательский ядерный университет «МИФИ», 2023. – 224 с. – ISBN 978-5-7262-2949-2.
9. *Таненбаум Э., Фимстер Н., Уэзеролл Д.* Компьютерные сети. – 6-е изд. – СПб.: Питер, 2023. – 992 с.
10. *Рембовский А.М., Ашихмин А.В., Козьмин В.А.* Радиомониторинг: задачи, методы, средства. – 3-е изд. – М.: Горячая линия – Телеком, 2012. – 640 с.
11. *Слободянюк П.В., Благодарный В.Г.* Радиомониторинг. Вчера, сегодня, завтра (Теория и практика построения системы радиомониторинга). – Прилуки: ООО «Издательство «Air-Поліграф», 2010. – 296 с.
12. *Захаров А.В., Кривицун А.В.* Имитация бурной деятельности, или каким не должен быть радиоконтроль в XXI веке // Информационно-методический журнал «Защита информации. Инсайд». – 2017. – № 1.
13. *Манелис В.Б., Сладких В.А., Козьмин В.А., Бизюков П.Е.* Адресное пеленгование базовых станций GSM, UMTS, LTE сетей сотовой связи // Системы управления, связи и безопасности. – 2021. – № 2.
14. *Карш А., Симонов Д.* Каждый инспектор желает знать. Причины возникновения радиопомех // Радиочастотный спектр. – 2013. – № 10. – С. 38-42.
15. Приказ Роскомнадзора от 02.02.2010 №78 «Об утверждении Инструкции по поиску и обнаружению источников радиопомех».
16. *Курков А., Пулин А.* Ловись, РЭС, большое и маленькое. Радиоконтроль сетей беспроводного доступа не дремлет // Радиочастотный спектр. – 2013. – № 12. – С. 34-37.
17. Кодекс Российской Федерации об административных правонарушениях от 30.12.2001 №195-ФЗ (ред. от 25.12.2023).
18. *Бикшанов П.* Маскировочная ШБД-сеть. Как владельцы неразрешённых беспроводных РЭС скрывают свои устройства // Радиочастотный спектр. – 2017. – № 8. – С. 44-46.
19. Постановление Правительства Российской Федерации от 31.12.2021 № 2607 «Об утверждении Правил оказания телематических услуг связи».
20. *Жаров А.А.* Публичный Wi-Fi выходит из тени // Радиочастотный спектр. – 2016. – № 10. – С. 15-16.
21. Постановление Правительства РФ от 14 мая 2014 г. № 434 «О радиочастотной службе».
22. Федеральный закон от 27.07.2006 № 149-ФЗ «Об информации, информационных технологиях и о защите информации».

REFERENCES

1. *Merrit M., Pollino D.* Bezopasnost' besprovodnykh setey [Security of wireless networks], transl. from engl. by Semenova A.V. Moscow: Kompaniya AyTi; DMK Press, 2004, 288 p.
2. Bank danykh ugroz bezopasnosti informatsii [Information security threats data bank], *FSTEC Rossii* [FSTEC of Russia]. Available at: <https://bdu.fstec.ru/threat> (data obrashcheniya: 27 January 2024).
3. Metodicheskiy dokument «Metodika otsenki ugroz bezopasnosti informatsii» (utv. Federal'noy sluzhboy po tekhnicheskomu i eksportnomu kontrolyu 5 fevralya 2021 g.) [Methodological document «Methodology for Assessing Information Security Threats» (approved by the Federal Service for Technical and Export Control on February 5, 2021)].
4. *Golembiovskaya O.M., Rytov M.Yu., Shinakov K.E., Gorlov A.P., Gubskov Yu.A., Golembiovskiy M.M., Kondrashova E.V.* Etapy formirovaniya modeli ugroz i modeli narushitelya informatsionnoy bezopasnosti s uchedom izmeneniy zakonodatel'stva Rossiyskoy Federatsii [Stages of formation of the threat model and the model of the information security intruder taking into account changes in the legislation of the Russian Federation]. 2nd ed. Saratov: Vuzovskoe obrazovanie, 2024, 265 p.
5. *Golembiovskaya O.M., Rytov M.Yu., Shinakov K.E., Golembiovskiy M.M., Kondrashova E.V.* Formalizatsiya podkhoda k opredeleniyu stepeni ushcherba i potentsiala narushitelya [Formalization of the approach to determining the degree of damage and potential of the intruder]. 2nd ed. Saratov: Vuzovskoe obrazovanie, 2024, 75 p.
6. *Poltavtseva M.A., Lavrova D.S.* Vysokoproizvoditel'nye sistemy obnaruzheniya vtorzheniy: ucheb. posobie [High-performance intrusion detection systems: a textbook]. 2nd ed. Moscow; Vologda: Infra-Inzheneriya, 2023, 152 c. ISBN 978-5-9729-1213-1.
7. *Shelukhin O.I., Rudnev A.N., Savelov A.V.* Sistemy obnaruzheniya vtorzheniy v komp'yuternye seti: ucheb. posobie [Intrusion detection systems in computer networks: a textbook]. Moscow: Moskovskiy tekhnicheskii universitet svyazi i informatiki, 2013, 88 p.
8. *Basynya E.A.* Setevaya informatsionnaya bezopasnost': uchebnik [Network information security: textbook]. Moscow: Natsional'nyy issledovatel'skiy yadernyy universitet «MIFI», 2023, 224 p. ISBN 978-5-7262-2949-2.
9. *Tanenbaum E., Fimster N., Uezeroll D.* Komp'yuternye seti [Computer Networks]. 6th ed. –Saint Petersburg.: Piter, 2023, 992 p.
10. *Rembovskiy A.M., Ashikhmin A.V., Koz'min V.A.* Radiomonitoring: zadachi, metody, sredstva [Radio monitoring: tasks, methods, means]. 3rd ed. Moscow: Goryachaya liniya – Telekom, 2012, 640 p.
11. *Slobodyanyuk P.V., Blagodarnyy V.G.* Radiomonitoring. Vchera, segodnya, zavtra (Teoriya i praktika postroyeniya sistemy radiomonitoringa) [Radiomonitoring. Yesterday, today, tomorrow (Theory and practice of building a radio monitoring system)]. Priluki: OOO «Izdatel'stvo «Aip- Poligraf», 2010, 296 p.
12. *Zakharov A.V., Krivtsun A.V.* Imitatsiya burnoy deyatelnosti, ili kakim ne dolzhen byt' radiokontrol' v XXI veke [Imitation of stormy activity, or what should not be radio monitoring in the XXI century], *Informatsionno-metodicheskiy zhurnal «Zashchita informatsii. Insayd»* [Information-methodical journal "Information Protection. Inside"], 2017, No. 1.
13. *Manelis V.B., Sladkikh V.A., Koz'min V.A., Bizyukov P.E.* Adresnoe pelengovanie bazovykh stantsiy GSM, UMTS, LTE setey sotovoy svyazi [Address direction finding of GSM, UMTS, LTE base stations of cellular communication networks], *Sistemy upravleniya, svyazi i bezopasnosti* [Control, Communication and Security Systems], 2021, No. 2.
14. *Karsh A., Simonov D.* Kazhdyy inspektor zhelaet znat'. Prichiny vozniknoveniya radio-pomekh [Every inspector wants to know. Causes of radio interference], *Radiochastotnyy spektr* [Radio Frequency Spectrum], 2013, No. 10, pp. 38-42.
15. Prikaz Roskomnadzora ot 02.02.2010 №78 «Ob utverzhdenii Instruktsii po poisku i obnaruzheniyu istochnikov radiopomekh» [Order of Roskomnadzor from 02.02.2010 № 78 "On approval of the Instruction on search and detection of radio interference sources"].
16. *Kurkov A., Pulin A.* Lovis', RES, bol'shoe i malen'koe. Radiokontrol' setey besprovodnogo dostupa ne dremlit [Catch, RPS, big and small. Radio control of wireless access networks does not slumber], *Radiochastotnyy spektr* [Radio Frequency Spectrum], 2013, No. 12, pp. 34-37.
17. Kodeks Rossiyskoy Federatsii ob administrativnykh pravonarusheniyakh ot 30.12.2001 №195-FZ (red. ot 25.12.2023) [Code of the Russian Federation on Administrative Offenses of 30.12.2001 №195-FL (ed. of 25.12.2023)].
18. *Bikshanov P.* Maskirovochnaya SHBD-set'. Kak vladel'tsy nerazreshennykh besprovodnykh RES skryvayut svoi ustroystva [Masking SHBD network. How owners of unauthorized wireless RPS hide their devices], *Radiochastotnyy spektr* [Radio Frequency Spectrum], 2017, No. 8, pp. 44-46.

19. Postanovlenie Pravitel'stva Rossiyskoy Federatsii ot 31.12.2021 № 2607 «Ob utverzhdenii Pravil okazaniya telematicheskikh uslug svyazi» [Resolution of the Government of the Russian Federation No. 2607 dated 31.12.2021 "On Approval of the Rules for the Provision of Telematic Communication Services"].
20. *Zharov A.A.* Publichnyy Wi-Fi vykhodit iz teni [Public Wi-Fi comes out of the shadows], *Radiochastotnyy spektr* [Radio Frequency Spectrum], 2016, No. 10, pp. 15-16.
21. Postanovlenie Pravitel'stva RF ot 14 maya 2014 g. № 434 «O radiochastotnoy sluzhbe» [Resolution of the Government of the Russian Federation of May 14, 2014, No. 434 "On Radio Frequency Service"].
22. Federal'nyy zakon ot 27.07.2006 № 149-FZ «Ob informatsii, informatsionnykh tekhnologiyakh i o zashchite informatsii» [Federal Law No. 149-FL dated 27.07.2006 "On Information, Information Technologies and Information Protection"].

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

Дьяков Антон Владимирович – Южный федеральный университет; e-mail: adyakov@sfedu.ru; г. Таганрог, Россия; аспирант.

Румянцев Константин Евгеньевич – e-mail: rke2004@mail.ru; тел.: +78634371902; кафедра информационной безопасности телекоммуникационных систем; зав. кафедрой; д.т.н.; профессор.

Dyakov Anton Vladimirovich – Southern Federal University; e-mail: adyakov@sfedu.ru; Taganrog, Russia; graduate student.

Rumyantsev Konstantin Evgenyevich – e-mail: rke2004@mail.ru; phone: +78634371902; the department of information security of telecommunication systems; head of department; dr. of eng.sc.; professor.

Раздел II. Электроника, нанотехнологии и приборостроение

УДК 629.7.02(083.3)

DOI 10.18522/2311-3103-2024-2-142-152

И.И. Дементьев, А.О. Костина

ОБ АНАЛИТИЧЕСКОЙ НЕОПРЕДЕЛЕННОСТИ КОМПОНЕНТ ТЕНЗОРА ЭНЕРГЕТИЧЕСКИХ КОЭФФИЦИЕНТОВ ЭЛЕКТРОМЕХАНИЧЕСКОЙ СВЯЗИ ДЛЯ ПЛЕНОЧНЫХ ПОЛИМЕРНЫХ ПЬЕЗОМАТЕРИАЛОВ

В настоящее время пьезоэлектрические материалы на основе пленок, например, из поливинилиденфторидов, нашли широкое применение в разных отраслях промышленности. Интерес к этим материалам обоснован, прежде всего, выгодными, по сравнению с материалами на основе пьезокерамических составов, величинами плотности пленочного материала, а также величинами эффективных пьезоэлектрических коэффициентов, что позволяет реализовать в изделиях новые принципы нагружения с целью преобразования механической энергии в электрическую энергию. Кроме того, меньшая жесткость пленочных полимерных материалов, по сравнению с жесткостью чувствительных элементов электроакустических преобразователей на основе пьезокерамических составов, обеспечивает возможность создания конструкций преобразователей большой площади или собранных из этих преобразователей апертур антенн, повторяющих обводы корпусов носителей. В статье описаны перспективы внедрения пленочных полимерных пьезоматериалов в изделия гидроакустической и электрогенераторной техники. Очевидно, что внедрение новых материалов в гидроакустическую и пьезогенераторную технику требует переосмысления методик расчетов характеристик изделий на основе пленочных пьезоактивных материалов. Одной из основных характеристик электроакустических преобразователей, применяющихся, как в гидроакустической технике, так и в изделиях генерации электрической энергии за счет использования волнения поверхности моря или океана, является энергетический коэффициент электромеханической связи. Указанный коэффициент является тензорной физической величиной, а, соответственно, для анизотропных материалов, которыми являются пьезоактивные пленки, например, из поливинилиденфторидов, необходим математический аппарат, который позволял бы определять компоненты тензора энергетического коэффициента электромеханической связи в ходе проектирования изделий гидроакустической и пьезогенераторной техники. Попытка разработки такого математического аппарата предпринята авторами статьи, но в ходе выполненных работ установлена аналитическая неопределенность компонент тензора энергетического коэффициента электромеханической связи. В статье представлено научное обоснование выявленной аналитической неопределенности компонент тензора энергетического коэффициента электромеханической связи из уравнений, устанавливающих зависимости указанных компонент от коэффициентов, характеризующих механические и пьезоэлектрические свойства пленочных материалов с учетом их анизотропии.

Гидроакустическая и электрогенераторная техника; пленочный полимерный пьезоэлектрический материал; анизотропия физико-механических и пьезоэлектрических свойств; компоненты тензора энергетических коэффициентов электромеханической связи; аналитическая неопределенность.

I.I. Dementev, A.O. Kostina

ABOUT ANALYTICAL UNCERTAINTY THE COMPONENT OF THE TENSOR OF POWER ELECTROMECHANICAL FORCE FACTORS FOR FILM POLYMERIC PIEZO MATERIAL

Currently, piezoelectric materials based on films, for example, polyvinylidene fluorides, have found wide application in various industries. The interest in these materials is justified, first of all, by the advantageous, in comparison with materials based on piezoceramic compositions, values of the density of the film material, as well as the values of the effective piezoelectric coefficients, which makes it possible to

implement new principles of loading in products in order to convert mechanical energy into electrical energy. In addition, the lower rigidity of film polymer materials, compared to the rigidity of sensitive elements of electroacoustic transducers based on piezoceramic compositions, makes it possible to create designs of large-area transducers or antenna apertures assembled from these transducers, repeating the contours of carrier housings. The article describes the prospects for the introduction of film polymer piezomaterials into hydroacoustic and electric generating equipment products. Obviously, the introduction of new materials into hydroacoustic and piezoelectric generator technology requires a rethinking of methods for calculating the characteristics of products based on film piezoactive materials. One of the main characteristics of electroacoustic transducers, used both in hydroacoustic technology and in products for generating electrical energy through the use of sea or ocean surface waves, is the energy coefficient of electromechanical coupling. The specified coefficient is a tensor physical quantity, and, accordingly, for anisotropic materials, which are piezoactive films, for example, from polyvinylidene fluorides, a mathematical apparatus is needed that would allow determining the components of the tensor of the energy coefficient of electromechanical coupling during the design of hydroacoustic and piezoelectric generator products. An attempt to develop such a mathematical apparatus was made by the authors of the article, but in the course of the work performed, the analytical uncertainty of the components of the tensor of the energy coefficient of electromechanical coupling was established. The article presents a scientific substantiation of the identified analytical uncertainty of the components of the tensor of the energy coefficient of electromechanical coupling from equations that establish the dependence of these components on coefficients characterizing the mechanical and piezoelectric properties of film materials, taking into account their anisotropy.

Hydroacoustic and electrogenerating equipment; film polymeric piezoelectric material; anisotropy of physicomachanical and piezoelectric properties; components of the tensor of power electromechanical force factors; analytical uncertainty.

Введение. На сегодняшний день пленочные полимерные материалы находят широкое применение в медицинской [1, 2], нефтехимической [3, 4], приборостроительной [5, 6], оборонной [7, 8] и других отраслях промышленности [9, 10].

Исследования свойств пленочных полимеров, проведенные японскими учеными, привели к открытию в 1969 году [11] пьезоэлектрического эффекта в пленках из поливинилиденфторида (ПВДФ), применявшегося в мире с начала 60-х годов для изготовления упаковочных материалов и защитных покрытий. Открытие пьезоэффекта в ПВДФ положило начало многочисленным работам в России и за рубежом по созданию чувствительных элементов на основе пьезоактивных пленок для приборов в интересах разных отраслей промышленности.

Выгодные, по сравнению с материалами на основе пьезокерамических составов, величины плотности пленок из ПВДФ заинтересовали разработчиков гидроакустической техники [7] и электрогенераторных устройств [12], принцип работы которых основан на использовании явления прямого пьезоэффекта. Указанное обстоятельство связано, прежде всего, с возможностью разрешения противоречия, заключающегося в необходимости, с одной стороны, создания средств с гибкими сплошными апертурами больших площадей и необходимости, с другой стороны, конструкторско-технологического обеспечения допустимых пороговых уровней по массе проектируемых средств. Технология сплошного заполнения крупногабаритных апертур чувствительными элементами позволит, применительно к гидроакустической антенной технике, повысить устойчивость к гидродинамическим помехам и, применительно к электрогенераторным устройствам, повысить их коэффициент полезного действия.

Внедрение преобразователей энергии на основе пленочных полимерных пьезоматериалов в конструкции гидроакустической и электрогенераторной техники требует переосмысления процессов ее проектирования. В работе [13] приведены методики и методы определения характеристик чувствительных элементов из пьезокерамики на основе химических составов, традиционно использующихся в отечественной гидроакустической технике, а также характеристик средств, в которых применены такие чувствительные элементы. Из указанных работ известно, что применяющиеся методики и методы адекватны для дискретных апертур гидроакустических средств. Таким образом, существует необходимость в разработке математического аппарата для расчетов характеристик пре-

образователей энергии, составляющих непрерывные нелинейные крупногабаритные апертуры, и характеристик изделий с указанными апертурами. Известны работы в области математического моделирования объектов на основе полимерных материалов, например, труды [6, 7, 14]. В указанных работах рассмотрены частные задачи механики деформируемого твердого тела для пьезоактивных конструкций с установленными под эти частные задачи ограничениями в виде граничных условий и принятых допущений. С целью создания теории расчетов электроакустических характеристик анизотропных конструкций произвольной геометрической конфигурации на основе полимерных пленок, обладающих пьезоэлектрическими свойствами, нами разработана методика математического моделирования напряженно-деформированного состояния таких конструкций [15], предложен математический аппарат для расчетов пьезоэлектрических характеристик и чувствительности пленочных электроакустических преобразователей для гидроакустической техники [16]. Важно отметить, что в указанных направлениях исследований существенные результаты получены авторами работ [17, 18]. Выполненные разработки позволили обосновать целесообразность внедрения в конструкции электроакустических преобразователей принципа всестороннего нагружения пьезоактивного материала из ПВДФ марки Ф-2МЭ на примере расчетов чувствительности преобразователей, основанных на использовании эффективного пьезоэлектрического модуля заряда. Применение принципа всестороннего нагружения пьезоматериала в составе электроакустических преобразователей имеет практическую значимость для создания гидрофонов, актуальность разработки которых обоснована в работе [19].

Коэффициент электромеханической связи. Определение и основные понятия.

В разрабатываемой нами теории расчетов электроакустических характеристик анизотропных конструкций произвольной геометрической конфигурации на основе полимерных пленок, обладающих пьезоэлектрическими свойствами, предусмотрен математический аппарат для определения энергетических коэффициентов электромеханической связи (КЭМС) преобразователей. Указанные коэффициенты являются мерой эффективности пьезоматериала, с точки зрения, преобразования механической энергии в электрическую энергию при прямом пьезоэффекте или, наоборот, электрической энергии в механическую энергию при обратном пьезоэффекте. Значения энергетических КЭМС отличаются друг от друга в разных направлениях системы координат, принятых для математического моделирования пьезоактивной конструкции, что обосновано анизотропией ее физикомеханических и пьезоэлектрических свойств. При этом значения энергетических коэффициента электромеханической связи в определенном направлении зависят от ориентации электрического поля в материале и граничных условий крепления конструкции, влияющих на ее напряженно-деформированное состояние. Очевидно, что энергетические КЭМС являются тензорной физической величиной. По определению тензор энергетических коэффициентов электромеханической связи представляет собой отношение преобразованной электрической или механической энергии к полной энергии, сообщенной пьезоматериалу. Физический смысл тензора энергетических КЭМС определяется через скалярное произведение двух одинаковых тензоров энергетических коэффициентов электромеханической связи, аналогично тому, как через квадрат энергетического КЭМС определяется физический смысл указанного коэффициента в одном из направлений, заданных в материале. Заключается физический смысл тензора энергетических КЭМС в том, что он отражает отношение генерируемой в пьезоэлектрическом материале механической энергии к полной запасаемой в материале электрической энергии или отношение генерируемой в пьезоэлектрическом материале электрической энергии при его деформациях к полной запасаемой в материале механической энергии. Исходя из определения и физического смысла тензора энергетических КЭМС, он эквивалентен коэффициенту полезного действия преобразователей энергии, поэтому исследования тензора энергетических коэффициентов электромеханической связи имеют практическую значимость, например, для повышения электрической мощности энергии, вырабатываемой электрогенераторными устройствами, или повышения чувствительности приемных элементов апертур гидроакустических антенн. Известно, что плотность полной энергии пьезомате-

риала при заданной напряженности электрического поля пропорциональна абсолютной диэлектрической проницаемости при отсутствии механических напряжений. При этом в материале, на который действует электрическое поле с такой же напряженностью, плотность полной энергии пропорциональна абсолютной диэлектрической проницаемости при отсутствии деформаций. Разность указанных плотностей энергии соответствует плотности механической энергии пьезоматериала, обусловленной его деформациями [20], что математически с учетом физического смысла тензора энергетических КЭМС можно записать в виде следующей формулы

$$\mathbf{k}_{ij} \cdot \mathbf{k}_{ij} = (\xi_{ij}^{\sigma} - \xi_{ij}^{\varepsilon}) \cdot [\xi_{ij}^{\sigma}]^{-1} = \mathbf{E}_{ij} - \xi_{ij}^{\varepsilon} \cdot [\xi_{ij}^{\sigma}]^{-1}, \quad (1)$$

где \mathbf{k}_{ij} – тензор энергетических КЭМС; \mathbf{E}_{ij} – единичный тензор; ξ_{ij}^{ε} – тензор абсолютной диэлектрической проницаемости при отсутствии деформаций материала; ξ_{ij}^{σ} – тензор абсолютной диэлектрической проницаемости материала, свободного от механических напряжений; $[\]^{-1}$ – обозначение обратного тензора; $i = 1 \dots 3; j = 1 \dots 3$.

Из уравнения (1) очевидно, что тензор энергетических коэффициентов электромеханической связи является тензором второго ранга.

Из работы [20] известны следующие математические зависимости:

$$\begin{aligned} \xi_k^{\varepsilon} \cdot [\xi_{ij}^{\sigma}]_j^{-1} &= \mathbf{S}_{kl}^{\mathbf{D}} \cdot [\mathbf{S}_{kl}^{\mathbf{E}}]^{-1}; \\ \xi_{ij}^{\varepsilon} \cdot [\xi_{ij}^{\sigma}]^{-1} &= \mathbf{E}_{ij} - \mathbf{d}_{ik} \cdot \left([\mathbf{S}_{kl}^{\mathbf{E}}]^{-1} \cdot \mathbf{d}_{ik}^{\mathbf{T}} \right) \cdot [\xi_{ij}^{\sigma}]^{-1}, \end{aligned} \quad (2)$$

где $\mathbf{S}_{ij}^{\mathbf{E}}$ – тензор упругой податливости материала при отсутствии воздействий электрического поля; $\mathbf{S}_{ij}^{\mathbf{D}}$ – тензор упругой податливости неполяризованного материала; \mathbf{d}_{ik} , $\mathbf{d}_{ik}^{\mathbf{T}}$ – тензор и транспонированный тензор пьезоэлектрических коэффициентов заряда в размерности $\frac{\text{Кл}}{\text{Н}}$; $k = 1 \dots 6; l = 1 \dots 6$.

Применяя математические зависимости (2) в формуле (1), получим следующее уравнение

$$\mathbf{k}_{ij} \cdot \mathbf{k}_{ij} = \mathbf{d}_{ik} \cdot \left([\mathbf{S}_{kl}^{\mathbf{E}}]^{-1} \cdot \mathbf{d}_{ik}^{\mathbf{T}} \right) \cdot [\xi_{ij}^{\sigma}]^{-1}. \quad (3)$$

Уравнение (3) позволяет математически описать каждую компоненту тензора энергетических КЭМС. Например, для определения компонент k_{31} и k_{33} , являющихся главными энергетическими коэффициентами электромеханической связи для электроакустических преобразователей из пьезокерамики на основе химических составов, традиционно используемых в отечественной гидроакустической технике, уравнение (3) будет иметь следующий вид:

$$k_{31}^2 = \frac{d_{31}^2}{S_{11}^{\mathbf{E}} \cdot \xi_{33}^{\sigma}}; \quad k_{33}^2 = \frac{d_{33}^2}{S_{33}^{\mathbf{E}} \cdot \xi_{33}^{\sigma}}. \quad (4)$$

Аналогично уравнениям (4) математически описываются энергетические КЭМС пьезоматериалов разных структур и классов. На сегодняшний день известны и широко применяются на практике результаты исследований [21, 22] физико-механических и пьезоэлектрических свойств кристаллических диэлектриков для определения их главных энергетических коэффициентов электромеханической связи. Достигнутый уровень результатов исследований [5, 23] структур полимерных пленочных материалов, обладающих пьезоэлектрическими свойствами, позволяет сделать вывод о том, что указанные материалы представляют практический интерес при их объемном нагружении для использования свойств прямого пьезоэффекта в гидроакустической и электрогенераторной технике. Таким образом, коэффициент полезного действия преобразователей энергии на основе полимерных пле-

ночных материалов одновременно зависит от каждой компоненты тензора энергетических КЭМС, а не только от энергетических коэффициентов электромеханической связи в определенных направлениях, называемых главными направлениями. При этом каждая компонента тензора энергетических КЭМС одновременно зависит от коэффициентов, характеризующих механические и пьезоэлектрические свойства анизотропного материала в разных направлениях. Определение компонент тензора энергетических коэффициентов электромеханической связи эмпирическими методами сопряжено с техническими трудностями [20, 21], а некоторые компоненты на сегодняшний день вообще не представляется возможным определить по причине технической нереализуемости экспериментально-измерительных установок. Кроме того, энергетические КЭМС, определенные эмпирическими методами, являются достоверными для конкретной структуры пленочных пьезоматериала. Принимая во внимание обширную номенклатуру пленочных материалов, обладающих пьезоэлектрическими свойствами, процесс проектирования гидроакустических или электрогенераторных средств, в которых предполагается применение указанных материалов, осложняется необходимостью проведения экспериментов по определению энергетических КЭМС, что влечет за собой дополнительные временные и финансовые затраты. В этой связи, актуальной представляется задача разработки математического аппарата для выполнения аналитических расчетов компонент тензора энергетических коэффициентов электромеханической связи полимерных анизотропных пленочных пьезоматериалов.

Система уравнений для определения коэффициентов электромеханической связи, как инструмент математического аппарата. Основываясь на принципе непротиворечивости математических уравнений законам физики, целесообразно для определения энергетических КЭМС использовать уравнение (3), вывод которого выполнен исходя из определения и физического смысла энергетического коэффициента электромеханической связи. Представим уравнение (3), приведенное в тензорной форме, в виде следующей системы алгебраических нелинейных уравнений:

$$\begin{aligned}
 \sum_{\substack{j=1; \\ i=1}}^{m=3; \\ n=3} k_{1j} \cdot k_{i1} &= \sum_{\substack{i=1; \\ g=1}}^{n=3; \\ m=3} \beta_{i1}^{\sigma} \cdot \Lambda_g; & \sum_{\substack{j=1; \\ i=1}}^{m=3; \\ n=3} k_{2j} \cdot k_{i2} &= \sum_{\substack{i=1; \\ g=4}}^{n=3; \\ m=6} \beta_{i2}^{\sigma} \cdot \Lambda_g; \\
 \sum_{\substack{j=1; \\ i=1}}^{m=3; \\ n=3} k_{3j} \cdot k_{i3} &= \sum_{\substack{i=1; \\ g=7}}^{n=3; \\ m=9} \beta_{i3}^{\sigma} \cdot \Lambda_g; & \sum_{\substack{j=1; \\ i=1}}^{m=3; \\ n=3} k_{1j} \cdot k_{i2} &= \sum_{\substack{i=1; \\ g=1}}^{n=3; \\ m=3} \beta_{i2}^{\sigma} \cdot \Lambda_g; \\
 \sum_{\substack{j=1; \\ i=1}}^{m=3; \\ n=3} k_{2j} \cdot k_{i1} &= \sum_{\substack{i=1; \\ g=4}}^{n=3; \\ m=6} \beta_{i1}^{\sigma} \cdot \Lambda_g; & \sum_{\substack{j=1; \\ i=1}}^{m=3; \\ n=3} k_{1j} \cdot k_{i3} &= \sum_{\substack{i=1; \\ g=1}}^{n=3; \\ m=3} \beta_{i3}^{\sigma} \cdot \Lambda_g; \\
 \sum_{\substack{j=1; \\ i=1}}^{m=3; \\ n=3} k_{3j} \cdot k_{i1} &= \sum_{\substack{i=1; \\ g=7}}^{n=3; \\ m=9} \beta_{i1}^{\sigma} \cdot \Lambda_g; & \sum_{\substack{j=1; \\ i=1}}^{m=3; \\ n=3} k_{2j} \cdot k_{i3} &= \sum_{\substack{i=1; \\ g=4}}^{n=3; \\ m=6} \beta_{i3}^{\sigma} \cdot \Lambda_g; \\
 & & \sum_{\substack{j=1; \\ i=1}}^{m=3; \\ n=3} k_{3j} \cdot k_{i2} &= \sum_{\substack{i=1; \\ g=7}}^{n=3; \\ m=9} \beta_{i2}^{\sigma} \cdot \Lambda_g,
 \end{aligned} \tag{5}$$

где β_{ij}^{σ} – коэффициенты, обратные абсолютной диэлектрической проницаемости свободного от механических напряжений материала; Λ_g ($g = 1 \dots 9$) – коэффициенты, которые определяются пьезоэлектрическими и физико-механическими свойствами материала в соответствии со следующими математическими зависимостями:

$$\begin{aligned}
 \Lambda_1 &= \sum_{k=1}^{n=6} \left(d_{1k} \cdot \sum_{l=1}^{m=6} C_{kl}^E \cdot d_{1l} \right); & \Lambda_2 &= \sum_{k=1}^{n=6} \left(d_{1k} \cdot \sum_{l=1}^{m=6} C_{kl}^E \cdot d_{2l} \right); \\
 \Lambda_3 &= \sum_{k=1}^{n=6} \left(d_{1k} \cdot \sum_{l=1}^{m=6} C_{kl}^E \cdot d_{3l} \right); & \Lambda_4 &= \sum_{k=1}^{n=6} \left(d_{2k} \cdot \sum_{l=1}^{m=6} C_{kl}^E \cdot d_{1l} \right); \\
 \Lambda_5 &= \sum_{k=1}^{n=6} \left(d_{2k} \cdot \sum_{l=1}^{m=6} C_{kl}^E \cdot d_{2l} \right); & \Lambda_6 &= \sum_{k=1}^{n=6} \left(d_{2k} \cdot \sum_{l=1}^{m=6} C_{kl}^E \cdot d_{3l} \right); \\
 \Lambda_7 &= \sum_{k=1}^{n=6} \left(d_{3k} \cdot \sum_{l=1}^{m=6} C_{kl}^E \cdot d_{1l} \right); & \Lambda_8 &= \sum_{k=1}^{n=6} \left(d_{3k} \cdot \sum_{l=1}^{m=6} C_{kl}^E \cdot d_{2l} \right); \\
 \Lambda_9 &= \sum_{k=1}^{n=6} \left(d_{3k} \cdot \sum_{l=1}^{m=6} C_{kl}^E \cdot d_{3l} \right),
 \end{aligned} \tag{6}$$

где C_{kl}^E – коэффициенты жесткости материала при отсутствии воздействий электрического поля.

Математический вывод формул для аналитического определения энергетических КЭМС с использованием системы алгебраических нелинейных уравнений (5) заключается в решении этой системы уравнений относительно искомых величин k_{ij} при заданных коэффициентах β_{ij}^σ и Λ_g . Аналогичный подход [24] используется, например, для преобразования уравнений обобщенного закона Гука из прямой математической формы, когда компоненты тензора относительных деформаций выражены через компоненты тензора механических напряжений, в обратную математическую форму, когда компоненты тензора напряжений выражены через компоненты тензора деформаций. Однако, в отличие от тензоров механических напряжений и относительных деформаций, объем теоретической и эмпирической информации о свойствах тензора энергетических КЭМС, как объекта математической физики, существенно ограничен [5, 20, 25]. Так, например, известно [24], что тензоры напряжений и деформаций являются симметричными тензорами второго ранга. Таким образом, полная информация о напряженно-деформированном состоянии материала представляется не девятью компонентами каждого из указанных тензоров, а шестью их компонентами, что существенно облегчает задачу преобразования уравнений обобщенного закона Гука из прямой математической формы в обратную математическую форму и, наоборот, из обратной формы в прямую форму. Напротив, исходя из определения и физического смысла тензора энергетических КЭМС, нельзя однозначно утверждать об его симметричности. Представляется возможным записать уравнения для определения главных энергетических коэффициентов электромеханической связи, например, для прямоугольного объекта, находящегося в условиях растягивающих и сжимающих внешних нагрузок, предполагая при этом попеременные изменения направлений механических воздействий и главных направлений в моделируемом объекте. Указанные уравнения имеют следующий вид:

$$\begin{aligned}
 k_{11} &= \frac{d_{11}}{\sqrt{S_{11}^E \cdot \xi_{11}^\sigma}}; & k_{12} &= \frac{d_{12}}{\sqrt{S_{22}^E \cdot \xi_{11}^\sigma}}; & k_{13} &= \frac{d_{13}}{\sqrt{S_{33}^E \cdot \xi_{11}^\sigma}}; \\
 k_{21} &= \frac{d_{21}}{\sqrt{S_{11}^E \cdot \xi_{22}^\sigma}}; & k_{22} &= \frac{d_{22}}{\sqrt{S_{22}^E \cdot \xi_{22}^\sigma}}; & k_{23} &= \frac{d_{23}}{\sqrt{S_{33}^E \cdot \xi_{22}^\sigma}}; \\
 k_{31} &= \frac{d_{31}}{\sqrt{S_{11}^E \cdot \xi_{33}^\sigma}}; & k_{32} &= \frac{d_{32}}{\sqrt{S_{22}^E \cdot \xi_{33}^\sigma}}; & k_{33} &= \frac{d_{33}}{\sqrt{S_{33}^E \cdot \xi_{33}^\sigma}}.
 \end{aligned} \tag{7}$$

При анализе уравнений (7) очевидны неравенства $k_{12} \neq k_{21}$, $k_{13} \neq k_{31}$, $k_{23} \neq k_{32}$, обоснованные условиями симметричности тензоров пьезоэлектрических коэффициентов заряда, упругой податливости и абсолютной диэлектрической проницаемости анизотропных материалов, согласно которым компоненты указанных тензоров $d_{12} \neq d_{21}$, $d_{13} \neq d_{31}$, $d_{23} \neq d_{32}$, $S_{22}^E \neq S_{11}^E$, $S_{33}^E \neq S_{11}^E$, $S_{33}^E \neq S_{22}^E$, $\xi_{11}^\sigma \neq \xi_{22}^\sigma$, $\xi_{11}^\sigma \neq \xi_{33}^\sigma$, $\xi_{22}^\sigma \neq \xi_{33}^\sigma$. Таким образом, можно сделать вывод о том, что тензор энергетических КЭМС является несимметричным тензором второго ранга.

Решение системы уравнений для определения коэффициентов электромеханической связи. Основные выкладки. Возвращаясь к решению системы алгебраических нелинейных уравнений (5) относительно искомым величин k_{ij} при известных коэффициентах β_{ij}^σ и Λ_g и принимая во внимание несимметричность тензора энергетических коэффициентов электромеханической связи, целесообразно отметить, что указанная система уравнений является статически определимой, так как имеет девять уравнений при девяти искомым величинах. Аналитическое решение системы уравнений (5) должно представлять собой девять уравнений, однозначно определяющих девять компонент тензора энергетических КЭМС через коэффициенты β_{ij}^σ и Λ_g . В результате аналитического вывода указанных уравнений получены следующие формулы:

$$\begin{aligned} k_{11} &= \sqrt{\beta_{11}^\sigma \cdot \Lambda_1 + \beta_{21}^\sigma \cdot \Lambda_2 + \beta_{31}^\sigma \cdot \Lambda_3 - k_{12} \cdot k_{21} - k_{13} \cdot k_{31}}; \\ k_{22} &= \sqrt{\beta_{12}^\sigma \cdot \Lambda_4 + \beta_{22}^\sigma \cdot \Lambda_5 + \beta_{32}^\sigma \cdot \Lambda_6 - k_{21} \cdot k_{12} - k_{23} \cdot k_{32}}; \\ k_{33} &= \sqrt{\beta_{13}^\sigma \cdot \Lambda_7 + \beta_{23}^\sigma \cdot \Lambda_8 + \beta_{33}^\sigma \cdot \Lambda_9 - k_{31} \cdot k_{13} - k_{32} \cdot k_{23}}; \\ k_{23} &= \frac{1}{k_{32}} \cdot \left(- \left[\frac{\beta_{12}^\sigma \cdot \Lambda_4 + \beta_{22}^\sigma \cdot \Lambda_5 + \beta_{32}^\sigma \cdot \Lambda_6 - k_{21} \cdot k_{12} - \frac{\beta_{12}^\sigma \cdot \Lambda_1 + \beta_{22}^\sigma \cdot \Lambda_2 + \beta_{32}^\sigma \cdot \Lambda_3 - k_{13} \cdot k_{32}}{k_{12}}}{-\sqrt{\beta_{11}^\sigma \cdot \Lambda_1 + \beta_{21}^\sigma \cdot \Lambda_2 + \beta_{31}^\sigma \cdot \Lambda_3 - k_{12} \cdot k_{21} - k_{13} \cdot k_{31}}} \right]^2 \right). \end{aligned} \quad (8)$$

Однако при последующем решении системы уравнений (5) относительно энергетических коэффициентов электромеханической связи k_{12} , k_{13} , k_{21} , k_{31} или k_{32} удается получить только квадратные уравнения. Например, уравнение относительно коэффициента k_{32} имеет следующий вид

$$k_{13} \cdot (k_{21} \cdot k_{12} + k_{31} \cdot k_{13}) \cdot k_{32}^2 + k_{12}^2 \cdot f_3 \cdot k_{32} - k_{12}^2 \cdot k_{31} \cdot (f_1 + f_2) = 0, \quad (9)$$

где символами f_g ($g = 1 \dots 3$) обозначены следующие математические выражения:

$$\begin{aligned} f_1 &= \frac{k_{12}^2 \cdot (\beta_{12}^\sigma \cdot \Lambda_4 + \beta_{22}^\sigma \cdot \Lambda_5 + \beta_{32}^\sigma \cdot \Lambda_6 - k_{21} \cdot k_{12}) - (\beta_{12}^\sigma)^2 \cdot \Lambda_1^2 - (\beta_{22}^\sigma)^2 \cdot \Lambda_2^2 +}{k_{12}^2} + \\ &+ \frac{-(\beta_{32}^\sigma)^2 \cdot \Lambda_3^2 - 2 \cdot (\beta_{12}^\sigma \cdot \Lambda_1 \cdot \beta_{22}^\sigma \cdot \Lambda_2 + \beta_{12}^\sigma \cdot \Lambda_1 \cdot \beta_{32}^\sigma \cdot \Lambda_3 + \beta_{22}^\sigma \cdot \Lambda_2 \cdot \beta_{32}^\sigma \cdot \Lambda_3)}{k_{12}^2}; \\ f_2 &= k_{12} \cdot k_{21} + k_{13} \cdot k_{31} - \beta_{11}^\sigma \cdot \Lambda_1 - \beta_{21}^\sigma \cdot \Lambda_2 - \beta_{31}^\sigma \cdot \Lambda_3 + \\ &+ 2 \cdot \left(\frac{\beta_{12}^\sigma \cdot \Lambda_1 + \beta_{22}^\sigma \cdot \Lambda_2 + \beta_{32}^\sigma \cdot \Lambda_3}{k_{12}} \right) \cdot \sqrt{\beta_{11}^\sigma \cdot \Lambda_1 + \beta_{21}^\sigma \cdot \Lambda_2 + \beta_{31}^\sigma \cdot \Lambda_3 - k_{12} \cdot k_{21} - k_{13} \cdot k_{31}}; \end{aligned} \quad (10)$$

$$f_3 = \beta_{11}^\sigma \cdot \Lambda_4 + \beta_{21}^\sigma \cdot \Lambda_5 + \beta_{31}^\sigma \cdot \Lambda_6 - \frac{k_{21}}{k_{12}} \cdot (\beta_{12}^\sigma \cdot \Lambda_1 + \beta_{22}^\sigma \cdot \Lambda_2 + \beta_{32}^\sigma \cdot \Lambda_3) -$$

$$-2 \cdot k_{31} \cdot \left(\frac{k_{13} \cdot [\beta_{12}^\sigma \cdot \Lambda_1 + \beta_{22}^\sigma \cdot \Lambda_2 + \beta_{32}^\sigma \cdot \Lambda_3]}{k_{12}^2} - \frac{k_{13}}{k_{12}} \cdot \sqrt{\beta_{11}^\sigma \cdot \Lambda_1 + \beta_{21}^\sigma \cdot \Lambda_2 + \beta_{31}^\sigma \cdot \Lambda_3 - k_{12} \cdot k_{21} - k_{13} \cdot k_{31}} \right).$$

Очевидно, что вывод формулы для определения коэффициента k_{32} из квадратного уравнения (9) сопряжен с необходимостью анализа дискриминанта, имеющего следующий вид

$$D = 4 \cdot k_{13} \cdot k_{31} \cdot (k_{21} \cdot k_{12} + k_{31} \cdot k_{13}) \cdot (k_{12}^2 \cdot [q_4 - q_3] - 2 \cdot q_6 - q_5) +$$

$$+ k_{12} \cdot \left(k_{12}^3 \cdot q_1 + q_2 \cdot \left[8 \cdot k_{13} \cdot k_{31} \cdot (k_{21} \cdot k_{12} + k_{31} \cdot k_{13}) \cdot \sqrt{q_3 - k_{12} \cdot k_{21} - k_{13} \cdot k_{31}} - \right. \right. \\ \left. \left. - k_{12} \cdot (k_{21} \cdot k_{12} + 2 \cdot k_{31} \cdot k_{13}) \right] + \right. \\ \left. + 2 \cdot k_{31} \cdot k_{13} \cdot k_{12} \cdot \left[k_{12} \cdot \sqrt{q_3 - k_{12} \cdot k_{21} - k_{13} \cdot k_{31}} + \right. \right. \\ \left. \left. + 2 \cdot k_{31} \cdot k_{13} \cdot (k_{21} \cdot k_{12} + k_{31} \cdot k_{13}) \right] \right), \quad (11)$$

где символами q_g ($g = 1 \dots 6$) обозначены следующие математические выражения:

$$q_1 = \beta_{11}^\sigma \cdot \Lambda_4 + \beta_{21}^\sigma \cdot \Lambda_5 + \beta_{31}^\sigma \cdot \Lambda_6; \quad q_2 = \beta_{12}^\sigma \cdot \Lambda_1 + \beta_{22}^\sigma \cdot \Lambda_2 + \beta_{32}^\sigma \cdot \Lambda_3;$$

$$q_3 = \beta_{11}^\sigma \cdot \Lambda_1 + \beta_{21}^\sigma \cdot \Lambda_2 + \beta_{31}^\sigma \cdot \Lambda_3; \quad q_4 = \beta_{12}^\sigma \cdot \Lambda_4 + \beta_{22}^\sigma \cdot \Lambda_5 + \beta_{32}^\sigma \cdot \Lambda_6;$$

$$q_5 = (\beta_{12}^\sigma)^2 \cdot \Lambda_1^2 + (\beta_{22}^\sigma)^2 \cdot \Lambda_2^2 + (\beta_{32}^\sigma)^2 \cdot \Lambda_3^2;$$

$$q_6 = \beta_{12}^\sigma \cdot \Lambda_1 \cdot \beta_{22}^\sigma \cdot \Lambda_2 + \beta_{12}^\sigma \cdot \Lambda_1 \cdot \beta_{32}^\sigma \cdot \Lambda_3 + \beta_{22}^\sigma \cdot \Lambda_2 \cdot \beta_{32}^\sigma \cdot \Lambda_3. \quad (12)$$

Обоснование аналитической неопределенности коэффициентов электромеханической связи. Исходя из определения и физического смысла энергетического КЭМС, уравнение (9) может иметь только вещественные корни, поэтому корректно для формулы (11) задать ограничение в виде неравенства $D \geq 0$. Однако при последующем аналитическом выводе формулы для определения коэффициента k_{32} в общей постановке получаем неопределенность в виде возможных двух корней квадратного уравнения (9), так как его дискриминант, что показано формулой (11), при аналитическом выводе не обращается в нуль, и, следовательно, отсутствует основание утверждать о существовании единственного корня уравнения (9). Таким образом, выполнить аналитический вывод единственной формулы для однозначного определения коэффициента k_{32} через известные коэффициенты β_{ij}^σ и Λ_g не представляется возможным. Следовательно, дальнейшее решение системы уравнений (5) с учетом формул (8) относительно коэффициентов k_{12} , k_{13} , k_{21} и k_{31} , при наличии двух формул для определения коэффициента k_{32} , как корней уравнения (9), не приведет к ожидаемому результату в виде девяти уравнений, однозначно определяющих девять компонент тензора энергетических КЭМС через коэффициенты β_{ij}^σ и Λ_g , что свидетельствует об аналитической неопределенности компонент тензора энергетических коэффициентов электромеханической связи из системы уравнений (5), устанавливающих зависимости энергетических КЭМС от коэффициентов, характеризующих механические и пьезоэлектрические свойства материалов с учетом их анизотропии.

Закключение. Подводя итог, целесообразно отметить, что с внедрением в гидроакустическую и электрогенераторную технику электроакустических преобразователей на основе полимерных пленок, обладающих пьезоэлектрическими свойствами, остро стоит проблема разработки теории расчетов электрических и акустических характеристик ани-

зотропных конструкций, в том числе их энергетических коэффициентов электромеханической связи. Наличие зависимостей каждой компоненты тензора энергетических КЭМС от коэффициентов, характеризующих механические и пьезоэлектрические свойства анизотропных материалов в разных направлениях, подчеркивает актуальность вопроса разработки математического аппарата для выполнения аналитических расчетов компонент тензора энергетических коэффициентов электромеханической связи полимерных пленочных пьезоматериалов. Более того, научно обоснованная в статье аналитическая неопределенность компонент тензора энергетических КЭМС из уравнений, непротиворечивость которых законам физики известна, свидетельствует о необходимости поиска новых путей и подходов к выполнению расчетов указанных коэффициентов применительно к новым для гидроакустической и электрогенераторной техники материалам.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Бежин А.И., Должиков А.А., Жуковский В.А., Нетьяга А.А., Плотников Р.В.* Экспериментальное обоснование применения новых поливинилиденфторидных эндопротезов с карбиновым покрытием для герниопластики // Вестник новых медицинских технологий. – 2007. – Т. XIV, № 1. – С. 99-101.
2. *Subochev P., Prudnikov M., Vorobyev V., Postnikova A., Sergeev E., Perekatova V., Orlova A., Kotomina V., Turchin I.* Wideband linear detector arrays for optoacoustic imaging based on polyvinylidene difluoride films // Journal of Biomedical Optics. – 2018. – Vol. 23 (9), 091408. – DOI: 10.1117/1.JBO.23.9.091408.
3. *Игнатик А.А.* Оценка напряженно-деформированного состояния и прочности трехслойного образца трубы из полипропилена с трещиноподобным дефектом // Оборудование и технологии для нефтегазового комплекса. – 2023. – № 2 (134). – С. 68-75.
4. *Anokhina T., Borisov I., Yushkin A., Vaganov G., Didenko A., Volkov A.* Phase separation within a thin layer of polymer solution as prompt technique to predict membrane morphology and transport properties // Polymers. – 2020. – Vol. 12, 2785. – DOI: 10.3390/polym12122785.
5. *Kim K.-B., Lee J.* A lead-free piezoelectric fiber generator with a high energy conversion constant material // Energies. – 2022. – Vol. 15, No. 6787. – <https://doi.org/10.3390/en15186787>.
6. *Berlincourt D.A., Curran D.R., Jaffe H.* Piezoelectric and piezomagnetic materials and their function in transducers // Physical Acoustics. Principles and Methods. – Elsevier, 1964. – P. 169-270.
7. *Kharat D.K., Mitra S., Akhtar S., Kumar V.* Polymeric piezoelectric transducers for hydrophone applications // Defence Science Journal. – 2007. – Vol. 57, No. 1. – P. 7-22.
8. *Шатохин А.В., Полканов К.И., Селезнев И.А., Жуков В.Б.* Направления развития антенн отечественных гидроакустических средств // Национальная оборона. – 2020. – № 5 (170). – С. 104-109.
9. *Прокимов А.А., Джурицкий К.Б., Смирнова Ю.А.* Перспективные изоляционные материалы для радиочастотных кабелей и соединителей // Компоненты и технологии. – 2017. – № 2 (187). – С. 107-115.
10. *Абдуллин И.Ш., Фатхутдинов Р.Х., Миронова О.Ю., Шалыминова Д.П., Сайфутдинова И.Ф.* Исследование полимерных материалов на основе полиамида для изготовления мембранной защитной одежды // Вестник Казанского технологического ун-та. – 2012. – Т. 15, № 14. – С. 143-145.
11. *Kawai H.* The piezoelectricity of polyvinylidene fluoride // Japanese Journal of Applied Physics. – 1969. – Vol. 8. – P. 975-976.
12. *Chen S.E., Yang R.Y., Wu G.K., Wu C.C.* A piezoelectric wave-energy converter equipped with a geared-linkage-based frequency up-conversion mechanism // Sensors. – 2021. – No. 21. – P. 204.
13. *Смарышев М.Д., Добровольский Ю.Ю.* Гидроакустические антенны: справочник по расчету направленных свойств гидроакустических антенн. – Л.: Судостроение, 1984. – 304 с.
14. *Moffett M.B., Ricketts D., Butler J.L.* The effect of electrode stiffness on the piezoelectric and elastic constants of a piezoelectric bar // Acoustical Society of America. – 1988. – No. 83 (2). – P. 805-811.
15. *Дементьев И.И., Шабанов В.А., Шабанова Н.С.* Подход к математическому моделированию пленочных анизотропных конструкций электроакустических преобразователей // Научно-технический сборника «Гидроакустика / Hydroacoustics». – 2022. – Вып. 51 (3). – С. 42-49.
16. *Дементьев И.И., Шабанов В.А., Шабанова Н.С.* Методика расчета чувствительности пленочного пьезоэлектрического преобразователя гидроакустической антенны // Тр. Всерос. конф. «Прикладные технологии гидроакустики и гидрофизики». – СПб., 2023. – С. 194-199.
17. *Саркисян А.А., Саркисян С.О.* Собственные колебания микрополярных упругих гибких пластин и пологих оболочек // Акустический журнал. – 2022. – Т. 68, № 2. – С. 139-151.
18. *Шарфарец Б.П., Дмитриев С.П., Курочкин В.Е., Сергеев В.А.* О методе акустоэлектрического преобразования на основе электрокинетических явлений // Акустический журнал. – 2022. – Т. 68, № 5. – С. 571-578.

19. Исаев А.Е., Хатамтаев Б.И. Акустический центр измерительного гидрофона // Акустический журнал. – 2023. – Т. 69, № 1. – С. 63-72.
20. Поплавко Ю.М., Якименко Ю.И. Физические механизмы пьезоэлектричества. – Киев: Аверс, 1997. – 153 с.
21. Рудь Н.А. Пьезоэлектрические и сегнетоэлектрические свойства кристаллических диэлектриков: метод. указания. – Ярославль: Изд-во Яросл. гос. ун-та, 2003. – 44 с.
22. Семенова О.Р. Кристаллофизика: учеб. пособие. – Пермь: Изд-во Перм. гос. нац. исслед. ун-та, 2019. – 179 с.
23. Aabid A., Raheman Md.A., Ibrahim Y.E., Anjum A., Hrairi M., Parveez B., Parveen N., Zayan J.M. A systematic review of piezoelectric materials and energy harvesters for industrial applications // *Sensors*. – 2021. – Vol. 21, No. 12. – <https://doi.org/10.3390/s21124145>.
24. Жилин П.А. Рациональная механика сплошных сред: учеб. пособие / под ред. Е.А. Ивановой. – СПб.: Изд-во Политехн. ун-та, 2012. – 584 с.
25. Марченков Н.В. Рентгенодифракционные исследования пьезоэлектрических кристаллов при воздействии внешних электрических полей: дис. ... канд. физ.-мат. наук: 01.04.18. – М., 2014. – 122 с.

REFERENCES

1. Bezhin A.I., Dolzhikov A.A., Zhukovskiy V.A., Netyaga A.A., Plotnikov R.V. Eksperimental'noe obosnovanie primeneniya novykh polivinilidenftoridnykh endoprotezov s karbinovym pokrytiem dlya gernioplastiki [Experimental substantiation of the use of new polyvinylidene fluoride endoprostheses with carbine coating for hernioplasty], *Vestnik novykh meditsinskikh tekhnologiy* [Bulletin of new medical technologies], 2007, Vol. XIV, No. 1, pp. 99-101.
2. Subochev P., Prudnikov M., Vorobyev V., Postnikova A., Sergeev E., Perekatova V., Orlova A., Kotomina V., Turchin I. Wideband linear detector arrays for optoacoustic imaging based on polyvinylidene difluoride films, *Journal of Biomedical Optics*, 2018, Vol. 23 (9), 091408. DOI: 10.1117/1.JBO.23.9.091408.
3. Ignatik A.A. Otsenka napryazhenno-deformirovannogo sostoyaniya i prochnosti trekhsloynogo obraztsa truby iz polipropilena s treshchinopodobnym defektom [Assessment of the stress-strain state and strength of a three-layer polypropylene pipe sample with a crack-like defect], *Oborudovanie i tekhnologii dlya neftegazovogo kompleksa* [Equipment and technologies for the oil and gas complex], 2023, No. 2 (134), pp. 68-75.
4. Anokhina T., Borisov I., Yushkin A., Vaganov G., Didenko A., Volkov A. Phase separation within a thin layer of polymer solution as prompt technique to predict membrane morphology and transport properties, *Polymers*, 2020, Vol. 12, 2785. DOI: 10.3390/polym12122785.
5. Kim K.-B., Lee J. A lead-free piezoelectric fiber generator with a high energy conversion constant material, *Energies*, 2022, Vol. 15, No. 6787. Available at: <https://doi.org/10.3390/en15186787>.
6. Berlincourt D.A., Curran D.R., Jaffe H. Piezoelectric and piezomagnetic materials and their function in transducers, *Physical Acoustics. Principles and Methods*. Elsevier, 1964, pp. 169-270.
7. Kharat D.K., Mitra S., Akhtar S., Kumar V. Polymeric piezoelectric transducers for hydrophone applications, *Defence Science Journal*, 2007, Vol. 57, No. 1, pp. 7-22.
8. Shatokhin A.V., Polkanov K.I., Seleznev I.A., Zhukov V.B. Napravleniya razvitiya antenno otechestvennykh gidroakusticheskikh sredstv [Directions for the development of antennas for domestic sonar systems], *Natsional'naya oborona* [National Defense], 2020, No. 5 (170), pp. 104-109.
9. Prokimov A.A., Dzhurinskiy K.B., Smirnova Yu.A. Perspektivnye izolyatsionnye materialy dlya radiochastotnykh kabeley i soediniteley [Promising insulating materials for radio frequency cables and connectors], *Komponenty i tekhnologii* [Components and Technologies], 2017, No. 2 (187), pp. 107-115.
10. Abdullin I.Sh., Fatkhutdinov R.Kh., Mironova O.Yu., Shalyminova D.P., Sayfutdinova I.F. Issledovanie polimernykh materialov na osnove poliamida dlya izgotovleniya membrannoy zashchitnoy odezhdy [Study of polyamide-based polymer materials for the manufacture of membrane protective clothing], *Vestnik Kazanskogo tekhnologicheskogo un-ta* [Bulletin of the Kazan Technological University], 2012, Vol. 15, No. 14, pp. 143-145.
11. Kawai H. The piezoelectricity of polyvinilidene fluoride, *Japanese Journal of Applied Physics*, 1969, Vol. 8, pp. 975-976.
12. Chen S.E., Yang R.Y., Wu G.K., Wu C.C. A piezoelectric wave-energy converter equipped with a geared-linkage-based frequency up-conversion mechanism, *Sensors*, 2021, No. 21, pp. 204.
13. Smaryshev M.D., Dobrovol'skiy Yu.Yu. Gidroakusticheskie anteny: spravochnik po raschetu napravlennykh svoystv gidroakusticheskikh anten [Hydroacoustic antennas: a guide to calculating the directional properties of hydroacoustic antennas]. Leningrad: Sudostroenie, 1984, 304 p.

14. *Moffett M.B., Ricketts D., Butler J.L.* The effect of electrode stiffness on the piezoelectric and elastic constants of a piezoelectric bar, *Acoustical Society of America*, 1988, No. 83 (2), pp. 805-811.
15. *Dement'ev I.I., Shabanov V.A., Shabanova N.S.* Podkhod k matematicheskomu modelirovaniyu plenochnykh anizotropnykh konstruktivnykh elektroakusticheskikh preobrazovateley [An approach to mathematical modeling of film anisotropic structures of electroacoustic transducers], *Nauchno-tekhnicheskiiy sbornik «Gidroakustika / Hydroacoustics»* [Scientific and technical collection "Hydroacoustics"], 2022, Issue 51 (3), pp. 42-49.
16. *Dement'ev I.I., Shabanov V.A., Shabanova N.S.* Metodika rascheta chuvstvitel'nosti plenochnogo p'ezoelektricheskogo preobrazovatelya gidroakusticheskoy anteny [Methodology for calculating the sensitivity of a film piezoelectric transducer of a hydroacoustic antenna], *Tr. Vseros. konf. «Prikladnye tekhnologii gidroakustiki i gidrofiziki»* [Proceedings of the All-Russian Conference "Applied Technologies of Hydroacoustics and Hydrophysics"]. St. Petersburg, 2023, pp. 194-199.
17. *Sarkisyan A.A., Sarkisyan S.O.* Sobstvennye kolebaniya mikropolyarnykh uprugikh gibkikh plastin i pologikh obolochek [Natural vibrations of micropolar elastic flexible plates and flat shells], *Akusticheskiiy zhurnal* [Acoustic Journal], 2022, Vol. 68, No. 2, pp. 139-151.
18. *Sharfaretz B.P., Dmitriev S.P., Kurochkin V.E., Sergeev V.A.* O metode akustoelektricheskogo preobrazovaniya na osnove elektrokineticheskikh yavleniy [On the method of acoustoelectric transformation based on electrokinetic phenomena], *Akusticheskiiy zhurnal* [Acoustic Journal], 2022, Vol. 68, No. 5, pp. 571-578.
19. *Isaev A.E., Khatamtaev B.I.* Akusticheskiiy tsentr izmeritel'nogo gidrofona [Acoustic center of the measuring hydrophone], *Akusticheskiiy zhurnal* [Acoustic Journal], 2023, Vol. 69, No. 1, pp. 63-72.
20. *Poplavko Yu.M., Yakimenko Yu.I.* Fizicheskie mekhanizmy p'ezoelektrichestva [Physical mechanisms of piezoelectricity], Kiev: Avers, 1997, 153 p.
21. *Rud' N.A.* P'ezoelektricheskie i segnetoelektricheskie svoystva kristallicheskikh dielektrikov: metod. ukazaniya [Piezoelectric and ferroelectric properties of crystalline dielectrics: method. Instructions]. Yaroslavl': Izd-vo Yarosl. gos. un-ta, 2003, 44 p.
22. *Semenova O.R.* Kristallofizika: ucheb. posobie [Crystal physics: textbook]. Perm': Izd-vo Perm. gos. nats. issled. un-ta, 2019, 179 p.
23. *Aabid A., Raheman Md.A., Ibrahim Y.E., Anjum A., Hrairi M., Parveez B., Parveen N., Zayan J.M.* A systematic review of piezoelectric materials and energy harvesters for industrial applications, *Sensors*, 2021, Vol. 21, No. 12. Available at: <https://doi.org/10.3390/s21124145>.
24. *Zhilin P.A.* Ratsional'naya mekhanika sploshnykh sred: ucheb. posobie [Rational mechanics of continuous media: textbook], ed. by E.A. Ivanovoy. St. Petersburg: Izd-vo Politekhn. un-ta, 2012, 584 p.
25. *Marchenkov N.V.* Rentgenodifraktsionnye issledovaniya p'ezoelektricheskikh kristallov pri vozdeystvii vneshnikh elektricheskikh poley: dis. ... kand. fiz.-mat. nauk [X-ray diffraction studies of piezoelectric crystals under the influence of external electric fields: cand. of phys. and math. sc. dis.]: 01.04.18. Moscow, 2014, 122 p.

Статью рекомендовал к опубликованию д.т.н., профессор Б.Г. Коноплев.

Дементьев Илья Игоревич – АО «Концерн «Океанприбор»; e-mail: iidementev@mail.ru; г. Санкт-Петербург, Россия; тел.: 88124997568; к.т.н.; начальник учебно-методического центра.

Костина Анна Олеговна – e-mail: mail@oceanpribor.ru, тел.: 88124997400 (доб. 1819); инженер первой категории.

Dementiev Ilya Igorevich – JSC «Concern «Oceanpribor»; e-mail: iidementev@mail.ru; Saint-Petersburg, Russia; phone: +78124997568; cand. of eng. sc.; chief of Educational center.

Kostina Anna Olegovna – e-mail: mail@oceanpribor.ru; phone: +78124997400 (ext. 1819); 1st class engineer.

А.В. Лабынцев, Н.Е. Пономарев, А.Н. Саржанов, Д.В. Харланов

ХАРАКТЕРИСТИКИ ДВУХРЕЗОНАТОРНОГО МИКРОПОЛОСКОВОГО ЗВЕНА ГРЕБЕНЧАТОГО ТИПА С ОПТИМАЛЬНОЙ ДЛИНОЙ ОБЛАСТИ СВЯЗИ МЕЖДУ РЕЗОНАТОРАМИ

При проектировании СВЧ-устройств важную роль играет введение в топологию фильтра дополнительных конструктивных элементов или параметров, соответствующий выбор которых позволяет оптимизировать характеристики устройства по заданному критерию. В работе рассматривается способ расширения полосы заграждения микрополоскового фильтра гребенчатого типа за счет введения продольного смещения между резонаторами. Показано, что определенный выбор длины области связи между резонаторами позволяет минимизировать или свести к нулю связь между ними в окрестности утроенной средней частоты основной полосы пропускания, и тем самым существенно подавить паразитный провал затухания в этом диапазоне частот. В процессе проектирования фильтра был использован метод синтеза, основанный на переходе от фильтра к соответствующему $2n$ -полюсу, где n - количество резонаторов фильтра. Такой подход позволяет рассматривать фильтр как набор отдельных резонаторов, соединенных друг с другом, и ввести понятие потенциальной полосы пропускания, которая может быть легко определена по характеристикам $2n$ -полюса и расположена в желаемом интервале частотной оси с помощью легко формализуемого алгоритма выбора геометрических размеров фильтра. Электрическими характеристиками $2n$ -полюсника, используемого для синтеза устройства в основной полосе частот и на этапе оптимизации параметров устройства вблизи утроенного значения частот основной полосы частот, являются собственные проводимости резонаторов и проводимости соединения между ними. Целью работы является исследование влияния длины области связи между резонаторами двухрезонаторного микрополоскового гребенчатого фильтра на его электрические характеристики. Показано, что выбор оптимального значения длины области связи между резонаторами позволяет расширить полосу заграждения фильтра по уровню минус 30-35 дБ более чем в два раза за счет подавления паразитной полосы пропускания, формирующейся в окрестности утроенной средней частоты основной полосы пропускания. Указанный положительный эффект, заключающийся в расширении полосы заграждения фильтра, подтвержден результатами проектирования гребенчатого фильтра на двух смещенных резонаторах в программе HFSS.

Микрополосковый гребенчатый фильтр; продольное смещение резонаторов; паразитная полоса пропускания; полоса заграждения.

A.V. Labyntsev, N.E. Ponomarev, A.N. Sarzhanov, D.V. Kharlanov

CHARACTERISTICS OF A TWO-RESONATOR COMB-TYPE MICROSTRIP ELEMENT WITH AN OPTIMAL LENGTH OF THE COUPLING REGION BETWEEN THE RESONATORS

When designing microwave devices, an important role is played by the introduction of additional structural elements or parameters into the filter topology, the appropriate choice of which allows you to optimize the characteristics of the device according to a given criterion. The paper considers a method for expanding the barrier band of a comb-type microstrip filter by introducing a longitudinal displacement between the resonators. It is shown that a certain choice of the length of the coupling region between the resonators makes it possible to minimize or reduce to zero the coupling between them in the vicinity of the tripled average frequency of the main bandwidth, and thereby significantly suppress the parasitic attenuation dip in this frequency range. In the process of designing the filter, a synthesis method was used based on the transition from the filter to the corresponding $2n$ -pole, where n is the number of filter resonators. This approach allows us to consider the filter as a set of individual resonators connected to each other, and to introduce the concept of potential bandwidth, which can be easily determined by the characteristics of the $2n$ -pole and positioned in the desired frequency axis interval by an easily formalized algorithm for selecting the geometric dimensions of the device. The electrical characteristics of the $2n$ -pole used to synthesize the device within the baseband and at the stage of optimizing the device parameters in the vicinity of the tripled value of the baseband frequencies are the intrinsic conductivities of the resonators and the conductivities of the connection between them. The aim of the work is to study the effect of the length of the coupling region between the

resonators of a two-resonator microstrip comb filter on its electrical characteristics. It is shown that the choice of the optimal value of the length of the coupling region between the resonators makes it possible to expand the barrier band of the filter at a level of minus 30-35 dB by more than two times due to the suppression of the parasitic bandwidth formed in the vicinity of the tripled average frequency of the main bandwidth. This positive effect, which consists in expanding the filter barrier band, is confirmed by the results of designing a comb filter on two offset resonators in the HFSS program.

Microstrip comb filter; longitudinal displacement of resonators; parasitic bandwidth; barrier strip.

Введение. Поиск новых технических решений в области разработки микрополосковых фильтров (МПФ) на данном этапе развития радиотехники и беспроводной связи не теряет своей актуальности [1–5]. Одним из важнейших аспектов улучшения электрической характеристики полосового фильтра является расширение его полосы заграждения. Для решения этой задачи используются различные подходы [6–14].

С точки зрения расширения полосы заграждения для построения фильтра надо использовать, прежде всего, не полуволновые резонаторы, у которых первая паразитная полоса пропускания сосредоточена в окрестности удвоенной средней частоты основной полосы пропускания (ОПП), а четвертьволновые резонаторы, которые обладают двумя очевидными достоинствами. Во-первых, длина четвертьволнового резонатора при той же резонансной частоте в два раза меньше полуволнового, что в два раза уменьшает габариты фильтра. Во-вторых, четвертьволновые резонаторы обладают первой паразитной полосой пропускания, сосредоточенной в окрестности не удвоенной, а утроенной средней частоты ОПП, что автоматически позволяет получить в два раза более широкую полосу заграждения фильтра на четвертьволновых резонаторах по сравнению с фильтром на полуволновых резонаторах.

К фильтрам на четвертьволновых резонаторах относятся хорошо известные фильтры гребенчатого типа и фильтры на встречных стержнях. Несмотря на очевидные указанные достоинства фильтров на четвертьволновых резонаторах, их недостатком является необходимость обеспечить короткое замыкание одного из концов каждого резонатора, которое реализуется либо с помощью металлизированного отверстия, либо путем соединения конца резонатора с металлизированным основанием по торцу подложки. Введение короткого замыкания представляет собой дополнительную технологическую операцию, усложняющую и удорожающую производство фильтра и несколько ухудшающую повторяемость его характеристик.

Тем не менее, расширение полосы заграждения фильтра является порой настолько важным фактором, что в ряде случаев осуществляется переход от полуволновых резонаторов, например в традиционном шпилечном фильтре, к совокупности двух четвертьволновых резонаторов за счет введения металлизированного отверстия в середину каждого резонатора [15]. Это позволяет более чем в два раза расширить полосу заграждения фильтра на шпилечных резонаторах при тех же габаритных размерах устройства.

Другим возможным решением, приводящим к расширению полосы заграждения фильтра, является смещение резонаторов по отношению друг к другу, позволяющее изменять длину области связи между ними [8, 9]. Соответствующий выбор длины области связи между резонаторами позволяет минимизировать или свести к нулю связь между ними в окрестности утроенной средней частоты ОПП, и тем самым существенно подавить паразитный провал затухания в этом диапазоне частот. В этом случае первая паразитная полоса пропускания будет сформирована в окрестности упятеренной средней частоты ОПП, что позволит заметно расширить полосу заграждения по сравнению с фильтром на несмещенных четвертьволновых резонаторах.

В настоящей работе приведены результаты исследования влияния длины области связи между резонаторами двухрезонаторного микрополоскового гребенчатого фильтра на его электрические характеристики. Показано, что выбор оптимального значения длины области связи между резонаторами позволяет расширить полосу заграждения фильтра по уровню минус 30–35 дБ более чем в два раза за счет подавления паразитной полосы пропускания, формирующейся в окрестности утроенной средней частоты ОПП.

Математическая модель исследуемого фильтра и методика синтеза в пределах ОПП. На рис. 1 представлены схематические изображения исследуемого двухрезонаторного микрополоскового звена гребенчатого типа и соответствующего ему $2n$ -полюсника.

В [7, 15–17] было показано, что успешное и сравнительно быстрое проектирование полосового фильтра практически любой планарной топологии можно осуществить путем перехода от фильтра к соответствующему $2n$ -полюснику, где n – число резонаторов фильтра. Звено микрополоскового фильтра гребенчатого типа на рис. 1 образовано двумя резонаторами, каждый из которых имеет на одном из своих концов металлизированное отверстие, обозначенное черным кружком. Эти отверстия соединяют резонатор с металлизированной стороной диэлектрической подложки и тем самым обеспечивают короткое замыкание резонаторов в точках расположения этих отверстий.

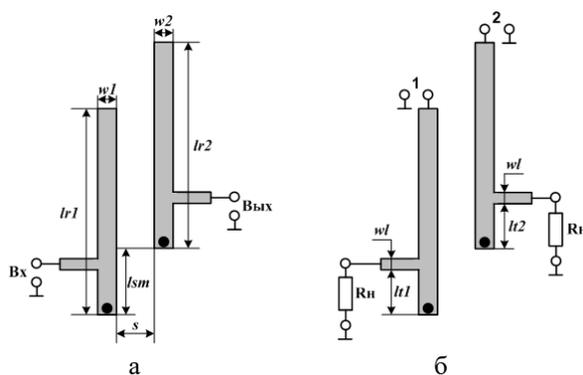


Рис. 1. Топология фильтра (а) и соответствующего ему 4-х полюсника (б)

На схемах рис. 1 указаны обозначения всех геометрических размеров устройства, значения которых должны быть либо заданы, либо определены в ходе решения задачи синтеза фильтра. Топология фильтра при ненулевом смещении резонаторов l_{sm} не обладает геометрической симметрией, поэтому ширины резонаторов w_1 , w_2 и их длины l_{r1} , l_{r2} , а также точки подключения l_{t1} , l_{t2} внешних цепей к резонаторам в большинстве случаев будут иметь неодинаковые значения.

Исходными данными для проектирования являются значения граничных частот ОПП, определяемой по уровню -3 дБ характеристики затухания фильтра, которая представляет собой зависимость элемента матрицы рассеяния $S_{21}(f)$, выраженного в дБ: $20 \lg S_{21}(f)$. На разных этапах проектирования для обозначения ОПП используется несколько различных терминов. На исходном этапе разработчик фильтра имеет дело с заданной полосой пропускания (ЗПП), границы которой определяются в техническом задании на проектирование. Под фактической полосой пропускания (ФПП) фильтра понимается диапазон частот, определяемый по уровню -3 дБ характеристики затухания $20 \lg S_{21}(f)$ спроектированного фильтра. Задача проектирования заключается в определении всех размеров фильтра на рис. 1,а, при которых ФПП фильтра, определенная в результате моделирования (например, в программе HFSS) или в результате проведения натурного эксперимента, будет максимально соответствовать ЗПП.

Осуществить успешное проектирование фильтра, исходя из анализа и оптимизации функции $20 \lg S_{21}(f)$, очень сложно, а при большом числе резонаторов практически невозможно. Это обусловлено тем, что функция $20 \lg S_{21}(f)$ не является монотонной, имеет сложный вид и ее поведение зависит в той или иной степени от всех геометрических размеров устройства, причем разработчик фильтра не может точно предсказать влияние отдельных геометрических параметров устройства на эту характеристику.

Существенно более простой является методика синтеза [15–17], основанная на переходе от фильтра к новому устройству, изображенному на рис. 1,б) и имеющему абсолютно аналогичные геометрические параметры, но описываемого не одной характери-

стикой $20lgS_{21}(f)$, а большим набором других электрических характеристик. Как видно на рис. 1 для перехода от фильтра (рис. 1,а) к новому 4-полоснику (рис. 1,б) необходимо: во-первых, в точках входа и выхода устройства разместить сосредоточенные резисторы с сопротивлением, равным входному сопротивлению на внешних зажимах фильтра, а, во-вторых, в середине разомкнутых концов каждого резонансного элемента надо создать новый виртуальный полюс. Эти полюсы, число которых равно числу резонаторов фильтра, являются виртуальными внешними зажимами для определения на них новых электрических характеристик резонансных элементов фильтра, а именно собственных проводимостей резонаторов $\dot{Y}_{ii}(f)$ и проводимости связи $\dot{Y}_{ij}(f)$ между ними. Важно отметить, что точка резонатора, в которой размещается его полюс, не может быть выбрана произвольно. Эта точка называется разомкнутым концом резонатора, так как при резонансе значение напряжения в ней максимально, а значение тока равно нулю, что соответствует режиму холостого хода.

Введенные в рассмотрение проводимости $\dot{Y}_{ij}(f)$, каждая из которых представляет собой комплексную функцию частоты, характеризуют свойства отдельного резонансного элемента во взаимосвязи с другими элементами $2n$ -полосника. В данном случае число резонаторов равно двум, поэтому матрица проводимостей 4-полосника на рис. 1,б имеет вид:

$$[\dot{Y}]_n = \begin{bmatrix} \dot{Y}_{11} & \dot{Y}_{12} \\ \dot{Y}_{21} & \dot{Y}_{22} \end{bmatrix}.$$

Требования к проводимостям 4-полосника формулируются в пределах ОПП, и на этом этапе вводится еще один важный термин для определения полосы пропускания – потенциальная полоса пропускания (ППП). Она представляет собой диапазон частот, в пределах которого выполняется совокупность условий для формирования ППП (1) и условий согласования в этой ППП (2):

условия формирования ППП

$$\left| \frac{Im\dot{Y}_{ii}(f)*\rho_i}{Im\dot{Y}_{i,i+1}(f)*\gamma} \right| \leq 1, \text{ при } f \in [f_n, f_b], i = 1,2; \quad (1)$$

условия согласования в ППП

$$\frac{\rho_1}{|Im\dot{Y}_{1,2}(f)|} = \frac{1}{Re\dot{Y}_{11}(f)}; \quad \frac{\rho_2}{|Im\dot{Y}_{2,1}(f)|} = \frac{1}{Re\dot{Y}_{22}(f)}, \text{ при } f \in [f_n, f_b]. \quad (2)$$

В соотношениях (1)-(2) использованы следующие обозначения:

$Im\dot{Y}_{ii}(f)$ – мнимая составляющая собственной проводимости i -го резонансного элемента;

$Im\dot{Y}_{i,i+1}(f)$ – мнимая составляющая проводимости связи между i -м и $i+1$ -м элементом;

$Re\dot{Y}_{ii}(f)$ – реальная составляющая собственной проводимости i -го резонансного элемента;

f_n – нижняя граничная частота ОПП;

f_b – верхняя граничная частота ОПП;

ρ_i, γ – вещественные константы, определяемые на основе параметров лестничного прототипа g_i по формулам [18, 19]:

$$\gamma = g_1, \quad \rho_j = \frac{g_1}{\sqrt{g_j * g_{j+1}}}, \quad j = 1,2, \dots, n-1. \quad (3)$$

Заметим, что функция $\frac{Im\dot{Y}_{ii}(f)*\rho_i}{Im\dot{Y}_{i,i+1}(f)*\gamma}$, фигурирующая в (1), называется резонансной характеристикой i -го резонатора.

Важно, что мнимые составляющие собственных проводимостей отдельных резонансных элементов $Im\dot{Y}_{ii}(f)$ зависят, прежде всего, от геометрических параметров именно того элемента, на зажиме которого они определяются. Другие элементы тоже оказывают определенное влияние на эту характеристику, но оно существенно меньше. Это позволяет удовлетворить требования (1) к проводимостям 4-полосника в пределах ППП в ходе итерационного подбора значений геометрических размеров при сравнительно небольшом числе итераций.

Выполнить условия согласования (2), фактически возможно лишь в одной или двух точках в пределах ППП, поскольку практически всегда функции $\frac{\rho_1}{|Im\dot{Y}_{1,2}(f)|}$, $\frac{\rho_2}{|Im\dot{Y}_{2,1}(f)|}$ и функции $\frac{1}{Re\dot{Y}_{11}(f)}$, $\frac{1}{Re\dot{Y}_{22}(f)}$ ведут себя неодинаково. Основное влияние на значение функции $\frac{1}{[Re\dot{Y}_{11}(f)+Re\dot{Y}_{22}(f)]}$ оказывают геометрические параметры резонаторов, к которым подключаются внешние цепи. Управлять значением функций $\frac{\rho_1}{|Im\dot{Y}_{1,2}(f)|}$ и $\frac{\rho_2}{|Im\dot{Y}_{2,1}(f)|}$ можно, прежде всего, величиной зазоров s_1 и s_2 , но надо учитывать, что значение $Im\dot{Y}_{1,2}(f)$ и $Im\dot{Y}_{2,1}(f)$ фигурирует и в условиях формирования ППП (1), влияя на наклон резонансной характеристики, а, следовательно, на ширину полосы пропускания. Поэтому основным способом обеспечить согласование фильтра с внешними цепями является верный выбор значения $\frac{1}{Re\dot{Y}_{11}(f)}$, $\frac{1}{Re\dot{Y}_{22}(f)}$.

Важно отметить, что условие (1) будут выполняться при любых геометрических размерах элементов фильтра, но в пределах некоторой произвольной ППП, границы которой не обязательно совпадают с границами ЗПП. То есть резонансная характеристика $\frac{Im\dot{Y}_{i,i}(f)*\rho_i}{Im\dot{Y}_{i,i+1}(f)*\gamma}$, определенная для каждого из резонансных элементов, всегда будет пересекать частотную ось на своей резонансной частоте, и в окрестности этого нуля резонансной характеристики будет формироваться ППП данного резонансного элемента. В ходе проектирования требуется подобрать значения геометрических размеров так, чтобы обеспечить максимальное совпадение ППП по обоим резонаторам, причем границы этих ППП должны с достаточной точностью совпадать с границами ЗПП. Если выполнить и условия согласования (2) в пределах этих ППП, то все три полосы пропускания (ЗПП, ППП и ФПП) совпадут. Этот факт в итоге проектирования обязательно требует подтверждения, которое осуществляется путем расчета характеристики фильтра $20lgS_{21}(f)$ и ее сопоставления с заданной характеристикой по ТЗ.

Оптимизация характеристики фильтра в окрестности утроенной частоты ОПП. Очевидно, что в ходе проектирования необходимо обязательно решить задачу синтеза в пределах ОПП. Но в рамках данной работы эта задача не является единственной, поскольку здесь также рассматривается и задача подбора оптимальной длины области связи между резонаторами, решение которой позволит расширить полосу заграждения фильтра в два раза.

Поскольку задача проектирования содержит два уровня оптимизации – в ОПП и в окрестности утроенной частоты ОПП, то и решать ее приходится поочередно по этим двум различным критериям. Предположим, что оптимальное значение смещения между резонаторами lsm , которое отвечает за поведение характеристик устройства в окрестности утроенной средней частоты ОПП, найдено. Очевидно, что это должно быть выполнено при фиксированных значениях всех прочих геометрических параметров фильтра. Далее приходится решать задачу синтеза устройства в пределах ОПП и на этом этапе проектирования значение смещения lsm , найденное ранее, должно быть зафиксировано. В процессе синтеза в пределах ОПП ширины резонаторов $w1$, $w2$, зазор между ними s , длины резонаторов $lr1$, $lr2$, а также точки подключения $lt1$, $lt2$ могут изменить свои значения. Это потребует на следующем этапе вернуться к оптимизации характеристик в окрестности утроенной средней частоты ОПП, в результате чего оптимальное значение lsm будет уточнено и изменится. Этот факт потребует уточнения результатов синтеза в окрестности ОПП... и так далее. Практика показывает, что этот двухступенчатый процесс оптимизации не является продолжительным, и требует не более 3-4 итераций.

Результаты исследования и оптимизации смещения между резонаторами. Задача проектирования устройства была решена в программе HFSS [20]. На рис. 2 представлена модель двухрезонаторного микрополоскового звена гребенчатого типа по схеме на рис. 1,б, разработанная в программе HFSS и предназначенная для анализа проводимостей 4-полюсника и решения задачи синтеза путем удовлетворения требований (1)–(2).

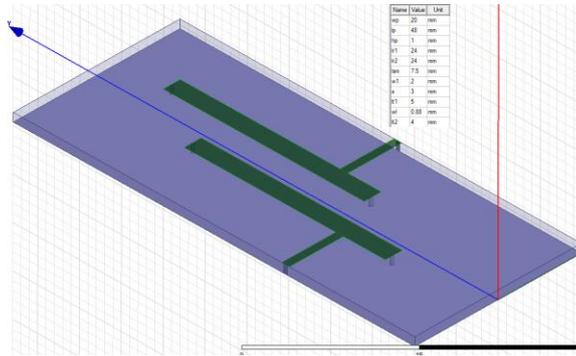


Рис. 2. Модель звена 4-полосника в HFSS с начальными значениями

Характеристики 4-полосника при смещении резонаторов $l_{sm}=7,5$ мм представлены на рис. 3 и содержат много полезной информации. Заметим, что малые значения проводимости связи в пределах основной ППП (отмечена на рис. 3 большими зелеными окружностями) говорят о большой крутизне резонансных характеристик в окрестности их нулей, а это обеспечивает малую ширину ОПП, что соответствует потенциальным возможностям фильтра гребенчатого типа. Окружностями голубого цвета отмечены частоты, в окрестностях которых потенциально могут формироваться паразитные полосы пропускания.

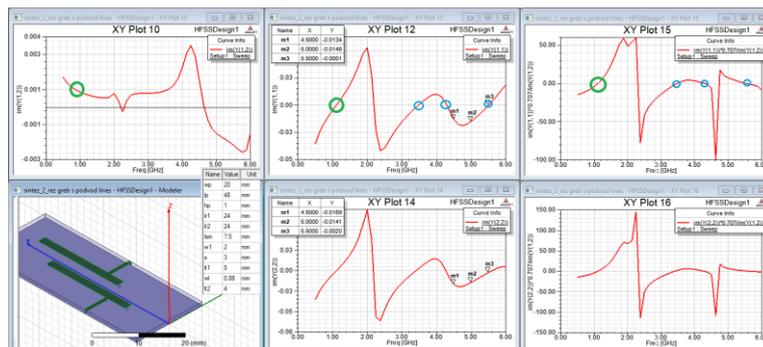


Рис. 3. Характеристики 4-полосника при $l_{sm} = 7,5$ мм

Результаты показали, что оптимальные значения l_{sm} находятся в пределах от 8 до 9 мм. Для наглядности на рис. 4 представлены характеристики 2n-полосника, а на рис. 5 приведены характеристики соответствующего фильтра при $l_{sm} = 8$ мм. Они показывают, что максимальное значение затухания в окрестности утроенной частоты ОПП не поднимается выше минус 28 дБ.

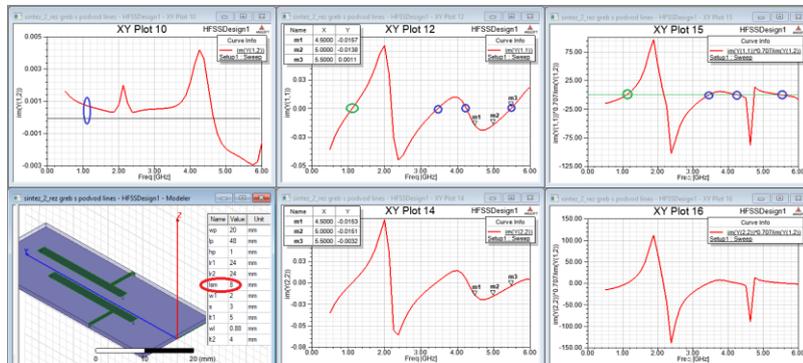


Рис. 4. Характеристики 4-полосника при $l_{sm} = 8,0$ мм

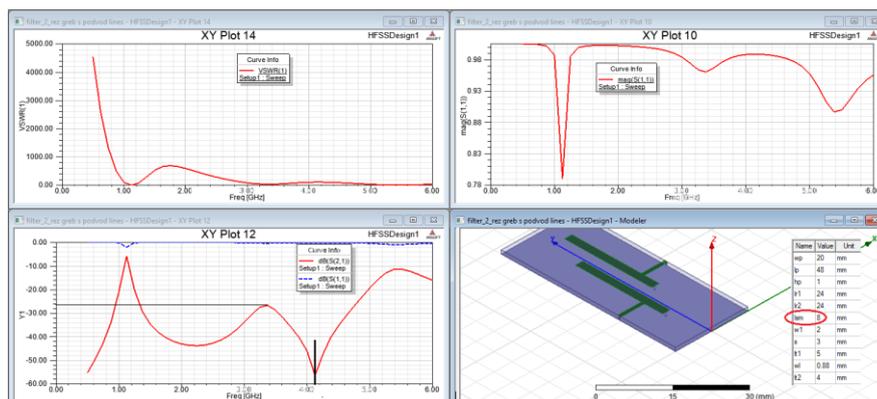


Рис. 5. Характеристики фильтра при $l_{sm} = 8,0$ мм

Исследование показало, что оптимальное значение смещения между резонаторами составляет $l_{sm} = 8,65$ мм при длине резонатора 24 мм. На рис. 6 представлены характеристики фильтра в окрестности утроенной частоты ОПП при $l_{sm} = 8,65$ мм. Они показывают, что значение затухания на утроенной частоте ОПП удалось повысить до минус 42 дБ.

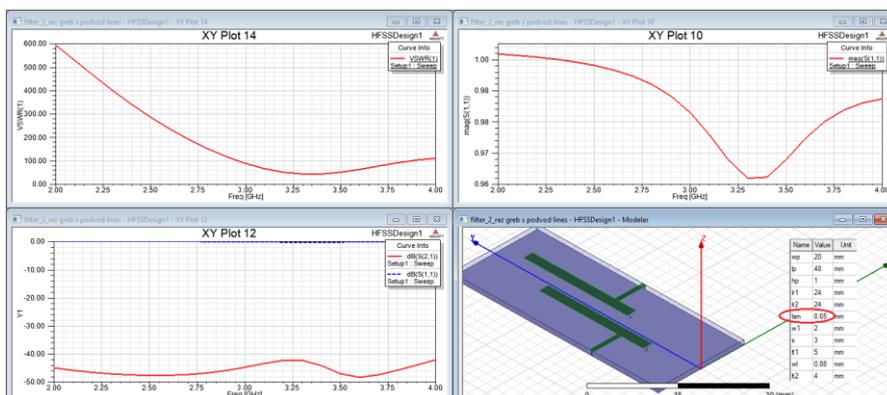


Рис. 6. Характеристики фильтра в окрестности 3 ГГц при $l_{sm} = 8,65$ мм

На этом первый этап оптимизации конструкции, связанный с определением оптимального значения длины области связи между резонаторами, закончен.

Синтез фильтра в пределах ОПП. Второй этап проектирования предполагает решение задачи синтеза в пределах ОПП, на частотах 1 ГГц.

Характеристики исходного 2n-полюсника, рассчитанные в окрестности ОПП при $l_{sm} = l_{sm\text{ опт}} = 8,65$ мм, показали, что ППП формируется около 1,12 ГГц, ширина ее составляет 0,04 ГГц (4%), и условия согласования (2) при $l_{t1} = 5$ мм, $l_{t2} = 4$ мм не выполнены. После подбора точек подключения внешних цепей и их новых значениях $l_{t1} = l_{t2} = 1,35$ мм удалось выполнить условия согласования в пределах ППП. Результат представлен на рис. 7. Заметим, что на данном этапе исследований положение и ширина ППП не задавалась, а просто определялась по резонансным характеристикам 4-полюсника.

Результат расчета характеристик фильтра для $l_{t1} = l_{t2} = 1,35$ мм представлен на рис. 8. Он подтверждает высокую степень соответствия между ППП и ФПП и говорит о достаточно хорошем согласовании, которое удалось обеспечить в пределах ФПП фильтра.

На этом первый этап синтеза фильтра в окрестности ОПП закончен.

Второй этап корректировки длины области связи. Изменение точек подключения внешних цепей к фильтру сказалось на подавлении паразитной полосы пропускания в окрестности утроенной частоты основной полосы пропускания. Результаты расчета характеристик в широком диапазоне частот приведены на рис. 9.

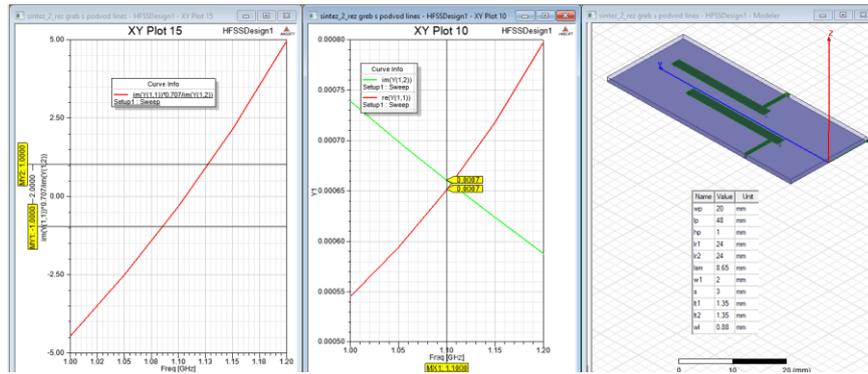


Рис. 7. Характеристики 4-полосника в ППП при $lt_1 = lt_2 = 1,35$ мм

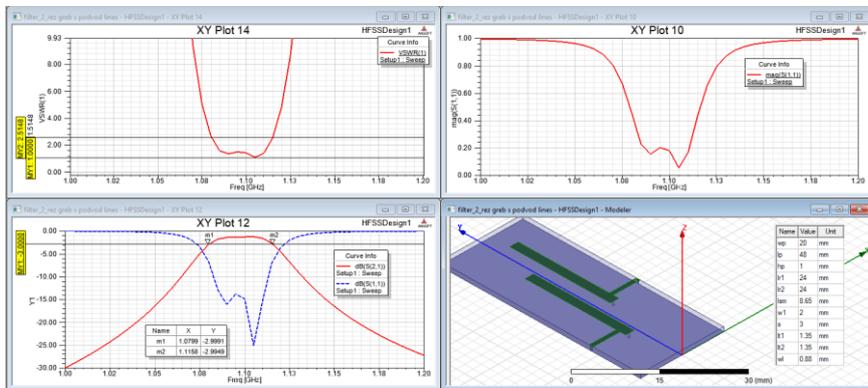


Рис. 8. Характеристики фильтра в окрестности ППП

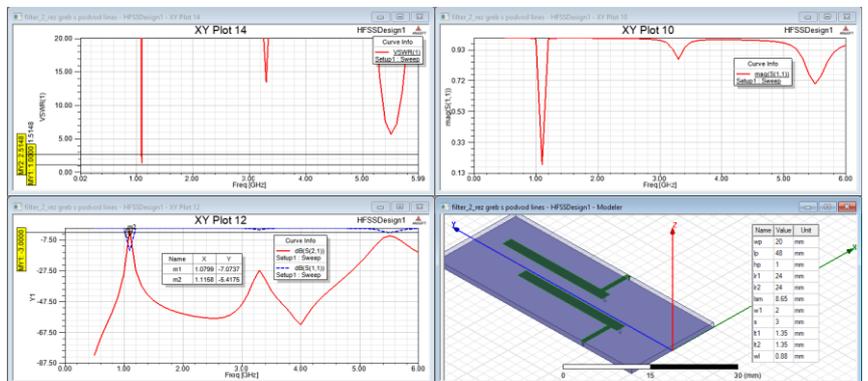


Рис. 9. Характеристики фильтра в широком диапазоне частот

Изменение точек подключения внешних цепей $lt_1 = lt_2$ привело к снижению уровня затухания на утроенной частоте ОПП с минус 42 дБ до минус 27,5 дБ. Следовательно, надо корректировать длину области связи, а затем снова возвращаться к синтезу фильтра в пределах ОПП.

В результате корректировки смещения между резонаторами, было найдено его новое оптимальное значение: $l_{sm\text{ опт}} = 8,76$ мм. Расчет характеристик фильтра показал, что уровень затухания в этом диапазоне составил не менее минус 33 дБ.

Второй этап синтеза в ОПП. Проверка влияния незначительного изменения смещения между резонаторами с $l_{sm} = 8,65$ мм на $l_{sm} = 8,76$ мм на поведения характеристик 4-полосника и характеристик фильтра в пределах ОПП показала, что точки подключения

внешних цепей к крайним резонаторам потребовалось незначительно скорректировать: было $t1 = 1,35$ мм, а $lt2 = 1,35$ мм, а стало $t1 = 1,3$ мм, а $lt2 = 1,355$ мм. За счет этого удалось добиться улучшения согласования и на этом проектирование устройства закончилось. Итоговые характеристики фильтра в окрестности основной полосы пропускания и в широком диапазоне частот представлены на рис. 10 и 11.

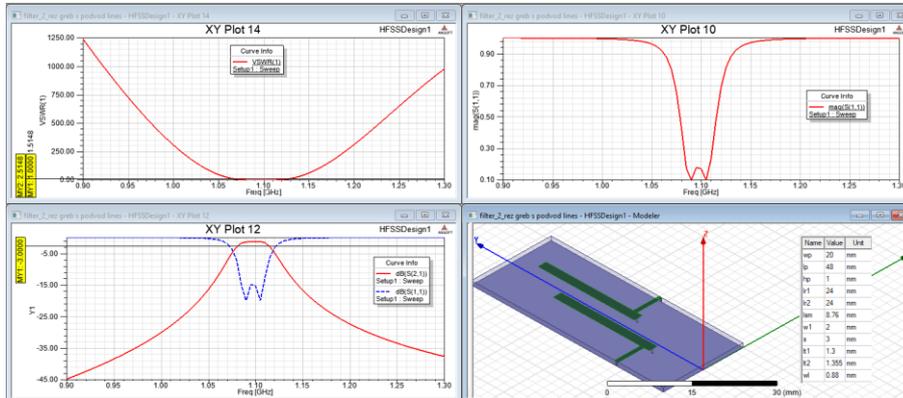


Рис. 10. Итоговые характеристики фильтра в окрестности ОПП

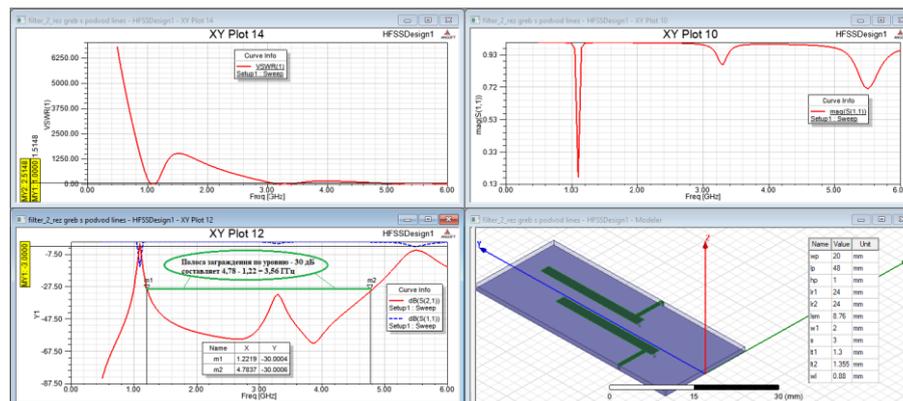


Рис. 11. Итоговые характеристики фильтра в широком диапазоне частот

Для сравнения и оценки положительного эффекта от введения смещения резонаторов фильтра на рис. 12 приведены характеристики аналогичного гребенчатого двухрезонаторного фильтра без смещения резонаторов.

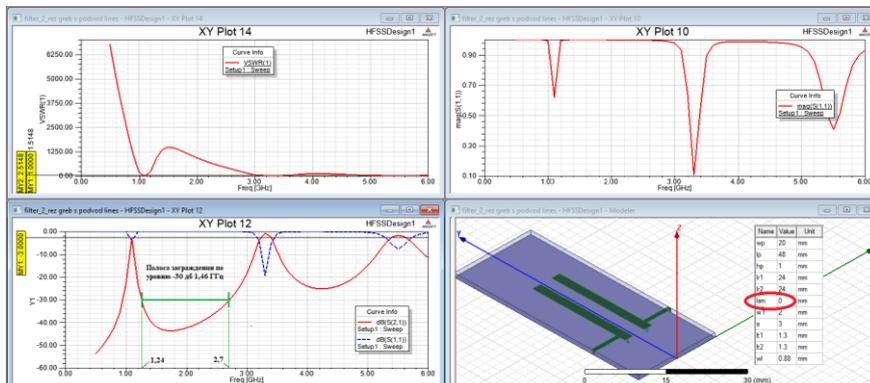


Рис. 12. Характеристики фильтра без смещения резонаторов

Сопоставление результатов на рис. 11 и 12 показывает, что полоса заграждения по уровню – 30 дБ в фильтре с оптимальным смещением резонаторов составляет 3,56 ГГц, а в фильтре без смещения резонаторов – всего 1,46 ГГц. Таким образом, за счет введения смещения между резонаторами удалось расширить полосу заграждения в 2,5 раза.

Заключение. В процессе исследования установлено положительное влияние длины области связи на механизм формирования характеристик фильтра в пределах полосы заграждения, а именно, уменьшение провала затухания в окрестности утроенного значения средней частоты основной полосы пропускания, найдено оптимальное значение длины области связи для максимального подавления паразитной полосы пропускания в окрестности утроенного значения средней частоты основной полосы пропускания. Выполнена оценка положительного эффекта, связанного с расширением полосы заграждения от смещения резонаторов в фильтре. За счет введения смещения между резонаторами удалось расширить полосу заграждения почти в 2,5 раза. Дополнительным положительным эффектом от введения смещения резонаторов явился рост крутизны обоих скатов характеристики в окрестности ОПП.

Материалы статьи подготовлены при поддержке проекта № ВнГр/23-01-РТ «Исследование программно определяемых радиосистем» в Южном федеральном университете".

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Аристархов Г.М., Кириллов И.Н.* Компактные микрополосковые фильтры с повышенной частотной избирательностью на основе полуволновых резонаторов // Системы синхронизации, формирования и обработки сигналов. – 2020. – Т. 11, № 2. – С. 40-44. – DOI: 10.1109/SYNCHRONINFO49631.2020.9166122.
2. *Аристархов Г.М., Кириллов И.Н., Корчагин А.И., Кувишинов В.В.* Компактные высокоизбирательные микрополосковые фильтры на свернутых сонаправленных шпилечных резонаторах // Радиотехника. – 2021. – Т. 85, № 4. – С. 126-137. – DOI: 10.18127/j00338486-202104-14.
3. *Аристархов Г.М., Аринин О.В., Кириллов И.Н.* Многополосные полосно-пропускающие микрополосковые фильтры на двух сонаправленных шпилечных резонаторах // Системы синхронизации, формирования и обработки сигналов. – 2023. – Т. 14, № 1. – С. 4-9.
4. Патент № 2775868 С1. Рос. Федерация, 2022. Высокоселективный микрополосковый полосно-пропускающий фильтр / *Беляев Б.А., Сержантов А.М., Ходенков С.А.*
5. Патент № 2797166 С1. Рос. Федерация, 2023. Микрополосковый полосно-пропускающий фильтр / *Беляев Б.А., Сержантов А.М., Ходенков С.А.*
6. *Сержантов А.М.* Резонансные полосковые структуры и частотно-селективные устройства на их основе с улучшенными характеристиками: дисс. ... д-ра техн. наук. – Красноярск. 2015.
7. *Лабынцев А.В.* Пассивные селективные устройства СВЧ на базе многомодовых микрополосковых линий: дисс. ... канд. техн. наук. – Таганрог: ТРТИ, 1987. – 201 с.
8. *Агафонов В.М., Лабынцев В.А., Лабынцев А.В.* Микрополосковый гребенчатый фильтр. А.С. СССР №1450018. БИ №1, 1989.
9. *Агафонов В.М., Лабынцев А.В.* Микрополосковый фильтр. А.С. СССР №1376140. БИ № 7, 1988.
10. *Labyntsev A.V.; Poveshenko L.I.; Kharlanov D.V.* Microstrip Filter with Extended Barrier // 2019 Radiation and Scattering of Electromagnetic Waves (RSEMW). – 2019. – P. 176-179. – <https://ieeexplore.ieee.org/document/8792722> doi: 10.1109/RSEMW.2019.8792722.
11. Patent US5066933. Band-pass filter. Publication Date 19.11.1991.
12. Patent US20020057143. High frequency filter, filter device, and electronic apparatus incorporating the same. Publication Date 16.05.2002.
13. *Xiao F., Norgren M.* Compact third-order microstrip bandpass filter using hybrid resonators // Progress in Electromagnetics Research C. – 2011. – Vol. 19. – P. 93-106.
14. *Агафонов В.М., Лабынцев В.А., Лабынцев А.В.* Микрополосковый гребенчатый фильтр. А.С. СССР №1262602. Оpubл. В БИ, 1986, № 37.
15. *Лабынцев А.В., Игнатьев В.В., Васильев В.В., Михайловский Вит.А., Харланов Д.В.* Проектирование фильтра из шпилечных резонаторов с металлизированными отверстиями с помощью программной системы // Программные продукты и системы. – 2023. – Т. 36, № 4. – С. 668-677. – DOI: 10.15827/0236-235X.142.668-677. – <http://www.swsys.ru/archive/2023-4.pdf>.
16. *Obukhovets Victor A., Labyntsev Alexey V., Kharlanov Dmitry V., Ponomarev Nikolay E., Sarzhanov Alexander N.* Example of Filter Synthesis on Two Anti-directional Stud Resonators with One Metalized Hole in Each Resonator // 2023 Radiation and Scattering of Electromagnetic Waves (RSEMW). – 2023. – P. 160-163. – DOI: 10.1109/rsemw 58451.2023.10202019.

17. *Labyntsev A.V., Poveshenko L.I., Kharlanov D.V.* Two-Stage Design of Microstrip Filters from Arbitrary Configuration Resonators // 2019 Radiation and Scattering of Electromagnetic Waves (RSEMW). – 2019. – P. 172-175. – <https://ieeexplore.ieee.org/document/8792770>. – DOI: 10.1109/RSEMW.2019.8792770.
18. *Маттей Д.Л., Янг Л., Джонс Е.М.Т.* Фильтры СВЧ, согласующие цепи и цепи связи: пер. с англ. / под ред. Л.В. Алексеева и Ф.В. Кушнера. – М.: Связь, 1971. – 440 с.
19. *Агафонов В.М.* Полиномиальные фильтры СВЧ // Радиотехника и электроника. – 1970. – Т. 15, № 10. – С. 2191-2193.
20. *Банков С.Е., Курушин А.А.* Расчёт антенн и СВЧ структур с помощью HFSS Ansoft. – М.: ЗАО «НПП „Родник“», 2009. – 256 с.

REFERENCES

1. *Aristarkhov G.M., Kirillov I.N.* Kompaktnye mikropoloskovye fil'try s povyshennoy chastotnoy izbiratel'nost'yu na osnove poluvolnovykh rezonatorov [Compact microstrip filters with increased frequency selectivity based on half-wave resonators.], *Sistemy sinkhronizatsii, formirovaniya i obrabotki signalov* [Systems of synchronization, signal generation and processing], 2020, Vol. 11, No. 2, pp. 40-44. DOI: 10.1109/SYNCHROINFO49631.2020.9166122.
2. *Aristarkhov G.M., Kirillov I.N., Korchagin A.I., Kuvshinov V.V.* Kompaktnye vysokoizbiratel'nye mikropoloskovye fil'try na svernutykh sonapravlennykh shpilechnykh rezonatorakh [Compact highly selective microstrip filters on folded co-directional stud resonators], *Radiotekhnika* [Radio engineering], 2021, Vol. 85, No. 4, pp. 126-137. DOI: 10.18127/j00338486-202104-14.
3. *Aristarkhov G.M., Arinin O.V., Kirillov I.N.* Mnogopolosnye polosno-propuskayushchie mikropoloskovye fil'try na dvukh sonapravlennykh shpilechnykh rezonatorakh [Multiband band-pass microstrip filters on two co-directional stud resonators], *Sistemy sinkhronizatsii, formirovaniya i obrabotki signalov* [Systems of synchronization, signal generation and processing], 2023, Vol. 14, No. 1, pp. 4-9.
4. *Belyaev B.A., Serzhantov A.M., Khodenkov S.A.* Patent № 2775868 C1. Ros. Federatsiya, 2022. Vysokoselektivnyy mikropoloskovyy polosno-propuskayushchiy fil'tr [Patent RF, No. 2775868 C1, 2022. Highly selective microstrip band-pass filter].
5. *Belyaev B.A., Serzhantov A.M., Khodenkov S.A.* Patent № 2797166 C1. Ros. Federatsiya, 2023. Mikropoloskovyy polosno-propuskayushchiy fil'tr [Patent RF, No. 2797166 C1, 2023. Microstrip band-pass filter].
6. *Serzhantov A.M.* Rezonansnye poloskovye struktury i chastotno-selektivnye ustroystva na ikh osnove s uluchshennymi kharakteristikami: diss. ... d-ra tekhn. nauk [Resonant strip structures and frequency-selective devices based on them with improved characteristics: dr. of eng. sc. diss.]. Krasnoyarsk. 2015.
7. *Labyntsev A.V.* Passivnye selektivnye ustroystva SVCH na baze mnogomodovykh mikropoloskovykh liniy: diss. ... kand. tekhn. nauk [Passive selective microwave devices based on multimode microstrip lines: cand. of eng. sc. diss.]. Taganrog: TRTI, 1987, 201 p.
8. *Agafonov V.M., Labyntsev V.A., Labyntsev A.V.* Mikropoloskovyy grebenchatyy fil'tr. A.S. SSSR №1450018. BI №1, 1989 [Microstrip comb filter. A.S. USSR No. 1450018. No. 1, 1989].
9. *Agafonov V.M., Labyntsev A.V.* Mikropoloskovyy fil'tr. A.S. SSSR № 1376140. BI № 7, 1988 [Microstrip Filter. A.S. USSR No. 1376140. No. 7, 1988].
10. *Labyntsev A.V.; Poveshenko L.I.; Kharlanov D.V.* Microstrip Filter with Extended Barrier, *2019 Radiation and Scattering of Electromagnetic Waves (RSEMW)*, 2019, pp. 176-179. – <https://ieeexplore.ieee.org/document/8792722> doi: 10.1109/RSEMW.2019.8792722.
11. Patent US5066933. Band-pass filter. Publication Date 19.11.1991.
12. Patent US20020057143. High frequency filter, filter device, and electronic apparatus incorporating the same. Publication Date 16.05.2002.
13. *Xiao F., Norgren M.* Compact third-order microstrip bandpass filter using hybrid resonators, *Progress in Electromagnetics Research C*, 2011, Vol. 19, pp. 93-106.
14. *Agafonov V.M., Labyntsev V.A., Labyntsev A.V.* Mikropoloskovyy grebenchatyy fil'tr. A.S. SSSR №1262602. Opubl. V BI, 1986, № 37 [Microstrip comb filter. USSR No. 1262602. Publ. In BI, 1986, No. 37].
15. *Labyntsev A.V., Ignat'ev V.V., Vasil'ev V.V., Mikhaylovskiy Vit.A., Kharlanov D.V.* Proektirovanie fil'tra iz shpilechnykh rezonatorov s metallizirovannymi otverstiyami s pomoshch'yu programmnoy sistemy [Software system for designing a filter made of hairpin resonators with metallized holes], *Programmnye produkty i sistemy* [Software products and systems], 2023, Vol. 36, No. 4, pp. 668-677. DOI: 10.15827/0236-235X.142.668-677. Available at: <http://www.swsys.ru/archive/2023-4.pdf>.
16. *Obukhovets Victor A., Labyntsev Alexey V., Kharlanov Dmitry V., Ponomarev Nikolay E., Sarzhanov Alexander N.* Example of Filter Synthesis on Two Anti-directional Stud Resonators with One Metallized Hole in Each Resonator, *2023 Radiation and Scattering of Electromagnetic Waves (RSEMW)*, 2023, pp. 160-163. DOI: 10.1109/rsemw 58451.2023.10202019.

17. *Labyntsev A.V., Poveshenko L.I., Kharlanov D.V.* Two-Stage Design of Microstrip Filters from Arbitrary Configuration Resonators, *2019 Radiation and Scattering of Electromagnetic Waves (RSEMW)*, 2019, pp. 172-175. Available at: <https://ieeexplore.ieee.org/document/8792770>. DOI: 10.1109/RSEMW.2019.8792770.
18. *Mattey D.L., Yang L., Dzhons E.M.T.* Fil'try SVCh, soglasuyushchie tsepi i tsepi svyazi [Microwave filters, impedance-matching networks, and coupling structures]: transl. from engl., ed. by L.V. Alekseeva i F.V. Kushnira. Moscow: Svyaz', 1971, 440 p.
19. *Agafonov V.M.* Polinomial'nye fil'try SVCh [Polynomial microwave filters], *Radiotekhnika i elektronika* [Radio Engineering and Electronics], 1970, Vol. 15, No. 10, pp. 2191-2193.
20. *Bankov S.E., Kurushin A.A.* Raschet antenn i SVCh struktur s pomoshch'yu HFSS Ansoft [Calculation of antennas and microwave structures using HFSS Ansoft]. Moscow: ZAO «NPP „Rodnik“», 2009, 256 p.

Статью рекомендовал к опубликованию д.т.н., профессор И.И. Турулин.

Лабынцев Алексей Викторович – Южный федеральный университет; e-mail: labyntsev@yandex.ru; г. Таганрог, Россия; тел.: 89085142003; кафедра теоретических основ радиотехники; к.т.н.; с.н.с.; доцент.

Пономарёв Николай Евгеньевич – e-mail: nikolain777@yandex.ru; г. Краснодар, Россия; тел.: 89094452610; кафедра антенн и радиопередающих устройств; соискатель.

Харланов Дмитрий Валентинович – e-mail: mitya999@mail.ru; г. Краснодар, Россия; тел.: 89620130677; кафедра антенн и радиопередающих устройств; соискатель.

Саржанов Александр Николаевич – Краснодарское Высшее военное училище; e-mail: sarjanov.a.n@gmail.com; г. Краснодар, Россия; тел. 89885593888; сотрудник КВВУ.

Labyntsev Alexey Viktorovich – Southern Federal University; e-mail: labyntsev@yandex.ru; Taganrog, Russia; phone: +79085142003; the department of theoretical foundations of radio engineering; cand. of eng. sc.; senior researcher; associate professor.

Ponomarev Nikolay Evgenyevich – e-mail: nikolain777@yandex.ru; Krasnodar, Russia; phone: +79094452610; the department of antennas and radio transmitting devices; the applicant.

Kharlanov Dmitry Valentinovich – e-mail: mitya999@mail.ru; Krasnodar, Russia; phone: +79620130677; the department of antennas and radio transmitting devices; the applicant.

Sarzanov Alexander Nikolaevich – Krasnodar Higher Military School; e-mail: sarjanov.a.n@gmail.com; Krasnodar, Russia; phone: +79885593888; KVVU employee.

УДК 621.38

DOI 10.18522/2311-3103-2024-2-164-175

Л.К. Самойлов, Д.Ю. Денисенко, А.Е. Титов

МАКСИМАЛЬНЫЕ ДИНАМИЧЕСКИЕ ПОГРЕШНОСТИ ФИЛЬТРОВ ЛЕЖАНДРА В СИСТЕМАХ КОНТРОЛЯ И УПРАВЛЕНИЯ*

В системах управления и контроля чаще всего используются фильтры нижних частот и полосовые фильтры. Для ограничения спектра сигналов от датчиков широко применяются аналоговые, дискретно-аналоговые и цифровые фильтры, амплитудно-частотные характеристики которых аппроксимируются различными математическими функциями, в т.ч. полиномами Лежандра. Применение фильтров Лежандра в контуре системы автоматического управления приводит к изменению ее динамических характеристик. Характер этого влияния зависит от порядка передаточной функции фильтра, а также от вида аппроксимации, которые выбираются при проектировании системы управления и контроля. Задержка информации в таких фильтрах является причиной появления динамической составляющей их погрешности, влияющей на общую погрешность системы управления и контроля, что снижает допустимую скорость её работы. В статье дается аналитическая оценка зависимости величины динамической погрешности для фильтров Ле-

* Исследование выполнено за счет гранта Российского научного фонда (проект № 23-79-10023, <https://rscf.ru/project/23-79-10023/>).

жандра нижних частот и полосовых. Это позволяет оперативно решать прямую и обратную задачи распределения погрешностей системы управления и контроля и обосновать скорость её работы. В статье проводится анализ схем полосового фильтра Лежандра первого, второго и третьего порядков, а далее, полученные результаты обобщаются на полосовой фильтр Лежандра произвольного порядка. Показано, что для фильтров нижних частот величины максимальных динамических погрешностей могут быть получены с высокой точностью. Для полосовых фильтров Лежандра погрешности аппроксимации математической зависимости максимальных динамических погрешностей от параметров фильтров определяются единицами процентов, но в некоторых случаях могут достигать 20%.

Динамическая погрешность; фильтр нижних частот; полосовой фильтр; фильтр Лежандра; система контроля и управления.

L.K. Samoilov, D.Yu. Denisenko, A.E. Titov

MAXIMUM DYNAMIC ERRORS OF LEGENDRE FILTERS IN CONTROL AND CONTROL SYSTEMS

In control and monitoring systems, low-pass filters and band-pass filters are most often used. To limit the spectrum of signals from sensors, analog, discrete-analog and digital filters are widely used, the amplitude-frequency characteristics of which are approximated by various mathematical functions, incl. Legendre polynomials. The use of Legendre filters in the circuit of an automatic control system leads to a change in its dynamic characteristics. The nature of this influence depends on the order of the filter transfer function, as well as on the type of approximation that are chosen when designing the control and monitoring system. The information delay in such filters causes the appearance of a dynamic component of their error, which affects the overall error of the control and monitoring system, which reduces the permissible speed of its operation. The article provides an analytical assessment of the dependence of the magnitude of the dynamic error for low-pass and bandpass Legendre filters. This allows you to quickly solve the direct and inverse problems of error distribution of the control and monitoring system and justify the speed of its operation. The article analyzes the Legendre bandpass filter circuits of the first, second and third orders, and then the results obtained are generalized to the Legendre bandpass filter of an arbitrary order. It is shown that for low-pass filters the values of maximum dynamic errors can be obtained with high accuracy. For Legendre bandpass filters, the errors in approximation of the mathematical dependence of the maximum dynamic errors on the filter parameters are determined in units of percent, but in some cases they can reach 20%.

Dynamic error; low-pass filter; bandpass filter; Legendre filter; control and monitoring system.

Введение. Рассматриваемые фильтры были синтезированы А. Папулисом [1–6]. Имя фильтрам дали полиномы Лежандра $L_n(x)$, которые использовались при синтезе фильтров [7–9].

Полином Лежандра является результатом решения уравнения:

$$\frac{d}{dx} \left[(1 - x^2) \frac{dL_n(x)}{dx} \right] + n(n + 1)L_n(x) = 0. \quad (1)$$

Для фильтров Лежандра нижних частот (ЛФН) порядка n квадрат амплитудно-частотной характеристики $M_n(\omega)$ выражается через полиномы Лежандра $L_n(\omega^2)$ как:

$$M_n^2(\omega) = \frac{1}{1 + L_n(\omega^2)}. \quad (2)$$

Фильтры Лежандра нижних частот имеют наибольшую крутизну спада амплитудно-частотной характеристики (АЧХ) среди всех типов фильтров в так называемой узкой полосе пропускания (рис. 1,а), что позволяет достигать более эффективного разделения частот, чем с помощью других линейных фильтров [1].

На рис. 1 приведены для сравнения АЧХ трех фильтров нижних частот: Баттерворта (В); Лежандра (Л) и Чебышева (Т).

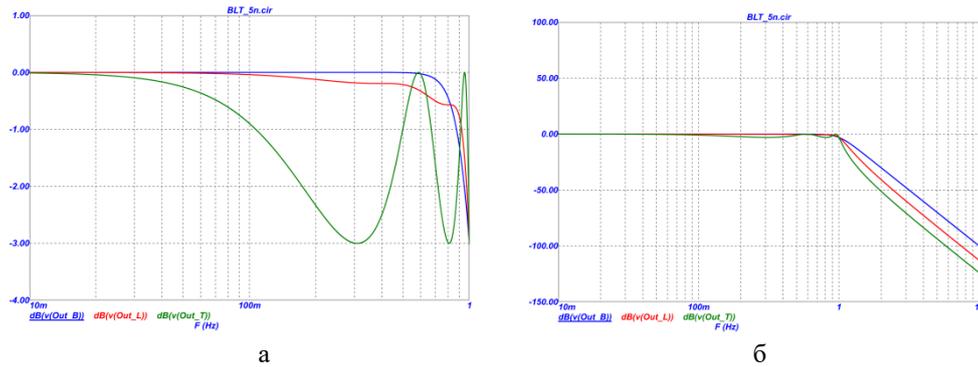


Рис. 1. Амплитудно-частотные характеристики фильтров Баттерворта (Out B), Лежандра (Out L) и Чебышева (Out T). а – в узкой полосе пропускания, б – в широкой полосе

Введем обозначения величин, используемых в формулах и выводах. На рис. 2 приведены АЧХ рассматриваемых фильтров Лежандра (ЛФ).

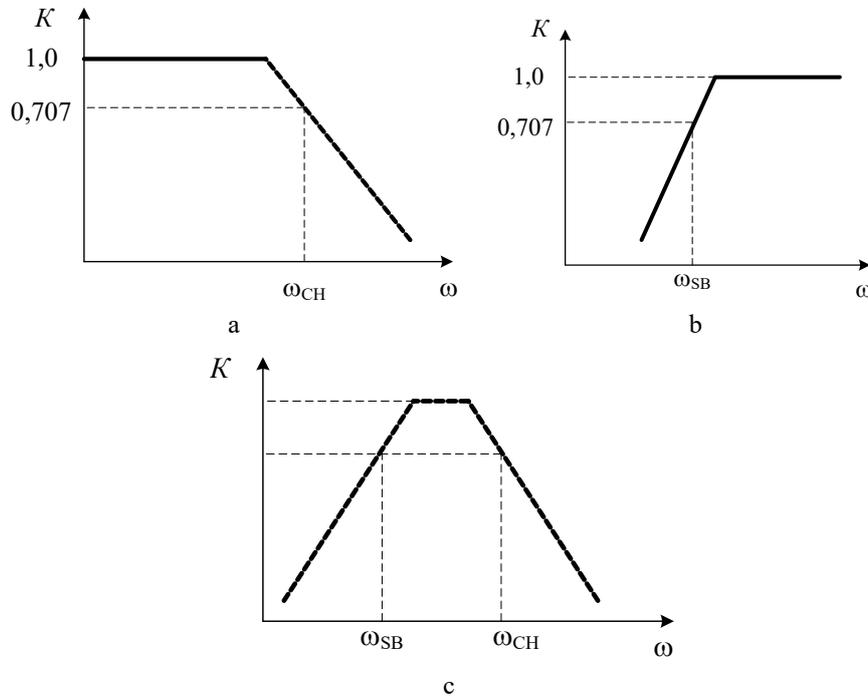


Рис. 2. АЧХ рассматриваемых ЛФ: а – нижних частот (ЛФН); б – верхних частот (ЛФВ); с – полосы частот (ЛФП)

АЧХ рассматриваемых фильтров (рис. 3) характеризуются частотами среза АЧХ на уровне 0,707: ЛФН имеет частоту среза ω_{SH} ; ЛФВ имеет частоту среза ω_{SB} , а ЛФП имеет две частоты среза – ω_{SH} и ω_{SB} [10–15].

Максимальная динамическая погрешность ЛФН n-го порядка ($\gamma_n^{HЧ}(\max)$) определяется круговой частотой среза АЧХ фильтра на уровне 0,707 (ω_{SH}) и временем задержки ($t_n^{HЧ}$) информации в ЛФН [12, 16]:

$$\gamma_n^{HЧ}(\max) = \omega_{SH} \cdot t_n^{HЧ}. \tag{3}$$

Фактически для вычисления $\gamma_n^{\text{НЧ}}(\text{max})$ необходимо получение аналитической зависимости $t_n^{\text{НЧ}}$ как функции n и ω_{SH} .

Полосовые фильтры являются устройством, получаемым путем последовательного соединения ЛФН и фильтра верхних частот (ЛФВ) [12].

Фильтры ЛФН, ЛФВ и ЛПФ n -го порядка получаются путем последовательного соединения фильтров первого и второго порядков.

Можно с большой долей уверенности утверждать, что зависимости для динамических погрешностей ЛФ будут иметь тот же характер, что и для других типов фильтров, например фильтров Чебышева.

Как показано в работе [13], динамическая погрешность ПФ Чебышева n -го порядка ($\gamma_n^{\text{ПФЧ}}$) для сигнала с круговой частотой ω_s будет (при $\omega_{\text{SB}} \leq \omega_s \leq \omega_{\text{SH}}$) определяется выражением:

$$\gamma_n^{\text{ПФЧ}}(\omega_s) = \frac{\omega_s}{\omega_{\text{SH}}} \cdot \gamma_n^{\text{НЧ}}(\text{при } \frac{\omega_s}{\omega_{\text{SH}}} = 1) \cdot \left(e^{-\frac{\omega_{\text{SB}}}{\omega_{\text{SH}}}} \right) \cdot 100\%. \quad (4)$$

Максимальная величина этой погрешности будет при $\omega_s = \omega_{\text{SH}}$:

$$\gamma_n^{\text{ПФЧ}}(\text{max}) = \gamma_n^{\text{НЧ}}(\text{при } \frac{\omega_s}{\omega_{\text{SH}}} = 1) \cdot \left(e^{-\frac{\omega_{\text{SB}}}{\omega_{\text{SH}}}} \right) \cdot 100\%. \quad (5)$$

Как показано в работах [12], [13] для повышения точности аппроксимации отрицательная экспонента должна быть скорректирована:

$$\gamma_n^{\text{ПФЧ}}(\text{max}) = \gamma_n^{\text{НЧ}}(\text{при } \frac{\omega_s}{\omega_{\text{SH}}} = 1) \cdot \left(e^{-\frac{\omega_{\text{SB}}}{\omega_{\text{SH}}}} - 0,2 \sin \pi \cdot \frac{\omega_{\text{SB}}}{\omega_{\text{SH}}} \right) \cdot 100\%. \quad (6)$$

В формулах (2) и (3) $\gamma_n^{\text{НЧ}}$ – динамическая погрешность ФНЧ, входящего в состав ПФ.

Основная цель настоящей работы. Можно с большой долей уверенности утверждать, что зависимости для ЛФ, аналогичные (5), (6), будут иметь идентичный вид, но в деталях будут отличаться. Конечно, предположение об идентичности требует проверки.

Получение конкретных значений динамических погрешностей для ЛФН и ЛФП в диапазоне $n = 1 \div 10$ позволит принимать первоначальные оперативные решения по проектированию СУ.

1. Методика получения аналитических зависимостей для величин динамических погрешностей ЛФ. Времена задержки сигналов в рассматриваемых ЛФ получают с использованием программы Nuhertz Filter Solutions. Для получения аналитических выражений времен задержки и динамических погрешностей используется программа Mathcad 14.

Фильтры Лежандра высоких порядков получаются путем соединения нужного числа ЛФ первого и второго порядков [17-20].

Ниже приводятся принципиальные схемы:

- ◆ ЛФН первого порядка (ЛФН1) (рис. 3,а);
- ◆ ЛФВ первого порядка (ЛФВ1) (рис. 3,б);
- ◆ ЛФН второго порядка (ЛФН2) (рис. 4) (выход фильтра LPF);
- ◆ ЛФВ второго порядка (ЛФВ2) (рис. 4) (выход фильтра HPF).

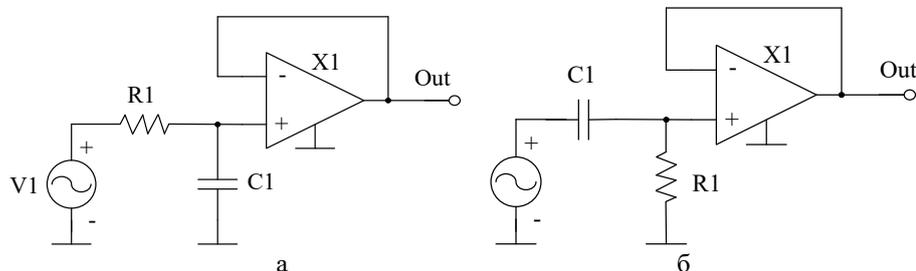


Рис. 3. Принципиальные схемы ЛФН1 (рис. 3,а) и ЛФВ1(рис. 3,б)

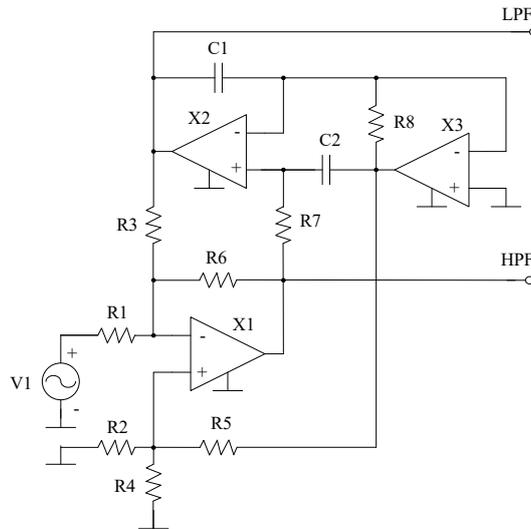


Рис. 4. Принципиальная схема ЛФ 2-го порядка, реализующая на выходе LPF фильтр ЛФН, а на выходе HPF фильтр ЛФВ

Полосовой фильтра в структуре системы управления и контроля состоит из последовательно соединенных ЛФН и ЛФВ (рис. 5).

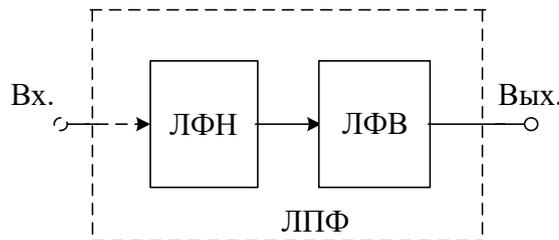


Рис. 5. Типовая структурная схема ЛПФ

Порядки ЛФН и ЛФВ, входящих в ЛПФ, могут быть различными, но в данной работе рассматривается классический вариант, когда порядки фильтров равны.

В соответствии с (3), максимальная динамическая погрешность ЛФН $\gamma_n^{HЧ}(\max)$ определяется круговой частотой среза АЧХ фильтра (ω_{SH}) и временем задержки ($t_{пф}$) информации в ЛФН.

Анализируя структурную схему ЛПФ (рис. 6), можно утверждать, что величина максимальной погрешности ЛПФ ($\gamma_n^{HЧ}(\max)$) будет при условии

$$\omega_{SH} \geq \omega_{BH}, \tag{7}$$

и будет равна или близка к величине $\gamma_n^{HЧ}(\max)$ ЛФН, входящего в ЛПФ.

В настоящей статье проводится анализ схем ЛПФ первого, второго и третьего порядков, а далее, полученные результаты обобщаются на ЛПФ произвольного порядка.

Для ЛПФ каждого порядка с использованием программы Nuhertz Filter Solutions получаются временные диаграммы переходных процессов при подаче ступеньки напряжения на вход ЛПФ. Для ЛПФ каждого порядка будет 5 временных диаграмм, которые получены с разными отношениями частот $\frac{\omega_{SB}}{\omega_{SH}}$:

$$\frac{\omega_{SB}}{\omega_{SH}} = 0; \quad \frac{\omega_{SB}}{\omega_{SH}} = 0,25; \quad \frac{\omega_{SB}}{\omega_{SH}} = 0,5; \quad \frac{\omega_{SB}}{\omega_{SH}} = 0,75; \quad \frac{\omega_{SB}}{\omega_{SH}} = 1,0. \tag{8}$$

Для определения $t_{пф}$ на вход ЭПФ подавался перепад напряжения и фиксировалось время задержка на уровне 0,5 от максимального значения для данной временной диаграммы.

Для моделирования процессов в ПФ выбраны фильтры с круговой частотой $\omega_{SH} = 2\pi \cdot 50\text{Гц}$.

2. Аналитические зависимости максимальных динамических погрешностей ЛФН ($\gamma_n^{HЧ}(\text{max})$) от порядка фильтра. В соответствии с [10]

$$\gamma_n^{HЧ}(\text{max}) = \omega_{SH} \cdot t_n^{HЧ}. \quad (9)$$

В табл. 1 приведены времена задержки ЛФН ($t_n^{HЧ}$) до 10 порядка при $\omega_{SH} = 2\pi \cdot 50,0$ Гц, полученные с использованием программы Nuhertz Filter Solutions.

Таблица 1

Расчетные значения времени задержки сигнала в зависимости от порядка ЛФН

n	1	2	3	4	5	6	7	8	9	10
$t_n^{HЧ}$ мС	2.203	4.576	7.712	10.42	13.56	16.44	19.66	22.54	25.68	28.73

На рис. 6 приведен график зависимости $t_n^{HЧ}$ от n.

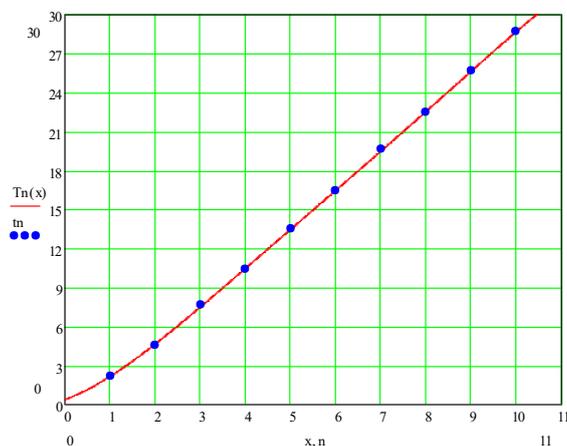


Рис. 6. График зависимости $t_n^{HЧ}$ от n для ЛФН

На графике: n – порядок ЛФН, Vn – порядок из таблицы (n), синий график $t_n^{HЧ}(n)$ – график по табл. 1, красный график $t_n^{HЧ}(n)$ – по аналитическому выражению.

Аналитическое выражение для аппроксимированной зависимости $t_n^{HЧ} = F(n)$ для фильтра с $\omega_{SH} = 2\pi \cdot 50,0$ Гц, полученное с помощью программы Mathcad 14, будет иметь вид:

$$t_n^{HЧ}(n) = 0,395 + 1,265n + 0,625n^2 - 0,11n^3 + 0,009281n^4 - 0,000301n^5. \quad (10)$$

3. Аналитические зависимости максимальных динамических погрешностей ЛПФ ($\gamma_n^{ПЧ}$) от порядка фильтра.

3а. Динамические погрешности ЛПФ первого порядка (ЛПФ1). В соответствии с принятым алгоритмом с помощью программы Nuhertz Filter Solutions были получены временные диаграммы переходных процессов в ЛПФ-1 для пяти значений $\frac{\omega_{SB}}{\omega_{SH}}$ (рис. 7):

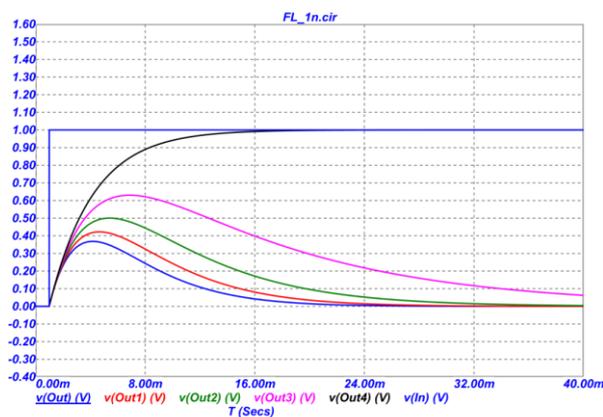


Рис. 7. Временные диаграммы переходных процессов ЛПФ-1

На рис. 7, а также на рис. 8, 9 временные диаграммы соответствуют следующим отношениям $\frac{\omega_{SB}}{\omega_{SH}}$:

- ◆ v(Out) (V) соответствует варианту $\frac{\omega_{SB}}{\omega_{SH}} = 1,0$;
- ◆ v(Out1) (V) соответствует варианту $\frac{\omega_{SB}}{\omega_{SH}} = 0,75$;
- ◆ v(Out2) (V) соответствует варианту $\frac{\omega_{SB}}{\omega_{SH}} = 0,5$;
- ◆ v(Out3) (V) соответствует варианту $\frac{\omega_{SB}}{\omega_{SH}} = 0,25$;
- ◆ v(Out4) (V) соответствует варианту $\frac{\omega_{SB}}{\omega_{SH}} = 0,00$.

Информация, полученная из временных диаграмм (рис. 7) и в результате расчетов для ЛПФ-1 сведена в Таблице 2.

Таблица 2

Графическая и расчетная информация для ЛПФ-1

$\frac{\omega_{SB}}{\omega_{SH}}$	$0,5 U_{\text{ВЫХ}}^{\text{max}}$	Задержка на уровне $0,5 U_{\text{ВЫХ}}^{\text{max}}$. $t_{\text{зад}}(1)$ (mC)	Относ. задер. $\frac{t_{\text{зад}}(1)}{t_0(1)}$	$e^{-\frac{\omega_{SB}}{\omega_{SH}}}$	$e^{-\frac{\omega_{SB}}{\omega_{SH}}} - 0,2 \sin \pi \cdot \frac{\omega_{SB}}{\omega_{SH}}$	Погреш. аппрок. в % с sin
0	0,5	2,202	1,0	1,00	1,00	0
0,25	0,315	1,387	0,629	0,780	0,704	11,9
0,5	0,250	1,000	0,454	0,607	0,407	10,1
0,75	0,211	0,833	0,378	0,473	0,397	4,5
1,0	0,184	0,744	0,337	0,369	0,369	8,9

3б. Динамические погрешности ЛПФ второго порядка (ЛПФ-2). Полученные с помощью программы Nuhertz Filter Solutions временные диаграммы переходных процессов для 5-ти ЛПФ-2 приведены на рис. 8.

Информация, полученная из временных диаграмм (рис. 8) и в результате расчетов для ЛПФ2 сведена в табл. 3.

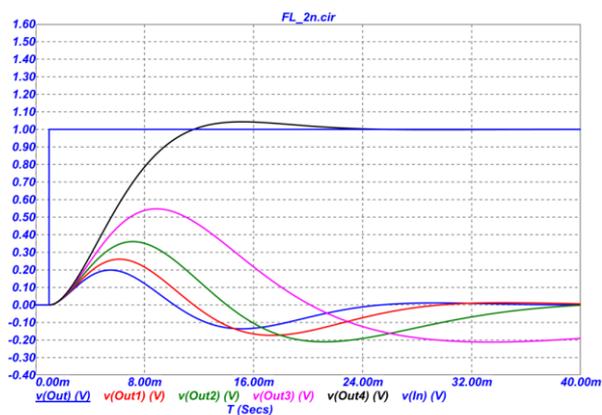


Рис. 8. Временные диаграммы переходных процессов в ЛПФ2 для пяти значений $\frac{\omega_{SB}}{\omega_{SH}}$ (0; 0,25; 0,5; 0,75; 1,0)

Таблица 3

Графическая и расчетная информация для ЛПФ2

$\frac{\omega_{SB}}{\omega_{SH}}$	$0,5 U_{\text{ВЫХ}}^{\text{max}}$	Задер. на уровне $0,5 U_{\text{ВЫХ}}^{\text{max}}$. $t_{\text{зад}}(1)$ (мс)	Относ. задер. $\frac{t_{\text{зад}}(1)}{t_0(1)}$	$e^{-\frac{\omega_{SB}}{\omega_{SH}}}$	$e^{-\frac{\omega_{SB}}{\omega_{SH}}} - 0.2 \sin \pi \cdot \frac{\omega_{SB}}{\omega_{SH}}$	Погреш. аппрок. в %.
0	0,521	4.721	1,0	1,00	1,00	0
0,25	0.274	3.257	0,689	0,780	0,704	2,8
0,5	0.180	2.007	0,440	0,607	0,407	7,5
0,75	0.130	2.220	0,470	0,473	0,397	14,4
1,0	0,995	1.926	0,407	0,369	0,369	8,8

3в. **Динамические погрешности ЛПФ третьего порядка (ЛПФ3).** Полученные с помощью программы Nuhertz Filter Solutions временные диаграммы переходных процессов для 5-ти ЛПФ3 приведены на рис. 9.

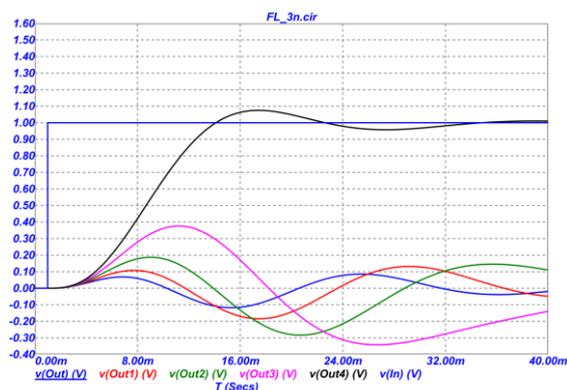


Рис. 9. Временные диаграммы переходных процессов в ЛПФ3 для пяти значений $\frac{\omega_{SB}}{\omega_{SH}}$ (0; 0,25; 0,5; 0,75; 1,0)

Информация, полученная из временных диаграмм (рис. 9) и в результате расчетов для ЭПФ-3 сведена в табл. 4.

Таблица 4

Графическая и расчетная информация для ЛПФ-3

$\frac{\omega_{SB}}{\omega_{SH}}$	$0,5 U_{ВЫХ}^{max}$	Задер. на уровне $0,5 U_{ВЫХ}^{max}$. $t_{зад}(1)$ (мс)	Относ. задер. $\frac{t_{зад}(1)}{t_0(1)}$	$e^{-\frac{\omega_{SB}}{\omega_{SH}}}$	$e^{-\frac{\omega_{SB}}{\omega_{SH}}}$ $-0.2 \sin$	Погреш. аппрок. в %.
0	0,537	7.980	1,0	1,00	1,00	0
0,25	0.188	5.506	0,690	0,780	0,704	2,0
0,5	0.093	4.050	0,508	0,607	0,407	19,6
0,75	0.054	3.691	0,462	0,473	0,397	14,0
1,0	0,034	3.168	0,397	0,369	0,369	7,0

Полученные результаты для ЛФН 1, 2 и 3 порядков, а также для ЛПФ 1, 2 и 3 порядков позволяют сделать вывод о том, что для произвольного n величины времен задержек, которые определяют динамические погрешности полосовых фильтров на качественном уровне, могут быть оценены функцией отрицательной экспоненты с поправочным коэффициентом. Начальные значения экспонент определяются временами задержки ЛФН, которые входят в ЛПФ.

В табл. 5 приведены времена задержки информации для обоих типов фильтров при $\omega_{SH} = 2\pi \cdot 50,0$ Гц.

Для ЛФН приведены результаты расчетов $t_n^{HЧ}$ для фильтров с 1-го по 10 порядок.

Для ЛПФ с порядком 1,2, и 3 приведены результаты, полученные в данной работе с последующей аппроксимацией. Для ЛПФ с порядком 4÷10 приведены аппроксимированные значения в соответствии с формулой (6).

Таблица 5

Результаты расчета времен задержки для ЛФН и ЛПФ с 1-го по 10 порядок

Порядок n ФНЧ, ПФ	$t_n^{HЧ}$	$t_n^{ПФ}(n; \frac{\omega_{SB}}{\omega_{SH}})$				
		$\frac{\omega_{SB}}{\omega_{SH}} = 0$	$\frac{\omega_{SB}}{\omega_{SH}} = 0,25$	$\frac{\omega_{SB}}{\omega_{SH}} = 0,5$	$\frac{\omega_{SB}}{\omega_{SH}} = 0,75$	$\frac{\omega_{SB}}{\omega_{SH}} = 1$
1	2,203	2,203	1,387	1,000	0,833	0,744
2	4.576	4.721	3.257	2.007	2.220	1.926
3	7.712	7.980	5.506	4.050	3.691	3.168
4	10.42	10.42	7.335	4.240	4.136	3.844
5	13.56	13.56	9.546	5.518	5.383	5.003
6	16.44	16.44	11.573	6.691	6.526	6.066
7	19.66	19.66	13.840	8.001	7.805	7.254
8	22.54	22.54	15.868	9.173	8.948	8.317
9	25.68	25.68	18.078	10.451	10.194	9.457
10	28.73	28.73	20.225	11.693	11.405	10.601

Заключение. В результате проведенного исследования можно сделать заключение, что значения максимальных динамических погрешностей фильтров Лежандра нижних частот и полосовых на качественном уровне подчиняются тем же зависимостям, что и для фильтров Баттерворта, Чебышева, Бесселя. Для фильтров Лежандра нижних частот величины максимальных динамических погрешностей могут быть рассчитаны с высокой точностью. Для полосовых фильтров Лежандра погрешности предложенной аппроксимации могут достигать 20%, что является приемлемым при качественном анализе характеристик проектируемых систем контроля и управления. Уточнение результатов может быть сделано расчетом конкретного варианта полосового фильтра Лежандра.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Улахович Д.А. Основы теории линейных электрических цепей. – СПб.: БХВ-Петербург, 2009. – 816 с.
2. Нестеров С.В. Синтез линейных электрических цепей. – Волгоград, Изд-во ВолГУ, 2013. – 84 с.
3. Andrei H., Andrei P.C., Oprea G. and Botea B. Basic equations of linear electric and magnetic circuits in quasi-stationary state based on principle of minimum absorbed power and energy // 2014 International Symposium on Fundamentals of Electrical Engineering (ISFEE), Bucharest, Romania, 2014. – P. 1-6.
4. Amgad A., AbdelAty A.M., Attia H.A. and Elbarawy M.T.M.M. Designing Of Fractional Order Legendre Filter Using Optimization Techniques // 2023 40th National Radio Science Conference (NRSC), Giza, Egypt, 2023. – P. 230-237.
5. Tseng C.-C. and Lee S.-L. Polynomial Graph Filter Design Using Legendre Polynomials // 2023 Sixth International Symposium on Computer, Consumer and Control (IS3C), Taichung, Taiwan, 2023. – P. 299-302.
6. Буренко Е.А. Аппроксимация характеристик электрических фильтров рациональными функциями // XLIV международная молодёжная научная конференция «Гагаринские чтения»: Сб. тезисов докладов. – М., 2018. – Т. 1. – С. 375-376.
7. Dörfler F., Simpson-Porco J.W. and Bullo F. Electrical Networks and Algebraic Graph Theory: Models, Properties, and Applications // Proceedings of the IEEE. – 2018. – Vol. 106, No. 5. – P. 977-1005.
8. Argypoulos P.E. and Lev-Ari H. Numerically stable, computationally efficient, filter bank based modeling of transfer functions for power system transient simulation // 2015 IEEE Power and Energy Conference at Illinois (PECI), Champaign, IL, USA, 2015. – P. 1-8.
9. Xie F., Yang Z., Yang C., Chen Y., Zhang B. and Qiu D. Construction and Experimental Realization of the Fractional-Order Transformer by Oustaloup Rational Approximation Method // IEEE Transactions on Circuits and Systems II: Express Briefs. – 2023. – Vol. 70, No. 4. – P. 1550-1554.
10. Самойлов Л.К. Ввод-вывод аналоговых сигналов в системах управления и контроля. – Таганрог: Изд-во ЮФУ, 2015. – 264 с.
11. Samoylov L.K., Denisenko D.Yu., Prokopenko N.N. The Function Approximation of the Signal Delay Time in the Anti-Alias Filter of the A/D Interface of the Instrumentation and Control System // 2018 IEEE International Conference on Electrical Engineering and Photonics. – 2018. – P. 1-4.
12. Palyuchenko Anton. RF signal chain discource – part 2: Essential building blok – Analog Devices. – 2021. – Vol. 55, No. 3.
13. Samojlov L.K., Denisenko D.Yu., Prokopenko N.N. Динамические погрешности процесса ввода аналоговых сигналов датчиков в системах управления и контроля: монография. – М.: Солон-Пресс, 2021. – 240 с. – ISBN 978-5-91359-444-0.
14. Samoylov Leontiy, Denisenko Darya, Prokopenko Nikolay, Bugakova Anna. Dynamic errors of broadband Chebyshev bandpass filters switched on at the ADC input in automatic control systems // 4th IEEE international conference on Design & Test of integrated micro & nano-Systems. – 2022. – P. 1-5.
15. Matejček M. and Šostronek M. Low-pass filter design with microcontroller // 2022 New Trends in Signal Processing (NTSP), Liptovský Mikuláš, Slovakia, 2022. – P. 1-8.
16. Nuhertz Filter Solutions 2019. Available: <https://www.ansys.com/products/electronics/ansys-nuhertz-filtersolutions>.
17. Rozaqi L., Nugroho A., Sanjaya K.H. and Simbolon A. Ivonita. Design of Analog and Digital Filter of Electromyography // 2019 International Conference on Sustainable Energy Engineering and Application (ICSEEA), Tangerang, Indonesia, 2019. – P. 186-192.
18. Grosinger P., Rybář J. and Ďuriš S. Design of a Fast Discrete Realtime Filters Based on Analog Models // 2021 22nd International Carpathian Control Conference (ICCC), Velké Karlovice, Czech Republic, 2021. – P. 1-5.
19. Sahu A. Kumar and Sahu A. Kumar. A Review on Different Filter Design Techniques and Topologies for Bio-potential Signal Acquisition Systems // 2018 3rd International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2018. – P. 934-937.
20. Zhang X. and Li Y. Evolutionary Design of an Analog Filter // 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, Hangzhou, China, 2013. – P. 474-477.

REFERENCES

1. Ulakhovich D.A. Osnovy teorii lineynykh elektricheskikh tsepey [Fundamentals of the theory of linear electrical circuits]. St. Petersburg: BKHV-Peterburg, 2009, 816 p.
2. Nesterov S.V. Sintez lineynykh elektricheskikh tsepey [Synthesis of linear electrical circuits]. Volgograd: Izd-vo VolGU, 2013, 84 p.
3. Andrei H., Andrei P.C., Oprea G. and Botea B. Basic equations of linear electric and magnetic circuits in quasi-stationary state based on principle of minimum absorbed power and energy, 2014 International Symposium on Fundamentals of Electrical Engineering (ISFEE), Bucharest, Romania, 2014, pp. 1-6.

4. *Amgad A., AbdelAty A.M., Attia H.A. and Elbarawy M.T.M.M.* Designing Of Fractional Order Legendre Filter Using Optimization Techniques, *2023 40th National Radio Science Conference (NRSC), Giza, Egypt, 2023*, pp. 230-237.
5. *Tseng C.-C. and Lee S.-L.* Polynomial Graph Filter Design Using Legendre Polynomials, *2023 Sixth International Symposium on Computer, Consumer and Control (IS3C), Taichung, Taiwan, 2023*, pp. 299-302.
6. *Burenko E.A.* Approksimatsiya kharakteristik elektricheskikh fil'trov ratsional'nymi funktsiyami. [Approximation of characteristics of electric filters by rational functions], *XLIV international youth scientific conference "Gagarin Readings". Collection of abstracts of reports.* Moscow, 2018, Vol. 1, pp. 375-376.
7. *Dörfler F., Simpson-Porco J.W. and Bullo F.* Electrical Networks and Algebraic Graph Theory: Models, Properties, and Applications, *Proceedings of the IEEE*, 2018, Vol. 106, No. 5, pp. 977-1005.
8. *Argyropoulos P.E. and Lev-Ari H.* Numerically stable, computationally efficient, filter bank based modeling of transfer functions for power system transient simulation, *2015 IEEE Power and Energy Conference at Illinois (PECI), Champaign, IL, USA, 2015*, pp. 1-8.
9. *Xie F., Yang Z., Yang C., Chen Y., Zhang B. and Qiu D.* Construction and Experimental Realization of the Fractional-Order Transformer by Oustaloup Rational Approximation Method, *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2023, Vol. 70, No. 4, pp. 1550-1554.
10. *Samoilov L.K.* Vvod-vyvod analogovykh signalov v sistemakh upravleniya i kontrolya. [Input-output of analog signals in control and monitoring systems]. Taganrog: Izd-vo YuFU, 2015, 264 p.
11. *Samoylov L.K., Denisenko D.Yu., Prokopenko N.N.* The Function Approximation of the Signal Delay Time in the Anti-Alias Filter of the A/D Interface of the Instrumentation and Control System, *2018 IEEE International Conference on Electrical Engineering and Photonics*, 2018, pp. 1-4.
12. *Patyuchenko Anton.* RF signal chain discourse – part 2: Essential building blok – Analog Devices, 2021, Vol. 55, No. 3.
13. *Samojlov L.K., Denisenko D.Yu., Prokopenko N.N.* Dinamicheskie pogreshnosti protsessa vvoda analogovykh signalov datchikov v sistemakh upravleniya i kontrolya: monografiya [Dynamic errors in the process of inputting analog signals of sensors in control and monitoring systems: monograph]. Moscow: Solon-Press, 2021, 240 p. ISBN 978-5-91359-444-0.
14. *Samoylov Leontiy, Denisenko Darya, Prokopenko Nikolay, Bugakova Anna.* Dynamic errors of broadband Chebyshev bandpass filters switched on at the ADC input in automatic control systems, *4th IEEE international conference on Design & Test of integrated micro & nano-Systems*, 2022, pp. 1-5.
15. *Matejček M. and Šostronek M.* Low-pass filter design with microcontroller, *2022 New Trends in Signal Processing (NTSP), Liptovský Mikuláš, Slovakia, 2022*, pp. 1-8.
16. Nuhertz Filter Solutions 2019. Available: <https://www.ansys.com/products/electronics/ansys-nuhertz-filtersolutions>.
17. *Rozaqi L., Nugroho A., Sanjaya K.H. and Simbolon A. Ivonita.* Design of Analog and Digital Filter of Electromyography, *2019 International Conference on Sustainable Energy Engineering and Application (ICSEEA), Tangerang, Indonesia, 2019*, pp. 186-192.
18. *Grosinger P., Rybář J. and Ďuriš S.* Design of a Fast Discrete Realtime Filters Based on Analog Models, *2021 22nd International Carpathian Control Conference (ICCC), Velké Karlovice, Czech Republic, 2021*, pp. 1-5.
19. *Sahu A. Kumar and Sahu A. Kumar.* A Review on Different Filter Design Techniques and Topologies for Bio-potential Signal Acquisition Systems, *2018 3rd International Conference on Communication and Electronics Systems (ICES), Coimbatore, India, 2018*, pp. 934-937.
20. *Zhang X. and Li Y.* Evolutionary Design of an Analog Filter, *2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, Hangzhou, China, 2013*, pp. 474-477.

Статью рекомендовал к опубликованию д.т.н., профессор И.И. Турулин.

Самойлов Леонтий Константинович – Южный федеральный университет; e-mail: lkamoilov@sfedu.ru; г. Таганрог, Россия; тел.: +78634371638; кафедра информационных измерительных технологий и систем; д.т.н.; профессор.

Денисенко Дарья Юрьевна – e-mail: d.y.denisenko@yandex.ru; тел.: +78634371689; кафедра систем автоматического управления; к.т.н.; доцент.

Титов Алексей Евгеньевич – e-mail: aetitov@sfedu.ru; тел.: +78634371689; кафедра систем автоматического управления; к.т.н.; доцент.

Samoilov Leonty Konstantinovich – Southern Federal University; e-mail: lksamoilov@sfnu.ru; Taganrog, Russia; phone: +78634371638; the department of information measurement technologies and systems; dr. of eng. sc.; professor.

Denisenko Darya Yurievna – e-mail: d.y.denisenko@yandex.ru; phone: +78634371689; the department of automatic control systems; cand. of eng. sc.; associate professor.

Titov Alexey Evgenievich – e-mail: aetitov@sfnu.ru; phone: +78634371689; the department of automatic control systems; cand. of eng. sc.; associate professor.

УДК 621.382.32: 621.314

DOI 10.18522/2311-3103-2024-2-175-183

Б.Г. Коноплев

АНАЛИЗ ВЫПРЯМИТЕЛЬНЫХ СВОЙСТВ НАНОМЕТРОВЫХ МОП-ТРАНЗИСТОРОВ В ДИОДНОМ ВКЛЮЧЕНИИ ПРИ СВЕРХНИЗКИХ НАПРЯЖЕНИЯХ

Успехи микроэлектроники, особенно развитие КМОП-технологии, обеспечили возможность создания приборов с чрезвычайно низкой потребляемой мощностью. Это позволило разрабатывать автономные беспроводные приборы, которые обеспечивают с использованием радиоволн не только прием, обработку и передачу информации, но и получение мощности питания от терминалов. Более того, для беспроводного и безбатарейного питания может применяться соби́рание радиочастотной энергии из окружающей среды: энергии излучения станций сотовой связи, радиотелевизионных станций, СВЧ-печей, Wi-Fi, Bluetooth и др. источников. Для преобразования радиочастотной энергии в напряжение питания чаще всего применяются выпрямители на основе нанометровых МОП-транзисторов в диодном включении. Когда устройства с беспроводным питанием находятся далеко от терминала или собирают энергию для питания из окружающей среды, плотность мощности электромагнитного поля и, следовательно, амплитуда входного напряжения может быть весьма малой. Поэтому актуальной является задача разработки и исследования таких устройств, способных работать при очень низких входных напряжениях. Целью исследования является анализ выпрямительных свойств диодов на основе нанометровых МОП-транзисторов в режиме слабой инверсии при сверхнизких входных напряжениях и выработка рекомендаций по выбору технологии и проектированию микросхем с беспроводным питанием. Получены выражения для оценки коэффициентов выпрямления диодов по току и по мощности. Выполнены расчеты по полученным выражениям и моделирование с использованием модели BSIM4v4.8.2 вольтамперных характеристик и зависимостей коэффициентов выпрямления диодов по току и по мощности от напряжения для типовой КМОП-технологии 90 нм. Показана возможность построения выпрямителей на основе МОП-транзисторов при сверхнизких напряжениях вплоть до единиц мВ. Даны рекомендации по обоснованию технологических и конструктивных параметров при проектировании модулей преобразования и соби́рания энергии беспроводных устройств.

Интернет вещей; беспроводная передача; соби́рание и преобразование энергии; диоды; нанометровые МОП-транзисторы; подпороговый режим; коэффициент выпрямления.

B.G. Konoplev

ANALYSIS OF THE RECTIFYING PROPERTIES OF NANOMETER MOS TRANSISTORS IN A DIODE CONNECTION AT ULTRALOW VOLTAGE

Advances in microelectronics, especially the development of CMOS technology, have made it possible to create devices with extremely low power consumption. This made it possible to develop autonomous wireless devices that, using radio waves, not only receive, process, and transmit information, but also receive power from the terminals. For wireless and battery-free power supply, harvesting of radio frequency energy from the environment can be used: radiation energy from cellular stations, radio and television stations, microwave ovens, Wi-Fi, Bluetooth, and other sources. To convert radio frequency energy into supply voltage, rectifiers based on nanometer diode-connected MOSFETs are most often used. When wireless powered devices are located far from the terminal or harvest energy from the environment, the power density of the electromagnetic field and therefore the amplitude of the input voltage can be quite small. The urgent task is to develop and study such devices capable of operating at very low input voltage.

es. The purpose of the study is to analyze the rectifying properties of diodes based on nanometer MOSFETs in weak inversion mode at ultra-low input voltages and to develop recommendations for the choice of technology and design of microcircuits with wireless power. Expressions are obtained for estimating the rectification coefficients of diodes in terms of current and power. Calculations using the obtained expressions and modeling using the BSIM4v4.8.2 model of current-voltage characteristics and dependences of diode rectification coefficients for current and power on voltage for a typical 90 nm CMOS technology were performed. The possibility of constructing rectifiers based on MOSFETs at ultra-low voltages down to units of mV has been demonstrated. Recommendations are given for justifying technological and design parameters when designing modules for converting and harvesting energy of wireless devices.

Internet of Things; wireless transmission; harvesting and conversion of energy; diodes; nanometer MOSFETs; subthreshold mode; rectification coefficient.

Введение. В конце 19-го века практически одновременно с изобретением А.С. Поповым и Г. Маркони радио для передачи информации Н. Тесла провел опыты по беспроводной передаче на большие расстояния электрической энергии значительной мощности посредством излучения и приема радиоволн [1]. Однако, несовершенство аппаратуры в то время и в последующие сто лет не позволило эффективно применять беспроводную передачу энергии для питания приборов вместо использования в качестве источников питания батарей или сети переменного тока.

Успехи микроэлектроники, особенно развитие КМОП-технологии, обеспечили возможность создания приборов с потребляемой мощностью порядка мили-, микро-, нано- и даже пиковатт [2, 3]. Это, в свою очередь, позволило разрабатывать полностью автономные беспроводные приборы, которые обеспечивают с использованием радиоволн не только прием, обработку и передачу информации, но и получение мощности питания от терминалов [4]. Более того, для беспроводного и безбатарейного питания может применяться соби́рание радиочастотной энергии из окружающей среды (Energy Harvesting – EH): энергии излучения станций сотовой связи, радиотелевизионных станций, СВЧ-печей, Wi-Fi, Bluetooth и др. источников [5–8]. Нас окружает «океан» бесплатной энергии [8–13].

Соби́рающие энергию питания из окружающей среды беспроводные устройства радиочастотной идентификации, контроля доступа, систем логистики, медицинских имплантатов, сенсорные сети «интернета вещей» (Internet of Things – IoT) и, далее, «интернета всего» (Internet of Everything – IoE) позволяют обеспечить новый уровень информатизации и повышения эффективности многих сфер жизни человека [14, 15].

Для преобразования радиочастотной энергии в напряжение питания чаще всего применяются выпрямители на основе нанометровых МОП-транзисторов в диодном включении [16, 17]. Это обусловлено полной технологической совместимостью таких выпрямителей в составе КМОП-микросхем, и, следовательно, минимальной стоимостью, что является определяющим при массовом производстве приборов.

Когда устройства с беспроводным питанием находятся далеко от терминала или собирают энергию для питания из окружающей среды, плотность мощности электромагнитного поля и, следовательно, амплитуда входного напряжения может быть весьма малой. Поэтому актуальной является задача разработки и исследования таких устройств, способных работать при очень низких входных напряжениях.

Целью исследования является анализ выпрямительных свойств диодов на основе нанометровых МОП-транзисторов в режиме слабой инверсии при сверхнизких входных напряжениях и выработка рекомендаций по выбору технологии и проектированию микросхем с беспроводным питанием.

Преобразователи радиочастотной энергии на основе МОП-транзисторов при сверхнизких входных напряжениях. На рис. 1 показана упрощенная схема соби́рания радиочастотной энергии [5, 16].

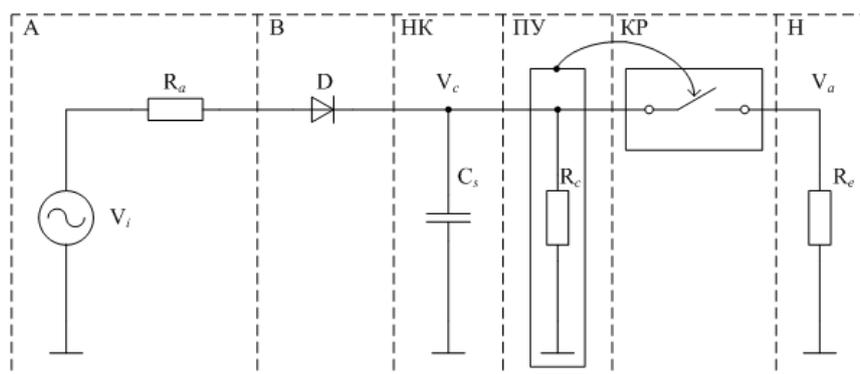


Рис. 1. Упрощенная схема сбора радиочастотной энергии: А – антенна; В – выпрямитель/умножитель напряжения; НК – накопительный конденсатор; ПУ – пороговое управляющее устройство; КР – ключ-регулятор напряжения; Н – нагрузка

Рассмотрим протокол работы микросистемы с сбором энергии, временная диаграмма накопления и расходования энергии которой представлена на рис. 2.

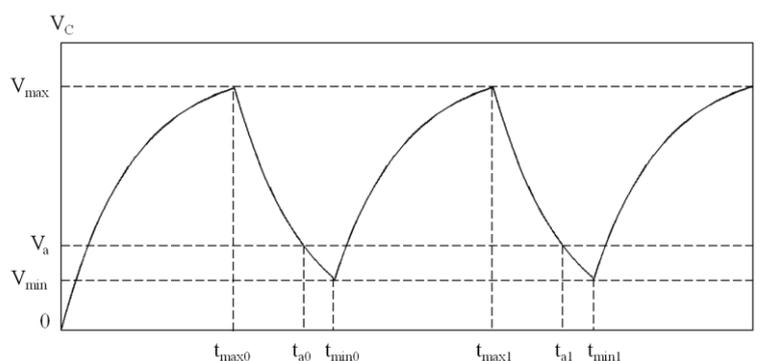


Рис. 2. Временная диаграмма накопления и расходования энергии

В фазе накопления энергии ключ *КР* (см. рис. 1) разомкнут. Эквивалентная схема антенны включает источник высокочастотного напряжения V_i и сопротивление антенны R_a . Принятый антенной сигнал подается на выпрямитель/умножитель напряжения (на рис. 1 условно показан в виде одного диода D). Для повышения чувствительности применяют многокаскадные выпрямители/умножители напряжения на МОП-транзисторах в диодном включении [18]. Обычно между антенной и выпрямителем/умножителем включается схема согласования импедансов [5] (на рис. 1 для упрощения не показана). Собранная из внешней среды энергия накапливается на конденсаторе C_s пока напряжение на нём не достигнет достаточного уровня, при котором возможна активная работа микросистемы (например, обработка и передача данных).

При достижении напряжения на конденсаторе уровня V_{max} пороговое управляющее устройство разрешает замыкание ключа, и микросистема переходит в фазу расходования энергии, передавая ее в нагрузку R_l . Ключ-регулятор в этой фазе поддерживает на нагрузке постоянное напряжение V_a .

Время t_{max0} заряда накопительного конденсатора до напряжения V_{max} зависит от мощности, поступающей в антенну, эффективности выпрямителя, ёмкости конденсатора и уровня мощности, потребляемой пороговым управляющим устройством (на рис. 1 отражено сопротивлением R_c).

В течение достаточно длительного времени даже при небольшой входной мощности P_s накопительный конденсатор может зарядиться собираемой энергией до требуемого для работы микросистемы напряжения. Время накопления энергии t_s равно t_{max0} при холодном старте или $t_{max1} - t_{min0}$ при горячем старте (см. рис. 2).

Накопленный заряд используется для питания микросистемы мощностью P_d в течение времени $t_d = t_{ai} - t_{maxi}$. При $P_d \gg P_s$ длительность фазы накопления энергии значительно больше длительности фазы расходования энергии $t_s \gg t_d$.

Как только напряжение на конденсаторе становится меньше V_{min} , система возвращается в режим собирания энергии. При малом удалении источника мощности, например, в системах оплаты транспортных услуг, $P_d < P_s$ и фазы накопления и расходования энергии совмещаются, тогда снижения напряжения на конденсаторе при $t > t_{maxi}$ (см. рис. 2) не происходит.

Исследование выпрямительных свойств диодов на основе МОП-транзисторов при сверхнизких входных напряжениях. Для повышения эффективности многокаскадных выпрямителей/умножителей напряжения выводы оснований МОП-транзисторов соединяют с выводами стоков, как показано на рис. 3 [18, 19]. Это позволяет для всех каскадов устранить влияние потенциала подложки на пороговое напряжение транзисторов [18]. Кроме того, такое соединение, в отличие от варианта соединения основания с истоком, предотвращает шунтирование МОП-транзистора прямосмещенным pn -переходом сток-основание при больших обратных напряжениях, подаваемых на выпрямитель [19].

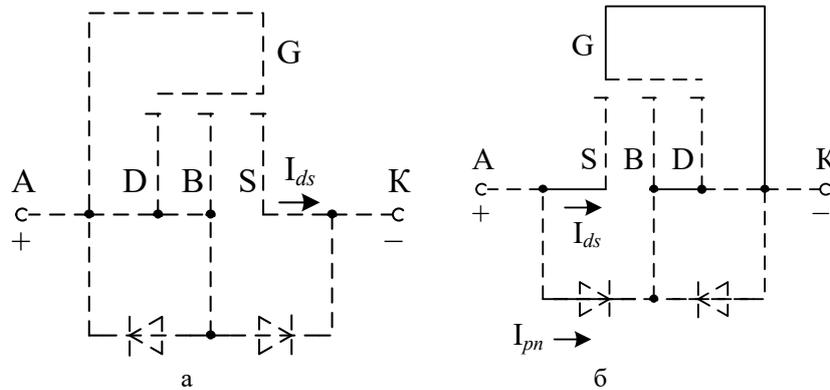


Рис. 3. Диодное включение МОП-транзисторов: а) nМОП; б) pМОП; S – исток; D – сток; B – основание; G – затвор; A – анод диода; K – катод диода; пунктирами показаны p - n -переходы исток-основание, сток-основание

При сверхнизких входных напряжениях МОП-транзисторы выпрямителя/умножителя работают в режиме слабой инверсии [18, 20], а при прямом смещении pn -перехода сток-основание (см. рис. 3) имеет место суперпозиция МОП-транзистора в режиме слабой инверсии и биполярного транзистора исток-основание-сток [21].

Подпороговый ток сток-исток nМОП-транзистора в режиме слабой инверсии определяется выражением [21, 22]:

$$I_{ds} = I_0 \cdot \exp\left(\frac{V_g - V_{th} - n \cdot V_s}{n \cdot \phi_T}\right) \cdot \left(1 - \exp\left(\frac{V_s - V_d}{\phi_T}\right)\right), \quad (1)$$

$$I_0 = 2 \cdot \mu \cdot \frac{\varepsilon \cdot \varepsilon_0}{d} \cdot \frac{W}{L} \cdot n \cdot \phi_T^2,$$

где n – параметр наклона вольтамперной характеристики транзистора в подпороговой области [2]; ϕ_T – тепловой потенциал; μ – подвижность носителей заряда в канале; ε – диэлектрическая проницаемость подзатворного диэлектрика; ε_0 – электрическая постоянная; d – толщина подзатворного диэлектрика; W – эффективная ширина канала; L – эффективная длина канала; V_{th} – пороговое напряжение; V_g , V_s , V_d – напряжения затвора, истока и стока, соответственно, относительно потенциала основания V_b .

В диодном включении (см. рис. 3,а) выводы затвора, стока и основания объединены, и, следовательно, их потенциалы одинаковы. Если положить потенциал основания равным нулю, то $V_b = V_g = V_d = 0$. Тогда из (1) получим:

$$I_{ds} = I_{ds0} \cdot \left(\exp\left(\frac{-V_s}{\varphi_T}\right) - 1 \right), \quad (2)$$

$$I_{ds0} = I_0 \cdot \exp\left(\frac{-V_{th}}{n \cdot \varphi_T}\right).$$

В выражении (2) значения $V_s < 0$ соответствуют прямому смещению диода (см. рис. 3,а) и прямому току I_{ds}^+ , а значения $V_s > 0$ – обратному смещению и обратному току I_{ds} . С учетом этого можно получить выражение для коэффициента выпрямления по току:

$$K_I = \frac{I_{ds}^+}{I_{ds}^-} = \left(\exp\left(\frac{V}{\varphi_T}\right) - 1 \right) / \left(1 - \exp\left(\frac{-V}{\varphi_T}\right) \right), \quad (3)$$

где V – напряжение на диоде.

При работе в выпрямителях в подпороговом режиме ток в цепи заряда конденсатора C_s определяется нелинейной вольтамперной характеристикой (ВАХ) диода на основе МОП-транзистора, так как сопротивление диода значительно превышает сопротивление других участков цепи. Пусть входной сигнал имеет форму меандра. За время положительной части входного импульса t^+ через прямосмещенный диод в конденсатор передается энергия Q^+ определяемая величиной мощности P^+ , выражение для которой можно получить из (2):

$$P^+ = \int_{V_{i-1}}^{V_i} I_{ds}(V) dV = I_{ds0} \cdot \left(\varphi_T \cdot \left(\exp\left(\frac{V_i}{\varphi_T}\right) - \exp\left(\frac{V_{i-1}}{\varphi_T}\right) \right) - V_i + V_{i-1} \right), \quad (4)$$

где V_{i-1} , V_i – напряжение на диоде до и после воздействия положительной части входного i -го импульса.

За время отрицательной части входного импульса t^- через обратносмещенный диод конденсатор разряжается на величину энергии Q^- определяемой мощностью P^- , выражение для которой можно получить из (2):

$$P^- = \int_{-V_i}^{-V_{i+1}} I_{ds}(V) dV = I_{ds0} \cdot \left(\varphi_T \cdot \left(\exp\left(\frac{-V_{i+1}}{\varphi_T}\right) - \exp\left(\frac{-V_i}{\varphi_T}\right) \right) + V_{i+1} - V_i \right), \quad (5)$$

где $-V_i$, $-V_{i+1}$ – напряжение на диоде до и после воздействия отрицательной части входного i -го импульса.

Оценим величину приращения напряжения ΔV_i на конденсаторе за период входного импульса T . За время T при $T \ll t_s$ в конденсатор передается энергия:

$$Q_i = Q_i^+ - Q_i^- \cong P_i^+ \cdot t^+ - P_i^- \cdot t^- \cong (P_i^+ - P_i^-) \cdot T/2 \cong C_s \cdot \Delta V_i^2/2. \quad (6)$$

Тогда из (6):

$$\Delta V_i \cong \sqrt{\left((P_i^+ - P_i^-) \cdot \frac{T}{C_s} \right)}. \quad (7)$$

Коэффициент выпрямления по мощности можно получить из (4, 5) как $K_P = P^+/P^-$. Этот параметр лучше характеризует выпрямительные свойства диодов на основе МОП-транзисторов по сравнению с параметром K_I , так как учитывает нелинейность ВАХ в диапазоне изменения напряжений на диоде.

Для оценки K_P можно положить в (4) $V_{i-1}=0$ и в (5) $V_i=0$. Тогда из (4, 5) получим:

$$K_P = P^+/P^- = \left(\varphi_T \cdot \left(\exp\left(\frac{V}{\varphi_T}\right) - 1 \right) - V \right) / \left(\varphi_T \cdot \left(\exp\left(\frac{-V}{\varphi_T}\right) - 1 \right) + V \right). \quad (8)$$

Из анализа (3, 8) следует, что при выпрямляемом напряжении $V > 0$ коэффициенты выпрямления $K_I > 1$ и $K_P > 1$, и, следовательно, диод на МОП-транзисторе потенциально может сохранять выпрямительные свойства при сколь угодно малых входных напряжениях.

Здесь следует отметить, что выражения (1-5, 8) получены для идеализированной модели, не учитывающей ряд физических процессов в реальных нанометровых МОП-транзисторах [22, 23], а также влияние pn -перехода исток-основание, включенного параллельно МОП-транзистору (см. рис. 3). В нанометровых МОП-транзисторах существенное влияние на токи утечки, кроме подпорогового тока, могут оказывать токи pn -переходов, короткоканальный эффект, ток затвора за счет туннельного эффекта и инжекции горячих носителей, GIDL-эффект (ток утечки, индуцированный затвором на сток или исток). На подпороговый ток существенное влияние оказывают DIBL-эффект (индуцированное стоком или истоком снижение порогового напряжения), эффект подложки, зависимость порогового напряжения от длины и ширины канала [23].

При проектировании СБИС применяют современные САПР, включающие достаточно точные подсистемы моделирования. В настоящей работе в среде схемотехнического моделирования Tanner EDA (T-Spice) использовалась модель BSIM4v4.8.2 [24], учитывающая особенности конструкции и физических процессов в наноразмерных МОП-транзисторах и включающая более двухсот параметров. Параметры модели соответствуют МОП-транзисторам с низкими токами утечки, изготовленным по технологии КМОП 90 нм [25]. В [18] показано, что в многокаскадных выпрямителях/умножителях, выполняемых по типовым КМОП-технологиям с изолированными карманами для размещения p МОП-транзисторов, целесообразно применять именно p -канальные транзисторы.

На рис. 4 представлены результаты расчетов по формулам (2, 3, 8) и моделирования в среде T-Spice характеристик диодов на основе p МОП-транзисторов.

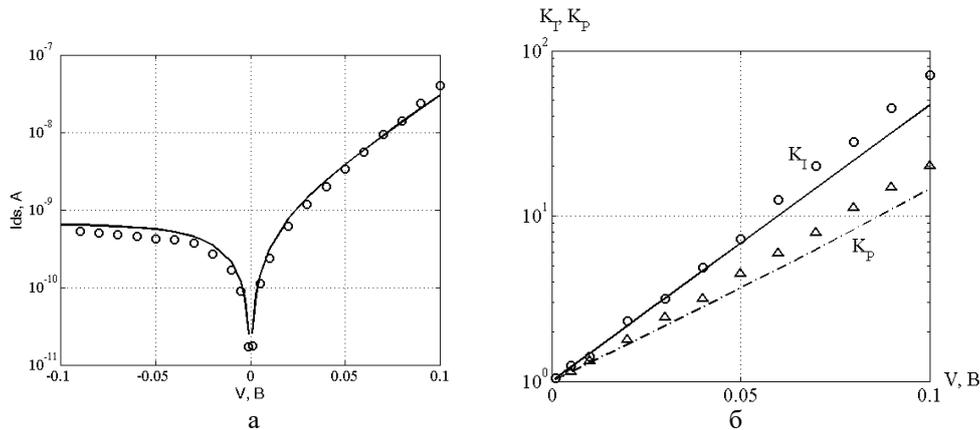


Рис. 4. Вольтамперные характеристики (а) и коэффициенты выпрямления (б) диода при сверхнизких напряжениях; ключевые параметры: $V_{th} = -0.32$ В, $\mu = 0.027$ м/В*с, $d = 1.2$ нм, $W/L = 14.8$, $n = 1.2$; результаты расчетов показаны сплошными и штрих-пунктирными кривыми, результаты моделирования – значками \circ и Δ

Для построения зависимости коэффициента выпрямления K_P от напряжения V применялось численное интегрирование результатов моделирования ВАХ. Отклонение результатов расчета по формулам от более точных результатов моделирования объясняется учетом в модели BSIM4v4.8.2 особенностей конструкции и физических процессов в наноразмерных МОП-транзисторах.

Анализ результатов расчетов и моделирования (см. рис. 4) и формул (4-7) показывает возможность построения выпрямителей на основе МОП-транзисторов при сверхнизких напряжениях вплоть до единиц мВ, при этом время накопления энергии t_s может быть весьма большим. Выбирая электрофизические, структурные и топологические параметры транзисторов в соответствии с выражениями (1, 2), можно в определенных пределах управлять величиной приращения напряжения ΔV_i и временем t_s .

Заключение. В работе проведено исследование выпрямительных свойств диодов на основе МОП-транзисторов при сверхнизких напряжениях, которые являются основой модулей преобразования и сбора энергии беспроводных устройств.

Получены выражения для оценки коэффициентов выпрямления диодов по току и по мощности. Выполнены расчеты по полученным выражениям и моделирование с использованием модели BSIM4v4.8.2 ВАХ и зависимостей коэффициентов выпрямления диодов по току и по мощности от напряжения для типовой КМОП-технологии 90 нм. Показана возможность построения выпрямителей на основе МОП-транзисторов при сверхнизких напряжениях вплоть до единиц мВ. Даны рекомендации по обоснованию технологических и конструктивных параметров при проектировании модулей преобразования и сбора энергии беспроводных устройств.

Результаты могут быть полезны для разработчиков беспроводных микроустройств.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Из истории изобретения и начального периода развития радиосвязи / под. ред. проф. В.Н. Ушакова. – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2008. – 288 с.
2. Wang A., Calhoun B.H., Chandrakasan A.P. Sub-threshold Voltage Circuit Design for Ultra Low Power Systems. – New York: Springer, 2006. – 209 p.
3. Reynders N., Dehaene W. Ultra-Low-Voltage Design of Energy-Efficient Digital Circuits. – New York: Springer, 2015. – 192 p.
4. Shinohara N. Wireless Power Transfer via Radiowaves. – London: ISTE Ltd., 2014. – 238 p.
5. Tran L.-G., Cha H.-K., Park W.-T. RF power harvesting: a review on designing methodologies and applications // Micro and Nano Systems Letters. – 2017. – Vol. 5, No. 14. – P. 1-16.
6. Clerckx B., Zhang R., Schober R., Ng D.W.K., Kim D.I., Poor H.V. Fundamentals of Wireless Information and Power Transfer: From RF Energy Harvester Models to Signal and System Designs // IEEE J. on selected areas in communications. – 2019. – Vol. 37, No. 1. – P. 4-33.
7. Smith J.R. Wirelessly Powered Sensor Networks and Computational RFID. – New York: Springer, 2013. – 271 p.
8. Luo Y., Pu L., Wang G., Zhao Y. RF Energy Harvesting Wireless Communications: RF Environment, Device Hardware and Practical Issues // Sensors. – 2019. – Vol. 19. – Article 3010. – 28 p.
9. Visser H.J., Reniers A.C.F., Theeuwes J.A.C. Ambient RF Energy Scavenging: GSM and WLAN Power Density Measurements // Proceedings of the 38th European Microwave Conference. – October 2008, The Netherlands, Amsterdam. – P. 721-724.
10. Pinuela M., Mitcheson P.D., Lucyszyn S. Ambient RF Energy Harvesting in Urban and Semi-Urban Environments // IEEE Transactions on Microwave Theory and Techniques. – 2013. – Vol. 61, No. 7. – P. 2715-2726.
11. Alhekail Z.O., Hadi M.A., Alkanhal M.A. Public safety assessment of electromagnetic radiation exposure from mobile base stations // J. of radiological protection. – 2012. – Vol. 32. – P. 325-337.
12. Pinuela M., Yates D. C., Mitcheson P.D., Lucyszyn S. London RF Survey for Radiative Ambient RF Energy Harvesters and Efficient DC-load Inductive Power Transfer // 7th European Conference on Antennas and Propagation (EUCAP 2013). – IEEE, 2013. – P. 2839-2843.
13. Vyas R.J., Cook B.B., Kawahara Y., Tentzeris M.M. E-WEHP: A Batteryless Embedded Sensor-Platform Wirelessly Powered from Ambient Digital-TV Signals // IEEE Transactions on microwave theory and techniques. – 2013. – Vol. 61, No 6. – P. 2491-2505.
14. International Roadmap for Devices and Systems. More than Moore White Paper. – IEEE, 2022. – 56 p.
15. International Roadmap for Devices and Systems. Executive Summary. – IEEE, 2022. – 76 p.
16. Коноплев Б.Г., Синюкин А.С. Исследование выпрямителей на основе наноразмерных МОП-транзисторов для микросистем с беспроводным питанием // Известия ЮФУ. Технические науки. – 2018. – № 2 (196). – С. 105-113.
17. Синюкин А.С., Коноплев Б.Г. Исследование влияния параметров наноразмерных МОП-транзисторов на характеристики преобразователей энергии для пассивных беспроводных устройств // Известия ЮФУ. Технические науки. – 2019. – № 6 (208). – С. 15-24.
18. Sinyukin A.S., Konoplev B.G. Integrated CMOS Microwave Power Converter for Passive Wireless Devices // Russian Microelectronics. – 2021. – Vol. 50, No. 3. – P. 189-196.
19. Shokrani M.R., Khoddam M., Hamidon M.N.B., Kamsani N.A., Rokhani F.Z., Shafie S.B. An RF Energy Harvester System Using UHF Micropower CMOS Rectifier Based on a Diode Connected CMOS Transistor // The Scientific World Journal. – 2014. – Vol. 2014. – Article 963709. – 11 p.

20. Enz C.C., Krummenacher F., Vittoz E.A. An Analytical MOS Transistor Model Valid in All Regions of Operation and Dedicated to Low-Voltage and Low-Current Applications // *Analog Integrated Circuits and Signal Processing*. – 1995. – No. 8. – P. 83-114.
21. Enz C.C., Vittoz E.A. Modeling for Low-Voltage and Low-Power Analog IC Design // *Microelectronic Engineering*. – 1997. – Vol. 39. – P. 59-76.
22. Enz C.C., Vittoz E.A. Charge-based MOS Transistor Modeling. The EKV model for low-power and RF IC design. – London: John Wiley & Sons Ltd., 2006. – 303 p.
23. Roy K., Mukhopadhyay S., Mahmoodi-Meimand H. Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicrometer CMOS Circuits // *Proceedings of the IEEE*. – 2003. – Vol. 91, No. 2. – P. 305-327.
24. Hu C., Niknejad A.M., Chauhan S.Y. BSIM4v4.8.2 MOSFET Model – User’s Manual. – USA, CA, Berkeley: University of California, 2020. – 176 p.
25. Sicard E., Bendhia S.D. Basics of CMOS Cell Design. – USA: McGraw-Hill, 2007. – 429 p.

REFERENCES

1. Iz istorii izobreteniya i nachal'nogo perioda razvitiya radiosvyazi [Iz istorii izobreteniya i nachalnogo perioda razvitiya radiosvyazi], ed. by prof. V.N. Ushakova. Saint Petersburg: Izd-vo SPbGETU «LETI», 2008, 288 p.
2. Wang A., Calhoun B.H., Chandrakasan A.P. Sub-threshold Voltage Circuit Design for Ultra Low Power Systems. New York: Springer, 2006, 209 p.
3. Reynders N., Dehaene W. Ultra-Low-Voltage Design of Energy-Efficient Digital Circuits. New York: Springer, 2015, 192 p.
4. Shinohara N. Wireless Power Transfer via Radiowaves. London: ISTE Ltd., 2014, 238 p.
5. Tran L.-G., Cha H.-K., Park W.-T. RF power harvesting: a review on designing methodologies and applications, *Micro and Nano Systems Letters*, 2017, Vol. 5, No. 14, pp. 1-16.
6. Clerckx B., Zhang R., Schober R., Ng D.W.K., Kim D.I., Poor H.V. Fundamentals of Wireless Information and Power Transfer: From RF Energy Harvester Models to Signal and System Designs, *IEEE J. on selected areas in communications*, 2019, Vol. 37, No. 1, pp. 4-33.
7. Smith J.R. Wirelessly Powered Sensor Networks and Computational RFID. New York: Springer, 2013, 271 p.
8. Luo Y., Pu L., Wang G., Zhao Y. RF Energy Harvesting Wireless Communications: RF Environment, Device Hardware and Practical Issues, *Sensors*, 2019, Vol. 19, Article 3010, 28 p.
9. Visser H.J., Reniers A.C.F., Theeuwes J.A.C. Ambient RF Energy Scavenging: GSM and WLAN Power Density Measurements, *Proceedings of the 38th European Microwave Conference. October 2008, The Netherlands, Amsterdam*, pp. 721-724.
10. Pinuela M., Mitcheson P.D., Lucyszyn S. Ambient RF Energy Harvesting in Urban and Semi-Urban Environments, *IEEE Transactions on Microwave Theory and Techniques*, 2013, Vol. 61, No. 7, pp. 2715-2726.
11. Alhekail Z.O., Hadi M.A., Alkanhal M.A. Public safety assessment of electromagnetic radiation exposure from mobile base stations, *J. of radiological protection*, 2012, Vol. 32, pp. 325-337.
12. Pinuela M., Yates D. C., Mitcheson P.D., Lucyszyn S. London RF Survey for Radiative Ambient RF Energy Harvesters and Efficient DC-load Inductive Power Transfer, *7th European Conference on Antennas and Propagation (EUCAP 2013)*. IEEE, 2013, pp. 2839-2843.
13. Vyas R.J., Cook B.B., Kawahara Y., Tentzeris M.M. E-WEHP: A Batteryless Embedded Sensor-Platform Wirelessly Powered from Ambient Digital-TV Signals, *IEEE Transactions on microwave theory and techniques*, 2013, Vol. 61, No 6, pp. 2491-2505.
14. International Roadmap for Devices and Systems. More than Moore White Paper. IEEE, 2022, 56 p.
15. International Roadmap for Devices and Systems. Executive Summary. IEEE, 2022, 76 p.
16. Konoplev B.G., Sinyukin A.S. Issledovanie vypyramiteley na osnove nanorazmernykh MOP-tranzistorov dlya mikrosistem s besprovodnym pitaniem [Research of rectifiers based on nanoscale MOS-devices for microsystems with wireless power supply], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2018, No. 2 (196), pp. 105-113.
17. Sinyukin A.S., Konoplev B.G. Issledovanie vliyaniya parametrov nanorazmernykh MOP-tranzistorov na kharakteristiki preobrazovateley energii dlya passivnykh besprovodnykh ustroystv [Research of the influence of parameters of nano-sized MOSFETs on the characteristics of energy converters for passive wireless devices], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2019, No. 6 (208), pp. 15-24.
18. Sinyukin A.S., Konoplev B.G. Integrated CMOS Microwave Power Converter for Passive Wireless Devices, *Russian Microelectronics*, 2021, Vol. 50, No. 3, pp. 189-196.

19. Shokrani M.R., Khoddam M., Hamidon M.N.B., Kamsani N.A., Rokhani F.Z., Shafie S.B. An RF Energy Harvester System Using UHF Micropower CMOS Rectifier Based on a Diode Connected CMOS Transistor, *The Scientific World Journal*, 2014, Vol. 2014, Article 963709, 11 p.
20. Enz C.C., Krummenacher F., Vittoz E.A. An Analytical MOS Transistor Model Valid in All Regions of Operation and Dedicated to Low-Voltage and Low-Current Applications, *Analog Integrated Circuits and Signal Processing*, 1995, No. 8, pp. 83-114.
21. Enz C.C., Vittoz E.A. Modeling for Low-Voltage and Low-Power Analog IC Design, *Microelectronic Engineering*, 1997, Vol. 39, pp. 59-76.
22. Enz C.C., Vittoz E.A. Charge-based MOS Transistor Modeling. The EKV model for low-power and RF IC design. London: John Wiley & Sons Ltd., 2006, 303 p.
23. Roy K., Mukhopadhyay S., Mahmoodi-Meimand H. Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicrometer CMOS Circuits, *Proceedings of the IEEE*, 2003, Vol. 91, No. 2, pp. 305-327.
24. Hu C., Niknejad A.M., Chauhan S.Y. BSIM4v4.8.2 MOSFET Model – User’s Manual. USA, CA, Berkeley: University of California, 2020, 176 p.
25. Sicard E., Bendhia S.D. Basics of CMOS Cell Design. USA: McGraw-Hill, 2007, 429 p.

Статью рекомендовал к опубликованию д.т.н. И.Е. Лысенко.

Коноплев Борис Георгиевич – Южный федеральный университет; e-mail: kbg@sfedu.ru; г. Таганрог, Россия; д.т.н.; профессор; научный руководитель Института нанотехнологий, электроники и приборостроения.

Konoplev Boris Georgievich – Southern Federal University; e-mail: kbg@sfedu.ru; Taganrog, Russia; dr. of eng. sc.; professor; scientific advisor, Institute of nanotechnologies, electronics, and equipment engineering.

УДК 621.373

DOI 10.18522/2311-3103-2024-2-183-191

А.Н. Зикий, А.С. Кочубей

ТРАНЗИСТОРНЫЙ ГЕНЕРАТОР С ДИЭЛЕКТРИЧЕСКИМ РЕЗОНАТОРОМ

Генераторы, стабилизированные диэлектрическими резонаторами, нашли широкое применение в связи, радиолокации, радионавигации и радиоэлектронной борьбе. Их исследованию посвящено большое число работ, однако повышение требований к ним заставляет разработчиков радиоаппаратуры искать новые технические решения. Особенно важна стабильность частоты генераторов СВЧ в условиях воздействия вибраций, ударов, перепадов температуры, изменения напряжения питания. Влияние внешних воздействующих факторов приводит к появлению паразитных параметров сигнала – уходам частоты, расширению спектральной линии, возрастанию гармоник и субгармоник. Целью данной работы является измерение паразитных параметров выходного сигнала. Объектом исследования является серийный генератор СВЧ. В работе даётся краткое его описание. К генератору предъявляются следующие требования: – расчёт диэлектрического резонатора; – моделирование микрополоскового фильтра из состава умножителя частоты; – рабочая частота $17490 \pm 3,5$ МГц; – выходная мощность не менее 10 дБм; – сопротивление нагрузки 50 Ом; напряжение питания 15 В; ток потребления не более 215 мА. В качестве результатов исследования представлены: – выбег частоты за 15 минут после включения; – зависимость частоты от напряжения питания; – спектр выходного сигнала в полосе 100 кГц; – спектр выходного сигнала в полосе 20 ГГц. Полученные результаты могут быть использованы студентами и преподавателями радиотехнических дисциплин, инженерами – разработчиками радиоаппаратуры. Статья дополняет известные результаты по экспериментальному исследованию генераторов в части паразитных параметров выходного сигнала. Для начинающих исследователей может быть полезна методика эксперимента.

Генератор; стабилизированный диэлектрическим резонатором; схема; конструкция; эксперимент; выбег частоты; зависимость частоты от напряжения питания; спектр; гармоники; субгармоники.

A.N. Zikiy, A.S. Kochubey

TRANSISTOR OSCILLATOR WITH DIELECTRIC RESONATOR

Generators stabilized by dielectric resonators have found wide application in communications, radar, radio navigation and electronic warfare. A large number of works have been devoted to their research, but increasing requirements for them forces radio equipment developers to look for new technical solutions. Especially important is the stability of the frequency of microwave generators under the influence of vibrations, shocks, temperature changes, changes in supply voltage. The influence of external influencing factors leads to the appearance of parasitic signal parameters – frequency departures, expansion of the spectral line, increase of harmonics and subharmonics. The purpose of this work is to measure the parasitic parameters of the output signal. The object of the study is a serial microwave generator. The paper gives a brief description of it. The generator has the following requirements: – calculation of a dielectric resonator; – simulation of a microstrip filter from the frequency multiplier; – operating frequency 17490 ± 3.5 MHz; – output power of at least 10 dBm; – load resistance 50 ohms; – supply voltage 15 V; – consumption current no more than 215 mA. As the results of the study are presented: – frequency run-out in 15 minutes after switching on; – frequency dependence on supply voltage; – the spectrum of the output signal in the 100 kHz band; – the spectrum of the output signal in the 20 GHz band. The results obtained can be used by students and teachers of radio engineering disciplines, engineers – developers of radio equipment. The article complements the well-known results on the experimental study of generators in terms of parasitic parameters of the output signal. For novice researchers, the experimental technique may be useful.

Generator stabilized by a dielectric resonator; circuit; design; experiment; frequency run-out; frequency dependence on supply voltage; spectrum; harmonics; subharmonics.

Введение. Гетеродины приёмников и задающие генераторы передатчиков определяют ряд важных параметров приёмно-передающей и измерительной аппаратуры. В первую очередь – стабильность частоты, диапазон рабочих частот. В ряде случаев параметры гетеродина существенно влияют на чувствительность приёмника, на полезную и паразитную модуляцию задающего генератора передатчика, поэтому их исследование является актуальным.

По генераторам имеется обширная литература, в том числе монографии [1–3], учебное пособие [4], обзоры [5–10], статьи [11–16], диссертации [17, 18], патенты [19, 20], однако на этом тема не исчерпана в связи с широким разнообразием требований, появлением новой элементной базы (транзисторов, микросхем) и новых материалов.

Постановка задачи. Объектом исследования является транзисторный генератор сантиметрового диапазона, стабилизированный диэлектрическим резонатором. Целью исследования является измерение паразитных параметров выходного сигнала, в том числе побочных колебаний в ближней и дальней зоне спектра, выбега частоты, зависимости частоты от напряжения питания. В известных работах такой информации явно недостаточно.

К генератору предъявляются следующие требования:

- ◆ рабочая частота $17490 \pm 3,5$ МГц;
- ◆ выходная мощность не менее 10 дБм;
- ◆ сопротивление нагрузки 50 Ом;
- ◆ напряжение питания 15 В;
- ◆ ток потребления не более 215 мА.

Схема и конструкция. Функциональная схема генератора приведена на рис. 1. Из этой схемы видно, что генератор включает автогенератор на частоту 8745 МГц, вентиль на эту частоту, умножитель частоты на два, вентиль на частоту около 17490 МГц, стабилизатор напряжения. Автогенератор построен на биполярном транзисторе 2Т647А-3 и стабилизирован диэлектрическим резонатором. Умножитель частоты на два построен на полевом транзисторе 3П604Б-2 и содержит в своём составе полосовой фильтр на полуволновых резонаторах.

Конструкцию генератора можно видеть на рис. 2. Корпус генератора выполнен из титана и разделён на 2 отсека. В первом отсеке находится автогенератор на поликоровой печатной плате. Во втором отсеке находится умножитель частоты, два вентиля и стаби-

лизатор напряжения. Диэлектрический резонатор находится под печатной платой автогенератора и связан микрополосковым резонатором через щелевую линию связи в печатной плате. Диэлектрический резонатор укрыт цилиндрической камерой, которая имеет винт плавной подстройки частоты генератора.

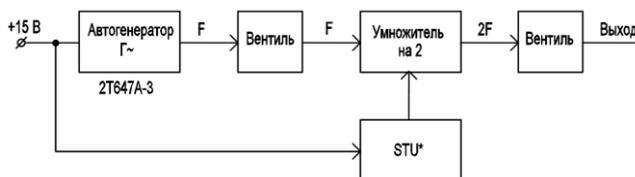


Рис. 1. Функциональная схема генератора



Рис. 2. Фото генератора со снятой крышкой

Расчёт диэлектрического резонатора (ДР). Расчёт ДР проведём по формуле из книги [21]. Резонансная частота равна:

$$f_p = \frac{\sqrt{\beta r^2 + \beta z^2}}{2\pi \cdot \sqrt{\epsilon_0 \cdot \mu_0 \cdot \epsilon_r}}$$

где $\beta r = 4,81/D$; $\beta z = \pi q/L$; D – диаметр резонатора; L – высота резонатора; q – число полу-волн, укладываемых по оси z резонатора, принимаем равным 1.

$$\epsilon_0 = 8,854 \cdot 10^{-12} \text{ Ф/м};$$

$$\mu_0 = 4\pi \cdot 10^{-7} \text{ Г/м}.$$

В качестве материала ДР применяется материал ТБНС с $\epsilon_r = 78$ [21]. В качестве первого приближения используем данные из книги [21]: при $D = 4,5$ мм; $L = 2,5$ мм резонансная частота ДР из ТБНС равна 8,5 ГГц.

Расчёт ведётся графоаналитическим методом. Для этого построим график зависимости f_p от L (рис. 3). На этом графике частоте $f_p = 8745$ МГц соответствует высота резонатора $L = 2,58$ мм (это второе приближение).

Поскольку резонансная частота ДР зависит от наличия вблизи проводящих поверхностей, точное значение резонансной частоты устанавливают в процессе регулировки автогенератора с помощью подстроечного винта, погруженного в камеру с ДР.

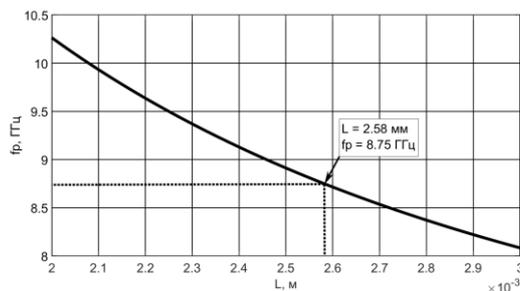


Рис. 3. Зависимость f_p от L для $\epsilon_r = 78$

Моделирование фильтра. Моделирование фильтра проводилось в программе AWR Design Environment. Фильтр должен выделять вторую гармонику сигнала на выходе умножителя частоты и подавлять первую, третью и последующие гармоники частоты 8745 МГц. Поскольку гармоники расположены достаточно далеко друг от друга, для выделения второй гармоники можно использовать простой фильтр, модель которого изображена на рис. 4. Исходные данные для расчёта и моделирования фильтра следующие:

- ◆ центральная частота фильтра 17500 МГц;
- ◆ потери на центральной частоте не более 2 дБ;
- ◆ волновое сопротивление 50 Ом;
- ◆ тип фильтра на полуволновых резонаторах с четвертьволновыми связями;
- ◆ технологическое ограничение – зазоры не менее 100мкм;
- ◆ подложка – поликор с $\varepsilon = 9,8$ и толщиной 0,5 мм;
- ◆ толщина меди 10 мкм.

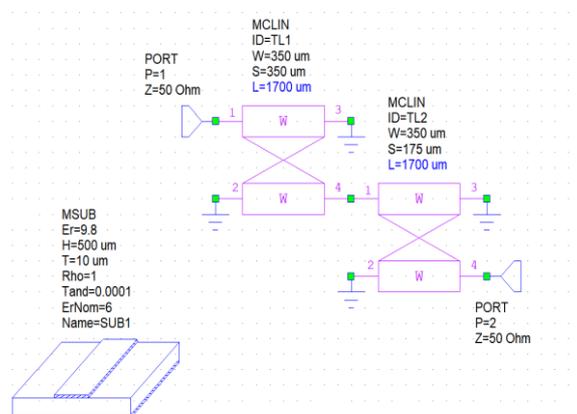


Рис. 4. Модель фильтра в программе AWR Design Environment

На рис. 5 можно видеть амплитудно-частотную характеристику (АЧХ) фильтра, полученную в результате моделирования. Из рис. 5 видно, что фильтр настроен на частоту 17500 МГц и имеет потери 1,176 дБ, что удовлетворяет требованиям.

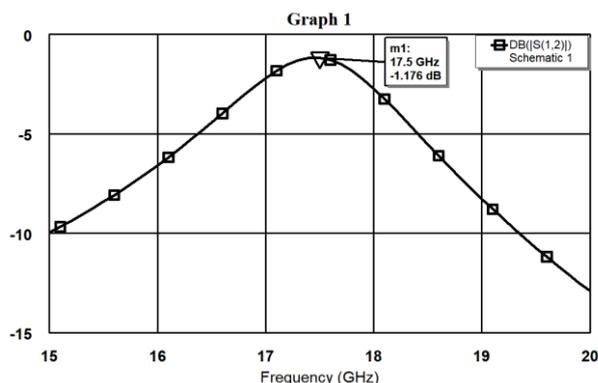


Рис. 5. АЧХ фильтра в программе AWR Design Environment

Эксперимент. Эксперименты проводились на установке, структурная схема которой приведена на рисунке 6.

В первом эксперименте проводилось измерение выбега частоты за 15 минут. Результаты измерений заносились в табл. 1. По данным табл. 1 построен график на рис. 7. Из этого графика видно, что за 15 минут частота генератора уменьшилась на 550 кГц.

Во втором эксперименте измерена зависимость частоты на выходе генератора от напряжения питания. Результаты измерений занесены в табл. 2. По данным табл. 2 построен график на рис. 8, из которого видно, что при изменении напряжения питания от 10,3 до 15 В частота увеличивается на 946 кГц. Средняя крутизна этой характеристики равна 225 кГц/В.

В третьем эксперименте проводилось фотографирование спектра выходного сигнала. На рис. 9 можно видеть спектр выходного сигнала в полосе 100 кГц. Ширина спектра сигнала на уровне минус 20 дБ от максимума не превышает 1 кГц. На рис. 10 показан спектр выходного сигнала в полосе 20 ГГц. Из него видно, что автогенератор работает на частоте 8745 МГц, а на выходе удвоителя частоты получается сигнал с частотой 17490 МГц. При этом частота 8745 МГц подавлена всего на 25 дБ, что недостаточно для качественного гетеродина. Кроме того, имеется третья гармоника исходного сигнала на частоте 26235 МГц, подавленная на 40 дБ.



Рис. 6. Структурная схема измерительной установки

Таблица 1

Выбег частоты при $U_{пит} = 15 В$

t , мин	$f_{вых}$, кГц	t , мин	$f_{вых}$, кГц
0	17489546	8	17489148
1	17489462	9	17489122
2	17489403	10	17489094
3	17489348	11	17489072
4	17489301	12	17489051
5	17489252	13	17489032
6	17489216	14	17489013
7	17489179	15	17488996

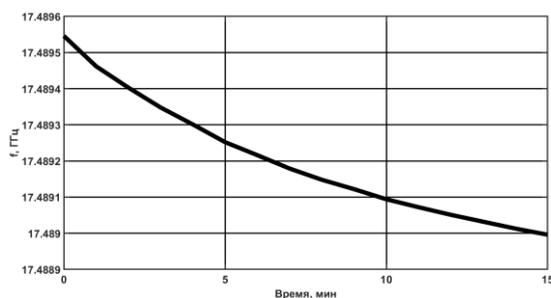


Рис. 7. Выбег частоты за 15 минут

Таблица 2

Зависимость частоты от напряжения питания

$U_{пит}$, В	$f_{вых}$, кГц
10.3	17488055
10.5	17488186
11	17488316

Окончание табл. 2

$U_{пит}, В$	$f_{вых}, кГц$
11.5	17488447
12	17488512
12.5	17488692
13	17488838
13.5	17488838
14	17488963
14.5	17489001
15	17488968

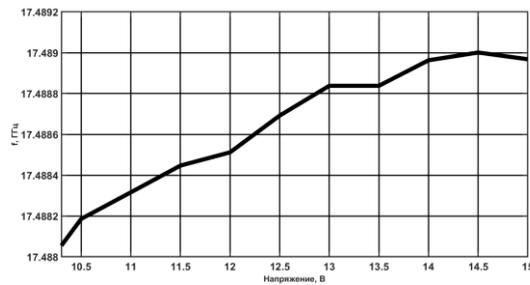


Рис. 8. Зависимость частоты от напряжения питания

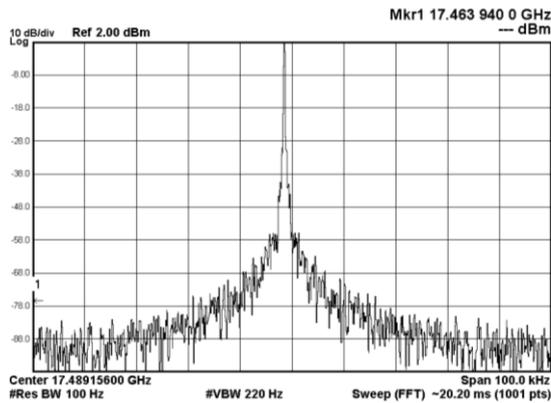


Рис. 9. Спектр сигнала на выходе генератора при полосе обзора 100 кГц

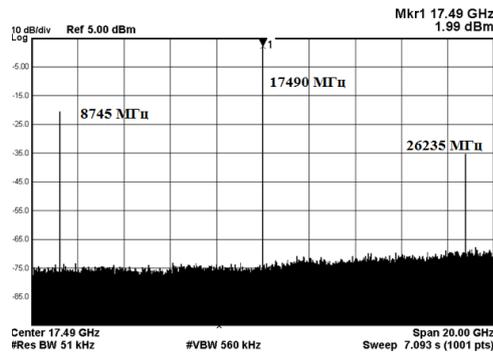


Рис. 10. Спектр сигнала на выходе генератора при полосе обзора 20 ГГц

Заключение. В процессе экспериментального исследования выявлены следующие паразитные параметры выходного сигнала:

- ◆ Ширина спектра сигнала на уровне минус 20 дБ от максимума, не более 1 кГц.
 - ◆ Уровень побочных спектральных составляющих на гармониках и субгармониках не более минус 25 дБ.
 - ◆ Выбег частоты за 15 минут после включения не более 550 кГц.
 - ◆ Уход частоты при изменении напряжения питания от 10,3 В до 15 В составил 946 кГц.
 - ◆ Температурный коэффициент частоты (ТКЧ) отрицательный.
- Основные параметры генератора:
- ◆ Выходная частота $17488 \pm 0,6$ МГц.
 - ◆ Выходная мощность 12 дБм.
 - ◆ Напряжение питания 15 В.
 - ◆ Ток потребления 180 мА.

При его использовании в качестве гетеродина приёмника рекомендуется использовать гетеродинный фильтр для подавления гармоник и субгармоник.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Баранов А.В., Кревский М.А. Транзисторные автогенераторы гармонических СВЧ колебаний. – М.: Горячая линия – Телеком, 2021. – 276 с.
2. Grebennikov A. RF and Microwave transistor oscillator design. – Chichester, England: John Wiley & Sons, Ltd. 2007. – 441 p.
3. Odinec M. RF and Microwave transistor oscillator design. – Artech House. Boston, London, 2002. – 398 p.
4. Зикий А.Н., Помазанов А.В. Стабильность частоты генераторов СВЧ: учеб. пособие. – Ростов-на-Дону – Таганрог: Изд-во ЮФУ, 2017. – 138 с.
5. Алексейчик Л.В., Бродуленко И.И., Гаврилюк Н.Г., Краюшкин В.В., Мальцев В.А. Параметры и методы анализа ДР и колебательных систем на их основе // Обзоры по электронной технике. Серия 1. Электроника СВЧ. Вып. 5 (1525). – М.: ЦНИИ «Электроника», 1990. – 64 с.
6. Алексейчик Л.В. и др. Состояние и перспективы применения миниатюрных ДР в радиоэлектронике. Ч. 1. Параметры миниатюрных ДР на СВЧ и методы их расчёта // Обзоры по электронной технике. Серия 1. Электроника СВЧ. – 1981. – Вып. 13 (832). – 98 с.
7. Алексейчик Л.В., Бродуленко И.И., Геворкян В.М., Казанцев Ю.А., Парышкин Л.А. Состояние и перспективы применения миниатюрных диэлектрических резонаторов в радиоэлектронике. Ч. II. Пассивные и активные СВЧ – устройства с миниатюрными диэлектрическими резонаторами // Обзоры по электронной технике. Серия I. Электроника СВЧ. Вып. 2 (865). – М.: ЦНИИ «Электроника», 1982. – 66 с.
8. Афанасьев А.И., Алексейчик Л.В., Бродуленко И.И., Гаврилюк Н.Г., Мальцев В.А. Состояние, тенденции развития и проектирование стабильных твердотельных генераторов и фильтров СВЧ малой мощности // Обзоры по электронной технике. Серия 1. Электроника СВЧ. Вып. 2 (1968). – М.: ЦНИИ «Электроника», 1993.
9. Бродуленко И.И., Абраменков А.И., Ковтунов Д.А., Лебедев В.Н., Сергиенко А.М. Стабильные и высокочастотные СВЧ – генераторы на диэлектрических резонаторах // Обзор по электронной технике. Серия 1. Электроника СВЧ. Вып. 10 (1454). – М.: ЦНИИ «Электроника», 1989. – 61 с.
10. Бродуленко И.И., Афанасьев А.И., Ковтунов Д.А., Лебедев В.Н., Мальцев В.А., Иванов В.В. Стабильные СВЧ генераторы малой мощности. (Проблемы качества и надёжности) // Обзоры по электронной технике. Серия 1. СВЧ техника. Вып. 1 (1696). – М.: ЦНИИ «Электроника», 1993. – 88 с.
11. Зикий А.Н., Пирогов Д.Ю., Чернышова О.А., Сальный И.А. Генератор с диэлектрическим резонатором // Вопросы специальной радиоэлектроники, серия Общие вопросы радиоэлектроники. – 2008. – № 1. – С. 142-146.
12. Зикий А.Н., Шитулин М.В. Транзисторный автогенератор, стабилизированный диэлектрическим резонатором // Матер. 4 международной научной конференции «Современные проблемы радиоэлектроники». – Ростов-на-Дону, ЮРГУЭС, 2012. – С. 285-287.
13. Бровченко С.П., Зикий А.Н., Чернышова О.А. СВЧ транзисторный автогенератор, стабилизированный диэлектрическим резонатором // Известия ЮФУ. Технические науки. – 2008. – № 3. – С. 182-184.

14. Зикий А.Н., Кочубей А.С. Экспериментальное исследование гетеродинного модуля // Известия ЮФУ. Технические науки. – 2020. – № 6. – С. 189-196.
15. Белов Л.А., Хилькевич В. Генераторы СВЧ с диэлектрическими резонаторами для стабилизации частоты // Электроника. Наука. Технология. Бизнес. – 2006. – № 7.
16. Зикий А.Н., Зламан П.Н., Сивокос Е.В. Экспериментальное исследование двух автогенераторов с диэлектрическими резонаторами // Успехи современной радиоэлектроники. – 2015. – № 10. – С. 198-209.
17. Бажилов В.А. Расчёт и исследование цилиндрических экранированных СВЧ и КВЧ колебательных систем на основе диэлектрических резонаторов: автореф. дисс. ... канд. техн. наук. – Нижний Новгород, НГТУ, 2007. – 20 с.
18. Царанкин Д.П. Методы генерирования СВЧ колебаний с минимальным уровнем фазовых шумов: автореф. дисс. ... д-ра техн. наук. – М.: МЭИ, 2004.
19. Патент США №6.714.089 High frequency signal source and method of generating same using dielectric resonator oscillator (DRO) circuit. 30.03.2004.
20. Патент США №4.386.326. Dielectric-resonator-tuned Microwave solid state oscillator. 30.05.1983.
21. Радиопередающие устройства / под ред. О.А. Челнокова. – М.: Радио и связь, 1982. – 256 с.

REFERENCES

1. Baranov A.V., Krevskiy M.A. Tranzistornyye avtogeneratory garmonicheskikh SVCh kolebaniy [Transistor autogenerators of harmonic microwave oscillations]. Moscow: Goryachaya liniya – Telekom, 2021, 276 p.
2. Grebennikov A. RF and Microwave transistor oscillator design. Chichester, England: John Wiley & Sons, Ltd. 2007, 441 p.
3. Odinec M. RF and Microwave transistor oscillator design. Artech House. Boston, London, 2002, 398 p.
4. Zikiy A.N., Pomazanov A.V. Stabil'nost' chastoty generatorov SVCh: ucheb. posobie [Frequency stability of microwave generators: tutorial. Rostov-on-Donu – Taganrog: Izd-vo YuFU, 2017, 138 p.
5. Alekseychik L.V., Brodulenko I.I., Gavrilyuk N.G., Krayushkin V.V., Mal'tsev V.A. Parametry i metody analiza DR i kolebatel'nykh sistem na ikh osnove [Parameters and methods of analysis of DR and oscillatory systems based on them], *Obzory po elektronnoy tekhnike. Seriya 1. Elektronika SVCh* [Reviews on electronic technology. Series 1. Electronics Microwave]. Issue 5 (1525). Moscow: TSNII «Elektronika», 1990, 64 p.
6. Alekseychik L.V. i dr. Sostoyanie i perspektivy primeneniya miniatyurnykh DR v radioelektronike. Ch. 1. Parametry miniatyurnykh DR na SVCh i metody ikh rascheta [The state and prospects of using miniature DRS in radio electronics. Part 1. Parameters of miniature microwave devices and methods of their calculation], *Obzory po elektronnoy tekhnike. Seriya 1. Elektronika SVCh* [Reviews on electronic technology. Series 1. Microwave Electronics], 1981, Issue 13 (832), 98 p.
7. Alekseychik L.V., Brodulenko I.I., Gevorkyan V.M., Kazantsev Yu.A., Paryshkuro L.A. Sostoyanie i perspektivy primeneniya miniatyurnykh dielektricheskikh rezonatorov v radioelektronike. Ch. II. Passivnye i aktivnye SVCh – ustroystva s miniatyurnymi dielektricheskimi rezonatorami [The state and prospects of using miniature dielectric resonators in radio electronics. Part II. Passive and active microwave devices with miniature dielectric resonators], *Obzory po elektronnoy tekhnike. Seriya 1. Elektronika SVCh* [Reviews on electronic technology. Series I. Microwave Electronics]. Issue 2 (865). Moscow: TSNII «Elektronika», 1982, 66 p.
8. Afanas'ev A.I., Alekseychik L.V., Brodulenko I.I., Gavrilyuk N.G., Mal'tsev V.A. Sostoyanie, tendentsii razvitiya i proektirovanie stabil'nykh tverdotel'nykh generatorov i fil'trov SVCh maloy moshchnosti [State, development trends and design of stable solid-state generators and low-power microwave filters], *Obzory po elektronnoy tekhnike. Seriya 1. Elektronika SVCh* [Reviews on electronic technology. Series 1. Microwave Electronics]. Issue 2 (1968). Moscow: TSNII «Elektronika», 1993.
9. Brodulenko I.I., Abramnikov A.I., Kovtunov D.A., Lebedev V.N., Sergienko A.M. Stabil'nye i vysokochastotnye SVCh – generatory na dielektricheskikh rezonatorakh [Stable and high-frequency microwave generators on dielectric resonators], *Obzor po elektronnoy tekhnike. Seriya 1. Elektronika SVCh* [An overview of electronic technology. Series 1. Microwave electronics]. Issue 10 (1454). Moscow: TSNII «Elektronika», 1989, 61 p.
10. Brodulenko I.I., Afanas'ev A.I., Kovtunov D.A., Lebedev V.N., Mal'tsev V.A., Ivanov V.V. Stabil'nye SVCh generatory maloy moshchnosti. (Problemy kachestva i nadezhnosti) [Stable microwave generators of low power. (Quality and reliability issues)], *Obzory po elektronnoy tekhnike. Seriya 1. SVCh tekhnika* [Reviews on electronic technology. Series 1. Microwave equipment]. Issue 1 (1696), Moscow: TSNII «Elektronika», 1993, 88 p.

11. Zikiy A.N., Pirogov D.Yu., Chernyshova O.A., Sal'nyy I.A. Generator s dielektricheskim rezonatorom [Generator with a dielectric resonator], *Voprosy spetsial'noy radioelektroniki, seriya Obshchie voprosy radioelektroniki* [Questions of special radio electronics, series General Questions of Radio Electronics], 2008, No. 1, pp. 142-146.
12. Zikiy A.N., Shipulin M.V. Tranzistornyy avtogenerator, stabilizirovanny dielektricheskim rezonatorom [Transistor autogenerator stabilized by a dielectric resonator], *Mater. 4 mezhdunarodnoy nauchnoy konferentsii «Sovremennye problemy radioelektroniki»* [Materials of the 4th International scientific conference "Modern problems of radio electronics"]. Rostov-on-Donu, YuRGUES, 2012, pp. 285-287.
13. Brovchenko S.P., Zikiy A.N., Chernyshova O.A. SVCh tranzistornyy avtogenerator, stabilizirovanny dielektricheskim rezonatorom [Microwave transistor generator stabilized by a dielectric resonator], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2008, No. 3, pp. 182-184.
14. Zikiy A.N., Kochubey A.S. Eksperimental'noe issledovanie geterodinnogo modulya [Experimental study of a heterodyne module], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2020, No. 6, pp. 189-196.
15. Belov L.A., Khil'kevich V. Generatory SVCh s dielektricheskimi rezonatorami dlya stabilizatsii chastoty [Microwave generators with dielectric resonators for frequency stabilization], *Elektronika. Nauka. Tekhnologiya. Biznes* [Electronics. The science. Technology. Business], 2006, No. 7.
16. Zikiy A.N., Zlaman P.N., Sivokoz E.V. Eksperimental'noe issledovanie dvukh avtogeneratorov s dielektricheskimi rezonatorami [Experimental study of two autogenerators with dielectric resonators], *Uspekhi sovremennoy radioelektroniki* [Successes of modern radio electronics], 2015, No. 10, pp. 198-209.
17. Bazhilov V.A. Raschet i issledovanie tsilindricheskikh ekranirovannykh SVCh i KVCh kolebatel'nykh sistem na osnove dielektricheskikh rezonatorov: avtoref. diss. ... kand. tekhn. nauk [Calculation and investigation of cylindrical shielded microwave and HF oscillatory systems based on dielectric resonators: abstract of the cand. of eng. sc. diss.]. Nizhniy Novgorod, NGTU, 2007, 20 p.
18. Tsarapkin D.P. Metody generirovaniya SVCh kolebaniy s minimal'nym urovnem fazovykh shumov: avtoref. diss. ... d-ra tekhn. nauk [Methods of generating microwave oscillations with a minimum level of phase noise: abstract of the dr. of eng. sc. diss.]. Moscow: MEI, 2004.
19. Patent SShA №6.714.089 High frequency signal source and method of generating same using dielectric resonator oscillator (DRO) circuit. 30.03.2004 [US Patent No. 6,714,089 A source of a high-frequency signal and a method for generating it using a dielectric resonator generator (DRO) circuit. 30.03.2004].
20. Patent SShA №4.386.326. Dielectric-resonator-tuned Microwave solid state oscillator. 30.05.1983 [US Patent No.4.386.326. Microwave solid-state generator tuned to a dielectric resonator. 30.05.1983].
21. Radiopere dayushchie ustroystva [Radio transmitting devices], ed. by O.A. Chelnokova. Moscow: Radio i svyaz', 1982, 256 p.

Статью рекомендовал к опубликованию к.т.н. М.И. Дулин.

Зикий Анатолий Николаевич – Таганрогский научно-исследовательский институт связи; e-mail: zikiy50@mail.ru; г. Таганрог, Россия; к.т.н.; с.н.с.; доцент.

Кочубей Алексей Сергеевич – e-mail: l.co4ubey@yandex.ru; тел.: 89896270939; магистр; инженер-конструктор 2 категории.

Zikiy Anatoliy Nikolaevich – Taganrog Research Institute of Communications; e-mail: zikiy50@mail.ru; Taganrog, Russia; cand. of eng. sc.; senior researcher; associate professor.

Kochubey Alexey Sergeevich – e-mail: l.co4ubey@yandex.ru; phone: +79896270939; master's degree; design engineer, category 2.

УДК 621.372.852

DOI 10.18522/2311-3103-2024-2-191-202

Д.В. Харланов, А.В. Лабынцев

ПОТЕНЦИАЛЬНЫЕ ВОЗМОЖНОСТИ ФИЛЬТРА НА ШПИЛЕЧНЫХ РЕЗОНАТОРАХ С МЕТАЛЛИЗИРОВАННЫМИ ОТВЕРСТИЯМИ

В задачах проектирования микрополосковых фильтров большую роль играет удачный выбор начального приближения для значений искомым геометрических размеров элементов фильтра. Особенно это важно, если проектируемый фильтр обладает новыми конструктивными признаками, и методы синтеза этого устройства еще находятся на стадии разработки. В работе рас-

сматриваются некоторые результаты проектирования полосовых микрополосковых фильтров, состоящих из противоположенных шпилечных резонаторов с металлизированным отверстием в середине каждого резонатора. Такое техническое решение позволяет существенно расширить полосу заграждения фильтра за счет подавления паразитной полосы пропускания, которая неизбежно формируется в традиционном фильтре на полуволновых шпильках в связи с возникновением резонанса на удвоенной частоте основной полосы пропускания. Введение в резонатор металлизированных отверстий приводит к необходимости исследования свойств этих резонаторов и к определению потенциальных возможностей полосового фильтра, состоящего из таких резонансных элементов. Важнейшей характеристикой полосового фильтра, задаваемой в любом техническом задании на проектирование, является ширина основной полосы пропускания и ее положение на частотной оси. В работе получены оценки относительной ширины полосы пропускания фильтра в зависимости от ширины реализуемых значений полосового проводника, образующего шпильку, и от зазора между плечами шпильки. В результате определяются потенциальные возможности рассматриваемых фильтров по реализуемой или относительной ширине основной полосы пропускания. Под физической реализуемостью фильтра понимается возможность технологически реализовать значения всех геометрических размеров элементов фильтра, исходя из накладываемых на них разумных ограничений. Эти ограничения определяются технологическими допусками для минимальных геометрических размеров и габаритными и частотными ограничениями для максимальных размеров. В работе определен диапазон реализуемых геометрических размеров. Методика решения задачи синтеза фильтра основана на переходе от фильтра к $2n$ -полюснику и в рамках данной работы она используется как инструмент для определения потенциальных возможностей исследуемых фильтров в электродинамической среде моделирования HFSS. Реализуемые значения относительной ширины полосы пропускания фильтра лежат в пределах от 10 % до 32 %, что позволяет отнести данный тип фильтра к устройствам с умеренной шириной полосы пропускания. Приводятся результаты проектирования фильтров на двух шпилечных резонаторах для случаев предельно узких и предельно широких полос пропускания.

Микрополосковый фильтр; шпилечный резонатор; металлизированное отверстие; полоса пропускания; полоса заграждения.

D.V. Kharlanov, A.V. Labyntsev

POTENTIAL CAPABILITIES OF THE FILTER ON HAIRPIN RESONATORS WITH METALLIZED HOLES

In the design tasks of microstrip filters, a good choice of the initial approximation for the values of the required geometric dimensions of the filter elements plays an important role. This is especially important if the filter being designed has new design features and synthesis methods for this device are still under development. The paper considers some results of the design of band-pass microstrip filters consisting of counter-directional hairpin resonators with a metallized hole in the middle of each resonator. Such a technical solution makes it possible to significantly expand the barrier band of the filter by suppressing the parasitic bandwidth, which is inevitably formed in a traditional filter on half-wave resonators due to the occurrence of resonance at twice the frequency of the main bandwidth. The introduction of metallized holes into the resonator leads to the need to study the properties of these resonators and to determine the potential capabilities of a bandpass filter consisting of such resonant elements. The most important characteristic of a bandpass filter, specified in any design specification, is the width of the main bandwidth and its position on the frequency axis. Estimates of the relative bandwidth of the filter are obtained depending on the width of the microstrip conductor forming the hairpin and on the gap between the half of the hairpin. As a result, the potential capabilities of the filters in question are determined by the relative bandwidth they implement. The physical feasibility of the filter is understood as the ability to technologically realize the values of all geometric dimensions of the filter elements, based on reasonable restrictions imposed on them. These limits are determined by technological tolerances for minimum geometric dimensions and dimensional and frequency limits for maximum dimensions. The paper defines a range of realizable geometric dimensions. The method of solving the problem of filter synthesis is based on the transition from a filter to a $2n$ pole, and in this work it is used as a tool to determine the potential capabilities of the studied filters in the electrodynamic environment of HFSS modeling. The realized values of the relative bandwidth of the filter range from 10% to 32%, which makes it possible to attribute this type of filter to devices with moderate bandwidth. The results of designing filters on two hairpin resonators for cases of extremely narrow and extremely wide bandwidth are presented.

Microstrip filter; hairpin resonator; metallized hole; bandwidth; barrier strip.

Введение. В настоящее время полосовые микрополосковые фильтры (МПФ) широко используются для частотной селекции сигналов малой мощности в диапазоне от 300 МГц до 20 ГГц. Анализ последних публикаций [1–6] в области разработки МПФ, а также методов их проектирования показал, что данная тематика по-прежнему актуальна.

В отдельный класс полосовых МПФ можно выделить фильтры, состоящие из так называемых шпилечных резонаторов, которые представляют собой полуволновые резонаторы, свернутые пополам. Существенным недостатком традиционных шпилечных фильтров является наличие второго резонанса на частоте воздействия, при которой на геометрической длине резонатора укладывается целая длина волны в линии. Это приводит к возникновению первой паразитной полосы пропускания в окрестности частоты, равной удвоенному значению средней частоты основной полосы пропускания (ОПП), а это, в свою очередь, существенно сужает верхнюю полосу заграждения фильтра.

Для расширения полосы заграждения ранее было предложено решение [7–12], позволяющее устранить указанную паразитную полосу пропускания путем преобразования простого полуволнового резонатора в совокупность двух четвертьволновых резонаторов за счет введения в центр резонатора металлизированного отверстия. Это позволило расширить полосу заграждения фильтра более чем в два раза.

С другой стороны, одним из основных требований, предъявляемых к характеристике фильтра, является значение относительной ширины ОПП устройства. Эту заданную ширину ОПП фильтр должен обеспечить при реализуемых и разумных геометрических параметрах конструкции. Например, гребенчатый микрополосковый фильтр даже при минимальных значениях зазоров между резонаторами, не позволяет реализовать полосу пропускания более 15%, а фильтр на встречных стержнях легко реализует полосу в 50–60%. Поэтому еще на начальном этапе проектирования важно знать потенциальные возможности конкретного фильтра по диапазону реализуемых им значений относительной ширины ОПП.

Кроме того, в начале проектирования разработчику фильтра желательно знать приблизительные первоначальные значения ширины резонаторов и зазоров между ними, при которых заданная ширина ОПП реализуется. Такая информация уже в самом начале разработки фильтра позволит выбрать хорошее начальное приближение для искомых геометрических размеров, что существенно сократит продолжительность проектирования и практически всегда гарантированно приведет к успешному результату.

В данной работе приводятся оценки реализуемых значений относительной ширины ОПП одного шпилечного резонатора с металлизированным отверстием в зависимости от ширины резонатора и зазора между плечами шпильки. Тем самым определяются потенциальные возможности рассматриваемых фильтров по реализуемой ими относительной ширине ОПП. Под реализуемостью фильтра понимается возможность технологически реализовать значения геометрических размеров его элементов, исходя из накладываемых на них разумных ограничений. Эти ограничения определяются технологическими допусками для минимальных геометрических размеров и габаритными и частотными ограничениями для максимальных размеров.

Постановка задачи и методика исследования. На рис. 1 приведен чертеж фильтра, образованного двумя шпилечными резонаторами с металлизированными отверстиями в каждом резонаторе.

Задача исследования заключается в определении относительной ширины полосы пропускания фильтра при заданных значениях ширины проводника w и зазора между резонаторами s_1 и s_2 . Чтобы определить фактическую ширину полосы пропускания фильтра при заданной ширине полосковых проводников и заданном зазоре между ними, необходимо полностью решить задачу синтеза этого фильтра. В противном случае характеристика фильтра не будет оптимальной, а, следовательно, ширина полосы пропускания устройства будет определена неправильно. Важно, что абсолютное значение длины резонансных элементов l_1 и l_2 на результаты решения поставленной задачи не влияет, поскольку относительная ширина полосы пропускания не зависит от длины резонаторов.

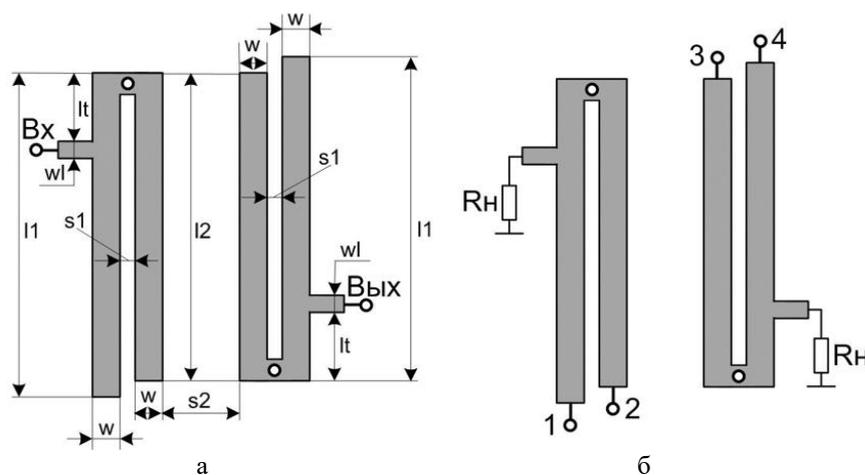


Рис. 1. Чертежи устройств из двух шпилек: а – фильтр; б – 8-полосник

Методика решения задачи синтеза фильтра (рис. 1,а) основана на переходе к $2n$ -полоснику (рис. 1,б) и рассмотрена в работах [13–19]. В данной работе эта методика дополнительно не рассматривается, а используется лишь как инструмент для определения потенциальных возможностей исследуемых фильтров в электродинамической среде моделирования HFSS [20]. На рис. 2 приведена конструктивная модель 8-полосника, реализованная в HFSS в соответствии с рис. 1,б).

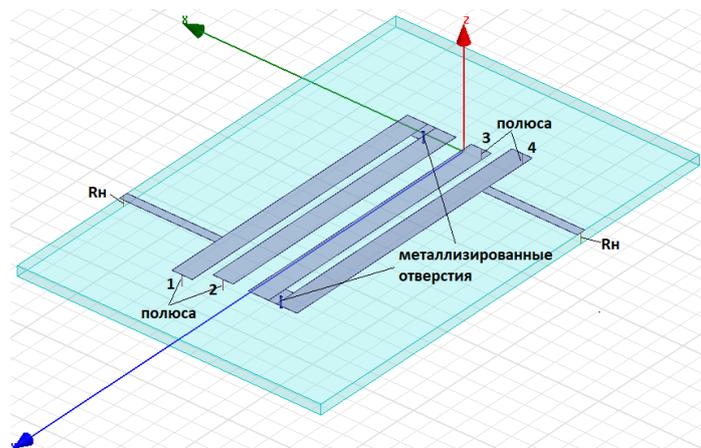


Рис. 2. Модель 8-полосника в HFSS

Определение диапазона реализуемых геометрических размеров. Вопрос реализуемости геометрических размеров в первую очередь относится к ширине и длине резонатора и к зазору между резонаторами.

Во-первых, минимальные значения ширины резонаторов w_{min} и зазоров s_{min} между ними определяются технологическими допусками на изготовление микрополосковых элементов. При любом способе изготовления платы, напыление это или травление, погрешность реализации линейных размеров полосковых проводников составляет около 5 микрон. Это говорит о том, что задавать размеры проводников с точностью более 5 микрон бессмысленно, так как это не удастся реализовать технологически в реальном фильтре. С другой стороны, те же абсолютные допуски в 5 микрон ограничивают минимальные размеры ширины полосковых проводников и зазоров между ними, поскольку

при соизмеримых с допусками значениях геометрических размеров мы получаем недопустимо большие относительные разбросы размеров, резко снижающие повторяемость характеристик фильтра. Как правило, полагают, что минимальная ширина полоскового проводника w_{min} и минимальный зазор между ними s_{min} составляют 100 мкм.

Во-вторых, максимальные размеры ширины полосковых элементов w_{max} и зазоров s_{max} тоже имеют разумные ограничения, поскольку они, прежде всего, определяют габариты фильтра. Кроме того, ширина резонатора должна быть в 4-5 раз меньше, чем его длина, которая, в свою очередь, определяется частотным диапазоном ОПП фильтра. Так для частоты 1 ГГц длина четвертьволнового резонатора (при относительной диэлектрической проницаемости подложки 10) составляет около 25 мм, и, следовательно, $w_{max} = 5$ мм. При частоте 5 ГГц $w_{max} = 1$ мм, а при 25 ГГц $w_{max} = 200$ микрон. Данный фактор в принципе ограничивает частотный диапазон использования микрополосковой технологии для реализации фильтров 50 гигагерцами, так как на этих частотах $w_{max} = w_{min} = 100$ мкм.

Увеличение зазоров между резонаторами s_{max} , как правило, обеспечивает уменьшение связи между ними, что дает возможность реализации фильтра с узкой полосой пропускания. Однако при проектировании узкополосных фильтров лучше идти по пути выбора подходящего топологического решения, например, использовать гребенчатый фильтр, который потенциально является узкополосным при небольших зазорах между резонаторами, а, следовательно, будет оптимальным результатом с точки зрения компактности устройства. В данной работе при исследовании свойств фильтра, образованного шпилечными резонаторами с металлизированными отверстиями, для определенности положим $s_{max} = 5$ мм.

Итак, диапазоны возможных значений ширины резонаторов и зазоров между ними:

$$w \in [0,1 \text{ мм}, 5 \text{ мм}]; s \in [0,1 \text{ мм}, 5 \text{ мм}]. \quad (1)$$

Конечно, все границы диапазонов (1) при специфических требованиях технического задания на проектирование можно расширить.

Неоднозначность в определении ширины ОПП. Прежде всего, надо определиться, по какому критерию будет определяться ширина полосы пропускания устройства. Здесь возможны различные случаи, которые приведены на рис. 3.

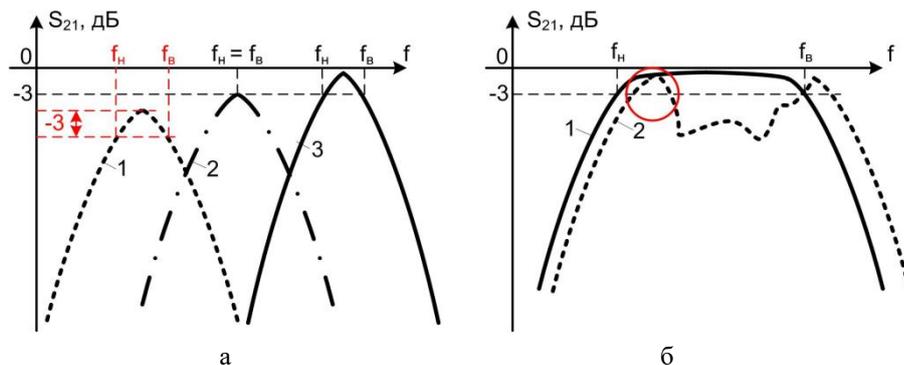


Рис. 3. К определению ширины ОПП в случае: а – узких полос пропускания; б – широких полос пропускания

Количественно относительная ширина ОПП выражается в процентах и определяется как отношение абсолютной ширины ОПП к средней частоте ОПП из соотношения:

$$\delta_{\text{П}} = \frac{2 \cdot (f_b - f_n)}{f_b + f_n} \cdot 100\%,$$

где f_n , f_b – соответственно нижняя граничная частота и верхняя граничная частота ОПП, определенные по уровню -3 дБ функции $20 \lg S_{21}(f)$.

Прежде всего, на ширину полосы пропускания, независимо от ее ширины, сильно влияет наличие диссипативных потерь (потери на тепло), которые обусловлены ненулевым значением тангенса угла потерь в реальной диэлектрической подложке. Этот фактор не позволяет обеспечить в реальном фильтре нулевой уровень затухания S_{21} ни на одной частоте в пределах полосы пропускания.

Особенно сильно это влияние на функцию $S_{21}(f)$ сказывается в случае узкополосных фильтров. На рис. 3,а представлено три различных варианта характеристик фильтра со сравнительно узкой ОПП. И уже здесь возникает вопрос о том, по какому уровню отсчитывать граничные частоты полосы пропускания – минус 3 дБ функции $20\lg S_{21}(f)$ от нулевого уровня затухания или от уровня начальных потерь. Например, по варианту 1 на рис. 3,а можно считать, что полоса пропускания на уровне минус 3 дБ функции $20\lg S_{21}(f)$, отсчитанного от нулевого уровня, отсутствует, а отсчитанная от уровня начальных потерь (изображено красным цветом) имеет вполне определенные граничные частоты. Важно заметить, что возникновение результирующей характеристики на рис. 3,а возможно не только в случае большого уровня диссипативных потерь, но и в случае, когда не обеспечено надлежащее согласование в пределах полосы пропускания. Видимо, неоднозначность в решении этого вопроса должна сниматься в конкретном ТЗ, где должно быть оговорен и максимальный уровень начальных потерь, и допустимая неравномерность затухания в полосе пропускания, и конкретный уровень затухания, по которому определяются границы полосы пропускания.

Чтобы снять эту неопределенность в нашем исследовании, будем считать, что граничные частоты полосы пропускания определяются по уровню минус 3 дБ от уровня начальных потерь при обеспечении достаточно хорошего согласования.

Что касается сравнительно широкой полосы пропускания, то два возможных варианта характеристики затухания для этого случая изображены на рис. 3,б. Первый случай, представленный сплошной кривой, иллюстрирует хороший результат проектирования фильтра, имеющего сравнительно широкую ФПП. А вот во втором случае, очевидно, что результат проектирования широкополосного фильтра явно не получился. Кстати, с этой ситуацией разработчик может столкнуться не один раз и именно при попытке проектирования узкополосного фильтра. Диапазон частот, выделенный красной окружностью на рис. 3,б, отмечен здесь как диапазон ОПП, заданный по ТЗ. Разработчик, увлекшись подбором геометрии для обеспечения этой необходимой по ТЗ полосы пропускания, не замечает, что в непосредственной близости от нее, чуть выше (как на рис. 3,б или ниже по частоте есть еще одна вредная ПП, наличие которой обесценивает все его усилия.

Приведенные выше рассуждения показывают, что определение ширины ПП фильтра и, тем более, ее предельных значений, не является простой и однозначной задачей.

Результаты исследования. В табл. 1 приведены результаты оценок относительной ширины ОПП одного шпилечного резонатора в зависимости от ширины полоскового проводника, образующего шпильку, и зазора между резонансными плечами шпильки. На рис. 4 представлены эти зависимости в виде семейства графиков.

Таблица 1

Оценки относительной ширины ОПП одной шпильки V , %

$S1=S$, мм \ W , мм	0,2	0,5	1,0	3,0	5,0
0,5	25	24,7	22,3	14	11
1,0	30	28	25,4	16,3	12,6
2,0	34	32	28,6	19,3	15,3
3,5	37	35,4	31,6	22,1	17,8
5,0	40	38	33,5	24	19,4

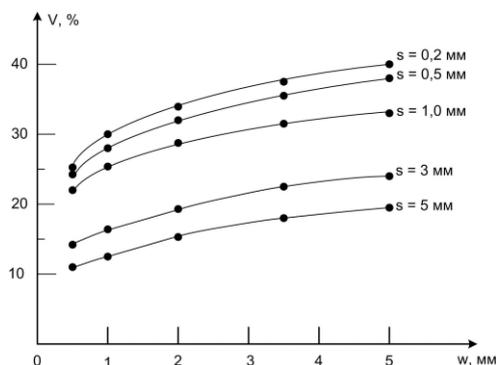


Рис. 4. Зависимости относительной ширины полосы пропускания шпилечного резонатора от ширины проводника и зазора между плечами

Представленные результаты показывают, что реализуемые ширины полос пропускания одного шпилечного резонатора с металлизированным отверстием на сгибе шпильки лежат в пределах от 10 до 40%. Предельные значения ширины полосы пропускания достижимы в двух крайних случаях: минимальная ширина ОПП ($\approx 10-11\%$) реализуется при максимальном зазоре $s_{max} = 5$ мм и минимальной ширине полоскового проводника $w_{min} = 0,5$ мм. И наоборот, максимальная ширина ОПП ($\approx 40\%$) реализуется при минимальном зазоре $s_{min} = 0,2$ мм и максимальной ширине полоскового проводника $w_{max} = 5$ мм.

Графики на рис. 4 позволяют разработчику обоснованно выбрать первоначальные геометрические размеры шпильки по относительной ширине ЗПП. Например, при заданной $V = 20\%$ и ширине полоскового проводника $w = 2$ мм можно выбрать начальное значение $s_0 = 2,5$ мм, а при $w = 4$ мм начальное значение s_0 уже будет $s_0 = 4$ мм. Очевидно, что такой различный выбор начальных приближений приведет к существенно различным результатам решения задачи синтеза, и габариты фильтра во втором случае будут в два раза больше, чем в первом. Это еще раз говорит о неоднозначности решения задачи синтеза, что в свою очередь дает возможность оптимизировать решение по различным критериям.

Ширина ОПП фильтра из двух шпилек. Ранее было показано, что минимальная ширина ОПП ($\approx 10\%$) реализуется одной шпилькой при максимальном зазоре $s_{max} = 5$ мм и минимальной ширине полоскового проводника $w_{min} = 0,5$ мм, а максимальная ширина ОПП ($\approx 40\%$) – при минимальном зазоре $s_{min} = 0,2$ мм и максимальной ширине полоскового проводника $w_{max} = 5$ мм. Приведем результаты исследования влияния зазора между двумя шпильками на ширину ОПП в этих двух крайних случаях.

При рассмотрении исследуемой конструкции устройства, состоящего из двух шпилек и представленной на рис. 1, для упрощения задачи будем считать, что фильтр обладает центральной геометрической симметрией, и что все плечи шпилек имеют одинаковую ширину w . Чертежи указанной конструкции фильтра и соответствующего 8-полосника, а также модель 8-полосника для осуществления синтеза фильтра, реализованная в HFSS, приведены на рис. 1.

Исследование предельно узких полос пропускания. Синтез фильтра при $s_1 = 5$ мм и $w = 0,5$ мм дал приемлемый результат при $s_2 = 3$ мм. Однако характеристики фильтра с соответствующими размерами, представленные на рис. 5, показывают, что начальные потери в ОПП составили 2,8 дБ, а ширина ОПП, определенная по уровню – 5,8 дБ, составила 7%.

Существенное увеличение зазора s_2 с 3 до 7 мм привело к росту начальных потерь с 3 до 10 дБ, и ОПП фильтра практически не сформировалась. Исходя из этих результатов, можно сделать вывод о том, что при узких полосковых проводниках, образующих шпильку, начальные вносимые потери слишком велики за счет низкой добротности резонаторов.

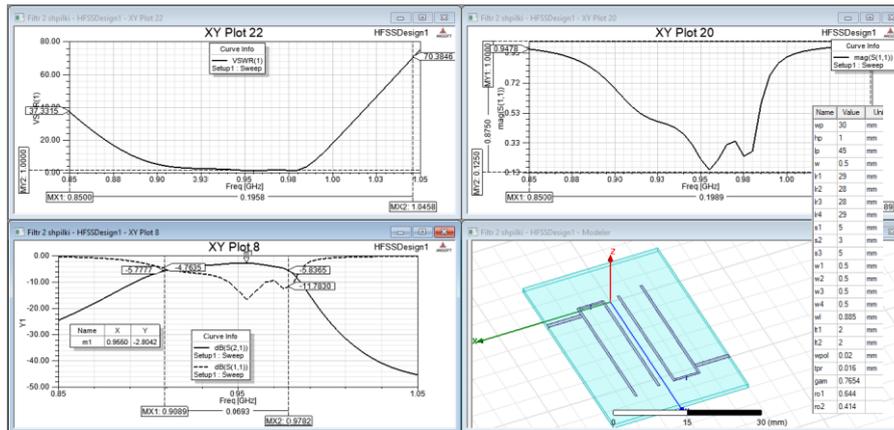


Рис. 5. Характеристики фильтра при $w = 0,5$ мм, $s_1 = 5$ мм и $s_2 = 3$ мм

Увеличение ширины полосковых проводников в два раза (с 0,5 до 1 мм) существенно изменило ситуацию. Результаты проектирования для $w = 1$ мм приведены на рис. 6.

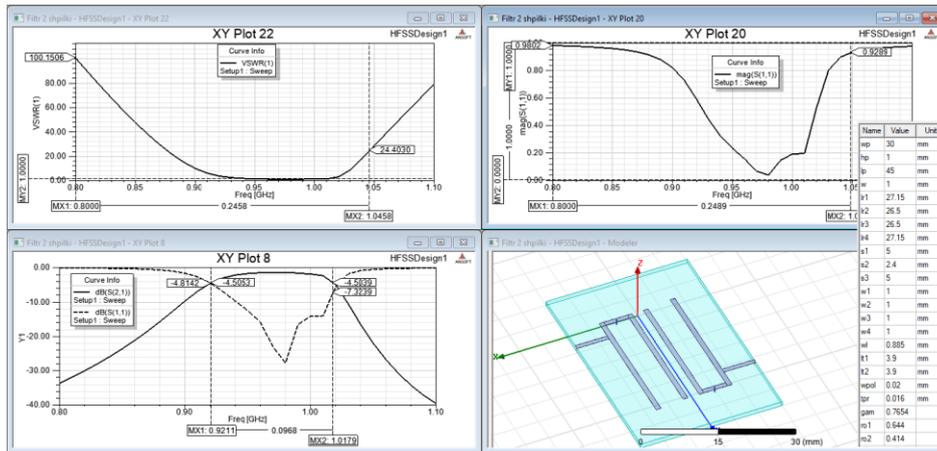


Рис. 6. Характеристики фильтра при $w = 1$ мм, $s_1 = 5$ мм и $s_2 = 2,4$ мм

Ширина ОПП, согласно результатам на рис. 8, составила 10,1%.

Таким образом, исследование предельно узких полос пропускания, реализуемых рассматриваемыми фильтрами, показало, что с ростом ширины полосковых проводников, образующих шпильку, при сравнительно больших зазорах ширина ОПП растет согласно графикам, приведенным на рис. 4. Результаты проектирования фильтров на двух резонаторах при $w = 2$ мм показали ширину ОПП $\delta_\phi = 14\%$, а при $w = 4$ мм относительная ширина ОПП составила $\delta_\phi = 23\%$.

Исследование предельно широких полос пропускания. Исходя из результатов табл. 1 фильтр с максимальной шириной ОПП должен иметь максимальную ширину резонаторов и минимальный зазор между ними. В результате проектирования 8-полосника при $s_1 = 0,2$ мм и $w = 5$ мм не удалось получить удовлетворительный результат, поэтому было взято значение $w = 4$ мм и после нескольких итераций был получен приемлемый результат, представленный на рис. 7.

На рис. 8 приведены характеристики фильтра, размеры которого соответствуют размерам 8-полосника на рис. 7.

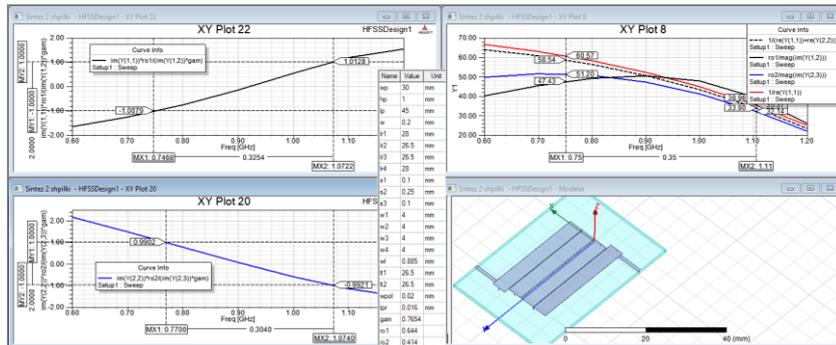


Рис. 7. Характеристики 8-полосника при $s_1 = 0,1$ мм и $w = 4$ мм

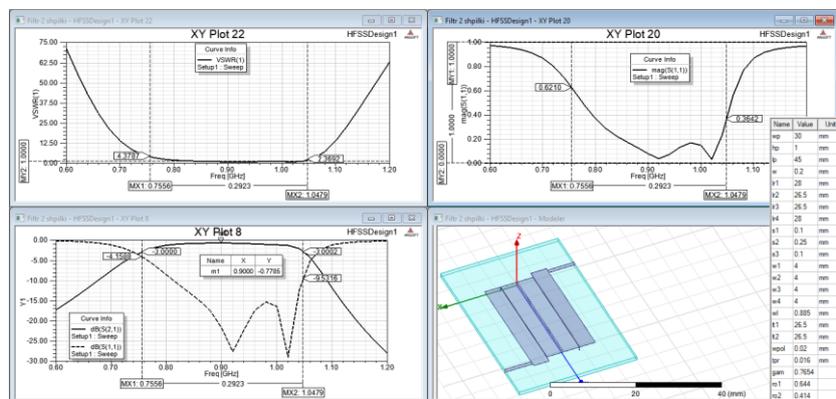


Рис. 8. Характеристики фильтра при $s_1 = 0,1$ мм и $w = 4$ мм

Ширина ФПП фильтра составляет

$$\delta_\phi = \frac{2 \cdot (1,05 - 0,76)}{1,05 + 0,76} \cdot 100\% = 32\%.$$

Ширина ОПП несколько меньше, чем предполагалось по данным из табл. 1, поэтому был выполнен синтез еще для случая $s_1 = 0,2$ мм и $w = 5$ мм. Результат синтеза приведен на рис. 9. Он показывает, что ширина ОПП фильтра практически не изменилась и составляет те же 32%.

$$\delta_\phi = \frac{2 \cdot (0,98 - 0,71)}{0,98 + 0,71} \cdot 100\% = 32\%.$$

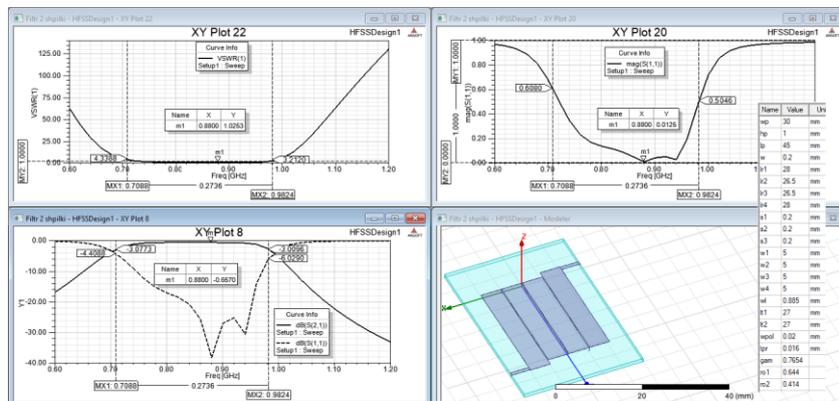


Рис. 9. Результирующие характеристики широкополосного фильтра

Заключение. В результате проведенного исследования были установлены потенциальные возможности фильтра на противонаправленных шпильках с металлизированными отверстиями. Реализуемые значения относительной ширины полосы пропускания фильтра лежат в пределах от 10 % до 32 %, то есть данный тип фильтра относится к устройствам с умеренной шириной полосы пропускания. Полученные результаты позволят разработчику данного фильтра осознанно выбрать начальные значения геометрических размеров устройства по заданной ширине полосы пропускания фильтра.

Материалы статьи подготовлены при поддержке проекта № ВнГр/23-01-РТ «Исследование программно определяемых радиосистем» в Южном федеральном университете".

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Аристархов Г.М., Кириллов И.Н.* Компактные микрополосковые фильтры с повышенной частотной избирательностью на основе полуволновых резонаторов // Системы синхронизации, формирования и обработки сигналов. – 2020. – Т. 11, № 2. – С. 40-44. – DOI: 10.1109/SYNCHROINFO49631.2020.9166122.
2. *Аристархов Г.М., Кириллов И.Н., Корчагин А.И., Кувишинов В.В.* Компактные высокоизбирательные микрополосковые фильтры на свернутых сонаправленных шпильчных резонаторах // Радиотехника. – 2021. – Т. 85, № 4. – С. 126-137. – DOI: 10.18127/j00338486-202104-14.
3. *Аристархов Г.М., Аринин О.В., Кириллов И.Н.* Многополосные полосно-пропускающие микрополосковые фильтры на двух сонаправленных шпильчных резонаторах // Системы синхронизации, формирования и обработки сигналов. – 2023. – Т. 14, № 1. – С. 4-9.
4. *Сержантов А.М.* Резонансные полосковые структуры и частотно-селективные устройства на их основе с улучшенными характеристиками: дисс. ... д-ра техн. наук. – Красноярск, 2015.
5. *Патент № 2775868 С1.* Рос. Федерация, 2022. Высокоселективный микрополосковый полосно-пропускающий фильтр / *Беляев Б.А., Сержантов А.М., Ходенков С.А.*
6. *Патент № 2797166 С1.* Рос. Федерация, 2023. Микрополосковый полосно-пропускающий фильтр / *Беляев Б.А., Сержантов А.М., Ходенков С.А.*
7. Patent US5066933. Band-pass filter. Publication Date 19.11.1991.
8. Patent US20020057143. High frequency filter, filter device, and electronic apparatus incorporating the same. Publication Date 16.05.2002.
9. *Xiao F., Norgren M.* Compact third-order microstrip bandpass filter using hybrid resonators // Clinics in colon and rectal surgery. – Thieme Medical Publishers, 2011. – Vol. 19. – P. 93-106.
10. *Агафонов В.М., Лабынцев В.А., Лабынцев А.В.* Микрополосковый гребенчатый фильтр. А.С. СССР №1262602. Опубл. В БИ, 1986, № 37.
11. *Лабынцев А.В., Игнатьев В.В., Васильев В.В., Михайловский Вит. А., Харланов Д.В.* Проектирование фильтра из шпильчных резонаторов с металлизированными отверстиями с помощью программной системы // Программные продукты и системы. – 2023. – Т. 36, № 4. – С. 668-677. – DOI: 10.15827/0236-235X.142.668-677. <http://www.swsys.ru/archive/2023-4.pdf>.
12. *Obukhovets V.A., Labyntsev A.V., Kharlanov D.V., Ponomarev N.E., Sarzhanov A.N.* Example of Filter Synthesis on Two Anti-directional Stud Resonators with One Metalized Hole in Each Resonator // 2023 Radiation and Scattering of Electromagnetic Waves (RSEMW). – 2023. – P. 160-163. – DOI: 10.1109/rsemw.58451.2023.10202019.
13. *Агафонов В.М., Лабынцев А.В.* Синтез микрополосковых гребенчатых фильтров // Известия Вузов СССР. Радиоэлектроника. – 1987. – № 2. – С. 95. Рукопись депонирована в ВИНТИ 29.08.86. №6286-В86.
14. *Агафонов В.М., Лабынцев А.В.* Синтез микрополосковых фильтров заданной структуры // Все-союзная научно-техническая конференция «Проблемы математического моделирования и реализации радиоэлектронных систем СВЧ на объемных интегральных схемах: Тезисы докладов. – М., 1987. – С. 73.
15. *Агафонов В.М.* Полиномиальные фильтры СВЧ // Радиотехника и электроника. – 1970. – Т. 15, № 10. – С. 2191-2193.
16. *Лабынцев А.В.* Пассивные селективные устройства СВЧ на базе многомодовых микрополосковых линий: дисс. ... канд. техн. наук. – Таганрог: ТРТИ, 1987. – 201 с.
17. *Маттей Д.Л., Янг Л., Джонс Е.М.Т.* Фильтры СВЧ, согласующие цепи и цепи связи: пер. с англ. / под ред. Л.В. Алексеева и Ф.В. Кушнера. – М.: Связь, 1971. – 440 с.
18. *Labyntsev A.V., Poveshenko L.I., Kharlanov D.V.* Two-Stage Design of Microstrip Filters from Arbitrary Configuration Resonators // 2019 Radiation and Scattering of Electromagnetic Waves (RSEMW). – 2019. – P. 172-175. – <https://ieeexplore.ieee.org/document/8792770>. – DOI: 10.1109/RSEMW.2019.8792770.

19. *Labyntsev A.V., Poveshenko L.I., Kharlanov D.V.* Microstrip Filter with Extended Barrier // 2019 Radiation and Scattering of Electromagnetic Waves (RSEMW). – 2019. – P. 176-179. – <https://ieeexplore.ieee.org/document/8792722>. – DOI: 10.1109/RSEMW.2019.8792722.
20. *Банков С.Е., Курушин А.А.* Расчёт антенн и СВЧ структур с помощью HFSS Ansoft. – М.: ЗАО «НПП „Родник“», 2009. – 256 с.

REFERENCES

1. *Aristarkhov G.M., Kirillov I.N.* Kompaktnye mikropoloskovye fil'try s povyshennoy chastotnoy izbiratel'nost'yu na osnove poluvolnovykh rezonatorov [Compact microstrip filters with increased frequency selectivity based on half-wave resonators], *Sistemy sinkhronizatsii, formirovaniya i obrabotki signalov* [Systems of synchronization, signal generation and processing], 2020, Vol. 11, No. 2, pp. 40-44. DOI: 10.1109/SYNCHROINFO49631.2020.9166122.
2. *Aristarkhov G.M., Kirillov I.N., Korchagin A.I., Kuvshinov V.V.* Kompaktnye vysokoizbiratel'nye mikropoloskovye fil'try na svernutykh sonapravlennykh shpilechnykh rezonatorakh [Compact highly selective microstrip filters on folded co-directional stud resonators], *Radiotekhnika* [Radio engineering], 2021, Vol. 85, No. 4, pp. 126-137. DOI: 10.18127/j00338486-202104-14.
3. *Aristarkhov G.M., Arinin O.V., Kirillov I.N.* Mnogopolosnye polosno-propuskayushchie mikropoloskovye fil'try na dvukh sonapravlennykh shpilechnykh rezonatorakh [Multiband band-pass microstrip filters on two co-directional stud resonators], *Sistemy sinkhronizatsii, formirovaniya i obrabotki signalov* [Systems of synchronization, signal generation and processing], 2023, Vol. 14, No. 1, pp. 4-9.
4. *Serzhantov A.M.* Rezonansnye poloskovye struktury i chastotno-selektivnye ustroystva na ikh osnove s uluchshennymi kharakteristikami: diss. ... d-ra tekhn. nauk [Resonant strip structures and frequency-selective devices based on them with improved characteristics: dr. of eng. sc. diss.]. Krasnoyarsk, 2015.
5. *Belyaev B.A., Serzhantov A.M., Khodenkov S.A.* Patent № 2775868 C1. Ros. Federatsiya, 2022. Vysokoselektivnyy mikropoloskovyy polosno-propuskayushchiy fil'tr [Patent RF, No. 2775868 C1, 2022. Highly selective microstrip band-pass filter].
6. *Belyaev B.A., Serzhantov A.M., Khodenkov S.A.* Patent № 2797166 C1. Ros. Federatsiya, 2023. Mikropoloskovyy polosno-propuskayushchiy fil'tr [Patent RF, No. 2797166 C1, 2023. Microstrip band-pass filter].
7. Patent US5066933. Band-pass filter. Publication Date 19.11.1991.
8. Patent US20020057143. High frequency filter, filter device, and electronic apparatus incorporating the same. Publication Date 16.05.2002.
9. *Xiao F., Norgren M.* Compact third-order microstrip bandpass filter using hybrid resonators, *Clinics in colon and rectal surgery*. Thieme Medical Publishers, 2011, Vol. 19, pp. 93-106.
10. *Agafonov V.M., Labyntsev V.A., Labyntsev A.V.* Mikropoloskovyy grebenchatyy fil'tr. A.S. SSSR №1262602. Opubl. V BI, 1986, № 37 [Microstrip comb filter. USSR No.1262602. Publ. In BI, 1986, No. 37].
11. *Labyntsev A.V., Ignat'ev V.V., Vasil'ev V.V., Mikhaylovskiy Vit. A., Kharlanov D.V.* Pro-ektirovanie fil'tra iz shpilechnykh rezonatorov s metallizirovannymi otverstiyami s pomoshch'yu programmnoy sistemy [Software system for designing a filter made of hairpin resonators with metallized holes], *Programmnye produkty i sistemy* [Software products and systems], 2023, Vol. 36, No. 4, pp. 668-677. DOI: 10.15827/0236-235X.142.668-677. Available at: <http://www.swsys.ru/archive/2023-4.pdf>.
12. *Obukhovets V.A., Labyntsev A.V., Kharlanov D.V., Ponomarev N.E., Sarzhanov A.N.* Example of Filter Synthesis on Two Anti-directional Stud Resonators with One Metalized Hole in Each Resonator, 2023 *Radiation and Scattering of Electromagnetic Waves (RSEMW)*, 2023, pp. 160-163. DOI: 10.1109/rsemw 58451.2023.10202019.
13. *Agafonov V.M., Labyntsev A.V.* Sintez mikropoloskovykh grebenchatykh fil'trov [Synthesis of microstrip comb filters], *Izvestiya Vuzov SSSR. Radioelektronika* [Izvestiya of Universities of the USSR, Radioelectronics], 1987, No. 2, pp. 95. The manuscript was deposited in VINITI on 08/29/86. No.6286-In 86.
14. *Agafonov V.M., Labyntsev A.V.* Sintez mikropoloskovykh fil'trov zadannoy struktury [Synthesis of microstrip filters of a given structure. – In the collection: Abstracts of reports], *Vsesoyuznaya nauchno-tehnicheskaya konferentsiya «Problemy matematicheskogo modelirovaniya i realizatsii radioelektronnykh sistem SVCh na ob'emnykh integral'nykh skhemakh: Tezisy dokladov* [All-Union Scientific and Technical Conference "Problems of mathematical modeling and implementation of microwave radioelectronic systems on volumetric integrated circuits: Abstracts of reports]. Moscow, 1987, pp. 73.
15. *Agafonov V.M.* Polinomial'nye fil'try SVCh [Polynomial microwave filters], *Radiotekhnika i elektronika* [Radio Engineering and Electronics], 1970, Vol. 15, No. 10, pp. 2191-2193.

16. *Labyntsev A.V.* Passivnye selektivnye ustroystva SVCh na baze mnogomodovykh mikro-poloskovykh liniy: diss. ... kand. tekhn. nauk [Passive selective microwave devices based on multimode microstrip lines: cand. of eng. sc. diss.]. Taganrog: TRTI, 1987, 201 p.
17. *Mattey D.L., Yang L., Dzhons E.M.T.* Fil'try SVCh, soglasuyushchie tsepi i tsepi svyazi [Microwave filters, matching circuits and communication circuits]: transl. from engl., ed. by L.V. Alekseeva and F.V. Kushnira. Moscow: Svyaz', 1971, 440 p.
18. *Labyntsev A.V., Poveshenko L.I., Kharlanov D.V.* Two-Stage Design of Microstrip Filters from Arbitrary Configuration Resonators, 2019 *Radiation and Scattering of Electromagnetic Waves (RSEMW)*, 2019, pp. 172-175. Available at: <https://ieeexplore.ieee.org/document/8792770>. DOI: 10.1109/RSEMW.2019.8792770.
19. *Labyntsev A.V., Poveshenko L.I., Kharlanov D.V.* Microstrip Filter with Extended Barrier, 2019 *Radiation and Scattering of Electromagnetic Waves (RSEMW)*, 2019, pp. 176-179. Available at: <https://ieeexplore.ieee.org/document/8792722>. DOI: 10.1109/RSEMW.2019.8792722.
20. Bankov S.E., Kurushin A.A. Raschet antenn i SVCh struktur s pomoshch'yu HFSS Ansoft [Calculation of antennas and microwave structures using HFSS Ansoft]. Moscow: ZAO «NPP „Rodnik“», 2009, 256 p.

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

Харланов Дмитрий Валентинович – Южный федеральный университет; e-mail: mitya999@mail.ru; г. Краснодар, Россия; тел.: 89620130677; кафедра антенн и радиопередающих устройств; соискатель.

Лабынцев Алексей Викторович – e-mail: labyntsev@yandex.ru; г. Таганрог, Россия; тел.: 89085142003; кафедра теоретических основ радиотехники; к.т.н.; с.н.с.; доцент.

Kharlanov Dmitry Valentinovich – Southern Federal University; e-mail: mitya999@mail.ru; Krasnodar, Russia; phone: +79620130677; the department of antennas and radio transmitting devices; the applicant.

Labyntsev Alexey Viktorovich – e-mail: labyntsev@yandex.ru; Taganrog, Russia; phone: 89085142003; department of theoretical foundations of radio engineering; cand. of eng. sc.; senior researcher; associate professor.

УДК 621.375.9

DOI 10.18522/2311-3103-2024-2-202-220

**А.В. Бугакова, О.В. Дворников, Н.Н. Прокопенко, В.А. Чеховский,
Д.В. Клейменкин**

СХЕМОТЕХНИЧЕСКИЕ ОСОБЕННОСТИ ПРОЕКТИРОВАНИЯ ВЫСОКОТЕМПЕРАТУРНЫХ АНАЛОГОВЫХ МИКРОСХЕМ НА GaN И GaAs ТРАНЗИСТОРАХ*

Высокотемпературные интегральные микросхемы (ИМС), сохраняющие работоспособность при температуре более 150°C, требуются во многих областях промышленности, например, в аэрокосмическом, авиационном и автомобильном приборостроении, нефтехимической промышленности, электроэнергетике, электронике военного назначения. В настоящее время зарубежные предприятия серийно выпускают несколько высокотемпературных аналоговых и аналого-цифровых ИМС на основе кремниевых КМОП КНИ структур – ADS1278-НТ, ADS1282-НТ, ADS8320-НТ, INA129-НТ, INA333-НТ, ОРА2333-НТ и др. В Российской Федерации также разработаны высокотемпературные кремниевые операционные усилители и АЦП. Однако максимальная рабочая температура таких изделий не превышает 200°C из-за наличия ограничений кремниевых технологий. По указанной причине в качестве полупроводников, предназначенных для высокотемпературных ИМС, чаще всего рассматриваются широкозонные, такие как карбид кремния (SiC), нитрид (GaN) и арсенид галлия (GaAs), которые обеспечивают ряд характеристик, необходимых для высокотемпературных применений: широкую запрещенную зону, высокую скорость насыщения носителей заряда и низкую концентрацию собственных носителей заряда. В статье

* Исследование выполнено за счет гранта Российского научного фонда № 23-79-10069, <https://rscf.ru/project/23-79-10069/>.

представлен аналитический обзор проблем разработки высокотемпературных GaN и GaAs микросхем. Рассмотрены особенности вольтамперных характеристик GaN и GaAs полевых транзисторов, работающих в режиме обеднения и обогащения, электрические схемы типовых аналоговых устройств (зарядочувствительных и операционных усилителей, компараторов, повторителей тока) и цифровых вентилях. Сделан вывод о том, что схемотехнический синтез GaAs аналоговых микросхем целесообразно выполнять на полевых транзисторах с каналом n-типа, работающих в режиме обеднения, и p-n-p гетероструктурных биполярных транзисторах. Приведены примеры таких схем. Актуальность вышеуказанных исследований связана с проблемами импортозамещения микросхем на широкозонных полупроводниках (GaN, GaAs), обеспечивающих широкий диапазон рабочих температур (свыше +150°C).

GaAs, GaN; транзистор с высокой подвижностью электронов; полевой транзистор, работающий в режиме обеднения; высокотемпературные аналоговые микросхемы.

A.V. Bugakova, O.V. Dvornikov, N.N. Prokopenko, V.A. Tchekhovski,
D.V. Kleimenkin

CIRCUIT FEATURES OF HIGH-TEMPERATURE ANALOG MICROCIRCUITS ON GaN AND GaAs TRANSISTORS

High-temperature integrated circuits, which remain operational at temperatures above 150°C, are required in many areas of industry: aerospace, aviation and automotive instrumentation, the petrochemical industry, electric power, and military electronics. Currently, foreign enterprises are mass-producing several high-temperature analog and analog-to-digital microcircuits based on silicon CMOS SOI structures – ADS1278-HT, ADS1282-HT, ADS8320-HT, INA129-HT, INA333-HT, OPA2333-HT, etc. High-temperature silicon operational amplifiers and ADCs have also been developed in the Russian Federation. However, the maximum operating temperature of such products does not exceed 200°C due to the limitations of silicon technologies. For this reason, wide-bandgap semiconductors such as silicon carbide (SiC), gallium nitride (GaN) and gallium arsenide (GaAs) are most often considered as semiconductors intended for high-temperature microcircuits, which provide a number of characteristics necessary for high-temperature applications: wide bandgap, high carrier saturation velocity and low concentration of intrinsic charge carriers. An overview of the problems of developing high-temperature analog microcircuits based on GaN and GaAs transistors is presented. The features of the current-voltage characteristics of GaN and GaAs field-effect transistors operating in depletion and enhancement modes, electrical circuits of typical analog devices (charge-sensitive and operational amplifiers, comparators, current followers) and logical gates are considered. It is concluded that it is advisable to carry out the circuit synthesis of GaAs analog microcircuits using field-effect transistors with an n-type channel operating in depletion mode and p-n-p heterostructure bipolar transistors. Examples of such schemes are given. The relevance of the above research is related to the problems of import substitution of microcircuits based on wide-bandgap semiconductors (GaN, GaAs), providing a wide range of operating temperatures (over +150°C).

GaAs; GaN; HEMT; depletion mode type FET; enhancement mode type FET; high-temperature analog microcircuits.

Введение. Во многих областях промышленности, таких как, аэрокосмическое, авиационное и автомобильное приборостроение, нефтехимическая промышленность, электроэнергетика, электроника военного назначения требуются высокотемпературные интегральные микросхемы (ИМС), сохраняющие работоспособность при температуре 150°C и более [1–20].

На сегодняшний день зарубежные предприятия серийно выпускают несколько высокотемпературных аналоговых и аналого-цифровых ИМС на основе кремниевых КМОП КНИ структур (ADS1278-HT, ADS1282-HT, ADS8320-HT, INA129-HT, INA333-HT, OPA2333-HT и др.). В Российской Федерации также разработаны высокотемпературные кремниевые операционные усилители и АЦП [1, 2]. Однако максимальная рабочая температура таких изделий не превышает 200°C из-за наличия ограничений кремниевых технологий. По указанной причине в качестве полупроводников, предназначенных для высокотемпературных ИМС, чаще всего рассматриваются широкозонные, такие как карбид кремния (SiC) [11], нитрид (GaN) [3–5, 11, 13, 15, 17, 20] и арсенид галлия (GaAs) [6–10, 12, 15, 18, 19], которые обеспечивают ряд характеристик, необходимых для высокотемпературных применений: широкую запрещенную зону, высокую скорость насыщения носителей заряда и низкую концентрацию собственных носителей заряда.

Подробный анализ полупроводниковых приборов для жестких условий эксплуатации позволил установить, что GaAs микросхемы [6–9, 10, 12, 15, 18, 19] рекомендуется применять до 300°C, на основе SiC целесообразно изготавливать мощные (высоковольтные) транзисторы, а на GaN [3–5, 11, 13, 15, 17, 20] – высокочастотные (быстродействующие) транзисторы и высокотемпературные ИМС с рабочим диапазоном температур до 500°C. Последнее обусловлено тем, что одним из главных преимуществ GaN по сравнению с SiC является меньшее температурное изменение параметров транзисторов и возможность формирования на одной подложке полевых транзисторов (field-effect transistor, FET), работающих в режиме обеднения (depletion mode type, D-FET) и обогащения (enhancement mode type, E-FET), что особенно важно при схемотехническом синтезе аналоговых и цифровых микросхем.

С нашей точки зрения, приоритет в разработке GaAs ИМС для диапазона температур до 300°C обусловлен тем, что арсенид-галлиевые технологические маршруты доступны, хорошо отработаны и обеспечивают удовлетворительные параметры транзисторов при относительно невысокой стоимости изготовления полупроводниковых пластин.

Целью настоящей статьи является рассмотрение особенностей схемотехнического проектирования высокотемпературных GaN и GaAs ИМС. Проблемы сборки кристаллов в корпус и создания технологических маршрутов изготовления GaN и GaAs ИМС будут рассмотрены отдельно.

1. Доступные для схемотехнического синтеза GaN и GaAs элементы. Типовые технологические маршруты изготовления GaN и GaAs ИМС обычно обеспечивают формирование на одной подложке D-FET с каналом n-типа и тонкопленочных резисторов с малым поверхностным сопротивлением от 50 до 500 Ом/квadrat [4, 5, 15], причем в качестве D-FET выступают транзисторы с переходом Шоттки (metal–semiconductor field-effect transistor, MESFET) разных конструкций.

Ограниченная номенклатура доступных элементов существенно усложняет схемотехнический синтез и эксплуатацию создаваемых микросхем, т.к. в ряде случаев приводит к тому, что аналоговые ИМС содержат только D-FET с различным отношением ширины затвора к длине W/L , а высокоомные резисторы являются внешними и располагаются на печатной плате. Кроме того, схемы, содержащие только D-FET требуют применения не симметричного биполярного напряжения питания.

Для упрощения схемотехнического синтеза GaAs аналоговых микросхем предприняты многочисленные попытки расширения номенклатуры активных элементов и создания технологических маршрутов для изготовления на одной подложке:

- ◆ транзисторов с высокой подвижностью электронов (high electron mobility transistor, HEMT) и n-p-n гетероструктурных биполярных транзисторов (heterojunction bipolar transistor, HBT) [6];
- ◆ MESFET с каналом n- и p- типа и n-p-n HBT [7];
- ◆ p-n-p и n-p-n HBT [8];
- ◆ D-FET с каналом n-типа (double-channel pseudomorphic high electron mobility transistor, DpHEMT) и p-n-p HBT [9].

Несмотря на то, что большинство исследований направлено на разработку технологических маршрутов изготовления на одной пластине MESFET с каналом n-типа и n-p-n HBT, с нашей точки зрения, для проектирования аналоговых GaAs ИМС следует применять комплементарную пару в виде D-FET с каналом n-типа и большой граничной частотой, например, DpHEMT, и p-n-p HBT, что позволит значительно улучшить характеристики аналоговых ИМС, а именно [10]:

- ◆ увеличить усиление напряжения входных каскадов и уменьшить уровень шумов, отнесенных к входу, за счет большой крутизны g_m при относительно малом токе стока I_D ;
- ◆ реализовать простые каскады сдвига уровня постоянного напряжения на p-n-p HBT;
- ◆ увеличить диапазон рабочего напряжения усилительных каскадов и сделать возможным выполнение двухтактных выходных каскадов при одновременном использовании DpHEMT и p-n-p HBT;

♦ обеспечить проектирование малошумящих зарядочувствительных усилителей (charge-sensitive amplifiers, CSA) благодаря применению головного DpHEMT с высоким отношением g_m к входной емкости.

Упрощение схемотехнического синтеза GaN ИМС осуществляется, в основном, за счет одновременного использования D-FET и E-FET [11].

Широкозонные D-FET обычно характеризуются большими рабочими токами стока, неиспользуемыми в аналоговых ИМС, поэтому в качестве основного активного элемента рекомендуется применять D-FET с минимальной длиной и шириной затвора, обеспечивающей получение воспроизводимых вольтамперных характеристик (ВАХ), а в качестве малошумящего D-FET – транзистор с отношением W/L не более чем в 10 раз превышающее W/L основного активного элемента.

Для увеличения крутизны допустимо осуществлять параллельное соединение D-FET, а для увеличения выходного малосигнального сопротивления и уменьшения тока стока при том же напряжении затвор-исток V_{GS} – последовательное соединение. Последние иллюстрируют зависимости нормированного тока стока (отношение тока стока к его величине при напряжении сток-исток $V_{DS} = 1,5$ В) на рис. 1.

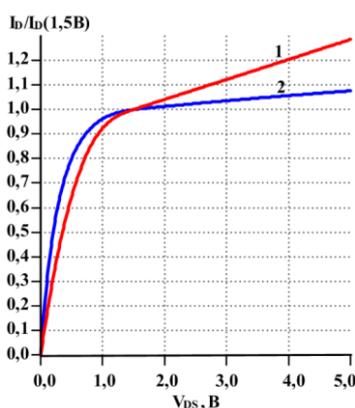


Рис. 1. Нормированная выходная ВАХ в схеме с общим истоком GaAs DpHEMT ($W/L=10$ мкм/0,2 мкм):
1 – одиночный транзистор при $I_D(1,5 В) = 197,45$ мкА, 2 – два последовательно соединенных транзистора при $I_D(1,5 В) = 48,58$ мкА [12]

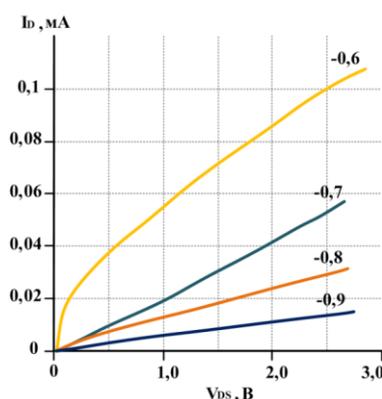


Рис. 2. Выходная ВАХ в схеме с общим истоком GaAs DpHEMT ($W/L=10$ мкм/0,2 мкм) около напряжения отсечки [12]

Известно, что некоторые конструкции широкозонных D-FET имеют две области ВАХ в схеме с общим истоком с резко различающимся выходным малосигнальным сопротивлением только при небольшом обратном напряжении затвор-исток, а вблизи отсечки V_{TH} выходная ВАХ подобна управляемому резистору и не рекомендуется для применения в усилительных каскадах [12]. Эти особенности ВАХ иллюстрирует рис. 2 для GaAs DpHEMT.

Для таких транзисторов рекомендуется выбор тока стока в диапазоне от максимальной величины (I_{DMAX} при $V_{GS}=0$, $V_{DS}=V_{TH}$) до 10-ти кратного уменьшения тока стока по сравнению с I_{DMAX} [12].

2. Схемы на основе D-FET с каналом n-типа. В качестве примера технологического маршрута изготовления D-FET с каналом n-типа рассмотрим GaN500 фирмы Canadian Photonics Fabrication Center [4]. Этот техмаршрут обеспечивает формирование на 3-х дюймовой SiC подложке толщиной 75 мкм НЕМТ с длиной затвора 0,5 мкм, двух уровней металлизированных межсоединений, нихромовых резисторов с поверхностным со-

противлением 50 Ом/квadrat, конденсаторов со структурой металл-диэлектрик-металл и удельной емкостью 0,19 фФ/мкм². ВАХ изготовленных НЕМТ при разных температурах, их совпадение с результатами моделирования иллюстрирует рис. 3, а на рис. 4, 5 – приведены температурные зависимости сопротивления нихрового резистора и емкости нескольких конденсаторов.

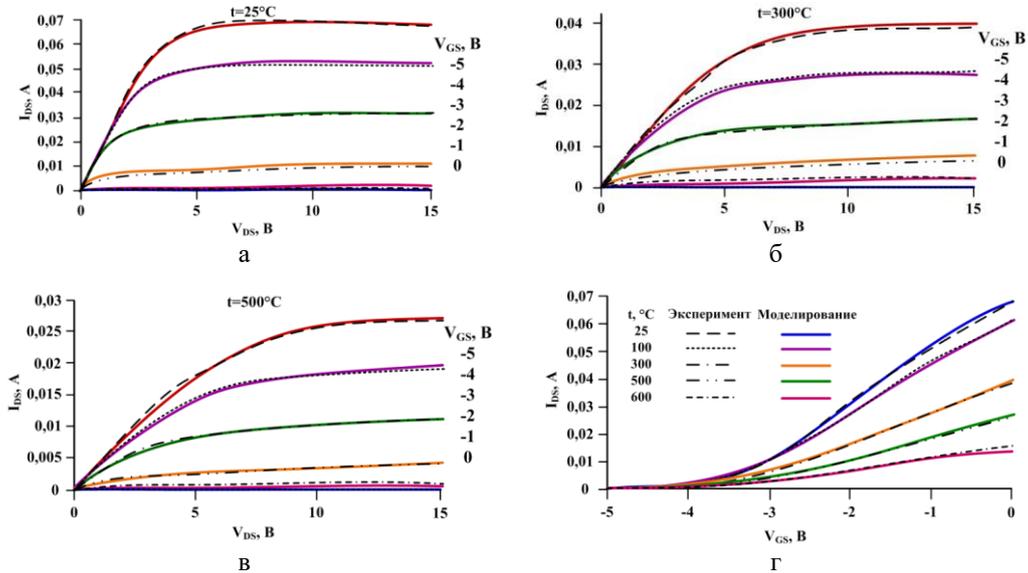


Рис. 3. Результаты измерений и моделирования ВАХ GaN500 НЕМТ [3]

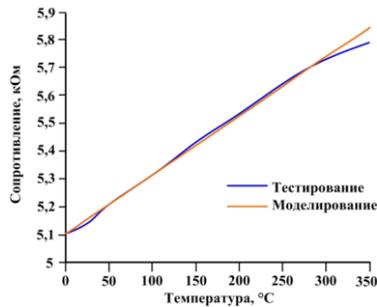


Рис. 4. Результаты измерений и моделирования сопротивления GaN500 резистора $R=5150 \text{ Ом}$, $W/L=5 \text{ мкм}/650 \text{ мкм}$ [4]

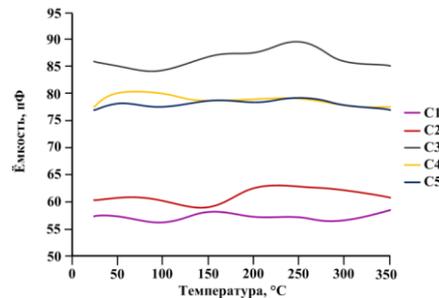


Рис. 5. Результаты измерений емкости GaN500 конденсаторов в диапазоне температур [4]

Как видно из ВАХ на рис. 3, в нормальных условиях $V_{TH} \approx -5 \text{ В}$, поэтому техмаршрут GaN500 позволяет реализовать логические вентили (рис. 6) с отрицательным входным уровнем при применении биполярного напряжения питания довольно большой величины $V_{DD}=14 \text{ В}$, $V_{SS}=-14 \text{ В}$, причем резисторы занимают большую часть площади кристалла.

GaN500 техмаршрут позволяет также проектировать типовые аналоговые схемы (рис. 7), в которых для минимизации тока потребления применяется большое количество последовательно соединенных транзисторов (Т6-Т9 на рис. 7,а).

Уменьшить площадь кристалла, занимаемую высокоомными резисторами можно за счет применения диодного соединения D-FET, что одновременно увеличивает коэффициент усиления напряжения (рис. 8), или применяя внешние резисторы для микросхем

малой степени интеграции. В последнем случае за основу при схемотехническом синтезе можно принять микросхемы для жестких условий эксплуатации на основе n-канальных полевых транзисторов, управляемых р-n-переходом (JFET).

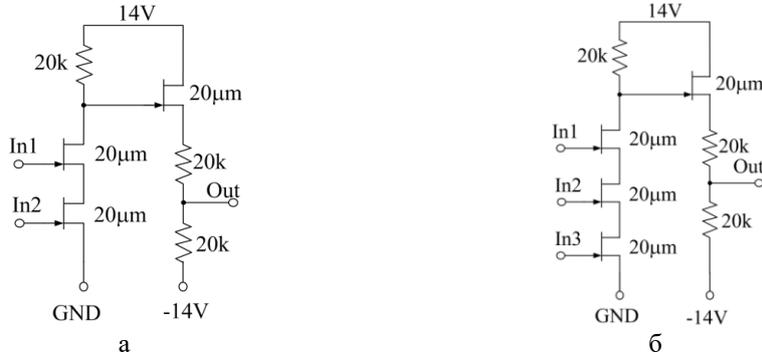


Рис. 6. Реализация цифровых вентилей по техмаршруту GaN500 [5]:
а – электрическая схема NAND2, б – электрическая схема NAND3

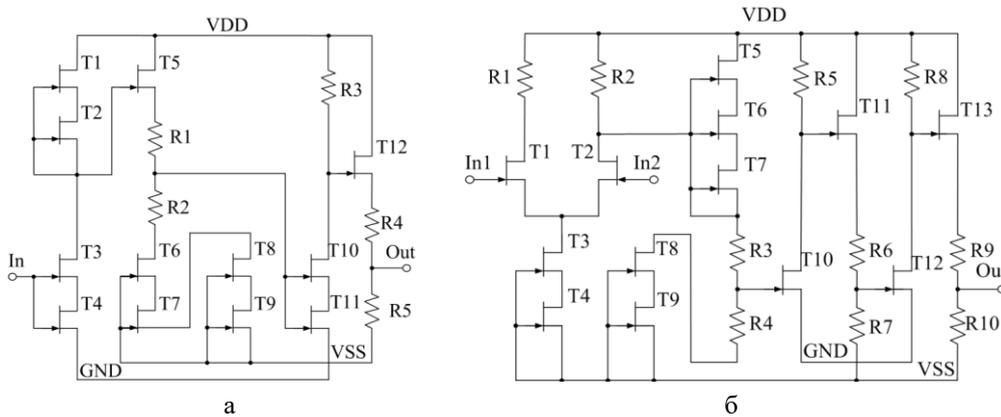


Рис. 7. Электрические схемы аналоговых устройств по техмаршруту GaN500 [4]:
а – усилитель напряжения; б – компаратор

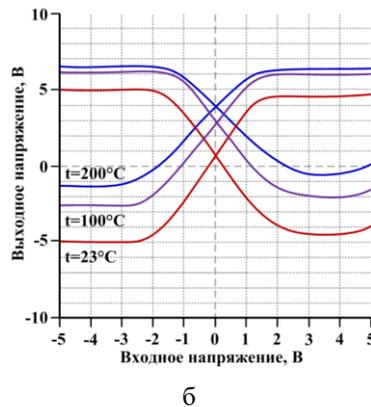
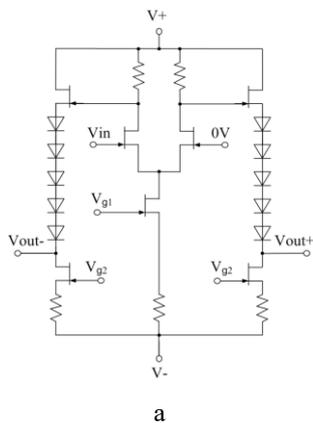


Рис. 8. Высокотемпературный GaN операционный усилитель с ионно-имплантированными резисторами [13]: а – электрическая схема; б – передаточная характеристика при разных температурах

Например, на рис. 9–12 приведены схемы радиационно-стойких CSA на кремниевых n-JFET [14], которые могут быть адаптированы к GaN и GaAs D-FET.

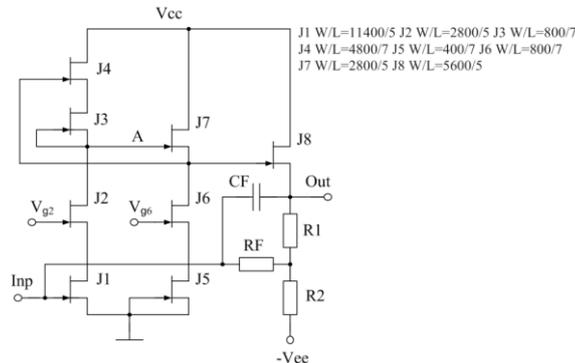


Рис. 9. Типовая схема включения CSA типа IPA3 с входной емкостью 60 пФ

Схема CSA типа IPA3 предназначена для обработки сигналов датчиков с емкостью от 100 до 1000 пФ. Отдельный вывод стока J_8 применяется для суммирования выходных токовых сигналов нескольких CSA. При соединении стока J_8 с шиной положительного питания в усилителе реализуется буферный каскад со схемой сдвига постоянного уровня напряжения (рис. 9). Несмотря на простоту схемотехнического решения, коэффициент усиления напряжения IPA3 в режиме холостого хода превышает 85 дБ.

Усилитель IPA4 (рис. 10) предназначен для обработки сигналов датчиков с емкостью от 10 до 100 пФ и отличается от IPA3, прежде всего, меньшими размерами «головного» транзистора J_1 ($W/L = 1820/5$ мкм) и расширенными возможностями подключения транзистора J_3 .

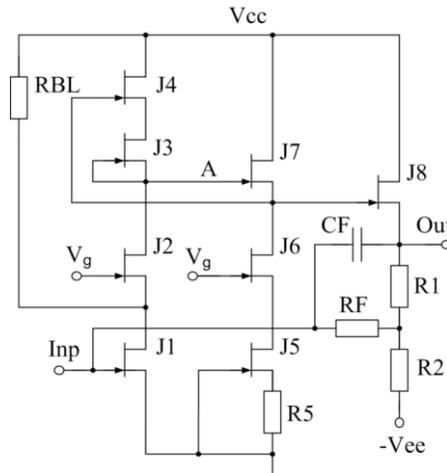


Рис. 10. Схема включения криогенного CSA типа IPA4 с входной емкостью 10 пФ

Источник тока на диодном включении транзистора J_3 обеспечивает ток через «головной» транзистор на уровне 600 мкА для уменьшения рассеиваемой мощности усилителя типа «А». При необходимости увеличения тока стока J_1 рекомендуется включение внешнего резистора R_{BL} . Ток стока J_3 определяет скорость перезарядки суммарной емкости в высокоимпедансном узле «А» и, следовательно, скорость нарастания выходного напряжения, для увеличения которой рекомендуется применение CSA типа «В» (рис. 11) с током перезарядки, определяемом резистором R_3 . Таким образом, в CSA типа «В» можно независимо установить скорость нарастания выходного напряжения (выбором сопротивления R_3) и уровень входного шума (сопротивлением R_{BL}).

Заметим, что в усилителе IPA4 возможно соединение стока J_2 с чисто резистивной нагрузкой при реализации трансрезистивного усилителя. Для уменьшения уровня шума разработано включение усилителя IPA4 (рис. 12) со схемой компенсации входного тока за счет прямого смещения p-n-перехода затвор-исток «головного» JFET. Транзистор J_1 работает при токе, превышающем максимальный ток стока, что возможно только в случае прямого смещения истокового перехода. Как показали экспериментальные исследования, компенсация входного тока позволяет исключить резистор R_F , что обеспечивает значительно меньший уровень шумов.

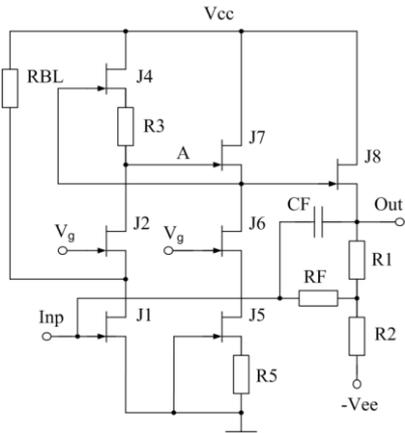


Рис. 11. Схема включения CSA типа «В» на основе рис. 10

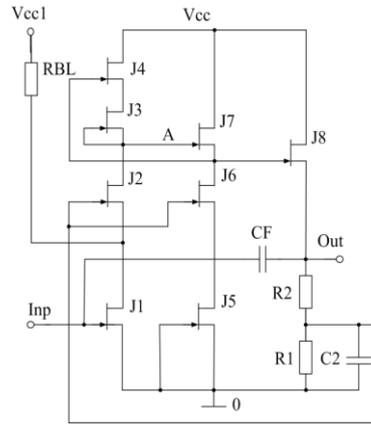


Рис. 12. Включение CSA типа IPA4 на основе рис. 10 без резистора R_F

В качестве примера адаптации схем CSA с кремниевым n-JFET к широкозонным полупроводниковым материалам является GaAs CSA, показанный на рис. 13 [15].

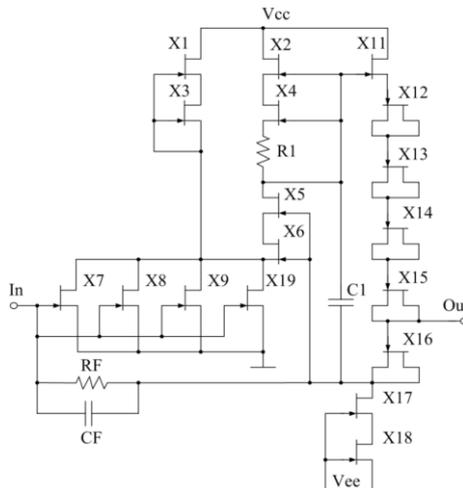


Рис. 13. CSA основе GaAs DpHEMT [15]

Известно, что при проектировании CSA необходимо максимально увеличивать крутизну головного FET и суммарное сопротивление R_{Σ} всех параллельных цепей в высокоимпедансном узле, а также уменьшать суммарную емкость C_{Σ} всех параллельных цепей в высокоимпедансном узле. Однако при этом надо учитывать, что увеличение крутизны за счет увеличения тока стока относительно слабое $g_m \sim \sqrt{I_D}$, но приводит к увеличению потребляемой мощности и уменьшению R_{Σ} , а увеличение крутизны увеличением W/L вызывает рост входной емкости CSA.

Частично указанные противоречия устранены в схеме CSA, показанной на рис. 13. Здесь все D-FET имеют размер затвора $W/L=10 \text{ мкм}/0,2 \text{ мкм}$, а X1, X3, X7-X10 – $W/L=100 \text{ мкм}/0,2 \text{ мкм}$. Эта схема имеет ряд преимуществ [15]:

- ◆ крутизна «головного» D-FET устанавливается за счет параллельного соединения транзисторов X7-X10 и величины их тока стока, в основном, определяемой цепью X1, X3,
- ◆ введение X5, X6 позволяет, с одной стороны, стабилизировать напряжение на стоке «головного» транзистора и уменьшить динамическую входную емкость CSA, а с другой стороны, при токе стока X3 существенно большем, чем X6, увеличить R_{Σ} и уменьшить C_{Σ} , что приводит к росту усиления и быстродействия усилителя напряжения;
- ◆ однокаскадная схема усилителя напряжения значительно упрощает частотную коррекцию подключением одного конденсатора C1 в высокоимпедансный узел.

При проектировании GaN и GaAs со схемой на рис. 13 рекомендуется применение методики схемотехнического моделирования, обеспечивающей получение минимального тока потребления, малого уровня шумов и высокого быстродействия, с учетом того, что основным варьируемым параметром является отношение ширины затвора W к его длине L , а суммарное сопротивление применяемых тонкопленочных резисторов минимально для уменьшения площади кристалла, занимаемой CSA [15].

Заметим, что необходимые во многих GaN и GaAs аналоговых устройствах фильтры могут быть созданы на основе конденсаторов и так называемых "токовых зеркал" (рис. 14) [16].

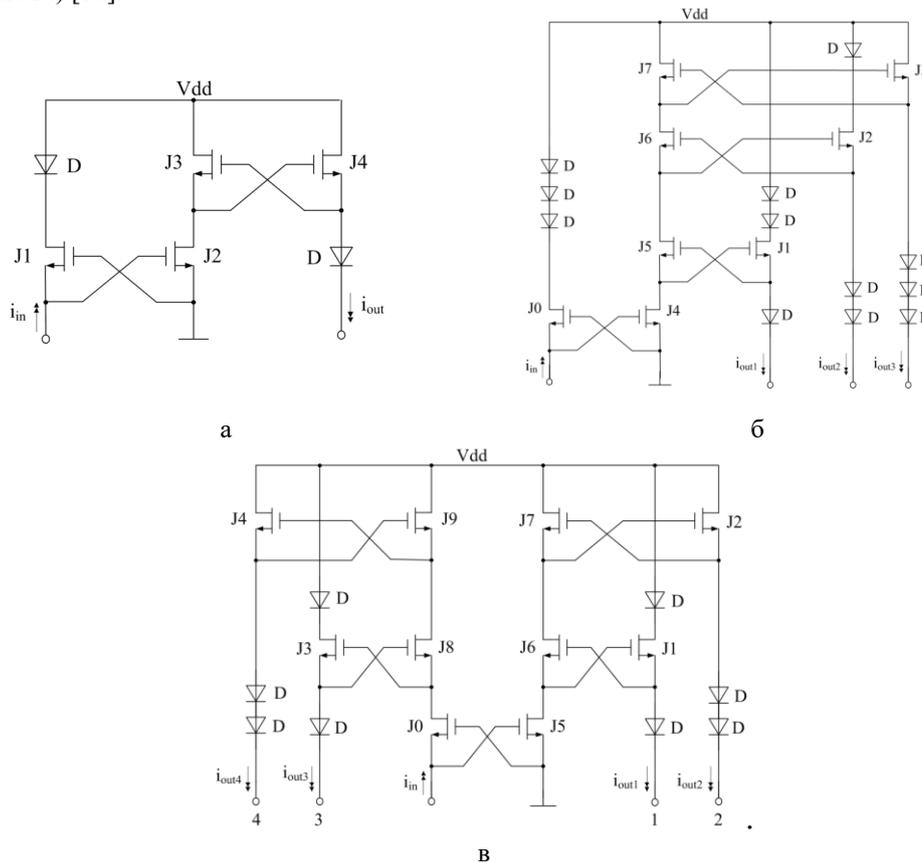


Рис. 14. "Токовые зеркала" основе GaAs MESFET [16]: а – одним выходом; б – с тремя выходами; в – с инвертирующими и не инвертирующими выходами

3. Схемы на основе D-FET и E-FET с каналом n-типа. Уменьшить площадь кристалла, занимаемую тонкопленочными резисторами, возможно при увеличении их поверхностного сопротивления, однако при этом увеличивается технологический разброс сопротивле-

ний. Кардинальным решением указанной проблемы является применение E-FET в качестве активного элемента и D-FET с соединенными выводами затвора и истока как нагрузочного резистора. Преимущества от применения такой пары FET объясняют большое количество работ, посвященных формированию E-FET и D-FET на одной подложке [4–6].

ВАХ для AlGaIn/GaN HEMT с $W/L=100$ мкм/1 мкм [4, 5] показаны на рис. 15 и обобщены в табл. 1.

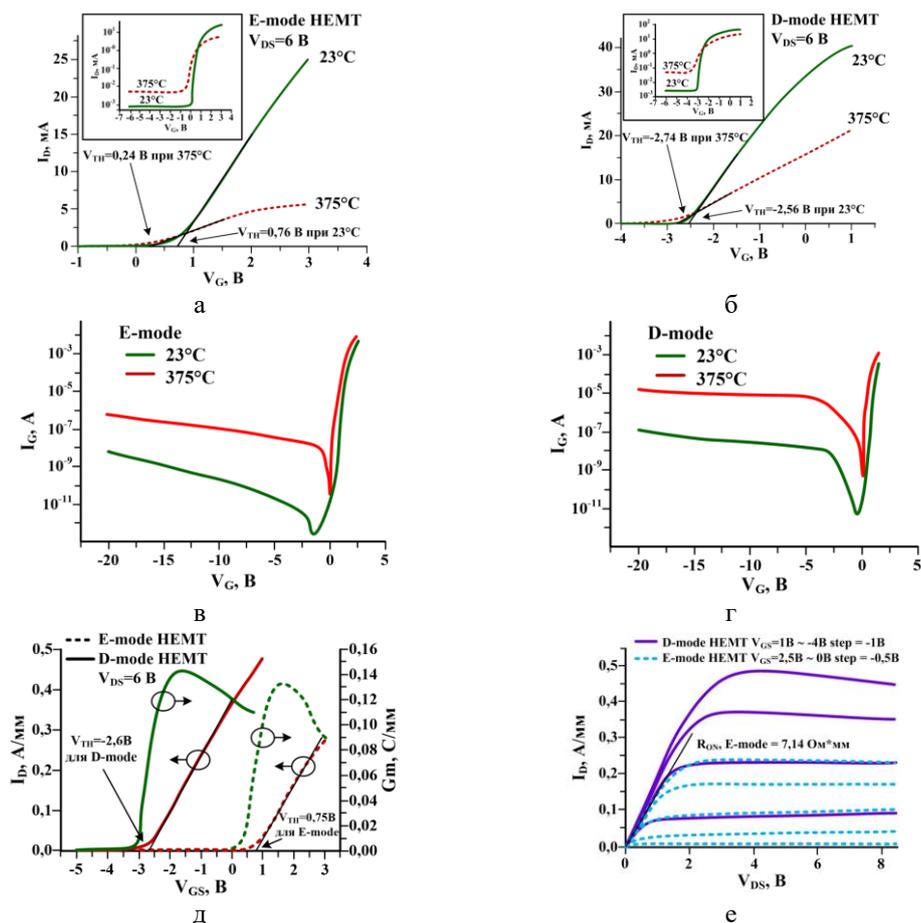


Рис. 15. Вольтамперные характеристики AlGaIn/GaN HEMT с $W/L=100$ мкм/1 мкм при разных температурах [4, 5]

Таблица 1

Параметры AlGaIn/GaN HEMT, работающих в режиме обеднения и обогащения [4, 5], при разных температурах

Наименование параметра	D-FET		E-FET	
	23°C	375°C	23°C	375°C
Напряжение отсечки V_{TH} , В	- 2,56	- 2,74	0,76	0,24
Температурное изменение напряжения отсечки dV_{TH}/dT , мВ/К	- 0,51		-1,5	
Максимальное значение крутизны на единицу ширины затвора g_{MMAX}/W , мС/мм	145	58	120	29
Максимальное значение тока стока на единицу ширины затвора I_{DMAX}/W , мА/мм	480	213	250	56
Граничная частота f_T , ГГц	14,3		10,7	

При проектировании схем на GaN HEMT следует учитывать следующие факторы:

- ♦ как показал ряд экспериментальных исследований, зависимость крутизны g_m от напряжения затвор-исток V_{GS} для GaN HEMT имеет немонотонный характер (рис. 15д), т.е. существует максимальное значение g_{mMAX} при некоторой величине V_{GSGMAX} . До начала схемотехнического проектирования необходимо выяснить приоритет в параметрах, ответить на вопрос - что важнее обеспечить максимальную крутизну при заданной температуре (рис. 15д) или минимальное изменение крутизны в диапазоне температур (рис. 16), и в зависимости от этого выбирать напряжение V_{GS} ;

- ♦ в известных конструкциях E-HEMT и D-HEMT ток стока и крутизна на единицу ширины затвора значительно отличается для разных типов HEMT. Обычно, E-HEMT применяется в качестве усилительного элемента, а D-HEMT – в качестве нагрузки. В этом случае, отношение W/L рекомендуется выбирать в диапазонах $(W/L)_{EMODE}/(W/L)_{DMODE}=7...50$ [4, 5] и уточнять при схемотехническом моделировании в зависимости от требуемых параметров ИМС;

- ♦ GaN HEMT имеют большой обратный ток затвора при высоких температурах (рис. 15в,г), поэтому при применении этих транзисторов во входных каскадах необходимо предусматривать формирование цепей следящей обратной связи, поддерживающих малое падение напряжения на переходах затвор-исток и затвор-сток;

- ♦ для GaN D-HEMT существует напряжение V_{GSGZTC} , соответствующее минимальному температурному коэффициенту крутизны (transconductance zero-temperature coefficient) [17].

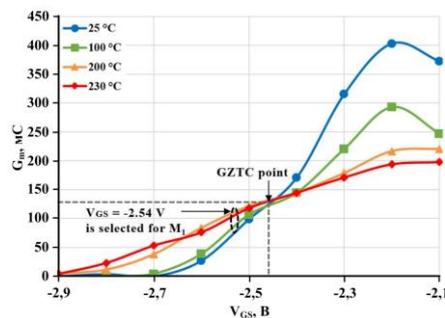


Рис. 16. Зависимость крутизны AlGaIn/GaN HEMT от напряжения затвор-исток при разных температурах [17]

Следует отметить, что при схемотехническом синтезе цифровых вентилях и аналоговых схем на основе широкозонных E-FET и D-FET за основу можно взять существующие схемотехнические решения для n -МОП с индуцированным каналом и резистивными нагрузками.

Разработанные аналоговые схемы на широкозонных E-FET и D-FET показаны на рис. 17–20.

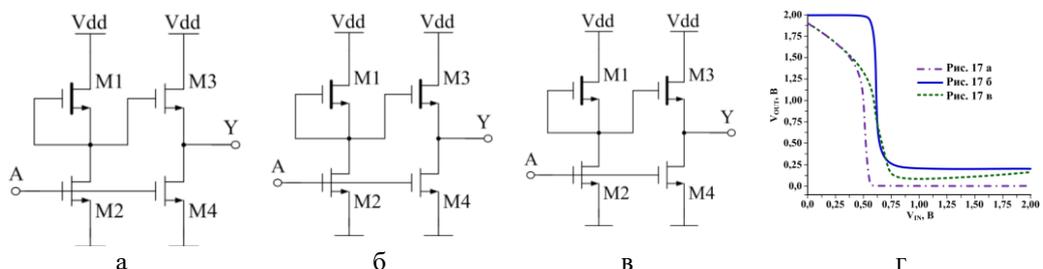


Рис. 17. Базовые схемы инвертирующих каскадов на GaN E-HEMT и D-HEMT (а, б, в) и результаты моделирования (г) их передаточных характеристик при $(W/L)_{EMODE}=50\text{мкм}/1,8\text{мкм}$, $(W/L)_{DMODE}=5\text{мкм}/1\text{мкм}$

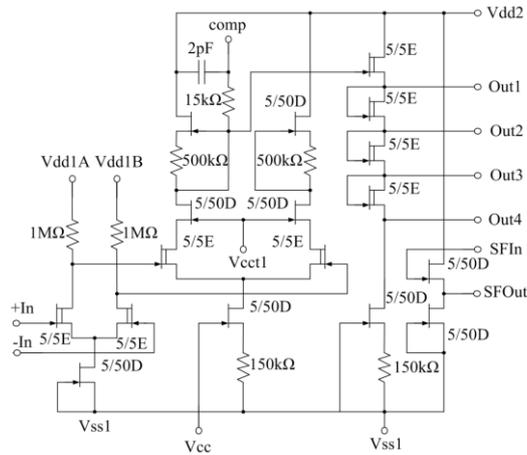


Рис. 18. Электрическая схема GaAs операционного усилителя [18]

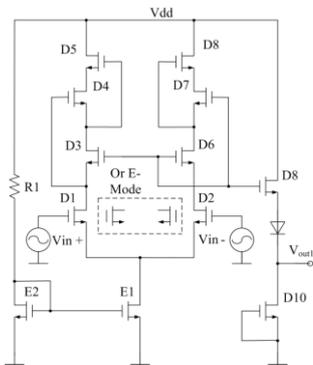


Рис. 19. Электрическая схема GaAs компаратора [19]

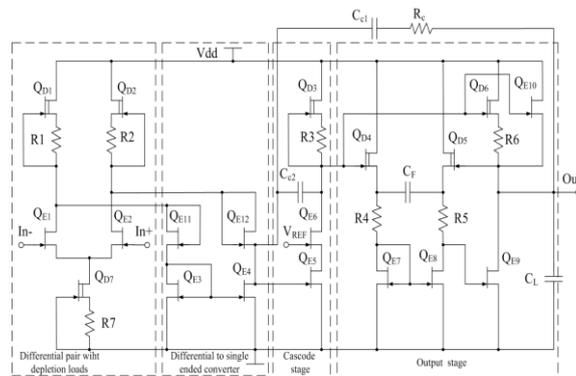


Рис. 20. Электрическая схема GaN операционного усилителя [20]:
QD – D-HEMT, QE – E-HEMT

Рассмотрим подробнее схему GaN операционного усилителя на рис. 20. В этой схеме все источники тока были реализованы на D-HEMT, которые имеют более высокое выходное малосигнальное сопротивление r_{OU} по сравнению с E-HEMT. Однако, т.к. D-HEMT с минимальной шириной и максимальной длиной канала не могли обеспечить необходимый небольшой ток, то в источниках тока на D-HEMT применены резисторы. Все источники тока на транзисторах Q_{D1} , Q_{D2} , Q_{D3} , Q_{D6} и Q_{D7} с встроенными резисторами имеют одинаковые размеры. Выходной ток источников тока около 20 мкА. При этом выходное малосигнальное r_{OUi} сопротивление i -того источников тока описывается выражением

$$r_{OUi} \approx g_{mDi} \cdot R_1 \cdot r_{Di}, \quad (1)$$

где g_{mDi} , r_{Di} – крутизна и сопротивление сток-исток i -того D-HEMT соответственно, R_1 – сопротивление в истоке.

Первый каскад представляет собой дифференциальную пару Q_{E1} , Q_{E2} на транзисторах E-HEMT с генератором стабильного тока Q_{D7} и активной нагрузкой Q_{D1} , Q_{D2} . Так как применяемый техмаршрут не позволяет сформировать диоды на p-n-переходе, то во втором каскаде E-HEMT Q_{E11} , Q_{E12} в диодном включении обеспечивают сдвиг постоянного уровня напряжения, а «токовое зеркало» на Q_{E3} , Q_{E4} позволяет перейти от дифференциального сигнала к синфазному.

Третий каскад на каскадном соединении E-HEMT Q_{E5} , Q_{E6} с нагрузкой в виде источника тока Q_{D3} , $R3$ является основным усилительным каскадом, усиление которого совместно с емкостью конденсатора C_{C2} обеспечивает (благодаря эффекту Миллера) коррекцию амплитудно-частотной характеристики.

Для уменьшения напряжения смещения нуля и увеличения коэффициента ослабления синфазного сигнала рекомендуется выполнение следующего отношения [20]

$$\frac{(W/L)_{E5}}{(W/L)_{E3}} = \frac{I_{D3}}{I_{E3}} \quad (2)$$

Выходной каскад представляет собой повторитель напряжения с малым выходным сопротивлением, в котором транзисторы работают в классе «AB», предпочтительном при работе на емкостную нагрузку. Q_{D4} , Q_{D5} , $R4$, $R5$, Q_{E7} и Q_{E8} образуют усилитель ошибки. Размеры транзисторов и сопротивления резисторов удовлетворяют условиям $(W/L)_{D4}=(W/L)_{D5}$, $R4 = R5$, $(W/L)_{E7}=(W/L)_{E8}$. Усилитель ошибки обеспечивает одинаковое напряжение на затворах Q_{D3} и Q_{D6} , поэтому выполняется равенство $I_{D6} = I_{D3}$. Более того, он гарантирует, что выходное напряжение соответствует напряжению затвора Q_{D4} как по постоянному, так и переменному току благодаря высокому коэффициенту усиления цепи, образованной транзистором Q_{E9} с общим истоком и активной нагрузкой на Q_{D6} . Стоит отметить, что для минимизации систематического напряжения смещения должно выполняться следующее отношение [20]

$$\frac{(W/L)_{E9}}{(W/L)_{E7}} = \frac{I_{D6}}{I_{E7}} \quad (3)$$

Действие отрицательной обратной связи, вызванной усилителем ошибки, приводит к дополнительному эффекту – уменьшению выходного малосигнального сопротивления выходного каскада [20]

$$r_o \approx (g_{MD6} + g_{MD9})^{-1} \quad (4)$$

Заметим, что транзистор Q_{E10} в состоянии покоя выключен, т.к. его напряжение затвор-исток равно нулю. Однако при большой амплитуде сигнала, когда выходной ток должен втекать в емкостную нагрузку Q_{E10} , включается и улучшает нагрузочную способность усилителя. С другой стороны, когда выходной ток должен вытекать, то улучшает нагрузочную способность транзистор Q_{E9} . Основные параметры операционного усилителя рис. 20 приведены в табл. 2.

Таблица 2

Основные параметры GaN операционного усилителя на рис. 20 [20]

Наименование параметра	Величина
Напряжение питания, В	6
Опорное напряжение V_{REF} , В	3
Ток потребления в режиме холостого хода, мА	0,2
Напряжение смещения нуля, мВ	1,0
Коэффициент усиления напряжения, дБ	135
Частота единичного усиления, МГц	0,5
Запас фазы при частоте единичного усиления, градусов	60

4. Схемы на основе D-FET с каналом n-типа и p-n-p HBT. Разработанные в [12] электрические схемы GaAs операционных усилителей показаны на рис. 21–23, а в табл. 3 – результаты моделирования параметров. GaAs_OAmp1 – реализован только на D-FET типа DrHEMT и приведен далее для сравнения с операционными усилителями с комплементарной парой транзисторов.

Таблица 3

Результаты схемотехнического моделирования ОУ [12]

Наименование параметра	Наименование ОУ		
	GaAs_OAmp1	GaAs_OAmp2	GaAs_OAmp3
Напряжение питания, В	-3/7	-5/5	-5/5
Ток потребления в режиме холостого хода, мА	0,570	0,646	1,141
Напряжение смещения нуля, мВ	1,3	-0,19	-0,18
Коэффициент усиления напряжения	$2,9 \cdot 10^3$	10^5	10^5
Максимальное выходное напряжение при сопротивлении нагрузки $R_{LOAD}=100$ кОм, В	-2,50/2,23	-4,41/4,52	-4,25/4,45
Частота единичного усиления, МГц	203 ¹	52 ²	52 ²
Запас фазы при частоте единичного усиления, градусов	54 ¹	54 ²	54 ²

Примечание: ¹C1=1,5 пФ; ²C1=20 пФ

Все DrHEMT на электрических схемах имеют размер затвора $W/L=10$ мкм/0,2 мкм, причем только входные транзисторы X12-X14 и X15-X17 имеют $W/L=100$ мкм/0,2 мкм. Требуемое значение сопротивления получено параллельно-последовательным соединением 5 кОм резисторов. Так резистор R2 на рис. 21 состоит из последовательного соединения одного 5 кОм резистора и 3-х параллельно соединенных по 5 кОм резисторов.

Во всех ОУ применен одинаковый дифференциальный каскад, состоящий из входных транзисторов с большой крутизной X12...X14 и X15...X17, каскодных транзисторов X10, X11, источников тока X18, R5 и X19, R6, нагрузки входных транзисторов X1, X5, R1 и X4, X8, R4, нагрузки каскодных транзисторов X2, X6, R2 и X3, X7, R3.

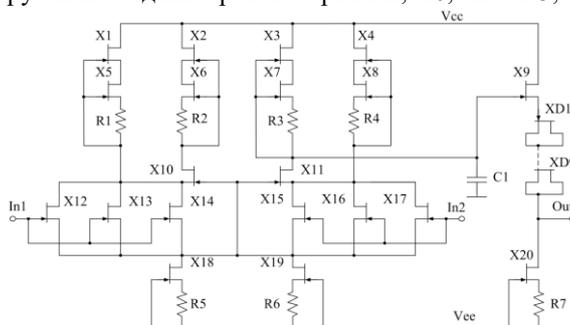


Рис. 21. Электрическая схема GaAs_OAmp1

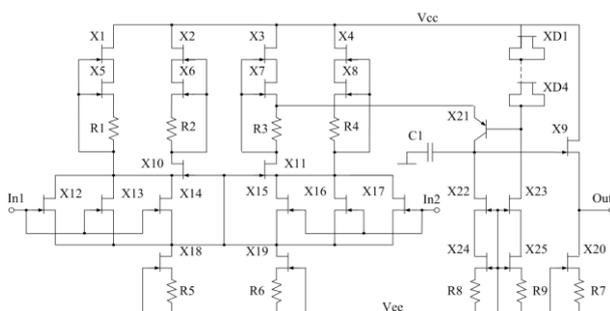


Рис. 22. Электрическая схема GaAs_OAmp2

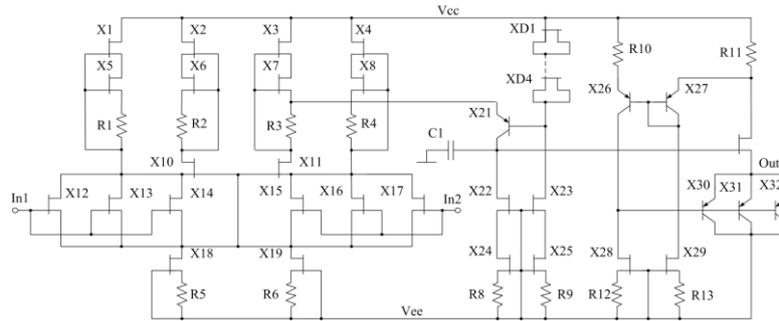


Рис. 23. Электрическая схема GaAs_OAmp3

Такие схемотехнические решения обычно применяются в каскодных FET для обеспечения большого усиления напряжения и использовано в CSA на рис. 13. Так, большая крутизна X12...X14 обеспечивается благодаря большому отношению W/L и значительной величине тока стока. В то же время, преобладающая часть постоянного тока стока транзисторов X12...X14 протекает через их нагрузку X1, X5, R1 в источник положительного напряжения V_{CC} , а меньшая часть – через X10, что позволяет включить в сток X10 большое нагрузочное сопротивление и увеличить усиление по напряжению. Заметим, что параметры элементов выбраны таким образом, что практически вся переменная составляющая тока стока X12...X14 протекает по низкому сопротивлению истока X10, а не высокому сопротивлению нагрузки X1, X5, R1. На транзисторах X9, X20 с цепочкой последовательно соединенных диодов XD1-XD9 реализован выходной истоковый повторитель. Конденсатор C1 применен для коррекции амплитудно-частотной характеристики и получения требуемого запаса по фазе для заданного коэффициента усиления.

Так как в GaAs_OAmp1 в качестве активных элементов применены только DpHEMT, типовое значение напряжения отсечки которых составляет около -1 В, то в этом случае для обеспечения симметричного биполярного выходного напряжения применено несимметричное напряжение питания (минус 3 В и плюс 7 В). Последний недостаток устранен в ОУ GaAs_OAmp2, GaAs_OAmp3, в которых для сдвига постоянного уровня и одновременного увеличения коэффициента усиления напряжения использован p-n-p НВТ X21, включенный по схеме с общей базой. Кроме того, в GaAs_OAmp3 предусмотрен двухтактный выходной каскад для работы с малым сопротивлением нагрузки R_{LOAD} .

Особое внимание следует уделять разработке выходных каскадов (рис. 24).

На рис. 24,а показан простой выходной каскад, примененный во всех схемах ОУ. К сожалению, его использование существенно увеличивает ток потребления ОУ. Для адекватного сравнения схем в модернизированном выходном каскаде (рис. 24,б) применено такое же количество НВТ, как и на рис. 24,а, а сопротивление резистора R5 выбрано таким образом, чтобы диапазон выходного напряжения каскадов при сопротивлении нагрузки $R_{LOAD}=2$ кОм был одинаков.

Как видно из результатов моделирования, показанных в табл. 4, модернизированный выходной каскад (рис. 24,б) имеет при указанных условиях почти в 4,1 раза меньше ток потребления, чем схема на рис. 24,а.

Постоянное выходное напряжение каскадов рис. 24,а,б не равно входному, а дифференциальный коэффициент передачи $K_D=\Delta V_{OUT}/\Delta V_{IN}$ зависит от уровня постоянного входного напряжения V_{IN} (иными словами, от величины и направления тока, протекающего через нагрузочный резистор R_{LOAD}). Такая особенность выходных каскадов не является недостатком при их применении в ОУ, в которых сдвиг выходного напряжения относительно входного в выходном каскаде обычно не вносит существенный вклад в напряжение смещения нуля по сравнению с другими факторами (несимметричной нагрузкой во входном дифференциальном каскаде и влиянием схемы преобразования дифференциального сигнала в синфазный), а зависимость $K_D=f(V_{IN})$ выходного каскада компенсируется действием отрицательной обратной связи, охватывающей ОУ.

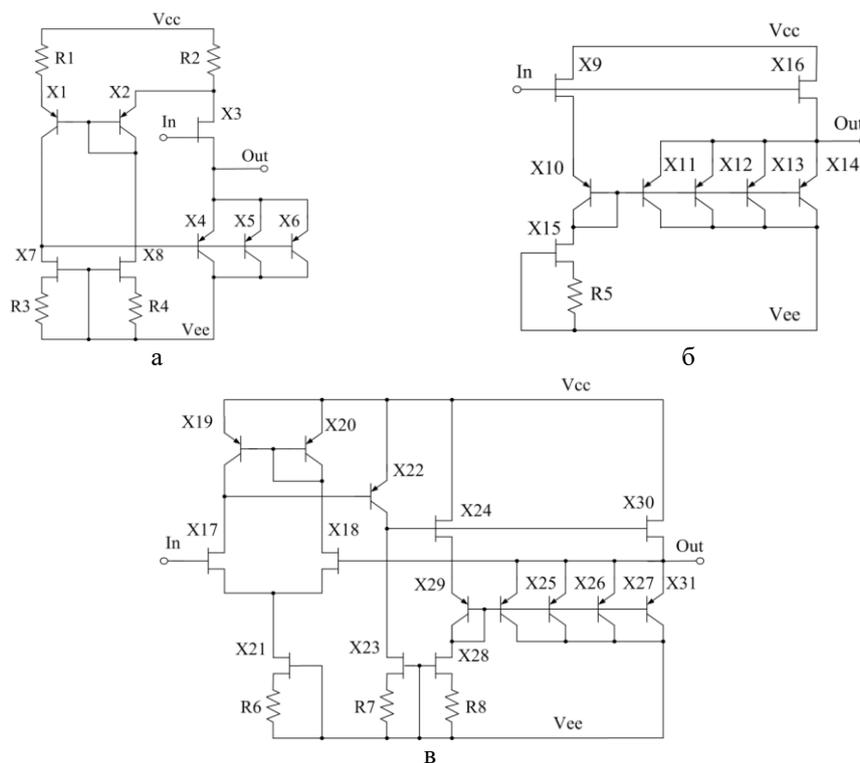


Рис. 24. Двухтактные выходные каскады GaAs аналоговых схем

Таблица 4

Результаты схемотехнического моделирования выходных каскадов [12]

Наименование параметра	Схема изображена на рисунке		
	27а	27б	27в
Напряжение питания, В	-5/5	-5/5	-5/5
Ток потребления в режиме холостого хода, мА	0,681	0,167	0,686
Максимальное выходное напряжение при $R_{LOAD}=2$ кОм, В	-2,85/4,43	-	-
Максимальное выходное напряжение при $R_{LOAD}=2$ кОм, В	-	2,76/4,47	2,76/3,89
Диапазон изменения коэффициента передачи при изменении входного напряжения от -2,7 В до 2,7 В и $R_{LOAD}=2$ кОм	от 0,737 до 0,946	от 0,766 до 0,910	от 0,997 до 0,998
Выходное напряжение при $V_{IN}=0$ и $R_{LOAD}=2$ кОм, мВ	-499,1	-565,2	-6,5

Выходной каскад рис. 24,в рекомендуется для использования в качестве отдельного повторителя напряжения. В нем применен простой ОУ (X17...X23), действие которого обеспечивает равенство потенциалов в узлах In и Ou3 (на входе и выходе повторителя напряжения) без зависимости от величины и направления тока, протекающего через нагрузку R_{LOAD} , подключаемую к узлу Ou3. К недостаткам такого повторителя можно отнести необходимость соединения коллектора X22 с конденсатором, обеспечивающим отсутствие самовозбуждения и увеличивающим площадь кристалла, занимаемую выходным каскадом.

Заключение. В статье выполнен сравнительный анализ базовых схемотехнических решений высокотемпературных аналоговых микросхем на GaN и GaAs транзисторах и их основных параметров. Это позволило установить, что в диапазоне рабочих температур до

300°C целесообразно проектирование аналоговых микросхем на GaAs D-FET с каналом n-типа и p-n-p HBT, а в диапазоне до 500°C – на GaN D-FET и E-FET с каналом n-типа. При выборе рабочей точки рекомендуется учитывать то, что максимальная крутизна и минимальное температурное изменение крутизны может достигаться при разном напряжении затвор-исток, а для некоторых конструкций D-FET следует применять область напряжений затвор-исток с не более, чем десятикратным уменьшением тока стока по сравнению с максимальным значением.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Коротков А.С., Морозов Д.В., Пилипко М.М., Енученко М.С.* 12-разрядный дельта-сигма АЦП для мониторинга состояния высокотемпературных объектов // Проблемы разработки перспективных микро- и нанoeлектронных систем (МЭС). – 2020. – Вып. 1. – С. 95-99.
2. *Балашов Е.В., Иванов Н.В., Ахметов Д.Б., Коротков А.С.* Высокотемпературный инструментальный усилитель // Наноиндустрия. – 2020. – № S96-1. – С. 160-163.
3. *Hassan A., Amer M., Savaria Y., Sawan M.* Towards GaN500-based High Temperature ICs: Characterization and Modeling up to 600°C // 2020 18th International New Circuits and Systems Conference (NEWCAS). IEEE. – 2020. – P. 275-278.
4. *Hassan A., Ali M., Trigui A., Savaria Y., Sawan M.* A GaN-Based Wireless Monitoring System for High-Temperature Applications // Sensors. – 2019. – Vol. 19 (8). – P. 1785.
5. *Hassan A., Noël J-P., Savaria Y., Sawan M.* Circuit Techniques in GaN Technology for High-Temperature Environments // Electronics. – 2022. – Vol. 11 (1). – P. 42.
6. *Fresina M.* Trends in GaAs HBTs for wireless and RF // Bipolar/BiCMOS Circuits and Technology Meeting J. IEEE. – 2011. – P. 150-153.
7. *Zampardi P.J., Sun M., Cismaru C., Li J.* Prospects for a BiCFET III-V HBT Process // Compound Semiconductor Integrated Circuit Symposium (CSICS). – 2012. – P. 1-3.
8. *Jena M.R., Panda A.K., Dash G.N.* A Comparative Analysis InP/InGaAs δ Doped based NPN and PNP HBT // International Journal of Engineering Research & Technology (IJERT). – 2019. – Vol. 8, Issue 09. – P. 819-823.
9. *Ловшенко И.Ю., Кратович П.С., Стемпцкиий В.Р., Дворников О.В., Кунц А.В., Павлючик А.А.* Гетеропереходный биполярный транзистор со структурой рnp-типа в арсенид-галлиевой технологии HBT-HEMT // Проблемы разработки перспективных микро- и нанoeлектронных систем – 2022: Сб. трудов / под общ. ред. академика РАН А.Л. Стемпковского. – М.: ИППМ РАН, 2022. – Вып. IV. – С. 149-154.
10. *Дворников О.В., Павлючик А.А., Прокопенко Н.Н., Чеховский В.А., Кунц А.В., Чумаков В.Е.* Арсенид-галлиевый аналоговый базовый кристалл // Проблемы разработки перспективных микро- и нанoeлектронных систем – 2021: Сб. трудов / под общ. ред. академика РАН А.Л. Стемпковского. – М.: ИППМ РАН, 2021. – С. 47-54.
11. *Cai Y., Cheng Z., Yang Z., Tang C., Lau K., Chen K.* High-Temperature Operation of AlGaIn/GaN HEMTs Direct-Coupled FET Logic (DCFL) Integrated Circuits // Electron Device Letters J. IEEE. – 2007. – Vol. 28, No. 5. – P. 328-331.
12. *Дворников О.В., Павлючик А.А., Прокопенко Н.Н., Чеховский В.А., Кунц А.В., Чумаков В.Е.* Унифицированные схемотехнические решения аналоговых арсенид-галлиевых микросхем // Известия вузов. Электроника. – 2022. – Т. 27, № 4. – С. 475-488.
13. *Nototo K., Hasegawa K., Satoh M., Nakamura T.* Reliability of High-Temperature Operation for GaN-Based OPAMP // MRS Online Proceedings Library (OPL). – 2009. – P. 1195-B08-02.
14. *Дворников О.В., Чеховский В.А., Дятлов В.Л., Прокопенко Н.Н.* Создание низкотемпературных аналоговых ИС для обработки импульсных сигналов датчиков. Ч. 3 // Современная электроника. – 2015. – № 6. – С. 34-39.
15. *Dvornikov O.V., Prokopenko N.N., Chumakov V.E., Tchekhovskiy V.A.* General-Purpose Simulation Technique for Charge-Sensitive Amplifiers Based on Wide-Band-Gap (GaAs, GaN) D-FETs // Proceedings of 16th International Conference of Actual Problems of Electronic Instrument Engineering (APEIE). IEEE. – 2023.
16. *Takai N., Fujii N.* GaAs MESFET multi-output current mirrors and their application to high frequency filters // International Conference on Electronics, Circuits and Systems. Surfing the Waves of Science and Technology (Cat. No. 98EX196). IEEE. – 1998. – Vol. 2. – P. 307-310.
17. *Ehteshamuddin M., Salem J.M., Ha D.S.* A high temperature variable gain amplifier based on GaN HEMT devices for downhole communications // 2017 IEEE International Symposium on Circuits and Systems (ISCAS). – 2017. – P. 1-4.

18. Fujiwara M., Nagata H., Hibi Y., Matsuo H., Sasaki M. Cryogenic low noise amplifier with GaAs JFETs // AIP Conference Proceedings J. American Institute of Physics. – 2009. – Vol. 1185, No. 1. – P. 267-270.
19. Lan D. Ning Y., Wang J., Jiang H. High performance two-stage bootstrapped GaAs comparator with gain enhancement // 16th Annual Wireless and Microwave Technology Conference (WAMICON). IEEE. – 2015. – P. 1-4.
20. Pennisi S., Pulvirenti F., Samperi K. Frequency Compensation Scheme for a Full GaN OpAmp driving 1-nF load // 2022 International Symposium on Circuits and Systems (ISCAS). IEEE. – 2022. – P. 2042-2046.

REFERENCES

1. Korotkov A.S., Morozov D.V., Pilipko M.M., Enuchenko M.S. 12-razryadnyy del'ta-sigma ATSP dlya monitoringa sostoyaniya vysokotemperaturnykh ob'ektov [12-bit delta-sigma ADC for monitoring the state of high-temperature objects], *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem (MES)* [Problems of development of advanced micro- and nanoelectronic systems stem (MES)], 2020, Issue 1, pp. 95-99.
2. Balashov E.V., Ivanov N.V., Akhmetov D.B., Korotkov A.S. Vysokotemperaturnyy instrumental'nyy usilitel' [High-temperature instrumental amplifier], *Nanoindustriya* [Nanoindustry], 2020, No. S96-1, pp. 160-163.
3. Hassan A., Amer M., Savaria Y., Sawan M. Towards GaN500-based High Temperature ICs: Characterization and Modeling up to 600°C, *2020 18th International New Circuits and Systems Conference (NEWCAS)*. IEEE, 2020, pp. 275-278.
4. Hassan A., Ali M., Trigui A., Savaria Y., Sawan M. A GaN-Based Wireless Monitoring System for High-Temperature Applications, *Sensors*, 2019, Vol. 19 (8), pp. 1785.
5. Hassan A., Noël J-P., Savaria Y., Sawan M. Circuit Techniques in GaN Technology for High-Temperature Environments, *Electronics*, 2022, Vol. 11 (1), pp. 42.
6. Fresina M. Trends in GaAs HBTs for wireless and RF, *Bipolar/BiCMOS Circuits and Technology Meeting J. IEEE*, 2011, pp. 150-153.
7. Zampardi P.J., Sun M., Cismaru C., Li J. Prospects for a BiCFET III-V HBT Process, *Compound Semiconductor Integrated Circuit Symposium (CSICS)*, 2012, pp. 1-3.
8. Jena M.R., Panda A.K., Dash G.N. A Comparative Analysis InP/InGaAs δ Doped based NPN and PNP HBT, *International Journal of Engineering Research & Technology (IJERT)*, 2019, Vol. 8, Issue 09, pp. 819-823.
9. Lovshenko I.Yu., Kratovich P.S., Stempitskiy V.R., Dvornikov O.V., Kunts A.V., Pavlyuchik A.A. Geteroperekhodnyy bipolyarnyy tranzistor so strukturoy pnp-tipa v arsenid-gallievoy tekhnologii HBT-HEMT [Heterojunction bipolar transistor with a pnp-type structure in arsenide gallium technology HBT-HEMT], *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem – 2022: Sb. trudov* [Problems of development of promising micro- and nanoelectronic systems - 2022. Collection of papers], edited by. ed. Academician of the RAS A.L. Stempkovskogo. Moscow: IPPM RAN, 2022, Issue IV, pp. 149-154.
10. Dvornikov O.V., Pavlyuchik A.A., Prokopenko N.N., Chekhovskiy V.A., Kunts A.V., Chumakov V.E. Arsenid-gallievyy analogovyy bazovyy kristall [Gallium Arsenide analog base crystal] *Problemy razrabotki perpektivnykh mikro- i nanoelektronnykh sistem – 2021: Sb. trudov* [Problems of development of promising micro- and nanoelectronic systems - 2021. Collection of works], edited by. ed. Academician of the RAS A.L. Stempkovskogo. Moscow: IPPM RAN, 2021, pp. 47-54.
11. Cai Y., Cheng Z., Yang Z., Tang C., Lau K., Chen K. High-Temperature Operation of AlGaIn/GaN HEMTs Direct-Coupled FET Logic (DCFL) Integrated Circuits, *Electron Device Letters J. IEEE*, 2007, Vol. 28, No. 5, pp. 328-331.
12. Dvornikov O.V., Pavlyuchik A.A., Prokopenko N.N., Chekhovskiy V.A., Kunts A.V., Chumakov V.E. Unifitsirovannyye skhemotekhnicheskie resheniya analogovykh arsenid-gallievyykh mikroskhem [Unified circuit solutions for analog gallium arsenide microcircuits], *Izvestiya vuzov. Elektronika* [News of universities. Electronics], 2022, Vol. 27, No. 4, pp. 475-488.
13. Nomoto K., Hasegawa K., Satoh M., Nakamura T. Reliability of High-Temperature Operation for GaN-Based OPAMP, *MRS Online Proceedings Library (OPL)*, 2009, pp. 1195-B08-02.
14. Dvornikov O.V., Chekhovskiy V.A., Dyatlov V.L., Prokopenko N.N. Sozdanie nizkotempera-turnykh analogovykh IS dlya obrabotki impul'snykh signalov datchikov. Ch. 3 [Creation of low-temperature analog ICs for processing pulse signals from sensors. Part 3], *Sovremennaya elektronika* [Modern electronics], 2015, No. 6, pp. 34-39.

15. *Dvornikov O.V., Prokopenko N.N., Chumakov V.E., Tchekhovskiy V.A.* General-Purpose Simulation Technique for Charge-Sensitive Amplifiers Based on Wide-Band-Gap (GaAs, GaN) D-FETs, *Proceedings of 16th International Conference of Actual Problems of Electronic Instrument Engineering (APEIE)*. IEEE, 2023.
16. *Takai N., Fujii N.* GaAs MESFET multi-output current mirrors and their application to high frequency filters, *International Conference on Electronics, Circuits and Systems. Surfing the Waves of Science and Technology (Cat. No. 98EX196)*. IEEE, 1998, Vol. 2, pp. 307-310.
17. *Ehteshamuddin M., Salem J.M., Ha D.S.* A high temperature variable gain amplifier based on GaN HEMT devices for downhole communications, *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2017, pp. 1-4.
18. *Fujiwara M., Nagata H., Hibi Y., Matsuo H., Sasaki M.* Cryogenic low noise amplifier with GaAs JFETs, *AIP Conference Proceedings J. American Institute of Physics*, 2009, Vol. 1185, No. 1, pp. 267-270.
19. *Lan D. Ning Y., Wang J., Jiang H.* High performance two-stage bootstrapped GaAs comparator with gain enhancement, *16th Annual Wireless and Microwave Technology Conference (WAMICON)*. IEEE, 2015, pp. 1-4.
20. *Pennisi S., Pulvirenti F., Samperi K.* Frequency Compensation Scheme for a Full GaN OpAmp driving 1-nF load, *2022 International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2022, pp. 2042-2046.

Статью рекомендовал к опубликованию д.т.н., профессор Б.Г. Коноплев.

Бугакова Анна Витальевна – Донской государственный технический университет; e-mail: annabugakova.1992@mail.ru; г. Ростов-на-Дону, Россия; тел.: +79094169380; к.т.н.; доцент кафедры «Информационные системы и радиотехника».

Прокопенко Николай Николаевич – e-mail: prokopenko@sssu.ru; тел.: +79281201984; д.т.н.; профессор; зав. кафедрой «Информационные системы и радиотехника».

Клейменкин Дмитрий Владимирович – e-mail: k-dima-01@mail.ru; тел.: +79281970049; магистрант кафедры «Информационные системы и радиотехника».

Дворников Олег Владимирович – ОАО «Минский научно-исследовательский приборостроительный институт»; e-mail: oleg_dvornikov@tut.by; г. Минск, Беларусь; д.т.н.; доцент; г.н.с.

Чеховский Владимир Алексеевич – Институт ядерных проблем Белорусского государственного университета; e-mail: vtchek@hep.by; г. Минск, Беларусь; зав. лабораторией электронных методов и средств эксперимента.

Bugakova Anna Vitalievna – Don State Technical University; e-mail: annabugakova.1992@mail.ru; Rostov-on-Don, Russia; phone: +79094169380; cand. of eng. sc.; associate professor of the department of information systems and radio engineering

Prokopenko Nikolay Nikolayevich – e-mail: prokopenko@sssu.ru; phone: +79281201984; dr. of eng. sc.; professor; head of the department of information systems and radio engineering.

Kleimenkin Dmitriy Vladimirovich – e-mail: k-dima-01@mail.ru; phone: +79281970049; master's student of the department of information systems and radio engineering.

Dvornikov Oleg Vladimirovich – “Minsk Scientific Research Instrument-Making Institute” JSC; e-mail: oleg_dvornikov@tut.by; Minsk, Belarus; dr. of eng. sc.; associate professor; chief researcher.

Tchekhovskiy Vladimir Alekseevich – Institute of Nuclear Problems, Belarusian State University; e-mail: vtchek@hep.by; Minsk, Belarus; head of the electronic methods and experimental means laboratory.

Раздел III. Алгоритмы обработки информации

УДК 004.023

DOI 10.18522/2311-3103-2024-2-221-228

М.В. Пикалов, А.М. Письмеров

НАСТРОЙКА ПАРАМЕТРОВ ГЕНЕТИЧЕСКОГО АЛГОРИТМА ПРИ ПОМОЩИ АНАЛИЗА ЛАНДШАФТА ФУНКЦИИ ПРИСПОСОБЛЕННОСТИ И МАШИННОГО ОБУЧЕНИЯ

*Выбор значений параметров в эволюционных алгоритмах сильно влияет на их производительность. Многие популярные методы настройки параметров ограничены максимальным числом вычислений целевой функции для поиска хорошего набора значений параметров. Недавно был предложен подход к выбору алгоритмов для решения оптимизационных задач, использующий анализ ландшафта функции приспособленности и машинное обучение для выбора оптимального алгоритма решения задачи на основе особенностей ее ландшафта. Подобное применение анализа ландшафта функции приспособленности мотивирует на дальнейшие исследования, в частности, применительно к настройке параметров эволюционных алгоритмов. Использование признаков ландшафта функции приспособленности позволяет выявлять похожие задачи и использовать данные о настройке параметров, полученные при тестировании на эталонных задачах, что значительно снижает число необходимых вычислений целевой функции при настройке. В этой работе на примере генетического алгоритма $(1 + (\lambda, \lambda))$ рассматривается подход к автоматическому выбору параметров с использованием анализа ландшафта целевой функции и машинного обучения. В предлагаемом решении оцениваются особенности ландшафта целевой функции поставленной задачи оптимизации и предлагаются оптимальные значения параметров алгоритма с помощью нейронной сети. Данная сеть была обучена на наборе данных об особенностях ландшафта, выраженных в виде числовых признаков и соответствующих им оптимальных наборов параметров алгоритма. В отличие от подходов к автоматическому выбору алгоритмов оптимизации для конкретной задачи, в данной работе рассматривается задача регрессии параметров алгоритма вместо проблемы классификации наиболее подходящего алгоритма из заданного набора. Результаты экспериментов на различных конфигурациях задачи *W-model*, а также на задаче *MAX-3SAT* показывают, что предлагаемый подход к автоматическому выбору параметров с учетом ландшафта целевой функции может помочь определить подходящие значения статических параметров генетического алгоритма $(1 + (\lambda, \lambda))$, так как алгоритм с предложенными значениями параметров превосходит другие рассмотренные варианты $(1 + (\lambda, \lambda))$ GA, в среднем требуя меньше вычислений целевой функции для нахождения оптимума, чем остальные рассмотренные алгоритмы.*

Эволюционные алгоритмы; настройка параметров; анализ ландшафта.

M.V. Pikalov, A.M. Pismirov

GENETIC ALGORITHM PARAMETER TUNING USING EXPLORATORY LANDSCAPE ANALYSIS AND MACHINE LEARNING

The choice of parameter values in evolutionary algorithms greatly affects their performance. Many popular parameter tuning methods are constrained by the maximum number of fitness function evaluations to find a good set of parameter values. Recently, an approach to algorithm selection for optimization problems has been proposed, which uses the analysis of the fitness function landscape and machine learning to select the optimal algorithm based on the characteristics of its landscape. Such application of fitness landscape analysis motivates further research, particularly in the context of parameter tuning in evolutionary algorithms. The use of landscape features allows for the identification of similar problems and the use of parameter tuning data obtained from testing on benchmark problems, significantly reducing the number of required fitness function evaluations during tuning. This work considers an approach to automatic parameter selection using landscape analysis of the objective function and machine learning, using

a (1 + (λ, λ)) genetic algorithm as an example. The proposed solution evaluates the characteristics of the landscape of the optimization problem's objective function and suggests optimal parameter values for the algorithm using a neural network. This network was trained on a dataset of landscape features expressed as numerical features and their corresponding optimal algorithm parameter sets. In contrast to approaches for automatic algorithm selection for a specific problem, this work addresses the problem of regressing algorithm parameters instead of classifying the most suitable algorithm from a given set. The results of experiments on different configurations of the W-model problem, as well as on the MAX-3SAT problem, show that the proposed approach to automatic parameter selection considering the landscape of the objective function can help determine appropriate values for the static parameters of the (1 + (λ, λ)) genetic algorithm. The algorithm with the proposed parameter values outperforms other considered (1 + (λ, λ)) GA options on average, requiring fewer evaluations of the objective function to find the optimum compared to the other algorithms considered.

Evolutionary algorithms; parameter tuning; landscape analysis.

Введение. Развитие методов выбора параметров в эволюционных алгоритмах является необходимым условием для их успешного применения на практике [1]. Существует несколько подходов к решению этой проблемы, которые обычно называются настройкой параметров и управлением параметрами.

Анализ ландшафта приспособленности [2] активно изучается как инструмент, который может помочь извлечь знания из экземпляров задач, позволяя настраивать используемые метаэвристики для рассматриваемой оптимизационной задачи. Исследовательский анализ ландшафта [3] рассматривает характеристики ландшафта приспособленности, которые могут быть извлечены из промежуточных решений и их значений приспособленности, что позволяет получить некоторое понимание проблемы, даже в условиях "черного ящика" [4, 5].

Исследовательский анализ ландшафта имеет многочисленные приложения [6]. Одно из перспективных направлений исследований – автоматический выбор алгоритма с помощью анализа ландшафта и машинного обучения [7], а также динамический выбор алгоритма [8].

В этой статье мы представляем наши первые шаги в разработке подхода к автоматическому выбору параметров, используя анализ ландшафта и нейронные сети. В предложенном подходе мы оцениваем особенности ландшафта задачи и предлагаем оптимальные значения параметров, используя нейронную сеть, обученную на наборе данных особенностей ландшафта и соответствующих оптимальных значений параметров. Мы рассматриваем проблему регрессии параметров алгоритма, а не классификации задачи по наилучшему подходящему алгоритму.

Для этого исследования мы собрали набор данных оптимальных значений параметров для генетического алгоритма (1 + (λ, λ)) [9] с четырьмя статическими параметрами на нескольких экземплярах тестовой задачи W-model [10]. Мы вычислили особенности ландшафта для тех же экземпляров задач, используя пакет flacco в R [11]. Используя нейронную сеть, обученную на этом наборе данных, мы можем рекомендовать значения параметров для генетического алгоритма (1 + (λ, λ)) для разных экземпляров задач оптимизации, для которых мы можем вычислить особенности ландшафта их функции приспособленности.

Экспериментальные результаты на различных экземплярах задачи W-model показывают, что предложенный подход, основанный на анализе ландшафта, может помочь определить хорошие значения для статических параметров генетического алгоритма (1 + (λ, λ)). Основным свидетельством этого является то, что алгоритм со значениями параметров, предложенными нашим методом, превосходит другие рассматриваемые варианты (1 + (λ, λ)) GA.

Генетический алгоритм (1 + (λ, λ)). Генетический алгоритм (1 + (λ, λ)), предложенный в [12], представляет собой двухэтапный генетический алгоритм, который показал свою эффективность на нескольких теоретических и практических задачах. Алгоритм пытается исследовать пространство поиска, используя высокие скорости мутаций на первом этапе. На втором этапе он применяет механизм скрещивания для противодействия

возможным негативным последствиям высокой скорости мутаций для отдельных особей. Выбор параметров в $(1 + (\lambda, \lambda))$ GA существенно влияет на его производительность и мотивирует исследования по его обобщению и, в частности, по настройке параметров.

В данной работе рассматривается генетический алгоритм $(1 + (\lambda, \lambda))$ с четырьмя параметрами $\lambda_1, \lambda_2, k, c$, аналогичный рассмотренному в [13]. Сам алгоритм работает следующим образом:

- ◆ на первом этапе каждой итерации алгоритм создает λ_1 мутантных особей, применяя оператор мутации со скоростью мутации k/n , где n – размер задачи;
- ◆ лучший мутант выбирается для скрещивания с родительской особью;
- ◆ на втором этапе создаются λ_2 особи с помощью оператора скрещивания, который берет биты у мутантной особи с вероятностью c ;
- ◆ лучшая особь заменяет родителя, если ее значение приспособленности равно или больше родительского;
- ◆ процесс повторяется до тех пор, пока не будет найден оптимум.

Тестовые задачи. В этой статье мы используем две тестовые задачи для разработки и оценки предлагаемого подхода: задачу W-model и задачу о выполнимости булевой формулы в максимизационной постановке MAX-3SAT.

Задача W-model[10] – это параметризованная задача оптимизации, основанная на известной задаче OneMax: поиске скрытого оптимума на основе количества совпадающих битов в двоичной строке.

W-model вводит несколько настраиваемых слоев, которые влияют на ландшафт целевой функции задачи и, таким образом, делают ее пригодной для сравнительного анализа. Настраиваемые слои задачи W-модели представляют из себя:

- ◆ нейтральность – вводит области с одинаковыми значениями функции пригодности [14];
- ◆ эпистаз – вводит зависимость между различными отдельными генами особей [15];
- ◆ жесткость – вводит так называемые горы и долины в ландшафт целевой функции [16];
- ◆ фиктивность – вводит переменные, не влияющие на значение приспособленности особей.

W-model позволяет нам создавать экземпляры задач с различными свойствами ландшафта функции пригодности, что важно для сбора набора данных о производительности алгоритма оптимизации, необходимого в нашем подходе.

Также, мы рассматриваем задачу о выполнимости булевой формулы в максимизационной постановке MAX-3SAT[17]. Эта задача изучается в эволюционных вычислениях как теоретически, так и практически, и было показано, что $(1 + 1)$ ЭА[18] и $(1 + (\lambda, \lambda))$ GA показывают время выполнения, подобное тому, которое они показывают на задаче OneMax [19].

Основная задача в задаче о выполнимости булевой формулы в максимизационной постановке состоит в том, чтобы найти такую битовую строку $x = x_i, i = 1, \dots, n$, чтобы максимизировать количество удовлетворяемых выражений в булевой формуле в КНФ (конъюнктивной нормальной форме) с тремя переменными в каждом выражении.

Обучающий набор данных. Набор обучающих данных состоит из среднего времени выполнения (числа вычислений целевой функции, необходимого для нахождения оптимума) параметризованного $(1 + (\lambda, \lambda))$ GA с разными значениями параметров на разных экземплярах задачи W-model.

Для $(1 + (\lambda, \lambda))$ GA рассматриваются следующие параметры:

$$\lambda_1 \in [2, 3, 4, \dots, 9]$$

$$\lambda_2 \in [2, 3, 4, \dots, 9]$$

$$k \in [1, 3, 5, 7, 9]$$

$$c \in [0.01, 0.03, 0.05, 0.07, 0.09]$$

Мы вычислили среднее время выполнения $(1 + (\lambda, \lambda))$ GA со всеми возможными комбинациями этих параметров на экземплярах задачи W-model со следующими параметрами:

$$n \in [8, 16, 32, 64, 128]$$

$$dummy \in [1, 1 - 1/n, 1 - 2/n, \dots, 0.8]$$

$$neutrality \in [0, 1, 2, \dots, n \times 0.2]$$

$$epistasis \in [0, 2, 3]$$

Для каждой комбинации параметров $(1 + (\lambda, \lambda)) GA$ $\lambda_1, \lambda_2, k, c$, и параметров задачи W-model $n, dummy, neutrality, epistasis$ 25 раз запускается генетический алгоритм $(1 + (\lambda, \lambda))$. Затем выбирается набор параметров с наименьшим средним временем выполнения в качестве наилучшего набора параметров для конкретного экземпляра задачи.

Для каждого экземпляра задачи W-model из набора данных о производительности мы использовали флассо для вычисления 25 отдельных векторов из 35 характеристик ландшафта функции приспособленности. Каждый вектор признаков рассчитывается из $n/4$ случайных особей, что следует учитывать при анализе эффективности предложенного подхода.

Комбинируя признаки с данными о производительности мы получаем набор данных лучших параметров $(1 + (\lambda, \lambda)) GA$ для определенных признаков ландшафта целевой функции, вычисленных флассо, и размера задачи $n: \{n, features\} \Rightarrow \{\lambda_1, \lambda_2, k, c\}$. Далее мы можем использовать эти данные для обучения нейронной сети, которая сможет предлагать определенные значения параметров для заданного набора признаков ландшафта функции приспособленности.

Обучение нейронной сети. Мы обучили FNN (нейронную сеть с прямой связью) с двумя скрытыми слоями с 32 и 16 узлами соответственно. Также рассматривались и другие архитектуры: сети с количеством скрытых слоев от 1 до 4, в которых было от 4 до 32 нейронов, результаты тестирования и валидации которых оказались хуже выбранной.

Были рассмотрены различные алгоритмы оптимизации, такие как Adam, SGD (стохастический градиентный спуск) и BFGS с ограниченной памятью (Бройден-Флетчер-Гольдфарб-Шанно). Алгоритм Adam показал наименьшую среднеквадратичную ошибку, поэтому он был выбран для окончательного обучения модели. Мы предоставляем вычисленные характеристики ландшафта функции приспособленности экземпляра задачи в качестве входных данных и получаем рекомендуемые значения параметров алгоритма оптимизации для $(1 + (\lambda, \lambda)) GA$ в качестве выходных данных.

Эксперименты. В этом разделе мы представляем наши эмпирические данные об эффективности предложенного подхода. Мы оцениваем время работы алгоритмов, т.е. количество оценок функции приспособленности, необходимых для нахождения оптимального решения.

Для каждой рассматриваемой тестовой задачи мы сравниваем производительность алгоритма с предложенными нейронной сетью параметрами, который мы называем $(1 + (\lambda, \lambda)) GA$ tuned, со следующими вариантами $(1 + (\lambda, \lambda)) GA$, рассмотренными в [20]:

$(1 + (\lambda, \lambda)), \lambda = 4$: $(1 + (\lambda, \lambda)) GA$ с выбором параметров по умолчанию

$\lambda_2 = \lambda_1, k = \lambda_1, c = 1/\lambda_1$ для $\lambda_1 = 4$;

$(1 + (\lambda, \lambda)), \lambda \leq 2 \log n$ и $(1 + (\lambda, \lambda)), \lambda \leq n$: $(1 + (\lambda, \lambda)) GA$ с динамическим λ , настроенный по правилу одной пятой с указанной верхней границей;

$(1 + (\lambda, \lambda)), \lambda \sim \text{pow}(2.5)$: $(1 + (\lambda, \lambda)) GA$ с λ , выбранным из степенного распределения с $\beta = 2.5$;

Мы также рассматриваем параметры в наборе обучающих данных, которые показывает наилучшую общую производительность среди всех вариантов выбора параметров. Эти значения параметров: $\lambda_1 = 2, \lambda_2 = 2, k = 8, c = 0.02$, мы обозначаем $(1 + (\lambda, \lambda)) GA$ с такими параметрами как $(1 + (\lambda, \lambda))$, single best. Для задачи W-model мы также рассмотрели эволюционный алгоритм (ЭА) (1+1) и алгоритм случайного локального поиска (RLS) в качестве дополнительной основы для сравнения алгоритмов.

Результаты на W-model. Сначала мы оценили производительность предложенного подхода на экземплярах W-model с размером задачи $n = 512$. Обратим внимание, что для обучающих данных мы использовали только экземпляры W-model с размером $n = 128$. Увеличение значения n также увеличивает количество различных значений в фиктивном слое и увеличивает верхнюю границу слоя нейтральности.

Примеры параметров для конфигураций W-model представлены в табл. 1. all-default представляет собой $(1 + (\lambda, \lambda))$ GA со статическими параметрами $\lambda_2 = \lambda_1, k = \lambda_1, c = 1/\lambda_1$ при $\lambda_1 = 4$, рекомендованными авторами алгоритма. all-singlebest – $(1 + (\lambda, \lambda))$ GA с параметрами, показавшими в среднем лучшую производительность среди всех рассмотренных наборов параметров на конфигурациях W-model из набора данных для обучения модели, то есть $\lambda_1 = 2, \lambda_2 = 2, k = 8, c = 0.02$. Данные значения параметров не меняются на различных экземплярах задачи. Каждая конфигурация W-model определяется параметрами: n – размерность, dum, neu, epi – параметры слоев dummy, neutrality, epistasis соответственно. Для экземпляров задачи W-model представлены рекомендуемые обученной нейронной сетью параметры.

Таблица 1

Параметры для W-model

Конфигурация задачи				Параметры $(1 + (\lambda, \lambda))$ GA			
n	dum	neu	epi	λ_1	λ_2	k	c
all-default				4	4	0.25	4
all-singlebest				2	2	0.02	8
256	0.8	50	1	4	2	0.071	9
256	0.95	0	0	3	5	0.052	1
512	0.90	80	1	5	4	0.012	4
512	0.99	25	2	4	5	0.069	9
256	1	0	3	3	3	0.034	3
512	0.85	0	2	5	7	0.045	2

Модель предлагает различные параметры $(1 + (\lambda, \lambda))$ GA для разных конфигураций задачи W-model. Анализируя зависимость предлагаемых параметров от конфигурации задачи можно заметить, что для больших значений размерности задачи n предлагаются большие значения параметров λ . Также при больших значениях параметра слоя нейтральности, модель рекомендует большие значения параметра k .

Результирующие показатели производительности показаны в табл. 2. Для каждого алгоритма приводится его средний ранг среди всех рассмотренных алгоритмов $Mean_r$, медиана ранга $Median_r$, стандартное отклонение ранга Std_r и средняя разница в числе вычислений функции приспособленности $Mean_d$ для алгоритма по сравнению с $(1 + (\lambda, \lambda))$ GA tuned, с учетом вычислений для сбора выборки для анализа ландшафта. Отрицательные значения разницы указывают на то, что конкретный алгоритм работает лучше, чем $(1 + (\lambda, \lambda))$ GA tuned, а положительные значения указывают на то, что алгоритм работает хуже.

Таблица 2

Сравнение алгоритмов на задаче W-model с размерностью 512

Алгоритм	$Mean_r$	$Median_r$	Std_r	$Mean_d$
RLS	2.93	3	1.83	-154.48
$(1 + (\lambda, \lambda))$, tuned	3.13	3	1.52	0
$(1 + (\lambda, \lambda))$, $\lambda \leq n$	3.40	3	1.91	204.54
$(1 + (\lambda, \lambda))$, $\lambda \leq 2 \log n$	3.86	4	1.58	247.22
$(1 + (\lambda, \lambda))$, $\lambda \sim \text{pow}(2.5)$	4.53	5	1.20	399.10
$(1 + (\lambda, \lambda))$, $\lambda = 4$	5.09	5	2.14	706.25
$(1 + 1)$ EA	6.32	6	1.35	886.95
$(1 + (\lambda, \lambda))$, single best	6.56	7	1.48	1272.10

Из результатов видно, что в среднем $(1 + (\lambda, \lambda))$ GA tuned является лучшим из всех рассмотренных вариантов $(1 + (\lambda, \lambda))$ GA. В большинстве случаев разница в производительности больше, чем количество оценок, необходимых для вычисления характеристик ландшафта функции приспособленности, и она увеличивается с размером задачи n . Вер-

сии $(1 + (\lambda, \lambda))$ GA с динамической настройкой параметров улучшают свою производительность с увеличением размера задачи, что может означать, что им либо нужно время, чтобы найти правильные значения параметров, либо они значительно выигрывают от изменения λ на разных этапах оптимизации. Тем не менее, $(1 + (\lambda, \lambda))$ GA tuned имеет лучшую производительность, чем они, что говорит о том, что он может выбирать хорошие значения параметров перед запуском оптимизации.

Кроме того, существует значительная разница в производительности $(1 + (\lambda, \lambda))$ GA tuned и $(1 + (\lambda, \lambda))$ GA single best. На $n = 512$ алгоритм single best показывает худшую производительность из всех, в то время как $(1 + (\lambda, \lambda))$ GA tuned имеет отличную производительность. Это наблюдение заставляет нас поверить, что нейронная сеть способна обучаться выбору параметров в зависимости от особенностей ландшафта целевой функции, а не просто выбирает лучшие в среднем параметры из обучающего набора данных.

Также были изучены конкретные диапазоны параметров W-model, чтобы понять преимущества и недостатки обученной модели и предлагаемого метода. Разделив диапазоны параметров на интервалы, был проведен сравнительный анализ производительности рассмотренных алгоритмов на соответствующих конфигурациях задач W-model на основе средних рангов алгоритмов и средней разницы числа вычислений функций приспособленности. В среднем предложенный подход имеет лучшую производительность на всех рассмотренных интервалах значений параметров.

Наконец, отметим, что алгоритмы $(1 + 1)$ EA начинают заметно снижать производительность с увеличением размерности задачи, а алгоритм RLS не в состоянии найти оптимум на нескольких экземплярах задачи W-model. Возможно, эти результаты объясняются тем, что параметры, предлагаемые нейронной сетью для алгоритма $(1 + (\lambda, \lambda))$, позволяют ему повторять поведение этих простых алгоритмов, тем самым быстрее решая конкретные экземпляры задач, на которых $(1 + (\lambda, \lambda))$ GA с его типичными значениями параметров уступает по производительности алгоритмам RLS и $(1 + 1)$ EA.

Результаты на MAX-3SAT. Также, мы оценили предложенный подход на задаче MAX-3SAT, которая также не использовалась для сбора обучающего набора данных. Выражения булевой формулы в опытах генерировались случайным образом и независимо друг от друга. Всего было рассмотрено около 1000 случайных конфигураций с размерностью задачи $n \in [256, 512]$. Результаты для задачи MAX-3SAT представлены в табл. 3.

Таблица 3

Сравнение алгоритмов на задаче MAX-3SAT

Алгоритм	$Mean_r$	$Median_r$	Std_r	$Mean_d$
$(1 + (\lambda, \lambda))$, tuned	2.45	2	1.10	0
$(1 + (\lambda, \lambda))$, $\lambda \leq n$	2.57	2	1.21	68.22
$(1 + (\lambda, \lambda))$, $\lambda \leq 2 \log n$	2.79	3	1.29	104.54
$(1 + (\lambda, \lambda))$, $\lambda \sim \text{pow}(2.5)$	3.82	4	1.21	209.10
$(1 + (\lambda, \lambda))$, $\lambda = 4$	4.90	5	0.98	299.50
$(1 + (\lambda, \lambda))$, single best	5.07	5	1.32	355.52

Как в других результатах, предложенный $(1 + (\lambda, \lambda))$ GA tuned оказывается более эффективен, чем другие варианты $(1 + (\lambda, \lambda))$ GA. В отличие от задачи W-model, наблюдаются более устойчивые результаты и меньший разброс рангов. Это можно еще раз объяснить тем, что экземпляры MAX-3SAT с разной конфигурацией хоть и имеют разные характеристики ландшафта, но все же больше похожи друг на друга, чем разные конфигурации W-model.

Выводы. В этой статье мы изучаем подход к автоматической настройке параметров с использованием анализа ландшафта функции приспособленности и машинного обучения. Этот подход направлен на использование анализа ландшафта целевой функции для переноса знаний об оптимальных параметрах из экземпляров задач, оцененных ранее, на новые экземпляры задач с аналогичными характеристиками ландшафта. В результате это

позволяет нам пропустить этап обучения, обычно присутствующий в подходах к настройке, за счет использования данных, полученных на задачах с различными особенностями ландшафта.

Мы оценили предложенный подход на генетическом алгоритме $(1 + (\lambda, \lambda))$ с четырьмя параметрами и на нескольких задачах оптимизации. Мы собрали данные о производительности алгоритма на разных экземплярах задачи W-model с различными характеристиками ландшафта функции приспособленности и обучили нейронную сеть на этих данных. Сеть способна предложить хорошие значения параметров для различных задач оптимизации на основе особенностей их ландшафта. Кроме того, мы отмечаем значительную разницу в производительности между параметрами, предложенными нейронной сетью, и лучшими в среднем параметрами из обучающего набора данных, что говорит о том, что алгоритм машинного обучения способен извлекать зависимости между особенностями ландшафта и выбором параметров.

Исходя из наших результатов, мы с оптимизмом смотрим на то, что применение анализа ландшафта целевой функции и машинного обучения для настройки параметров может помочь нам разработать более эффективные инструменты для решения задачи выбора параметров. В дальнейшем мы планируем расширить этот подход к управлению параметрами путем переоценки характеристик ландшафта и рекомендуемых параметров во время процесса оптимизации.

Для будущих исследований у нас есть несколько целей. Прежде всего, важно дополнительно оценить предлагаемый подход на большем числе задач. Это может не только помочь нам лучше понять его преимущества и недостатки, но и предоставить больше данных для обучения, что всегда хорошо в машинном обучении. Кроме того, мы считаем важным изучить различные стратегии выборки для оценки характеристик ландшафта и их влияния на наши результаты. Наконец, разработка простого инструмента для настройки параметров с использованием нашего подхода сделает наши результаты легко доступными для других, а также значительно упростит сбор данных о производительности различных алгоритмов, что является важным требованием для эффективности нашего подхода.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Lobo F.J., Lima C.F., Michalewicz Z. (ed.). Parameter setting in evolutionary algorithms. Springer Science & Business Media, 2007, Vol. 54.
2. Mersmann O., Preuss M., Trautmann H. Benchmarking evolutionary algorithms: Towards exploratory landscape analysis, 2010.
3. Mersmann O. et al. Exploratory landscape analysis, *Proceedings of the 13th annual conference on Genetic and evolutionary computation*, 2011, pp. 829-836.
4. Lopez-Ibanez, M., Dubois-Lacoste J., Caceres L.P., Stutzle T., Birattari M. The irace package: Iterated racing for automatic algorithm configuration, *Operations Research Perspectives*, 2016, 3, pp. 43-58.
5. Hutter F., Hoos H.H., Leyton-Brown K. Sequential model-based optimization for general algorithm configuration, *Proceedings of Learning and Intelligent Optimization*. Springer, 2011, pp. 507-523.
6. Ochoa G., Malan K. Recent advances in fitness landscape analysis, *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 2019, pp. 1077-1094.
7. Kerschke P., Trautmann H. Automated algorithm selection on continuous black-box problems by combining exploratory landscape analysis and machine learning, *Evolutionary computation*, 2019, pp. 99-127.
8. Janković A., Doerr C. Adaptive landscape analysis, *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 2019, pp. 2032-2035.
9. Bassin A., Buzdalov M. The $(1 + (\lambda, \lambda))$ Genetic Algorithm for Permutations, *Proceedings of Genetic and Evolutionary Computation Conference Companion*. ACM, 2020, pp. 1669-1677.
10. Weise T., Wu Z. Difficult features of combinatorial optimization problems and the tunable w-model benchmark problem for simulating them, *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 2018, pp. 1769-1776.
11. Kerschke P., Trautmann H. The R-Package FLACCO for exploratory landscape analysis with applications to multi-objective optimization problems, *2016 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2016, pp. 5262-5269.
12. Doerr B., Doerr C., Ebel F. From black-box complexity to designing new genetic algorithms, *Theoretical Computer Science*, 2015, Vol. 567, pp. 87-104.
13. Dang N., Doerr C. Hyper-parameter tuning for the $(1+(\lambda, \lambda))$ GA, *Proceedings of the Genetic and Evolutionary Computation Conference*, 2019, pp. 889-897.

14. Kimura M. [et al.]. Evolutionary rate at the molecular level, *Nature*, 1968, Vol. 217, No. 5129, pp. 624-626.
15. Cordell H.J. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans, *Human molecular genetics*, 2002, Vol. 11, No. 20, pp. 2463-2468.
16. Weise T. [et al.]. A tunable model for multiobjective, epistatic, rugged, and neutral fitness landscapes, *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, 2008, pp. 795-802.
17. Garay M.R. Computers and intractability. A guide to the theory of NP-completeness, 1979.
18. Beyer H.G. Evolution strategies, *Scholarpedia*, 2007, Vol. 2, No. 8, pp. 1965.
19. Eshelman L. On crossover as an evolutionarily viable strategy, *Proceedings of the Fourth International Conference on Genetic Algorithms*, 1991, pp. 61-68.
20. Schumer M., Steiglitz K. Adaptive step size random search, *IEEE Transactions on Automatic Control*, 1968, Vol. 13, No. 3, pp. 270-276.

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

Пикалов Максим Вадимович – Национальный исследовательский университет ИТМО; e-mail: pikmaksim@gmail.com; г. Санкт-Петербург, Россия; тел.: 88126070283; аспирант.

Письмеров Алексей Максимович – e-mail: alekseipismerov@gmail.com; аспирант.

Pikalov Maxim Vadimovich – ITMO University; e-mail: pikmaksim@gmail.com; St. Petersburg, Russia; phone: +78126070283; graduate student.

Pismerov Aleksei Maximovich – e-mail: alekseipismerov@gmail.com; graduate student.

УДК 004.932.2

DOI 10.18522/2311-3103-2024-2-228-236

П.А. Гессен, А.И. Лизин, В.А. Павлова, В.Д. Саенко, В.А. Тупиков

РЕКУРСИВНЫЙ АЛГОРИТМ АНАЛИЗА И ВОССТАНОВЛЕНИЯ КОНТУРОВ В СИСТЕМАХ НАВИГАЦИИ И НАВЕДЕНИЯ

В целях разработки алгоритма обнаружения объектов для встраиваемых вычислительных систем оптико-электронных комплексов, был проведен анализ имеющегося мирового научно-технического опыта, направленного на улучшение процесса выделения контуров. На основе проведенного анализа, авторами статьи был разработан новый метод коррекции контурных изображений. Этот метод реализует подход, позволяющий объединять разорванные контуры и применять фильтрацию по различным параметрам для оптимальной анализа контуров. Первым этапом работы алгоритма является применение размытия изображения, за которым следует применение алгоритма детекции границ Кенни. Затем происходит утоньшение контуров и фильтрация контурного изображения для удаления самых слабых контуров. Следующими этапами являются создание и обработка каждого отдельного контура, а также фильтрация выбросов. Заключительным этапом является соединение и поиск точек перегиба контура. В рамках работы выделены как преимущества, так и недостатки классических методов выделения контуров в контексте использования их в алгоритмах обнаружения объектов. Авторами исследования был проведен анализ двух классических морфологических операторов - дилатации и эрозии, а также существующих основных вариаций их применения таких как открытие и закрытие, как методов объединения контуров. В результате сравнительного анализа результатов работы морфологических операторов дилатации и эрозии, а также основных вариаций их применения, с рекурсивным алгоритмом анализа и восстановления контуров было выявлено преимущество последнего в части сохранения целостности морфологических признаков объектов. Авторами предложены также идеи дальнейшего развития рекурсивного алгоритма анализа и восстановления контуров, а также его дальнейшего применения в задачах обнаружения объектов на изображениях.

Контурный анализ; обработка изображений; восстановление контуров.

P.A. Gessen, A.I. Lizin, V.A. Pavlova, V.D. Saenko, V.A. Tupikov

RECURSIVE ANALYSIS ALGORITHM AND RESTORATION OF CONTOURS IN NAVIGATION AND GUIDANCE SYSTEMS

In order to develop an object detection algorithm for embedded computing systems of optical-electronic complexes, an analysis of the existing world scientific and technical experience was carried out, aimed at improving the process of identifying contours. Based on the analysis, the authors of the article developed a new method for correcting contour images. This method implements an approach that allows you to merge broken contours and apply filtering based on various parameters for optimal contour analysis. The first step of the algorithm is to apply blur to the image, followed by the application of the Kenny edge detection algorithm. Then the contours are thinned and the contour image is filtered to remove the weakest contours. The next steps are the creation and processing of each individual contour, as well as filtering outliers. The final stage is to connect and search for inflection points of the contour. The work highlights both the advantages and disadvantages of classical edge extraction methods in the context of their use in object detection algorithms. The authors of the study analyzed two classical morphological operators - dilatation and erosion, as well as the existing basic variations of their use, such as opening and closing, as methods for combining contours. As a result of a comparative analysis of the results of the work of morphological operators of dilatation and erosion, as well as the main variations of their application, with a recursive algorithm for analyzing and restoring contours, the advantage of the latter in terms of preserving the integrity of the morphological characteristics of objects was revealed. The authors also proposed ideas for further development of a recursive algorithm for analysis and restoration of contours, as well as its further application in problems of detecting objects in images.

Contour analysis; image processing; contour restoration.

Введение. В системах компьютерного зрения при выделении контуров стандартными алгоритмами существует проблема низкой точности и наличия большого количества шума. Целью данной статьи является разработка алгоритма обнаружения объектов по характеристикам их формы, то есть по морфологическим признакам. Данный алгоритм будет особенно полезен при обнаружении объектов на изображениях в инфракрасном спектре.

В данной статье рассматривалось множество алгоритмов выделения контуров [1–7], но были выбраны классические алгоритмы, такие как детектор границ Кэнни [8], оператор Собеля [9] и векторный оператор Лапласа [10] из-за их быстродействия. Классические алгоритмы имеют недостатки, связанные с разрывом контуров объектов и появлением шума. Все эти недостатки могут негативно влиять на результаты алгоритмов, которые базируются на контурных изображениях.

Детектор границ Кэнни – начальный этап включает применение оператора первой производной Гауссиана с целью сглаживания изображения и снижения чувствительности к шуму. Далее происходит отбор точек с максимальным значением градиента на контуре границы. Для сохранения целостности границы в окрестности локальных максимумов градиента используются четыре фильтра, направленных на выявление горизонтальных, вертикальных и диагональных границ. В последующем этапе применяются два порога для удаления границ низкой интенсивности. Таким образом, алгоритм Кэнни обеспечивает подавление краев, не связанных с сильными границами, что позволяет сохранить четкость и структуру границ на изображении.

Оператор Собеля представляет собой метод аппроксимации градиента яркости изображения, который вычисляет градиент яркости в каждой точке изображения, определяя величину изменения яркости и ее направление. Этот оператор находит широкое применение в анализе изображений, позволяя выявить границы объектов и определить их ориентацию.

Результатом применения оператора Собеля являются значения изменения яркости в каждой точке изображения, что позволяет оценить вероятность нахождения точки на границе объекта, а также определить направление изменения яркости вдоль этой границы.

Оператор Собеля основан на применении фильтрации изображения путем свертки по горизонтали и вертикали с использованием матрицы размером 3x3. Этот подход обеспечивает эффективное вычисление приближенных производных по горизонтальным и вертикальным направлениям, что делает оператор Собеля легко применимым для обработки изображений.

Векторный оператор Лапласа основанным на анализе нулевых значений. Он использует вторые производные для выявления границ на изображениях. В отличие от методов, основанных на градиентном подходе, оператор Лапласа выражается в виде скалярной функции. Его применение для выделения границ состоит в том, что оператор Лапласа достигает максимальных значений в точках, где функция яркости имеет "перегибы".

Недостатком этого алгоритма является возможное раздваивание границ в случае недостаточно резких перепадов яркости, когда присутствуют участки с постоянным наклоном. В таких ситуациях необходима специальная обработка для предотвращения раздваивания линии.

Вследствие большого количества шума и контуров ширина которых превышает один пиксель при выделении их операторами Собеля и Лапласа, алгоритм, представленный в данной статье, работает только с контурами, выделенными детектором границ Кэнни. На рис. 1 представлены результаты работы алгоритмов выделения контуров.

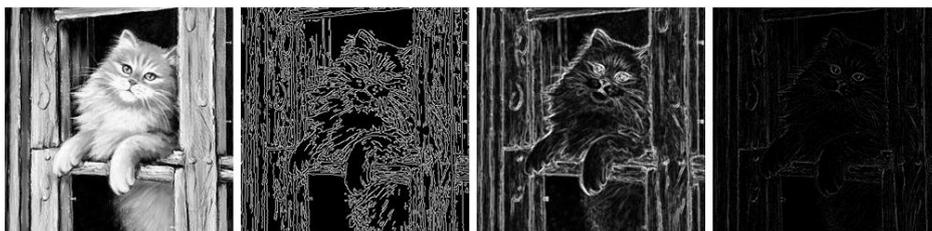


Рис. 1. Исходное изображение и результаты работы алгоритмов выделения контуров (Кэнни, Собель, Лаплас)

На основе анализа описанных алгоритмов авторами был разработан новый алгоритм анализа контуров, позволяющий соединять замкнутые и разорванные контура, фильтровать шум на контурном изображении делать контуры толщиной в один пиксель и выделять ключевые особенности контуров.

Применяемый подход. Алгоритм анализа контуров, представленный в данной статье, базируется на идее рекурсивного обхода графа в глубину [11, 12]. Рекурсивный обход графа в глубину выполняется следующим образом:

1. Перебор всех рёбер, исходящих из рассматриваемой вершины.
2. Если ребро ведёт в вершину, которая не была рассмотрена ранее, то запускается алгоритм от этой нерассмотренной вершины.
3. Возвращение и продолжение перебора рёбер.

Также в алгоритме анализа контуров, представленном в данной статье, используется алгоритм проверки на восьмисвязность [13]. Данный алгоритм проверяет окрестность Мура первого порядка [14] у рассматриваемого пикселя. Подходящие под яркостные параметры будут «соседями» рассматриваемого пикселя. Применяв данный алгоритм к каждому «соседу», возможно обойти все пиксели, состоящие в одном объекте на бинаризованном изображении.

Два вышеописанных алгоритма используются в данной статье для обхода бинаризованного контурного изображения для выделения отдельных контуров как объектов.

На рис. 2 представлена общая блок-схема разработанного алгоритма.

Блок предобработки. В качестве предобработки изображения перед работой алгоритма анализа контуров выполняются следующие шаги:

1. Размытие входного изображения. Данный шаг позволяет повысить точность работы алгоритмов по выделению контуров.
2. Применение детектора границ Кэнни.
3. Утоньшение контуров посредством разности контурного изображения и размытого контурного изображения. Данный шаг позволяет уменьшить толщину контуров до одного пикселя, но данный метод имеет погрешность и не всегда контуры после данного шага имеют толщину в один пиксель, дальнейшее утоньшение будет выполнено алгоритмом анализа контуров.

4. Фильтрация контурного изображения от самых слабых контуров. В следствии третьего шага предобработки контура будут выделены по силе с помощью яркости.

Где слабые контура это контура чья яркость меньше 3 по шкале от 0 до 255 (0 – чёрный, а 255 – белый).

После блока предобработки перебираются все пиксели на изображении в поисках пикселя принадлежащего к контуру какого-либо объекта. Как только такой пиксель будет найден начнётся, обработка контура, содержащего данный пиксель.

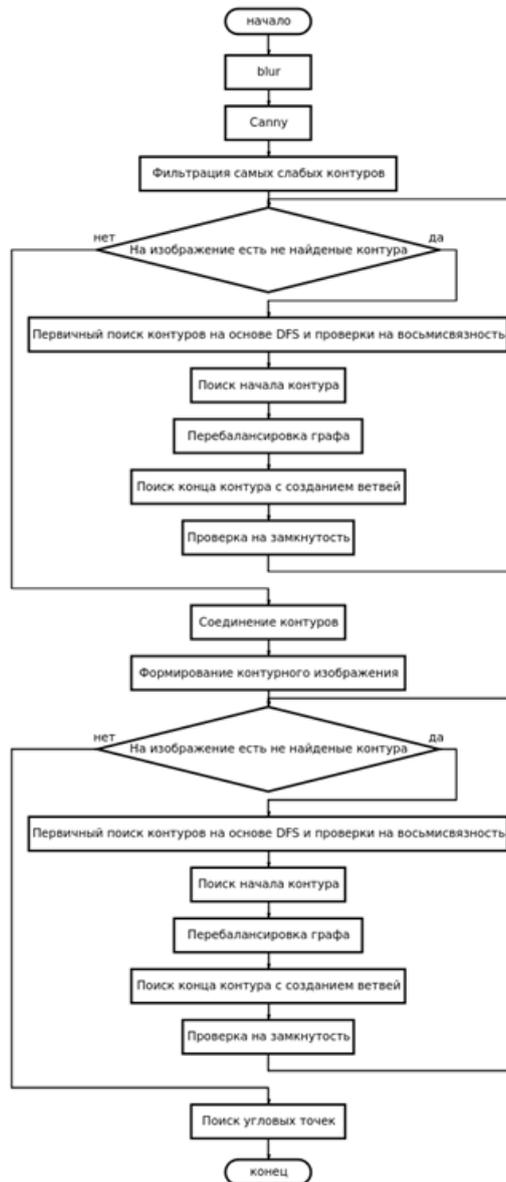


Рис. 2. Блок-схема алгоритма анализа контуров

Блок создания и обработки отдельного контура. Обработка контура строится на рекурсивном обходе контура как единого объекта с помощью проверки на восьмисвязность. Первый найденный пиксель контура будет выбран как ключевой. Все соседи ключевого пикселя будут являться частью одного контура. Во время одной итерации рекурсивного алгоритма обнаруживаются все «соседи» данного пикселя. Следующая итерация

запускается по первому из непосещённых «соседей». Как только у рассматриваемого пикселя не находится не посещённых «соседей», происходит возврат из рекурсии. Пиксель, у которого не осталось непосещённых соседей, является краевым пикселем. При выходе из рекурсии, для рассматриваемого пикселя считается расстояние до самого дальнего из известных краевых пикселей. Строится направленный граф, где узлами являются пикселями контура, а рёбрами графа связи между рассматриваемым пикселем и его «соседями». Корнем графа является ключевой пиксель. Направленность ребра графа задаётся относительно удалённости рассматриваемого пикселя до ключевого пикселя.

По рассчитанному расстоянию удалённости от конечного пикселя на графе фильтруются «выбросы» (рис. 3). В следствии фильтрации «выбросов» алгоритм анализа контуров, описанный в данной статье, позволяет получить контура шириной в один пиксель.

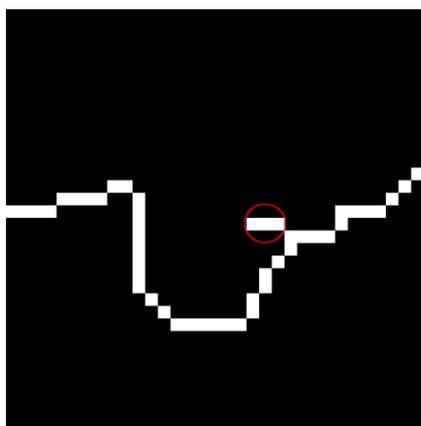


Рис. 3. Пример фильтрации выбросов (в красный круг обведён «выброс»)

Дальнейшим этапом является поиск базовой части контура, т.е. самой длинной части контура без ответвлений. Длина базовой части контура считается от одного конечного пикселя до другого. Первым этапом будет поиск самого удалённого от корня графа узла. Данный узел станет стартовым в рассматриваемом контуре. Во время поиска стартового узла происходит перебалансировка графа, таким образом, что корнем графа станет стартовый узел графа. Далее выполняется обход графа в глубину с целью пересчёта расстояний до самого далёкого узла графа. По пересчитанным расстояниям ищется пиксель конца базовой части контура (рис. 4), как самый удалённый узел графа от его корня.

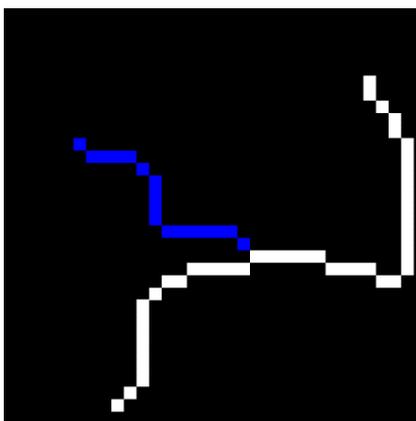


Рис. 4. Пример базовой части контура и его ветви (белым цветом показана основная часть контура, синим – ветвь контура)

По аналогии с базовой частью контура рассматриваются его ветви. У каждой ветви уже известен стартовый узел, им является точка разветвления контура, в которой начинается наша ветвь. Выполняется поиск конца ветви таким же образом как поиск конца основной ветви. Данная операция выполняется до тех пор, пока не будут пройдены все ветви контура.

Для удобства представления формируется новый граф. Корневым узлом нового графа является базовая часть контура. Узлами же представлены ветви контура. Рёбрами представлены пиксели разветвления контура.

Следующим этапом анализа отдельного контура является проверка данного контура на замкнутость. Выполняется данная проверка путём сравнения расстояния между началом и концом базовой части контура с заданными порогом, если начало и конец базовой части контура ближе друг к другу чем заданный порог, то с помощью алгоритма Брезенхэма данный контур дополняется недостающими до замкнутости контура пикселями.

Последним этапом обработки отдельного контура является выделение ключевых точек контура. Ключевыми точками являются пиксели конца ветвей контура и пиксели начала и конца базовой части контура. Также к ключевым точкам относятся точки разветвления контура.

Блок соединения контуров. После анализа всех контуров на изображении по отдельности пришло время для анализа всех контуров в совокупности. Главным этапом, представленным в данной статье, является объединение контуров по ключевым точкам. Объединение происходит по принципу полного перебора ключевых точек контуров. Сначала перебираются все возможные пары контуров, в этих парах перебираются все возможные пары ключевых точек, таким образом, что одна точка из пары принадлежит первому выбранному контуру, а вторая точка соответственно второму контуру. Рассчитывается расстояние между выбранными ключевыми точками, если расстояние меньше заданного порога, то контур дополняется с помощью алгоритма Брезенхэма.

Далее формируется изображение содержащие соединённые и очищенные контуры. Для избавлений от шума и огрех соединения контуров заново выполняется блок создания и обработки отдельных контуров

Блок поиска точек перегиба контура. В следствии специфического метода хранения контура алгоритм анализа контуров, предложенный авторами способен находить точки перегиба и точки углов контура. Поиск точек перегиба выполняется путём осмотра трёх последовательно идущих пикселей через некоторый шаг, между этими тремя пикселями рассчитывается угол. Если данный угол больше некоторого порогового значения, то принимается решение, что угол находится где-то между первым и третьим пикселем. Для подтверждения более точных координат угла, перебираются все возможные тройки пикселей в промежутке между первым и третьим пикселями, только шаг между пикселями из новой тройки должен быть меньше, чем изначальный шаг.

Вариативные настройки алгоритма анализа контуров. Алгоритм анализа контуров, представленный в данной статье, предоставляет множество вариативных настроек, таких как:

1. Пороговое значение длины «выбросов».
2. Пороговое значение длины для проверки контура на замкнутость.
3. Пороговое значение длины для объединения контуров.

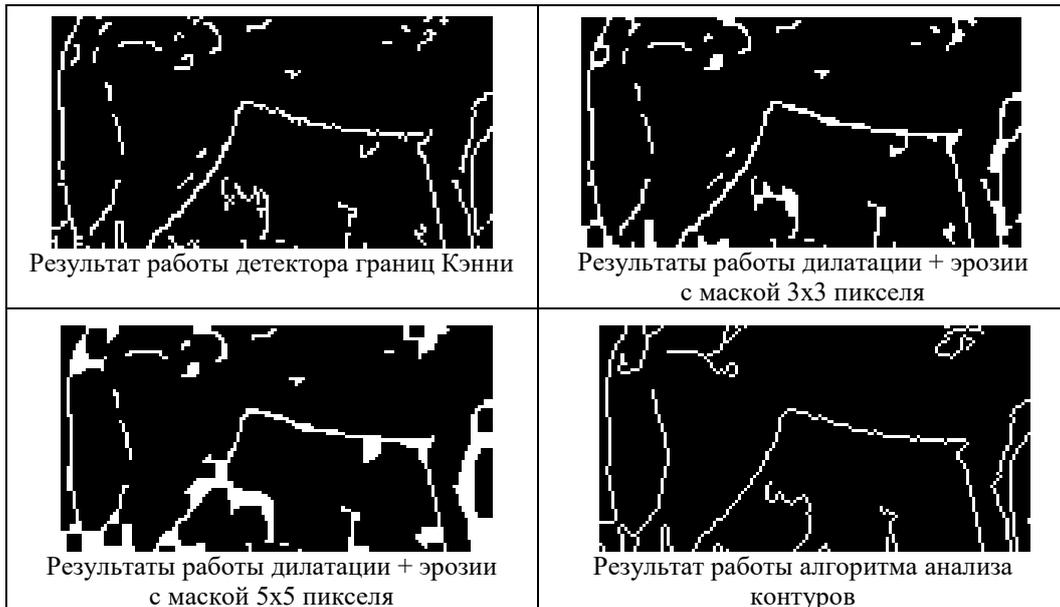
Параметры фильтрации контуров. В алгоритме, предложенном в данной статье, имеется возможность фильтровать контура по таким параметрам как:

1. Длина основной части контура.
2. Соотношение длины основной части контура и суммы длин ветвей данного контура.
3. Общая удалённость контура от других контуров.
4. Соотношение длины основной части контура к хорде контура.

Сравнение результатов работы алгоритма анализа контуров с эрозией и дилатацией. Для объединения контуров классическими методами довольно часто используются дилатация и эрозия [15]. Первым этапом к контурному изображению применяется дилатация. Вторым этапом к изображению после дилатации применяется эрозия.

Таблица 1

 <p>Оригинальное изображение в чёрно-белом формате</p>	 <p>Зона интереса на оригинальном изображении</p>
 <p>Результат работы детектора границ Кэнни</p>	 <p>Результаты работы дилатации + эрозии с маской 3x3 пикселя</p>
 <p>Результаты работы дилатации + эрозии с маской 5x5 пикселя</p>	 <p>Результат работы алгоритма анализа контуров</p>
 <p>Оригинальное изображение в чёрно-белом формате</p>	 <p>Зона интереса на оригинальном изображении</p>



По результатам визуального сравнения (табл. 1) можно сделать выводы, что алгоритмы дилатации и эрозии, а также существующие основные вариации их применения, такие как открытие и закрытие, не позволяют получить морфологические признаки объектов без потерь и существенных искажений в отличие от предложенного алгоритма.

Заключение. Резюмируя изложенное, можно сделать вывод, что разработанный алгоритм позволяет выполнять обработку контурных изображений, в ходе которой уменьшается количество контурного шума, и соединяются контура, разделённые неидеальной работой алгоритмов контурного выделения. Такая обработка даёт повысить качество работы алгоритмов обнаружения, работающих по контурным изображениям.

В продолжение данных исследований планируется оптимизировать процесс объединения контуров и избавиться от повторного обнаружения и анализа отдельных контуров на изображении после их объединения, для уменьшения затрачиваемых вычислительных ресурсов и повышения качества объединения контуров. Планируется также исследовать возможность фильтрации контуров, используя зависимость от времени с помощью методов оптического потока [16–20].

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Amer G.M.H., and Abushaala A.M. Edge Detection Methods, in *2015 2nd World Symposium on Web Applications and Networking (WSWAN) (IEEE)*, 2015, pp. 1-7.
2. Arbelaez P., Maire M., Fowlkes C., and Malik J. Contour Detection and Hierarchical Image Segmentation, *IEEE Trans. Pattern Anal. Mach Intell.*, 2010, 33 (5), pp. 898-916. DOI: 10.1109/TPAMI.2010.161.
3. Deng R., and Liu S. Deep Structural Contour Detection, in *Proceedings of the 28th ACM International Conference on Multimedia*, 2020, pp. 304-312.
4. Duan R.L., Li Q.X., and Li Y.H. Summary of Image Edge Detection, *Opt. Tech.*, 2005, 3 (3), pp. 415-419.
5. Marr D., and Hildreth E. Theory of Edge Detection, *Proc. R. Soc. Lond. B Biol. Sci.*, 1980, 207 (1167), pp. 187-217.
6. Xiaofeng R., and Bo L. Discriminatively Trained Sparse Code Gradients for Contour Detection, *Adv. Neural Inf. Process. Syst.*, 2012, 25.
7. Davis L.S. A Survey of Edge Detection Techniques, *Computer graphics image Process*, 1975, Vol. 4, No. 3, pp. 248-270.
8. Canny J. A Computational Approach to Edge Detection, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1986, 8 (6), pp. 679-698.
9. Kittler J. On the accuracy of the Sobel edge detector, *Image and Vision Computing*, 1983, Vol. 1, No. 1, pp. 37-42.
10. Marr D. *Vision*, Freeman, 1982, Chap. 2, pp. 54-78.
11. Shimon E. *Graph Algorithms*, Computer Science Press, 1979. ISBN 0-7167-8044-5.

12. Reif J.H. Depth-first search is inherently sequential, *Information Processing Letters*, 1985, 20 (5), pp. 229-234.
13. Gonzalez R.C., Woods R.E. Digital image processing. Boston, MA Addison-Wesley, 2001, pp. 90-93.
14. Moore E.F. Machine models of self-reproduction, 1962, pp. 17-31.
15. Vernon D. Machine Vision. Prentice-Hall, 1991, pp. 63-66, 76-78.
16. Lucas B.D. and Kanade T. An iterative image registration technique with an application to stereo vision, *International Joint Conference on Artificial Intelligence*, 1981, pp. 674-679.
17. Horn B.K.B.; Schunck B.G. Determining Optical Flow, *Artif. Intell.*, 1981, 17, pp. 185-203.
18. Hans-Hellmut Nagel. On the Estimation of Optical Flow: Relations between Different Approaches and Some New Results, *Artificial intelligence*, 1981, pp. 299-324.
19. Beauchemin S.S., Barron J.L. The Computation of Optical Flow, *ACM Computing Surveys*, 1995, Vol. 27, No. 3, pp. 433-467,
20. Denis Fortun, Patrick Bouthemy, Charles Kervrann. Optical flow modeling and computation: a survey, *Computer Vision and Image Understanding*, 2015, 134, pp. 21.

Статью рекомендовал к опубликованию к.т.н. П.К. Шульженко.

Тупиков Владимир Алексеевич – АО «Научно-производственное предприятие «Авиационная и морская электроника»; e-mail: tupikov@nppame.ru; г. Санкт-Петербург, Россия; тел.: 88123274667; д.т.н.; профессор.; зам. ген. Директора; директор научно-производственного комплекса робототехнических систем специального назначения.

Павлова Валерия Анатольевна – e-mail: pavlova@nppame.ru; тел.: 88123274667; к.т.н.; зам. директора по НИОКР.

Лизин Алексей Игоревич – e-mail: lizin_a@nppame.ru; начальник центра средств интеллектуальной обработки изображений.

Гессен Павел Алексеевич – e-mail: gessen_pa@nppame.ru; зам. начальника центра средств интеллектуальной обработки изображений.

Саенко Владислав Дмитриевич – e-mail: gessen_pa@nppame.ru; младший программист.

Tupikov Vladimir Alekseevich – SPE "Research and Production Enterprise "Air and Marine Electronics", e-mail: tupikov@nppame.ru; Saint Petersburg, Russia; phone: +78123274667; dr. of eng. sc.; professor; deputy director general; director of Research and production complex of special purpose robotic systems.

Pavlova Valeria Anatolyevna – e-mail: pavlova@nppame.ru; cand. of eng. sc.; deputy director of research at Research and production complex of special purpose robotic systems.

Lizin Aleksey Igorevich – e-mail: lizin_a@nppame.ru; chief of intellectual image processing center.

Gessen Pavel Alekseevich – e-mail: gessen_pa@nppame.ru; deputy chief of intellectual image processing center.

Saenko Vladislav Dmitrievich – e-mail: eltsova@nppame.ru; junior programmer.

УДК 004.896

DOI 10.18522/2311-3103-2024-2-236-246

И.О. Шепель

МЕТОД ГЕНЕРАЦИИ КАРТЫ ЗАНЯТОСТИ МОБИЛЬНОГО РОБОТА ПО ДАННЫМ МУЛЬТИСПЕКТРАЛЬНОЙ СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ

Рассматривается проблема генерации карт занятости окружающего пространства для робототехнических платформ по данным от мультиспектральной системы технического зрения. Целью работы является качественное улучшение генерируемой модели проходимости путем комлексирования геометрических и семантических данных от лидаров и стереокамер, а также прямых измерений скорости от радаров миллиметрового диапазона. Представленные алгоритмы и их модификации являются универсальными по отношению к источнику данных и не требуют физической синхронизации сенсоров. В исследовании решается задача построения как статических

априорных, так и работающих в реальном масштабе времени динамических карт занятости. Предложен подход по объединению априорной семантической карты с генерируемой в процессе автономного движения робота. Описаны подходы к накоплению и обновлению семантической информации в картах. Также рассматривается задача обнаружения динамических препятствий в картах занятости на основе модифицированного алгоритма фильтрации частиц. Описанный в статье комбинированный метод увеличивает точность определения динамических препятствий и позволяет корректно детектировать препятствие даже в случае ошибки алгоритма определения динамики. Выделены метрики по количественной оценке карт занятости. Разработанный алгоритм был протестирован на открытых датасетах *Semantic KITTI*, *nuScenes* в автомобильном домене данных, а также на малом сервисном роботе-уборщике как в симуляторе *CARLA*, так и в реальных условиях с активным пешеходным движением. Программная реализация алгоритма работает в реальном масштабе времени на встраиваемых вычислителях *Jetson AGX Xavier* и *Jetson AGX Orin*.

Карта занятости; лидар; радар; стереокамера; комплексирование данных; обнаружение препятствий; динамический объект; семантическая карта.

I.O. Shepel

METHOD FOR GENERATING A MOBILE ROBOT OCCUPANCY MAP FROM MULTISPECTRAL VISION SYSTEM DATA

This paper addresses the problem of generating occupancy maps of the surrounding space for robotic platforms using data from a multispectral vision system. The aim of the work is to qualitatively improve the generated passability model by combining geometric and semantic data from lidars and stereo cameras, as well as direct velocity measurements from millimeter-wave radars. The presented algorithms and their modifications are universal to the data source and do not require physical synchronization of sensors. The paper solves the problem of constructing both static a priori and real-time dynamic occupancy maps. An approach for combining an a priori semantic map with the one generated in the runtime is proposed. Approaches for accumulating and updating semantic information in the maps are described. The problem of detecting dynamic obstacles in occupancy maps based on a modified particle filtering algorithm is also considered. The combined method described in the paper increases the accuracy of dynamic obstacle detection and enables correct obstacle detection even if the dynamics detection algorithm fails. Metrics for quantifying occupancy maps are defined. The developed algorithm has been tested on Semantic KITTI, nuScenes open datasets in the automotive data domain, and on a small service cleaning robot both in the CARLA simulator and in real-world conditions with active pedestrian traffic. The software implementation of the algorithm runs in real time on Jetson AGX Xavier and Jetson AGX Orin embedded computers.

Occupancy map; lidar; radar; stereo camera; data complexing; obstacle detection; dynamic object; semantic map.

Введение. В настоящее время активно развиваются различные робототехнические платформы (РТП) от беспилотных автомобилей до малых сервисных роботов. При этом возникает целый ряд нерешенных проблем: необходимо обеспечить безопасность работы в среде с активным движением (в случае с пешеходными зонами это движение имеет хаотичный характер), роботы должны функционировать при разных световых и погодных условиях, а также их стоимость постоянно минимизируется из экономии, что в свою очередь упрощает состав системы технического зрения (СТЗ) и ограничивает мощность бортовых вычислителей. Основными сенсорами, используемыми для определения препятствий в робототехнике являются лидары, камеры (в том числе стерео) и радары. При этом на конечных платформах не всегда присутствуют все три типа сенсоров, поэтому существует необходимость в универсальном по отношению к составу СТЗ способе обнаружения препятствий.

В работе представлен алгоритм построения карты занятости по облакам точек от лидаров, радаров и стереокамер, способный обнаруживать динамические препятствия. Ключевые отличия разработанного алгоритма следующие:

1) универсальность по отношению к составу и типу сенсоров СТЗ, их взаимному расположению и временной синхронизации;

2) устойчивость алгоритма обнаружения динамических препятствий к ложным срабатываниям фильтра частиц;

3) использование семантической информации от стереокамер и лидаров для построения априорной статической карты и семантического слоя накапливаемой карты занятости.

Также в работе представлен новый подход к генерации эталонных карт проходимости для их дальнейшей оценки.

Обзор существующих методов. Карта занятости или карта проходимости [1] – это представление окружающего пространства в виде дискретной решетки, каждой ячейке которой ставится в соответствие вероятность нахождения в ней препятствия. Карты проходимости являются наиболее используемым методом, поскольку позволяют гибко настраивать параметры препятствий в зависимости от используемой платформы [2], а также предоставляют всю необходимую информацию для алгоритмов планирования траектории. Большинство существующих методов предполагает работу с одним источником данных [2, 3], либо данные комплексированы на низком уровне, что накладывает жесткие требования на расположение сенсоров и их временную синхронизацию [4, 5], что для некоторых платформ является невыполнимым требованием. Также с помощью карт занятости можно решать другую принципиально важную задачу – обнаружение динамических объектов [6], что является критически важным для обеспечения безопасного автономного движения в средах с активным движением автомобилей и пешеходов. Для этого используются либо подходы основанные на фильтрации частиц [6, 7], либо с помощью нейросетевых методов [8]. При этом последние как правило сильно переобучаются на домен данных и требуют переобучения для каждого конкретной РТП, а существующие методы с фильтрацией частиц не используют всю доступную информацию от различных сенсоров, а также требовательны к вычислительным ресурсам бортовых вычислителей. Отдельным направлением исследований являются HD-карты и семантические карты. Первые используются при разработке беспилотных автомобилей и являются по сути детализированным векторным описанием известной дорожной инфраструктуры от полос движения до знаков и светофоров [9]. При этом аналогичного подхода, который может быть использован для сервисных роботов, перемещающихся в недетерминированных средах, разделяемых с людьми, не существует из-за разнообразности и сложности решаемых задач. Семантические карты занятости отличаются тем, что в дополнении к геометрическим признакам используют результаты сегментации как для лидаров [10], так и для стереокамер [11], что как и в случае с HD-картами можно использовать в качестве априорной информации. Большинство существующих методов занимаются построением трёхмерной воксельной семантической карты и используют тяжеловесные нейросетевые архитектуры [12, 13], что делает их неприменимыми для бортовых вычислителей малых РТП. При этом отдельно следует выделить работу [14], авторы которой вводят модель для обновления и накопления семантической информации, которая может быть использована и для двумерных карт. Основной проблемой существующих подходов при этом является отсутствие универсального подхода, который сочетал бы возможности объединения как информации о динамических объектах, так и семантических меток и был бы способен работать со всеми возможными сенсорами.

Алгоритм построения накапливаемой карты проходимости. Представленный в статье алгоритм является развитием предыдущих работ автора в области построения карт проходимости и обнаружения препятствий [1, 15, 16]. Общая схема алгоритма представлена на рис. 1. Ниже приведено его краткое описание.

Облака точек от каждого сенсора обрабатываются по отдельности соответствующими компонентами генерации моментальных карт. Точки проецируются в ячейки карты, затем по набору метрик для каждой ячейки независимо определяется является ли она препятствием или нет. В зависимости от задачи эти метрики позволяют выделять препятствия по высоте, по уклону, препятствия-ступени, а также ямы [2, 18]. После этого выполняется трассировка лучей от местоположения сенсора. Все ячейки на лучах до препятствия помечаются как свободные, а после – как неизвестные. Каждой метке соответ-

ствует свое значение вероятности: p_F для свободных, p_U для неизвестных, а для препятствий – p_O , при этом верно неравенство: $p_O > p_U > p_F$. Подробнее этот алгоритм описан автором в статье [1]. Затем все моментальные карты занятости, построенные по одиночным кадрам от сенсоров, объединяются в компоненте построения накапливаемой карты проходимости, который решает ряд задач:

- 1) синхронизация и объединение нескольких моментальных карт;
- 2) накопление и обновление данных в карте;
- 3) обнаружение динамических объектов с помощью фильтра частиц;
- 4) накопление семантической информации;
- 5) добавление в карту априорной информации.

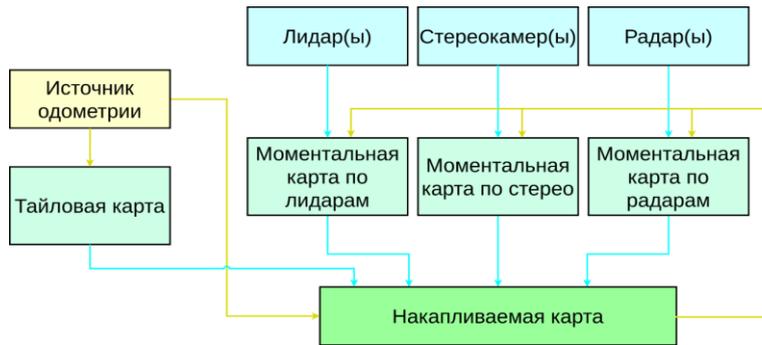


Рис. 1. Общая схема алгоритма

Для синхронизации моментальных карт от разных источников используется следующая схема. Один из сенсоров, как правило самый высокочастотный, выбирается в качестве основного, а остальные – вспомогательными. По приходу моментальной карты от основного источника все пришедшие до этого вспомогательные карты объединяются с основной. При объединении карт выбирается максимальное значение в соответствующих ячейках, т.е. если по облаку точек хотя бы одного из сенсоров зарегистрировано препятствие, то и в объединенной карте эта ячейка также будет препятствием. Необходимо отметить, что такое поведение может быть заменено взвешенным усреднением, где вес рассчитывается исходя из модели ошибок сенсора. Также отметим, что важным вопросом является преобразование данных от источника одометрии в одну общую систему координат. Для этого компонент накапливаемой карты по приходу данных одометрии перемещает свою собственную систему координат, и при поступлении моментальных карт фиксируется её положение. Таким образом учитывается движение РТП во время между регистрацией данных от вспомогательных и основного датчика.

Накопление и обновление происходит с помощью Байесовского фильтра [19] по формуле:

$$L^{i+1} = \begin{cases} \log \frac{p^{i+1}}{1-p^{i+1}}, & d^{i+1} = 1, \\ L^0 = 0, & d^{i+1} = 0 \wedge d^i = 1, \\ L^i + \log \frac{p^{i+1}}{1-p^{i+1}}, & d^{i+1} = 0 \wedge d^i = 0, \end{cases} \quad (1)$$

где L^i – логарифм отношения вероятностей после i -ой итерации, p^i – значение вероятности из объединённой моментальной карты, $d^i = 1$ – условие, характеризующее ячейки с динамическими препятствиями. Способ их обнаружения будет описан ниже.

Обнаружение динамических препятствий. После шагов объединения и накопления полученная накапливаемая карта используется в фильтре частиц для определения динамических препятствий. Можно выделить следующие шаги алгоритма:

- 1) предсказание положения частиц;
- 2) распределение частиц по ячейкам;

- 3) расчёт обновлённой вероятности занятости ячеек;
- 4) обновление весов частиц с предыдущего шага;
- 5) инициализация новых частиц;
- 6) вычисление среднего и дисперсии двухмерной скорости частиц;
- 7) повторная выборка и отсеивание частиц.

Подробно разработанный алгоритм описан автором в работе [16], здесь же приведем основные отличия от базовой версии, представленной в [6]. Во-первых, в базовом алгоритме не учитывается собственное движение РТП на шаге предсказания их положения, что уменьшает точность определения скорости объектов, даже для высокочастотных датчиков и небольшой собственной скорости платформы. В модифицированной версии компенсация производится до первого шага. Во-вторых, оригинальный алгоритм генерирует ложноположительные ошибки для частиц с высокой скоростью. Максимальная начальная скорость частиц является одним из гиперпараметров алгоритма и определяет порог по скорости для обнаружения препятствий. Маленькое значение не позволит найти скоростные препятствия, а высокое приводит к тому, что некоторые частицы «перепрыгивают» большое расстояние с одного статического препятствия на другое, что и приводит к ошибке. В оригинальном алгоритме от ошибок подобного рода избавляются с помощью использования большого числа частиц, что значительно увеличивает нагрузку на графический процессор, а значит не применимы на малопроизводительных вычислителях сервисных роботов. Для фильтрации таких выбросов была разработана следующая эвристика: все ячейки, у которых масса только что рождённых частиц больше порога и среди соседних нет динамических ячеек, помечаются как статические. С такой модификацией возможно существенно уменьшить общее число частиц без потерь точности. Третьей существенной модификацией является добавление алгоритма связанных компонент для более точной аппроксимации границ: при движении объектов их проекции на карту меняются, что также может приводить к ошибкам. Алгоритм связанных компонент позволяет точно определить такие скопления ячеек и полностью пометить их динамическими. Использование алгоритма обнаружения динамических препятствий вместе с правилом обновления вероятности в накапливаемой карте (1) позволяет избавиться от эффекта «хвостов», что повышает общую точность определения препятствий (рис. 2).

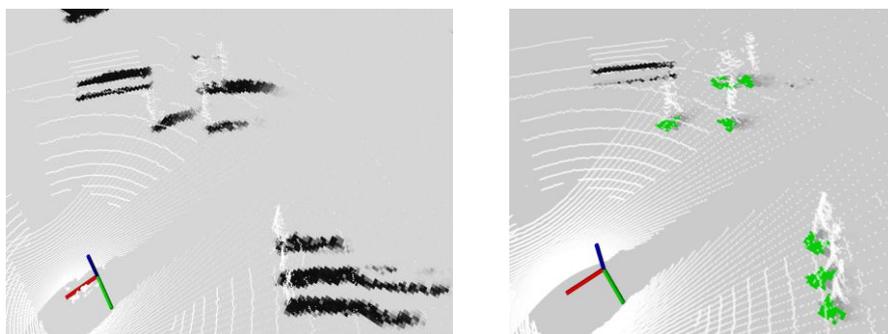


Рис. 2. «Хвосты» за идущими пешеходами на накапливаемой карте (слева), и результат работы алгоритма определения динамических (справа). Черным показаны статические препятствия, а зеленым – динамические

Последняя модификация исходного алгоритма затрагивает шаг инициализации новых частиц. При использовании семантически раскрашенных облаков точек в моментальной карте возможно выделить потенциально динамические препятствия, как это показано в автором в работе [17]. В этом случае засеивание частиц можно производить только в таких ячейках, что также позволяет существенно снизить количество частиц и общую вычислительную нагрузку. Аналогичным образом, в картах построенных по облакам точек от радара можно использовать прямые измерения радиальной скорости объектов, и засеивать только ячейки-препятствия с ненулевой измеренной скоростью.

Априорная и семантическая накапливаемая карта. Другая принципиальная модификация алгоритма – использование семантической информации как дополнительного критерия проходимости ячеек. Кроме того такая информация позволяет разметить зоны, запрещенные к движению для робота, но при этом физически проходимые (например, газоны). В зависимости от ограничений конкретной РТП и поставленной задачи добавление семантической информации в накапливаемую карту может происходить как во время работы робота, так и с помощью априорной карты. Алгоритм ее построения состоит из следующих шагов:

- 1) семантическая сегментация одиночного облака точек;
- 2) распределение точек от последовательных снимков по дискретам общей карты;
- 3) классификация ячеек на свободные и занятые;
- 4) постобработка полученной карты и разбиение её на тайлы.

Поскольку этот алгоритм не выполняется непосредственно во время работы РТП, то на него не накладываются ограничения по быстродействию и возможно использовать тяжёлые и наиболее точные нейросетевые модели для сегментации. Затем в зависимости от итогового класса ячейки принимается решение о её занятости. Отметим, что точки потенциально динамических классов (пешеходов и автомобилей) удаляются из итоговой карты. На этапе постобработки с помощью морфологических операций эрозии и дилатации удаляются одиночные выбросы. Затем карта разбивается на небольшие тайлы, которые могут быть быстро загружены как дополнительный слой в накапливаемую карту занятости при работе на реальных РТП. На рис. 3 показан пример итоговой карты.



Рис. 3. Вид сверху на офисное здание (слева сверху). Накопленное семантическое облако точек (справа сверху), оранжевым показаны точки классов препятствий, фиолетовым – свободных и априорная карта занятости офисного помещения (снизу)

Использование априорной карты возможно только в заранее определенной и подготовленной области, при изменении ландшафта (например строительных работах) модель теряет свою актуальность. Для устранения таких недостатков предлагается алгоритм онлайн построения семантического слоя карты. На первом шаге этого алгоритма по семантическому облаку точек строится моментальная карта проходимости с семантическим слоем [17], а в каждой ячейке дополнительно сохраняется информация о количестве точек каждого представленного класса. Затем в компоненте накапливаемой карты производится обновление классов ячеек с помощью представленной в работе [14] модели сенсора с семантичес-

ским подсчетом. Это модель также использует Байесовский фильтр, только при этом классы определяются с использованием ядра в некоторой окрестности, что позволяет получить более плотную карту при разреженных входных данных и отфильтровать одиночные ошибки сегментации. Исходный алгоритм накопления был модифицирован аналогичным формуле (1) образом, для того чтобы обновление ячеек динамических классов происходило без задержки, вызванной Байесовским накоплением. Для этого статические классы обновляются отдельно от динамических, что позволяет не терять информацию о классе ячейки, накопленную на предыдущих итерациях, даже если на текущей по ней перемещается динамический объект. Результат работы алгоритма представлен на рис. 4.



Рис. 4. Изображение с камеры (сверху), семантический слой моментальной карты проходимости (слева снизу) и накапливаемой (справа снизу).
Разные цвета соответствуют разным классам ячеек

Экспериментальные результаты. Несмотря на длительное использование карт проходимости не выработан подход к их количественной оценке, поскольку параметры для обнаружения препятствий зависят от конкретной задачи. Поэтому для оценки разработанных алгоритмов использовались классические метрики точности P и полноты R [18] (считались по ячейкам-препятствиям), а также корреляцию C , «оценки карт» S и анализ возможных ложноположительных FP и ложноотрицательных FN траекторий, предложенные в статье [20]. Метрики точности, полноты и корреляции оценивают карты как изображения, и чем они выше, тем более качественной является модель. Наоборот, «оценка карт» и анализ траекторий оценивают их с точки зрения дальнейшего применения для алгоритмов планирования, и чем ниже эти показатели, тем лучше для алгоритмов планирования. Для подсчета всех выбранных оценок требуются эталонные карты проходимости, и для их генерации был предложен следующий подход: моментальные карты генерировались только по эталонным семантическим данным без учёта каких-либо геометрических признаков, а по ним в свою очередь строились накапливаемые. Для проведения экспериментов были выбраны следующие открытые наборы данных: SemanticKITTI [21], nuScenes [22] в автомобильном домене, а также в симуляторе CARLA [23] был сгенерирован синтетический датасет с использованием модели реального уличного робота-уборщика. Все эксперименты проводились на картах размером 40 на 40 метров, и разрешением ячеек 0.1 метра. Приведем здесь результаты экспериментов только для накапливаемых карт на синтетических данных (табл. 1 и 2).

Таблица 1

Метрики алгоритма обнаружения динамических препятствий

Тип карты	P , %	R , %	C , %	S	FP , %	FN , %
Базовая реализация	87.85	41.98	73.20	2.84	4.08	16.47
Алгоритм обнаружения динам. преп.	89.13	41.80	73.23	2.94	4.01	16.10

Отметим, что количество ячеек динамических препятствий составляет лишь небольшую часть от общего числа занятых ячеек, поэтому увеличение точности на 1.28%, которое достигается только за счёт удаления эффекта «хвостов», показывает высокую эффективность разработанного метода. Также незначительно улучшаются все остальные метрики кроме полноты.

Таблица 2

Метрики алгоритма построения карты с семантическим слоем

Тип карты	P , %	R , %	C , %	S	FP , %	FN , %
Базовая реализация	49.23	69.94	59.82	2.2231	11.20	15.07
С семантическим слоем	57.31	68.81	61.23	2.3240	7.63	14.44

Для оценки алгоритма построения карт с семантическим слоем эталонные карты генерировались по данным лидара, которые отличаются высокой точностью измерения дальности, а оцениваемые карты генерировались по облакам точек от стереокамеры (использовались методы с открытым исходным кодом для генерации диспаратетов [24] и семантической сегментации [25]). Результаты эксперимента показывают, что добавление семантической информации для определения степени проходимости ячейки значительно повышает точность (8.08%) и корреляцию (1.41%), а также уменьшает количество ложноположительных (3.57%) и ложноотрицательных траектории (0.63%), хотя при этом наблюдается и некоторое снижение оценки S , что говорит о более сильном различии между эталонной и расчётной картой. Подобные результаты наблюдаются и на реальных наборах данных, так для SemanticKITTI повышение точности от использования фильтра частиц для определения динамики составило 0.35%, а для nuScenes – 1.2%, при этом другие метрики значительно не меняются. В свою очередь добавление семантики на SemanticKITTI (в nuScenes отсутствуют стереокамеры) также сильно улучшает метрики: точность (10.16%) и корреляцию (4.38%), а также уменьшает количество ложноположительных и ложноотрицательных траектории (около 1%).

Также на этих датасетах измерялась и производительность предложенного метода. В качестве целевых платформ были выбраны рабочая станция с процессором Intel i7-9700KF и видеокартой NVidia RTX 3070, а также встраиваемый компьютер Nvidia Jetson AGX Orin. Усреднённое суммарное время работы пайплайна (моментальная и накапливаемая карта) без и с использованием семантического слоя для рабочей станции составило 22 и 26 мс, а аналогичные показатели для Jetson Orin – 42 и 47 мс соответственно. Такие метрики быстродействия позволяют обрабатывать данные на частоте около 20 Гц для сервисных роботов с низкопроизводительными вычислителями, и более 30 Гц для беспилотных автомобилей с более мощными блоками, то есть разработанный алгоритм может применяться на подобных РТП.

Выводы и будущая работа. В статье представлены модификации алгоритма карта занятости, в которых решаются проблемы комплексирования данных мультиспектральной СТЗ, обнаружения динамических препятствий с помощью фильтра частиц и добавления дополнительной семантической информации в карту при определении проходимости как в виде оффлайн, так и онлайн решения. Разработанные модификации позволяют генерировать карту занятости по асинхронным сенсорным данным; уменьшают ошибки определения динамических препятствий при использовании малого числа частиц, а также существенно увеличивают метрики получаемых карт при добавлении семантической информации, что подтверждается проведёнными экспериментальными исследованиями. Все описанные выше алгоритмы были апробированы на реальном сервисном роботоподобнике, оснащённом лидарами, стереокамерой и радаром (рис. 5).



Рис. 5. Внешний вид сервисного робота-уборщика

Благодарность. Экспериментальные исследования, в том числе оснащение робототехнической платформы и запись наборов данных, проведены совместно с ООО «ИнтеграНТ» и АО «168роботикс», г. Москва.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Шенель И.О. Модифицированный алгоритм построения карты занятости по облаку точек от нескольких лидаров // Программные продукты и системы. – 2020. – Т. 33, № 2.
2. Fleischmann P., Berns K. A Stereo Vision Based Obstacle Detection System for Agricultural Applications // Field and Service Robotics: Results of the 10th International Conference / под ред. D.S. Wettergreen, T.D. Barfoot. – Cham: Springer International Publishing, 2016. – P. 217-231.
3. Huang J. et al. An online multi-lidar dynamic occupancy mapping method // 2019 IEEE Intelligent Vehicles Symposium (IV). – IEEE, 2019. – P. 517-522.
4. Jaspers H., Himmelsbach M., Wuensche H.J. Multi-modal local terrain maps from vision and lidar // 2017 IEEE Intelligent Vehicles Symposium (IV). – IEEE, 2017. – P. 1119-1125.
5. Cho H. [u dp.]. A multi-sensor fusion system for moving object detection and tracking in urban driving environments // 2014 IEEE International Conference on Robotics and Automation (ICRA). – IEEE, 2014. – P. 1836-1843.
6. Nuss D. et al. A random finite set approach for dynamic occupancy grid maps with real-time application // The International Journal of Robotics Research. – 2018. – Vol. 37, No. 8. – P. 841-866.
7. Danescu R., Nedevschi S. A Particle-Based Solution for Modeling and Tracking Dynamic Digital Elevation Maps // IEEE Transactions on Intelligent Transportation Systems. – 2014. – Vol. 3, No. 15. – P. 1002-1015.
8. Mohajerin N., Rohani M. Multi-step prediction of occupancy grid maps with recurrent neural networks // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2019. – P. 10600-10608.
9. Liu R., Wang J., Zhang B. High definition map for automated driving: Overview and analysis // The Journal of Navigation. – 2020. – Vol. 73, No. 2. – P. 324-341.
10. Nikdel P., Chen M., Vaughan R. Recognizing and Tracking High-Level, Human-Meaningful Navigation Features of Occupancy Grid Maps // 2020 17th Conference on Computer and Robot Vision (CRV). – IEEE, 2020. – P. 62-69.
11. Roddick T., Cipolla R. Predicting semantic map representations from images using pyramid occupancy networks // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2020. – P. 11138-11147.
12. Li Y. [u dp.]. Voxformer: Sparse voxel transformer for camera-based 3d semantic scene completion // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2023. – P. 9087-9098.
13. Zhang Y., Zhu Z., Du D. Occformer: Dual-path transformer for vision-based 3d semantic occupancy prediction // Proceedings of the IEEE/CVF International Conference on Computer Vision. – 2023. – P. 9433-9443.
14. Gan L. [u dp.]. Bayesian spatial kernel smoothing for scalable dense semantic mapping // IEEE Robotics and Automation Letters. – 2020. – Vol. 5, No. 2. – P. 790-797.
15. Шенель И.О. Алгоритм комплексирования нескольких источников данных в общую карту занятости // Известия ЮФУ. Технические науки. – 2021. – № 3 (220). – С. 64-71.

16. Шепель И.О. Обнаружение динамических объектов на карте занятости с накоплением на основе фильтра частиц // Известия ЮФУ. Технические науки. – 2022. – № 3. – С. 192-201.
17. Shepel I. [u ðp.]. Occupancy Grid Generation With Dynamic Obstacle Segmentation in Stereo Images // IEEE Transactions on Intelligent Transportation Systems. – 2022. – Vol. 23, No. 9. – P. 14779-14789.
18. Шепель И.О. Построение модели проходимости окружающей среды по облаку точек стереокамеры с использованием иерархической карты высот // Инженерный вестник Дона. – 2018. – 1 (48).
19. Thrun S. Learning occupancy grid maps with forward sensor models // Autonomous robots. – 2003. – Vol. 15, No. 2. – P. 111-127.
20. Collins T., Collins J., Ryan D. Occupancy grid mapping: An empirical evaluation // 2007 mediterranean conference on control & automation. – IEEE, 2007. – P. 1-6.
21. Behley J. [u ðp.]. SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences // Proc. of the IEEE International Conf. on Computer Vision (ICCV). – 2019.
22. Caesar H. [u ðp.]. nuScenes: A Multimodal Dataset for Autonomous Driving // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 06.2020.
23. Dosovitskiy A. [u ðp.]. CARLA: An Open Urban Driving Simulator // Proceedings of the 1st Annual Conference on Robot Learning. – 2017. – P. 1-16.
24. Mihelich P., Konolige K., Leibs J. stereo_image_proc / Open Source Robotics Foundation. – 2024. – URL: https://github.com/ros-perception/image_pipeline/tree/rolling/stereo_image_proc.
25. Zhu Y. [u ðp.]. Improving semantic segmentation via video propagation and label relaxation // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2019. – P. 8856-8865.

REFERENCES

1. Shepel' I.O. Modifitsirovanny algorithm postroeniya karty zanyatosti po oblaku toчек ot neskolk'kikh lidarov [Modified algorithm for constructing an occupancy map using a point cloud from several lidars], *Programmnyye produkty i sistemy* [Software products and systems], 2020, Vol. 33, No. 2.
2. Fleischmann P., Berns K. A Stereo Vision Based Obstacle Detection System for Agricultural Applications, *Field and Service Robotics: Results of the 10th International Conference*, ed. by D.S. Wettergreen, T.D. Barfoot. Cham: Springer International Publishing, 2016, pp. 217-231.
3. Huang J. et al. An online multi-lidar dynamic occupancy mapping method, *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2019, pp. 517-522.
4. Jaspers H., Himmelsbach M., Wuensche H.J. Multi-modal local terrain maps from vision and lidar, *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2017, pp. 1119-1125.
5. Cho H. [et al.]. A multi-sensor fusion system for moving object detection and tracking in urban driving environments, *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 1836-1843.
6. Nuss D. et al. A random finite set approach for dynamic occupancy grid maps with real-time application, *The International Journal of Robotics Research*, 2018, Vol. 37, No. 8, pp. 841-866.
7. Danescu R., Nedeveschi S. A Particle-Based Solution for Modeling and Tracking Dynamic Digital Elevation Maps, *IEEE Transactions on Intelligent Transportation Systems*, 2014, Vol. 3, No. 15, pp. 1002-1015.
8. Mohajerin N., Rohani M. Multi-step prediction of occupancy grid maps with recurrent neural networks, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 10600-10608.
9. Liu R., Wang J., Zhang B. High definition map for automated driving: Overview and analysis, *The Journal of Navigation*, 2020, Vol. 73, No. 2, pp. 324-341.
10. Nikdel P., Chen M., Vaughan R. Recognizing and Tracking High-Level, Human-Meaningful Navigation Features of Occupancy Grid Maps, *2020 17th Conference on Computer and Robot Vision (CRV)*. IEEE, 2020, pp. 62-69.
11. Roddick T., Cipolla R. Predicting semantic map representations from images using pyramid occupancy networks, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 11138-11147.
12. Li Y. [et al.]. Voxformer: Sparse voxel transformer for camera-based 3d semantic scene completion, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 9087-9098.
13. Zhang Y., Zhu Z., Du D. Occformer: Dual-path transformer for vision-based 3d semantic occupancy prediction, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 9433-9443.
14. Gan L. [et al.]. Bayesian spatial kernel smoothing for scalable dense semantic mapping, *IEEE Robotics and Automation Letters*, 2020, Vol. 5, No. 2, pp. 790-797.

15. Shepel' I.O. Algoritm kompleksirovaniya neskol'kikh istochnikov dannykh v obshchuyu kartu zanyatosti [Algorithm for integrating several data sources into a common employment map], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2021, No. 3 (220), pp. 64-71.
16. Shepel' I.O. Obnaruzhenie dinamicheskikh ob"ektov na karte zanyatosti s nakopleniem na osnove filtra chastits [Detection of dynamic objects on an occupancy map with accumulation based on a particle filter], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2022, No. 3, pp. 192-201.
17. Shepel I. [et al.]. Occupancy Grid Generation With Dynamic Obstacle Segmentation in Stereo Images, *IEEE Transactions on Intelligent Transportation Systems*, 2022, Vol. 23, No. 9, pp. 14779-14789.
18. Shepel' I.O. Postroenie modeli prokhodimosti okruzhayushchey sredy po oblaku tochek stereokamery s ispolzovaniem ierarkhicheskoy karty vysot [Construction of a model of environmental cross-country ability using a point cloud of a stereo camera using a hierarchical height map], *Inzhenernyy vestnik Dona* [Engineering Bulletin of the Don], 2018, 1 (48).
19. Thrun S. Learning occupancy grid maps with forward sensor models, *Autonomous robots*, 2003, Vol. 15, No. 2, pp. 111-127.
20. Collins T., Collins J., Ryan D. Occupancy grid mapping: An empirical evaluation, *2007 mediterranean conference on control & automation*. IEEE, 2007, pp. 1-6.
21. Behley J. [et al.]. SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences, *Proc. of the IEEE International Conf. on Computer Vision (ICCV)*, 2019.
22. Caesar H. [et al.]. nuScenes: A Multimodal Dataset for Autonomous Driving, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 06.2020.
23. Dosovitskiy A. [et al.]. CARLA: An Open Urban Driving Simulator, *Proceedings of the 1st Annual Conference on Robot Learning*, 2017, pp. 1-16.
24. Mihelich P., Konolige K., Leibs J. stereo_image_proc / Open Source Robotics Foundation, 2024. Available at: https://github.com/ros-perception/image_pipeline/tree/rolling/stereo_image_proc.
25. Zhu Y. [et al.]. Improving semantic segmentation via video propagation and label relaxation, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 8856-8865.

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

Шепель Илья Олегович – Южный федеральный университет; e-mail: chepsiilya@gmail.com; г. Москва, Россия; тел.: +79085180350; кафедра математического обеспечения и применения ЭВМ; соискатель.

Shepel Ilya Olegovich – Southern Federal University; e-mail: chepsiilya@gmail.com; Moscow, Russia; phone: +79085180350; the department of mathematical support and computer applications; post-graduate.

УДК 629.3.051

DOI 10.18522/2311-3103-2024-2-246-253

И.А. Шипов, Е.В. Ветошкин

ПОДХОДЫ К МОДУЛЬНОМУ ПРОЕКТИРОВАНИЮ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ НАВИГАЦИОННЫХ СИСТЕМ НАЗЕМНЫХ ПОДВИЖНЫХ ОБЪЕКТОВ

Цель исследования – сформировать и обобщить подходы, применяемые при проектировании программного обеспечения навигационных систем наземных подвижных объектов. В статье описан опыт разработки программного обеспечения навигационных систем различных наземных объектов. В зависимости от типа шасси, характера решаемых функциональных задач, требований к эргономике, степени автономности и ряда других факторов будет определяться сложность принимаемых программно-алгоритмических решений. Алгоритмы функционирования системы навигации со всеми входящими в нее компонентами, как правило, не являются требовательными к вычислительным мощностям, за исключением реализации работы с цифровыми картами местности. В связи с этим, не смотря на значительное отставание отечественной элементной базы от зарубежных аналогов, формирование комплексированного навигационного решения может быть реализовано на ее базе. Еще одной особенностью проектирования программного обеспечения навигационных систем является необходимость функционирования их в режиме реального времени. Это необходимо для синхронной обработки, поступающей от различных источников первичной информации, и обеспечения формирования потребителю навигационного решения со стабильной

частотой. Модульный подход проектирования программного обеспечения построен на принципах унификации внутренних функций навигационной системы и переносимости полученных ранее решений между проектами. При таком подходе каждый элемент выполнен в форме самостоятельного модуля, который представляет законченную подпрограмму с набором входных и выходных параметров. Реализация взаимодействия между модулями зависит от типа его исполнения и может быть выполнена как на уровне встраивания исходного кода, так и на уровне обмена параметрами через каналы информационного взаимодействия. Применение модульного подхода при проектировании программного обеспечения навигационных систем наземных объектов позволяют за короткий период создавать полнофункциональные технические решения, обеспечивающие потребности конечного пользователя. При этом снижается риск последующих конструктивных ошибок и доработок в связи с увеличенным объемом проверок и накопленным опытом применения разработанного ранее алгоритма.

Наземная навигация; навигационная система; программное обеспечение; математическая модель; техническая реализация.

I.A. Shipov, E.V. Vetoshkin

APPROACHES TO MODULAR DESIGN OF SOFTWARE FOR LANDINGS NAVIGATION SYSTEMS

The purpose of the study is to formulate and generalize the approaches used in the design of software for navigation systems of ground-based moving objects. The article describes the experience of developing software for navigation systems of various ground objects. Depending on the type of chassis, the nature of the functional tasks being solved, ergonomic requirements, degree of autonomy and a number of other factors, the complexity of the software and algorithmic solutions used will be determined. Algorithms for the operation of a navigation system with all its components, as a rule, are not demanding on computing power, with the exception of the implementation of work with digital terrain maps. In this regard, despite the significant lag of the domestic element base from foreign analogues, the formation of an integrated navigation solution can be implemented on its basis. Another feature of the design of software for navigation systems is the need for them to function in real time. This is necessary for synchronous processing coming from various sources of primary information and ensuring the formation of a navigation solution for the consumer with a stable frequency. The modular approach to software design is built on the principles of unifying the internal functions of the navigation system and the portability of previously obtained solutions between projects. With this approach, each element is made in the form of an independent module, which is a complete subroutine with a set of input and output parameters. The implementation of interaction between modules depends on the type of its execution and can be performed both at the level of embedding source code and at the level of exchanging parameters through information interaction channels. The use of a modular approach to the design of software for navigation systems of ground objects allows us to create fully functional technical solutions that meet the needs of the end user in a short period of time. At the same time, the risk of subsequent design errors and modifications is reduced due to the increased volume of checks and the accumulated experience in using the previously developed algorithm.

Land navigation; navigation system; software; mathematical model; technical implementation.

Введение. Разработка программно-алгоритмического обеспечения является ключевым этапом создания всех современных навигационных систем (НС) наземных подвижных объектов. В зависимости от типа выбранной аппаратной платформы, набора решаемых функциональных задач, сложности алгоритмов, количества внутренних компонентов и других факторов будет определяться сложность программных решений, применяемых при проектировании. При этом аппаратная часть вычислителей современных НС может быть построена на базе практически любой из существующих архитектур (x86, ARM и другие) [1].

В свою очередь сложность программных решений возрастает вместе со сложностью структуры и расширением набора функциональных задач навигационной системы.

Основное противоречие развития направления алгоритмов функционирования современных НС заключается в том, что для достижения максимально возможных выходных технических характеристик необходима комплексная обработка данных от всех имеющихся источников первичной информации. При этом оптимальная обработка сигналов может быть реализована только на уровне каждого элемента системы.

Цель данного исследования – систематизация подходов проектирования программно-алгоритмического обеспечения навигационных систем наземных объектов на основе модульного принципа. Для достижения цели исследования были решены задачи:

- ◆ сформулирована концепция модульной структуры программного обеспечения комплексированной системы навигации;
- ◆ сформированы подходы к проектированию взаимодействия между программными модулями;
- ◆ сформирован банк программных решений, позволяющий ускорить создание навигационных систем наземных объектов.

Описанные подходы являются результатом обобщения опыта, полученного при проектировании средств навигации для наземных подвижных объектов различного типа и назначения.

Роль программно-алгоритмического обеспечения. Современные навигационные системы являются сложными техническими системами. Они состоят из нескольких компонентов, взаимодействие между которыми реализовано аппаратно-программными средствами [2]. Ключевая роль программного обеспечения обусловлена тем, что преобладающее количество функций НС реализовано именно на программном уровне. При этом кроме реализации самих компонентов важную роль играет тип взаимосвязей между ними и интерфейс взаимодействия с пользователем.

Для разработки систем навигации и ориентирования нового поколения [3–6] требуются программно-алгоритмические решения, которые кроме выполнения основных функций позволяют:

- ◆ ускорить и упростить отладку алгоритмов функционирования;
- ◆ максимально глубоко анализировать функционирование системы на протяжении всего жизненного цикла;
- ◆ проводить исследования и наработки с целью развития и модернизации имеющихся систем;
- ◆ снизить вероятность ошибок в программной части конечного изделия.

Критическая зависимость навигационных систем от программно-алгоритмических решений не только дает возможность исключения ряда компонентов и подсистем и их последующей программной реализации, но и задает высокие требования к эффективности применяемых алгоритмов. Методика разработки НС по «классической» структуре, при которой узконаправленные группы разработчиков создают отдельные блоки, а затем ведущий интегратор строит из них конечную навигационную систему, может не обеспечить эффективный результат. Разработка современных навигационных систем предъявляет особые требования к уровню вовлеченности всех участников разработки и делает ключевой роль взаимодействия их между собой. Происходит размытие зон ответственности отдельных разработчиков, а результатом их интеллектуальной деятельности становится не отдельный блок или подсистема, а система навигации в целом. При должном уровне эффективности управления такой многофункциональной командой существенно повышается заинтересованность всех участников проекта в успешности завершения разработки в целом и как следствие качество конечной продукции [7].

Модульный подход создания программного обеспечения позволяет создавать программное обеспечение навигационной системы за короткий период за счет использования набора отлаженных ранее решений, программных библиотек и функций.

Проектирование программно-алгоритмического обеспечения. Под программным обеспечением навигационной системы, как правило, понимают совокупность алгоритмов функционирования всех ее составных частей. Типовая внутренняя структура НС изображена на рис. 1 и состоит из нескольких элементов.

При такой структуре НС определяющим для потребителя является «вычислитель навигационной информации», в котором будет также реализован интерфейс внешнего взаимодействия и формироваться пакет навигационных данных.

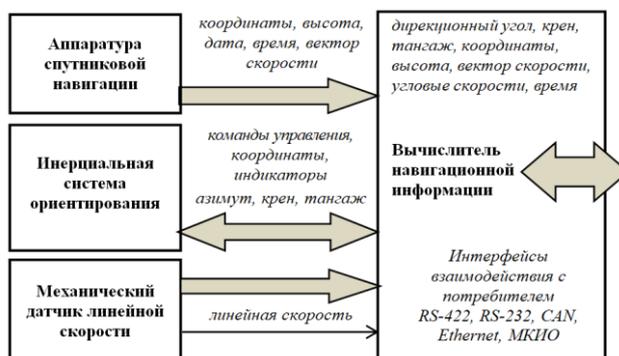


Рис. 1. Типовая структура навигационной системы

К ключевым функциям типовой НС в общем виде можно отнести:

- ◆ решение задачи ориентирования;
- ◆ определение местоположения объекта в автономном режиме;
- ◆ определение местоположения объекта в комплексированном режиме;
- ◆ информационное взаимодействие с элементами системы;
- ◆ диспетчеризация внутренних потоков и подпрограмм;
- ◆ взаимодействие с потребителем;

Все обозначенные функции реализовываются в виде программных модулей с возможностью сопряжения друг с другом:

- ◆ модуль функционирования навигационной системы;
- ◆ модуль аппаратуры спутниковой навигации;
- ◆ модуль взаимодействия с одометрическими датчиками;
- ◆ модуль блока чувствительных элементов инерциальной системы;
- ◆ модуль функционирования инерциальной системы ориентирования;
- ◆ модуль взаимодействия с потребителем навигационной информации.

На рис. 2 приведен типовой алгоритм формирования пакета навигационной информации для потребителя в современной системе навигации. Также дополнительно может быть реализована фильтрация, в том числе на основе фильтра Калмана [8, 9].

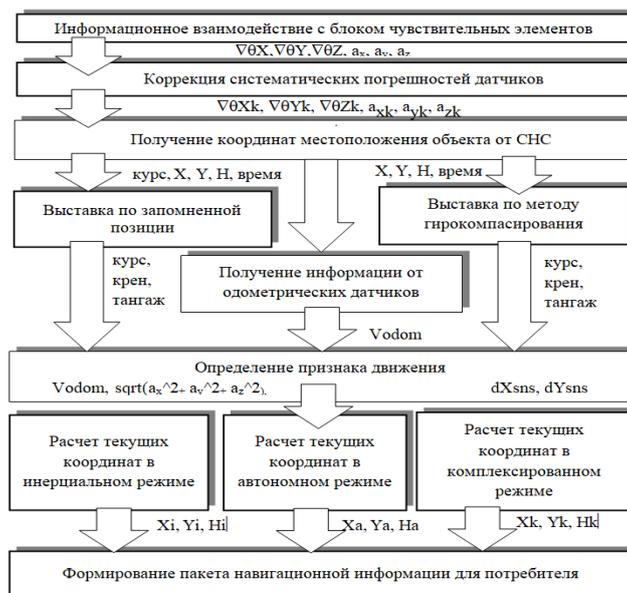


Рис. 2. Типовая структура навигационной системы

Применение современных электронных компонентов при построении вычислительных систем позволяет обеспечить высокий уровень унификации программных решений [10–12]. Типовое программно-алгоритмическое обеспечение функционирования навигационной системы условно состоит из следующих основных составных частей:

- ◆ модуль взаимодействия с сопряженными устройствами;
- ◆ модуль первичной обработка поступающей информации;
- ◆ блок решения навигационных задач;
- ◆ модуль мониторинга технического состояния и самодиагностики;
- ◆ модуль работы с цифровой картой местности;
- ◆ модуль решения навигационных и сервисных задач;
- ◆ модуль взаимодействия с FLASH памятью.

Одним из наиболее распространённых языков программирования, применяемых при создании современных средств автоматизации, является C++. При использовании одного языка программирования при разработке модулей повышается переносимость кода между элементами, однако для обеспечения встраиваемости также может быть реализован механизм сборки программных решений методом библиотек.

При модульном проектировании программного обеспечения одним из ключевых требований является обеспечение возможности встраивания всех модулей в контур будущей структуры НС без дополнительных доработок. Таким образом, должно быть реализовано программное и аппаратное соответствие программных протоколов и аппаратных интерфейсов информационного взаимодействия всех элементов.

Рассмотрим такой подход на примере модуля программного инерциальной системы ориентирования. Он может быть реализован в форме унифицированного алгоритма определения углов ориентации объекта на основе первичных данных от двух триад гироскопических и акселерометрических датчиков. Еще одним параметром необходимым для обеспечения функционирования модуля является время получения показаний датчиков. Выходные параметры могут иметь сокращенный вид и представлять себя значения азимута, крена и тангажа или расширенный включающий в себя другие представления угловой ориентации, такие как например кватернион поворота.

На рис. 3 представлена обобщенная схема [13–15] взаимодействия модулей навигационной системы наземного подвижного объекта.

При этом функционирование модулей может осуществляться как в рамках одного адресного пространства вычислителя, так и быть выполнено в виде нескольких приложений или выполняться на разных аппаратных платформах.

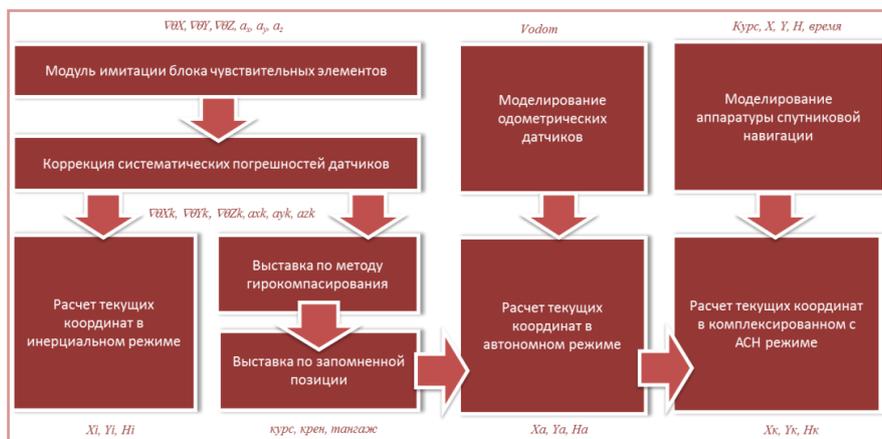


Рис. 3. Схема взаимодействия программно-алгоритмических модулей

Описанный в статье [17] подход показывает пример распределения программных функций внутри одного вычислителя на базе нескольких процессоров. Аналогичным образом может быть выполнено разделение на базе ядер одного процессора, потоков одного программного модуля, подпрограмм и даже программ.

Формирование навигационного решения для потребителя происходит на основе набора исходных данных [18], которые формируются несколькими источниками (инерциальная система, аппаратура спутниковой навигации, датчики скорости и другие) в разные моменты времени. Современные навигационные системы функционируют в режиме реального времени, в связи с чем съем информации происходит с высокой частотой, а обработка данных происходит синхронно.

Применение описанного модульного принципа к построению программно-алгоритмического обеспечения, за счет оптимального сочетания элементов и реализации взаимосвязей между ними, позволяет обеспечить наилучшие выходные тактико-технические характеристики создаваемой системы в целом.

Опыт проектирования современных НС [19, 20] показал возможность реализации полного функционала на базе встраиваемых 32 разрядных микроконтроллеров. При этом успешно были реализованы режимы работы, алгоритм расчета углов ориентации, фильтрация первичной информации, взаимодействие с потребителем по нескольким каналам обмена. Дополнительные мощности и высокопроизводительная аппаратная часть необходима в случае решения задач электронной картографии и развернутого графического интерфейса взаимодействия с пользователем. Таким НС включают в своем составе, как правило, панельный компьютер на базе операционной системы семейства Linux.

Заключение. Учитывая сложность структуры и специфику построения современных навигационных систем наземных объектов, их создание необходимо вести с применением эффективных подходов к проектированию. Разработка программно-алгоритмического обеспечения на базе модульного принципа позволяет снизить трудозатраты при создании новых систем и уменьшить риски конструктивных ошибок. При этом должны быть обеспечены:

- ◆ адаптивность разработанных ранее модулей и возможность их быстрого встраивания в структуру создаваемой НС;
- ◆ возможность масштабируемости структуры программно-алгоритмического обеспечения навигационной системы;
- ◆ накопление опыта использования алгоритмических решений в нескольких сценариях применения.

Создание же каждого модуля необходимо вести с учетом нескольких основных принципов, а именно:

- ◆ унификации внешних интерфейсов взаимодействия;
- ◆ переносимость кода между различными средами программирования;
- ◆ адаптивность и встраиваемость внутренней структуры;
- ◆ наличие изменяемых входных и выходных параметров.

Описанные подходы к проектированию программного обеспечения навигационных систем наземных объектов позволяют за короткий период создавать полнофункциональные технические решения обеспечивающие потребности конечного потребителя.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Белов А.В. Конструирование устройств на микроконтроллерах. – СПб.: Наука и Техника, 2005. – 256 с.
2. Алямов А.Э., Баласов И.Ю., Бажанов В.А. Импортзамещение электронной компонентной базы в оборонном производстве // Всероссийский экономический журнал ЭКО. – 2015. – № 11 (497). – С. 19-29.
3. Шипов И.А., Ветошкин Е.В. Комплексированная навигация наземных робототехнических комплексов // Робототехника и техническая кибернетика. – 2021. – Т. 9, № 2 (225). – С. 127-132.
4. Брозгуль Л.И., Зайцев А.В. Состояние и перспективы развития инерциальных навигационных систем // Мехатроника, автоматизация, управление. – 2006. – № 3.

5. Шипов И.А., Ветошкин Е.В., Морозов А.В. Интегрированные инерциально-спутниковые системы наземных робототехнических комплексов // Сб. трудов XXXII конференции памяти Н.Н. Острякова в направлении «Гироскопические и интегрированные инерциально-спутниковые системы».
6. Stallings W. Operating Systems: Internals and Design Principles. – 7th ed. – Prentice Hall, 2011.
7. Systems Engineering Handbook. Version 2a. INCOSE, 2004.
8. Васильев К.К., Аникин А.А. Калмановское комплексирование и моделирование навигационных систем // Электронная техника: Межвузовский сборник научных трудов / под ред. Д.В. Андреева. – Ульяновск: УлГТУ, 2005.
9. Grewal M., Henderson V., Miyasako R. Application of Kalman filtering to the calibration and alignment of inertial navigation systems // IEEE Transactions, Automatic Control. – 1991. – Vol. 36.
10. Vincent Mahout. Assembly Language Programming: Arm Cortex-M3, 2011.
11. Белочкин П.Е., Кацай Д.А. Особенности моделирования бесплатформенной системы ориентации по уравнениям Эйлера в среде Matlab и Mathcad // Современные наукоемкие технологии. – 2014. – № 5. – С. 18-20.
12. Жидкова Н.В., Волков В.Л. Моделирование бесплатформенной системы ориентации // Современные проблемы науки и образования. – 2015. – № 1. – URL: <http://www.science-education.ru/121-17099>.
13. Tel G. Introduction to Distributed Algorithms. Second edition. – Cambridge University Press, 2000.
14. Tranquilla J.M., Cam J.P., Al-Rizzo H.M. Analysis of a Choke Ring Groundplane for Multipath Control in Global Positioning System (GPS) Applications // IEEE Transactions on Antennas and Propagation. – 1994. – Vol. 42, No. 7. – P. 905-911. – DOI: <https://doi.org/10.1109/8.299591>.
15. Матвеев В.В., Распопов В.Я. Основы построения бесплатформенных инерциальных навигационных систем / под ред. В.Я. Распопова. – СПб.: ГНЦ РФ ОАО «Концерн «ЦНИИ «Электроприбор», 2009. – 280 с.
16. Volkov V.L. Mathematical modeling of inertial measurement systems // International Journal of Applied and Fundamental Research. – 2015. – No. 2. – URL: www.science-sd.com/457-24631.
17. Шипов И.А. Реализация распределенных вычислений на отечественных микропроцессорных устройствах // Известия ЮФУ. Технические науки. – 2022. – № 1 (225). – С. 218-226.
18. Woodman O.J. An introduction to inertial navigation // Technical reports published by the University of Cambridge Computer Laboratory are freely available via the Internet: <http://www.cl.cam.ac.uk/techreports/> 2007. – 37 p.
19. Bilich A., Larson K. M., Axelrad P. SNR-Based Multipath Corrections to GPS Phase Measurements: Improving the Accuracy of Permanent GPS Stations // AGU Fall Meeting Abstracts. – 2002. – Vol. 1. – P. 06.
20. Филиппов С.И., Матахин В.В., Шипов И.А., Петров А.В. Применение методик и инструментов системной инженерии при создании систем наземной навигации // Оборонная техника. – 2017. – № 7-8. – С. 38-44.

REFERENCES

1. Belov A.V. Konstruirovaniye ustroystv na mikrokontrollerakh [Design of devices on microcontrollers]. St. Petersburg: Nauka i Tekhnika, 2005, 256 p.
2. Alyamov A.E., Balasov I.Yu., Bazhanov V.A. Importozameshchenie elektronnoy komponentnoy bazy v oboronnom proizvodstve [Import substitution of electronic components in defense production], *Vserossiyskiy ekonomicheskiy zhurnal EKO* [All-Russian Economic Journal ECO], 2015, No. 11 (497), pp. 19-29.
3. Shipov I.A., Vetoshkin E.V. Kompleksirovannaya navigatsiya nazemnykh robototekhnicheskikh kompleksov [Integrated navigation of ground robotic systems], *Robototekhnika i tekhnicheskaya kibernetika* [Robotics and technical cybernetics], 2021, Vol. 9, No. 2 (225), pp. 127-132.
4. Brozgul' L.I., Zaytsev A.V. Sostoyaniye i perspektivy razvitiya inertsial'nykh navigatsionnykh sistem [State and prospects for the development of inertial navigation systems], *Mekhatronika, avtomatizatsiya, upravlenie* [Mechatronics, automation, control], 2006, No. 3.
5. Shipov I.A., Vetoshkin E.V., Morozov A.V. Integrirovannyye inertsial'no-sputnikovyye sistemy nazemnykh robototekhnicheskikh kompleksov [Integrated inertial-satellite systems of ground-based robotic complexes], *Sb. trudov XXXII konferentsii pamyati N.N. Ostryakova v napravlenii «Giroskopicheskie i integrirovannyye inertsial'no-sputnikovyye sistemy»* [Collection of proceedings of the XXXII conference in memory of N.N. Ostryakov in the direction of “Gyroscopic and integrated inertial-satellite systems”].
6. Stallings W. Operating Systems: Internals and Design Principles. 7th ed. Prentice Hall, 2011.

7. Systems Engineering Handbook. Version 2a. INCOSE, 2004.
8. *Vasil'ev K.K., Anikin A.A.* Kalmanovskoe kompleksirovanie i modelirovanie navigatsionnykh sistem [Kalman integration and modeling of navigation systems], *Elektronnaya tekhnika: Mezhvuzovskiy sbornik nauchnykh trudov* [Electronic technology: Interuniversity collection of scientific papers], pod red. D.V. Andreeva. Ul'yanovsk: UIGTU, 2005.
9. *Grewal M., Henderson V., Miyasako R.* Application of Kalman filtering to the calibration and alignment of inertial navigation systems, *IEEE Transactions, Automatic Control*, 1991, Vol. 36.
10. *Vincent Mahout.* Assembly Language Programming: Arm Cortex-M3, 2011.
11. *Belochkin P.E., Katsay D.A.* Osobennosti modelirovaniya besplatformennoy sistemy orientatsii po uravneniyam Eylera v srede Matlab i Mathcad [Features of modeling a strapdown orientation system using Euler's equations in Matlab and Mathcad], *Sovremennye naukoemkie tekhnologii* [Modern high technology], 2014, No. 5, pp. 18-20.
12. *Zhidkova N.V., Volkov V.L.* Modelirovanie besplatformennoy sistemy orientatsii [Simulation of a strapdown orientation system], *Sovremennye problemy nauki i obrazovaniya* [Modern problems of science and education], 2015, No. 1. Available at: <http://www.science-education.ru/121-17099>.
13. *Tel G.* Introduction to Distributed Algorithms. Second edition. Cambridge University Press, 2000.
14. *Tranquilla J.M., Cam J.P., Al-Rizzo H.M.* Analysis of a Choke Ring Groundplane for Multipath Control in Global Positioning System (GPS) Applications, *IEEE Transactions on Antennas and Propagation*, 1994, Vol. 42, No. 7, pp. 905-911. DOI: <https://doi.org/10.1109/8.299591>.
15. *Matveev V.V., Raspopov V.Ya.* Osnovy postroeniya besplatformennykh inertsial'nykh navigatsionnykh sistem [Fundamentals of constructing strapdown inertial navigation systems], ed. by V.Ya. Raspopova. St. Petersburg: GNTS RF OAO «Kontsem «TSNII «Elektropribor», 2009, 280 p.
16. *Volkov V.L.* Mathematical modeling of inertial measurement systems, *International Journal of Applied and Fundamental Research*, 2015, No. 2. Available at: www.science-sd.com/457-24631.
17. *Shipov I.A.* Realizatsiya raspredelennykh vychisleniy na otechestvennykh mikroprotsessornykh ustroystvakh [Implementation of distributed computing on domestic microprocessor devices], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2022, No. 1 (225), pp. 218-226.
18. *Woodman O.J.* An introduction to inertial navigation [An introduction to inertial navigation], *Technical reports published by the University of Cambridge Computer Laboratory are freely available via the Internet: http://www.cl.cam.ac.uk/techreports/ 2007* [Technical reports published by the University of Cambridge Computer Laboratory are freely available via the Internet: <http://www.cl.cam.ac.uk/techreports/ 2007>], 37 p.
19. *Bilich A., Larson K. M., Axelrad P.* SNR-Based Multipath Corrections to GPS Phase Measurements: Improving the Accuracy of Permanent GPS Stations, *AGU Fall Meeting Abstracts*, 2002, Vol. 1, pp. 06.
20. *Filippov S.I., Matakhin V.V., Shipov I.A., Petrov A.V.* Primenenie metodik i instrumentov sistemnoy inzhenerii pri sozdanii sistem nazemnoy navigatsii [Application of systems engineering methods and tools in the creation of land navigation systems], *Oboronnaya tekhnika* [Defense technology], 2017, No. 7-8, pp. 38-44.

Статью рекомендовал к опубликованию д.т.н. А.И. Новиков.

Шипов Илья Александрович – АО «Всероссийский научно-исследовательский институт «Сигнал»; e-mail: pil14@inbox.ru; г. Ковров, Россия; тел.: +79100905025; заместитель начальника отдела.

Ветошкин Евгений Владимирович – e-mail: vetoshkin@vniisignal.ru; тел.: +79206244967; главный конструктор средств навигационно-геодезического обеспечения.

Shipov Ilya Aleksandrovich – All-Russian Scientific Research Institute “Signal” JSC; e-mail: pil14@inbox.ru; Kovrov, Russia; phone: +79100905025; deputy head of department.

Vetoshkin Evgeniy Vladimirovich – e-mail: vetoshkin@vniisignal.ru; phone: +79206244967; chief designer of navigation and geodetic systems.

Е.А. Титенко, И.Е. Чернецкая, М.А. Титенко, Э.В. Мельник, Д.А. Трокоз

РАСШИРЕННАЯ ПРОДУКЦИОННАЯ МАШИНА ВЫВОДА ДЛЯ РЕАЛИЗАЦИИ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ

Актуальность. В работе развивается теоретический подход организации параллельных вычислений на основе производственной модели управления потоком данных. Производственная парадигма параллельных вычислений имеет необходимые условия построения новых архитектур и организации высокопроизводительных параллельных вычислений. Рассматриваются производственные системы, управляющие наборами левых частей производств (образцами). **Цель** – повышение эффективности параллельного вывода решений за счет сокращения непродуктивных затрат времени на перебор возможных альтернатив в пространстве графа вывода. Метод решения основан на создании расширенной машины символического вывода для реализации параллельных вычислений. Машина символического вывода – это абстрактная система, систематизирующая производственный вывод как последовательность четырех вычислительно-поисковых этапов. Машина вывода задает общий вид однородной вычислительной системы. Главное отличие – декомпозиция базы производственных правил на отдельные подмножества на основе алгебры производств и структуризации отношений между производствами. Вместо единой «плоской» структуры предлагается базу производств декомпозировать на части – ввести систему независимых подмножеств производств. Параллельный вывод реализуется по отдельным подмножествам без потери общности, при этом перебор возможных альтернатив является сокращенным. Каждое подмножество производств имеет специальное слово-маркер, по значению которого активизируется только одно подмножество производств. Оно загружается в операционную часть однородной вычислительной системы для параллельного исполнения. **Результаты.** Показано, что количественные оценки сокращения времени вывода зависят от общего числа производств, количества образуемых подмножеств и их размера. Моделирование показало, что даже простейшая декомпозиция на два подмножества (одно подмножество состоит из 2-х производств) дает временной выигрыш (1,07-1,52) раз, пропорциональный общему числу производств. **Выводы.** Построенная расширенная машина символического вывода является основой для последующего создания архитектуры однородной вычислительной системы с комбинацией централизованного и локального управления, что позволяет вычислительным блокам однородной операционной части параллельно работать без избыточного обращения к общей памяти.

Производственная система; параллельные вычисления; независимые подмножества производств; конфликтные слова.

E.A. Titenko, I.E. Chernetskaya, M.A. Titenko, E.V. Melnik, D.A. Trokoz

ADVANCED PRODUCTION OUTPUT ENGINE FOR IMPLEMENTING PARALLEL COMPUTING

Relevance. The paper discusses a theoretical approach to organizing parallel computing based on a production model of data flow control. The production paradigm of parallel computing has the necessary conditions for building new architectures and organizing high-performance parallel computing. We consider production (mathematical) systems that control sets of left-hand sides of productions (samples). The goal is to increase the efficiency of parallel inference of solutions by reducing unproductive time spent searching through possible alternatives in the inference graph space. The research is based on the creation of an extended symbolic computation machine for implementing parallel steps. A symbolic computing machine is an abstract system that systematizes production output as a sequence of four computational and search stages. The inference engine defines the general appearance of a homogeneous computing system. The main difference is the decomposition of the base of production rules into separate subsets based on the algebra of production and the structuring of relations between products. Instead of a single “flat” structure, it is proposed to decompose the product base into parts - to introduce a system of independent subsets of products. Parallel inference is implemented for individual subsets without loss of generality, while the search for possible alternatives is reduced. Each subset of productions has a special marker word, the value of which activates only one subset of productions. It is loaded into the operating part of a homogeneous computing system for parallel execution. **Results.** It is shown that quantitative estimates of the reduction in output time depend on the total number of productions, the number of subsets formed and their size. Simulation has shown that even the simplest decomposition into two subsets (one

subset consists of 2 productions) gives a time gain of (1.07-1.52) times, proportional to the total number of productions. Conclusions. The created extended symbolic computing machine is the basis for the subsequent creation of the architecture of a homogeneous computing system with a combination of centralized and local control. This property allows computational units of a homogeneous operating part to work in parallel without excessive access to shared memory.

Production system; parallel computing; independent subsets of production; conflict words.

Введение. Перспективные однородные вычислительные системы (ОВС) в значительной степени ориентируются на новый класс проблемно-поисковых задач, связанных с выполнением расчетно-поисковых, переборных, оптимизационных вычислений [1–4]. Значимыми примерами таких задач являются задачи вычислительной химии, физики, лингвистики, алгебраической комбинаторики слов, символьного кодирования/декодирования временных рядов, организации вычислений на графах, обеспечения реконфигурации многопроцессорных систем, вычислительно трудоемкие задачи структуризации Big Data и другие проблемно-поисковые задачи и модельные задачи, использующие модели и технологии искусственного интеллекта [5–9]. Их отличительная особенность – ориентация на обработку высокоуровневой информации (числовые, символьные, реляционные, смысловые данные), комбинация числовых и символьных преобразований, необходимость генерации множества вариантов для последующего отбора наиболее подходящих из них. Другая отличительная особенность таких задач – существенная доля временных затрат операций обращения к памяти в общем объеме вычислений, нерегулярный характер обращений к ячейкам памяти, что не позволяет использовать конвейеризацию аппаратных схем декодирования адресов, принятую при обработке массивов числовых данных.

Решение проблемно-поисковых задач преимущественно имеет недетерминированный характер. Вычислительно-поисковые действия связаны с генерацией, анализом, перебором множества промежуточных вариантов, отбором наиболее приоритетных путей в пространстве альтернатив и их проверкой на предмет совпадения с желаемыми или эталонными значениями. Такие вычисления традиционно выполняются в рамках моделей обработки знаний (логическая, фреймовая, продукционная модели, модель семантических сетей и др.) [10–13].

Традиционный путь разработки однородных ВС для таких задач считается экстенсивным. Он основан на методе «грубой силы»: введение структурной избыточности в архитектуру ОВС, повышение производительности за счет повышения тактовой частоты процессоров и устройств, распараллеливание преимущественно потока команд, создание системы макрокоманд, выполняющих крупноблочные числовые операции. Тем не менее используемая модель вычислений не имеет связи со структурой ОВС и функциональными блоками в ее составе. Как следствие, возникает семантический разрыв, порождающий набор избыточных вычислительных, коммуникационных, пересылочных операций, операций перезаписи данных между блоками памяти вместо их содержательной обработки. Как правило, такие операции не меняют структуру обрабатываемых данных, но входят в общее время решения задачи. Данный семантический разрыв позволяет достигать ОВС лишь около 10-15% от пиковой производительности на задачах реального уровня сложности [14]. Особенно ощутимы временные потери на задачах с преобладанием существенной доли обменных операций, что характерно для задач искусственного интеллекта (ИИ).

В противовес экстенсивному пути однородные ОВС целесообразно развивать на основе «сквозной» технологии, совмещающей методы и средства проектирования и реализации параллельных вычислений. «Сквозная» технология вычислений основана на преемственности математических, алгоритмических, архитектурных и схемотехнических решений в составе ОВС. Она является своеобразным «интеллектуальным мостиком» между теоретическими и аппаратно-программными средствами параллельных вычислений. Интеллектуализация вычислений является эффективным инструментом повышения производительности ОВС для проблемно-поисковых задач реального уровня сложности (количество анализируемых элементов $\approx 10^{14}$ - 10^{15}) за счет извлечения дополнительной информации о свойствах задачи и ходе вычислительного процесса.

В общем случае, под интеллектуализацией вычислений подразумеваются процессы автоматической (автоматизированной):

- ◆ адаптации структуры ОВС к информационному графу задачи;
- ◆ выполнения вычислительного процесса по мере готовности данных;
- ◆ эволюционного изменения алгоритма работы в зависимости от достигнутых состояний;
- ◆ аппаратной реализации крупноблочных символьных операций;
- ◆ динамической реконфигурации ОВС в зависимости от информационного графа задачи;
- ◆ выделения и накопления типовых (шаблонных) описаний и их последующего применения и др.

Модели и методы обработки знаний являются основой интеллектуализации вычислений в перспективных однородных ВС с параллельной обработкой высокоуровневой информации [10, 12, 15, 16].

Продукционные системы (ПС) – это общепризнанный и достаточно исследованный математический аппарат, использующий вычислительную схему «условие – реакция» для управления вычислительным процессом. Целесообразность выбора ПС для повышения производительности работы современных ОВС определяется следующими свойствами:

- ◆ однородность правил;
- ◆ легкость расширения базы продукций;
- ◆ гибкость реализации различных схем управления;
- ◆ естественный параллелизм исполнения набора продукций на основе потока данных (левых частей продукций);
- ◆ преемственность математических и алгоритмических операций;
- ◆ возможность расширения структуры правила типовыми числовыми обработчиками.

ПС предназначены для обработки символьной информации (ОСИ), под которой понимаются вычисления от словесного и выше уровня абстракции данных [17, 18]. Формальной основой ОСИ является алгебра продукций. Она позволяет исследовать соотношения между строковыми операндами, выполнять эквивалентные преобразования продукций, декомпозировать систему продукций, вести параллельные вычисления на основе стратегий выводов. Главное преимущество таких продукционных систем – это естественная ориентация на управление потоком данных, т.е. на параллельное исполнение подмножества правил (продукций) по мере их готовности. Условие готовности для продукций из ПС считается унифицированным – активация продукции, т.е. положительный результат поиска по образцу в описании внешней среды. Семантическая связь ПС с ОВС реализуется путем отображения продукций на вычислительные блоки, аппаратно реализующие базовые операции над строковыми данными. В этом случае создаются необходимые условия для параллельного исполнения множества продукций на основе принципа управления потоком данных.

Таким образом, актуальность исследования определяется применением продукционных моделей и стратегий параллельного вывода для их последующей аппаратной реализации на ОВС.

Материалы и методы. Пусть заданы $n \in \mathbb{N}$, рабочий алфавит A , метасимволы $\rightarrow \notin A$, $* \notin A$. Для описания продукционных систем принято вводить расширение M над рабочим алфавитом A . Расширение необходимо, чтобы выделять служебные символы и особые (предметные) свойства в данных и шагах работы продукционной системы

$$B = \{ A, \rightarrow, \bullet, \xi, \eta, \dots, \mu, \alpha, \beta, \dots, \delta, \nabla, \diamond \dots \}, \quad (1)$$

где ξ, η, \dots, μ – алфавитные (символьные) переменные, $\alpha, \beta, \dots, \delta$ – переменные-челноки, $\nabla, \diamond \dots$ – переменные-указатели.

Объекты $O_i, P_i \in A^*$ – слова, составленные по правилам конкатенации, где $i=1-n$.

Пусть задано обрабатываемое слово $S \in A^*$.

Продукция – это математическое выражение в алфавите M вида [12]

$$O \xrightarrow{\Theta} P, \quad (1)$$

где O – левая часть (слово-условие) в алфавите B ; P – правая часть (слово-реакция) в алфавите B , Θ – индикатор типа продукции (текущая или заключительная).

Тогда продукционная система (*Rule System Based – RSB*) задается как система-тройка:

$$RSB = \{ B, \Theta, C \}, \quad (2)$$

где C – стратегия (схема) вывода, Θ – определяющее множество продукций, управляемых множеством образцов продукций (поток данных):

$$\Theta : \begin{cases} O_1 \rightarrow P_1 \\ \dots \\ O_i \rightarrow P_i \\ \dots \\ O_n \rightarrow P_n \end{cases} \quad (3)$$

Содержательно ПС представляется множеством продукций, объединенных стратегией вывода C , которая задает количество и порядок срабатывания продукций на каждом шаге вывода. Согласно принципу управления потоком данных шаг вывода состоит в срабатывании подмножества продукций (активированные продукции) и формировании набора промежуточных слов. ПС завершает работу на основе двух условий:

- ◆ срабатывание заключительной продукции;
- ◆ отсутствие вхождения левых частей продукций в слово S .

Известно, что в архитектурах и машинах, управляемых потоком данных (*DataFlows Machine*) [15], используется принцип потенциального распараллеливания вычислений по данным. Это означает, что операционная часть однородной ВС содержит множество однотипных вычислительных блоков, реализующих все базовые операции. Тем не менее лишь некоторое подмножество продукций из выражения (3) будет потенциально активно для вывода на текущем ярусе графа вывода. Соответственно из операционной части ОВС необходимо динамически выбирать только такие продукции, число которых изменяется по шагам вывода. В результате, в операционной части однородной ВС возникают конфликты на размещение подмножеств продукций.

Для снижения ресурсных затрат однородной ВС на работу ПС (объем памяти, время перебора продукций) предлагается ее декомпозиция на обособленные подмножества, выделяемые на основе операций алгебры продукций [19]. В этом случае работа ПС выполняется по каждому подмножеству с сохранением корректности конечного результата.

Декомпозиция ПС включает операции конструктивной логики пересечения, объединения, дополнения слов и др. Эти операции попарно устанавливают свойства структурной связи между образцами и подстановками продукций вида (3) [12]. Над ПС выполняются подготовительные вычислительные действия, позволяющие исследовать соотношения между продукциями и выполнить разбиение их на t независимых подмножеств D_1, D_2, \dots, D_t , обеспечивающие параллельный вывод по каждому подмножеству отдельно.

Тем не менее характер подготовительных действий имеет комбинаторную сложность. Количество проверок на парные пересечения левых и правых частей из выражения (3) имеет квадратичную зависимость, что при больших n или частых изменениях состава ПС является достаточно критичным для эффективности работы ПС.

Обеспечение параллельной работы ПС вида (3) обеспечивается поиском такого подмножества продукций D_y ($y \in 1 \dots t$), которое имеет максимальный размер. Соответственно целевая функция подготовительных действий имеет вид

$$f\{\Theta\} = D, |D| \rightarrow \max. \quad (4)$$

Построенное подмножество максимального размера позволяет задать в операционной части ОВС минимально необходимое число функциональных блоков, реализующих продукции, и обеспечить тем самым параллельное исполнение множества путей на текущем уровне графа вывода.

С точки зрения теоретического проектирования параллельный вывод в системе продукции описывается системой вида [20]

$$\{R, \Theta, I_M\}, \quad (5)$$

где R – исходные данные в алфавите M ; Θ – определяющее множество продукции (база продукции); I_M – машина вывода.

Машина вывода – это абстрактная система, имеющая циклическую вычислительную схему на основе управления потоком данных (активированных продукции). Будучи впервые созданной для проектирования экспертных систем, машина вывода в дальнейшем стала инструментом для проектирования и программирования однородных ВС под класс задач на основе продукционной парадигмы вычислений [21, 22].

Основу работы машины вывода составляет вычислительный цикл, содержащий четыре вычислительных этапа: выборка данных, активация подмножества продукционных правил, разрешение конфликта и срабатывания продукции, изменения состояний внешней рабочей памяти R .

Теоретико-множественное описание машины вывода I_M имеет вид

$$I_M = \{V, SMP, K, W\}, \quad (6)$$

где V – блок выборки из R и B множества активных данных и множества активных продукции; SMP – блок сопоставления, определяющий множество вхождений образов продукции во входные данные; K – блок разрешение конфликтов; W – блок выполнение выбранных продукции.

Структура стандартной машины вывода I_M представлена на рис. 1 в виде циклической последовательности 4-х вычислительных-поисковых блоков, собственно базы продукции и рабочей памяти для записи исходных данных и результатов вывода.

Цикл работы машины вывода состоит в последовательном выполнении выше обозначенных модулей: 1 – выборка, 2 – сопоставление, 3 – разрешение конфликтов, 4 – выполнения продукции и модификация рабочей памяти.

Использование продукционной модели приводит к управлению данными (образцами активированных продукции) в общем цикле работы. Машина вывода на каждом цикле формально порождает новые слова (промежуточные или конечные решения), которые записываются в рабочую память. Циклы вывода повторяются до тех пор, пока не будет получен целевой результат или ПС будет не применима к обрабатываемым словам.

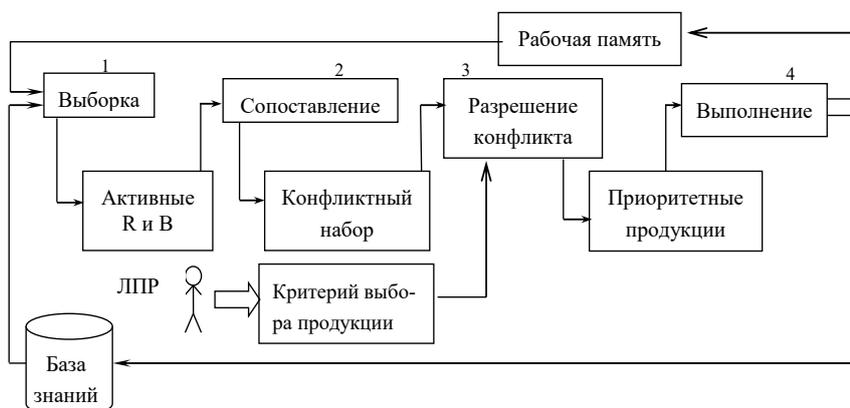


Рис. 1. Структура стандартной машины вывода

Разнообразие алгоритмических, технических вариантов реализаций блоков выборки, сопоставления, разрешения конфликтов и выполнения определяет стратегию вывода решений. При этом блоки сопоставления и разрешения конфликтов являются определяющими во временных затратах работы машины вывода.

Недостаток машины вывода со стандартной структурой заключается в последовательном характере вычислений как по блокам, так и внутри блоков, а формируемый при этом граф вывода, понимается как лабиринт альтернативных путей. Разрешение конфликта определяется лицом, принимающим решения (ЛПР), директивно. Выбор ЛПР приоритетных продуктов из конфликтного набора преимущественно проводится в рамках имеющихся вычислительных ресурсов и не зависит от машины вывода. Он проводится без содержательного анализа множества структурных пересечений образцов между собой, что является системным ограничением для обеспечения параллельного вывода.

Для систем продуктов с недетерминированным характером их срабатывания на этапе сопоставления, в общем случае, формируется некоторое подмножество активированных продуктов и активных данных, готовых к применению и модификации соответственно. Тем не менее, блок разрешения конфликта, в основном, осуществляет выбор из конфликтного множества единственной приоритетной продукции для выполнения. Остальные активированные продукты и активные данные будут рассмотрены на следующем шаге вывода или отложены по исполнению.

Сущность предлагаемого подхода к реализации недетерминизма в работе машины вывода заключается в понимании графа вывода не как лабиринта, а как сетевого пространства равных возможностей, в котором все пути в пределах текущего уровня могут быть сгенерированы параллельно.

Для формализации данных рассуждений выполняется модификация машины вывода путем добавления в ее состав новых блоков: блока синтеза конфликтных слов и блока обеспечения параллельных выводов.

Блок синтеза конфликтных слов основан на циклическом анализе парных пересечений образцов системы продуктов (3) и полученных конфликтных слов K . Он обеспечивает формирование полного списка слов, для которых складывается конфликтная ситуация выбора. Проверка на конечность такого списка слов является логическим индикатором возможности параллельного вывода. Эту работу выполняет блок 5 обеспечения параллельных выводов.

Основу формирования списка конфликтных слов составляют конструктивные дизъюнкции вида

$$\left\{ \begin{array}{l} (O_i^H = O_j^K) \vee (O_i^K = O_i^H) \vee (O_i \subset O_j) \vee (O_j \subset O_i) \\ (K_{i1}^H = K_{j1}^K) \vee (K_{i1}^K = K_{i1}^H) \vee (K_{i1} \subset K_{j1}) \vee (K_{j1} \subset K_{i1}) \end{array} \right. , \quad (7)$$

где O_i^H, O_i^K – собственные начало и окончания образца O_i ; O_j^H, O_j^K – собственные начало и окончания образца O_j ; K_{i1}^H, K_{i1}^K – собственные начало и окончания $i1$ -го конфликтного K_{i1} ; K_{j1}^H, K_{j1}^K – собственные начало и окончания $j1$ -го конфликтного K_{j1} ; $i \neq j$, ($i=1 \dots n, j=1 \dots n$), $i1 \neq j1$, ($i=1 \dots t, j=1 \dots t$).

Истинность конструктивных дизъюнций (7) позволяет структурировать продукции из базы, разбить их на обособленные подмножества.

Блок 5 обеспечения параллельных выводов представлен набором параллельных стратегий: И- параллельный вывод, ИЛИ- параллельный вывод и др. [23, 24]. Данные стратегии обеспечивают эффективное движение в сетевом пространстве поиска на основе анализа текущей конфликтной ситуации, состава обрабатываемых слов в рабочей памяти, что создает предпосылки для динамического выбора наиболее эффективной стратегии.

На рис. 2 представлена структура модифицированной машины вывода. Практическая ценность данной машины вывода определяется тем, что вычисления в блоке синтеза конфликтных слов осуществляются однократно и в дальнейшем используются параллельными стратегиями для параллельной генерации текущих состояний.

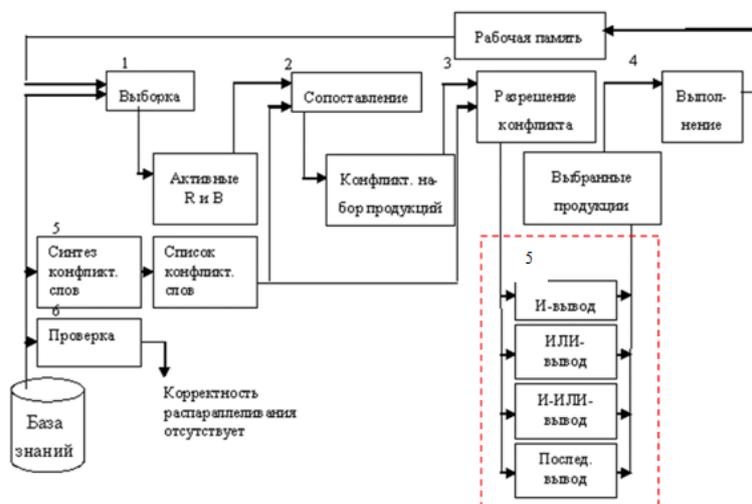


Рис. 2. Структура модифицированной машины вывода

Новизна модифицированной машины вывода заключается в формировании множества конфликтных слов на основе анализа пересечений пар образцов продукции и выборе приоритетных продукции из конфликтного множества без участия оператора (ЛПР). Комбинаторно-вычислительный процесс формирования множества конфликтных слов позволит сократить число возвратных переходов и тем самым сократить время построения графа вывода.

Дополнительные блоки проверки и синтеза конфликтных слов собственно в цикл работы модифицированной машины вывода не входят, поскольку обеспечивают необходимые вычисления однократно без использования обрабатываемых данных из рабочей памяти.

Вместе с тем модифицированная машина вывода имеет «плоскую» структуру базы продукции. Это означает, что в каждом цикле работы машины вывода на этапе сопоставления анализируются все продукции из базы, хотя фактически потенциально активным будет лишь некоторое подмножество продукции.

Для сокращения временных затрат до этапа сопоставления обрабатываемых слов из рабочей памяти и продукции из базы продукции выполняется анализ и декомпозиция базы продукции на независимые подмножества. Рис. 3 демонстрирует структуру расширенной машины вывода.

Теоретической основой декомпозиции базы продукции является конструктивные дизъюнкции вида (7). На основе пар номеров продукции, имеющих непустое пересечение, составляется список так называемых связанных продукции, составляющих подмножество D продукции и имеющих общий дескриптор, идентифицирующий данное подмножество. Этим дескриптором является расширенное конфликтное слово, построенное на основе конструктивных дизъюнций (7).

Коммутативность пересечений образцов продукции по (7) позволяет разбить исходную базу на t независимых подмножеств D_1, D_2, \dots, D_t , что обеспечит в дальнейшем выполнение этапа сопоставления только по этим подмножествам D_y , ($y=1 \dots t$).

Следующая возможность повышения гибкости стратегии вывода для ПС – расширительная модификация состава продукции. Известный подход [24] состоит во введении пред- и постусловий применения продукции. В отличие от традиционного подхода в предлагаемых производственных системах для уменьшения количества рассматриваемых вариантов вводятся не условия, а временные и количественные оценки самих продукции. Эта информация позволяет ранжировать продукции и при неопределённости выбора руководствоваться только наиболее приоритетной продукцией.

В качестве таких оценок в состав продукции вводятся следующие показатели:

- ◆ номер продукции (z_0);
- ◆ частота активации продукции (z_1);
- ◆ временной отсчет последнего срабатывания продукции (z_2);
- ◆ частота попадания в конфликтный набор продукции (z_3);
- ◆ длина левой части (z_4);
- ◆ длина правой части (z_5);
- ◆ количество переменных в левой части (z_6);
- ◆ приоритет заключения (z_7);
- ◆ условие модификации левой части (z_8);
- ◆ условие модификации право части (z_9);

Тогда продукция вида (2) приобретает вид

$$z_0 (z_1, \dots, z_6) O \rightarrow P(z_7, \dots, z_9). \quad (8)$$

Новизна модификации структуры продукции вида (8) связана с введением в ее состав элементов, инвариантных к типу решаемых задач, не только учитывающих предысторию состояний в графе вывода, но и условия модификации самих продукции в процессе выполнения, что уменьшает размер конфликтного множества продукции и повышает тем самым временную эффективность вывода решений в ПС

В итоге, расширенная машина вывода (рис. 3) представляет собой объект-семерку вида:

$$I_M = \{D, CW, AR, V, CMP, K, W\}, \quad (6)$$

где D – конфигуратор для декомпозиции на независимые подмножества продукции (слоты);

CW – набор конфликтных слов, маркирующих независимые подмножества продукции;

AR – арбитр, осуществляющий выбор подмножества продукции и стратегии параллельного вывода.

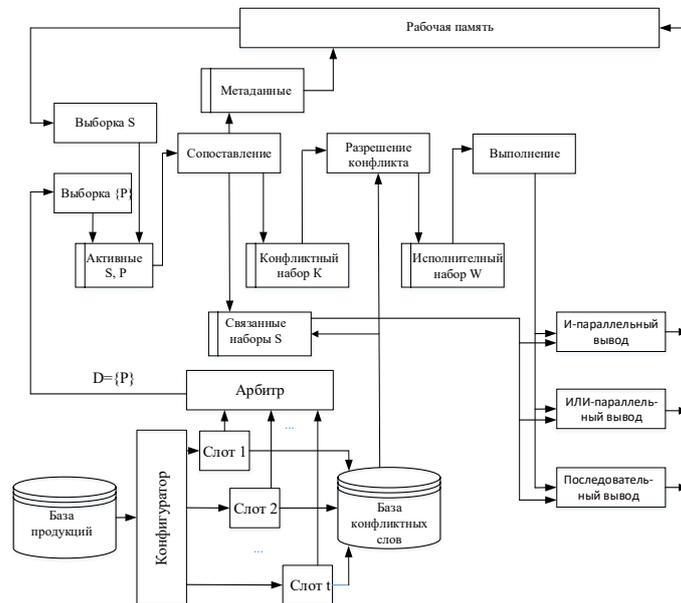


Рис. 3. Структура расширенной машина вывода

Результаты и их обсуждение. В качестве анализируемых ПС рассматриваются типовые системы продукционной обработки символьных данных: удвоение слова, обращения слова [25].

Пусть в рабочем алфавите A задано слово S в виде $S = \xi_1 \dots \xi_i \xi_{i+1} \dots \xi_L$, где $\xi_i \in A$ ($i=1-L$), $L=|S|$ – длина слова.

Тогда результат работы алгоритма U удвоения слова представляется как

$$U(S) = SS. \tag{9}$$

При удвоении S в качестве служебных метасимволов выделяются:

- ♦ метасимвол-челнок α , реализующий движение и выделяющий конечную область обработки рядом с челноком;

- ♦ признак удвоения текущей буквы ξ слова S – метасимвол-маркер β .

Определяющее множество Θ имеет вид

$$\left\{ \begin{array}{l} \alpha\xi \rightarrow \xi\beta\xi\alpha \\ \beta\xi\eta \rightarrow \eta\beta\xi \\ \beta \rightarrow \lambda \\ \alpha \rightarrow \bullet\lambda \\ \lambda \rightarrow \alpha \end{array} \right., \tag{10}$$

где λ – обозначение пустого символа (слова).

Первая продукция из (10) выполняет удвоение текущей буквы из S , расположенной справа от челнока α . Копия буквы ξ помечается маркером β , располагаемым слева от нее. Вторая продукция (10) выполняет движение копии буквы ξ вправо, реализуя тем самым посимвольное удвоение слова. Третья продукция из (10) удаляет растиражированные по обрабатываемому слову маркеры β . Четвертая продукция (10) заключительная, она удаляет челнок α , формируя тем самым окончательный результат и завершая работу алгоритма удвоения. Пятая продукция (10) срабатывает первой, она присоединяет слева к слову S челнок α , иницилируя тем самым этап алгоритма посимвольного удвоения. Этап посимвольного удвоения реализуется собственно первой продукцией (10), которая срабатывает L раз.

Трассировка работы алгоритма для $L=3$ приведена ниже в табл. 1. Анализ системы продукций удвоения слова на пересечения по (7) показывает, что продукции первая и вторая продукции из (10) имеют общую часть. Это означает, что система продукций (10) может быть декомпозирована на 2 отдельных подмножества:

- $D_1 = \{1, 2\}$;
- $D_2 = \{3, 4, 5\}$.

Далее вычислительный процесс может быть реализован отдельно по каждому из подмножеств продукций, что существенно сократит количество холостых шагов на перебор продукций в каждом из подмножеств. Трассировочная табл. 1 подтверждает, что продукции из подмножества $D_1 = \{1, 2\}$ составляют единую последовательность, не имеющую вставок из подмножества D_2 .

Таблица 1

Трассировка шагов алгоритма $U(S) = SS$

Обрабатываемое слово	Номер сработавшей продукции	Номер шага ПС	Примечание
Processed word	Number of triggered production	Step number PS	Note
$\xi_1 \xi_2 \xi_3$	5	1	Исходное слово
$\alpha \xi_1 \xi_2 \xi_3$	1	2	
$\xi_1 \beta \xi_1 \alpha \xi_2 \xi_3$	1	3	
$\xi_1 \beta \xi_1 \xi_2 \beta \xi_2 \alpha \xi_3$	1	4	

$\xi_1 \beta \xi_1 \xi_2 \beta \xi_2 \xi_3 \beta \xi_3 \alpha$	2	5	Движение вправо
$\xi_1 \xi_2 \beta \xi_1 \beta \xi_2 \xi_3 \beta \xi_3 \alpha$	2	6	
$\xi_1 \xi_2 \beta \xi_1 \xi_3 \beta \xi_2 \beta \xi_3 \alpha$	2	7	
$\xi_1 \xi_2 \xi_3 \beta \xi_1 \beta \xi_2 \beta \xi_3 \alpha$	3	8	Удаление маркера
$\xi_1 \xi_2 \xi_3 \xi_1 \beta \xi_2 \beta \xi_3 \alpha$	3	9	
$\xi_1 \xi_2 \xi_3 \xi_1 \xi_2 \beta \xi_3 \alpha$	3	10	Удаление челнока
$\xi_1 \xi_2 \xi_3 \xi_1 \xi_2 \xi_3 \alpha$	4	11	
$\xi_1 \xi_2 \xi_3 \xi_1 \xi_2 \xi_3$	-		STOP

Известен оригинальный алгоритм обращения, предложенный А.А. Марковым [25]. Оригинальным алгоритмическим решением является отсутствие внешнего маркера конца слова. Для прекращения фазы движения символов слова вправо А.А. Марков ввел внутренний маркер, обозначаемый удвоенным челноком $\alpha\alpha$ как признак обработки всех символов слова S .

Основу ПС обращения слова составляют продукции

$$\alpha\xi\eta \rightarrow \eta\alpha\xi,$$

$$\lambda \rightarrow \alpha.$$

Первая продукция выполняет движение вправо текущего символа ξ слова с помощью челнока α . Вторая продукция осуществляет левую конкатенацию челнока α к слову S . Тогда многократные применения продукции движения вправо приводят к инверсному расположению символов исходного слова. Рабочие символы попутно размечаются метасимволом α , который теперь понимается как маркер обработанной буквы.

Для перехода к следующей фазе вычислений необходимо ввести новый челнок движения вправо β . Челнок β при движении будет удалять метасимвол α , пока не останется один. В этой ситуации необходимо заключительной продукцией завершить работу ПС $\beta \rightarrow \bullet\lambda$.

Итоговое определяющее множество ПС обращения слова имеет вид

$$\left\{ \begin{array}{l} \alpha\alpha \rightarrow \beta\alpha \\ \beta\alpha \rightarrow \beta \\ \beta \rightarrow \bullet\lambda \\ \alpha\xi\eta \rightarrow \eta\alpha\xi \\ \lambda \rightarrow \alpha \end{array} \right. \quad (12)$$

Пусть $A=\{a,b,c\}$ и пусть в A задано слово $S=abc$. Тогда работа (12) над S имеет вид

$$\begin{aligned} abc &\xrightarrow{5} \alpha abc \xrightarrow{4} \beta aac \xrightarrow{4} \beta caa \xrightarrow{5} \alpha \beta caa \xrightarrow{4} \alpha \beta aca \xrightarrow{5} \alpha \beta aca \alpha \xrightarrow{5} \\ \alpha \alpha c a b a a &\xrightarrow{1} \beta \alpha c a b a a \xrightarrow{2} \beta c a b a a \xrightarrow{3} c \beta a b a a \xrightarrow{2} c \beta b a a \xrightarrow{3} c b \beta a a \xrightarrow{2} \\ c b \beta a &\xrightarrow{3} c b a \end{aligned}$$

Анализ работы (12) по шагам показывает, что четвертая и пятая продукции имеют непустое пересечение, что позволяет объединить их в одно подмножество. Исходная система продукций разбивается на подмножества:

$$\begin{aligned} - D_1 &= \{1, 2, 3\}; \\ - D_2 &= \{4, 5\}. \end{aligned}$$

Трассировка сработавших продукций из (12) показывает, что продукции действительно выполняются отдельно по подмножествам D_1 и D_2 . Значит, предварительная декомпозиция (12) на части является новым рабочим шагом в расширенной машине вывода, позволяющим сократить количество холостых срабатываний и повысить эффективность вывода решений.

Оценочно, декомпозиция исходной продукционной базы на 2 подмножества позволяет сократить число холостых срабатываний продукции до $(n / d_{max} + f(i) \cdot k)$ раз, где k – число уровней в графе вывода решений, $f(i)$ – модельная оценка точки разрыва ПС на подмножества.

Для моделирования временных затрат работы ПС задаются следующие параметры:

- ◆ количество продукции – n ;
- ◆ количество уровней в графе вывода – k ;
- ◆ количество декомпозируемых подмножеств – d ;
- ◆ номера точек разрыва ПС – i_1, i_2, \dots, i_d ;
- ◆ временная пропорция операций сопоставления и подстановки $\tau_1 : \tau_2$.

Пусть рассматривается декомпозиция базы продукции на 2 подмножества, одно из которых состоит из 2-х продукции. Пусть принята пропорция сопоставления и подстановки $5\tau : \tau$. Единственная точка разрыва ПС будет иметь варьируемое значение. Минимальный номер точки разрыва – 3, максимальный – $(n-2)$.

На рис. 4 представлены результаты моделирования времени работы $T(n)$ продукционной системы при $i=3, i=n-2$, для количества уровней в графе вывода $k=10$. На рис. 5 представлены результаты моделирования времени работы $T(n)$ продукционной системы при $i=3, i=n-2$ для количества уровней в графе вывода $k=15$.

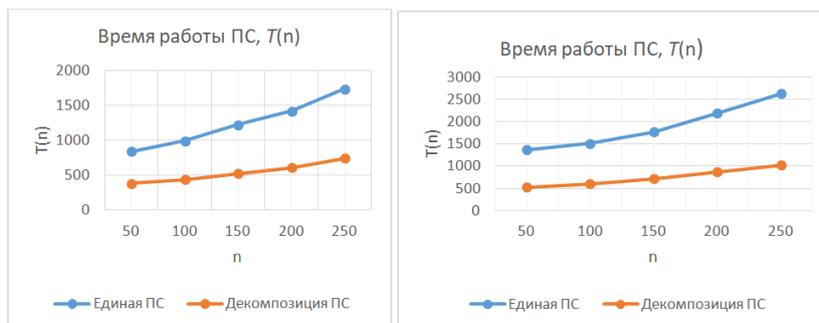


Рис. 4. Время зависимости времени работы ПС при $i=3, i=n-2$ для $k=10$

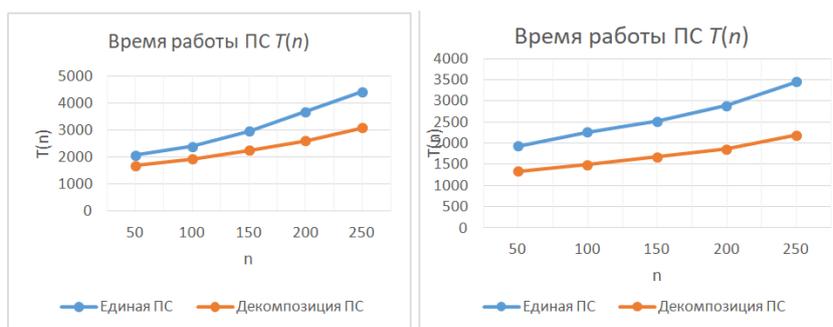


Рис. 5. Время зависимости времени работы ПС при $i=3, i=n-2$ для $k=15$

Графики демонстрируют существенное сокращение затрат времени при работе ПС, разбитой на независимые подмножества, так как в каждой итерации (уровень графа вывода) сокращается число рассматриваемых альтернатив в среднем на 25-30% от мощности наибольшего подмножества. При этом с увеличением количества продукции в 5 раз работа ПС при ее декомпозиции на 2 подмножества (одно из которых – две продукции) относительный выигрыш по времени описывается интервалами $(1,07 \div 1,52)$ и $(1,14 \div 1,74)$ раз, что подтверждает потенциальное преимущество работы ПС при ее разбиении на независимые подмножества.

Выводы:

1. Организация высокопроизводительных параллельных вычислений на основе продукционной парадигмы связывается с созданием теоретических и прикладных инструментов, направленных на сокращение временных затрат параллельного вывода.

2. В теоретическом аспекте перспективный путь организации параллельных вычислений – создание продукционной машины вывода, имеющей однородную исполнительную часть и средства декомпозиции базы продукции на подмножества с сохранением корректности результата. Созданная расширенная машина вывода отличается введением ее структуру блоков синтеза конфликтных слов, декомпозиции на независимые подмножества и параллельного выполнения. Анализ отношений пересечения продукции своими левыми частями позволяет логически разбить систему продукции на независимые подмножества и вести параллельный вывод по каждому из подмножеств с сохранением корректности результата. Продукционная база представлена независимыми подмножества продукции, имеющими специальные слова-маркеры, которые участвуют в выборе стратегии вывода.

3. Временной выигрыш декомпозиции продукционной базы достигается за счет сокращения количества рассматриваемых вариантов (альтернатив) на этапах сопоставления активных продукции и данных, разрешения конфликтов, срабатывания продукции. Экспертная оценка показала, что в худшем случае декомпозиция исходной продукционной базы на два подмножества позволяет сократить число холостых срабатываний продукции на $(n / d_{max} + f(i) \cdot k)$ раз, где k – число уровней в графе вывода решений, d_{max} – мощность максимального подмножества. Моделированием установлено сокращение затрат времени при работе продукционной системы в среднем на 25-30% от мощности наибольшего подмножества. При этом с увеличением количества продукции до 5 раз работа продукционной системы при ее декомпозиции на 2 подмножества (одно из которых – две продукции) относительный выигрыш по времени в худшем случае составил $(1,07 \div 1,52)$ раз, что подтверждает преимущество создания расширенной машины вывода, которая имеет однородную структуру, и способна выполнять параллельные вычисления по независимым подмножествам продукции.

4. Структуризация вида продукции является прикладным инструментом повышения эффективности параллельного вывода за счет избирательной выборки активных продукции из базы. Для этого в дополнение к традиционным показателям выборки активных продукции вводятся условия изменения структуры продукции в процессе вывода, что позволяет зафиксировать в продукционной базе машины вывода общезначимые свойства и отношения, а в дальнейшем – использовать их для сокращения числа шагов работы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Воеводин В.В.* Математические модели и методы в параллельных процессах. – М.: Наука, 1986. – 296 с.
2. *Гузик В.Ф., Каляеа И.А., Левин И.И.* Реконфигурируемые вычислительные системы: учеб. пособие. – Таганрог: Изд-во ЮФУ, 2016. – 472 с.
3. *Булыга Ф.С., Курейчик В.М.* Сравнительный анализ методов векторизации текстовых данных большой размерности // Известия ЮФУ. Технические науки. – 2023. – № 2 (232). – С. 212-226.
4. *Титенко Е.А., Емельянов С.Г., Зерин И.С.* Однородные вычислительные структуры для параллельных символьных вычислений // Известия Юго-Западного государственного университета. – 2011. – № 6 (39). – Ч. 2. – С. 77-82.
5. *Lothaire M.* Algebraic Combinatorics on Words. – Cambridge: Cambridge University Press; 2002. (In Eng). – DOI: <https://doi.org/10.1017/CBO9781107326019>.
6. *Lothaire M.* Applied Combinatorics on Words. In: Encyclopedia of Mathematics and its Applications. – Cambridge: Cambridge University Press, 2005.
7. *Люгер Дж.Ф.* Искусственный интеллект: стратегии и методы решения сложных проблем. – М.: Издательский дом «Вильямс», 2003. – 864 с.
8. *Огнев И.В., Борисов В.В., Сутула Н.А.* Ассоциативные память, среды, системы, – М.: Горячая линия – Телеком, 2016. – 420 с.

9. *Гладков Л.А., Курейчик В.В., Курейчик В.М., Сороколетов П.В.* Биоинспирированные методы в оптимизации. – М.: Физматлит, 2009. – 384 с.
10. *Адамов А.А., Эйсымонт Л.К.* Варианты архитектурных решений ЭКБ для систем искусственного интеллекта // Проектирование будущего. Проблемы цифровой реальности: Тр. 3-й Международной конференции. – М.: ИПМ им. М.В. Келдыша, 2020. – С. 112-131.
11. *Ва Б.У., Лоуррай М.Б., Гоцзе Ли.* ЭВМ для обработки символьной информации // ТИИЭР. – 1989. – Т. 77, № 4. – С. 5-40.
12. *Добрица В.П., Титенко Е.А., Халин Ю.А., Киселев А.В.* Системы искусственного интеллекта. – Курск: ЗАО «Университетская книга», 2023. – 143 с.
13. *Эйсымонт Л.К., Моляков А.С., Заборовский В.С., Федоров С.А.* Символьная обработка: эпизоды отечественной истории и перспективы // Матер. 2-й Всероссийской научно-технической конференции «Суперкомпьютерные технологии (СКТ-2012)». с. Дивноморское. 2012. – С. 202-206.
14. *Бурцев В.С.* Параллелизм вычислительных процессов и развитие архитектур суперЭВМ: Сб. статей. – М.: ТОРУС ПРЕСС, 2006. – 416 с.
15. *Wichert A.* Artificial intelligence and a universal quantum computer // AI Communications. – 2016. – Vol. 29, Issue 4. – P. 537-543.
16. *Бова В.В., Кравченко Ю.А., Родзин С.И.* Методы и алгоритмы кластеризации текстовых данных (обзор) // Известия ЮФУ. Технические науки. – 2022. – № 4 (228). – С. 122-143.
17. *Рыбина Г.В.* Основы построения интеллектуальных систем. – М.: Финансы и статистика. 2010. – 430 с.
18. *Попов Э.В.* Статические и динамические экспертные системы. – М.: Финансы и статистика. 1996. – 211 с.
19. *Довгаль В.М.* Методы модификации формальных систем обработки символьной информации. – Курск: Изд-во Курск. гос. техн. ун-т, 1996. – 115 с.
20. *Титенко Е.А.* Продукционная модель и абстрактная машина для реализации распределенных параллельных вычислений // Телекоммуникации. – 2012. – № 6. – С. 7-11.
21. *Тиликин А.П., Титенко Е.А.* Модификация цикла работы машины вывода для параллельных вычислительных устройств // Известия Юго-Западного государственного университета. – 2011. – № 6-2 (39). – С. 92-96.
22. *Тутов Е.Б., Титенко Е.А., Атакищев А.О., Цуканов К.С.* Применение контекстно-свободных грамматик для обработки текстов на естественном языке // Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение. – 2012. – № 2-1. – С. 97-101.
23. *Довгаль В.М., Титенко Е.А.* Параллельные стратегии выводов для систем поддержки принятия решения // Вестник новых медицинских технологий. – 2006. – Т. 13, № 1. – С. 143-144.
24. *Добрица В.П., Титенко Е.А., Халин Ю.А., Катыхин А.И.* Модели представления и обработки знаний в информационно-аналитических системах. – Курск: ЗАО «Университетская книга», 2023. – 172 с.
25. *Kapri X.B.* Основания математической логики: пер. с англ. – М.: Мир, 1969. – 568 с.

REFERENCES

1. *Voevodin V.V.* Matematicheskie modeli i metody v parallel'nykh protsessakh [Mathematical models and methods in parallel processes]. Moscow: Nauka, 1986, 296 p.
2. *Guzik V.F., Kalyaeva I.A., Levin I.I.* Rekonfiguriruemye vychislitel'nye sistemy: ucheb. posobie [Reconfigurable computing systems: textbook; manual]. Taganrog: Izd-vo YuFU, 2016, 472 p.
3. *Bulyga F.S., Kureychik V.M.* Sravnitel'nyy analiz metodov vektorizatsii tekstovykh dannykh bol'shoy razmernosti [Comparative analysis of methods for vectorizing high-dimensional text data], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2023, No. 2 (232), pp. 212-226.
4. *Titenko E.A., Emel'yanov S.G., Zerin I.S.* Odnородnye vychislitel'nye struktury dlya parallel'nykh simvol'nykh vychisleniy [Homogeneous computational structures for parallel symbolic computations], *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta* [News of the South-West State University], 2011, No. 6 (39), Part 2, pp. 77-82.
5. *Lothaire M.* Algebraic Combinatorics on Words. Cambridge: Cambridge University Press; 2002. (In Eng). – DOI: <https://doi.org/10.1017/CBO9781107326019>.
6. *Lothaire M.* Applied Combinatorics on Words. In: Encyclopedia of Mathematics and its Applications. Cambridge: Cambridge University Press, 2005.
7. *Lyuger Dzh.F.* Iskusstvennyy intellekt: strategii i metody resheniya slozhnykh problem [Artificial intelligence: strategies and techniques for solving complex problems]. Moscow: Izdatel'skiy dom «Vil'yams», 2003, 864 p.

8. *Ognev I.V., Borisov V.V., Sutula N.A.* Assotsiativnye pamyat', sredy, sistemy [Associative memory, environments, systems], Moscow: Goryachaya liniya – Telekom, 2016, 420 p.
9. *Gladkov L.A., Kureychik V.V., Kureychik V.M., Sorokoletov P.V.* Bioinspirirovannye metody v optimizatsii [Bioinspired optimization methods]. Moscow: Fizmatlit, 2009, 384 p.
10. *Adamov A.A., Eysymont L.K.* Varianty arkhitekturnykh resheniy EKB dlya sistem iskusstvennogo intellekta [Options for ECB architectural solutions for artificial intelligence systems], *Proektirovanie budushchego. Problemy tsifrovoy real'nosti: Tr. 3-y Mezhdunarodnoy konferentsii* [Designing the future. Problems of digital reality: proceedings of the 3rd International Conference]. Moscow: IPM im. M.V. Keldysha, 2020, pp. 112-131.
11. *Va B.U., Louray M.B., Gotsze Li.* EVM dlya obrabotki simvol'noy informatsii [Computer for processing symbolic information], *TIIER* [TIIER], 1989, Vol. 77, No. 4, pp. 5-40.
12. *Dobritsa V.P., Titenko E.A., KHalin Yu.A., Kiselev A.V.* Sistemy iskusstvennogo intellekta [Artificial intelligence systems]. Kursk: ZAO «Universitetskaya kniga», 2023, 143 p.
13. *Eysymont L.K., Molyakov A.S., Zaborovskiy V.S., Fedorov S.A.* Simvol'naya obrabotka: epizody otechestvennoy istorii i perspektivy [Symbolic processing: episodes of Russian history and prospects], *Mater. 2-y Vserossiyskoy nauchno-tekhnicheskoy konferentsii «Superkomp'yuternye tekhnologii (CKT-2012)»* [Materials of the 2nd All-Russian Scientific and Technical Conference “Supercomputer Technologies (SKT-2012)”]. Divnomorskoe, 2012, pp. 202-206.
14. *Burtsev V.S.* Parallelizm vychislitel'nykh protsessov i razvitie arkhitektur superEVM: Sb. statey [Parallelism of computing processes and development of supercomputer architectures: collection of articles: Articles]. Moscow: TORUS PRESS, 2006, 416 p.
15. *Wichert A.* Artificial intelligence and a universal quantum computer, *AI Communications*, 2016, Vol. 29. Issue 4, pp. 537-543.
16. *Bova V.V., Kravchenko Yu.A., Rodzin S.I.* Metody i algoritmy klasterizatsii tekstovykh dannykh (obzor) [Methods and algorithms for clustering text data (review)], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2022, No. 4 (228), pp. 122-143.
17. *Rybina G.V.* Osnovy postroeniya intellektual'nykh sistem [Fundamentals of building intelligent systems]. Moscow: Finansy i statistika. 2010, 430 p.
18. *Popov E.V.* Statische i dinamicheskie ekspertnye sistemy [Static and dynamic expert systems]. Moscow: Finansy i statistika. 1996, 211 p.
19. *Dovgal' V.M.* Metody modifikatsii formal'nykh sistem obrabotki simvol'noy informatsii [Methods for modifying formal systems for processing symbolic information]. Kursk: Izd-vo Kursk. gos. tekhn. unt, 1996, 115 p.
20. *Titenko E.A.* Produktsionnaya model' i abstraktnaya mashina dlya realizatsii raspredelennykh parallel'nykh vychisleniy [Production model and abstract machine for implementing distributed parallel computing], *Telekommunikatsii* [Telecommunications], 2012, No. 6, pp. 7-11.
21. *Tipikin A.P., Titenko E.A.* Modifikatsiya tsikla raboty mashiny vyvoda dlya parallel'nykh vychislitel'nykh ustroystv [Modification of the output machine operation cycle for parallel computing devices], *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta* [News of the South-West State University], 2011, No. 6-2 (39), pp. 92-96.
22. *Tutov E.B., Titenko E.A., Atakishchev A.O., TSukanov K.S.* Primenenie kontekstno-svobodnykh grammatik dlya obrabotki tekstov na estestvennom yazyke [Application of context-free grammars for processing texts in natural language], *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta. Seriya: Upravlenie, vychislitel'naya tekhnika, informatika. Meditsinskoe priborostroenie* [News of the South-Western State University. Series: Management, computer technology, computer science. Medical instrumentation], 2012, No. 2-1, pp. 97-101.
23. *Dovgal' V.M., Titenko E.A.* Parallelnye strategii vyvodov dlya sistem podderzhki prinyatiya resheniya [Parallel inference strategies for decision support systems], *Vestnik novykh meditsinskikh tekhnologiy* [Bulletin of new medical technologies], 2006, Vol. 13, No. 1, pp. 143-144.
24. *Dobritsa V.P., Titenko E.A., Khalin Yu.A., Katykhin A.I.* Modeli predstavleniya i obrabotki znaniy v informatsionno-analiticheskikh sistemakh [Models of representation and processing of knowledge in information and analytical systems]. Kursk: ZAO «Universitetskaya kniga», 2023, 172 p.
25. *Karri Kh.B.* Osnovaniya matematicheskoy logiki [Foundations of mathematical logic]: transl. from english. Moscow: Mir, 1969, 568 p.

Статью рекомендовал к опубликованию д.т.н. Ю.А. Кравченко.

Титенко Евгений Анатольевич – Юго-Западный государственный университет; e-mail: johntit@mail.ru; г. Курск, Россия; тел.: +79051588904; к.т.н.; доцент; в.н.с.

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

Титенко Михаил Андреевич – e-mail: mikhail-titenko@mail.ru; аспирант.

Мельник Эдуард Всеволодович – Федеральный исследовательский центр Южный научный центр Российской академии наук; e-mail: evml7@mail.ru; г. Таганрог, Россия; тел.: 88634615459; д.т.н., г.н.с.

Трокоз Дмитрий Анатольевич – Пензенский государственный технологический университет, e-mail: trokoz@penzgtu.ru; г. Пенза, Россия; проректор по научной работе.

Titenko Evgeny Anatolievich – South-West State University; e-mail: johntit@mail.ru; Kursk, Russia; phone: +79051588904; candidate of technical sciences, associate professor, leading researcher.

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

Titenko Mikhail Andreevich – e-mail: mikhail-titenko@mail.ru; postgraduate of the department of computer science.

Melnik Eduard Vsevolodovich – Federal Research Center, The Southern Scientific Center of the Russian Academy of Sciences; e-mail: evml7@mail.ru; Taganrog, Russia; phone: +78634615459; dr. of eng. sc.; chief researcher.

Trokoz Dmitry Anatolievich – Penza State Technological University; e-mail: trokoz@penzgtu.ru; Penza, Russia; dr. of eng. sc.; associate professor; vice-rector for scientific work.

УДК 004.9

DOI 10.18522/2311-3103-2024-2-268-278

А.А. Белевцев, А.М. Белевцев, В.А. Балыбердин**МЕТОДИКА ПОСТРОЕНИЯ ДОРОЖНЫХ КАРТ РАЗВИТИЯ
ТЕХНОЛОГИЧЕСКИХ ТРЕНДОВ И ТЕХНОЛОГИЙ В УСЛОВИЯХ
ОГРАНИЧЕНИЙ**

В современных условиях перехода к шестому технологическому укладу и обострения глобальной конкуренции первостепенное значение приобретает проблема обеспечения технологического суверенитета и технологического превосходства на национальном и корпоративном уровнях. В основе решения этой проблемы лежит организация эффективного развития научных исследований и, в первую очередь, исследований, направленных на создание новых высокоэффективных технологий в различных сферах деятельности. В работе предлагается методика анализа и развития технологических трендов и технологий в условиях неполноты и нечеткости информации. Методика обеспечивает прогнозирование и оценку развития общемировых технологически трендов для заданных предметных областей, а также определение возможных альтернатив технологического развития и разработки стратегии достижения технологического паритета и технологического превосходства в условиях ограничений. Реализация предлагаемых в методике процедур позволяет определить организационную стратегию достижения технологического паритета на различных уровнях рассмотрения, а также сформировать перечень технологий развития технологических трендов, обеспечивающих технологическое превосходство. Определено, что методика построения дорожных карт развития технологических трендов и технологий в условиях ограничений должна обеспечивать: 1. Анализ развития общемировых технологических трендов для заданных предметных отраслей на национальном (на пример в РФ), отраслевом и корпоративном уровнях. 2. Анализ и разработку стратегии обеспечения технологического паритета и технологического превосходства в условиях ограничений. В отличие от имеющихся исследований представленная методика позволяет: – получать количественные прогнозные оценки развития технологи-

ческих трендов и составляющих технологий, с учетом их связности, на основе конструирования будущих событий; – обеспечивать возможность построения дорожных карт развития технологических трендов и технологий на национальном, отраслевом и корпоративном уровнях в условиях ограничений. В основу проведенного исследования по созданию методического аппарата положен системный подход, при котором стратегия построения дорожных карт развития технологических трендов и технологий рассматривается как цельный комплекс взаимосвязанных элементов, обеспечивающих решение поставленных задач. В результате реализации предложенной процедуры возможно определить: – организационную стратегию достижения технологического паритета; – время достижения технологического паритета; – сформировать перечень технологий развития технологических трендов, обеспечивающих технологическое превосходство.

Технология; технологические тренды; приоритеты; функциональное назначение; дорожная карта.

A.A. Belevtsev, A.M. Belevtsev, V.A. Balyberdin

METHODOLOGY FOR CONSTRUCTING ROADMAPS FOR THE DEVELOPMENT OF TECHNOLOGICAL TRENDS AND TECHNOLOGIES UNDER CONDITIONS OF RESTRICTIONS

Under modern conditions when the transfer to the sixth waves of innovation is realized and the global competition is escalated the problem of technological parity and technological superiority becomes very important for various levels of review. The main way to solve this problem is to develop the science investigations and first of all – to develop the investigations in the area of creation the new and effective technologies everywhere. The roadmap procedure for technological trends and technologies development under information incompleteness and fuzziness is represented. The procedure provides the carrying of analysis for world technological trends in the set of predetermined areas, the definition of possible technological development alternatives and the strategy for technological parity and technological superiority achievement under predetermined restrictions. The realization of the procedures proposed allows to determine the strategy to get the technological parity for various levels of review as well as to form the list for technological trends to insure the technological superiority. It is determined that the roadmap methodology development for technological trends and technologies under restrictions must include: 1. The development analyses of the worldwide technological trends for the predetermined subject areas in the national (namely in Russia), branch and corporativ levels. 2. The analyses and construction of the strategy for support the technological priority and technological superiority under existing restrictions. In difference of other investigations the exist methodology submitted permits: – to build the quantity forecasts estimations for technological trends and compiling technologies on the base of future events constructing; – to support the possibilities to construct the roadmaps for technological trends and technologies on national, branch and corporation levels under the restrictions. As a bases of the investigation made on the methodology creation the system approach is used when the strategy for technological trends and technologies roadmaps construction is considered as a whole complex of associated elements made to solve the tasks put. As a result of the supposed procedure realization there is a chance to define: – the organization strategy to reach the technological parity; – the time to reach the parity; – the methodology list to form the technological trends to reach the technological superiority.

Technologies; technological trends; priorities; functional duty; roadmap.

Введение. В условиях перехода к шестому технологическому укладу, наличие неопределенностей, связанных с формированием направлений и траекторий развития технологий нового поколения, смены социально-экономической модели мировой экономики и резкого обострения глобальной конкуренции задача обеспечения технологической суверенитета и технологического превосходства приобретает первостепенное значение.

В этой связи задача разработки собственной стратегии развития на национальном и корпоративном уровнях с учетом прогноза общемирового научно-технологического развития в условиях нечеткости и неполноты информации встречает значительные затруднения [2, 4, 5–10], в этой связи является исключительно актуальной.

Информационно аналитической основой для решения поставленной задачи являются [1, 3, 13, 16, 17, 19]:

- ◆ Проведение глобального мониторинга, кластеризация и динамическое формирование технологических трендов по заданным предметным областям.
- ◆ Формирование взаимосвязей, анализируемых технологий, технологических направлений (трендов) и областей применения.
- ◆ Определение новых (виртуальных) технологий и технологических направлений формируемых, как новые элементы анализируемого направления и имеющее причинно-следственные связи с технологиями, выявленными в ходе мониторинга.

В работах [1, 11, 21] разработана методика формирования прогноза развития технологических трендов и построения дорожных карт которая позволяет:

- ◆ получить количественные прогнозные оценки развития технологических трендов и составляющих технологий, с учетом их связности, на основе конструирования будущих событий;
- ◆ провести прогнозные оценки времени реализации технологий и технологических трендов в условиях неопределенности и неполноты информации, отсутствия апробированных количественных оценок и прототипов.
- ◆ построить дорожные карты (этапов) развития технологических трендов и технологий для заданной предметной области и групп предметных областей;

На основе предложенного подхода могут быть получены исходные данные для решения задачи построения дорожных карт развития технологических трендов и технологий на национальном, отраслевом и корпоративном уровнях в условиях ограничений.

Общий подход. На основании проведения исследований и разработок в рассматриваемой области [12, 14, 15, 18, 20, 21] методика построения дорожных карт развития технологических трендов и технологий в условиях ограничений должна обеспечивать:

1. Анализ развития общемировых технологических трендов для заданных предметных отраслей на национальном (на пример в РФ), отраслевом и корпоративном уровнях.
2. Анализ и разработка стратегии обеспечения технологического паритета и технологического превосходства в условиях ограничений.

Постановка задачи. Пусть заданы: уровень проводимого анализа (национальный, отраслевой, корпоративный), предметная область $P = \{P_m\}$ для которой на основе [] сформированы:

- ◆ общемировой технологический тренд развития $C_n^{Pm} = \{c_{ij}^n\} n = \overline{1, N}$
- ◆ сделан переход к логическому графу технологического тренда G_n^{Pm} , а затем к динамическому графу $G_n^{Pm}(t)$ [1].
- ◆ определены характеристики технологии $c_{ij}^n \in C_n^{Pm}$ и представлены в виде картежей.

$$c_{ij}^n \rightarrow \langle \Phi_{i_j}^n, t_{ij}^n, S_{ij}^n, \rho_{ij}^n \rangle \quad (1)$$

- ◆ построена дорожная карта прогноза развития технологического тренда в заданной предметной области - D_n^{Pm} , которая представляет собой прогноз объективного развития технологий в заданной предметной области.

- ◆ определен вектор приоритетов ρ_{ij}^n для всех технологий c_i технологического тренда.

Последовательность этапов решения задачи. Рассмотрим алгоритм реализации первого этапа методики, который можно представить в виде следующей последовательности действий:

1. Выбираем предметную область P_m для которой построены:

- ◆ $G_n^{Pm}(t)$ и D_n^{Pm} [1] (рис. 1) и определены характеристики технологии c_{ij}^n (табл. 1)

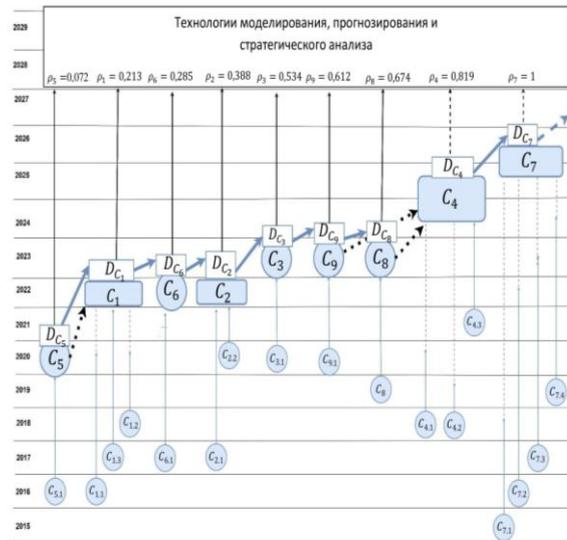


Рис. 1. Динамический граф и дорожная карта прогноза развития технологического тренда $G_n^{Pm}(t)$ за рубежом

Таблица 1

Характеристика технологий c_i технологического тренда $G_n^{Pm}(t)$ и D_n^{Pm}

C_n	C_n	ΦH_i^n	t_i^n [год]	ρ_i^n
C_1	Автоматизация средств анализа поддержки принятия решений человеком с минимальным негативным элементом предвзятости	ΦH_1^n	6	$\rho_1 = 0,141$
C_2	Технологии прогнозирование	ΦH_2^n	4	$\rho_2 = 0,103$
C_3	Технологии построения эмпирических моделей	ΦH_3^n	6	$\rho_3 = 0,146$
C_4	Технологии моделирования социального поведения	ΦH_4^n	6	$\rho_4 = 0,146$
C_5	Решение аналитических задач на основе краусорсинга	ΦH_5^n	3	$\rho_5 = 0,072$
C_6	Технологии идентификации угроз	ΦH_6^n	3	$\rho_6 = 0,072$
C_7	Технологии интерпретации событий	ΦH_7^n	8	$\rho_7 = 0,179$
C_8	Прогнозирование контрфактуалов (альтернативных событий)	ΦH_8^n	3	$\rho_8 = 0,061$
C_9	Технологии социальной инженерии	ΦH_9^n	3	$\rho_9 = 0,078$

2. На основании анализа $G_n^{Pm}(t)$ и D_n^{Pm} определены 4 списка технологий (L_1-L_4) (табл. 2):

- ◆ технологий, разработка которых проводится в РФ.
Формирование Списка 1 - L_1
- ◆ оригинальных технологий, которые разрабатываются в РФ для которых справедливо $c_{ij} \neq c_{ij}^*$ и $\Phi H_{ij} \equiv \Phi H_{ij}^*$.
Формирование Списка 2 - L_2

♦ оригинальные технологии, которые разрабатываются в РФ и не разрабатываются за рубежом, для которых справедливо $c_{ij} \neq c_{ij}^*$ и $\Phi H_{ij} \neq \Phi H_{ij}^*$.

Формирование Списка 3 - L_3

♦ определение перечня технологий, работы по которым в РФ не проводятся

Формирование Списка 4 - L_4

Таблица 2

Списки технологий технологического тренда $G_n^{Pm}(t)$ для РФ

Список L_1 :	$C_2(c_{2.2}), C_5(c_{5.2})$
Список L_2 :	$C_3(c_{3.2}), C_2(c_{2.3}, c_{2.4}), C_1(c_{1.4}, c_{1.5}, c_{1.6}), C_6(c_{6.2}), C_9(c_{9.2}), C_4(c_{4.4}, c_{4.5}, c_{4.6})$
Список L_3 :	$C_{10}(c_{10.1}), C_{11}(c_{11.1}, c_{11.2}), C_{12}(c_{12.1}, c_{12.2}),$
Список L_4 :	$C_2(c_8), C_7(c_{7.1}, c_{7.2}, c_{7.3}, c_{7.4})$

3. Проводим оценку приоритетов для технологий вошедших в Список 3 [11].

Для технологий из списков L_1, L_2 и L_4 сохраняются приоритеты, рассчитанные на предыдущих этапах анализа в соответствии.

4. Проводим оценку времени и стоимости работ по созданию технологий, вошедших в списки L_1, L_2 и L_4 .

5. Формируем текущую проекцию тренда на его развитие в РФ.

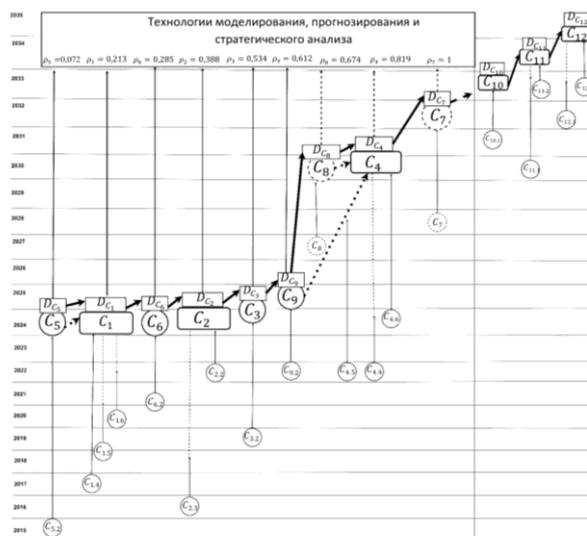


Рис. 2. Проекция динамического графа и дорожной карты развития технологического тренда P_m в РФ

6. В соответствии с методикой [1] осуществляется построение $\bar{G}_n^{Pm}(t)$ и дорожной карты \bar{D}_n^{Pm} развития технологического тренда C_n^{Pm} в РФ (рис. 2).

7. Переформатируем список L_4 в список L_4^* , в котором размещаем технологии в порядке уменьшения значений вектора приоритетов.

8. Формируем вектор \bar{X}_n , который определяет количество вхождений технологий c_i^n в технологии более высокого ранга.

Определение значений вектора \bar{X}_n и проводится на основе матрицы инцидентий K_n логического графа G_{ij}^{Pm} .

$$K_n = \{K_{ij}^n\} \text{ порядка } I \times (I + L),$$

где $K_{ij}=1$, если имеется дуга, соединяющая i вершину с j вершиной, т.е. технология i должна быть реализована до того, как начнутся исследования по технологии j .

Если такое условие отсутствует, то $K_{ij}^n=0$.

В силу того, что технологии $\{c_{ij}^n\}$ технологического тренда C_n^{Pm} могут иметь связи не только с технологиями внутри тренда, но и с технологиями входящими в состав других технологических трендов C_{n+m}^{Pm} , размерность матрицы инцидентий равна $I \times (I + L)$, где L общее вхождение технологий в другие тренды.

На основании матрицы K_n можно построить вектор \bar{X}_n определяющий количество вхождений технологий c_{ij}^n в технологии более высокого ранга в составе анализируемого тренда, так и других трендов предметной области P .

9. Проводим анализ дорожных карт развития технологического тренда D_n^{Pm} за рубежом и \bar{D}_n^{Pm} в РФ в условиях отсутствия ограничений на ресурсы для реализации технологий списка L_4 .

В результате получаем оценки:

- ◆ изменений функциональной полноты по годам;
- ◆ изменения значений вектора приоритетов по годам;
- ◆ времени отставания (или опережения) развития траектории обоих дорожных карт (рис. 3)

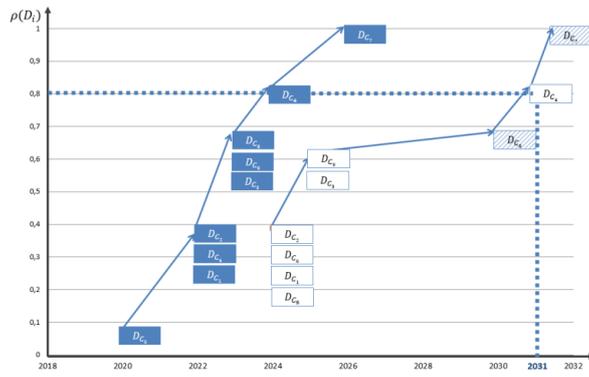


Рис. 3. Дорожная карта развития технологического тренда P_m в РФ и за рубежом

В силу того, что каждый этап D_{c_i} реализации дорожной карты будет представлять собой создание технологий или взаимосвязанных групп технологий со своим значением вектора приоритетов, то

$$\sum \rho_{ij}(t) \rightarrow 1. \tag{2}$$

Тогда меру отставания траекторий развития технологических трендов $G_n^{Pm}(t)$ и $\bar{G}_n^{Pm}(t)$ можно определить на основе выражения

$$\sum \rho_{ij}(t) - \sum \bar{\rho}_{ij}(t). \tag{3}$$

Переход к п.1. Анализ технологических трендов для следующей области применения P_{m+1} , предметной области P .

8. Формирование, упорядоченных списков технологий – списков L_4^* , полученных для всех областей применения предметной области $P = \{P_m\}$ с учетом изменения значений их приоритетов.

9. Формирование векторов \bar{X}_n полученных для всех технологических трендов всех областей применения предметной области $P = \{P_m\}$

Таким образом, в результате выполнения первого этапа будут сформированы:

- ◆ Списки – L_4 , и L_4^* ;
- ◆ вектор \bar{X}_n определяющего количество вхождений технологий c_{ij}^n в технологии более высокого ранга.

Результаты выполнения первого этапа являются основой для выполнения второго этапа методики.

Рассмотрим алгоритм реализации второго этапа.

Цель – разработка стратегии достижения паритета и технологического превосходства в условиях ограничений на ресурсы.

Для создания технологии из списка L_4^* и достижения технологического паритета возможны следующие направления действий:

1. Разработка в РФ аналога технологии с вариантами:

- ◆ ФН, ρ соответствие мировому уровню;
- ◆ f , ФН, ρ приближение к мировому уровню;

Следует отметить, что затраты и время реализации отдельных технологий зависят от степени их готовности. При этом возможна замена некоторых технологий их имеющимися аналогами с определенной потерей качества реализуемых процессов.

2. Разработка технологий в РФ, которые имеют $\Phi H_{ij} \equiv \Phi H_{ij}^*$ при $c_{ij} \neq c_{ij}^*$ т.е. другой технологический способ реализации - (список L_2).

3. Разработка технологий, которых нет в общемировых трендах, но которые имеют новое функциональное назначение $\Phi H_{ij}^* \notin \{\Phi H_{ij}\}$. При этом они расширяют функциональные возможности технологии и всего технологического тренда в целом и которые можно рассматривать как прямые альтернативы технологии из Списка L_4^*

4. Разработка технологий, которых нет в общемировых трендах, но которые имеют новое функциональное назначение $\Phi H_{ij}^* \notin \{\Phi H_{ij}\}$ и не являются прямыми альтернативами технологий из Списка L_4 . При этом они расширяют функциональные возможности всего тренда в целом и при наличии паритета с общемировыми трендами могут обеспечить технологическое превосходство.

В этой связи встает задача разработки стратегии достижения паритета и определение возможных траекторий обеспечения технологического превосходства.

Общая процедура решения этой задачи может быть представлена в следующем виде (рис. 4)

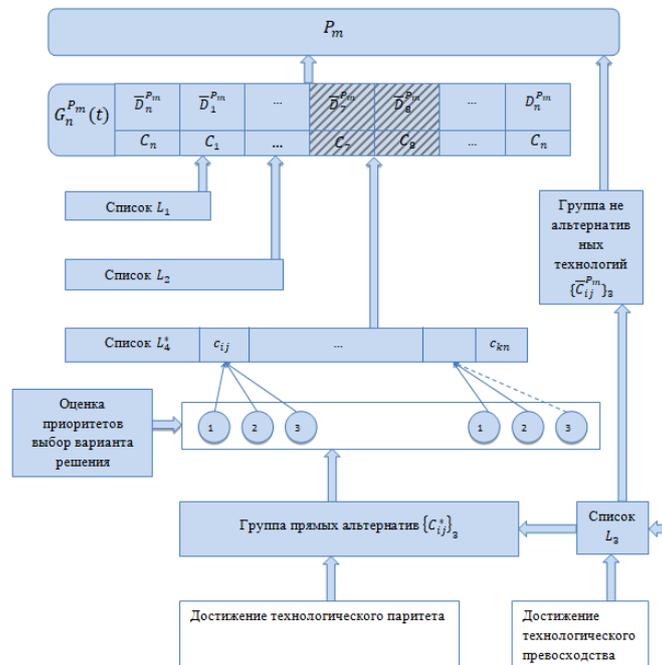


Рис. 4. Общая процедура достижения технологического паритета и технологического превосходства.

1. Для каждой технологии c_{ij}^n Списка L_4^* проверяем наличие технологий в Списке L_3 ($c_{ij} \neq c_{ij}^*$ и $\Phi H_{ij} \neq \Phi H_{ij}^*$) с уровнем готовности $TRL=\{1,2\}$, которые можно рассматривать как её прямую альтернативу и включаем в состав группы прямых альтернатив $\{C_{ij}^*\}_3$ для каждой c_{ij} (рис. 4).

2. Для каждой технологии Списка L_4^* сформируем группы возможных вариантов создания технологий $\{1, 2, 3\}$ или $\{1,2\}$ (в случае отсутствия предложений в группе прямых альтернатив) и по критерию $\{\Phi H_{ij}, t_{ij}, S_{ij}\}$ определяем вектор приоритетов [11]. Принимаем решение о вариантах её создания способами $\{1,2,3\}$.

3. В случае принятия решения о реализации хотя бы одной технологии c_{ij} по варианту 3, то для всей совокупности технологий тренда $\bar{C}_n^{Pm}(t)$ проводится перерасчет вектора приоритетов $\bar{\rho}_{ij}^*(C_i)$.

4. Для технологий списка L_4^* проводим оценку времени реализации t_{ij}^* [21]

5. Формируем группу вершин $\{D_i^*(C_i)\}$ в порядке уменьшения $\bar{\rho}_{ij}^*(C_i)$.

На графе \bar{D}_i^{Pm} (рис.2) выделяем подграфы для вершин $D_i^*(C_i)$ списка L_4^* .

6. Для группы $\{D_i^*(C_i)\}$ формируем группу $\{c_{ij}\}$ с учетом ограничений

$$\sum_i S_4(c_{ij}) \leq S_{max} . \quad (4)$$

где S_{max} – максимальное значение общей стоимости этапа.

7. Задаем момент времени начала работ t_{n1} .

8. Для всех технологий, вошедших в первую группу $\{c_{ij}\}$, начинаем пошаговое движение по подграфам $D_i^*(C_i)$ снизу вверх (шаг – год) [1].

9. На каждом шаге контролируем завершение работ по технологиям по спискам L_1 и L_2 и определяем объем дополнительных средств

$$S_4(t) = S_{4max} + \Delta S_{1,2} . \quad (5)$$

10. В случае выполнения п.9, появление дополнительного финансирования в состав стартовой группы $\{c_{ij}\}$ включаем следующую технологию из списка L_4^* .

Фиксируем момент времени начала работ на созданием очередной технологии t_{n2} .

11. Выполняем п.8 – п.10 и определяем момент времени достижения $C_i \in \{D_n^*(C_i)\}$.

12. Включаем в состав группы $\{c_{ij}\}$ следующую технологию и переходим к п.8.

13. Определяем текущее изменение значения вектора приоритетов $\bar{\rho}_{ij}^*(C_i)$ всего технологического тренда $D_n^*(C_i)$. Если $\sum \bar{\rho}_{ij}^* = 0,8$, то фиксируем момент времени достижения технологического паритета.

14. Фиксируем все значения времени, полученные для всех $C_i \in \{D^*(C_i)\}$, и проводим корректировку всей дорожной карты $\bar{D}_n^*(t)$.

Дорожная карта работ по группе $\{C_{ij}^{Pm}\}_3$ не альтернативных технологий может быть организована по аналогичной процедуре. Время её реализации будет определяться выделенными ресурсами.

Списки L_3 формируются для всех технологических трендов. Объединение этих списков позволит получить список L_5 , который будет содержать перечень технологий для формирования научно-технического задела и определения новых трендов предметной области $P = \{P_m\}$.

Заключение. В результате проведенных исследований определены основные положения перспективной методики построения дорожных карт развития технологических трендов и технологий в условиях ограничений.

В отличие от имеющихся исследований представленная методика позволяет:

- ◆ получать количественные прогнозные оценки развития технологических трендов и составляющих технологий, с учетом их связности, на основе конструирования будущих событий;

♦ обеспечивать возможность построения дорожных карт развития технологических трендов и технологий на национальном, отраслевом и корпоративном уровнях в условиях ограничений.

В основу проведенного исследования по созданию методического аппарата положен системный подход, при котором стратегия построения дорожных карт развития технологических трендов и технологий рассматривается как цельный комплекс взаимосвязанных элементов, обеспечивающих решение поставленных задач.

В результате реализации предложенной процедуры возможно определить:

- ♦ организационную стратегию достижения технологического паритета;
- ♦ время достижения технологического паритета;
- ♦ сформировать перечень технологий развития технологических трендов, обеспечивающих технологическое превосходство.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Белевцев А.А., Белевцев А.М., Бальбердин В.А.* Методика прогнозирования развития технологических трендов и построения дорожных карт на основе конструирования будущих событий // Известия ЮФУ. Технические науки. – 2023. – № 3. – С. 56-64.
2. Методика формирования национального плана развития науки и технологий в авиастроении. – <https://mai.ru/upload/iblock/b42/b428c607ca999dbfe1f2f12c84cff0a.pdf>. – 2020. – 23 с.
3. *Анферова М.С., Белевцев А.А., Белевцев А.М.* Методика анализа развития зарождающихся технологий и технологических фронтов // Известия ЮФУ. Технические науки. – 2023. – № 4. – С. 87-96.
4. *Буренок В.М., Ивлев А.А., Корчак В.Ю.* Развитие военных технологий XXI века: Проблемы, планирование, реализация. – Тверь: Изд. Купол, 2009. – 624 с.
5. *Saritas O.* Systemic foresight methodology. In D. Meissner, L. Gokhberg, & A. Sokolov (Eds.) // Foresight and science, technology and innovation policies: Best practices. – Berlin: Springer, 2013. – P. 83-117.
6. *Shibata N., Kajikawa Y., & Sakata I.* Detecting potential technological fronts by comparing scientific papers and patents // Foresight. – 2011. – 13 (5). – P. 51-60.
7. *Fuller C.M., Biros D.P. and Delen D.* An investigation of data and text mining methods for real world deception detection // Expert Systems with Applications. – 2011. – No. 38. – P. 8392-8398.
8. *Cobo M.J., Lopez-Herrera A.G., Herrera-Viedma E., Herrera F.* An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field // Journal of Informetrics. – 2011. – Vol. 5. – P. 146-166.
9. *Gokhberg L., Fursov K., Miles I., Perani G.* Developing and using indicators of emerging and enabling technologies // Handbook of Innovation Indicators and Measurement / Ed. F. Gault. Cheltenham: Edward Elgar. – 2013. – P. 349-380.
10. *Palomino M.A., Vincenti A., Owen R.* Optimising web-based information retrieval methods for horizon scanning // Foresight. – 2013. – Vol. 15, No. 3. – P. 159-176.
11. *Белевцев А.А., Белевцев А.М., Бальбердин В.А.* Методика анализа и оценки приоритетов технологических трендов и технологий // Известия ЮФУ. Технические науки. – 2022. – № 6.
12. *Бальбердин В.А., Белевцев А.М., Бендерский Г.П.* Прикладные методы оценки и выбора решений в стратегических задачах инновационного менеджмента. – 5-е изд. – М.: ИТК «Дашков и К», 2022. – 340 с.
13. *Белевцев А.М., Бальбердин В.А., Белевцев А.А., Садреев Ф.Г.* О разработке интеллектуальных сервисов мониторинга технологических трендов в информационно-аналитических комплексах // Научные технологии. – 2019. – № ..
14. *Белевцев А.М., Бальбердин В.А., Белевцев А.А.* Об оценке направлений развития информационных технологий для систем сетевидного управления // Матер. Всероссийской научно-технической конференции с международным участием «Компьютерные и информационные технологии в науке, инженерии и управлении»: Сб. докладов. Таганрог «КомТех 2017». – 2017. – С. 8-12.
15. *Белевцев А.М., Бальбердин В.А., Белевцев А.А., Маркелов Е.Б.* Некоторые тенденции развития информационных технологий для систем сетевидного управления // Вооружение и экономика. – 2021. – № 3.
16. *Саати Т.Л.* Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1993. – 340 с.
17. *Саати Т.Л.* Принятие решений при зависимостях и обратных связях. Аналитические сети. – М.: ЛИБРОКОМ, 2009. – 358 с.

18. Белевцев А.М., Бальбердин В.А., Бендерский Г.П., Белевцев А.А. Анализ направлений развития нано- и IT-технологий для построения специализированных сетевых коммуникационных систем нового поколения // Известия ЮФУ. Технические науки. – 2015. – № 3.
19. Белевцев А.М., Садреев Ф.Г., Пахомов С.Н. Организация параллельного мониторинга направлений развития науки, техники и технологий в открытых источниках информации // Известия ЮФУ. Технические науки. – 2013. – № 5. – С. 86-95.
20. Kim Y., Tian Y., Jeong Y., Jihee R., & Myaeng S.-H. Automatic discovery of technology trends from patent text // Proceedings of the 2009 ACM symposium on applied computing. 2009. – P. 1480-1487.
21. Белевцев А.М., Бальбердин В.А., Белевцев А.А. Методика оценки времени и стоимости реализации технологических трендов в условиях неопределенности и не полноты информации // Научно-технические технологии. – 2019. – № 5.

REFERENCES

1. Belevtsev A.A., Belevtsev A.M., Balyberdin V.A. Metodika prognozirovaniya razvitiya tekhnologicheskikh trendov i postroeniya dorozhnykh kart na osnove konstruirovaniya budushchikh sobyitiy [Methodology for forecasting the development of technological trends and building road map on the basis of constructing future events], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2023, No. 3, pp. 56-64.
2. Metodika formirovaniya natsional'nogo plana razvitiya nauki i tekhnologiy v aviastroenii [Methodology for the formation of a national plan for the development of science and technology in the aircraft industry]. Available at: <https://mai.ru/upload/iblock/b42/b428c607ca999dbfe1f2f12c84cfff0a.pdf>, 2020, p. 23.
3. Anferova M.S., Belevtsev A.A., Belevtsev A.M. Metodika analiza razvitiya zarozhdayushchih tekhnologiy i tekhnologicheskikh frontov [Methodology for analyzing the development of emerging technologies and technological fronts], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2023, No. 4, pp. 87-96.
4. Burenok V.M., Ivlev A.A., Korchak V.Yu. Razvitie voennykh tekhnologiy XXI veka: Problemy, planirovanie, realizatsiya [Development of military technologies of the XXI century: Problems, planning, implementation]. Tver': Izd-vo «Kupol», 2009, 624 p.
5. Saritas O. Systemic foresight methodology. In D. Meissner, L. Gokhberg & A. Sokolov (Eds.), *Foresight and science, technology and innovation policies: Best practices*. Berlin: Springer, 2013, pp. 83-117.
6. Shibata N., Kajikawa Y., & Sakata I. Detecting potential technological fronts by comparing scientific papers and patents, *Foresight*, 2011, 13 (5), pp. 51-60.
7. Fuller C.M., Biros D.P. and Delen D. An investigation of data and text mining methods for real world deception detection, *Expert Systems with Applications*, 2011, No. 38, pp. 8392-8398.
8. Cobo M.J., Lopez-Herrera A.G., Herrera-Viedma E., Herrera F. An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field, *Journal of Informetrics*, 2011, Vol. 5, pp. 146-166.
9. Gokhberg L., Fursov K., Miles I., Perani G. Developing and using indicators of emerging and enabling technologies, *Handbook of Innovation Indicators and Measurement* / ed. F. Gault. Cheltenham: Edward Elgar, 2013, pp. 349-380.
10. Palomino M.A., Vincenti A., Owen R. Optimising web-based information retrieval methods for horizon scanning, *Foresight*, 2013, Vol. 15, No. 3, pp. 159-176.
11. Belevtsev A.A., Belevtsev A.M., Balyberdin V.A. Metodika analiza i otsenki prioritetrov tekhnologicheskikh trendov i tekhnologiy [Methods of analysis and evaluation of the priorities of technological trends and technologies], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2022, No. 6.
12. Balyberdin V.A., Belevtsev A.M., Benderskiy G.P. Prikladnye metody otsenki i vybora resheniy v strategicheskikh zadachakh innovatsionnogo menedzhmenta [Applied Methods of Evaluation and Choice of Solutions in the Strategic Problems of Innovation Management]. 5th ed. Moscow: ITK «Dashkov i K», 2022, 340 p.
13. Belevtsev A.M., Balyberdin V.A., Belevtsev A.A., Sadreev F.G. O razrabotke intellektual'nykh servisov monitoring tekhnologicheskikh trendov v informatsionno-analiticheskikh kompleksakh [On the development of intelligent services for monitoring technological trends in information and analytical complexes], *Naukoemkie tekhnologii* [Science-intensive technologies], 2019, No. 3.
14. Belevtsev A.M., Balyberdin V.A., Belevtsev A.A. Ob otsenke napravleniy razvitiya informatsionnykh tekhnologiy dlya sistem setetsentricheskogo upravleniya [On the assessment of directions for the development of information technologies for network-centric control systems], *Mater. Vserossiyskoy nauchno-tekhnicheskoy konferentsii s mezhdunarodnym uchastiem «Komp'yuternye i informatsionnye*

- tekhnologii v nauke, inzhenerii i upravlenii: Sb. dokladov* [Proceedings of the All-Russian Scientific and Technical Conference with International Participation "Computer and Information Technologies in Science, Engineering and Management: Collection of Reports]. Taganrog, 2017, pp. 8-12.
15. Belevtsev A.M., Balyberdin V.A., Belevtsev A.A., Markelov E.B. Nekotorye tendentsii razvitiya informatsionnykh tekhnologiy dlya sistem setetsentricheskogo upravleniya [Some trends in the development of information technologies for network-centric control systems], *Vooruzhenie i ekonomika* [Armament and Economics], 2021, No. 3.
 16. Saati T.L. Prinyatie resheniy. Metod analiza ierarkhiy [Making decisions. Hierarchy analysis method]. Moscow: Radio i svyaz', 1993, 340 p.
 17. Saati T.L. Prinyatie resheniy pri zavisimostyakh i obratnykh svyazyakh. Analiticheskie seti [Decision making under dependencies and feedbacks. Analytical networks]. Moscow: LIBROKOM, 2009, 358 p.
 18. Belevtsev A.M., Balyberdin V.A., Benderskiy G.P., Belevtsev A.A. Analiz napravleniy razvitiya nano-IT-tekhnologiy dlya postroeniya spetsializirovannykh setevykh kommuni-katsionnykh sistem novogo pokoleniya [Analysis of the directions of development of nano-and IT-technologies for the construction of specialized network communication systems of a new generation], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, No. 3.
 19. Belevtsev A.M., Sadreev F.G., Pakhomov S.N. Organizatsiya parallel'nogo monitoringa napravleniy razvitiya nauki, tekhniki i tekhnologiy v otkrytykh istochnikakh informatsii [Organization of parallel monitoring of directions of development of science, technology and technology in open sources of information], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2013, No. 5, pp. 86-95.
 20. Kim Y., Tian Y., Jeong Y., Jihee R., & Myaeng S.-H. Automatic discovery of technology trends from patent text, *Proceedings of the 2009 ACM symposium on applied computing*, 2009, pp. 1480-1487.
 21. Belevtsev A.M., Balyberdin V.A., Belevtsev A.A. Metodika otsenki vremeni i stoimosti realizatsii tekhnologicheskikh trendov v usloviyakh neopredelennosti i nepolnoty informatsii [Methodology for estimating the time and cost of implementing technological trends in conditions of uncertainty and incompleteness of information], *Naukoemkie tekhnologii* [Science-intensive technologies], 2019, No. 5.

Статью рекомендовал к опубликованию д.т.н. Ю.А. Кравченко.

Белевцев Андрей Андреевич – ПАО«Сбербанк»; e-mail: andrey.belevtsev@gmail.com; г. Москва, Россия; старший вице-президент.

Белевцев Андрей Михайлович – Московский авиационный институт (национальный исследовательский университет); e-mail: ambelevtsev@yandex.ru; г. Москва, Россия; тел.: +79037691788; д.т.н.; профессор.

Балыбердин Валерий Алексеевич – 3 Центральный научно-исследовательский институт МО РФ; e-mail: balyberdinvaleri@yandex.ru; г. Москва, Россия; д.т.н.; профессор; заслуженный деятель науки РФ; в.н.с.

Belevtsev Andrey Andrevich – PJSC "Sberbank"; e-mail: andrey.belevtsev@gmail.com; Moscow, Russia; senior vice president.

Belevtsev Andrey Mikhaylovich – Moscow Aviation Institute (National Research University); e-mail: ambelevtsev@yandex.ru; Moscow, Russia; phone: +79037691788; dr. of eng. sc.; professor.

Balyberdin Valeriy Alexeevich – 3 Central Research Institute of the Ministry of Defense of the Russian Federation; e-mail: balyberdinvaleri@yandex.ru; Moscow, Russia; dr. of eng. sc.; professor; Honored Scientist of the Russian Federation; leading researcher.

УДК 004.93'1

DOI 10.18522/2311-3103-2024-2-278-289

А.В. Руденко, М.А. Руденко

ПОДДЕРЖКА ПРИНЯТИЯ ВРАЧЕБНЫХ РЕШЕНИЙ ПРИ ПЛАНИРОВАНИИ ПРОВЕДЕНИЯ ПРОЦЕДУРЫ ЛАЗЕРНОЙ ЛИТОТРИПСИИ

При подготовке к проведению процедуры лазерной литотрипсии, выбирая параметры лазерной установки, врач учитывает множество факторов, таких как масса и плотность обнаруженных в почке конкрементов, расположение камней в почке, близость кровеносных сосудов. Также важным

параметром является время воздействия на камень лазерным лучом до разрушения камня. При этом расчет времени разрушения камня – достаточно трудоемкая процедура, время разрушения зависит от массы камня и параметров энергии лазера и его частоты. Поэтому актуальным является создание системы по поддержке принятия врачебных решений при проведении процедуры лазерной литотрипсии, позволяющей выполнить расчет времени разрушения камней и осуществить выбор значений параметров лазера. В статье предлагается алгоритм поддержки выбора врачом-урологом режима работы лазера во время проведения процедуры лазерной литотрипсии при лечении мочекаменной болезни человека, являющейся частью системы поддержки принятия врачебного решения в хирургии и урологии с использованием технологий компьютерного зрения. Предложенный алгоритм нечеткой оценки параметров лазера при выборе его режима работы в зависимости от массы камня и выбранного времени разрушения камня и других факторов (распределение плотности камня, расположение камня в почке, близость стенок и сосудов) осуществляет формирование рекомендаций по установлению параметров работы лазера. Система поддержки принятия врачебного решения позволила сократить время на принятие решения врачом, избежать ошибок при выборе параметров работы лазерной установки по дроблению камней в почках.

Поддержка принятия решения; нечеткая оценка; лазерная литотрипсия; масса камня; время разрушения; параметры лазера.

A.V. Rudenko, M.A. Rudenko

SUPPORT FOR MEDICAL DECISION-MAKING WHEN PLANNING THE LASER LITHOTRIPSY PROCEDURE

When preparing for the laser lithotripsy procedure, choosing the parameters of the laser installation, the doctor considers many factors, such as the mass and density of concretions found in the kidney, the location of kidney stones, and the proximity of blood vessels. Another important parameter is the time of exposure to the stone with a laser beam before the stone is destroyed. At the same time, calculating the time of destruction of a stone is a rather time-consuming procedure, the time of destruction depends on the mass of the stone and the parameters of the laser energy and its frequency. Therefore, it is relevant to create a system to support medical decision-making during the laser lithotripsy procedure, which allows you to calculate the time of stone destruction and select the values of laser parameters. The article proposes an algorithm to support the choice of the laser operating mode by a urologist during the laser lithotripsy procedure in the treatment of human urolithiasis, which is part of the medical decision support system in surgery and urology using computer vision technologies. The proposed algorithm for fuzzy estimation of laser parameters when choosing its operating mode, depending on the mass of the stone and the selected time of destruction of the stone and other factors (distribution of stone density, location of the stone in the kidney, proximity of walls and vessels) generates recommendations for setting the parameters of the laser. The medical decision support system made it possible to reduce the time for a doctor to decide, to avoid mistakes when choosing the parameters of the laser installation for crushing kidney stones.

Fuzzy estimation; laser lithotripsy; stone mass; destruction time; laser energy; laser frequency.

Введение. Врачам при работе с пациентами во время постановки диагноза и выбора правильного лечения или планирования хирургического вмешательства приходится учитывать большое число факторов. В помощь приходят информационные компьютерные технологии в виде различных систем поддержки принятия решений.

Под системой поддержки принятия врачебных решений (СППВР) понимается программное обеспечение, позволяющее собирать, анализировать информацию, влиять на выбор врачом решения при постановке диагноза, выбора лечения с целью снижения ошибок и повышения качества оказываемой медицинской помощи [1]. СППВР может представлять собой сложный программный комплекс, использующий математические модели, компьютерное зрение, искусственный интеллект, базы данных.

На настоящий момент разработано и внедряется в медицинскую практику достаточное большое число разнообразных систем ППВР. В большинстве своем – это справочные системы, системы ведения медицинской документации, системы поддержки диагностики отдельных заболеваний.

Целью исследования являлась разработка СППВР при планировании хирургического вмешательства при лечении мочекаменной болезни (МКБ). Авторами был проведен анализ публикаций по теме СППВР в медицине по лечению МКБ, был сделан вывод о том, что СППВР в данном направлении медицины в большинстве своем находятся в стадии разработки и проектирования [2–13]. В настоящий момент отсутствуют системы, которые бы помогали врачу-урологу в планировании хирургического вмешательства с применением лазерной литотрипсии для удаления камней в почках при МКБ.

Лазерная литотрипсия - дробление камня почки или мочеточника с использованием различных видов энергии, является современным методом лечения мочекаменной болезни, лазер эффективно удаляет камни во всех участках мочевыводящей системы, позволяет вылечить пациентов при значительно небольшой травматичности самой операции [8, 9, 11, 14]. В ходе процедуры лазерные импульсы генерируют ударную волну, которые на фоне локального повышения температуры разрушают камни потенциально до мелких частиц [10, 11, 15–18].

До проведения процедуры лазерного дробления камней при лечении МКБ врачам необходимо знать массу найденного камня, распределение плотности камня от центра к периферии, расположение камня в почке, определить параметры энергии и частоты работы лазера, которые зависят от времени разрушения камня, которое, в свою очередь, зависит от перечисленных выше факторов. Правильное определение параметров работы лазера определяет безопасность пациента во время операции. Поэтому существует необходимость разработки системы поддержки принятия врачебного решения для выбора оптимального и безопасного воздействия лазером на камень, находящийся внутри почки, чтобы его разрушение происходило за минимально короткое время, при этом исключалась возможность нанесения травм внутренним органам пациента.

Описание метода. Выбор режима работы лазера при проведении процедуры лазерной литотрипсии определяется следующими факторами: временем, в течение которого произойдет разрушение камня, расположением камня внутри почки, его формой, значением плотности и распределением плотности по объему камня, близостью кровеносных сосудов, близостью тонких стенок внутренних органов.

Для определения величин энергии и частоты работы лазера, а также для формирования рекомендаций по выбору режимов работы при проведении операции лазерной литотрипсии, режимы работы лазерной установки можно условно распределить на следующие категории: режим для дробления осколков камней – характеризуется низким значением величины энергии лазера и низкой частотой; режим работы для камней с низкой плотностью – характеризуется низким значением величины энергии лазера и более высокими применяемыми частотами; режим работы при неудобных подходах инструмента к камню – характеризуется значением величины энергии лазера чуть ниже среднего и средними и чуть выше частотами работы лазера; оптимальный режим работы – применяется в большинстве процедур дробления камней и характеризуется средними значениями величин энергии лазера и частоты работы лазера; режим работы в зоне высокой плотности камня – характеризуется значением энергии лазера выше среднего, значения частоты лазера находятся в диапазоне средних и чуть выше средних; режим работы с высоким риском нанесения травмы – характеризуется значениями энергии лазера выше среднего и гораздо выше среднего, частотами работы лазера выше среднего.

Для формирования рекомендаций по выбору режима дробления камней в почках и установлению величин энергии и частоты работы лазерной установки в данном исследовании был использован нечеткий логический вывод. Данный метод формирования рекомендаций был выбран из-за неоднозначности определения режима работы лазерной установки врачом, который при установлении величин энергии и частоты работы лазера отталкивается не только от времени разрушения камня, зависящего от массы камня, но и от расположения камня внутри почки, близостью кровеносных сосудов, расположением самой почки в теле пациента, размерами камней и их количеством, а также диапазонами значений энергии лазера и частоты его работы. Поэтому однозначно установить величину энергии лазера и частоты его работы по рассчитанной массе камня не представляется возможным.

Алгоритм формирования рекомендаций по нечеткому выбору категории режима работы лазера из указанных выше и установлению величин энергии и частоты его работы использует нечеткий вывод, который определяется функциями принадлежности двух переменных: энергия лазера и частота. Были определены следующие наборы градаций для указанных переменных.

Для энергии лазера: очень высокая – 0,8–0,9 от максимальной; нормальная – 0,6–0,8 от максимальной; щадящая – 0,45–0,6 от максимальной; низкая – ниже 0,45 от максимальной.

Для частоты: очень высокая – 0,9–1 от максимальной; высокая – 0,75–0,9 от максимальной; нормальная – 0,5–0,75 от максимальной; щадящая 0,35–0,5; низкая – ниже 0,35 от максимальной.

Время разрушения камня зависит от массы камня, выбранной энергии и частоты работы лазера. Расчет времени разрушения камня лазером выполняется по следующей формуле [19]:

$$t_{\text{разр}} = \frac{m_{\text{камня}}}{(\gamma * E * F)}, \quad (1)$$

где $m_{\text{камня}}$ – масса камня, грамм; γ – коэффициент гамма, равный 0,4; E – энергия лазера, Дж, задается в диапазоне от 0,6 до 2,4; F – частота излучения лазера, Гц, задается в диапазоне от 8 до 15.

Каждому из режимов работы лазера соответствует определенное значение времени разрушения камня.

Априорное правило является нечетким описанием в n -мерном пространстве факторов \vec{y} , последовательность правил является нечеткой меткой класса из множества R режимов проведения операций [20]:

$$\vec{y} = [y_1, y_2, \dots, y_j, \dots, y_n], j = 1, 2, \dots, n. \quad (2)$$

Здесь n обозначает количество факторов, определяющих выбор режима, которые формируют \vec{y} - входной вектор.

Степень активации i -го правила из множества R вычисляется как:

$$\beta_i(\vec{y}) = \prod_{j=1}^n B_{ij}(y_j), i = 1, 2, \dots, R, \quad (3)$$

где B_{ij} – функции принадлежности нечетких множеств, которые являются нечеткими отношениями вывода i -го правила (режима) и входного вектора \vec{y} .

Вывод классификатора определяет режим, который имеет наивысшую степень активации β_i :

$$z = v(i^*), \quad i^* = \arg \max_{1 \leq i \leq R} \beta_i. \quad (4)$$

Степень уверенности в выбранном режиме задается нормализованной степенью запуска правила:

$$Conf_{i^*} = \frac{\beta_{i^*}}{\sum_{i=1}^R \beta_i}. \quad (5)$$

Данные для проведения исследований и расчётов предоставлены Клиническим медицинским многопрофильным центром имени Святителя Луки г. Симферополь (КММЦ), который обладает современной базой радиологии и уникальными методиками лечения МКБ.

В соответствии с приведенной формулой (1) и данными, полученными от медицинских экспертов, была разработана таблица расчета времени разрушения камней в зависимости от частоты и энергии лазера. Были проведены экспериментальные расчеты с помощью MS Excel.

Исходные данные для расчетов: величина гамма-излучения лазера – 0,4; частота лазера, Гц – изменяется в диапазоне от 8 до 15 с шагом 1; энергия лазера, Дж – изменяется в диапазоне от 0,6 до 2,4 с шагом 0,1.

На рис. 1–3 приведены результаты расчетов времени для камней с массой до 1 грамма, до 10 грамм, свыше 20 грамм (мелкий, средний и крупный камни).

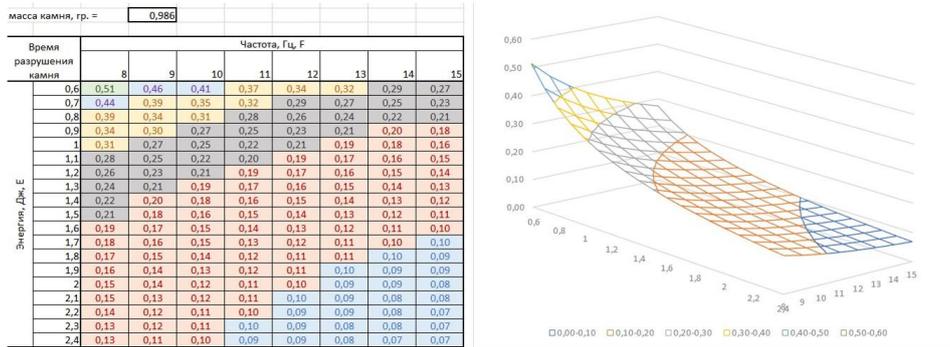


Рис. 1. Расчет времени разрушения камня массой 0,98 грамм во время лазерной литотрипсии при различных режимах работы лазера

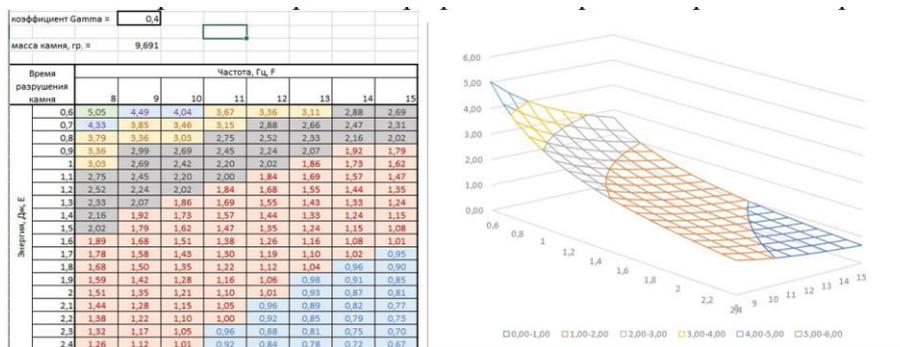


Рис. 2. Расчет времени разрушения камня массой 9,7 грамма во время лазерной литотрипсии при различных режимах работы лазера

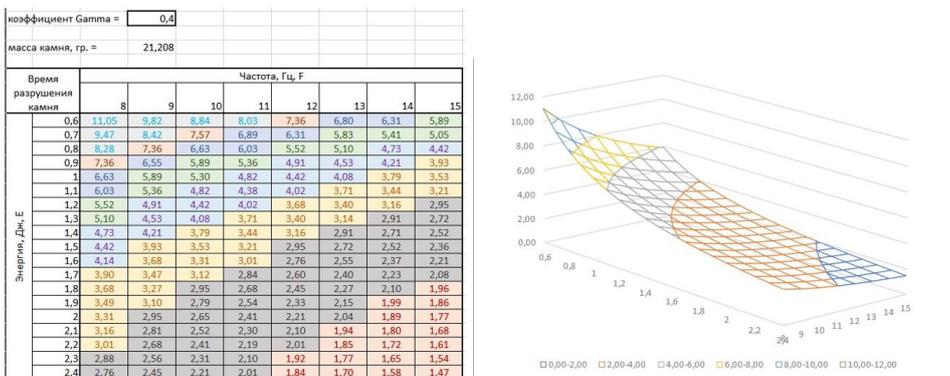


Рис. 3. Расчет времени разрушения камня массой 21,2 грамма во время лазерной литотрипсии при различных режимах работы лазера

Как видно на рисунках, диапазоны значения времени разрушения камней и значения величин энергии лазера и его частоты окрашены в различные цвета, что соответствует установленным ранее режимам работы лазера. Подбирая время разрушения камня, и руководствуясь данными пациента, врач сможет выбрать соответствующий режим работы лазера и установить значения величин энергии и частоты работы на лазерном аппарате.

Результаты. Расчет времени разрушения камня в зависимости от его массы и параметров работы лазера, а также нечеткие рекомендации по подбору режима работы лазерной установки для воздействия на камень в почке были реализованы в программном мо-

дуле поддержки принятия врачебных решений в прототипе системы поддержки принятия врачебных решений в хирургии и урологии с использованием технологий компьютерного зрения. В процессе работы модуля поддержки принятия врачебных решений прототипа системы производится расчет времени разрушения камня в зависимости от его массы и установленных по умолчанию диапазонов параметров работы лазера – значений энергии и частоты, а также в зависимости от величины шагов изменения этих величин и установленной величины гамма для лазера. После проведения расчетов производится формирование таблицы с результатами расчетов.

Для определения массы камня необходимо было вычислить плотность и объем камня.

Для определения плотности камня была использована следующая формула расчета физической плотности камней [20]:

$$\rho_{\text{камень}} = HU_{\text{камень}} * 0,000485 + 1,539, \quad (6)$$

где $HU_{\text{камень}}$ – значение светимости камня по Хаунсфилду (HU), по данным компьютерной томографии.

Значение плотности вычислялось для каждого вокселя, принадлежащего объёму камня.

Объем одного вокселя вычислялся в соответствии с параметрами проведения процедуры компьютерной томографии, сохраненных в DICOM файлах, как произведение толщины среза по координатам X и Y на расстояние между срезами по координате Z:

$$V_{\text{vox}} = x_{\text{thin}} * y_{\text{thin}} * z_{\text{thin}}, \quad (7)$$

где V_{vox} – значение объема одного вокселя, $x_{\text{thin}}, y_{\text{thin}}$ – толщины среза по координатам X и Y, z_{thin} – расстояние между срезами по координате Z.

Объем камня вычислялся путем суммирования объемов всех вокселей в камне:

$$V = \sum_{i=0}^I \sum_{j=0}^J \sum_{k=0}^K V_{\text{vox}}, \quad (8)$$

где I, J, K – количество вокселей по осям x, y, z соответственно.

Масса камня рассчитывалась как произведение объема одного вокселя на сумму плотностей всех вокселей, принадлежащих объёму камня:

$$m = V_{\text{vox}} * \sum_{i=0}^I \sum_{j=0}^J \sum_{k=0}^K \rho_{ijk}, \quad (9)$$

где m – масса камня, ρ_{ijk} – значение плотности текущего вокселя в объеме камня.

После расчета массы камня производится расчет времени его разрушения в соответствии с формулой (1). Расчет времени выполняется для каждого значения энергии лазера и частоты из задаваемого диапазона значений и соответствующего шага изменения каждой из этих величин. Таким образом, для одного и того же камня получается матрица значений времени его разрушения.

Далее производится нечеткий анализ диапазонов значений величин энергии лазера и его частоты в соответствии с установленными градациями значений.

На основании заданных видов режимов работы лазера, применяемых при проведении процедуры литотрипсии, рассчитанной массы камня, значением плотностей его частей выполняется формирование рекомендаций по подбору величин энергии лазера и его частоты. С целью предоставления полной информации для планирования операционного воздействия модулем формируется вывод для всех допустимых режимов работы лазера со значениями настроек энергии и частоты в виде таблицы значений времени разрушения камня. Для наглядности представления ячейки таблицы окрашиваются определенным цветом, характеризующим каждый из режимов работы лазера в соответствии с установленной цветовой гаммой (табл. 1).

Таблица 1

Цвета окрашивания ячеек в таблице расчета времени разрушения камней

Название режима	Цвет
Режим для дробления осколков камней	
Режим работы для камней с низкой плотностью	
Режим работы при неудобных подходах инструмента	
Оптимальный режим работы	
Режим работы в зоне высокой плотности камня	
Режим работы с высоким риском нанесения травмы	
Малозффективный режим работы	
Малозффективный режим работы – не рекомендовано к применению	

В результате хирург, планирующий операцию, может видеть рекомендуемые режимы работы лазерной установки, время разрушения камня, величины энергии и частоты, характерных для различных режимов работы лазера.

Интерфейс модуля поддержки принятия врачебных решений при выборе режима воздействия лазером на камень представляет собой окно с полями ввода диапазонов значений величин энергии, частоты лазера и шага изменения данных величин и кнопок, выполняющих действия по расчету и сохранению данных (рис. 4). По умолчанию установлены следующие значения параметров лазера: - величина гамма-излучения – 0,4; частота – в диапазоне от 8 до 15, шаг изменения 1; энергия – в диапазоне от 0,6 до 2,4, шаг изменения 0,1.

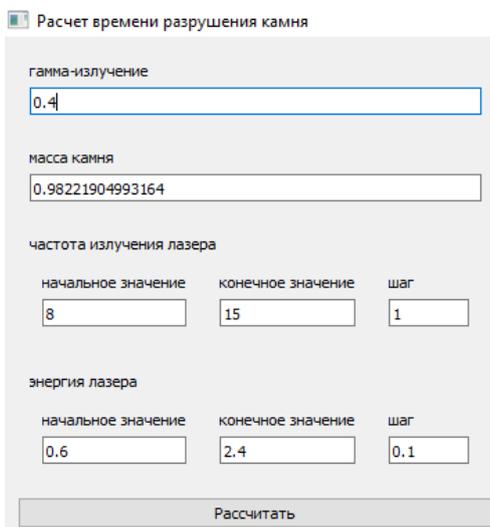


Рис. 4. Окно ввода параметров и расчета времени разрушения камней

Рис. 4. Окно ввода параметров и расчета времени разрушения камней

В интерфейсе модуля также присутствует таблица с уже рассчитанными из заданных по умолчанию диапазонов значений энергии и частоты лазера времени разрушения для конкретного камня (рис. 5).

	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0
0.6	0.51	0.45	0.41	0.37	0.34	0.31	0.29	0.27
0.7	0.44	0.39	0.35	0.32	0.29	0.27	0.25	0.23
0.8	0.38	0.34	0.31	0.28	0.26	0.24	0.22	0.2
0.9	0.34	0.3	0.27	0.25	0.23	0.21	0.19	0.18
1.0	0.31	0.27	0.25	0.22	0.2	0.19	0.18	0.16
1.1	0.28	0.25	0.22	0.2	0.19	0.17	0.16	0.15
1.2	0.26	0.23	0.2	0.19	0.17	0.16	0.15	0.14
1.3	0.24	0.21	0.19	0.17	0.16	0.15	0.13	0.13
1.4	0.22	0.19	0.18	0.16	0.15	0.13	0.13	0.12
1.5	0.2	0.18	0.16	0.15	0.14	0.13	0.12	0.11
1.6	0.19	0.17	0.15	0.14	0.13	0.12	0.11	0.1
1.7	0.18	0.16	0.14	0.13	0.12	0.11	0.1	0.1
1.8	0.17	0.15	0.14	0.12	0.11	0.1	0.1	0.09
1.9	0.16	0.14	0.13	0.12	0.11	0.1	0.09	0.09
2.0	0.15	0.14	0.12	0.11	0.1	0.09	0.09	0.08
2.1	0.15	0.13	0.12	0.11	0.1	0.09	0.08	0.08
2.2	0.14	0.12	0.11	0.1	0.09	0.09	0.08	0.07
2.3	0.13	0.12	0.11	0.1	0.09	0.08	0.08	0.07

Рис. 5. Таблица с рассчитанными значениями времени разрушения камня

Ячейки таблицы раскрашены в соответствии с цветами режимов работы лазера, рекомендованных для использования при проведении операции литотрипсии. Под таблицей располагается визуализация рассчитанных величин времени разрушения камня и рекомендации по выбору режима работы лазера. Диапазоны значений времени разрушения камня окрашены в соответствии с значением градации нечеткой переменной «режим лазера» (рис. 6).

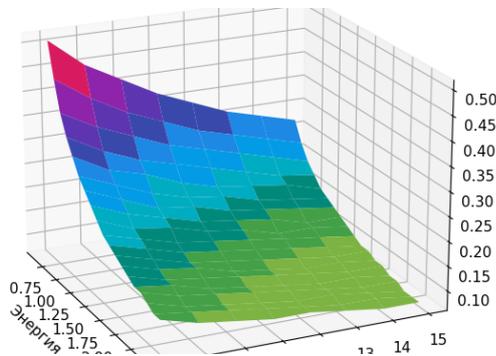


Рис. 6. Рекомендации по выбору режима работы лазера

Врач, осуществляя подготовку операции лазерной литотрипсии, анализирует параметры ранее найденных системой камня в почке: масса, значения плотности по НУ и физической плотности камня, по трехмерной визуализации оценивает расположение камня внутри почки, выбирает оптимальный или иной режим работы лазерной установки в зависимости от данных пациента и предлагаемой цветовой таблицы, по времени разрушения камня из ячейки он получает из данной таблицы величины энергии лазера и частоту его работы, которую необходимо установить на лазерном аппарате. Дополнительно врач может изменить начальные значения параметров лазера и выполнить расчет времени разрушения с новыми параметрами. Модуль выполнит перерасчет времени, осуществит формирование рекомендаций по выбору режима в соответствии с введенными значениями энергии лазера и частоты.

Таким образом осуществляется поддержка врачебного решения при осуществлении выбора режима работы лазера, сокращается время, затрачиваемое врачом, на выбор частоты и энергии лазера, уменьшается риск нанесения травм внутренних органов от неправильной установки частоты или энергии лазерного воздействия на камень.

Заключение. В данной статье предложен алгоритм нечеткой оценки параметров лазера при выборе режима работы лазерной установки для проведения процедуры лазерной литотрипсии. В алгоритме использован механизм нечеткого логического вывода из-за неоднозначности выбора режима работы лазерной установки, связанной с оценкой врачом не только времени разрушения камня, но и субъективных факторов. В процессе работы алгоритма производится вычисление времени разрушения камня в зависимости от его массы, диапазонов изменения значений энергии и частоты лазера, формируется матрица значений времени разрушения. Далее формируются рекомендации в виде таблицы значений времени разрушения и окраски диапазонов этих значений в цвета, которые соответствуют заданным режимам разрушения камней.

Алгоритм был реализован в виде программного модуля и был интегрирован в прототип системы поддержки принятия врачебных решений в хирургии и урологии с использованием технологий компьютерного зрения, проведено тестирование на базе КММЦ им. Святителя Луки г. Симферополь. Разработанный алгоритм и его программная реализация позволили сократить время на принятие решения врачом по выбору параметров лазера, снизить риски нанесения повреждений внутренним органам пациента при проведении хирургической процедуры лазерной литотрипсии. Данный алгоритм может быть использован при создании других подобных систем поддержки принятия решений в медицине.

Благодарности. Данные исследования выполнены при финансовой поддержке ФГБУ «Фонд содействия развитию малых форм предприятий в научно-технической сфере», программа «Старт», конкурс «Старт-Искусственный интеллект-1» (I очередь), заявка С1ИИ-112266, договор № 27ГС1ИИС12-D7/71365.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Гусев А.В., Зарубина Т.В. Поддержка принятия врачебных решений в медицинских информационных системах медицинской организации // *Врач и информационные технологии*. – 2017. – № 2. – С. 60-72.
2. Омирова Н.И., Тишков А.В. Система поддержки принятия решений врача-уролога в биотехнической системе диагностики мочекаменной болезни // *Биотехносфера*. – 2020. – № 5(65). – С. 43-48. – DOI: 10.25960/bts.2020.5.43.
3. Щамхалова К.К., Меринов Д.С., Артемов А.В. и др. Искусственный интеллект и нейронные сети в урологии // *Экспериментальная и клиническая урология*. – 2023. – Т. 16, № 2. – С. 32-37. – DOI: 10.29188/2222-8543-2023-16-2-32-37.
4. Галченков А.С., Новицкий В.О., Кондратьев Е.А. Сервисы поддержки принятия решений по диагностике и лечению заболеваний и их практическое применение на примере ХБП 5Д // *Врач и информационные технологии*. – 2020. – № 1. – С. 45-51. – DOI: 10.37690/1811-0193-2020-1-45-51.
5. Левенков К.О., Коровин Е.Н., Новикова Е.И. Нейросетевое моделирование процесса выбора схемы лечения пациентов с хроническим пиелонефритом и мочекаменной болезнью // *Моделирование, оптимизация и информационные технологии*. – 2018. – Т. 6, № 4(23). – С. 61-71. – DOI: 10.26102/2310-6018/2018.23.4.005.
6. Коцарь А.Г. Математическое моделирование и алгоритмизация прогнозирования, диагностики, профилактики и лечения мочекаменной болезни: дисс. ... д-ра мед. наук. – 2014. – 302 с.
7. Зубарев Д.А., Корневская С.Н., Коцарь А.Г. и др. Использование нечетких математических моделей при ведении больных с мочекаменной болезнью // *Системный анализ и управление в биомедицинских системах*. – 2017. – Т. 16, № 2. – С. 345-350.
8. Чернега В.С., Глуховская-Степаненко Н.П., Еременко А. и др. Оценка скорости фрагментации моче-вых камней при контактной литотрипсии гольмиевым лазером // *Урология*. – 2018. – № 5. – С. 69-73. – DOI: 10.18565/urology.2018.5.69-72.
9. Коробков Д.М., Мосина Л.М., Степанов Н.Ю. Применение дистанционной ударно-волновой литотрипсии при мочекаменной болезни // *Международный научно-исследовательский журнал*. – 2021. – № 2-3 (104). – С. 31-33. – DOI: 10.23670/IRJ.2021.103.2.067.

10. Степанов Н.Ю., Дуваров З.А., Бояркин Е.В. и др. Применение экстракорпоральной ударно-волновой литотрипсии и контактной уретеролитотрипсии при лечении конкрементов в дистальном отделе мочеточника // *Международный научно-исследовательский журнал*. – 2020. – № 6 (98). – С 116-118.
11. Чернега В.С., Глуховская-Степаненко Н.П., Еременко С.Н. и др. Сетевая модель для оценки длительности медицинского технологического процесса лазерной контактной литотрипсии // *Врач и информационные технологии*. – 2018. – № 4. – С. 75-82.
12. Борисов Д.Д., Кульнев С.В., Севрюков В.В. Применение искусственного интеллекта при организации лечебно-диагностических мероприятий // *Известия российской военно-медицинской академии*. – 2019. – 38 (4). – С. 122-125. – URL: https://journals.eco-vector.com/RMMAreP/article/view/26037/20184/ru_RU.
13. Fitri L.A., Haryanto F., Arimura H., YunHao C., Ninomiya K., Nakano R. Automated classification of urinary stones based on microcomputed tomography images using convolutional neural network // *Physica Medica*. – 2020. – 78. – P. 201-208. – URL: <https://doi.org/10.1016/j.ejmp.2020.09.007>.
14. Parakh A., Lee H., Lee J.H., Eisner B.H., Sahani D.V., Do S. Urinary Stone Detection on CT Images Using Deep Convolutional Neural Networks: Evaluation of Model Performance and Generalization // *Radiol Artif Intell*. – 2019 Jul 24. – 1(4):e180066. – URL: <https://doi.org/10.1148/ryai.2019180066>.
15. Пранович А.А., Исмаилов А.К., Карельская Н.А. и др. Искусственный интеллект в диагностике и лечении мочекаменной болезни // *Российский журнал телемедицины и электронного здравоохранения*. – 2022. – 8 (1). – С. 42-57. – URL: <https://doi.org/10.29188/2712-9217-2022-8-1-42-57>.
16. Jendeborg J., Thunberg P., Lide n M. Differentiation of distal ureteral stones and pelvic phleboliths using a convolutional neural network // *Urolithiasis*. – 2020. – URL: <https://doi.org/10.1007/s00240-020-01180-z>.
17. Nithya A., Appathurai A., Venkatadri N., Ramji D.R., Anna Palagan C. Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images // *Meas J Int Meas Confed*. – 2020. – 149. – URL: <https://doi.org/10.1016/j.measurement.2019.106952>.
18. Кудряш В.Л., Маршев С.В., Габлия М.Ю. и др. Практические аспекты применения дистанционной литотрипсии у больных мочекаменной болезнью // *Урология*. – 2013. – № 2. – С. 12-17.
19. Чернега В.С., Еременко С.Н., Еременко А.Н. и др. Прогнозирование времени трансуретральной гольмиевой литотрипсии в лечении уролитиаза // *Врач и информационные технологии*. – 2020. – № 2. – С. 72-80. – DOI: 10.37690/1811-0193-2020-2-72-80.
20. Руденко М.А., Руденко А.В. Нечеткая модель классификации медицинских изображений на основе нейронных сетей // *Международная конференция по мягким вычислениям и измерениям*. – 2021. – Т. 1. – С. 336-339.

REFERENCES

1. Gusev A.V., Zarubina T.V. Podderzhka prinyatiya vrachebnykh resheniy v meditsinskikh informatsionnykh sistemakh meditsinskoy organizatsii [Support for medical decision-making in medical information systems of a medical organization], *Vrach i informatsionnye tekhnologii* [Doctor and information technologies], 2017, No. 2, pp. 60-72.
2. Omirova N.I., Tishkov A.V. Sistema podderzhki prinyatiya resheniy vracha-urologa v biotekhnicheskoy sisteme diagnostiki mochekamennoy bolezni [Decision support system of a urologist in a biotechnical system for diagnosing urolithiasis], *Biotekhnosfera* [Biotechnosphere], 2020, No. 5(65), pp. 43-48. DOI: 10.25960/bts.2020.5.43.
3. Shchamkhalova K.K., Merinov D.S., Artemov A.V. i dr. Iskusstvennyy intellekt i neyronnye seti v urologii [Artificial intelligence and neural networks in urology], *Eksperimental'naya i klinicheskaya urologiya* [Experimental and clinical urology], 2023, Vol. 16, No. 2, pp. 32-37. DOI: 10.29188/2222-8543-2023-16-2-32-37.
4. Galchenkov A.S., Novitskiy V.O., Kondrat'ev E.A. Servisy podderzhki prinyatiya resheniy po diagnostike i lecheniyu zabolevaniy i ikh prakticheskoe primeneniye na primere KhBP 5D [Decision support services for the diagnosis and treatment of diseases and their practical application on the example of CKD 5D], *Vrach i informatsionnye tekhnologii* [Doctor and information technology], 2020, No. 1, pp. 45-51. DOI: 10.37690/1811-0193-2020-1-45-51.
5. Levenkov K.O., Korovin E.N., Novikova E.I. Neyrosetevoye modelirovaniye protsessa vybora skhemy lecheniya patsientov s khronicheskim pielonefritom i mochekamennoy boleznyu [Neural network modeling of the process of choosing a treatment regimen for patients with chronic pyelonephritis and urolithiasis], *Modelirovaniye, optimizatsiya i informatsionnye tekhnologii* [Modeling, optimization and information technology], 2018, Vol. 6, No. 4 (23), pp. 61-71. – DOI: 10.26102/2310-6018/2018.23.4.005.

6. Kotsar' A.G. Matematicheskoe modelirovanie i algoritimizatsiya prognozirovaniya, diagnostiki, profilaktiki i lecheniya mochekamennoy bolezni: diss. ... d-ra med. nauk [Mathematical modeling and algorithmization of forecasting, diagnosis, prevention and treatment of urolithiasis: dr. of med. sc. diss.], 2014, 302 p.
7. Zubarev D.A., Korenevskaya S.N., Kotsar' A.G. i dr. Ispol'zovanie nechetkikh matematicheskikh modeley pri vedenii bol'nykh s mochekamennoy boleznyu [The use of fuzzy mathematical models in the management of patients with urolithiasis], *Sistemnyy analiz i upravlenie v biomeditsinskikh sistemakh* [System analysis and management in biomedical systems], 2017, Vol. 16, No. 2, pp. 345-350.
8. Chernega V.S., Tlukhovskaya-Stepanenko N.P., Eremenko A. i dr. Otsenka skorosti fragmentatsii mochevykh kamney pri kontaktnoy litotripsii gol'mievym lazerom [Assessment of the rate of fragmentation of urinary stones during contact lithotripsy with a holmium laser], *Urologiya* [Urology], 2018, No. 5, pp. 69-73. DOI: 10.18565/urology.2018.5.69-72.
9. Korobkov D.M., Mosina L.M., Stepanov N.Yu. Primenenie distantsionnoy udarno-volnovoy litotripsii pri mochekamennoy bolezni [The use of remote shock wave lithotripsy in urolithiasis], *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal* [International Scientific Research Journal], 2021, No. 2-3 (104), pp. 31-33. DOI: 10.23670/IRJ.2021.103.2.067.
10. Stepanov N.Yu. Duvayarov Z.A., Boyarkin E.V. i dr. Primenenie ekstrakorporal'noy udarno-volnovoy litotripsii i kontaktnoy ureterolitotripsii pri lechenii konkrementov v distal'nom otdel'noy mochetochke [The use of extracorporeal shock wave lithotripsy and contact ureterolithotripsy in the treatment of concretions in the distal ureter], *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal* [International Scientific Research Journal], 2020, No. 6 (98), pp 116-118.
11. Chernega V.S., Tlukhovskaya-Stepanenko N.P., Eremenko S.N. i dr. Setevaya model' dlya otsenki dlitel'nosti meditsinskogo tekhnologicheskogo protsessa lazernoy kontaktnoy litotripsii [A network model for estimating the duration of the medical technological process of laser contact lithotripsy], *Vrach i informatsionnye tekhnologii* [Doctor and Information Technologies], 2018, No. 4, pp. 75-82.
12. Borisov D.D., Kul'nev S.V., Sevryukov V.V. Primenenie iskusstvennogo intellekta pri organizatsii lechebno-diagnosticskikh meropriyatii [The use of artificial intelligence in the organization of therapeutic and diagnostic measures], *Izvestiya rossiyskoy voenno-meditsinskoy akademii* [Proceedings of the Russian Military Medical Academy], 2019, 38 (4), pp 122-125. Available at: https://journals.eco-vector.com/RMMAREp/article/view/26037/20184/ru_RU.
13. Fitri L.A., Haryanto F., Arimura H., YunHao C., Ninomiya K., Nakano R. Automated classification of urinary stones based on microcomputed tomography images using convolutional neural network, *Physica Medica*, 2020, 78, pp. 201-208. Available at: <https://doi.org/10.1016/j.ejmp.2020.09.007>.
14. Parakh A., Lee H., Lee J.H., Eisner B.H., Sahani D.V., Do S. Urinary Stone Detection on CT Images Using Deep Convolutional Neural Networks: Evaluation of Model Performance and Generalization, *Radiol Artif Intell.*, 2019 Jul 24, 1(4):e180066. Available at: <https://doi.org/10.1148/ryai.2019180066>.
15. Pranovich A.A., Ismailov A.K., Karel'skaya N.A. i dr. Iskusstvennyy intellekt v diagnostike i lechenii mochekamennoy bolezni [Artificial intelligence in the diagnosis and treatment of urolithiasis], *Rossiyskiy zhurnal telemeditsiny i elektronnoy zdravookhraneniya* [Russian Journal of Telemedicine and E-Health], 2022, 8 (1), pp. 42-57. Available at: <https://doi.org/10.29188/2712-9217-2022-8-1-42-57>.
16. Jendeborg J., Thunberg P., Lide n M. Differentiation of distal ureteral stones and pelvic phleboliths using a convolutional neural network, *Urolithiasis*, 2020. Available at: <https://doi.org/10.1007/s00240-020-01180-z>.
17. Nithya A., Appathurai A., Venkatadri N., Ramji D.R., Anna Palagan C. Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images, *Meas J Int Meas Confed.*, 2020, 149. Available at: <https://doi.org/10.1016/j.measurement.2019.106952>.
18. Kudryash V.L., Marshev S.V., Gabliya M.Yu i dr. Prakticheskie aspekty primeneniya distantsionnoy litotripsii u bol'nykh mochekamennoy boleznyu [Practical aspects of the use of remote lithotripsy in patients with urolithiasis], *Urologiya* [Urology], 2013, No. 2, pp. 12-17.
19. Chernega V.S., Eremenko S.N., Eremenko A.N. i dr. Prognozirovaniye vremeni transuretral'noy gol'mievoy litotripsii v lechenii urolitiaza [Predicting the time of transurethral holmium lithotripsy in the treatment of urolithiasis], *Vrach i informatsionnye tekhnologii* [Doctor and Information Technologies], 2020, No. 2, pp. 72-80. DOI: 10.37690/1811-0193-2020-2-72-80.
20. Rudenko M.A., Rudenko A.V. Nechetkaya model' klassifikatsii meditsinskikh izobrazheniy na osnove neyronnykh setey [Fuzzy model of classification of medical images based on neural networks], *Mezhdunarodnaya konferentsiya po myagkim vychisleniyam i izmereniyam* [International Conference on Soft Computing and Measurements], 2021, Vol. 1, pp. 336-339.

Статью рекомендовала к опубликованию д.т.н., профессор И.Л. Каширина.

Руденко Андрей Владимирович – Крымский федеральный университет имени В.И. Вернадского; e-mail: rudenkoandre@mail.ru; г. Симферополь, Россия; преподаватель; ORCID: <https://orcid.org/0009-0004-6297-2742>.

Руденко Марина Анатольевна – e-mail: rudenko.ma@cfuv.ru; к.т.н.; доцент кафедры компьютерной инженерии и моделирования Физико-технического института; ORCID: <https://orcid.org/0000-0002-8334-8453>.

Rudenko Andrei Vladimirovich – V.I. Vernadsky Crimean Federal University; e-mail: rudenkoandre@mail.ru; Simferopol, Russia; assistant; ORCID: <https://orcid.org/0009-0004-6297-2742>.

Rudenko Marina Anatol'evna – e-mail: rudenko.ma@cfuv.ru; cand. of eng. sc.; associate professor of the department of computer engineering and modeling of the Institute of Physics and Technology; ORCID: <https://orcid.org/0000-0002-8334-8453>.

Раздел IV. Сообщение об отзыве публикации

УДК 21474

DOI 10.18522/2311-3103-2022-6-140-147

Алмашаал Мохаммад Джалаль

К ОЦЕНКЕ ОБЛАСТИ ПРИТЯЖЕНИЯ ПОЛОЖЕНИЯ РАВНОВЕСИЯ НЕЛИНЕЙНЫХ СИСТЕМ УПРАВЛЕНИЯ

Алмашаал Мохаммад Джалаль. К оценке области притяжения положения равновесия нелинейных систем управления // Известия ЮФУ. Технические науки. – 2022. – № 6 (230). – С. 140-147.

Согласно решению редакции журнала «Известия ЮФУ. Технические науки» данная статья отозвана.

Section IV. Report of retraction

Almashaal Mohammad Jalal

TO ESTIMATION OF ATTRACTION AREA OF EQUILIBRIUM OF NONLINEAR CONTROL SYSTEMS

Almashaal Mohammad Jalal. To estimation of attraction area of equilibrium of nonlinear control systems *Izvestiya YuFU. Tekhnicheskije nauki* [Izvestiya SFedU. Engineering Sciences], 2022, No. 6 (230), pp. 140-147.

According to the "Izvestiya SFedU. Engineering Sciences" Editorial decision this article has been retracted.

ПРАВИЛА ОФОРМЛЕНИЯ РУКОПИСЕЙ

1. Объем статьи должен быть не менее 12 и не более 18 страниц. Формат (А 4). Редактор *Word 7 for Windows*, шрифт Times New Roman, размер 14, интервал 1,5. Авторы представляют в редакцию 1 экз. статьи и идентичный электронный вариант.

2. Названию статьи предшествует индекс УДК, соответствующий заявленной теме.

3. Текст статьи начинается с названия статьи (на русском и английском языках), фамилии, имени и отчества автора (полностью) и снабжается аннотацией на русском и английском языках объемом *не менее 250-300 слов*. В тексте аннотации указывается цель, задачи исследования и краткие выводы. В аннотации *не следует* давать ссылки на номер публикации в списке литературы к статье. После аннотаций приводятся ключевые слова (словосочетания), несущие в тексте основную смысловую нагрузку (на русском и английском языках).

4. В тексте статьи следует использовать минимальное количество таблиц и иллюстраций. Рисунок должен иметь объяснения значений всех компонентов, порядковый номер, название, расположенное под рисунком. В тексте на рисунок дается ссылка. Таблица должна иметь порядковый номер, заголовок, расположенный над ней. Данные таблиц и рисунков не должны дублировать текст. Формулы должны быть набраны *в редакторе формул Word 7 for Windows*.

5. Цитаты тщательно сверяются с первоисточником и визируются автором на обратной стороне последней страницы: "Цитаты и фактический материал сверены". Подпись, дата.

6. Наличие пристатейного библиографического списка на русском и английском языках обязательно. *Ссылок должно быть не менее 20-ти*, из них на зарубежные источники – не менее 35 %. В тексте ссылки должны быть в квадратных скобках.

Примеры оформления литературы: а) для книг: фамилия, инициалы автора(ов), полное название книги, место, год издания, страницы; б) для статей: фамилия и инициалы автора(ов), полное название сборника, книги, газеты, журнала, где опубликована статья, место и год издания (сборника, книги), номер (для журнала), год и дата (для газеты), выпуск, часть (для сборника), страницы, на которых опубликована статья. Иностранная литература оформляется по тем же правилам.

Ссылки на неопубликованные работы не допускаются.

7. Рукопись должна быть тщательно вычитана. Редакционная коллегия оставляет за собой право при необходимости сокращать статьи, редактировать и отсылать авторам на доработку.

8. Статьи сопровождаются сведениями об авторе(ах) (фамилия, имя, отчество, ученое звание, должность, место работы, адрес, электронный адрес и номер телефона) на русском и английском языках.

9. Плата с аспирантов за публикацию рукописей не взимается.

Адрес журнала в Интернете: <http://izv-tn.tti.sfedu.ru/>.