

СОДЕРЖАНИЕ

ФУНДАМЕНТАЛЬНЫЕ ОСНОВЫ ПРОБЛЕМ НАДЕЖНОСТИ И КАЧЕСТВА

Собиров Б. Ш., Карамергенов Т. Т., Северцев Н. А., Юрков Н. К.
КВАНТОВЫЕ ТЕХНОЛОГИИ – ГЛОБАЛЬНЫЙ ВЫЗОВ, ТРЕБУЮЩИЙ ОЧЕРЕДНОЙ РЕВОЛЮЦИИ МЕТОДОВ ОБРАБОТКИ И ПЕРЕДАЧИ ИНФОРМАЦИИ 6

Иванов А. И., Ермакова А. И.
ОЦЕНКА КАЧЕСТВА РАБОТЫ МНОГОКРИТЕРИАЛЬНОГО НЕЙРОСЕТЕВОГО СТАТИСТИЧЕСКОГО АНАЛИЗА НА ПРИМЕРЕ ПРОВЕРКИ ГИПОТЕЗЫ НЕЗАВИСИМОСТИ ДАННЫХ МАЛЫХ ВЫБОРОК 13

Годунов А. И., Куканов С. А., Бакланов Д. В.
ФИЗИЧЕСКИЕ ОСОБЕННОСТИ ПРОЦЕССОВ ДЕФОРМАЦИИ И РАЗРУШЕНИЯ МАТЕРИАЛА МЕТОДОМ КОНЕЧНЫХ УЗЛОВ 19

Макеева Г. С.
МОДЕЛИРОВАНИЕ УПРАВЛЯЕМОСТИ ХАРАКТЕРИСТИК РЕШЕТОК ПЛАЗМОННЫХ ГРАФЕНОВЫХ НАНОАНТЕНН В СРЕДНЕМ ИНФРАКРАСНОМ ДИАПАЗОНЕ 30

Старостин И. Е.
МЕТОД МАТЕМАТИЧЕСКОГО ПРОТОТИПИРОВАНИЯ ЭНЕРГЕТИЧЕСКИХ ПРОЦЕССОВ КАК ИНСТРУМЕНТ СИНТЕЗА МАТЕМАТИЧЕСКОГО ЯДРА ЦИФРОВЫХ ДВОЙНИКОВ РАЗЛИЧНЫХ ФИЗИКО-ХИМИЧЕСКИХ СИСТЕМ 41

Тимошкин М. О., Романова Е. Г.
ПРИМЕНЕНИЕ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ КЛАССИФИКАЦИИ ОПУХОЛЕЙ ГОЛОВНОГО МОЗГА 51

**ПРОЕКТИРОВАНИЕ И ТЕХНОЛОГИИ ПРИБОРОСТРОЕНИЯ
И РАДИОЭЛЕКТРОННОЙ АППАРАТУРЫ**

Фокина Е. А., Трофимов А. А., Пономарев В. Н., Здобнов С. А.
ПРОЕКТИРОВАНИЕ ИМИТАЦИОННОЙ МОДЕЛИ ДАТЧИКА ТЕМПЕРАТУРЫ ПРИ ВОЗДЕЙСТВИИ КРИТИЧЕСКИХ ТЕМПЕРАТУР И СИНУСОИДАЛЬНОЙ ВИБРАЦИИ 66

Фролов С. И., Данилова Е. А., Кочегаров И. И.
КОМПОНОВКА БЕЗРЕЗОНАНСНЫХ БЛОКОВ БОРТОВЫХ РАДИОЭЛЕКТРОННЫХ СИСТЕМ С ИСПОЛЬЗОВАНИЕМ ДВУХСЛОЙНЫХ ПЕЧАТНЫХ ПЛАТ С ЖЕСТКОЙ ОСНОВОЙ 73

Якимов А. Н.
МОДЕЛЬНОЕ ИССЛЕДОВАНИЕ ВЛИЯНИЯ ФОРМЫ ПЕРЕМЕННО-ФАЗНЫХ УЧАСТКОВ РАСПРЕДЕЛЕНИЯ ПОЛЯ НА ИЗЛУЧЕНИЕ АНТЕННЫ 82

**ДИАГНОСТИЧЕСКИЕ МЕТОДЫ ОБЕСПЕЧЕНИЯ
НАДЕЖНОСТИ И КАЧЕСТВА СЛОЖНЫХ СИСТЕМ**

Макеева Г. С.
ЭЛЕКТРОННО-УПРАВЛЯЕМЫЕ НЕЛИНЕЙНЫЕ УСТРОЙСТВА
ДАЛЬНОГО И СРЕДНЕГО ИНФРАКРАСНОГО ДИАПАЗОНА
НА ПЛАТФОРМЕ ГРАФЕНОВОЙ ПЛАЗМОНИКИ. ОБЗОР 90

Годунов А. И., Куканов С. А., Суздаальцев П. С.
НАВИГАЦИОННЫЕ ЭЛЕМЕНТЫ ПОЛЕТА БЕСПИЛОТНОГО ЛЕТАТЕЛЬНОГО АППАРАТА 104

Васильева А. В.
ИССЛЕДОВАНИЕ АЛГОРИТМОВ ИЗМЕРЕНИЯ СОСТАВЛЯЮЩИХ
ВЕКТОРА СКОРОСТИ РАДИОВЫСОТОМЕРНЫХ СИСТЕМ,
АДАПТИВНЫХ К ПОДСТИЛАЮЩЕЙ ПОВЕРХНОСТИ..... 112

Зырянов Ю. Т., Калхиташвили Д. Ш., Хрипунов С. П.
РАЗРАБОТКА ПРОТОТИПА АРХИТЕКТУРЫ ИНТЕЛЛЕКТУАЛЬНОГО
МИКРОКОНТРОЛЛЕРА ПРИ РЕАЛИЗАЦИИ ПРИЛОЖЕНИЙ ИНТЕРНЕТА ВЕЩЕЙ..... 118

Егжов Д. В., Надейкина Л. А., Терешко П. О.
СИСТЕМА ПОДДЕРЖКИ АВИАПАССАЖИРОВ С ИСПОЛЬЗОВАНИЕМ
ТЕХНОЛОГИЙ МОНИТОРИНГА ИХ МЕСТОПОЛОЖЕНИЯ..... 124

БЕЗОПАСНОСТЬ В ЧРЕЗВЫЧАЙНЫХ СИТУАЦИЯХ

Masloboev A. V.
AN INDEX-BASED METHOD FOR INTEGRAL ESTIMATION OF REGIONAL
CRITICAL INFRASTRUCTURE RESILIENCE USING FUZZY CALCULATIONS
(PART 2. RESILIENCE CAPACITY MODELS AND BACKBONE CAPABILITIES) 130

CONTENT

FUNDAMENTALS OF RELIABILITY AND QUALITY ISSUES

- Sobirov B.Sh., Karamergenov T.T., Severtsev N.A., Yurkov N.K.**
QUANTUM TECHNOLOGIES ARE A GLOBAL CHALLENGE, REQUIRING ANOTHER
REVOLUTION OF METHODS PROCESSING AND TRANSMISSION OF INFORMATION 6
- Ivanov A.I., Ermakova A.I.**
ASSESSING THE PERFORMANCE QUALITY OF MULTICRITERIA NEURAL NETWORK
STATISTICAL ANALYSIS USING THE EXAMPLE OF TESTING THE HYPOTHESIS
OF INDEPENDENCE OF SMALL SAMPLE DATA 13
- Godunov A.I., Kukanov S.A., Baklanov D.V.**
PHYSICAL FEATURES OF THE PROCESSES OF DEFORMATION
AND DESTRUCTION OF THE MATERIAL BY THE FINITE NODE METHOD..... 19
- Makeeva G.S.**
MODELING THE CONTROLLABILITY OF CHARACTERISTICS OF PLASMON
GRAPHENE NANOANTENNA GRATINGS IN THE MID-IR RANGE 30
- Starostin I.E.**
METHOD OF MATHEMATICAL PROTOTYPING OF ENERGY PROCESSES
AS A TOOL FOR SYNTHESIS OF THE MATHEMATICAL KERNEL
OF DIGITAL TWINS OF DIFFERENT PHYSICAL AND CHEMICAL SYSTEMS 41
- Timoshkin M.O., Romanova E.G.**
APPLICATION OF DEEP LEARNING NEURAL NETWORK
ALGORITHMS FOR BRAIN TUMOR CLASSIFICATION..... 51

DESIGN AND TECHNOLOGY OF INSTRUMENTATION AND ELECTRONIC EQUIPMENT

- Fokina E.A., Trofimov A.A., Ponomarev V.N., Zdobnov S.A.**
DESIGNING A SIMULATION MODEL OF A TEMPERATURE SENSOR
UNDER CRITICAL TEMPERATURE AND SINUSOIDAL VIBRATION CONDITIONS..... 66
- Frolov S.I., Danilova E.A., Kochegarov I.I.**
THE LAYOUT OF ON-BOARD ELECTRONIC SYSTEMS WITHOUT RESONANT BLOCKS
USING DOUBLE-LAYER PRINTED CIRCUIT BOARDS WITH A RIGID BASE 73
- Yakimov A.N.**
A MODEL STUDY OF THE INFLUENCE OF THE SHAPE OF VARIABLE-PHASE
FIELD DISTRIBUTION AREAS ON ANTENNA RADIATION 82

**DIAGNOSTIC METHODS FOR ENSURING RELIABILITY
AND QUALITY OF COMPLEX SYSTEMS**

Makeeva G.S.

ELECTRONICALLY CONTROLLED NONLINEAR DEVICES BASED ON THE GRAPHENE
PLASMONICS PLATFORM IN THE FAR AND MID-INFRARED RANGE. REVIEW..... 90

Godunov A.I., Kukanov S.A., Suzdaltsev P.S.

FLIGHT NAVIGATION ELEMENTS AN UNMANNED AERIAL VEHICLE 104

Vasilieva A.V.

INVESTIGATION OF ALGORITHMS FOR MEASURING THE COMPONENTS OF THE VELOCITY
VECTOR OF RADIO ALTIMETER SYSTEMS ADAPTIVE TO THE UNDERLYING SURFACE 112

Zyryanov Yu.T., Kalkhitashvili D.Sh., Khripunov S.P.

DEVELOPMENT OF A PROTOTYPE OF INTELLIGENT MICROCONTROLLER
ARCHITECTURE WHEN IMPLEMENTING INTERNET OF THINGS APPLICATIONS..... 118

Egzhov D.V., Nadeykina L.A., Tereshko P.O.

AIR PASSENGER SUPPORT SYSTEM USING LOCATION MONITORING TECHNOLOGIES..... 124

SAFETY IN EMERGENCY SITUATIONS

Masloboev A.V.

AN INDEX-BASED METHOD FOR INTEGRAL ESTIMATION OF REGIONAL
CRITICAL INFRASTRUCTURE RESILIENCE USING FUZZY CALCULATIONS
(PART 2. RESILIENCE CAPACITY MODELS AND BACKBONE CAPABILITIES) 130

30 сентября 2024 г. ученому-классику, стоявшему у истоков теории надежности в СССР, лауреату Государственной премии СССР, обладателю ордена «Знак Почета», выдающемуся члену редколлегии нашего журнала, доктору физико-математических наук, профессору НИУ ВШЭ Виктору Алексеевичу Каштанову исполнилось 90 лет.

Долгое время В. А. Каштанов проработал деканом факультета прикладной математики (ФПМ) Московского института электронного машиностроения, который позже был переименован в Московский институт электроники и математики, ныне Институт им. А. Н. Тихонова, который является структурным подразделением Высшей школы экономики, которым он руководил 32 года.

Родился в Москве, после окончания средней школы с серебряной медалью поступил на мехмат. На 3 курсе Виктор Алексеевич познакомился с математическим гением, заведующим кафедрой математического анализа Андреем Колмогоровым, преподававшим математику в МГУ, который определил его дальнейшую судьбу. 9 мая 1957 г. Виктор Алексеевич



успешно защитил диплом под руководством Колмогорова. Сразу по окончании он был распределен в лабораторию математики НИИ-17. Через год после начала работы Виктор Алексеевич приступил к решению задач в области надежности. За время обучения в аспирантуре НИИ-17 его научным руководителем являлся А. Д. Соловьев, который сформулировал задачу по теории надежности, определивший путь профессора Каштанова. Виктор Алексеевич изучал локальные свойства процессов с независимыми приращениями. Первая книга, где системно была изложена теория надежности, вышла в 1965 г., где отсутствовал этап эксплуатации. В 1983 г. вышла вторая версия книги, где, наряду с уже классическими к тому времени разделами появились новые, связанные с моделированием и эксплуатацией, и Виктор Алексеевич стал соавтором этой монографии. В дальнейшем сложился творческий тандем с Евгением Юрьевичем Барзиловичем, с которым в 1971 г. была издана монография «Некоторые математические вопросы теории обслуживания сложных систем». Научная специализация Виктора Алексеевича – математические методы экстремальных задач надежности и эффективности, на основе которой 29 июня 1976 г. была защищена докторская диссертация. В 1968 г. Виктор Алексеевич перешел на работу в МИЭМ. В 1969 г. Виктор Алексеевич стал преподавать по совместительству на кафедре теории вероятностей и математической статистики недавно открытого ФПМ (факультета прикладной математики) МИЭМ (Московского института электронного машиностроения). А уже в 1979 г. он стал деканом факультета и проработал в этой должности 32 года. Параллельно с этим он работал в научной группе, организатором которой был Николай Алексеевич Северцев. В группу входили математики, представители военной отрасли и промышленности. Они занимались серьезными исследованиями, связанными со стратегическими государственными задачами. В составе этой группы Виктор Алексеевич получил Государственную премию СССР. В настоящее время Виктор Алексеевич работает ординарным профессором.

Желаем всегда оставаться оптимистом, не заикливаться на плохом, всегда настраивать себя на оптимистический лад, не позволять себе впасть в уныние.

Редколлегия

ФУНДАМЕНТАЛЬНЫЕ ОСНОВЫ ПРОБЛЕМ НАДЕЖНОСТИ И КАЧЕСТВА

FUNDAMENTALS OF RELIABILITY AND QUALITY ISSUES

УДК 338.45:622.3

doi: 10.21685/2307-4205-2024-3-1

КВАНТОВЫЕ ТЕХНОЛОГИИ – ГЛОБАЛЬНЫЙ ВЫЗОВ, ТРЕБУЮЩИЙ ОЧЕРЕДНОЙ РЕВОЛЮЦИИ МЕТОДОВ ОБРАБОТКИ И ПЕРЕДАЧИ ИНФОРМАЦИИ

Б. Ш. Собиров¹, Т. Т. Карамергенов², Н. А. Северцев³, Н. К. Юрков⁴

¹ Сургутский государственный университет, Сургут, Россия

² Военный институт Сил воздушной обороны имени дважды Героя Советского Союза Т. Я. Бегельдинова, Актобе, Республика Казахстан

³ Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Россия

⁴ Пензенский государственный университет, Пенза, Россия

¹ elenan_27@mail.ru, ² sabinak@mail.ru, ³ severs@ccas.ru, ⁴ yurkov_NK@mail.ru

Аннотация. *Актуальность и цели.* Актуальность исследования обусловлена тем, что сегодня один из наиболее серьезных фундаментальных вызовов для технологических компаний – это то, что потребности в вычислительных мощностях постоянно растут. Бизнес-среда все быстрее приближается к новой вычислительной парадигме. Одним из ответов на этот вызов могут стать квантовые технологии. Цель исследования – оценка текущего состояния квантовых вычислений и определение направления развития этого направления, а также препятствий и ограничений применения квантовых технологий. *Материалы и методы.* Эмпирической базой исследования послужили исследования института искусственного интеллекта AIRI, а также исследования СаихаBank, TotalEnergies, Quantinuum авторов В. G. Ibrahimov, А. О. Orujov, А. Н. Hasanov, К. М. Tahirova в статье «Research and analysis efficiency fiber optical communication lines using quantum technology». *Результаты.* В данном исследовании представлено главное преимущество квантовых вычислений, отмечается, что они основаны совершенно на других принципах, чем традиционные, и будут использовать собственную аппаратную и математическую базу. Представлены результаты исследований и экспериментов в области квантовых вычислений, достигнутые нефтяными компаниями. Прделанная работа и накопленный опыт позволили сформулировать ключевые вызовы развития квантовых вычислений в индустриальных компаниях. *Выводы.* Практическая значимость данного исследования заключается в том, что позволит бизнес-сообществу, а также органам государственной власти выработать стратегию развития в новой вычислительной парадигме. Позволит компаниям вырабатывать эффективные стратегии при решении задач оптимизации и перспективных вычислительных архитектур.

Ключевые слова: квантовые технологии, искусственный интеллект, бизнес-процессы, бизнес-модели, эффективность

Для цитирования: Собиров Б. Ш., Карамергенов Т. Т., Северцев Н. А., Юрков Н. К. Квантовые технологии – глобальный вызов, требующий очередной революции методов обработки и передачи информации // Надежность и качество сложных систем. 2024. № 3. С. 6–12. doi: 10.21685/2307-4205-2024-3-1

QUANTUM TECHNOLOGIES ARE A GLOBAL CHALLENGE, REQUIRING ANOTHER REVOLUTION OF METHODS PROCESSING AND TRANSMISSION OF INFORMATION

B.Sh. Sobirov¹, T.T. Karamergenov², N.A. Severtsev³, N.K. Yurkov⁴

¹ Surgut State University, Surgut, Russia

² Military Institute of Air Defense Forces named after twice Hero of the Soviet Union T.Ya. Begeldinov, Aktobe, Republic of Kazakhstan

³ Federal Research Center "Informatics and Management" of the Russian Academy of Sciences, Moscow, Russia

⁴ Penza State University, Penza, Russia

¹ elenan_27@mail.ru, ² sabinak@mail.ru, ³ severs@ccas.ru, ⁴ yurkov_NK@mail.ru

Abstract. *Background.* The relevance of the study is due to the fact that today one of the most serious fundamental challenges for technology companies is that the needs for computing power are constantly growing. The business environment is increasingly moving towards a new computing paradigm. Quantum technologies may be one of the answers to this challenge. The purpose of the study is to assess the current state of quantum computing and determine the direction of development of this area, as well as obstacles and limitations in the use of quantum technologies. *Materials and methods.* The empirical basis of the study was research from the AIRI Institute of Artificial Intelligence, as well as research from CaixaBank, TotalEnergies, Quantinuum, authors B.G. Ibrahimov, A.O. Orujov, A.H. Hasanov, K.M. Tahirova in the article "Research and analysis efficiency fiber optical communication lines using quantum technology". *Results.* The author of this study presents the main advantage of quantum computing, noting that they are based on completely different principles than traditional ones, and will use their own hardware and mathematical base. The study presents the results of research and experiments in the field of quantum computing achieved by oil companies. The work done and the experience gained made it possible to formulate the key challenges for the development of quantum computing in industrial companies. *Conclusions.* The practical significance of this research lies in the fact that it will allow the business community, as well as government authorities, to develop a development strategy in the new computing paradigm. Will allow companies to develop effective strategies when solving optimization problems and advanced computing architectures.

Keywords: quantum technologies, artificial intelligence, business processes, business models, efficiency

For citation: Sobirov B.Sh., Karamergenov T.T., Severtsev N.A., Yurkov N.K. Quantum technologies are a global challenge, requiring another revolution of methods processing and transmission of information. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):6–12. (In Russ.). doi: 10.21685/2307-4205-2024-3-1

Введение

Квантовые технологии становятся все более популярными у бизнеса – растет число инвестиций как в разработку и создание собственных исследовательских подразделений, так и в стартапы, специализирующиеся в этой области.

В последние несколько лет технологии квантовых вычислений все больше перемещаются из научных лабораторий в бизнес-среду, превращаясь из научно-исследовательской задачи в реальные приложения и инструменты. Появляются прототипы коммерческих квантовых компьютеров [1, 2]. В «квантовую гонку» включаются все больше игроков, которые рассчитывают найти дисрапт, способный изменить рынки и бизнес-модели.

Экспотенциальный рост вложений – интеллектуальных и финансовых – в область квантовых вычислений во многом обусловлен тем, что:

- почти достигнут предел кремниевой парадигмы, что требует принципиально новых решений;
- развитие искусственного интеллекта требует более мощных вычислительных возможностей;
- квантовые технологии несут новые риски информационной безопасности, что требует создания опережающих систем защиты.

Квантовые технологии могут качественно ускорить работу нейросетей, на порядки увеличив количество одновременно обрабатываемых данных, преодолеть существующие пределы наращивания сложности анализируемых моделей, в том числе экономия времени и электроэнергии на вычисление [3, 4].

Сочетание технологий искусственного интеллекта и квантовых вычислений – это именно тот драйвер, который необходим для ускорения технологического развития, поскольку последние

обещают быть более эффективными в обработке больших объемов данных и решении сложных задач.

Методика исследования

Стандартизация – необходимое условие использования новых технологий, и она активно формируется в отношении квантовых вычислений.

Канадский стартап Xanadu продемонстрировал квантовое превосходство при решении задачи выборки гауссовых бозонов (проблема выборки распределения вероятностей) на фотонном квантовом компьютере.

По мере развития рынка квантовых вычислений все более актуальной задачей становится введение единых отраслевых стандартов. Так, в апреле 2020 г. технические советы Европейского комитета по стандартизации (CEN) и Европейского комитета по электротехнической стандартизации (CENELEC) создали Оперативную группу по квантовым технологиям (FGQT).

В 2023 г. эта группа опубликовала Дорожную карту стандартизации, учитывающую совокупность потребностей Европы в регламентировании квантовых вычислений и квантовой метрологии [5–7].

Индустрия технологий квантового искусственного интеллекта, как любая зарождающаяся отрасль, сталкивается с острым дефицитом кадров и высококвалифицированных сотрудников. В этой связи очевидным является, что для практического применения квантовых технологий критически важно преодолеть кадровый голод – выстроить эффективную систему подготовки специалистов новой индустрии.

Инвестиции в стартапы технологий квантового искусственного интеллекта по имущественному типу инвесторов за 2001–2022 гг. представлены на рис. 1.

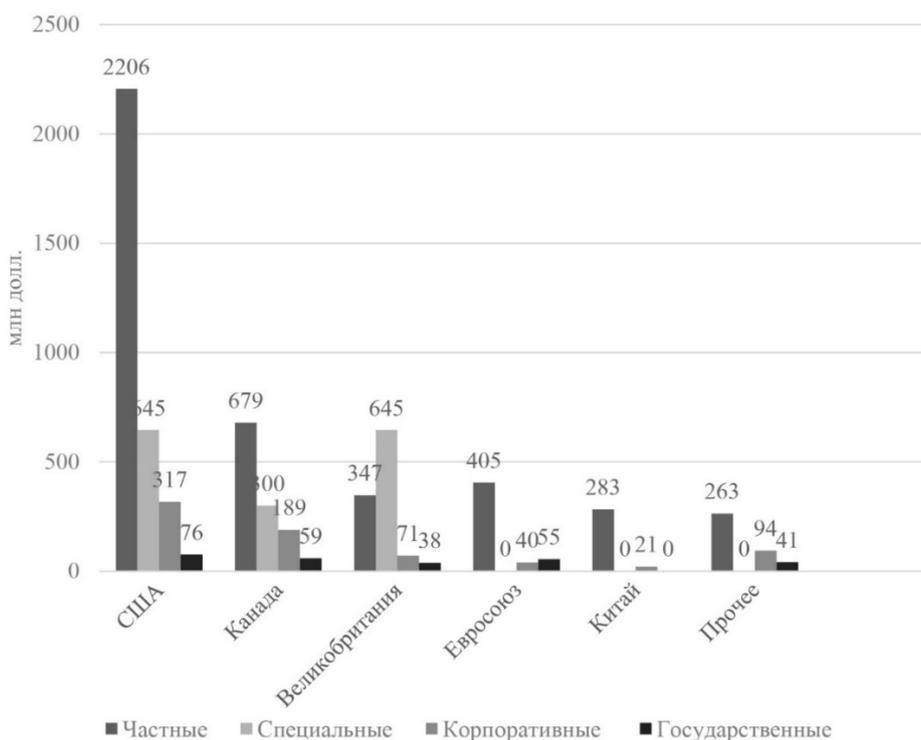


Рис. 1. Инвестиции в стартапы технологий квантового искусственного интеллекта по имущественному типу инвесторов за 2001–2022 гг., млн долл. [8, 9]

В неограниченный потенциал квантового искусственного интеллекта продолжают верить инвесторы – в 2022 г. общий объем вложений в стартапы в области технологий квантового искусственного интеллекта достиг 2,35 млрд долл., что на 1 % выше, чем в 2021 г.

Именно в 2021–2022 гг. были заключены крупнейшие инвестиционные сделки в стартапы технологий квантового искусственного интеллекта. 8 из 10 крупнейших сделок 2022 г. были заключены с компаниями-производителями оборудования, которые являются наиболее капиталоемкой частью цепочки создания стоимости квантовых технологий.

Бизнес-задачи в финансовой отрасли, требующие больших вычислительных возможностей

Финтех – одна из индустрий, которая может существенно видоизмениться под влиянием технологий квантового искусственного интеллекта. Первые эксперименты в этой области финансовые компании начали уже несколько лет назад: они разрабатывают алгоритмы, пользуются облачным доступом к квантовым компьютерам, совмещают квантовые вычисления с классическими [7], ускоряя таким образом расчеты.

Задачи для квантовых вычислений в финансах:

1. Оптимизационные модели.

Структура активов в оптимальном портфеле (оптимизационные модели):

- оптимальная торговая траектория;
- оптимальные возможности арбитража;
- оптимальный выбор функций в кредитном скоринге.

2. Модели глубокого обучения.

Обнаружение возможностей получения прибыли среди активов при торговле:

- прогнозирование рынков;
- классификация данных;
- регрессия.

3. Методы Монте-Карло.

Оценка риска и доходности портфеля или целевой компании:

- ценообразование финансовых деривативов;
- анализ риска;
- другие задачи.

Один из пионеров по изучению технологий квантового искусственного интеллекта для финтеха – CaixaBank. В 2022 г. готовый облачный квантовый сервис компании D-Wave позволил ускорить расчеты оптимального инвестиционного портфеля на 90 %. В 2020 г. банк разработал первый квантовый алгоритм машинного обучения для классификации рисков в банковском деле страны. Он объединил квантовые и облачные вычисления на разных этапах процесса расчета – для классификации профилей кредитного риска.

Технологии квантового искусственного интеллекта – другое не менее важное направление экспериментов банка. Их протестировали в целях минимизации финансовых рисков двух видов портфелей: ипотечных кредитов и казначейских векселей. Эксперимент показал, что технологии вычисления квантового искусственного интеллекта позволяют прийти к тем же выводам, что и традиционный, но при этом требуется лишь несколько десятков симуляций вместо тысяч и миллионов, проводимых в случае классических вычислений.

Российские банки также экспериментируют с технологиями квантового искусственного интеллекта и ищут для них практические бизнес-задачи. В частности, Сбербанк активно реализует различные инициативы в области квантовых вычислений.

В 2021 г. был запущен совместный исследовательский проект банка с ООО СП «Квант», направленный на развитие методологии оценки квантового влияния на бизнес, его устойчивость и конкурентоспособность.

Задачи, решаемые бизнесом с помощью технологий квантового искусственного интеллекта

Бизнес нефтяных компаний связан с решением большого количества вычислительно сложных задач, к которым, например, относятся обработка и интерпретация сейсморазведочных работ, моделирование геологии и гидродинамики течения флюида в пласте, логистика доставки продукции до потребителя и многие другие [8].

Запущенная программа направлена в первую очередь на разработку приложений, позволяющих решать практические задачи бизнеса. Исследования фокусируются на разработке квантовых алгоритмов, ИИ и программного обеспечения.

Нефтяная индустрия потенциально может получить выгоды от применения технологий квантового искусственного интеллекта в решении широкого спектра задач: начиная от универсальных, таких как планирование деятельности, и заканчивая специфическими, например, моделирование на уровне отдельных химических элементов.

Так, на практике участники рынка выделяют четыре ключевые категории прикладных задач, актуальных для нефтяной индустрии:

- 1) моделирование физических систем: они зачастую описываются дифференциальными уравнениями и сводятся к решению линейных систем;
- 2) оптимизационные задачи, включающие задачи дискретной, непрерывной и смешанной дискретно-непрерывной оптимизации;
- 3) молекулярное моделирование и квантовая химия;
- 4) квантовое машинное обучение.

Например, для улучшения плана разбуривания месторождения была построена модель квадратичной оптимизации. Результаты ее тестирования на устройствах квантового отжига D-Wave показали увеличение эффективности планирования на 15 %.

TotalEnergies исследует возможности технологий квантового искусственного интеллекта по улучшению процессов улавливания углекислого газа. Вместе с командой Quantinuum компания представила предпринт, в котором подробно говорится о потенциальном использовании квантовых вычислений для смягчения последствий изменения климата.

Исследовательская группа разработала методологию квантовых вычислений, описывающую связывание молекулярного диоксида углерода с материалом, который активно исследуется для улавливания углерода – металлоорганическим каркасом.

Задачи нефтяной индустрии, решение которых даст наибольший эффект в случае применения алгоритмов квантового искусственного интеллекта:

- гидродинамическое моделирование в пористой среде;
- полноволновая сейсмическая инверсия;
- сейсмическая томография;
- регуляризация сейсмических данных;
- сейсмическая миграция;
- оптимальное планирование производства нефтепродуктов;
- динамическая оптимизация работы установок нефтеперерабатывающего завода;
- оптимизация логистики перемещения буровых установок;
- моделирование каталитического разложения метана;
- оптимизация работы химического реактора.

Для скорейшего перевода квантовых вычислений в практическую плоскость необходимо активизировать усилия по разработке квантово-вдохновленных устройств для оптимизации.

Заключение

В настоящее время в России и мире наблюдается стремительный рост интереса к технологиям квантового искусственного интеллекта. Они имеют революционный потенциал для различных отраслей экономики и научных направлений.

Одним из вариантов реализации этого направления могло бы стать объединение усилий ведущих бизнес- и исследовательских групп в формате «квантовой сети» из национальных точек роста на базе университетов и технологических центров при активном участии бизнес-лидеров в постановке задач и внедрении на рынки.

Важность задач требует более активного внедрения соответствующих образовательных инициатив и, главное, обеспечения возможностей для профессиональной реализации молодых специалистов.

В этой связи важным является не пропустить момент «квантового превосходства». Первые компании, которые его достигнут, получат уникальное стратегическое конкурентное преимущество, что приведет к качественному скачку в эффективности существующих и потенциальных бизнес-моделей.

Список литературы

1. Майстер В. А., Ширинкина Е. В. Цифровая трансформация финансовой отрасли как фактор конкурентоспособности // Труды Международного симпозиума Надежность и качество. 2018. Т. 2. С. 269–271.
2. Шевчук Д. С., Федоров А. К., Ильин И. В. Синергия технологических и организационных инноваций: развитие экосистемы квантовых технологий для противостояния вызовам человечества // Экономика и предпринимательство. 2021. № 2 (127). С. 188–191.

3. Крылов В. С., Бекирова Э. А., Ветеранова Д. С. Сквозная квантовая технология цифровой экономики // Ученые записки Крымского инженерно-педагогического университета. 2020. № 1 (67). С. 129–134.
4. Серов Е. Р., Васильев С. А. Применение квантовых технологий в банковском бизнесе // Экономика и управление. 2023. Т. 29, № 3. С. 248–255.
5. Салыгин В. И., Лобов Д. С. Перспективы применения технологий квантового распределения ключей на примере объектов нефтегазовой отрасли // Друкерровский вестник. 2023. № 1 (51). С. 246–253.
6. Ширинкина Е. В. Оценка трендов цифровой трансформации финансовой отрасли // Надежность и качество сложных систем. 2019. № 2 (26). С. 114–120.
7. Ширинкина Е. В. Цифровые кадры как новый показатель качества человеческого капитала // Современная научная мысль. 2018. № 4. С. 181–184.
8. Ibrahimov B. G., Orujov A. O., Hasanov A. H., Tahirova K. M. Research and analysis efficiency fiber optical communication lines using quantum technology // T-Comm. 2021. Vol. 15, № 10. P. 50–54.
9. Ulyanov S. V., Reshetnikov A. G., Zrelova D. P. Industrial robotic intelligent robust control system: applying quantum soft computing technologies and quantum software engineering in unpredicted control situations // Software & Systems. 2023. № 1. P. 26–45.

References

1. Mayster V.A., Shirinkina E.V. Digital transformation of the financial industry as a factor of competitiveness. *Trudy Mezhdunarodnogo simpoziuma Nadezhnost' i kachestvo = Proceedings of the International Symposium Reliability and Quality*. 2018;2:269–271. (In Russ.)
2. Shevchuk D.S., Fedorov A.K., Il'in I.V. Synergy of technological and organizational innovations: the development of an ecosystem of quantum technologies to confront the challenges of humanity. *Ekonomika i predprinimatel'stvo = Economics and entrepreneurship*. 2021;(2):188–191. (In Russ.)
3. Krylov V.S., Bekirova E.A., Veteranova D.S. End-to-end quantum technology of digital economy. *Uchenye zapiski Krymskogo inzhenerno-pedagogicheskogo universiteta = Scientific notes of the Crimean Engineering Pedagogical University*. 2020;(1):129–134. (In Russ.)
4. Serov E.R., Vasil'ev S.A. Application of quantum technologies in banking business. *Ekonomika i upravlenie = Economics and Management*. 2023;29(3):248–255. (In Russ.)
5. Salygin V.I., Lobov D.S. Prospects for the application of quantum key distribution technologies on the example of oil and gas industry facilities. *Drukerovskiy vestnik = Drucker's Bulletin*. 2023;(1):246–253. (In Russ.)
6. Shirinkina E.V. Assessment of trends in the digital transformation of the financial industry. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems*. 2019;(2):114–120. (In Russ.)
7. Shirinkina E.V. Digital personnel as a new indicator of the quality of human capital. *Sovremennaya nauchnaya mysl' = Modern scientific thought*. 2018;(4):181–184. (In Russ.)
8. Ibrahimov B.G., Orujov A.O., Hasanov A.H., Tahirova K.M. Research and analysis efficiency fiber optical communication lines using quantum technology. *T-Comm*. 2021;15(10):50–54.
9. Ulyanov S.V., Reshetnikov A.G., Zrelova D.P. Industrial robotic intelligent robust control system: applying quantum soft computing technologies and quantum software engineering in unpredicted control situations. *Software & Systems*. 2023;(1):26–45.

Информация об авторах / Information about the authors

Бежан Шукриллоевич Собиров

старший преподаватель кафедры менеджмента и бизнеса,
Сургутский государственный университет
(Россия, г. Сургут, пр-т Ленина, 1)
E-mail: elenan_27@mail.ru

Bezhan Sh. Sobirov

Senior lecturer of the sub-department of management and business,
Surgut State University
(1 Lenin avenue, Surgut, Russia)

Тимур Тыныштыкович Карамергенов

начальник кафедры воздушной навигации и боевого управления авиации учебно-методического управления, Военный институт Сил воздушной обороны имени дважды Героя Советского Союза Т. Я. Бегельдинова
(Республика Казахстан, Актобе, пр-т Алии Молдагуловой, 16)
E-mail: sabinak@mail.ru

Timur T. Karamergenov

Head of the sub-department of air navigation and combat aviation control of the educational and methodological department, Military Institute of Air Defense Forces named after twice Hero of the Soviet Union T.Ya. Begeldinov
(16 Aliya Moldagulova avenue, Aktobe, Republic of Kazakhstan)

Николай Алексеевич Северцев

доктор технических наук, профессор,
главный научный сотрудник отдела управления
робототехническими устройствами,
Федеральный исследовательский центр
«Информатика и управление»
Российской академии наук (Вычислительный центр
имени А. А. Дородницына РАН)
(Россия, г. Москва, ул. Вавилова, 40)
E-mail: severs@ccas.ru

Николай Кондратьевич Юрков

доктор технических наук, профессор,
заслуженный деятель науки РФ,
заведующий кафедрой конструирования
и производства радиоаппаратуры,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: yurkov_NK@mail.ru

Nikolay A. Severtsev

Doctor of technical sciences, professor,
chief researcher of the department of control
of robotic devices,
Federal Research Center "Informatics and Management"
of the Russian Academy of Sciences
(Dorodnitsyn computer center
of the Russian Academy of Sciences)
(40 Vavilova street, Moscow, Russia)

Nikolay K. Yurkov

Doctor of technical sciences, professor,
the honoured worker of science
of the Russian Federation,
head of the sub-department
of radio equipment design and production,
Penza State University
(40 Krasnaya street, Penza, Russia)

Авторы заявляют об отсутствии конфликта интересов /

The authors declare no conflicts of interests.

Поступила в редакцию/Received 30.05.2024

Поступила после рецензирования/Revised 27.06.2024

Принята к публикации/Accepted 29.07.2024

ОЦЕНКА КАЧЕСТВА РАБОТЫ МНОГОКРИТЕРИАЛЬНОГО НЕЙРОСЕТЕВОГО СТАТИСТИЧЕСКОГО АНАЛИЗА НА ПРИМЕРЕ ПРОВЕРКИ ГИПОТЕЗЫ НЕЗАВИСИМОСТИ ДАННЫХ МАЛЫХ ВЫБОРОК

А. И. Иванов¹, А. И. Ермакова²

¹ Пензенский научно-исследовательский электротехнический институт, Пенза, Россия

² Пензенский государственный университет, Пенза, Россия

¹ ivan@pniei.penza.ru, ² aermakova27061992@mail.ru

Аннотация. *Актуальность и цели.* Целью статьи является создание критерия качества оценки эффективности нейросетевой обработки малых выборок при многокритериальной проверке гипотезы независимости. *Материалы и методы.* На сегодняшний день известно порядка 200 статистических критериев проверки гипотезы независимости малой выборки. Используя их совместно, удастся существенно повысить качество принимаемых нейросетью решений. В связи с этим возникает проблема эффективной сравнительной оценки достигнутого качества той или иной нейросетью. *Результаты и выводы.* Предложено оценивать качество нейросетевого решения по достигнутому увеличению выборки при ее анализе одним-единственным критерием Пирсона – Эдлтона – Эджурта конца 19 в. на выборке большого объема в наиболее сложной ситуации проверки гипотезы независимости.

Ключевые слова: проверка гипотезы независимости, коэффициент корреляции, малые выборки, нейросетевое объединение большого числа критериев

Для цитирования: Иванов А. И., Ермакова А. И. Оценка качества работы многокритериального нейросетевого статистического анализа на примере проверки гипотезы независимости данных малых выборок // Надежность и качество сложных систем. 2024. № 3. С. 13–18. doi: 10.21685/2307-4205-2024-3-2

ASSESSING THE PERFORMANCE QUALITY OF MULTICRITERIA NEURAL NETWORK STATISTICAL ANALYSIS USING THE EXAMPLE OF TESTING THE HYPOTHESIS OF INDEPENDENCE OF SMALL SAMPLE DATA

A.I. Ivanov¹, A.I. Ermakova²

¹ Penza Scientific Research Electrotechnical Institute, Penza, Russia

² Penza State University, Penza, Russia

¹ ivan@pniei.penza.ru, ² aermakova27061992@mail.ru

Abstract. *Background.* The purpose of the article is to create a quality criterion for assessing the effectiveness of neural network processing of small samples in multicriteria testing of the independence hypothesis. *Materials and methods.* To date, about 200 statistical criteria for testing the hypothesis of independence of a small sample are known. Using them together can significantly improve the quality of decisions made by the neural network. In this regard, the problem of effective comparative assessment of the achieved quality by one or another neural network arises. *Results and conclusions.* It is proposed to evaluate the quality of a neural network solution by the achieved increase in the sample when analyzing it using a single Pearson-Edleton-Edgeworth criterion of the late 19th century on a large sample in the most difficult situation of testing the independence hypothesis.

Keywords: testing the independence hypothesis, correlation coefficient, small samples, neural network combination of a large number of criteria

For citation: Ivanov A.I., Ermakova A.I. Assessing the performance quality of multicriteria neural network statistical analysis using the example of testing the hypothesis of independence of small sample data. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):13–18. (In Russ.). doi: 10.21685/2307-4205-2024-3-2

Постановка задачи

В конце 19 в. основатели современной математической статистики Пирсон, Эдлтон, Эджуорт для оценки коэффициентов корреляции активно использовали ставшую сегодня классической формулу [1]:

$$r_1(x, y) \approx \sum_{i=1}^{16} \frac{(x_i - M(x)) \cdot (y_i - M(y))}{\sigma(x) \cdot \sigma(y)}, \quad (1)$$

где $M(\cdot)$ – функционал вычисления математического ожидания малой выборки в 16 опытов; $\sigma(\cdot)$ – функционал вычисления стандартного отклонения.

К сожалению, классическая формула (1) дает требуемую практикой достоверную вероятность оценок только при больших выборках. При малых выборках в 16 опытов для слабо коррелированных данных $r \approx 0.0$, ошибка может достигать значений $\Delta r = \pm 0.7$. Столь значительный интервал ошибок недопустим при анализе биометрических данных медицины, биологии, экономики, психологии, педагогики. Сегодня статистические оценки, выполненные на малых выборках, воспринимаются физико-математической общественностью как недостоверные. Последнее сильно влияет на материальные затраты, связанные с обучением следующего поколения исследователей, так как увеличивает срок сбора реальных данных в аспирантуре или адъюнктуре. Так, аспирант биолог за три года обучения в аспирантуре вполне способен вырастить и исследовать 16 кроликов, опираясь только на собственные усилия. При этом к концу своего обучения он столкнется с недоверием коллег к его статистическим оценкам в его диссертации. Ситуация меняется кардинально, если он успеет вырастить и исследовать 160 кроликов, однако при этом сильно увеличиваются материально-технические затраты и время на проведение натурального эксперимента.

Положение кардинально меняется, если отказаться от практики прошлого века, когда исследователи-одиночки старались под свою конкретную задачу создать свой наилучший статистический критерий [1], учитывающий некоторую совокупность граничных условий. В 21 в. парадигма вычислительных оценок изменилась, появилась возможность использовать большие сети искусственных нейронов, т.е. каждому из известных статистических критериев (их более 200 приведено в справочнике [1]). Каждому из известных статистических критериев можно поставить в соответствие свой искусственный нейрон, далее несколько нейронов могут быть объединены в одну нейросеть [2].

Практика показала, что нейросетевое объединение примерно 20 классических статистических критериев позволяет малую исходную выборку реальных данных в 16 опытов виртуально увеличить, добавив к ней еще примерно 20 виртуальных опытов (каждый искусственный нейрон увеличивает выборку примерно на один виртуальный опыт). Соответственно, диссертация аспиранта-биолога могла бы выглядеть более убедительной, если бы он обрабатывал данные, полученные не только от 16 реальных кроликов, а от существенно большей выборки в 36 виртуально-реальных кроликов. Именно это и дает некоторый дополнительный эффект роста доверия к многокритериальной статистической обработке данных в сравнении с однокритериальной обработкой.

Приведенный выше тезис доказан численными экспериментами для достаточно больших сетей, собранных из бинарных искусственных нейронов [3]. При этом на сегодня по реальным данным нейрофизиологов [4] у живых существ нет бинарных естественных нейронов. Все живые существа используют более сложные q-арные естественные нейроны [5].

Живые существа не используют бинарные естественные нейроны в силу того, что они способны работать, выполняя параллельный анализ только двух статистических гипотез [2, 3]. Если искусственные нейроны делать троичными ($q = 3$), то они становятся способными выполнять параллельный анализ сразу трех статистических гипотез. Это приводит к существенному росту доверия для новых нейросетевых решений с трехуровневыми выходными квантователями [6, 7].

Выигрыш от первого шага замены бинарных нейронов на троичные нейроны подтвержден имитационным моделированием. Преимущества троичных нейронов над бинарными очевидны [6, 7], и именно по этой причине природа не использует бинарные нейроны в живых организмах [4, 5]. Однако и троичные естественные нейроны, видимо, используются природой достаточно редко. Предположительно дальнейшее повышение числа состояний выходного квантователя должно приводить к росту качества решений, принимаемых сетями q-арных искусственных нейронов [8, 9]. Рисунок 1 иллюстрирует этот тезис, на нем отображен 9-арный искусственный нейрон, полученный 9-арным квантованием выходных состояний функционала оценки коэффициентов корреляции (1) для малой выборки в 16 опытов.

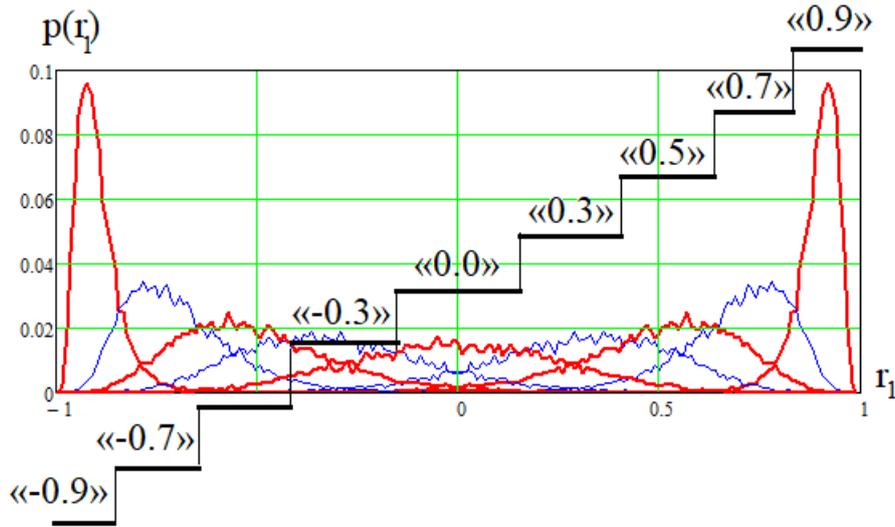


Рис. 1. Распределения плотности вероятности значений – $p(r_1)$ коэффициентов корреляции, вычисленных по формуле (1) на малой выборке в 16 опытов

Из рисунка видно, что центр каждого стабильного участка 9-арного квантователя искусственного нейрона Пирсона – Эдлтона – Эджуорта совпадает с математическим ожиданием 9 использованных для настройки плотностей распределения. Очевидно, что столь узкие интервалы стабильности или округления данных 9-арного квантователя должны давать значительные ошибки оценок для одного искусственного нейрона. Однако использование десятков подобных искусственных нейронов [3] после усреднения их откликов должно давать достаточно точные итоговые результаты нейровычислений. В этой ситуации возникает задача сравнения данных разных нейросетевых вычислителей корреляции в рамках приведения их данных к одной шкале, соответствующей классической формуле Пирсона – Эдлтона – Эджуорта.

**Эталонная точка шкалы Пирсона – Эдлтона – Эджуорта,
удобная для сравнения качества работы тестируемых нейросетей**

Из данных рис. 1 следует, что интервал ошибок Δr существенно зависит от математического ожидания контролируемого параметра – $M(r)$. Так, для пяти положительных значений математических ожиданий интервалы ошибок и амплитуда ошибок составят:

$$\begin{cases} M(r) = 0,0 \rightarrow -0,7 < \Delta r < +0,7 \rightarrow A(r) \approx 1,4, \\ M(r) = +0,3 \rightarrow -0,3 < \Delta r < +0,8 \rightarrow A(r) \approx 1,1, \\ M(r) = 0,5 \rightarrow +0,0 < \Delta r < +0,85 \rightarrow A(r) \approx 0,85, \\ M(r) = 0,7 \rightarrow +0,3 < \Delta r < +0,95 \rightarrow A(r) \approx 0,65, \\ M(r) = 0,9 \rightarrow +0,6 < \Delta r < +1,0 \rightarrow A(r) \approx 0,4, \end{cases} \quad (2)$$

где $A(r) = (\max(r) - \min(r))$ – амплитуда или размах интервала ошибок Δr .

При любом конечном объеме выборки реальных данных оценка коэффициентов корреляции по формуле (1) хуже всех (имеет наибольшую амплитуду ошибок) при $M(r) = 0,0$. По мере увеличения значений коэффициентов корреляции ошибка его оценок падает (точность увеличивается примерно в три раза).

Несмотря на то, что шкала оценок коэффициентов корреляции Пирсона – Эдлтона – Эджуорта имеет максимальную амплитуду ошибок в точке $M(r) = 0,0$, именно эту точку следует выбирать как эталонную. Это связано с тем, что обеспечить попадание в эту точку проще всего. Для этого достаточно двукратного программного обращения к вызову программного генератора псевдослучайных чисел. Существующие программные генераторы дают независимые (некоррелированные) данные. Кроме того, эта точка стабильна и не меняется для любого объема малых выборок.

На рис. 2 представлены результаты численного эксперимента по имитации плотностей распределения коэффициентов корреляции независимых данных для выборок разного объема.

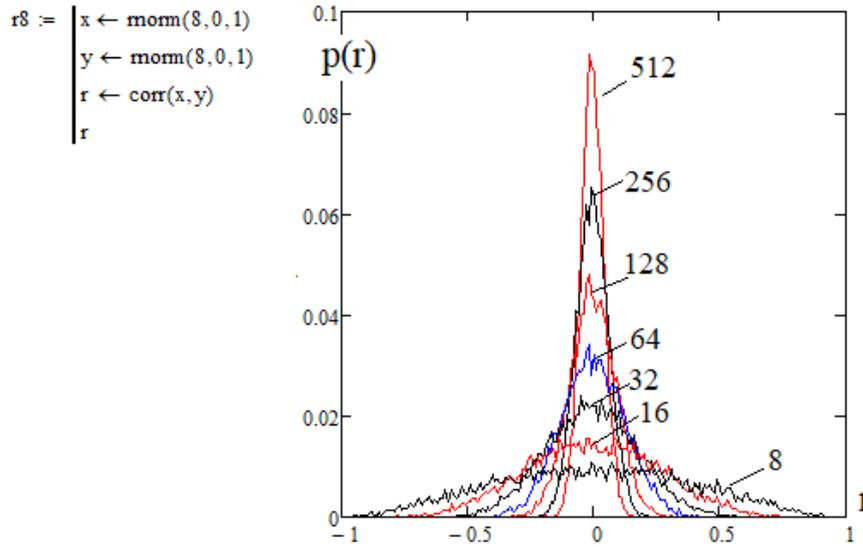


Рис. 2. Распределения значений коэффициентов корреляции для выборок разного объема в 8, 16, 32, 64, 128, 256, 512 опытов

Еще одной причиной выбора точки $M(r) = 0,0$ при тестировании качества работы нейросетевых вычислителей является то, что именно в этой точке закон плотности распределения значений коэффициентов корреляции близок к нормальному для выборки любого объема. Во всех иных точках закон распределения оказывается асимметричным и, соответственно, уже не описывается двумя первыми статистическими моментами.

В левой части рис. 2 приведена программа на языке MathCAD, позволившая получить данные, отображенные в правой части рисунка.

Только в одной точке $M(r) = 0,0$ мы имеем самое простое статистическое описание распределения данных, чье стандартное отклонение $\sigma(r)$ является простой функцией, описывающей как связь объема выборки с интервалом неопределенности Δr , так и связи объема выборки с амплитудой интервала неопределенности при доверительной вероятности 0,99:

$$\begin{cases} \Delta r \approx 6\sigma(r), \\ A(r) \approx 6\sigma(r). \end{cases} \quad (3)$$

Оба статистических параметра совпадают $\Delta r = A(r)$ только в точке $M(r) = 0,0$. Формально по любому из статистических параметров системы уравнений (3) может быть найден объем эквивалентной входной выборки. Графики связи стандартного отклонения с объемом выборки приведены на рис. 3.

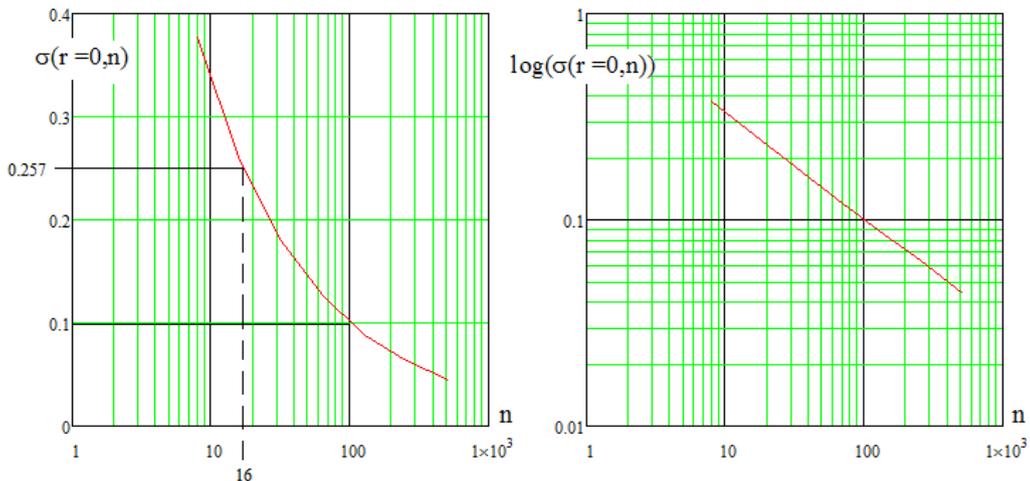


Рис. 3. Обратная оценка объема выборки по значению стандартного отклонения в двух типах систем координат (полулогарифмической системе и полностью логарифмической)

Приведенными на рис. 3 номограммами (как правой, так и левой) следует пользоваться для решения обратной задачи пересчета стандартного отклонения в размер эквивалентной выборки. В качестве примера рассмотрим ситуацию стандартного отклонения $\sigma(r = 00, n = 16) = 0,257$. Для этой ситуации мы имеем дело с простейшим нейросетевым анализатором с одним нейроном Пирсона – Эдлтона – Эджуорта, анализирующим малую выборку объемом в 16 опытов.

Однако, если мы наблюдаем ситуацию $\sigma(r = 00, n = 16) = 0,1$, то по номограмме рис. 3 мы должны сделать вывод о том, что для одного эталонного нейрона Пирсона – Эдлтона – Эджуорта малая выборка должна иметь объем $N = 100$ вместо реальной выборки объемом $n = 16$. Это вполне возможно из-за добавления в сеть нескольких дополнительных нейронов. Даже если мы не знаем число добавленных дополнительных нейронов и их тип, мы должны прийти к выводу, что тестируемая нейросеть по сравнению с эталонной формулой (1) снизила вероятность ошибок в 2,57 раза, что эквивалентно увеличению объема входной выборки с 16 до 100 опытов. При этом 16 опытов являются реальными, а дополнительные 84 виртуальными. Дополнительные виртуальные опыты появляются из-за более глубокой нейросетевой обработки исходных данных.

Также следует отметить, представленные на рис. 3 зависимости аналитически описываются в разных формах. Зависимость в левой части рисунка хорошо приближается квадратичным полиномом. Зависимость в правой части удобна тем, что хорошо описывается линейно. Параллельно с зависимостями $\sigma(r = 00, n)$ и $\log(\sigma(r = 00, n))$ могут быть использованы аналогичные зависимости $A(r = 00, n)$ и $\log(A(r = 00, n))$ в соответствии с соотношением (3). Использование двух зависимостей позволяет поднять точность оценок через их усреднения. Корреляционная связь для этих двух разных статистических параметров отсутствует.

Список литературы

1. Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников. М. : ФИЗМАТЛИТ, 2006. 816 с.
2. Иванов А. И. Искусственные математические молекулы: повышение точности статистических оценок на малых выборках (программы на языке MathCAD) : препринт. Пенза : Из-во ПГУ, 2020. 36 с.
3. Иванов А. И. Нейросетевой многокритериальный статистический анализ малых выборок. Проверка гипотезы независимости : справочник. Пенза : Изд-во ПГУ, 2022. 218 с.
4. Николлс Д., Мартин Р., Валлас Б., Фукс П. От нейрона к мозгу / пер. с англ. П. М. Балабана, А. В. Галкина, Р. А. Гиниатуллина [и др.]. М. : Едиториал УРСС, 2003. 672 с.
5. Иванов А. И. Малые выборки, нейроморфные вычисления: быстрые алгоритмы оценки энтропии Шеннона – Пирсона квадратичной сложности : справочник. Пенза : Изд-во ПГУ, 2023. 32 с.
6. Технический отчет для ТК 164. «Защищенные приложения искусственного интеллекта: модификация алгоритма автоматического обучения бинарных перцептронов по ГОСТ Р 52633.5-2011 под тройные искусственные нейроны». Пенза : Пензенский государственный университет, 2023. Инв. № 1/16-23 от 06.07.23.
7. Иванов А. И., Юнин А. П., Бояршинов М. А., Иванов А. П. Перспектива совместного использования двоичных и троичных искусственных нейронов при анализе качества «белого» шума в пространстве сверток Хэмминга, вычисленных по разным модулям // Интеллектуальные системы в производстве. 2022. Т. 20, № 3. С. 88–92. doi: 10.22213/2410-9304-2022-3-88-93
8. Иванов А. И., Иванов А. П., Савинов К. Н., Еременко Р. В. Виртуальное усиление эффекта распараллеливания вычислений при переходе от бинарных нейронов к использованию Q-арных искусственных нейронов // Надежность и качество сложных систем. 2022. № 4. С. 89–97.

References

1. Kobzar' A.I. *Prikladnaya matematicheskaya statistika. Dlya inzhenerov i nauchnykh rabotnikov = Applied mathematical statistics. For engineers and scientists*. Moscow: FIZMATLIT, 2006:816. (In Russ.)
2. Ivanov A.I. *Iskusstvennye matematicheskie molekuly: povyshenie tochnosti statisticheskikh otsenok na malykh vyborokakh (programmy na yazyke MathCAD): preprint = Artificial mathematical molecules: improving the accuracy of statistical estimates on small samples (programs in MathCAD) : preprint*. Penza: Iz-vo PGU, 2020:36. (In Russ.)
3. Ivanov A.I. *Neyrosetevoy mnogokriterial'nyy statisticheskiy analiz malykh vyborok. Proverka gipotezy nezavisimosti: spravochnik = Neural network multicriteria statistical analysis of small samples. Testing the independence hypothesis: handbook*. Penza: Izd-vo PGU, 2022:218. (In Russ.)
4. Nikolls D., Martin R., Vallas B., Fuks P. *Ot neyrona k mozgu: per. s angl. P.M. Balabana, A.V. Galkina, R.A. Giniatullina [i dr.] = From neuron to brain : translated from English by P.M. Balaban, A.V. Galkin, R.A. Giniatullina [et al.]*. Moscow: Editorial URSS, 2003:672. (In Russ.)

5. Ivanov A.I. *Malye vyborki, neyromorfnye vychisleniya: bystrye algoritmy otsenki entropii Shennona – Pirsona kvadrachnoy slozhnosti: spravochnik = Small samples, neuromorphic calculations: fast algorithms for estimating Shannon–Pearson entropy of quadratic complexity: handbook*. Penza: Izd-vo PGU, 2023:32. (In Russ.)
6. *Tekhnicheskij otchet dlya TK 164. «Zashchishchennyye prilozheniya iskusstvennogo intellekta: modifikatsiya algoritma avtomaticheskogo obucheniya binarnykh perseptronov po GOST R 52633.5-2011 pod troichnye iskusstvennye neyrony» = Technical report for TC 164. "Protected applications of artificial intelligence: modification of the algorithm for automatic training of binary perceptrons according to GOST R 52633.5-2011 for ternary artificial neurons"*. Penza: Penzenskiy gosudarstvennyy universitet, 2023;(Inv. № 1/16-23 ot 06.07.23). (In Russ.)
7. Ivanov A.I., Yunin A.P., Boyarshinov M.A., Ivanov A.P. The prospect of joint use of binary and ternary artificial neurons in the analysis of the quality of "white" noise in the space of Hamming convolutions calculated using different modules. *Intellektual'nye sistemy v proizvodstve = Intelligent systems in production*. 2022;20(3):88–92. (In Russ.). doi: 10.22213/2410-9304-2022-3-88-93
8. Ivanov A.I., Ivanov A.P., Savinov K.N., Eremenko R.V. Virtual enhancement of the effect of parallelization of calculations during the transition from binary neurons to the use of Q-ary artificial neurons. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems*. 2022;(4):89–97. (In Russ.)

Информация об авторах / Information about the authors

Александр Иванович Иванов

доктор технических наук, профессор,
научный консультант,
Пензенский научно-исследовательский
электротехнический институт
(Россия, г. Пенза, ул. Советская, 9)
E-mail: ivan@pniei.penza.ru

Aleksandr I. Ivanov

Doctor of technical sciences, professor,
scientific consultant,
Penza Scientific Research Electrotechnical Institute
(9 Sovetskaya street, Penza, Russia)

Анна Игоревна Ермакова

преподаватель кафедры радио- и спутниковой связи,
Военный учебный центр,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: aermakova27061992@mail.ru

Anna I. Ermakova

Lecturer of the sub-department of radio
and satellite communications,
Military Training Center,
Penza State University
(40 Krasnaya street, Penza, Russia)

Авторы заявляют об отсутствии конфликта интересов /

The authors declare no conflicts of interests.

Поступила в редакцию/Received 10.06.2024

Поступила после рецензирования/Revised 05.07.2024

Принята к публикации/Accepted 30.08.2024

ФИЗИЧЕСКИЕ ОСОБЕННОСТИ ПРОЦЕССОВ ДЕФОРМАЦИИ И РАЗРУШЕНИЯ МАТЕРИАЛА МЕТОДОМ КОНЕЧНЫХ УЗЛОВ

А. И. Годунов¹, С. А. Куканов², Д. В. Бакланов³

¹ Пензенский государственный университет, Пенза, Россия

^{2,3} Филиал Военной академии материально-технического обеспечения имени генерала армии А. В. Хрулева в г. Пензе, Пенза, Россия

¹ avitelpgu@mail.ru, ² Kuk_@mail.ru, ³ baklanov90@yandex.ru

Аннотация. *Актуальность и цели.* Изучены деформации и разрушение материала композитной гильзы в области контакта двух различных материалов (стали и полиамида) с помощью математической модели разрывного метода Галеркина. Особенностью исследования является то, что данный метод сохраняет массовые и кинетические характеристики в узлах модели, разрушение происходит при достижении предела текучести. Также для адекватного описания разрушения для каждого материала задано уравнение состояния вещества по Ми – Грюнейзену. *Материалы и методы.* Рассмотрен порядок проведения моделирования с целью оптимизации и выбора оптимальных технических решений с помощью программных комплексов моделирования разрушения на основе метода конечных узлов и сопоставления полученных результатов с математическими расчетами на основе математической модели разрывного метода Галеркина. *Результаты и выводы.* По итогу исследования решена задача расчета на прочность разработанной композитной гильзы не только по всей ее длине, но и на стыке двух различных материалов с учетом перерезывающего усилия в области их адгезивного соединения.

Ключевые слова: разрывной метод Галеркина, функция формы EFG, локальное граничное интегрирование, эрозия

Для цитирования: Годунов А. И., Куканов С. А., Бакланов Д. В. Физические особенности процессов деформации и разрушения материала методом конечных узлов // Надежность и качество сложных систем. 2024. № 3. С. 19–29. doi: 10.21685/2307-4205-2024-3-3

PHYSICAL FEATURES OF THE PROCESSES OF DEFORMATION AND DESTRUCTION OF THE MATERIAL BY THE FINITE NODE METHOD

A.I. Godunov¹, S.A. Kukanov², D.V. Baklanov³

¹ Penza State University, Penza, Russia

^{2,3} Branch of the Military Academy of Logistics named after Army General A.V. Khrulev in Penza, Penza, Russia

¹ avitelpgu@mail.ru, ² Kuk_@mail.ru, ³ baklanov90@yandex.ru

Abstract. *Background.* In this article, deformations and destruction of the composite sleeve material in the contact area of two different materials (steel and polyamide) are studied using a mathematical model of the Galerkin discontinuous method. The peculiarity of the study is that this method preserves the mass and kinetic characteristics in the nodes of the model, destruction occurs when the yield point is reached. Also, for an adequate description of the destruction, the Mi-Grüneisen equation of state of matter is given for each material. *Materials and methods.* The procedure of modeling is considered in order to optimize and select optimal technical solutions using software systems for modeling fracture based on the finite node method and comparing the results obtained with mathematical calculations based on the mathematical model of the discontinuous Galerkin method. *Results and conclusions.* As a result of the study, the problem of calculating the strength of the developed composite sleeve was solved not only along its entire length, but also at the junction of two different materials, taking into account the shearing force in the area of their adhesive connection.

Keywords: discontinuous Galerkin method, EFG shape function, local boundary integration, erosion

For citation: Godunov A.I., Kukanov S.A., Baklanov D.V. Physical features of the processes of deformation and destruction of the material by the finite node method. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):19–29. (In Russ.). doi: 10.21685/2307-4205-2024-3-3

Расчет на прочность композитной гильзы (рис. 1) проводился в квазистатическом приближении для момента времени, при котором реализуется максимальное давление пороховых газов $P = 295$ МПа. Допустимость квазистатического расчета обосновывается достаточно большими временами протекающих процессов (максимальные деформации развиваются в изделии за время $\sim 74 \cdot 10^{-5}$ с) по сравнению со временем распространения возмущений в материале изделия (прохождение упругих волн по длине стандартной конструкции занимает время, порядка единиц микросекунд). Кроме этого, предполагается возможность использования свойств полиамидного материала, полученных в статических испытаниях. Это предположение сделано как для металлов, так и для композитов, для которых при реализующихся скоростях деформаций свойства (предел текучести/прочности, модули упругости) меняются незначительно.

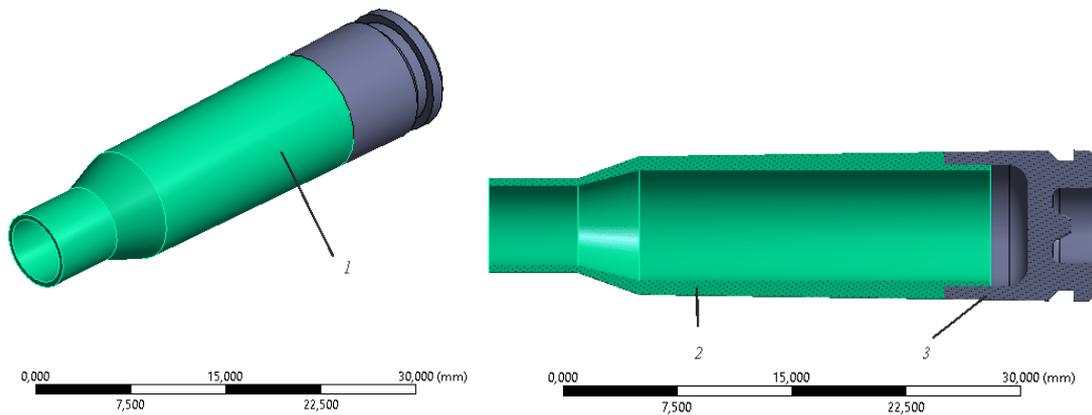


Рис. 1. Композитная гильза:
1 – общий вид гильзы; 2 – тело гильзы из полиамида PA612; 3 – донная часть из стали

Таким образом, материал должен работать в упругой зоне. При проектировании композитного тела гильзы, изготовленного из полиамида, критериями являются пределы прочности материала, а при рассмотрении металлической донной части критерием является предел текучести. Соединение конечных элементов и узлов представлено на рис. 2.

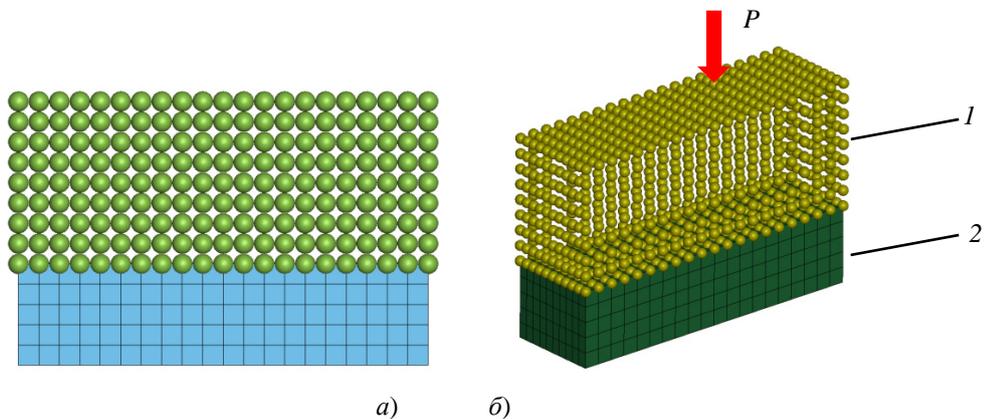


Рис. 2. Соединение конечных элементов и узлов:
а – вид оболочки; б – вид объема: 1 – тело гильзы из полиамида;
2 – донная часть гильзы из стали; P – давление пороховых газов 295 МПа

Разрывной метод Галеркина [2] сочетает в себе свойства конечно-объемных и конечно-элементных методов, при этом под сеточной функцией понимается функция, на каждом интервале (x_j, x_{j+1}) которой имеется многочлен порядка $p \in \mathbb{N} \cup \{0\}$. На каждом интервале можно выбрать любой базис в пространстве многочленов порядка p и продолжить эти функции нулем вне интервала. Объединение таких функций образует базис в пространстве сеточных функций [1].

Решение нахождения параметров разрушения материала тела композитной гильзы при таком подходе можно описать в виде линейных комбинаций базисных функций с зависящими от времени коэффициентами:

$$\vartheta(t, x) = \sum_{k=0}^{N_{x-1}} \sum_{m=0}^P \vartheta_{k,m}(t) \phi_k^m(x). \quad (1)$$

Проинтегрируем уравнение (1) по x . Поскольку базисная функция равна нулю вне одного отрезка, получим

$$\int_{x_j}^{x_{j+1}} \phi_j^m(x) \frac{du}{dt}(t, x) dx + \mu \int_{x_j}^{x_{j+1}} \phi_j^m(x) \frac{du}{dt}(t, x) dx = 0. \quad (2)$$

Проинтегрируем второе слагаемое (2) по частям:

$$\frac{d}{dt} \int_{x_j}^{x_{j+1}} \phi_j^m(x) u(t, x) dx + \mu \phi_j^m(x_{j+1}) u(t, x_{j+1}) - \mu \phi_j^m(x_j) u(t, x_j) - \mu \int_{x_j}^{x_{j+1}} \frac{d\phi_j^m(x)}{dx} u(t, x) dx = 0. \quad (3)$$

Обратим внимание, что при давлении $P = 0$ каждая из базисных функций является нулевой константой внутри своей ячейки и после деления на эту константу выражение (3) примет вид

$$\frac{d}{dt} \int_{x_j}^{x_{j+1}} u(t, x) dx + \mu u(t, x_{j+1}) - \mu u(t, x_j) = 0. \quad (4)$$

Далее приступим к замене $\mu u(t, x_{j+1})$ на некоторые функции, называемые численными потоками. После преобразования (3) уравнение примет вид

$$\int_{x_j}^{x_{j+1}} \phi_j^m(x) \frac{d\upsilon}{dt}(t, x) dx + \phi_j^m(x_{j+1}) F_{j+1}(t) - \phi_j^m(x_j) F_j(t) - \mu \int_{x_j}^{x_{j+1}} \frac{d\phi_j^m(x)}{dx} \upsilon(t, x) dx = 0, \quad (5)$$

где

$$F_j(t) = \vartheta(t, x_j - 0) \frac{1 + \sigma \text{sign} \mu}{2} + \vartheta(t, x_j + 0) \frac{1 - \sigma \text{sign} \mu}{2}, \quad (6)$$

а $0 \leq \sigma \leq 1$. Символами $\vartheta(t, x_j \pm 0)$ обозначены предельные значения слева и справа от точки x_j . Часто полагают $\sigma \equiv 1$, но оставим возможность варьирования этого параметра [3].

Уравнение для описания поля деформации и перемещений:

$$\bar{U}^h(x, t) = \sum_{K=1}^{NP} W_K(x) U^h(x_K, t) \equiv \sum_{K=1}^{NP} \phi_K(x) d_K(t). \quad (7)$$

Связи частиц до момента разрушения представлены на рис. 3, их можно описать как

$$\begin{aligned} \bar{U}^h(x_1, t) = & \phi_1(x_1) d_1(t) + \phi_2(x_1) d_2(t) + \phi_4(x_1) d_4(t) + \phi_5(x_1) d_5(t) + \\ & + \phi_6(x_1) d_6(t) + \phi_7(x_1) d_7(t) + \phi_8(x_1) d_8(t) + \phi_{10}(x_1) d_{10}(t). \end{aligned} \quad (8)$$

$$\bar{U}^h(x_2, t) = \phi_1(x_2) d_1(t) + \phi_2(x_2) d_2(t) + \phi_3(x_2) d_3(t) + \phi_8(x_2) d_8(t) + \phi_9(x_2) d_9(t) + \phi_{10}(x_2) d_{10}(t). \quad (9)$$

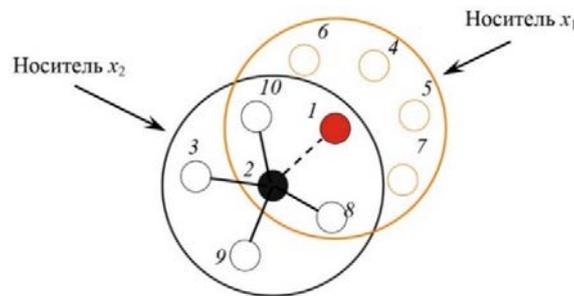


Рис. 3. Схема взаимодействия двух частиц с носителями связей:

1 – частица с математическим носителем $\phi_1(x_2) d_1(t)$;

2 – частица с математическим носителем $\phi_2(x_1) d_2(t)$

При выполнении условия $\phi_1(x_2)d_2(t)=0$ и $\phi_2(x_1)d_2(t)=0$ происходит разрушение связей и, соответственно, разрушение материала. В описанном методе разрушение по пластической деформации связей позволяет описать аппроксимацию ядром сглаживания, при котором приближенное представление функций и их производных осуществляется через ядро сглаживания и ее производные [4]. При этом методе нет потери энергии и массы материала (в отличие от метода конечных элементов), который подвергается разрушению. Высокие порядки аппроксимации способны лучше описывать высокие градиенты.

Для описания сглаживания (рис. 4) используется кубическая В-сплайновая функция ядра в качестве весовой функции:

$$W_a(x-x_1) = \left\{ \begin{array}{l} \frac{2}{3} - 4\left(\frac{\|x-x_1\|}{a}\right)^2 + 4\left(\frac{\|x-x_1\|}{a}\right)^3 \quad 0 \leq \frac{\|x-x_1\|}{a} \leq \frac{1}{2} \\ \frac{4}{3} - 4\left(\frac{\|x-x_1\|}{a}\right) + 4\left(\frac{\|x-x_1\|}{a}\right)^2 - \frac{4}{3}\left(\frac{\|x-x_1\|}{a}\right)^3 \quad \frac{1}{2} < \frac{\|x-x_1\|}{a} \leq 1 \end{array} \right. \quad (10)$$

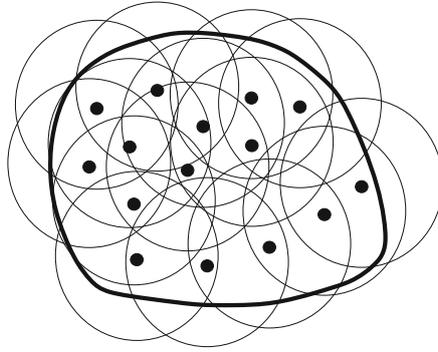


Рис. 4. Графическое представление без сеточной дискретизации

Метод скользящего наименьшего квадрата заключается в минимизации взвешенной L_2 -нормы:

$$J = \sum_{i=1}^{NP} W_a(x-x_1) \left[\sum_{i=1}^n H_i(x) b_i(x) - u(x_1) \right]^2, \quad (11)$$

где NP – число узлов в поддержке X , для которых $W_a(x-x_1) \neq 0$.

В неявном виде безсеточный метод использует комбинацию гладких базисных функций (примитивов) для нахождения скалярной функции, такой, что все узлы данных близки к изоконтуре этой скалярной функции в глобальном смысле [1]. В реальности построение поверхности оболочки с помощью трехмерного безсеточного метода оказывается неадекватным. Это связано с тем, что топология реальной поверхности может быть очень сложной в трех измерениях. Без информации об упорядоченности или связности узлов реконструированная поверхность не сможет отобразить пересечения оболочек, внешних границ и углов формы.

Уравнение (11) можно записать в виде

$$J = (Hb - u)^T W_a(x) (Hb - u), \quad (12)$$

где

$$U^T = (u_1, u_2, \dots, u_{NP}), \quad (13)$$

$$H = \begin{bmatrix} \{H(x_1)\}^T \\ \dots \\ \{H(x_{NP})\}^T \end{bmatrix}, \quad (14)$$

$$\{H(x_1)\}^T = \{H_1(x_1), \dots, W_a(x - x_{NP})\}, \quad (15)$$

$$W_a = \text{diag}[w_a(x - x_1), \dots, w_a(x - x_{NP})]. \quad (16)$$

Чтобы найти коэффициент b , получим экстремум J по формуле

$$\frac{dJ}{db} = M^{[n]}(x)b(x) - B(x)u = 0, \quad (17)$$

где $M^{[n]}(x)$ называется матрицей моментов $w_a(x - x_1)$ и имеет вид

$$M^{[n]}(x) = H^T W_a(x) H. \quad (18)$$

Итак, у нас есть

$$B(x) = M^{[n]}(x)B(x)u. \quad (19)$$

Чтобы $M^{[n]}(x)$ была инвертируемой, поддержка $w_a(x - x_1)$ должна быть больше минимального размера, который связан с порядком базисных функций. Используя решение уравнений (6), (17), аппроксимация получается следующим образом:

$$u^h(x) = \sum_{i=1}^{NP} \Psi_i(x) u_i, \quad (20)$$

где функции формы $\Psi_i(x)$ задаются следующим образом:

$$\Psi_i(x) = H^T(x) M^{[n]}(x) B(x) \quad (21)$$

и $\Psi_i(x)$ являются полными n -го порядка, т.е.

$$\sum_{i=1}^{NP} \Psi_i(x) x_{11}^p x_{21}^q = x_1^p x_2^q \text{ для } p + q = 0, \dots, n. \quad (22)$$

Сходимость метода Галеркина для дифференциального уравнения определяется аппроксимацией неизвестных и численным интегрированием в слабой форме. Функции формы Element Free Galerkin Method (далее EFG) с линейной согласованностью могут быть получены из линейной аппроксимации поля перемещения с линейными базисными функциями. Использование линейно согласованных бессеточных функций формы в аппроксимации Галеркина не гарантирует линейной точности решения [2]. Для линейной точности решения в приближении Галеркина необходимы два ограничения на интегрирование:

$$\sum_{L=1}^{NIT} \nabla \Psi_i(x_L) A_L \text{ для } \{I : \text{supp}(\Psi_i) \cap G = 0\}, \quad (23)$$

$$\sum_{L=1}^{NIT} \nabla \Psi_i(x_L) A_L = \sum_{L=1}^{NITh} n \Psi_i(\tilde{x}_L) S_L \text{ для } \{I : \text{supp}(\Psi_i) \cap G^h \neq 0\}, \quad (24)$$

где G^h – естественная граница; G – общая граница; n – нормаль к поверхности G^h ; x_L и A_L – пространственная координата и вес точки интегрирования области соответственно; \tilde{x}_L и S_L – пространственная координата и вес точки интегрирования естественной границы, NIT – количество точек интегрирования для доменного интегрирования; $NITh$ – количество точек интегрирования для интегрирования естественной границы.

Метод сглаживания деформаций в качестве регуляризации нестабильности материала при локализации деформаций был расширен в методе узлового интегрирования для удовлетворения ограничений на интегрирование. Используя аналогичную концепцию для доменного интегрирования со сглаживанием деформации в представительной области точки Гаусса, получим

$$\tilde{\nabla} u_i^h(x_L) = \frac{1}{A_L} \int_{\Omega_L} \nabla u_i^h(x_L) d\Omega, \quad A_L = \int_{\Omega_L} d\Omega, \quad (25)$$

где Ω_L – представительный домен в каждой точке Гуасса, а $\tilde{\nabla}$ – сглаженный градиент:

$$\tilde{\nabla} u_i^h(x_L) = \frac{1}{A_L} \int_{G_L} n u_i^h(x_L) dG, \quad (26)$$

где G_L – граница представительной области точки Гуасса L :

$$\tilde{\nabla} u_i^h(x_L) = \sum_i \frac{1}{A_L} \int_{G_L} \Psi_i(x) n dG \cdot d_{ii} \equiv \sum_i \tilde{\nabla} \Psi_i(x_L) d_{ii}. \quad (27)$$

Можно сказать, что сглаженный градиент функции формы EFG· удовлетворяет целочисленным ограничениям в уравнениях (23) и (24) независимо от используемого численного интегрирования.

Разработаны два проекционных метода для создания средней поверхности оболочки с использованием приближения скользящих наименьших квадратов. Для обработки произвольно больших вращений с умеренными деформационными реакциями оболочечных структур используется метод вращения с обновленным лагранжианом [6]. Метод локального граничного интегрирования в сочетании с селективным методом сокращенного интегрирования вводится для обеспечения линейной точности и снятия блокировки сдвига. Это позволяет описать физический процесс разрушения материала композитной гильзы.

Восстановление поверхности по неорганизованным узлам является очень сложной задачей в трех измерениях. При разработке бессеточных оболочек предполагается, что поверхность оболочки описывается сеткой конечных элементов. Этого можно легко добиться, преобразовав часть конечных элементов оболочки в зону без сетки (рис. 5). Благодаря связности узлов, обеспечиваемой сеткой конечных элементов, поверхность оболочки может быть восстановлена с помощью бессеточной интерполяции из узловых положений.

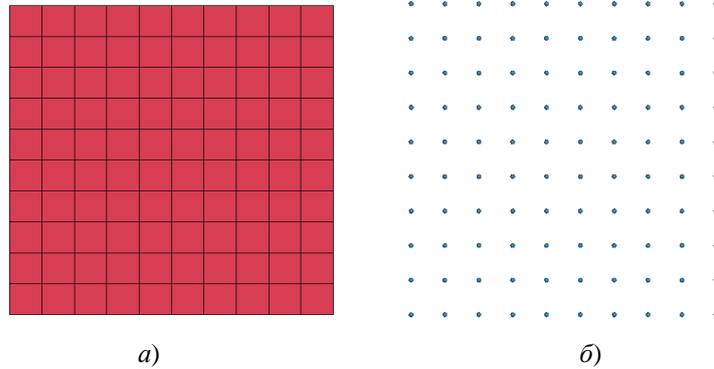


Рис. 5. Преобразование конечных элементов элементарной оболочки:
 а – массив конечных элементов; б – массив конечных узлов

Немаловажную роль в математическом моделировании соединенных конечных элементов и конечных узлов (рис. б) необходимо отвести их совместному вращению в пространстве. Геометрия и кинематические поля оболочки могут быть описаны с помощью опорной поверхности в направлении движения. Модифицированное предположение Мидлина – Рейсснера требует, чтобы движение и смещение оболочки было линейным в направлении волокон материала [5]. Предположим, что опорной поверхностью является средняя поверхность оболочки, глобальные координаты и смещения в произвольной точке внутри тела оболочки можно задать как

$$x = \bar{x} + \xi \frac{h}{2} V_3; \quad (28)$$

$$u = \bar{u} + \xi \frac{h}{2} U, \quad (29)$$

где x и u – вектор положения и смещение опорной поверхности; V_3 – направление волокна материала; U – смещение, возникающее в результате вращения материала; h – длина волокна.

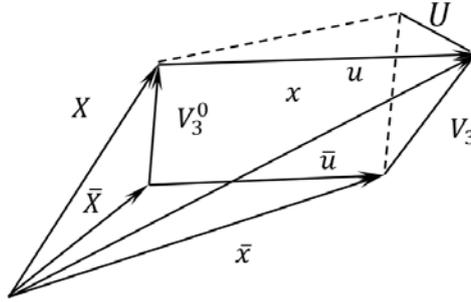


Рис. 6. Векторное представление деформации оболочки

В приближении без сетки движение и деформация задаются как

$$x(\xi, \eta, \zeta) = \bar{x}(\xi, \eta) + V(\xi, \eta, \zeta) \approx \sum_{i=1}^{NP} \tilde{\Psi}_1(\xi, \eta) x_i + \sum_{i=1}^{NP} \tilde{\Psi}_1(\xi, \eta) \frac{\zeta h_i}{2} V_{3i}, \quad (30)$$

$$u(\xi, \eta, \zeta) = \bar{u}(\xi, \eta) + U(\xi, \eta, \zeta) \approx \sum_{i=1}^{NP} \tilde{\Psi}_1(\xi, \eta) u_i + \sum_{i=1}^{NP} \tilde{\Psi}_1(\xi, \eta) \frac{\zeta h_i}{2} [-V_{2i} V_{1i}] \begin{Bmatrix} \alpha_i \\ \beta_i \end{Bmatrix}, \quad (31)$$

где x и u – глобальные координаты и смещения в свободной от сетки узлов области i ; V_{3i} – единичный вектор направления волокна; V_{1i} и V_{2i} – базовые векторы узловой системы координат в узле i ; α_i и β_i – повороты вектора направления V_{3i} вокруг осей V_{2i} и V_{1i} , а $\tilde{\Psi}_1$ – двумерные функции формы.

Локальная система координат совместного вращения $\hat{x}, \hat{y}, \hat{z}$ определяется в каждой точке интегрирования на опорной поверхности оболочки с касательными \hat{x} и \hat{y} [3]. На основании вышеизложенного базис векторов примет вид

$$\hat{e}_1 = \frac{x_{,\xi}}{\|x_{,\xi}\|}, \hat{e}_3 = \frac{x_{,\xi} \times x_{,\eta}}{\|x_{,\xi} \times x_{,\eta}\|}, \hat{e}_2 = \hat{e}_3 \times \hat{e}_1. \quad (32)$$

Чтобы описать повороты волокон узла без сетки в оболочке, введем узловую систему координат (рис. 7), тремя базовыми векторами которой являются V_1, V_2, V_3 , где направление волокна в узле V_1, V_2 определено как

$$V_1 = \frac{\hat{x} \times V_3}{|\hat{x} \times V_3|}, V_2 = V_3 \times V_1. \quad (33)$$

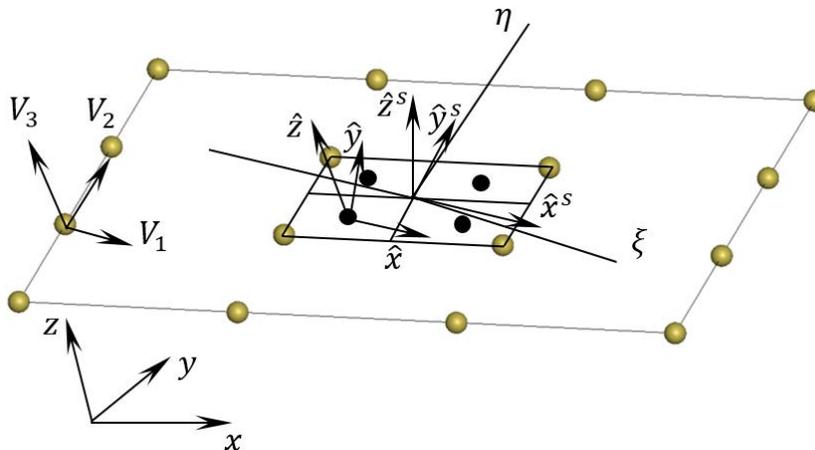


Рис. 7. Локальная и узловая системы координат с одновременным вращением

Затем вращение направляющего материала описываем из глобальных вращений:

$$\begin{Bmatrix} \alpha \\ \beta \end{Bmatrix} = \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \Delta\Theta, \quad \Delta\Theta = [\Delta\Theta_1 \Delta\Theta_2 \Delta\Theta_3]^T. \quad (34)$$

В локальной системе координат с одновременным вращением движения и смещение аппроксимируются функциями формы без сетки вида

$$\hat{x}_i = \sum_{I=1}^{NP} \tilde{\Psi}_I \hat{x}_{iI} + \zeta \sum_{I=1}^{NP} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{3iI}, \quad (35)$$

$$\hat{u}_i = \sum_{I=1}^{NP} \tilde{\Psi}_I \hat{u}_{iI} + \zeta \sum_{I=1}^{NP} \tilde{\Psi}_I \frac{h_I}{2} \left[-\hat{V}_{2iI} \hat{V}_{1iI} \right] \begin{Bmatrix} \alpha_i \\ \beta_i \end{Bmatrix}. \quad (36)$$

Лагранжевая деформация сглаживается как

$$\tilde{\varepsilon}^m = \sum_I \tilde{B}_I^m \hat{d}_I, \quad \tilde{\varepsilon}^b = \zeta \sum_I \tilde{B}_I^b \hat{d}_I, \quad \tilde{\varepsilon}^s = \sum_I \tilde{B}_I^s \hat{d}_I, \quad (37)$$

где сглаженные операторы деформации вычисляются путем усреднения согласованных операторов деформации по области вокруг оцениваемой точки:

$$\tilde{B}_I^m(x_I) = \frac{1}{A_I} \int_{\Omega_I} \tilde{B}_I^m dA, \quad \tilde{B}_I^b(x_I) = \frac{1}{A_I} \int_{\Omega_I} \tilde{B}_I^b dA, \quad \tilde{B}_I^s(x_L) = \frac{1}{A_L} \int_{\Omega_L} \tilde{B}_I^s dA, \quad (38)$$

$$\tilde{B}_I^m = \begin{bmatrix} \tilde{\Psi}_{I,x} & 0 & 0 & -J_{13}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{2xI} & -J_{13}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{1xI} \\ 0 & \tilde{\Psi}_{I,y} & 0 & -J_{23}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{2yI} & -J_{23}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{1yI} \\ \tilde{\Psi}_{I,x} & \tilde{\Psi}_{I,y} & 0 & -J_{23}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{2xI} - J_{13}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{2yI} & -J_{23}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{1xI} + J_{13}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{1yI} \end{bmatrix},$$

$$\tilde{B}_I^b = \begin{bmatrix} 0 & 0 & 0 & -\tilde{\Psi}_{I,x} \frac{h_I}{2} \hat{V}_{2xI} & \tilde{\Psi}_{I,x} \frac{h_I}{2} \hat{V}_{1xI} \\ 0 & 0 & 0 & -\tilde{\Psi}_{I,y} \frac{h_I}{2} \hat{V}_{2yI} & \tilde{\Psi}_{I,y} \frac{h_I}{2} \hat{V}_{1yI} \\ 0 & 0 & 0 & -\tilde{\Psi}_{I,y} \frac{h_I}{2} \hat{V}_{2xI} - \tilde{\Psi}_{I,x} \frac{h_I}{2} \hat{V}_{2yI} & \tilde{\Psi}_{I,y} \frac{h_I}{2} \hat{V}_{1xI} + \tilde{\Psi}_{I,x} \frac{h_I}{2} \hat{V}_{1yI} \end{bmatrix},$$

$$\tilde{B}_I^s = \begin{bmatrix} 0 & 0 & \tilde{\Psi}_{I,y} & -J_{33}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{2yI} - J_{23}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{2zI} & J_{33}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{1yI} + J_{23}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{1zI} \\ 0 & 0 & \tilde{\Psi}_{I,x} & -J_{33}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{2xI} - J_{13}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{2zI} & J_{33}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{1xI} + J_{13}^{-1} \tilde{\Psi}_I \frac{h_I}{2} \hat{V}_{1zI} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

где J^{-1} является обратной матрицей Якоби в точке интегрирования. Локальная степень свободы выражается как

$$\hat{d}_I = [\hat{u}_{xI} \quad \hat{u}_{yI} \quad \hat{u}_{zI} \quad \alpha_I \quad \beta_I]^T. \quad (39)$$

Внутренний узловый вектор силы равен

$$\hat{F}_I^{\text{int}} = \int_{\Omega} \tilde{B}_I^{mT} \hat{\sigma} d\Omega + \int_{\Omega} \zeta \tilde{B}_I^{bT} \hat{\sigma} d\Omega + \int_{\Omega} \tilde{B}_I^{sT} \hat{\sigma} d\Omega. \quad (40)$$

Использование обновленной формулировки Лагранжа подразумевает, что система координат отсчета определяется системой совместного вращения в конфигурации в момент времени t [7]. Следовательно, локальные узловые векторы силы и перемещения, относящиеся к этой системе координат, должны быть преобразованы в глобальную систему координат при построении. Математическая модель процесса разрушения и деформации материала 5,45-мм композитной гильзы представлена на рис. 8.

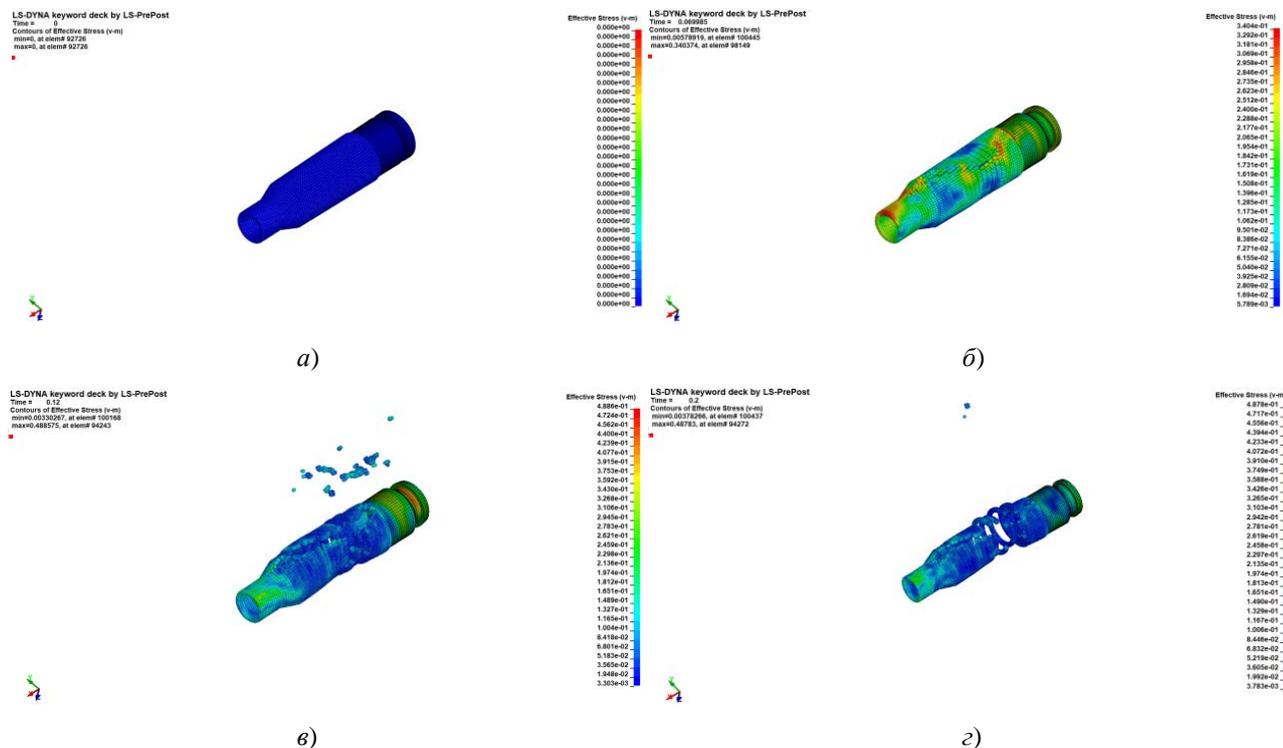


Рис. 8. Математическая модель процесса разрушения и деформации материала:

a – концентрация силового воздействия на материал; *б* – распределение стресса по фон Мизесу в материале; *в* – физическое образование потери связи массива узлов разрушаемого материала; *г* – разрушение материала

Заключение

Таким образом, был определен наиболее рациональный математический метод, который решался в лагранжевой постановке задач при использовании явной схемы интегрирования разрывным методом конечных узлов Галеркина. Также решение задачи динамического нагружения композитной гильзы методом конечных узлов позволит исключить такое понятие, как «эрозия» элемента, сохранить массу и энергию на всем цикле интегрирования. Высокие порядки аппроксимации позволили лучше описать высокие градиенты, однако стоит отметить как недостаток большую вычислительную сложность. Кроме того, был выбран восьмиузловой конечный элемент, позволяющий с достаточной точностью описать геометрические размеры элементов системы.

Список литературы

1. Каплун А. Б., Морозов Е. М., Олферьева М. А. ANSYS в руках инженера : практ. руководство. М. : Ленанд/URSS, 2017. 272 с.
2. Галеркин Б. Г. Собрание сочинений. М. : Изд.-во АН СССР, 1952. Т. 1. 391 с.
3. Johnson G. R., Cook W. H. Fracture characteristics of three metals subjected to various strains, strain rates, temperatures and pressures // Eng Fract Mech. 1985. Vol. 21, № 1. P. 31–48.
4. Плитов И. С., Полилов А. Н., Татусь Н. А. Компьютерное моделирование рациональной структуры криволинейного армирования профилированных композитных элементов // Проблемы машиностроения и автоматизации. 2013. № 4. С. 73–79.
5. Chang F. K., Chang K. Y. A progressive damage model for laminated composites containing stress concentrations // J Compos Mater. 1987. Vol. 21, № 9. P. 834–855.

6. Wang Y. X., Long R. R., Zhang Q. M. [et al.]. Numerical simulation of deformation characteristics of two-layer metal-plastic cylindrical shells during internal explosive immersion // *Acta Armamentarii*. 2020. Vol. 41, № S2. С. 189–199.
7. Моссаковский П. А., Костырева Л. А. О новом способе экспериментального исследования материалов на динамический сдвиг при высокоскоростном деформировании // *Проблемы прочности и пластичности*. 2018. Т. 80, № 1. С. 127–135.
8. Седов Л. И. *Механика сплошной среды*. М. : Наука, 1970. Т. 2. 492 с.

References

1. Kaplun A.B., Morozov E.M., Olfer'eva M.A. *ANSYS v rukakh inzhenera: prakt. rukovodstvo = ANSYS in the hands of an engineer : a practical guide*. Moscow: Lenand/ URSS, 2017:272. (In Russ.)
2. Galerkin B.G. *Sobranie sochineniy = Collected works*. Moscow: Izd.-vo AN SSSR, 1952;1:391. (In Russ.)
3. Johnson G.R., Cook W.H. Fracture characteristics of three metals subjected to various strains, strain rates, temperatures and pressures. *Eng Fract Mech*. 1985;21(1):31–48.
4. Plitov I.S., Polilov A.N., Tatus' N.A. Computer modeling of the rational structure of curved reinforcement of profiled composite elements. *Problemy mashinostroeniya i avtomatizatsii = Problems of mechanical engineering and automation*. 2013;(4):73–79. (In Russ.)
5. Chang F.K., Chang K.Y. A progressive damage model for laminated composites containing stress concentrations. *J Compos Mater*. 1987;21(9):834–855.
6. Wang Y.X., Long R.R., Zhang Q.M. et al. Numerical simulation of deformation characteristics of two-layer metal-plastic cylindrical shells during internal explosive immersion. *Acta Armamentarii*. 2020;41(S2):189–199.
7. Mossakovskiy P.A., Kostyreva L.A. On a new method of experimental investigation of materials for dynamic shear during high-speed deformation. *Problemy prochnosti i plastichnosti = Problems of strength and plasticity*. 2018;80(1):127–135. (In Russ.)
8. Sedov L.I. *Mekhanika sploshnoy sredy = Mechanics of a continuous medium*. Moscow: Nauka, 1970;2:492. (In Russ.)

Информация об авторах / Information about the authors

Анатолий Иванович Годунов

доктор технических наук, профессор,
заслуженный деятель науки РФ,
профессор кафедры автоматизации и телемеханики,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: avitelpgu@mail.ru

Сергей Анатольевич Куканов

кандидат технических наук,
доцент кафедры средств ближнего боя,
Филиал Военной академии
материально-технического
обеспечения имени генерала армии
А. В. Хрулева в г. Пензе
(Россия, г. Пенза, Военный городок)
E-mail: Kuk_@mail.ru

Дмитрий Владимирович Бакланов

адъюнкт,
Филиал Военной академии
материально-технического
обеспечения имени генерала армии
А. В. Хрулева в г. Пензе
(Россия, г. Пенза, Военный городок)
E-mail: baklanov90@yandex.ru

Anatoly I. Godunov

Doctor of technical sciences, professor,
honored scientist of the Russian Federation,
professor of the sub-department
of automation and telemechanics,
Penza State University
(40 Krasnaya street, Penza, Russia)

Sergei A. Kukanov

Candidate of technical sciences, associate professor
of the sub-department of melee weapons,
Branch of the Military Academy
of Logistics named after Army General
A.V. Khrulev in Penza
(Military town, Penza, Russia)

Dmitry V. Baklanov

Adjunct,
Branch of the Military Academy
of Logistics named after Army General
A.V. Khrulev in Penza
(Military town, Penza, Russia)

Авторы заявляют об отсутствии конфликта интересов /

The authors declare no conflicts of interests.

Поступила в редакцию/Received 08.06.2024

Поступила после рецензирования/Revised 20.07.2024

Принята к публикации/Accepted 25.08.2024

МОДЕЛИРОВАНИЕ УПРАВЛЯЕМОСТИ ХАРАКТЕРИСТИК РЕШЕТОК ПЛАЗМОННЫХ ГРАФЕНОВЫХ НАНОАНТЕНН В СРЕДНЕМ ИНФРАКРАСНОМ ДИАПАЗОНЕ

Г. С. Макеева

Пензенский государственный университет, Пенза, Россия
radiotech@pnzgu.ru

Аннотация. *Актуальность и цели.* Внедрение наноантенн для оптической беспроводной связи в инфракрасном (ИК) и видимом диапазонах обеспечивает более высокую скорость передачи данных при уменьшении размера антенны. Возможность перестройки характеристик графена химическим легированием или напряжением смещения актуальна при разработке реконфигурируемых наноантенн. Целью работы является исследование характеристик (S -параметров, диаграммы направленности (ДН)) решеток плазмонных графеновых наноантенн (ПГНА), их управляемости и возможности сканирования по частоте при изменении химического потенциала графена (приложением внешнего электрического поля) в среднем ИК-диапазоне. *Материалы и методы.* Применение графена, обладающего хорошей электропроводностью, управляемой проводимостью и плазмонными свойствами в терагерцовом, дальнем и среднем ИК-диапазонах, является одной из наиболее многообещающих альтернатив благородным металлам (Au и Ag) как плазмонным материалам только на оптических частотах. Моделирование характеристик решеток ПГНА проведено с помощью программы электродинамического моделирования CST Microwave Studio 2019, что позволяет решить научные проблемы, связанные с проектированием графеновых антенн в ИК-диапазоне волн. *Результаты.* Получены результаты моделирования управляемости характеристик (S -параметров, ДН) единичного элемента антенной решетки (ПГНА прямоугольной геометрии) и решеток ПГНА на резонансных частотах основной моды поверхностных плазмон-поляритонов, и показана возможность сканирования по частоте при изменении химического потенциала графена (приложением внешнего электрического поля) в среднем ИК-диапазоне. *Выводы.* Из результатов моделирования характеристик решеток ПГНА следует, что с увеличением химического потенциала графена (в интервале значений 0,3–1 эВ) происходит перестройка рабочих частот (сканирование по частоте) в среднем ИК-диапазоне, возрастает коэффициент усиления решетки ПГНА, происходит сужение главного лепестка ДН, уменьшение уровня боковых лепестков; при увеличении числа единичных элементов ($N = 256$) уровни боковых лепестков ДН уменьшаются и наблюдается лучшая управляемость главного лепестка ДН.

Ключевые слова: средний ИК-диапазон, решетки плазмонных графеновых наноантенн, химический потенциал, диаграмма направленности, сканирование по частоте

Для цитирования: Макеева Г. С. Моделирование управляемости характеристик решеток плазмонных графеновых наноантенн в среднем инфракрасном диапазоне // Надежность и качество сложных систем. 2024. № 3. С. 30–40. doi: 10.21685/2307-4205-2024-3-4

MODELING THE CONTROLLABILITY OF CHARACTERISTICS OF PLASMON GRAPHENE NANOANTENNA GRATINGS IN THE MID-IR RANGE

G.S. Makeeva

Penza State University, Penza, Russia
radiotech@pnzgu.ru

Abstract. *Background.* The introduction of nanoantennas for optical wireless communications in the infrared (IR) and visible bands enables higher data rates while reducing antenna size. The ability to tune the characteristics of graphene by chemical doping or bias voltage is relevant in the development of reconfigurable nanoantennas. The goal of the work is to study the characteristics (S -parameters, radiation patterns (RP)) of plasmonic graphene nanoantenna (PGNA) arrays, their controllability and the possibility of frequency scanning when changing the chemical potential of graphene (by applying an external electric field) in the mid-IR range. *Materials and methods.* The use of graphene, which has good electrical conductivity, controlled conductivity and plasmonic properties in the terahertz (THz), far-IR and mid-IR ranges, is one of the most promising alternatives to noble metals (Au and Ag) as plasmonic materials at optical frequencies only. Modeling of the characteristics of PGNA arrays was carried out using the electrodynamic modeling program CST Microwave Studio 2019, which makes it possible to solve scientific problems associated with

the design of graphene antennas in the IR wavelength range. *Results.* The results of modeling the controllability of the characteristics (S-parameters, DP) of a element of the antenna array (PGNA of rectangular geometry) and PGNA arrays at the resonant frequencies of the fundamental mode of surface plasmon polaritons (SPP) are obtained and the possibility of scanning in frequency when changing the chemical potential of graphene (by applying an external electric fields) in the mid-IR range. *Conclusions.* From the results of modeling the characteristics of PGNA arrays, it follows that with an increase in the chemical potential of graphene (in the range of values 0,3–1 eV), the operating frequencies are adjusted (frequency scanning) in the mid-IR range, the gain of the PGNA array increases, the RP main lobe narrows, and the RP side lobe level decrease; with an increase in the number of single elements ($N = 256$), the levels of the RP side lobes decrease and better controllability of the RP main lobe is observed.

Keywords: mid-IR range, plasmonic graphene nanoantenna arrays, chemical potential, radiation pattern, frequency scanning

For citation: Makeeva G.S. Modeling the controllability of characteristics of plasmon graphene nanoantenna gratings in the mid-ir range. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):30–40. (In Russ.). doi: 10.21685/2307-4205-2024-3-4

Введение

За последние несколько лет быстро растет внедрение наноантенн для оптической беспроводной связи в инфракрасном (ИК) и видимом диапазонах, что обеспечивает более высокую скорость передачи данных при уменьшении размера антенны [1]. Патч-наноантенна впервые была разработана для ближнего ИК-диапазона [2]. Эта наноантенна возбуждалась гибридным плазмонным волноводом на основе благородных металлов (Au и Ag), обладающих плазмонными свойствами только на оптических частотах, а в среднем и дальнем ИК-диапазонах являющихся проводниками, а не плазмонными материалами [3].

Впоследствии графен продемонстрировал плазмонные свойства в терагерцовом (ТГц), дальнем и среднем ИК-диапазонах. Главным преимуществом графена являются перестраиваемые характеристики [4, 5]. Его можно перестраивать либо внутренними, либо внешними средствами, т.е. химическим легированием или напряжением смещения. Эта возможность перестройки используется при разработке реконфигурируемых наноантенн [5].

Графен, обладающий хорошей электропроводностью, управляемой проводимостью, и плазмонными свойствами, является одной из наиболее многообещающих альтернатив [6, 7]. Если графен интегрирован в антенны, то графеновые антенны показывают лучшие характеристики излучения, чем их типичные металлические аналоги. Это обусловлено хорошей проводимостью графена и тем, что его химический потенциал, а значит, и поверхностную проводимость можно изменять, приложив напряжение смещения, что, в свою очередь, изменяет значение поверхностного импеданса графена [8].

Чтобы функционировать на наномасштабах, традиционные антенны должны работать на очень высоких, т.е. оптических частотах, что делает их выбор непрактичным. Однако отличительная черта графена – более медленная подвижность электронов – позволяет работать в ТГц, дальнем и среднем ИК-диапазонах, что делает графен жизнеспособным кандидатом на роль наноразмерных антенн [9]. Как следствие, материалы на основе графена предлагают потенциал в разработке реконфигурируемых антенн ТГц [4–6, 8, 9] и среднего ИК-диапазонов [1, 10, 11].

Усиленное электромагнитное удержание оптических полей и последующая адекватная концентрация оптического излучения с помощью оптических наноантенн имеют невероятное значение в различных фотонных приложениях, включая ИК-визуализацию, зондирование, спектроскопию и фотодетектирование [10, 12].

Концепция ограничения и усиления оптического поля чрезвычайно многообещающая в ИК и ТГц диапазонах частот, что открывает возможности для таких важных приложений, как визуализация [13], зондирование [14] и спектроскопия [15]. Однако слабое ограничение оптического поля металлическими плазмонными структурами, обусловленное малым проникновением электромагнитных полей в металл на низких частотах, сильно снижает эффективность [16].

Омические потери графена уменьшаются при воздействии сильного химического легирования, кроме того, графеновые поверхностные плазмон-поляритоны (ППП) обладают более высоким удержанием, обеспечивая более сильные взаимодействия излучения с материей [17, 18]. Необычайные электрические свойства графена и технологии изготовления высококристаллических листов графена увеличивают время жизни ППП сотен оптических циклов, что позволяет решить проблемы, связанные с плазмонными антеннами из благородных металлов [19].

В этой связи актуальной является задача исследования методами автоматизированного моделирования характеристик (S -параметров, диаграммы направленности) решеток плазмонных графеновых наноплазмонных антенн (ПГНА), их управляемости и возможности сканирования по частоте при изменении химического потенциала графена (приложением внешнего электрического поля) в среднем ИК-диапазоне.

Результаты моделирования управляемости характеристик (элемента матрицы рассеяния S_{11} , ДН) плазмонной графеновой наноплазмонной антенны в ИК-диапазоне

С помощью пакета прикладных программ CST Studio Suite 2023 [20] на первом этапе проведено моделирование единичного элемента графеновой антенной решетки среднего ИК-диапазона – плазмонной графеновой наноплазмонной антенны прямоугольной геометрии. На рис. 1 представлена модель ПГНА, разработанная в CST Microwave Studio 2023 [20, 21], которая состоит из графенового элемента прямоугольной геометрии (длиной l , шириной w), расположенного на металлизированной диэлектрической подложке (диоксид кремния SiO_2 , диэлектрическая проницаемость $\epsilon = 2,2$) с геометрическими размерами a, b, h . Эта модель с возбуждением дискретным портом использована для решения электродинамической задачи с помощью пакета прикладных программ CST Microwave Studio 2023 методом конечного интегрирования во временной области [21].

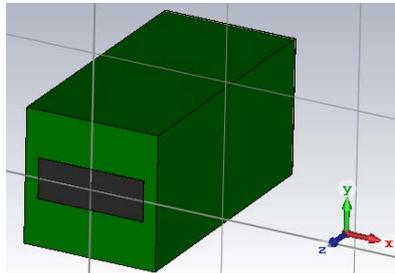


Рис. 1. Модель единичного элемента решетки – ПГНА прямоугольной геометрии с металлизированной диэлектрической подложкой, возбуждаемой дискретным портом, в пакете прикладных программ CST Studio Suite 2019: $l = 400$ нм, $w = 150$ нм, $a = b = 500$ нм

С помощью пакета прикладных программ CST Microwave Studio 2023 проведен электродинамический расчет частотных зависимостей элемента матрицы рассеяния $|S_{11}|$ – коэффициента отражения на входе ПГНА (с геометрическими размерами $l = 400$ нм; $w = 150$ нм, $a = b = 500$ нм) для различных значений химического потенциала графена μ_c в ИК-диапазоне частот $f = 15\text{--}60$ ТГц, результаты которого представлены на рис. 2, 3. Графен имеет следующие параметры: температура $T = 300$ К, параметр рассеяния $\tau = 1$ пс [22].

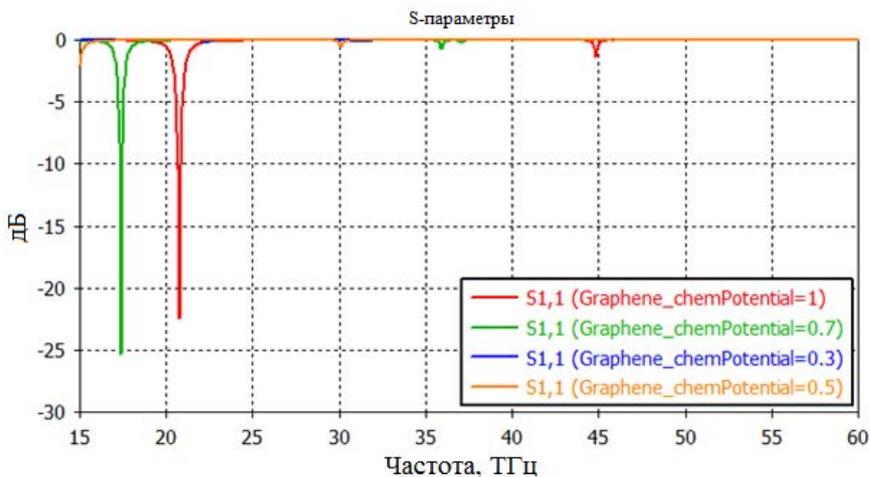


Рис. 2 Частотные зависимости элемента матрицы рассеяния $|S_{11}|$ ПГНА в ИК-диапазоне частот $f = 15\text{--}60$ ТГц для различных значений химического потенциала μ_c графена: 1 – синий цвет $\mu_{c1} = 0,3$ эВ; 2 – оранжевый цвет $\mu_{c2} = 0,5$ эВ; 3 – зеленый цвет $\mu_{c3} = 0,7$ эВ; 4 – красный цвет $\mu_{c4} = 1$ эВ; $l = 400$ нм; $w = 150$ нм; $a = b = 500$ нм; $T = 300$ К, $\tau = 1$ пс

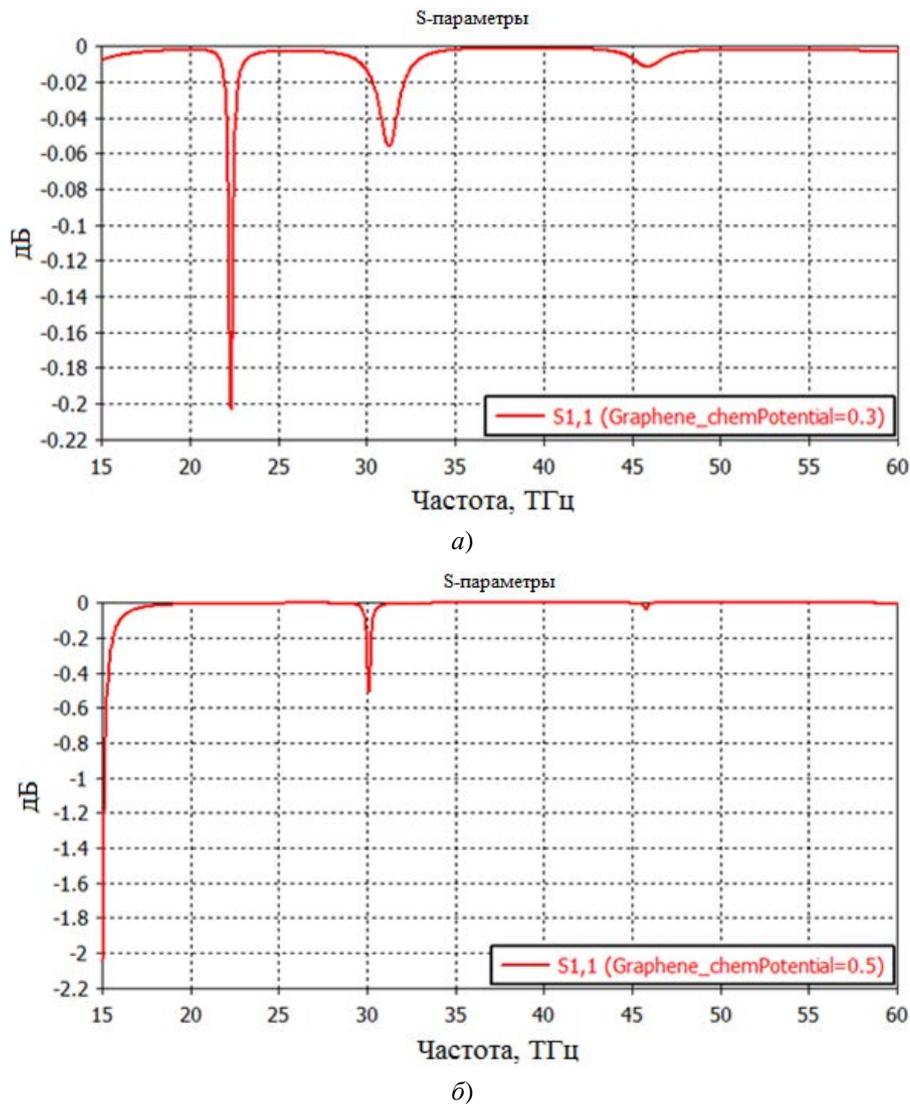


Рис. 3. Частотные зависимости элемента матрицы рассеяния $|S_{11}|$ ПГНА в ИК-диапазоне частот $f = 15\text{--}60$ ТГц для различных значений химического потенциала: $a - \mu_c = 0,3$ эВ; $b - \mu_c = 0,5$ эВ; $l = 400$ нм; $w = 150$ нм, $a = b = 500$ нм; $T = 300$ К, $\tau = 1$ пс

Графен характеризуется поверхностной проводимостью σ , определяемой формулой Кубо [22]. Поверхностная проводимость графена в соответствии с моделью Кубо зависит от химического потенциала графена μ_c , который можно изменять химическим легированием или приложением внешнего электрического поля [22, 23].

Из результатов моделирования (рис. 2, 3) следует, что при изменении химического потенциала μ_c положение по частоте и значения минимумов коэффициента отражения $|S_{11}|$ изменяются. Это объясняется тем, что в ТГц и дальнем, среднем ИК-диапазонах в графене, расположенном на поверхности диэлектрического слоя, могут распространяться поверхностные плазмон-поляритоны. В ограниченном графеновом элементе наблюдается резонанс ППП, при котором резко возрастает плотность поверхностного электрического тока и поглощение [24]. Минимумы коэффициента отражения $|S_{11}|$ на входе ПГНА обусловлены максимумами коэффициента поглощения P в графеновом элементе [24, 25].

Положению каждого минимума $|S_{11}|$ соответствует определенная частота плазмонного резонанса мод ППП основного и высших типов в графеновом элементе прямоугольной геометрии. При этом первый минимум $|S_{11}|$ соответствует резонансу основной моды ППП, другие же минимумы $|S_{11}|$ соответствуют резонансам высших мод ППП [24].

Из анализа полученных результатов можно констатировать, что при увеличении значения химического потенциала μ_c ($0,3\text{--}1$ эВ) частоты плазмонного резонанса $f_{\text{рез}}$ увеличиваются (рис. 2, 3), а именно резонансные частоты $f_{\text{рез}}$ основной моды ППП изменяются следующим образом: $f_{\text{рез1}} = 15$ ТГц

при $\mu_c = 0,5$ эВ; $f_{рез2} = 17,385$ ТГц при $\mu_c = 0,7$ эВ; $f_{рез3} = 20,715$ ТГц при $\mu_c = 1$ эВ; при этом минимальное значение $|S_{11}|$ равно $-25,4$ дБ при $\mu_c = 0,7$ эВ (см. рис. 2).

Резкое возрастание поверхностного электрического тока на графеновом элементе при плазмонном резонансе мод ППП приводит к усилению эффективности процесса излучения на резонансных частотах $f_{рез}$. Положению каждого минимума $|S_{11}|$ на частотах резонансов основной и высших мод ППП соответствуют определенные рабочие частоты эффективного излучения антенны, при этом минимальные значения коэффициента отражения $|S_{11}|$ определяют максимальную эффективность излучения.

Из результатов расчета (см. рис. 2) следует, что оптимальной рабочей точкой ПГНА является резонансная частота $f_{рез2} = 17,385$ ТГц при значении химического потенциала $\mu_c = 0,7$ эВ, где коэффициент отражения $|S_{11}|$ имеет минимальное значение $-25,4$ дБ и при этом максимальный коэффициент поглощения. Следовательно, на данной резонансной частоте возможна максимальная эффективность излучения ПГНА.

При увеличении химического потенциала ($\mu_c = 0,3-1$ эВ) резонансные частоты $f_{рез}$ минимумов коэффициента отражения $|S_{11}|$ сдвигаются в сторону более высоких частот ИК-диапазона, т.е. рабочие частоты ПГНА управляются приложением внешнего электрического поля и ПГНА становится частотно-сканируемой.

Результаты моделирования с помощью пакета прикладных программ CST Microwave Studio 2023 ДН ПГНА на резонансных частотах $f_{рез}$ основной моды ППП для различных значений химического потенциала μ_c (0,5 эВ, 0,7 эВ, 1 эВ) и сканирование по частоте в среднем ИК-диапазоне приведены на рис. 4–6.

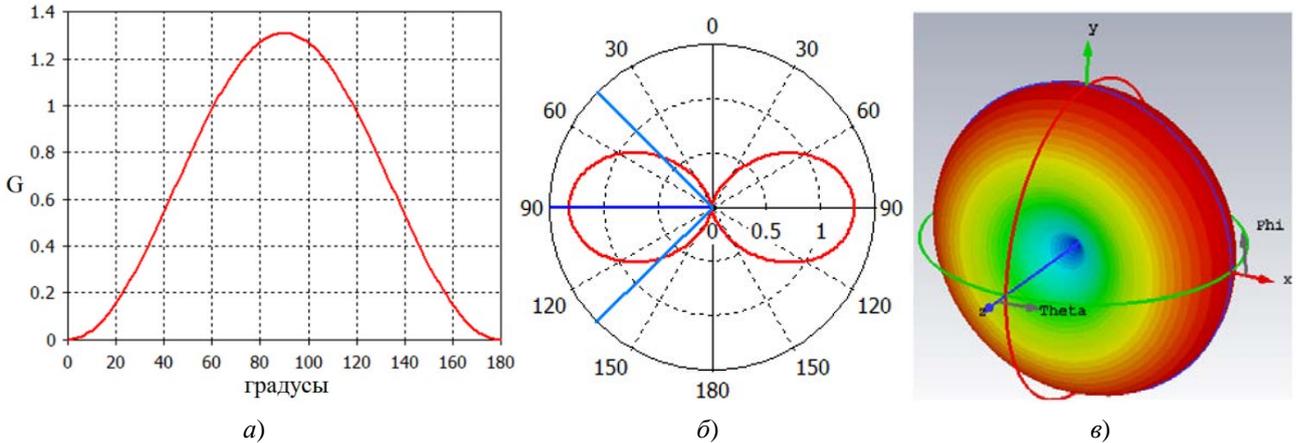


Рис. 4. ДН ПГНА на резонансной частоте $f_{рез1}$ основной моды ППП в среднем ИК-диапазоне: $f_{рез1} = 15$ ТГц; $\mu_c = 0,5$ эВ в декартовой, полярной в E -плоскости (в зависимости от θ , при $\varphi = 0$ град.), сферической системах координат; коэффициент усиления (амплитуда главного лепестка) $G = 1,31$, направление главного лепестка 90 град., $\Theta_{0,5} = 89,8$ град.

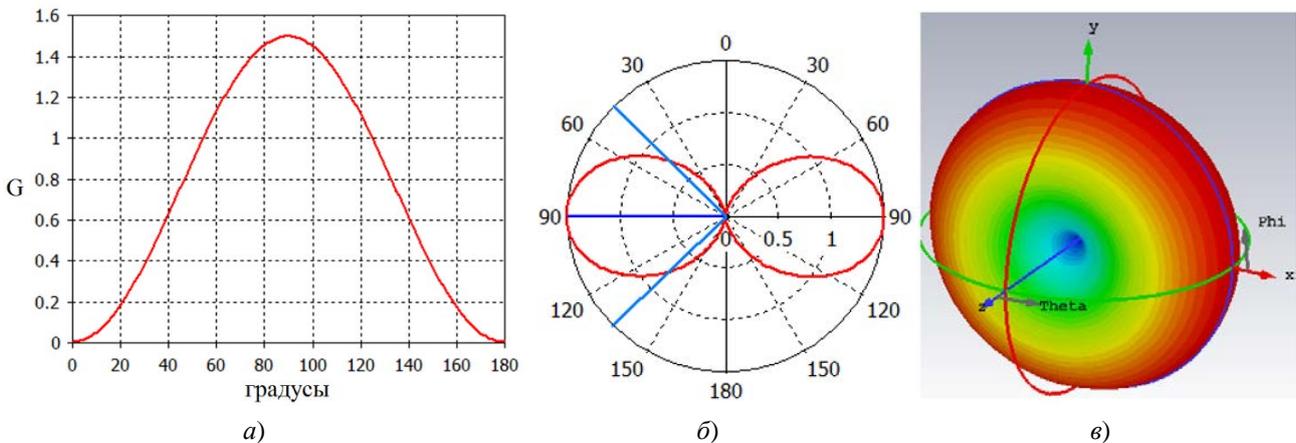


Рис. 5. ДН ПГНА на резонансной частоте $f_{рез2}$ основной моды ППП в среднем ИК-диапазоне: $f_{рез2} = 17,385$ ТГц; $\mu_c = 0,7$ эВ в декартовой, полярной в E -плоскости (в зависимости от θ , при $\varphi = 0$ град.), сферической системах координат; коэффициент усиления (амплитуда главного лепестка) $G = 1,5$, направление главного лепестка 90 град., $\Theta_{0,5} = 90$ град.

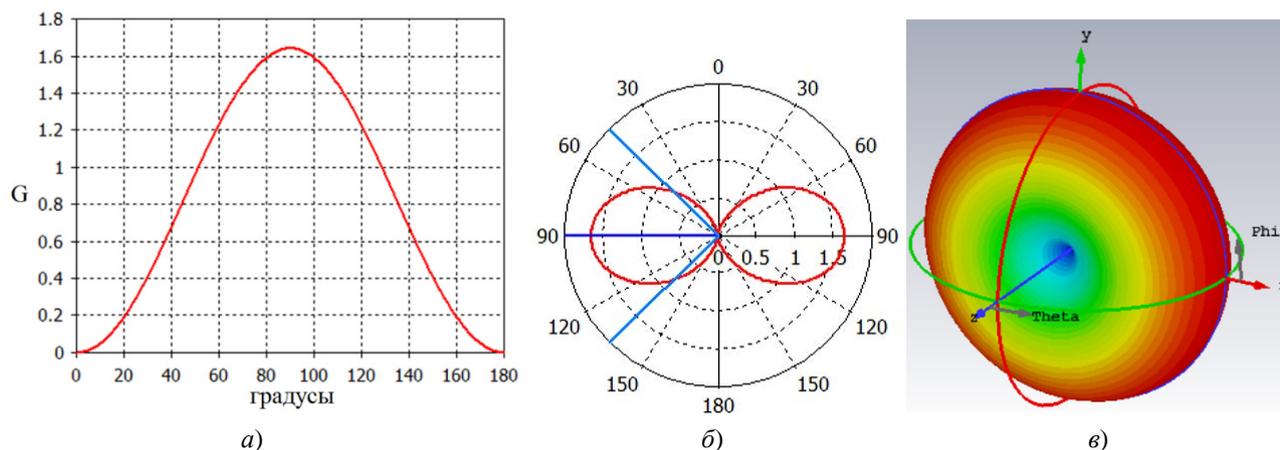


Рис. 6. ДН ПГНА на резонансной частоте $f_{рез3}$ основной моды ППП в среднем ИК-диапазоне: $f_{рез3} = 20.715$ ТГц; $\mu_c = 1$ эВ в декартовой, полярной в E -плоскости (в зависимости от θ , при $\varphi = 0$ град.), сферической системах координат; коэффициент усиления (амплитуда главного лепестка) $G = 1,64$, направление главного лепестка 90 град., $\Theta_{0,5} = 89,8$ град.

Результаты, представленные на рис. 4–7, демонстрируют возможность перестройки рабочей частоты ПГНА (сканирование по частоте) на резонансных частотах $f_{рез}$ основной моды ППП в среднем ИК-диапазоне при изменении химического потенциала μ_c в интервале значений 0,5–1 эВ (приложением внешнего электрического поля).

На резонансных частотах $f_{рез1}$, $f_{рез2}$, $f_{рез3}$ основной моды ППП наблюдается резонанс электрического тока, создаваемого стоячей полуволной ППП вдоль широкой стороны прямоугольного графенового элемента [24] и, следовательно, полуволновое распределение тока j_s по длине графенового элемента. Форма 3D ДН излучения, создаваемого этим полуволновым (по длине волны ППП) электрическим излучателем, тороидальная (рис. 3–6, в). Ось 2D ДН ориентирована в экваториальной плоскости параллельно относительно продольного излучающего тока, ДН в E -плоскости (в зависимости от θ , при $\varphi = 0$ град) (рис. 3–6, б) соответствует ДН полуволнового симметричного вибратора [26].

Из результатов моделирования (рис. 4–6) следует, что при изменении значения химического потенциала μ_c графена в интервале от 0,5 до 1 эВ происходит изменение рабочих частот, т.е. сканирование по частоте в среднем ИК-диапазоне и изменяются основные характеристики ПГНА: направление главного лепестка ДН изменяется с 88 до 90 град., коэффициент усиления (амплитуда главного лепестка) с $G = 1,64$ по 1,66, ширина ДН по уровню половинной мощности $\Theta_{0,5}$ с 90 по 89,8 град.

Результаты моделирования управляемости ДН решетки ПГНА и сканирование по частоте в среднем ИК-диапазоне при изменении химического потенциала

На втором этапе с помощью пакета прикладных программ CST Studio Suite 2019 проведено моделирование характеристик решетки ПГНА (с числом единичных элементов 16×16 $N = 256$): элемента матрицы рассеяния $|S_{11}|$ и 2D, 3D ДН в ИК-диапазоне частот для различных значений химического потенциала графена μ_c .

На рис. 7 представлена модель решетки ПГНА ($N = 256$) среднего ИК-диапазона. Модель построена в пакете прикладных программ CST Studio Suite 2023.

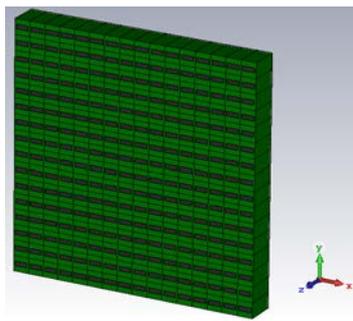


Рис. 7. Модель решетки ПГНА среднего ИК-диапазона: $N = 256$; $l = 400$ нм; $w = 150$ нм; $a = b = 500$ нм

Результаты моделирования с помощью пакета прикладных программ CST Microwave Studio 2023 ДН решетки ПГНА на резонансных частотах основной моды ППП для различных значений химического потенциала μ_c (0,5; 0,7; 1 эВ) приведены на рис. 8–11.

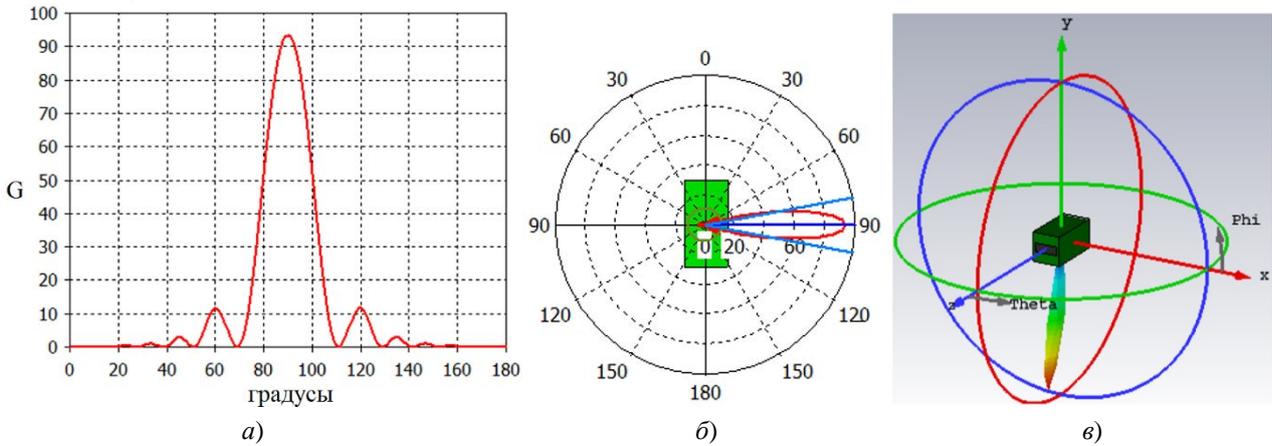


Рис. 8. ДН решетки ПГНА на резонансной частоте $f_{рез1}$ основной моды ППП и сканирование по частоте в среднем ИК-диапазоне:
 а – в декартовой; б – полярной в E -плоскости (в зависимости от θ , при $\varphi = 0$ град.);
 в – сферической системах координат; $N = 256$; $f_{рез1} = 15$ ТГц; $\mu_c = 0,5$ эВ;
 коэффициент усиления (амплитуда главного лепестка) $G = 93,4$,
 направление главного лепестка 90 град., $\Theta_{0,5} = 21,3$ град., уровень боковых лепестков $-9,1$ дБ

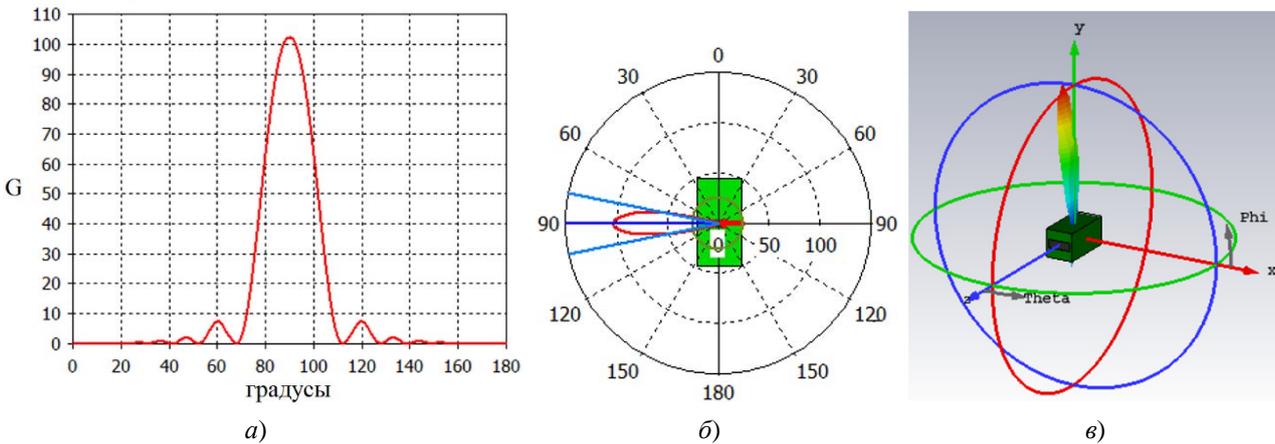


Рис. 9. ДН решетки ПГНА на резонансной частоте $f_{рез2}$ основной моды ППП и сканирование по частоте в среднем ИК-диапазоне:
 а – в декартовой; б – полярной в E -плоскости (в зависимости от θ , при $\varphi = 0$ град.);
 в – сферической системах координат; $N = 256$; $f_{рез2} = 17,385$ ТГц; $\mu_c = 0,7$ эВ;
 коэффициент усиления (амплитуда главного лепестка) $G = 102$,
 направление главного лепестка 90 град., $\Theta_{0,5} = 23,2$ град., уровень боковых лепестков -6 дБ

Из результатов моделирования следует, что при изменении химического потенциала μ_c графена в интервале значения от 0,5 до 1 эВ происходит изменение рабочих частот, т.е. сканирование по частоте в среднем ИК-диапазоне, и изменяются основные характеристики решетки ПГНА: направление главного лепестка ДН с 72 до 90 град., коэффициент усиления (амплитуда главного лепестка) с $G = 76,3$ по 102 (максимальное значение $G = 102$ наблюдается на частоте основного минимума $|S_{11}|$ при $\mu_c = 0,7$ эВ), ширина ДН по уровню половинной мощности $\Theta_{0,5}$ с 19,8 по 28,7 град., уровень боковых лепестков с $-2,1$ до -14 дБ.

Результаты моделирования с помощью пакета прикладных программ CST Microwave Studio 2023 ДН решетки ПГНА на резонансной частоте моды ППП второго порядка $f_{рез}^2 = 22,335$ ТГц (см. рис. 3,а) для значения химического потенциала $\mu_c = 0,3$ эВ приведены на рис. 11.

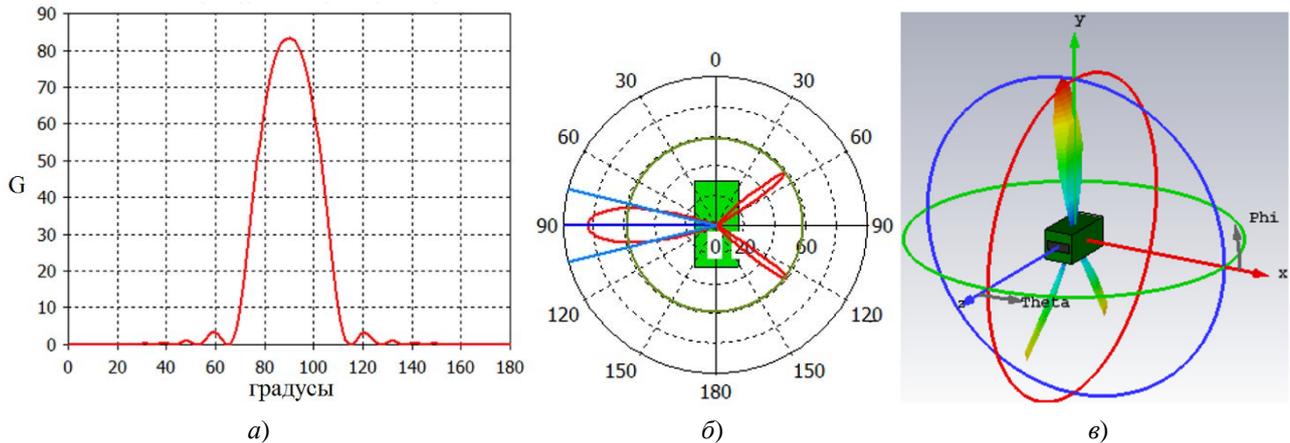


Рис. 10. ДН решетки ПГНА на резонансной частоте $f_{\text{рез}3}$ основной моды ППП и сканирование по частоте в среднем ИК-диапазоне:
a – в декартовой; *б* – полярной в *E*-плоскости (в зависимости от θ , при $\varphi = 0$ град.);
в – сферической системах координат; $N = 256$; $f_{\text{рез}3} = 20,715$ ТГц; $\mu_c = 1$ эВ;
 коэффициент усиления (амплитуда главного лепестка) $G = 83,3$,
 направление главного лепестка 90 град., $\Theta_{0,5} = 28,7$ град., уровень боковых лепестков $-1,5$ дБ

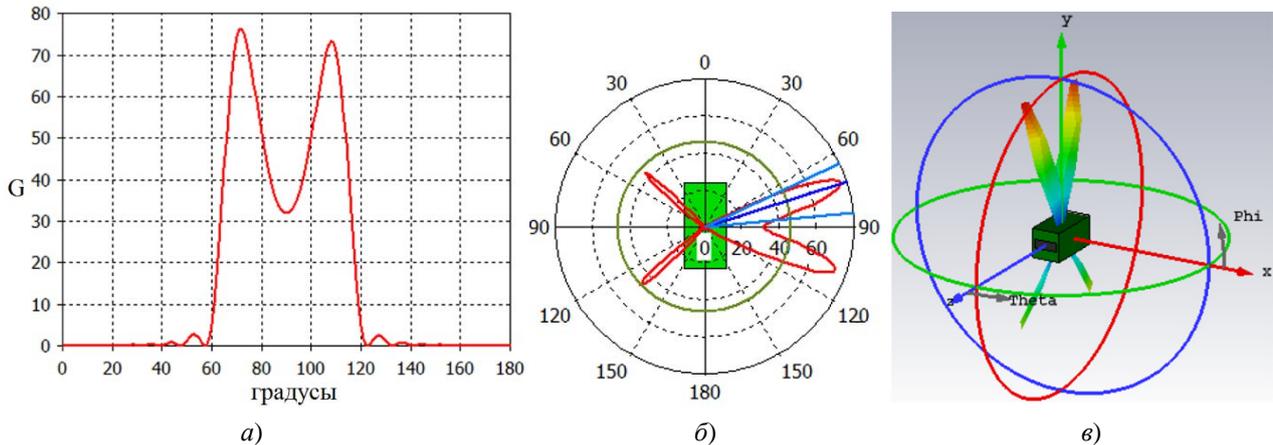


Рис. 11. ДН решетки ПГНА среднего ИК-диапазона:
a – в декартовой; *б* – полярной в *E*-плоскости (в зависимости от θ , при $\varphi = 0$ град.);
в – сферической системах координат; $N = 256$; $f_{\text{рез}}^2 = 22,335$ ТГц; $\mu_c = 0,3$ эВ;
 коэффициент усиления (амплитуда главного лепестка) $G = 76,3$,
 направление главного лепестка 72 град., $\Theta_{0,5} = 19,8$ град., уровень боковых лепестков $-2,1$ дБ

При переходе от резонансной частоты $f_{\text{рез}}^1$ основной моды ППП к резонансной частоте $f_{\text{рез}}^2$ второй моды ППП происходит изменение формы главного лепестка ДН (рис. 11). На резонансной частоте $f_{\text{рез}}^2$ второй моды ППП наблюдается резонанс электрического тока, создаваемого двумя стоячими полуволнами ППП вдоль широкой стороны прямоугольного графенового элемента [24], и, следовательно, имеется волновое распределение поверхностного электрического тока по длине графенового элемента. Форма главного лепестка ДН излучения, создаваемого этим волновым (по длине волны ППП) электрическим излучателем, 2-лучевая (рис. 11), и это изменение соответствует характеру изменения формы ДН симметричного вибратора при изменении его электрической длины [26].

Заключение

Из результатов моделирования решетки ПГНА среднего ИК-диапазона с помощью пакета прикладных программ CST Microwave Studio 2023 следует, что с увеличением химического потенциала μ_c от 0,5 до 1 эВ возрастает коэффициент усиления (амплитуда главного лепестка), происходит сужение главного лепестка ДН, уменьшение уровня боковых лепестков; с увеличением числа

единичных элементов N уровни боковых лепестков ДН уменьшаются, а также наблюдается лучшая управляемость главного лепестка ДН. Уровень главного лепестка на основном минимуме на частоте основного минимума $|S_{11}|$ при $\mu_c = 0,7$ эВ имеет максимальное значение, поскольку величина коэффициента отражения S_{11} минимальна (т.е. максимальный коэффициент прохождения S_{21}), следовательно, эффективность излучения возрастает.

Для реконфигурируемой решетки ПГНА показана возможность перестройки рабочих частот (сканирование по частоте) в среднем ИК-диапазоне и формирование 2-лучевой формы главного лепестка ДН путем изменения химического потенциала графена (при приложении внешнего электрического поля) без изменения их геометрии и размеров.

Исследованные решетки ПГНА среднего ИК-диапазона допускают динамическую реконфигурацию и изменение формы ДН посредством приложенного внешнего электрического поля и могут найти применение в визуализации, зондировании и спектроскопии, ИК-системах беспроводной связи, а также ИК сенсорных системах.

Список литературы

1. Zhaoyi Li, Nanfang Yu. Modulation of mid-infrared light using graphene-metal plasmonic antennas // *Applied Physics Letters*. 2013. Vol. 102. P. 131108. doi: 10.1063/1.4800931
2. Yousefi L., Foster A. C. Waveguide-fed optical hybrid plasmonic patch nano-antenna // *Optics Express*. 2012. Vol. 20, № 16. P. 18327–18335.
3. Hosseiniadjad S. E., Komjani N., Noghani M. T. A comparison of graphene and noble metals as conductors for plasmonic one-dimensional waveguides // *IEEE Transactions on Nanotechnology*. 2015. Vol. 14. P. 829–836.
4. Tamagnone M., Gomez-Diaz J. S., Mosig J. R., Perruisseau-Carrier J. Reconfigurable terahertz plasmonic antenna concept using a graphene stack // *Applied Physics Letters*. 2012. Vol. 101, № 21. P. 1–4.
5. Rodrigues N. R. N. M., de Oliveira R. M. S., Dmitriev V. Smart terahertz graphene antenna // *IEEE Antennas and Propagation Magazine*. 2018. Vol. 60, № 5. P. 26–40.
6. Lamri I. E., Ahmad S., Nakmouche M. F. [et al.]. Design and development of a graphene-based reconfigurable patch antenna array for THz applications // *Frequenz*. 2023. Vol. 77, № 3-4. P. 219–228. doi: 10.1515/freq-2022-0051
7. Jorner J. M., Akyildiz I. F. Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band // *IEEE Trans. Wireless Commun.* 2011. Vol. 10, № 10. P. 3211–3221.
8. Shi L. P., Zhang Q. H., Zhang S. H. [et al.]. Efficient graphene reconfigurable reflectarray antenna electromagnetic response prediction using deep learning // *IEEE Access*. 2021. Vol. 9. P. 22671–22678.
9. Perruisseau-Carrier J. Graphene for antenna applications: opportunities and challenges from microwaves to THz // 2012 Loughborough Antennas & Propagation Conference (LAPC). Loughborough, UK, LAPC, 2012. P. 1–4.
10. Ullah Z., Nawi I., Witjaksono G. [et al.]. Electrically enhanced graphene-metal plasmonic antenna for infrared sensing // *Optik*. 2021. Vol. 241. P. 166961. doi: 10.1016/j.ijleo.2021.166961
11. Mohanta J., Meher P. R., Mishra S. K. Nano-antennas Using Single and Double Isolated Graphene for Mid-infrared Applications // 2019 Photonics & Electromagnetics Research Symposium – Spring (PIERS-Spring). Rome, 2019. P. 4292–4299. doi: 10.1109/PIERS-Spring46901.2019.9017471
12. Dong Z., Sun C., Si J., Deng X. A tunable plasmonic nano-antenna based on metal – graphene double-nanorods // *Laser Phys. Lett.* 2018. Vol. 15. P. 056202.
13. Hajian H., Ghobadi A., Butun B. [et al.]. Tunable, omnidirectional, and nearly perfect resonant absorptions by a graphene-hBN- based hole array metamaterial // *Optics Express*. 2018. Vol. 26. P. 16940–16954.
14. Hosseiniadjad S. E., Neshat M., Faraji-dana R., Lemme M. Reconfigurable THz plasmonic antenna based on few-layer graphene with high radiation efficiency // *Nanomaterials*. 2018. Vol. 8. P. 577. doi: 10.3390/nano8080577
15. Baqir M. A., Choudhury P. K., Farmani A. [et al.]. A novel method to detect functional microRNA regulatory modules by bicliques merging // *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 2016. Vol. 13. P. 549–556. doi: 10.1109/JPHOT.2019.2931586
16. Ullah Z., Nawi I., Witjaksono G. [et al.]. Dynamic absorption enhancement and equivalent resonant circuit modeling of tunable graphene-metal hybrid antenna // *Sensors*. 2020. Vol. 20. P. 3187.
17. Wu P. C., Pala R. A., Kafae Shirmanesh G. [et al.]. Dynamic beam steering with all-dielectric electro-optic III–V multiple-quantum-well metasurfaces // *Nat. Commun.* 2019. Vol. 10. P. 1–9. doi: 10.1038/s41467-019-11598-8
18. Bandurina D. A., Svintsov D., Gayduchenko I. [et al.]. Resonant terahertz detection using graphene plasmons // *Nat. Commun.* 2018. Vol. 9. P. 4–11. doi: 10.1038/s41467-018-07848-w
19. Sherrott M. C., Hon P. W. C., Fountaine K. T. [et al.]. Experimental demonstration of $>230^\circ$ phase modulation in gate-tunable graphene-gold reconfigurable mid-infrared metasurfaces // *Nano Lett.* 2017. Vol. 17. P. 3027–3034. doi: 10.1021/acs.nanolett.7b00359
20. CST Microwave Studio 2023. URL: <https://www.3ds.com/products/simulia>

21. Курушин А. А. Школа проектирования СВЧ устройств в CST STUDIO SUITE. М. : One-Book, 2014. 433 с.
22. Hanson G. W. Dyadic Green's functions and guided surface waves for a surface conductivity model of graphene // *J. of Appl. Phys.* 2008. Vol. 103. P. 064302.
23. Голованов О. А., Макеева Г. С., Вареница В. В. Проводимость графена в терагерцовом и инфракрасном диапазонах частот // *Надежность и качество сложных систем.* 2014. № 4 (8). С. 26–33.
24. Лерер А. М., Макеева Г. С. Поляризационные эффекты и резонансное поглощение при дифракции терагерцовых волн на графеновых метаповерхностях // *Оптика и спектроскопия.* 2018. Т. 125, вып. 6. С. 838–843.
25. Голованов О. А., Макеева Г. С., Ринкевич А. Б. Взаимодействие электромагнитных волн с периодически решетками микро- и нанолент графена в терагерцовом диапазоне // *Журнал технической физики.* 2016. Т. 86, вып. 2. С. 119–126.
26. Драбкин А. Л., Зузенко В. Л. Антенно-фидерные устройства. М. : Сов. радио, 1961. 816 с.

References

1. Zhaoyi Li, Nanfang Yu. Modulation of mid-infrared light using graphene-metal plasmonic antennas. *Applied Physics Letters.* 2013;102:131108. doi: 10.1063/1.4800931
2. Yousefi L., Foster A.C. Waveguide-fed optical hybrid plasmonic patch nano-antenna. *Optics Express.* 2012;20(16):18327–18335.
3. Hosseininejad S.E., Komjani N., Noghani M.T. A comparison of graphene and noble metals as conductors for plasmonic one-dimensional waveguides. *IEEE Transactions on Nanotechnology.* 2015;14:829–836.
4. Tamagnone M., Gomez-Diaz J.S., Mosig J.R., Perruisseau-Carrier J. Reconfigurable terahertz plasmonic antenna concept using a graphene stack. *Applied Physics Letters.* 2012;101(21):1–4.
5. Rodrigues N.R.N.M., de Oliveira R.M.S., Dmitriev V. Smart terahertz graphene antenna. *IEEE Antennas and Propagation Magazine.* 2018;60(5):26–40.
6. Lamri I.E., Ahmad S., Nakmouche M.F. et al. Design and development of a graphene-based reconfigurable patch antenna array for THz applications. *Frequenz.* 2023;77(3-4):219–228. doi: 10.1515/freq-2022-0051
7. Jornet J.M., Akyildiz I.F. Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band. *IEEE Trans. Wireless Commun.* 2011;10(10):3211–3221.
8. Shi L.P., Zhang Q.H., Zhang S.H. et al. Efficient graphene reconfigurable reflectarray antenna electromagnetic response prediction using deep learning. *IEEE Access.* 2021;9:22671–22678.
9. Perruisseau-Carrier J. Graphene for antenna applications: opportunities and challenges from microwaves to THz. *2012 Loughborough Antennas & Propagation Conference (LAPC).* Loughborough, UK, LAPC, 2012:1–4.
10. Ullah Z., Nawi I., Witjaksono G. et al. Electrically enhanced graphene-metal plasmonic antenna for infrared sensing. *Optik.* 2021;241:166961. doi: 10.1016/j.ijleo.2021.166961
11. Mohanta J., Meher P.R., Mishra S.K. Nano-antennas Using Single and Double Isolated Graphene for Mid-infrared Applications. *2019 Photonics & Electromagnetics Research Symposium – Spring (PIERS-Spring).* Rome, 2019:4292–4299. doi: 10.1109/PIERS-Spring46901.2019.9017471.
12. Dong Z., Sun C., Si J., Deng X. A tunable plasmonic nano-antenna based on metal – graphene double-nanorods. *Laser Phys. Lett.* 2018;15:056202.
13. Hajian H., Ghobadi A., Butun B. et al. Tunable, omnidirectional, and nearly perfect resonant absorptions by a graphene-hBN- based hole array metamaterial. *Optics Express.* 2018;26:16940–16954.
14. Hosseininejad S.E., Neshat M., Faraji-dana R., Lemme M. Reconfigurable THz plasmonic antenna based on few-layer graphene with high radiation efficiency. *Nanomaterials.* 2018;8:577. doi: 10.3390/nano8080577
15. Baqir M.A., Choudhury P.K., Farmani A. et al. A novel method to detect functional microRNA regulatory modules by bicliques merging. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 2016;13:549–556. doi: 10.1109/JPHOT.2019.2931586
16. Ullah Z., Nawi I., Witjaksono G. et al. Dynamic absorption enhancement and equivalent resonant circuit modeling of tunable graphene-metal hybrid antenna. *Sensors.* 2020;20:3187.
17. Wu P.C., Pala R.A., Kafaie Shirmanesh G. et al. Dynamic beam steering with all-dielectric electro-optic III–V multiple-quantum-well metasurfaces. *Nat. Commun.* 2019;10:1–9. doi: 10.1038/s41467-019-11598-8
18. Bandurin D.A., Svinsov D., Gayduchenko I. et al. Resonant terahertz detection using graphene plasmons. *Nat. Commun.* 2018;9:4–11. doi: 10.1038/s41467-018-07848-w
19. Sherrott M.C., Hon P.W.C., Fountaine K.T. et al. Experimental demonstration of >230° phase modulation in gate-tunable graphene-gold reconfigurable mid-infrared metasurfaces. *Nano Lett.* 2017;17:3027–3034. doi: 10.1021/acs.nanolett.7b00359
20. *CST Microwave Studio 2023.* Available at: <https://www.3ds.com/products/simulia>
21. Kurushin A.A. *Shkola proektirovaniya SVCh ustroystv v CST STUDIO SUITE = School of microwave device design in CST STUDIO SUITE.* Moscow: One-Book, 2014:433. (In Russ.)
22. Hanson G.W. Dyadic Green's functions and guided surface waves for a surface conductivity model of graphene. *J. of Appl. Phys.* 2008;103:064302.

23. Golovanov O.A., Makeeva G.S., Varenitsa V.V. Graphene conductivity in terahertz and infrared frequency ranges. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems*. 2014;(4):26–33. (In Russ.)
24. Lerer A.M., Makeeva G.S. Polarization effects and resonant absorption in the diffraction of terahertz waves on graphene metasurfaces. *Optika i spektroskopiya = Optics and spectroscopy*. 2018;125(6):838–843. (In Russ.)
25. Golovanov O.A., Makeeva G.S., Rinkevich A.B. Interaction of electromagnetic waves with periodic lattices of micro- and nanofilaments of graphene in the terahertz range. *Zhurnal tekhnicheskoy fiziki = Journal of Technical Physics*. 2016;86(2):119–126. (In Russ.)
26. Drabkin A.L., Zuzenko V.L. *Antenno-fidernye ustroystva = Antenna-feeder devices*. Moscow: Sov. radio, 1961:816. (In Russ.)

Информация об авторах / Information about the authors

Галина Степановна Макеева

доктор технических наук, профессор, профессор
кафедры радиотехники и радиоэлектронных систем,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: radiotech@pnzgu.ru

Galina S. Makeeva

Doctor of technical sciences, professor,
professor of the sub-department
of radio engineering and radio-electronic systems,
Penza State University
(40 Krasnaya street, Penza, Russia)

Автор заявляет об отсутствии конфликта интересов /

The author declares no conflicts of interests.

Поступила в редакцию/Received 16.06.2024

Поступила после рецензирования/Revised 26.07.2024

Принята к публикации/Accepted 31.08.2024

МЕТОД МАТЕМАТИЧЕСКОГО ПРОТОТИПИРОВАНИЯ ЭНЕРГЕТИЧЕСКИХ ПРОЦЕССОВ КАК ИНСТРУМЕНТ СИНТЕЗА МАТЕМАТИЧЕСКОГО ЯДРА ЦИФРОВЫХ ДВОЙНИКОВ РАЗЛИЧНЫХ ФИЗИКО-ХИМИЧЕСКИХ СИСТЕМ

И. Е. Старостин

Московский государственный технический университет гражданской авиации, Москва, Россия
starostinigo@yandex.ru

Аннотация. *Актуальность и цели.* Цифровой двойник по своей сути представляет собой цифровую копию объекта, благодаря непрерывному обмену данными с объектом воспроизводящую актуальное его состояние. Это достигается путем применения в качестве математического ядра цифрового двойника математической модели, основывающейся на реальном поведении объекта, параметры которой обновляются в соответствии с актуальной информацией, получаемой с объекта. Для построения моделей систем различной физической и химической природы авторами был предложен метод математического прототипирования энергетических процессов, основывающийся на методах современной неравновесной термодинамики, механики, электродинамики. Благодаря чему упомянутый метод позволяет создать модели динамики процессов в системе, не противоречащие общим физическим законам, а также особенностям протекания процессов в конкретной физической или химической системе. Затем полученные модели преобразовываются к моделям систем, применимых для решения практических задач, с точностью до определяемых из экспериментальных данных постоянных коэффициентов. Однако здесь актуальным является вопрос выбора базиса, в котором строится преобразованная модель рассматриваемой системы и выбора алгоритма обучения предложенной модели. Решению приведенных вопросов посвящена настоящая работа. *Материалы и методы.* Синтез уравнений динамики физических и химических процессов осуществляется на базе метода математического прототипирования энергетических процессов. Преобразованные модели строятся путем качественного анализа динамик системы в различных режимах ее работы с последующим анализом связей между параметрами ее динамик и параметрами входных воздействий с дальнейшим построением интерполяционной преобразованной модели. Интерполяционные методы выбираются на основе вышеупомянутых результатов качественного анализа. *Результаты.* Предложенная концепция построения моделей системы дает корректные модели системы, вбирающие в себя качественное ее поведение. Это обеспечивает корректность математического ядра цифровых двойников при работе систем в различных режимах работы системы, для которых строятся вышеупомянутые цифровые двойники. *Выводы.* Предложенная концепция программной части цифровых двойников имеет широкую область их применения благодаря широте областей применимости метода математического прототипирования энергетических процессов.

Ключевые слова: метод математического прототипирования энергетических процессов, цифровые двойники, интерполяция, качественный анализ

Для цитирования: Старостин И. Е. Метод математического прототипирования энергетических процессов как инструмент синтеза математического ядра цифровых двойников различных физико-химических систем // Надежность и качество сложных систем. 2024. № 3. С. 41–50. doi: 10.21685/2307-4205-2024-3-5

METHOD OF MATHEMATICAL PROTOTYPING OF ENERGY PROCESSES AS A TOOL FOR SYNTHESIS OF THE MATHEMATICAL KERNEL OF DIGITAL TWINNINGS OF DIFFERENT PHYSICAL AND CHEMICAL SYSTEMS

I.E. Starostin

Moscow State Technical University of Civil Aviation, Moscow, Russia
starostinigo@yandex.ru

Abstract. *Background.* A digital twin is essentially a digital copy of an object, which, thanks to the continuous exchange of data with the object, reproduces its current state. This is achieved by using a mathematical model as the mathematical core of the digital twin, based on the real behavior of the object, the parameters of which are updated in accordance with the current information received from the object. To build models of systems of various physical

and chemical natures, the authors proposed a method for mathematical prototyping of energy processes, based on the methods of modern nonequilibrium thermodynamics, mechanics, and electrodynamics. Due to this, the mentioned method provides models of the dynamics of processes in a system that do not contradict general physical laws, as well as the peculiarities of the course of processes in a specific physical or chemical system. Then the resulting models are converted to models of systems applicable to solve practical problems, up to constant coefficients determined from experimental data. However, the relevant issue here is the choice of basis in which the transformed model of the system under consideration is built and the choice of training algorithm for the proposed model. This work is devoted to solving the above questions. *Materials and methods.* The synthesis of equations for the dynamics of physical and chemical processes is carried out on the basis of the method of mathematical prototyping of energy processes. Transformed models are built by qualitative analysis of the dynamics of the system in various modes of its operation, followed by analysis of the relationships between the parameters of its dynamics and the parameters of input influences with the further construction of an interpolation transformed model. Interpolation methods are selected based on the above-mentioned results of qualitative analysis. *Results.* The proposed concept of constructing system models provides correct models of the system that incorporate its qualitative behavior. This ensures the correctness of the mathematical core of digital twins when operating systems in various operating modes of the system for which the above-mentioned digital twins are built. *Conclusions.* The proposed concept of the software part of digital twins has a wide range of application due to the breadth of areas of applicability of the method of mathematical prototyping of energy processes.

Keywords: method of mathematical prototyping of energy processes, digital twins, interpolation, qualitative analysis

For citation: Starostin I.E. Method of mathematical prototyping of energy processes as a tool for synthesis of the mathematical kernel of digital twins of different physical and chemical systems. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):41–50. (In Russ.). doi: 10.21685/2307-4205-2024-3-5

Введение

Концепция цифрового двойника призвана помочь предприятиям быстрее обнаруживать физические проблемы, точнее предсказывать их результаты и производить более качественные продукты [1–4]. Также цифровые двойники, позволяющие в режиме реального времени диагностировать и прогнозировать различные дефекты, аварийные режимы, имеют важное значение в процессе эксплуатации и технического обслуживания объекта [1–4]. Важнейшей составляющей математического ядра (МЯ) цифрового двойника является математическая модель, воспроизводящая реальное поведение объекта [5].

Для построения упомянутых математических моделей (ММ) объектов, характеризующихся протеканием в них физических и химических процессов, автором был предложен в рамках современной неравновесной термодинамики, механики и электродинамики метод математического прототипирования энергетических процессов (ММПЭП) [6–8]. Построенные ММПЭП модели динамики процессов различной физической и химической природы не противоречат общим физическим законам (законам сохранения, началам термодинамики, и т.д.), а также физическим особенностям протекания процессов в рассматриваемом объекте [6, 7]. ММ системы, которая является основой математического ядра цифровых двойников, получается путем преобразования уравнений ММПЭП [7, 8], задав аналитическое выражение общего решения систем дифференциальных уравнений ММПЭП, вбирающего в себя физические особенности протекания процессов в рассматриваемом объекте [7].

Упомянутые аналитические выражения преобразованной ММ могут быть заданы с использованием методов интерполяции опорных динамик (также вбирающих в себя физические особенности функционирования конкретной системы) [7, 9]. Это также гарантирует корректность преобразованных аналитических моделей [7]. Также использование методов интерполяции дает возможность строить ММ непосредственно на опорных экспериментальных данных [7, 9]. Настоящая работа посвящена построению ММ, положенных в основу МЯ цифрового двойника.

Материалы и методы

В соответствие с ММПЭП состояние системы характеризуется независимо от ее предыстории параметрами состояния, изменяющимися в результате протекания физических и химических процессов в ней [6]. Причиной протекания процессов в системе являются динамические силы, определяемые через частные производные свободной энергии по координатам состояния, расходуемой на протекание процессов в системе [6]. Однако динамические силы однозначно не определяют протекание процессов в системе; независимо от динамических сил протекание процессов в системе определяется кинетическими свойствами, «шкалой» которых является кинетическая матрица (рис. 1) [6]. Динамика процессов в системе в свою очередь определяет динамику измеряемых пара-

метров (ИП) и контролируемых параметров (КП) (см. рис. 1) [6]. Приведенные на рис. 1 факторы определяют общую структуру дифференциальных уравнений динамики процессов, а также динамики ИП и КП, получаемых ММПЭП [6].

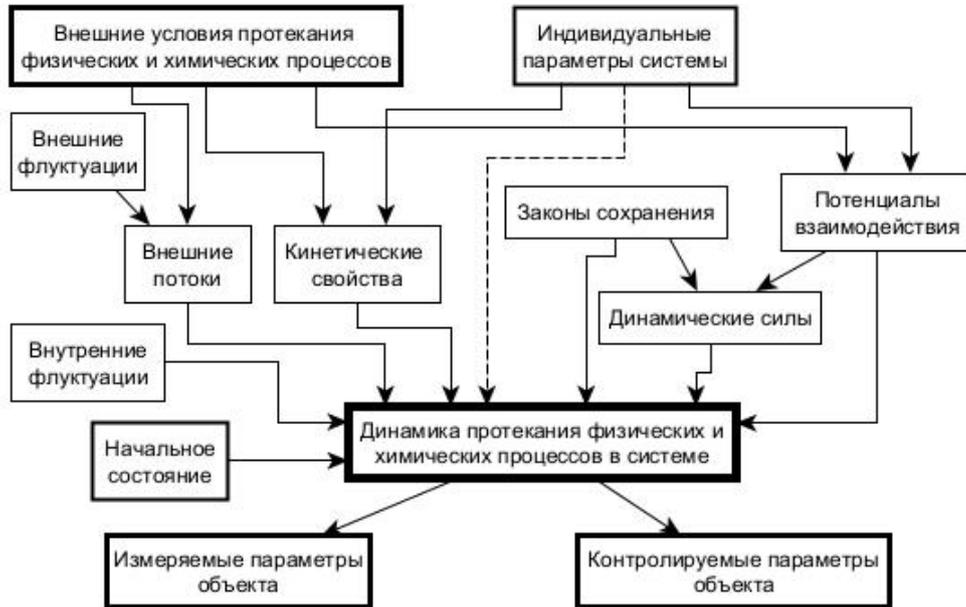


Рис. 1. Факторы, определяющие динамику физических и химических процессов

В общем случае ММ объекта (рис. 2) представляет собой алгоритм определения из экспериментальных данных ИП объекта начального состояния и индивидуальных параметров системы с последующим определением КП системы (с учетом внешних условий протекания процессов) (видно из рис. 1) [6, 8].

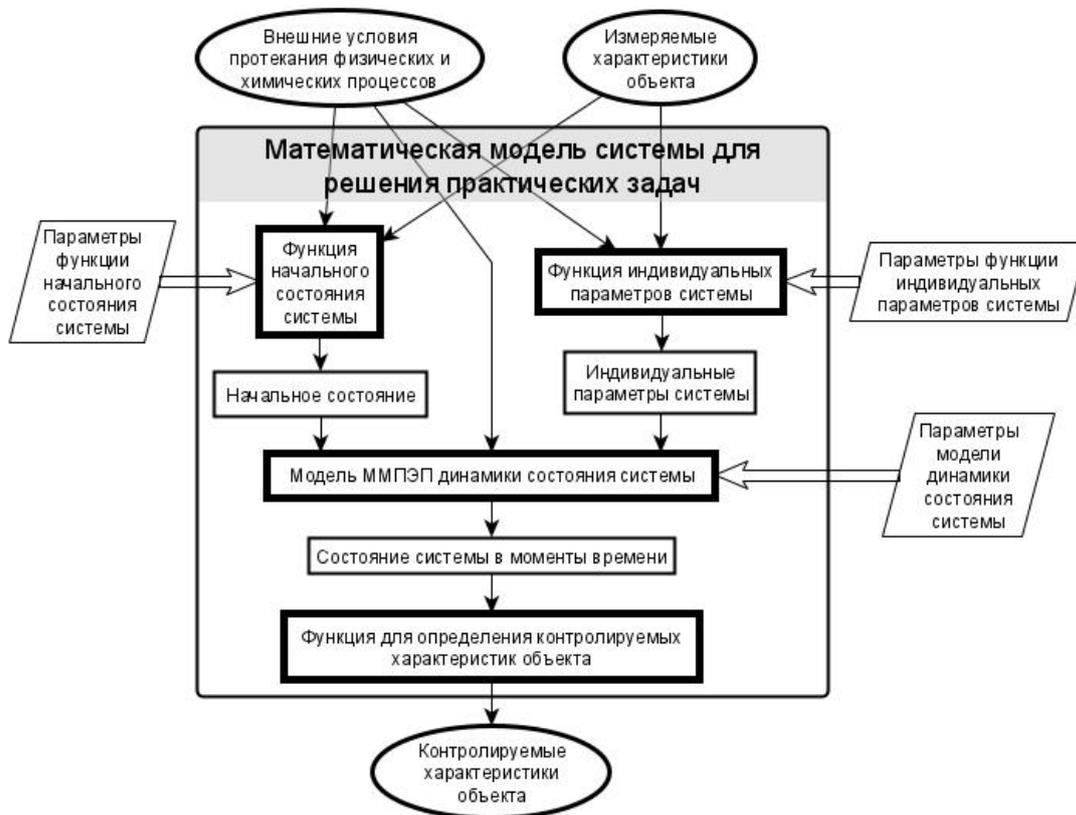


Рис. 2. Математическая модель системы для решения практических задач. В параллелограммах приведены обучаемые параметры модели

Следует отметить, что в КП системы могут войти и ее ИП в последующие моменты времени [8]. Структура модели объекта (рис. 2) строится на основе аналитического выражения динамики (вбирающего в себя физику рассматриваемого объекта) и вышеупомянутых функций [7, 8].

Аналитически общее решение системы дифференциальных уравнений ММПЭП представимо в виде [7]:

$$\Delta \mathbf{x}(t) = \Delta \tilde{\mathbf{x}}^* \left(\varepsilon(\varepsilon_0, t), \hat{\mathbf{u}}(t, \hat{\mathbf{u}}), \mathbf{q}_{\text{дисс}}, \tilde{\mathbf{s}}(t) \right), \quad \hat{\mathbf{u}} = \hat{\mathbf{u}}[U(t)], \quad (1)$$

$$\mathbf{x}(t) = \tilde{\mathbf{x}}^* \left(\Delta \mathbf{x}(t), \hat{\mathbf{u}}(t, \hat{\mathbf{u}}), \mathbf{q}_{\text{топ}}, \hat{\mathbf{s}}^* \left(\varepsilon(\varepsilon_0, t), \hat{\mathbf{u}}(t, \hat{\mathbf{u}}), \tilde{\mathbf{s}}(t), \mathbf{q}_{\text{упр}} \right) \right), \quad (2)$$

$$\tilde{\mathbf{s}}(t) = \hat{\mathbf{s}} \left(\varepsilon(\varepsilon_0, t), \hat{\mathbf{u}}(t, \hat{\mathbf{u}}), \mathbf{q}_{\text{упр}} \right), \quad \varepsilon_0 = \tilde{\mathbf{x}}^{*-1} \left(\mathbf{x}_0, \hat{\mathbf{u}}(0, \hat{\mathbf{u}}), \mathbf{q}_{\text{топ}}, \mathbf{q}_{\text{дисс}}, \mathbf{q}_{\text{упр}} \right), \quad (3)$$

$$\varepsilon(\varepsilon_0, 0) = \varepsilon_0, \quad \varepsilon(\varepsilon_0, t + \tau) = \varepsilon(\varepsilon(\varepsilon_0, \tau), t), \quad \dim(\varepsilon) = \dim(\mathbf{x}), \quad (4)$$

$$\forall \varepsilon_0, U, \mathbf{q}_{\text{дисс}}, \mathbf{s} \exists \lim_{t \rightarrow +\infty} \left(\Delta \tilde{\mathbf{x}}^* \left(\varepsilon(\varepsilon_0, t), U, \mathbf{q}_{\text{дисс}}, \mathbf{s} \right) \right) = \Delta \tilde{\mathbf{x}}^{**} \left(\varepsilon_0, U, \mathbf{q}_{\text{дисс}}, \mathbf{s} \right), \quad (5)$$

$$\text{из } \left(\frac{d\mathbf{x}(t)}{dt} \right)_{\text{ext}} \equiv 0 \text{ следует } \hat{\mathbf{s}}^* \left(\varepsilon, U, \mathbf{s}, \mathbf{q}_{\text{упр}} \right) \equiv \hat{\mathbf{s}}^{**} \left(U, \mathbf{s}, \mathbf{q}_{\text{упр}} \right), \quad (6)$$

$$\text{из } U(t) \equiv U = \text{const} \text{ следует } \hat{\mathbf{u}}(t, \hat{\mathbf{u}}) \equiv \hat{\mathbf{u}}(U), \quad (7)$$

где $\mathbf{x}(t)$ – динамика параметров состояния системы; $\Delta \mathbf{x}(t)$ – динамика величин, характеризующих независимые параметры состояния (не связанные между собой законами сохранения); $U(t)$ – динамика характеристик системы, не изменяющихся в результате протекания процессов внутри системы, а изменяющихся только в результате внешних воздействий на нее; $(d\mathbf{x}(t)/dt)_{\text{ext}}$ – внешние потоки.

Как видно из уравнений (1)–(7), функция $\tilde{\mathbf{x}}^* \left(\Delta \mathbf{x}, U, \mathbf{q}_{\text{топ}}, \hat{\mathbf{s}}^* \right)$ характеризует топологическую составляющую, функция $\Delta \tilde{\mathbf{x}}^* \left(\varepsilon, U, \mathbf{q}_{\text{дисс}}, \mathbf{s} \right)$ – диссипативную составляющую, а функции $\hat{\mathbf{s}}^* \left(\varepsilon, U, \mathbf{s}, \mathbf{q}_{\text{упр}} \right)$ и $\hat{\mathbf{s}} \left(\varepsilon, U, \mathbf{q}_{\text{упр}} \right)$ – управляющую составляющую системы [7, 9]. Для любого аналитического задания динамики системы в виде (1)–(7) найдется система уравнений ММПЭП, общим решением которого будет (1)–(7) [7]. Отсюда вытекает корректность задания аналитического выражения (1)–(7) общего решения уравнений ММПЭП. Квадратные скобки в выражении (1) означают взятие функционала динамики $U(t)$.

Измеряемые $\mathbf{y}(t)$ и контролируемые $\mathbf{z}(t)$ параметры системы [6, 7]:

$$\mathbf{y}(t) = \mathbf{Y}(t, \mathbf{c}_y, \mathbf{p}), \quad \mathbf{z}(t) = \mathbf{Z}(t, \mathbf{c}_z, \mathbf{p}), \quad \mathbf{c}_y = \mathbf{c}_y[\mathbf{x}(t), U(t)], \quad \mathbf{c}_z = \mathbf{c}_z[\mathbf{x}(t), U(t)], \quad (8)$$

где \mathbf{p} – индивидуальные параметры системы. Для задания аналитического выражения динамики рассматриваемой системы используется локальное упрощение уравнений ММПЭП [6, 7, 10] с последующим заданием соответствующих упрощенных аналитических выражений динамик $\mathbf{x}(t)$, из которых формируется полное общее решение дифференциальных уравнений ММПЭП в виде (1)–(7) [6, 7, 10]. На основе системы уравнений (1)–(8) строится искомая преобразованная модель системы [7].

Общее решение системы дифференциальных уравнений ММПЭП в виде (1)–(7) задается путем интерполяции опорных общих решений уравнений ММПЭП $\left\{ \mathbf{x}_i(\mathbf{x}_0, t) \right\}_{i=1}^{n_{\text{оп}}}$, полученных при соответствующих фиксированных значениях \mathbf{p}_i индивидуальных параметров \mathbf{p} , $\hat{\mathbf{u}}_i$ параметров $\hat{\mathbf{u}}$, \mathbf{a} параметров \mathbf{a} функциональных разложений для свойств веществ и процессов [7]:

$$\mathbf{x}(t) = \mathbf{F}_x \left(\left\{ \mathbf{x}_i \left(\hat{\varepsilon}_i(t), \hat{\mathbf{u}}_i(t) \right) \right\}_{i=1}^{n_{\text{он}}}, \hat{\mathbf{b}} \right), \quad \varepsilon_0 = \mathbf{F}_x^{-1}(\mathbf{x}_0, \hat{\mathbf{b}}), \quad \hat{\mathbf{b}} = \hat{\mathbf{b}}(\mathbf{b}, \hat{\mathbf{u}}, \mathbf{p}), \quad (9)$$

$$\hat{\varepsilon}_i(t) = \hat{\varepsilon}_i^* \left(\varepsilon(\varepsilon_0, t), \hat{\mathbf{b}} \right), \quad \hat{\mathbf{u}}_i(t) = \hat{\mathbf{u}}_i^* \left(\hat{\mathbf{u}}(t, \hat{\mathbf{u}}_i), \hat{\mathbf{b}} \right), \quad i = \overline{1, n_{\text{он}}}, \quad \mathbf{b} = \mathbf{b}(\mathbf{a}), \quad (10)$$

где опорные динамики $\left\{ \mathbf{x}_i(\mathbf{x}_0, t) \right\}_{i=1}^{n_{\text{он}}}$ аппроксимируются соответствующими функциями $\mathbf{x}_i(\hat{\varepsilon}_i(t), \hat{\mathbf{u}}_i(t))$, $i = 1, n_{\text{он}}$, заданными с учетом (10) в виде уравнений (1)–(7). Функции \mathbf{F}_x , $\hat{\mathbf{b}}$, \mathbf{b} представляют собой выбранные методы интерполяции [7]. Общее решение дифференциальных уравнений ММПЭП, заданное в виде (9), (10), удовлетворяет (1)–(7) [7], а значит, является корректным [7].

С учетом выражений (8)–(10) получим для динамики ИП $\mathbf{y}(t)$ и КП $\mathbf{z}(t)$ [7]:

$$\mathbf{y}(t) = \tilde{\mathbf{g}}_y(\mathbf{x}_0, t, \hat{\mathbf{b}}_y), \quad \hat{\mathbf{b}}_y = \hat{\mathbf{b}}_y(\hat{\mathbf{u}}, \tilde{\mathbf{b}}(\hat{\mathbf{p}}, \hat{\mathbf{u}}), \mathbf{b}), \quad \hat{\mathbf{p}} = \hat{\mathbf{p}}(\mathbf{a}, \mathbf{p}), \quad \mathbf{b} = \mathbf{b}(\mathbf{a}), \quad (11)$$

$$\mathbf{z}(t) = \tilde{\mathbf{g}}_z(\mathbf{x}_0, t, \hat{\mathbf{b}}_z), \quad \hat{\mathbf{b}}_z = \hat{\mathbf{b}}_z(\hat{\mathbf{u}}, \tilde{\mathbf{b}}(\hat{\mathbf{p}}, \hat{\mathbf{u}}), \mathbf{b}), \quad (12)$$

где параметры $\tilde{\mathbf{b}} = \tilde{\mathbf{b}}(\hat{\mathbf{p}}, \hat{\mathbf{u}})$, $\hat{\mathbf{p}} = \hat{\mathbf{p}}(\mathbf{a}, \mathbf{p})$ меняются независимо с изменением параметров \mathbf{a} и \mathbf{p} . Это дает возможность ввести функционалы $\tilde{\mathbf{g}}_{\mathbf{x}_0}$ и $\tilde{\mathbf{g}}_{\tilde{\mathbf{b}}}$ для \mathbf{x}_0 и $\tilde{\mathbf{b}}$ соответственно [7]:

$$\mathbf{x}_0 = \tilde{\mathbf{g}}_{\mathbf{x}_0}(\hat{\mathbf{u}}, \tilde{\mathbf{b}}_{\mathbf{x}_0}), \quad \tilde{\mathbf{b}}_{\mathbf{x}_0} = \tilde{\mathbf{g}}_{\tilde{\mathbf{b}}_{\mathbf{x}_0}}[\mathbf{y}_r(t)], \quad \tilde{\mathbf{b}} = \tilde{\mathbf{g}}_{\tilde{\mathbf{b}}}(\hat{\mathbf{u}}, \tilde{\mathbf{b}}_r), \quad \tilde{\mathbf{b}}_r = \tilde{\mathbf{g}}_{\tilde{\mathbf{b}}_r}[\mathbf{y}_r(t)], \quad (13)$$

$$\mathbf{y}_r(t) \equiv \tilde{\mathbf{g}}_y \left(\tilde{\mathbf{g}}_{\mathbf{x}_0}[\mathbf{y}_r(t)], t, \hat{\mathbf{b}}_y \left(\hat{\mathbf{u}}, \tilde{\mathbf{g}}_{\tilde{\mathbf{b}}}[\mathbf{y}_r(t)], \mathbf{b} \right) \right),$$

где $\mathbf{y}_r(t)$ – репрезентативный участок динамики ИП $\mathbf{y}(t)$. Определив согласно уравнению (13) параметры \mathbf{x}_0 и $\tilde{\mathbf{b}}$, из известных из экспериментальных данных репрезентативных участков $\mathbf{y}_r(t)$ динамик ИП $\mathbf{y}(t)$, мы в соответствии с уравнениями (11), (12) определяем индивидуальные параметры $\hat{\mathbf{p}}$ и параметры функциональных разложений \mathbf{b} [7]. Затем в силу выражений (11)–(13) мы определим КП [7].

Отсюда на основе выражений (11)–(13) может быть задана функция, определяющая по репрезентативным участкам контрольных динамик ИП $(\mathbf{y}_{r,i}^{\text{кон}}(t), \hat{\mathbf{u}}_i)$, $i = 1, N_{\text{кон}}$ индивидуальные параметры системы $\hat{\mathbf{p}}$ ($N_{\text{кон}}$ – число контрольных динамик рассматриваемой системы) [7, 8]:

$$\hat{\mathbf{p}} = \tilde{\mathbf{g}}_{\hat{\mathbf{p}}} \left(\left\{ \tilde{\mathbf{b}}_i^{\text{кон}}, \hat{\mathbf{u}}_i \right\}_{i=1}^{N_{\text{кон}}}, \mathbf{b} \right), \quad \tilde{\mathbf{b}}_i^{\text{кон}} = \tilde{\mathbf{g}}_{\tilde{\mathbf{b}}_i^{\text{кон}}}(\hat{\mathbf{u}}_i, \tilde{\mathbf{b}}_{r,i}^{\text{кон}}), \quad \tilde{\mathbf{b}}_{r,i}^{\text{кон}} = \tilde{\mathbf{g}}_{\tilde{\mathbf{b}}_{r,i}^{\text{кон}}}[\mathbf{y}_{r,i}^{\text{кон}}(t)], \quad i = 1, N_{\text{кон}}. \quad (14)$$

Выражения (9)–(14) представляют собой задание аналитического выражения динамики состояния системы, позволяющее получить динамику в зависимости от параметров контрольных динамик $\tilde{\mathbf{b}}_i^{\text{кон}} = \tilde{\mathbf{g}}_{\tilde{\mathbf{b}}_i^{\text{кон}}}(\hat{\mathbf{u}}_i, \tilde{\mathbf{b}}_{r,i}^{\text{кон}})$, $\tilde{\mathbf{b}}_{r,i}^{\text{кон}} = \tilde{\mathbf{g}}_{\tilde{\mathbf{b}}_{r,i}^{\text{кон}}}[\mathbf{y}_{r,i}^{\text{кон}}(t)]$, $i = 1, N_{\text{кон}}$. При этом функция $\tilde{\mathbf{b}} = \tilde{\mathbf{b}} \left(\left\{ \tilde{\mathbf{b}}_i^{\text{кон}}, \hat{\mathbf{u}}_i \right\}_{i=1}^{N_{\text{кон}}}, \hat{\mathbf{u}}, \mathbf{b} \right)$ в правой части (11), (12), как и функция $\mathbf{x}_0 = \tilde{\mathbf{g}}_{\mathbf{x}_0}(\hat{\mathbf{u}}, \tilde{\mathbf{b}}_{\mathbf{x}_0})$, $\tilde{\mathbf{b}}_{\mathbf{x}_0} = \tilde{\mathbf{g}}_{\tilde{\mathbf{b}}_{\mathbf{x}_0}}[\mathbf{y}_r(t)]$, также может представлять собой выбранный метод интерполяции (видно из работы [7]). С учетом выражений (11)–(14) получим выражения для динамик ИП $\mathbf{y}(t)$ и КП $\mathbf{z}(t)$, аналогичные (11), (12):

$$\mathbf{y}(t) = \tilde{\mathbf{g}}_y(\mathbf{x}_0, t, \hat{\mathbf{b}}_y), \quad \hat{\mathbf{b}}_y = \hat{\mathbf{b}}_y \left(\hat{\mathbf{u}}, \tilde{\mathbf{b}} \left(\tilde{\mathbf{g}}_{\hat{\mathbf{p}}} \left(\left\{ \tilde{\mathbf{b}}_i^{\text{кон}}, \hat{\mathbf{u}}_i \right\}_{i=1}^{N_{\text{кон}}}, \mathbf{b} \right), \hat{\mathbf{u}}, \mathbf{b} \right) \right), \quad (15)$$

$$\mathbf{z}(t) = \tilde{\mathbf{g}}_z(\mathbf{x}_0, t, \hat{\mathbf{b}}_z), \hat{\mathbf{b}}_z = \hat{\mathbf{b}}_z \left(\hat{\mathbf{u}}, \tilde{\mathbf{b}} \left(\tilde{\mathbf{g}}_p \left(\left\{ \tilde{\mathbf{b}}_i^{\text{кон}}, \hat{\mathbf{u}}_i \right\}_{i=1}^{N_{\text{кон}}}, \mathbf{b} \right), \hat{\mathbf{u}}, \tilde{\mathbf{b}} \right) \right). \quad (16)$$

Входящие в выражения (15), (16) параметры $\tilde{\mathbf{b}} = \tilde{\mathbf{b}}(\mathbf{a})$, $\tilde{\mathbf{b}} = \tilde{\mathbf{b}}(\mathbf{a})$ определяются из статистики ИП $\mathbf{y}_{i,j}^{\text{исп}}(t)$, $\hat{\mathbf{u}}_{i,j}$, $i=1, N_{\text{исп},j}$, $j=1, n_{\text{экз}}$ испытаний различных экземпляров одного бренда ($n_{\text{экз}}$ – число испытываемых экземпляров лабораторной системы; $N_{\text{исп},j}$, $j=1, n_{\text{экз}}$ – число испытаний каждого j -го испытываемого экземпляра), используя (15), (16), в соответствии с [7, 8]:

$$\tilde{\mathbf{b}} = \tilde{\mathbf{b}} \left(\left\{ \left\{ \tilde{\mathbf{b}}_{i,j}^{\text{исп}}, \hat{\mathbf{u}}_{i,j} \right\}_{i=1}^{N_{\text{исп},j}} \right\}_{j=1}^{n_{\text{экз}}} \right), \tilde{\mathbf{b}}_{i,j}^{\text{исп}} = \tilde{\mathbf{g}}_{\tilde{\mathbf{b}},j}^{\text{исп}}(\hat{\mathbf{u}}_{i,j}, \tilde{\mathbf{b}}_{r,i,j}^{\text{исп}}), \tilde{\mathbf{b}}_{r,i,j}^{\text{исп}} = \tilde{\mathbf{g}}_{\tilde{\mathbf{b}},j}^{\text{исп}}[\mathbf{y}_{r,i,j}^{\text{исп}}(t)],$$

$$i = 1, N_{\text{исп},j}, j = 1, n_{\text{экз}}. \quad (17)$$

Правая часть полученного выражения (17) может также представлять собой выбранный метод интерполяции [7].

Выражения (15)–(17) представляют собой ММ системы, которая может быть положена в основу МЯ цифровых двойников [7, 8]. Такая ММ получается из статистики испытаний экземпляров системы рассматриваемого бренда, а также из статистики испытаний лабораторных систем. Нетрудно видеть, что для выбора интерполяционных методов, входящих в (15)–(17) для $\tilde{\mathbf{b}} = \tilde{\mathbf{b}} \left(\left\{ \tilde{\mathbf{b}}_i^{\text{кон}}, \hat{\mathbf{u}}_i \right\}_{i=1}^{N_{\text{кон}}}, \hat{\mathbf{u}}, \mathbf{b} \right)$, $\tilde{\mathbf{b}}$, $\tilde{\mathbf{g}}_{\mathbf{x}_0}$ соответственно, необходимо, выполнив моделирование динамик в системе для совокупностей случайно заданных значений \mathbf{x}_0 , \mathbf{p} и \mathbf{a} , получить качественный характер динамик, а также качественный характер связи между параметрами полученных динамик. На основе упомянутого качественного анализа смоделированных динамик системы выбираются опорные динамики, входящие в (9), (10), и выражения для их задания, удовлетворяющие (1)–(7) [7], а также опорные значения параметров динамик системы и выбираются интерполяционные методы $\tilde{\mathbf{b}}$, $\tilde{\mathbf{b}}$, $\tilde{\mathbf{g}}_{\mathbf{x}_0}$.

Использование интерполяционных методов (т.е. методов, непосредственно базирующихся на реперных значениях [9]) обусловлено тем, что для построения преобразованной ММ в общем случае необходимо локально упрощать систему уравнений ММПЭП [6, 7, 11]. Коэффициенты локально упрощенной системы уравнений ММПЭП, из которых определяются коэффициенты исходных уравнений ММПЭП, определяются значениями ИП в соответствующей области локального упрощения [6, 7, 11] – идеология методов интерполяции [9]. В силу нелинейности реальных систем полученных областей локальных упрощений уравнений ММПЭП достаточно много [6, 7], что также обуславливает применение методов интерполяции для построения ММ систем, положенных в основу МЯ их цифровых двойников [9].

Функционал $\tilde{\mathbf{g}}_{\tilde{\mathbf{b}}}$ может быть построен с помощью различных методов идентификации [12]. Аналогично функционалу $\tilde{\mathbf{g}}_{\tilde{\mathbf{b}}}$ методами [12] могут быть введены функционалы $\hat{\mathbf{b}}_y = \tilde{\mathbf{g}}_{\hat{\mathbf{b}}_y}[\mathbf{y}(t)]$ и $\hat{\mathbf{b}}_z = \tilde{\mathbf{g}}_{\hat{\mathbf{b}}_z}[\mathbf{z}(t)]$. С помощью введенных функционалов $\tilde{\mathbf{g}}_{\hat{\mathbf{b}}_y}$ и $\tilde{\mathbf{g}}_{\hat{\mathbf{b}}_z}$ определяются из результатов вычислительного эксперимента соответствующие параметры $\hat{\mathbf{b}}_y$ и $\hat{\mathbf{b}}_z$, входящие в выражения (15) и (16) соответственно. Из качественного анализа полученных из результатов вычислительного эксперимента параметров $\hat{\mathbf{b}}_y$ и $\hat{\mathbf{b}}_z$ и задаются функции для параметров $\hat{\mathbf{b}}_y$ и $\hat{\mathbf{b}}_z$ и независимые параметры $\tilde{\mathbf{b}}$.

Построение моделей для математического ядра цифровых двойников

Итак, полученные ММ из уравнений ММПЭП, являющиеся составляющими МЯ цифровых двойников, могут быть двух видов:

- ММ систем, которые строятся из статистики испытаний систем рассматриваемого бренда;
- ММ систем, которые строятся на основе статистики испытаний лабораторных систем.

Рассмотрим первый случай (случай построения ММ систем, которые строятся из статистики испытаний систем рассматриваемого бренда). В рассматриваемом случае на основе уравнения (15) получаем из ИП $y(t)$ параметры \hat{b}_y , а затем из функции для \hat{b}_y получаем параметры \tilde{b} и \bar{b} для каждой динамики ИП $y(t)$. Затем в силу одинаковости для всех экземпляров рассматриваемой системы параметров \bar{b} берем в качестве значения параметров \tilde{b} их среднее арифметическое значений, полученных вышеописанным путем для каждой динамики ИП $y(t)$. Затем получаем на основе выражения (14) контрольные параметры $\tilde{b}_i^{\text{кон}}$, $i=1, N_{\text{кон}}$, соответствующие параметрам внешних воздействий \hat{u}_i , $i=1, N_{\text{кон}}$, и задаем интерполяционную зависимость для \tilde{b} (исходя из (17)), входящую в выражения (15), (16):

$$\tilde{b} = \tilde{b} \left(\left\{ \tilde{b}_k^{\text{кон}}, \hat{u}_k \right\}_{k=1}^{N_{\text{кон}}}, \hat{u}, \left\{ \left\{ \tilde{b}_{i,j}^{\text{реп}}, \hat{u}_{i,j} \right\}_{i=1}^{N_{\text{реп},j}} \right\}_{j=1}^{n_{\text{экз}}} \right), \quad (18)$$

где аналогично (17):

$$\tilde{b}_{i,j}^{\text{реп}} = \bar{g}_{b,i,j}^{\text{реп}}(\hat{u}_{i,j}, \tilde{b}_{r,i,j}^{\text{реп}}), \tilde{b}_{r,i,j}^{\text{реп}} = \bar{g}_{b,r,i,j}^{\text{реп}}[y_{r,i,j}^{\text{реп}}(t)], \quad i=1, N_{\text{реп},j}, \quad j=1, n_{\text{экз}}, \quad (19)$$

используя значения контрольных параметров $\tilde{b}_{r,i,j}^{\text{реп}}$, $\hat{u}_{i,j}$, $i=1, N_{\text{реп},j}$, $j=1, n_{\text{экз}}$ репрезентативных участков динамик $y_{r,i,j}^{\text{реп}}(t)$, $i=1, N_{\text{реп},j}$, $j=1, n_{\text{экз}}$, репрезентативных экземпляров рассматриваемой системы. Также параметры $\tilde{b}_{i,j}^{\text{реп}}$, $i=1, N_{\text{реп},j}$, $j=1, n_{\text{экз}}$ могут либо явно входить в уравнение (18), либо явно в уравнение (18) могут входить параметры \tilde{b} , которые определяются через $\tilde{b}_{r,i,j}^{\text{реп}}$, $\hat{u}_{i,j}$, $i=1, N_{\text{реп},j}$, $j=1, n_{\text{экз}}$, используя методы теории идентификации [12]. Второй подход позволяет сглаживать случайные погрешности экспериментальных данных [12, 13]. Общая структура ММ, положенной в основу МЯ цифрового двойника, показана на рис. 3.

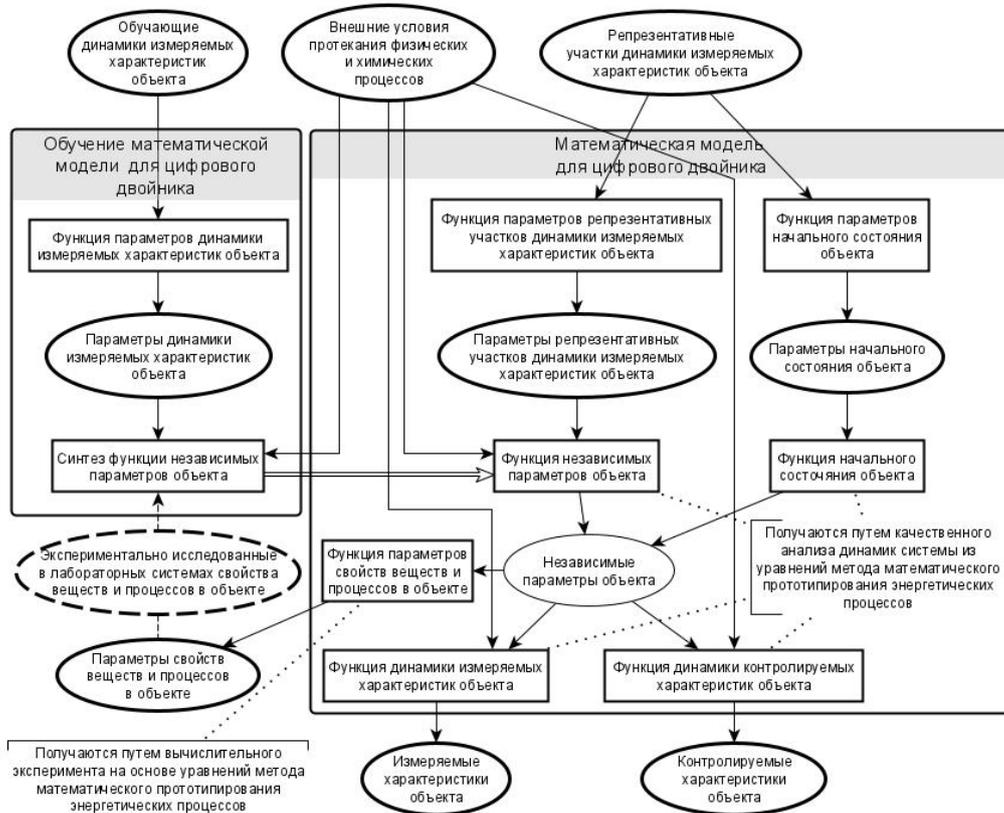


Рис. 3. Модель системы, получаемая на основе статистики испытаний экземпляров рассматриваемой системы с учетом испытаний лабораторных систем

Как нетрудно видеть из выражений (14)–(16), из полученных значений независимых параметров $\tilde{\mathbf{b}}$ можно определять параметры свойств веществ и процессов (СВП) в системе [7]. Для этого необходимо ввести функцию, определяющую параметры СВП в системе (рис. 3), которая получается также методами интерполяции с использованием вычислительного эксперимента на основе системы уравнений ММПЭП [7]. С использованием упомянутой функции, построив вышеописанным образом ММ для лабораторных систем, мы получим параметры СВП в лабораторной системе. Однако часть этих свойств могут быть справедливыми и для рассматриваемой системы, что делает целесообразным их учет в интерполяционном выражении (18) (рис. 3) [7]. Так реализуется схема построения ММ на основе статистики испытаний лабораторных систем, структура которой также показана на рис. 3.

Функции $\bar{\mathbf{g}}_{\mathbf{b}}^{\text{кон}}$, входящие в уравнение (14), а также функции $\bar{\mathbf{g}}_{\mathbf{b},j}^{\text{реп}}$, $j = 1, n_{\text{экз}}$, входящие в уравнение (19), могут быть заданы интерполяционными методами, а также с использованием методов наименьших квадратов на основе функций динамик ИП (15) [12]. При этом с использованием интерполяционных методов могут быть заданы приближенные реализации упомянутых функций, результаты вычислений которых доводятся методом наименьших квадратов [12]. Аналогичное касается и вышеупомянутых функций $\bar{\mathbf{g}}_{\mathbf{b}_y}$ и $\bar{\mathbf{g}}_{\mathbf{b}_z}$.

Результаты

Класс ММ объектов, зашиваемых в основу МЯ цифровых их двойников, структура которых показана на рис. 3, представляет собой развитие статистических методов построения ММ [13]. Упомянутое развитие заключается в том, что класс ММ, в котором ищется ММ [12, 13], заложена реальная физика функционирования объекта, что гарантирует корректность построенной ММ [12, 13]. Таким образом, предложенный подход построения ММ систем вбирает в себя как преимущества подхода черного ящика (простоту), так и преимущества подхода, основанного на моделировании физических и химических процессов в объекте (экстраполируемость модели на режимы, для которых имеется малое число экспериментальных данных) [13].

Благодаря использованию методов интерполяции предложенная схема построения моделей может быть положена в основу алгоритмов машинного обучения, в частности нейронных и гибридных сетей [7]. Пример построения ММ литий-ионного аккумулятора US18650VTC6 Series в соответствии с предложенной схемой приведен в работе [14].

Обсуждение

Итак, задача построения ММ системы, положенной в основу МЯ цифровых двойников объектов различной физической и химической природы, разбивается на два класса подзадач:

1. Качественный анализ на основе уравнений ММПЭП динамик рассматриваемого объекта и перекрестных связей между параметрами динамик объекта в различных режимах его функционирования и параметрами внешних воздействий на него.
2. Построение ММ системы в сформированном на основе качественного анализа ММПЭП системы классе ММ.

Таким образом, построение искомой ММ в соответствии с предложенной в настоящей работе схемой представляет собой использование известных методов [1–5, 9, 12, 13], примененных к классу ММ, построенного путем качественного анализа на основе ММПЭП рассматриваемого объекта.

Заключение

Качественный анализ динамики процессов в объектах различной физической и химической природы на основе ММПЭП является важнейшим способом задания класса ММ, в котором строятся ММ объекта известными методами [1–5, 9, 12, 13]. Отсюда вытекает важнейшая задача планирования вычислительного эксперимента на основе ММПЭП с последующим качественным анализом полученных ММПЭП динамик рассматриваемого объекта.

Список литературы

1. Васильев А. Н., Тархов Д. А., Малыхина Г. Ф. Методы создания цифровых двойников на основе нейросетевого моделирования // Современные информационные технологии и ИТ-образование. 2018. Т. 4, № 3. С. 521–532.

2. Кокорев Д. С., Юрин А. А. Цифровые двойники: понятие, типы и преимущества для бизнеса // *Colloquium-journal*. 2019. № 10 (34). С. 31–35.
3. Гончаров А. С., Саклаков В. М. Цифровой двойник: обзор существующих решений и перспективы развития технологии // *Материалы Всероссийской научно-практической конференции «Информационно-телекоммуникационные системы и технологии (ИТСИТ-2018)»* (г. Кемерово, 11–13 октября 2018 г.). Кемерово, 2018. С. 24–26.
4. Курганова Н. В., Филин М. А., Черняев Д. С. [и др.]. Внедрение цифровых двойников как одно из ключевых направлений цифровизации производства // *International Journal of Open Information Technologies*. 2019. Vol. 7, № 5. С. 105–115.
5. Петров А. В. Имитация как основа технологии цифровых двойников // *Вестник ИрГТУ*. 2018. Т. 22, № 10. С. 56–66. doi: 10.21285/1814-3520-2018-10-56-66
6. Khalyutin S. P., Starostin I. E., Agafonkina I. V. Generalized Method of Mathematical Prototyping of Energy Processes for Digital Twins Development // *Energies*. 2023. Vol. 16. P. 1933–1958. doi: 10.3390/en16041933
7. Старостин И. Е. Построение на основе интерполяции моделей различных физических и химических систем методом математического прототипирования энергетических процессов // *Надежность и качество сложных систем*. 2024. № 1 (45). С. 49–59. doi: 10.21685/2307-4205-2024-1-6
8. Старостин И. Е., Дружинин А. А., Гавриленков С. И. Использование машинного обучения с учителем для построения математических моделей систем методом математического прототипирования энергетических процессов // *Труды Международного симпозиума Надежность и качество*. 2023. Т. 1. С. 66–72.
9. Калиткин Н. Н. Численные методы. СПб. : БХВ-Петербург, 2011. 592 с.
10. Шапиро С. В. Основы синергетики. Уфа : УГАЭС, 2012. 200 с.
11. Ланцов В. Н. Методы понижения порядка моделей сложных систем. Владимир : Изд-во ВлГУ, 2017. 84 с.
12. Eukhoff P. *Systems identification: parameters and state estimation*. Eindhoven, Netherlands : University of technology, 1975. 680 p.
13. Антонов А. В. Системный анализ. М. : Высш. шк., 2004. 454 с.
14. Starostin I. E., Khalyutin S. P., Davidov A. O. [et al.]. Obtaining a model for the voltage and temperature of the US18650VTC6 Series lithium-ion battery in constant current discharge mode from the analysis of physical and chemical processes in the accumulator // 18th technical scientific conference on aviation dedicated to the memory of N.E. Zhukovsky, TSCZH 2021. 2021. P. 109–117.

References

1. Vasil'ev A.N., Tarkhov D.A., Malykhina G.F. Methods of creating digital twins based on neural network modeling. *Sovremennye informatsionnye tekhnologii i IT-obrazovanie = Modern information technologies and IT education*. 2018;4(3):521–532. (In Russ.)
2. Kokorev D.S., Yurin A.A. Digital twins: concept, types and advantages for business. *Colloquium-journal*. 2019;(10):31–35. (In Russ.)
3. Goncharov A.S., Saklakov V.M. Digital twin: an overview of existing solutions and prospects for technology development. *Materialy Vserossiyskoy nauchno-prakticheskoy konferentsii «Informatsionno-telekommunikatsionnye sistemy i tekhnologii (ITSIT-2018)»* (g. Kemerovo, 11–13 oktyabrya 2018 g.) = *Materials of the All-Russian scientific and practical conference "Information and telecommunication systems and technologies (ITSIT-2018)"* (Kemerovo, October 11-13, 2018). Kemerovo, 2018:24–26. (In Russ.)
4. Kurganova N.V., Filin M.A., Chernyaev D.S. et al. The introduction of digital twins as one of the key directions of digitalization of production. *International Journal of Open Information Technologies*. 2019;7(5):105–115. (In Russ.)
5. Petrov A.V. Imitation as the basis of digital twins technology. *Vestnik IrGTU = Bulletin of IrSTU*. 2018;22(10):56–66. (In Russ.). doi: 10.21285/1814-3520-2018-10-56-66
6. Khalyutin S.P., Starostin I.E., Agafonkina I.V. Generalized Method of Mathematical Prototyping of Energy Processes for Digital Twins Development. *Energies*. 2023;16:1933–1958. doi: 10.3390/en16041933
7. Starostin I.E. Construction on the basis of interpolation of models of various physical and chemical systems by the method of mathematical prototyping of energy processes. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems*. 2024;(1):49–59. (In Russ.). doi: 10.21685/2307-4205-2024-1-6
8. Starostin I.E., Druzhinin A.A., Gavrilencov S.I. Using machine learning with a teacher to build mathematical models of systems by mathematical prototyping of energy processes. *Trudy Mezhdunarodnogo simpoziuma Nadezhnost' i kachestvo = Proceedings of the International Symposium Reliability and Quality*. 2023;1:66–72. (In Russ.)
9. Kalitkin N.N. *Chislennyye metody = Numerical methods*. Saint Petersburg: BKhV-Peterburg, 2011:592. (In Russ.)
10. Shapiro S.V. *Osnovy sinergetiki = Fundamentals of synergetics*. Ufa: UGAES, 2012:200. (In Russ.)
11. Lantsov V.N. *Metody ponizheniya poryadka modeley slozhnykh system = Methods of lowering the order of models of complex systems*. Vladimir: Izd-vo VIGU, 2017:84. (In Russ.)

12. Eykhoff P. *Systems identification: parameters and state estimation*. Eindhoven, Netherlands: University of technology, 1975:680.
13. Antonov A.V. *Sistemnyy analiz = System analysis*. Moscow: Vyssh. shk., 2004:454. (In Russ.)
14. Starostin I.E., Khalyutin S.P., Davidov A.O. et al. Obtaining a model for the voltage and temperature of the US18650VTC6 Series lithiumion battery in constant current discharge mode from the analysis of physical and chemical processes in the accumulator. *18th technical scientific conference on aviation dedicated to the memory of N.E. Zhukovsky, TSCZH 2021*. 2021:109–117.

Информация об авторах / Information about the authors

Игорь Евгеньевич Старостин

доктор технических наук, профессор,
профессор кафедры электротехники
и авиационного электрооборудования,
Московский государственный технический
университет гражданской авиации
(Россия, г. Москва, Кронштадтский бульвар, 20)
E-mail: starostinigo@yandex.ru

Igor E. Starostin

Doctor of technical sciences, professor, professor
of the sub-department of electrical engineering
and aviation electrical equipment,
Moscow State Technical University of Civil Aviation
(20 Kronshtadtskiy boulevard, Moscow, Russia)

Автор заявляет об отсутствии конфликта интересов /

The author declares no conflicts of interests.

Поступила в редакцию/Received 15.05.2024

Поступила после рецензирования/Revised 15.07.2024

Принята к публикации/Accepted 25.08.2024

ПРИМЕНЕНИЕ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ КЛАССИФИКАЦИИ ОПУХОЛЕЙ ГОЛОВНОГО МОЗГА

М. О. Тимошкин¹, Е. Г. Романова²

¹ Московский государственный университет имени М. В. Ломоносова, Москва, Россия

² Пензенский государственный университет, Пенза, Россия

¹ max.timoshkin@inbox.ru, ² romanova.elenar2016@yandex.ru

Аннотация. *Актуальность и цели.* Глубокое обучение – это активно развивающаяся область машинного обучения, которая позволяет находить зависимости в слабоструктурированных данных. Актуальность работы заключается в том, что в настоящее время автоматическая классификация типов тканей играет важную роль в компьютерной диагностике. Цель работы заключается в автоматизации процесса определения вида опухоли головного мозга по ее изображению с помощью методов глубокого обучения. *Материалы и методы.* В работе применяются методы глубокого обучения для автоматической классификации вида опухоли головного мозга по ее изображению. *Результаты и выводы.* Было произведено сравнение нейросетевых моделей при различных архитектурах, в различных режимах обучения и при наличии и отсутствии обогащения выборки. В процессе обучения с помощью лучшей архитектуры удалось достичь качества 96 % на валидационной выборке.

Ключевые слова: классификация опухолей головного мозга, глубокое обучение, архитектуры нейронных сетей, обогащение выборки, МРТ-изображения, режимы обучения

Для цитирования: Тимошкин М. О., Романова Е. Г. Применение нейросетевых алгоритмов глубокого обучения для классификации опухолей головного мозга // Надежность и качество сложных систем. 2024. № 3. С. 51–65. doi: 10.21685/2307-4205-2024-3-6

APPLICATION OF DEEP LEARNING NEURAL NETWORK ALGORITHMS FOR BRAIN TUMOR CLASSIFICATION

M.O. Timoshkin¹, E.G. Romanova²

¹ Moscow State University named after M.V. Lomonosov, Moscow, Russia

² Penza State University, Penza, Russia

¹ max.timoshkin@inbox.ru, ² romanova.elenar2016@yandex.ru

Abstract. *Background.* Deep learning is a rapidly developing area of machine learning that allows you to find dependencies in semi-structured data. The relevance of the work lies in the fact that currently automatic classification of tissue types plays an important role in computer diagnostics. The goal of the work is to automate the process of determining the type of brain tumor from its image using deep learning methods. *Materials and methods.* The work uses deep learning methods to automatically classify the type of brain tumor based on its image. *Results and conclusions.* A comparison was made of neural network models with different architectures, in different training modes, and with and without sample enrichment. During the training process, using the best architecture, it was possible to achieve a quality of 96 % on the validation set.

Keywords: brain tumor classification, deep learning, neural network architectures, sample enrichment, MRI images, learning modes

For citation: Timoshkin M.O., Romanova E.G. Application of deep learning neural network algorithms for brain tumor classification. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):51–65. (In Russ.). doi: 10.21685/2307-4205-2024-3-6

Введение

Существуют медицинские компании, занимающиеся лечением опухолей головного мозга. Автоматическая классификация типов тканей играет важную роль в компьютерной диагностике. Существуют различные виды опухолей: менингиома, глиома, опухоли гипофиза и др. Дано большое количество изображений трех видов опухолей головного мозга. Требуется автоматизировать

процесс определения вида опухоли, т.е. по изображению опухоли определить ее вид. Автоматизированный процесс определения вида опухоли должен быть как можно более точнее, так как в случае неверно определенного вида опухоли пациенту будет предоставлено несоответствующее медицинское лечение, которое потребует излишних финансовых затрат и может привести даже к смерти пациента.

Описание данных

Исходные данные представляют собой набор Т1-взвешенных МРТ-изображений с контрастным усилением, содержащих образцы опухолей головного мозга: менингиомы (1), глиомы (2) и опухоли гипофиза (3) [1].

Набор данных опухолей головного мозга содержит 3064 Т1-взвешенных изображения с контрастным усилением от 233 пациентов с тремя видами опухолей головного мозга:

- менингиома (708 изображений-срезов);
- глиома (1426 изображений-срезов);
- опухоль гипофиза (930 изображений-срезов).

Генеральной совокупностью в данном случае будет являться совокупность всевозможных изображений, содержащих один из трех видов опухоли головного мозга.

Объектом является изображение-срез одного (из трех возможных) вида опухоли головного мозга. Совокупность всех объектов образует выборочную совокупность.

Выборка данных была получена после прохождения МРТ и получения различных Т1-взвешенных изображений с контрастным усилением 233 пациентами с опухолью головного мозга. Данное множество изображений (выборка) является подмножеством всевозможных изображений, содержащих один из трех видов опухоли головного мозга (генеральная совокупность).

Можно считать, что данные пациенты были выбраны случайным образом из всего множества больных с опухолью головного мозга. Следовательно, можно считать, что выборка была получена методом бесповторного отбора данных из генеральной совокупности. Кроме того, были соблюдены принципы «равновозможности» и «случайности» попадания единиц в выборочную совокупность. Следовательно, обеспечена репрезентативность собранных данных относительно генеральной совокупности.

В выборку входят изображения-срезы только 233 пациентов с известным видом опухоли. Планируется поступление новых данных – МРТ Т1-взвешенных изображений-срезов опухоли головного мозга для новых пациентов (из генеральной совокупности), для которых вид опухоли необходимо определить (классифицировать) по изображению. Можно считать, что новые пациенты также будут поступать случайным образом из генеральной совокупности, т.е. методом бесповторного отбора из генеральной совокупности. Следовательно, репрезентативность новых данных относительно генеральной совокупности будет обеспечена по тем же причинам, что и для собранных данных.

Цель исследователей, которые осуществляли сбор данных, заключается в разработке подхода, который может автоматически классифицировать опухоли головного мозга по изображению ее среза [2].

Был выполнен анализ пропусков и анализ правдоподобности. Исходные данные не содержат пропусков, все изображения являются правдоподобными, что было выявлено путем их просмотра.

Каждое изображение в выборочной совокупности представляет собой двумерный срез. Набор данных изображений-срезов головного мозга с Т1-взвешиванием был получен в больнице Наньфан, Гуанчжоу, Китай и больнице общего профиля Тяньцзинского медицинского университета, Китай, с 2005 по 2010 г. Было собрано 3064 среза от 233 пациентов, содержащих 708 менингиом, 1426 глиом и 930 опухолей гипофиза. Изображения имеют плоскостное разрешение 512×512 с размером пикселя $0,49 \times 0,49$ мм². Толщина среза – 6 мм, зазор между срезами – 1 мм. Граница опухоли была очерчена вручную тремя опытными радиологами [2].

Таким образом, каждый объект представляется растровым изображением (трехмерным тензором) фиксированного размера. Целевой признак (вид опухоли) является категориальным признаком с тремя различными возможными значениями: менингиома (1), глиома (2), опухоль гипофиза (3).

Набор данных изображений опухолей головного мозга, доступный по ссылке [1], содержит три папки «1», «2», «3», соответствующие изображениям различных классов опухолей: менингиомы, глиомы и опухоли гипофиза соответственно.

Формальная задача

Дано: множество МРТ-изображений опухолей головного мозга и соответствующих видов опухоли $\{(x_i, y_i)\}_{i=1}^l$, где x_i – трехмерный тензор размерности $[512, 512, 4]$, $y_i \in Y$, $Y = \{0, 1, 2\}$.

Найти: стратегию, которая классифицирует любой объект из генеральной совокупности $x \in X$, т.е. определяет целевую переменную $y \in Y$.

Показатель качества: точность классификации

$$Accuracy = \frac{P}{N},$$

где P – количество верно классифицированных объектов; N – общее количество элементов в выборке.

Все части формальной постановки задачи полностью соответствуют содержательной постановке задачи:

– множество изображений опухолей головного мозга определено как $\{x_i\}_{i=1}^l$;

– множество видов опухолей: (менингиома, глиома, опухоль гипофиза) определено как $= \{0, 1, 2\}$.

Показатель качества был выбран таким, так как именно он отражает качество классификации, т.е. долю верно классифицированных видов опухоли головного мозга. Большое качество классификатора соответствует большому количеству людей, которым вид опухоли диагностирован верно и, следовательно, большей части больных может быть предоставлено соответствующее медицинское лечение.

Разбиение выборки

Для сохранения репрезентативности были соблюдены принципы «равновозможности» и «случайности» попадания единиц в выборочную совокупность; был выбран метод бесповторного отбора (табл. 1).

Таблица 1

Разбиение выборки на обучающую, валидационную и контрольную

Выборка	Абсолютные размеры выборки	Относительные размеры выборки
Исходная	3064	100 %
Обучающая	1839	60 %
Валидационная	612	20 %
Контрольная	613	20 %

Распознавание

Архитектура распознавания для одного объекта (рис. 1) включает шаги:

1. переход от слабоструктурированного описания к первичному структурированному представлению;
2. переход от первичного структурированного представления к основному структурированному представлению;
3. получение выхода для объекта. (Оценка, неявно преобразуемая в апостериорное распределение);
4. принятие окончательного бизнес-решения.

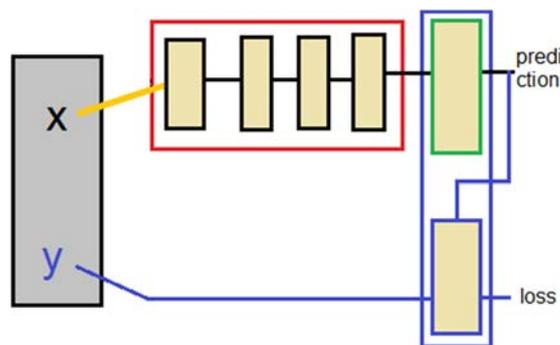


Рис. 1. Архитектура решения для распознавания одного объекта: unstructured x \rightarrow structured \rightarrow representation \rightarrow posterior \rightarrow prediction

0. На вход подается слабоструктурированное наблюдаемое описание объекта – МРТ-изображение опухоли головного мозга, x -трехмерный тензор размерности [512, 512, 4], где каждый элемент тензора представляет собой натуральное значение в интервале [0; 255].

1. От слабоструктурированного описания выполняется переход к первичному структурированному представлению. Изображение преобразуется в формат RGB и нормализуется. Структурированное представление – x -трехмерный тензор размерности [512, 512, 3], где каждый элемент тензора представляет собой вещественное число в интервале [0; 1].

2. От первичного структурированного представления происходит переход к основному структурированному представлению. Переход происходит через несколько промежуточных представлений. Каждое промежуточное представление – это выход некоторого слоя нейронной сети в процессе применения. Основное структурированное представление – выход предпоследнего слоя нейронной сети, который представляет собой вектор некоторого размера и отражает скрытые высокоуровневые свойства данного объекта x .

3. Получение выхода для объекта. Выход для объекта – это выход последнего слоя нейронной сети в процессе применения, который представляет собой вектор размера 3 (в работе классифицируется три вида опухолей головного мозга) и оценку апостериорного распределения на классы (виды опухолей).

4. Принятие окончательного решения для соответствующего МРТ-изображения опухоли (объекта x) определяется по виду опухоли, доставляющему максимум апостериорному распределению (выходу для данного объекта). Математически, $\bar{y} = \arg \max_{i \in Y} a_i(x)$, где \bar{y} – окончательное решение; $Y = \{0, 1, 2\}$ – множество видов опухолей; $a(x)$ – выход для объекта x (оценка апостериорного распределения).

Для прецедентов (объектов с известным скрытым состоянием y) дополнительно можно рассчитать фактическое значение функции потерь:

$$L(y, \bar{y}) = [y = \bar{y}] = \begin{cases} 1, & y = \bar{y}, \\ 0, & y \neq \bar{y}. \end{cases}$$

Переход к структурированному представлению

Дано: слабоструктурированное наблюдаемое описание объекта – МРТ-изображение опухоли головного мозга, x -трехмерный тензор размерности [512, 512, 4], где каждый элемент тензора представляет собой натуральное значение в интервале [0; 255].

Изображение преобразуется в формат RGB. Структурированное представление – x -трехмерный тензор размерности [512, 512, 3].

Полученный тензор нормализуется, т.е. каждый элемент полученного тензора делится на значение 255. Каждый элемент преобразованного тензора представляет собой вещественное число в интервале [0; 1].

Обогащение выборки

Обогащение выборки проводится как этап предобработки объекта, т.е. отдельные объекты генерируются очень быстро. Обогащение выборки производится как часть предобработки с помощью библиотеки Keras.

Были выделены следующие допустимые преобразования для обогащения выборки для данной задачи:

– случайное отображение изображения (Random Flip) (рис. 2). В данном случае не уместно только отображение по горизонтали;

– случайный сдвиг изображения (Random Translation) (рис. 3) по горизонтали и вертикали. В данном случае уместны небольшие сдвиги (смещение до 20 %);

– случайный поворот изображения (Random Rotation) (рис. 4). В данном случае уместны только небольшие повороты (менее 45 градусов);

– случайное масштабирование изображения (Random Zoom) (рис. 5) по горизонтали и вертикали. В данном случае уместен небольшой процент масштабирования (менее 20 %);

- случайное изменение контрастности изображения (Random Contrast) (рис. 6). Допустим небольшой коэффициент изменения контрастности;
- случайное изменение яркости изображения (Random Brightness) (рис. 7). Допустим небольшой коэффициент изменения яркости.

Таким образом, для данной содержательной задачи было выделено множество допустимых преобразований для обогащения выборки (рис. 8).

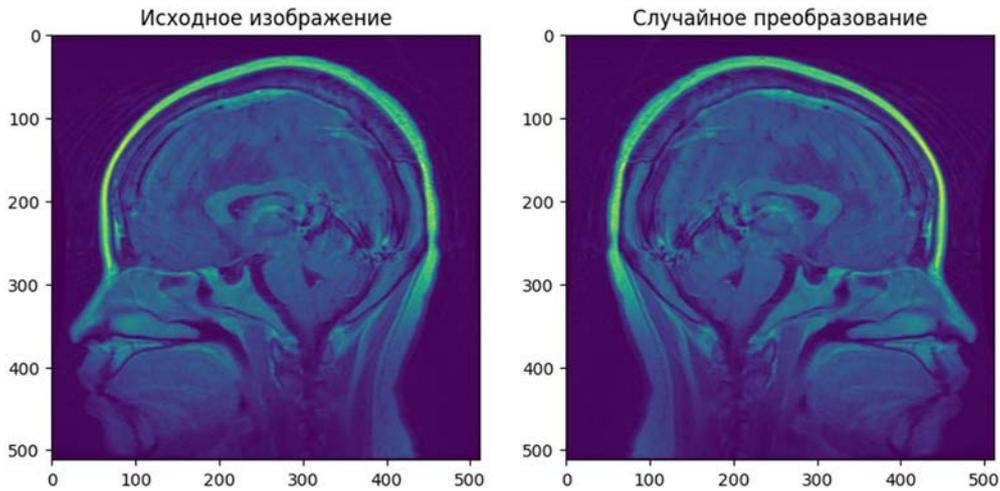


Рис. 2. Случайное отображение изображения по горизонтали

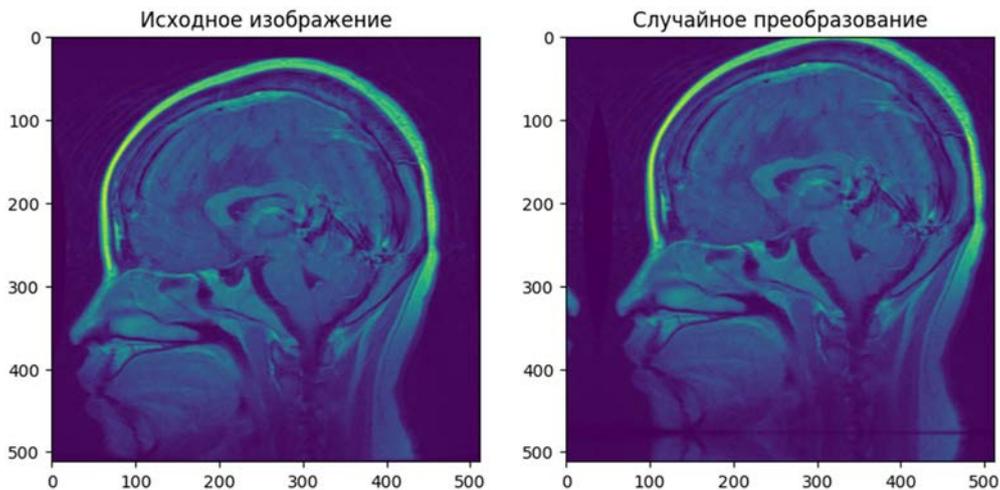


Рис. 3. Случайный сдвиг изображения

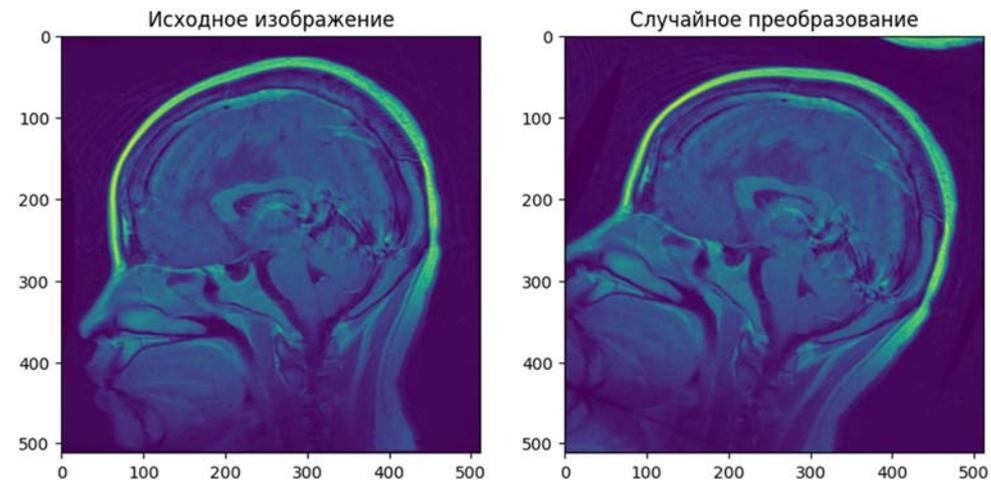


Рис. 4. Случайный поворот изображения

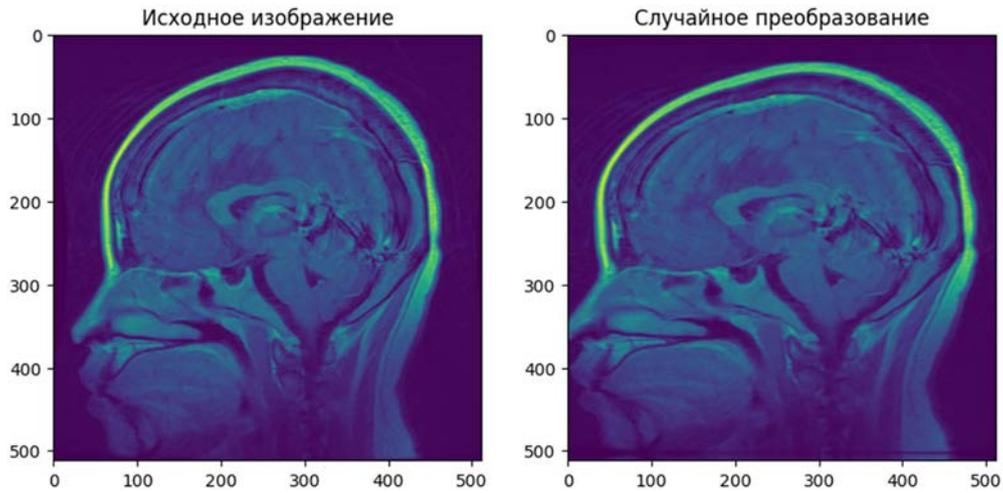


Рис. 5. Случайное масштабирование по горизонтали и вертикали

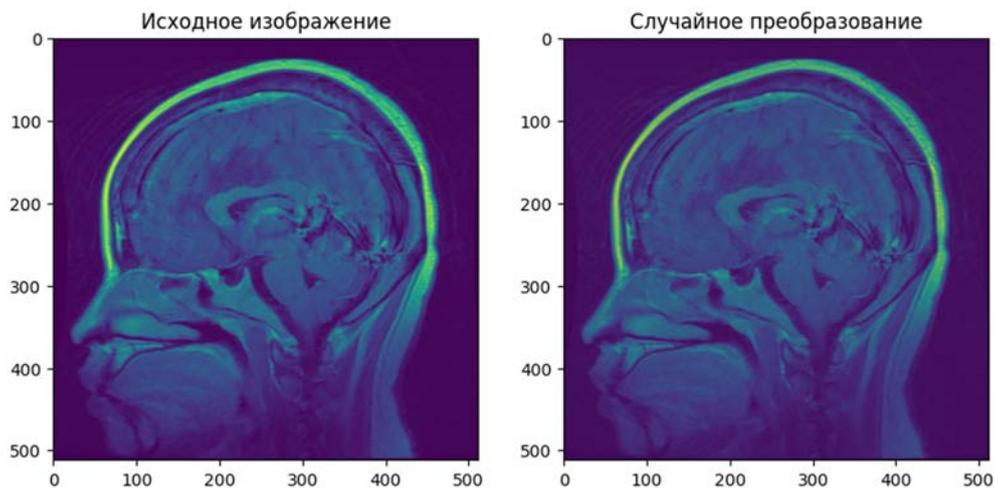


Рис. 6. Случайное изменение контрастности изображения

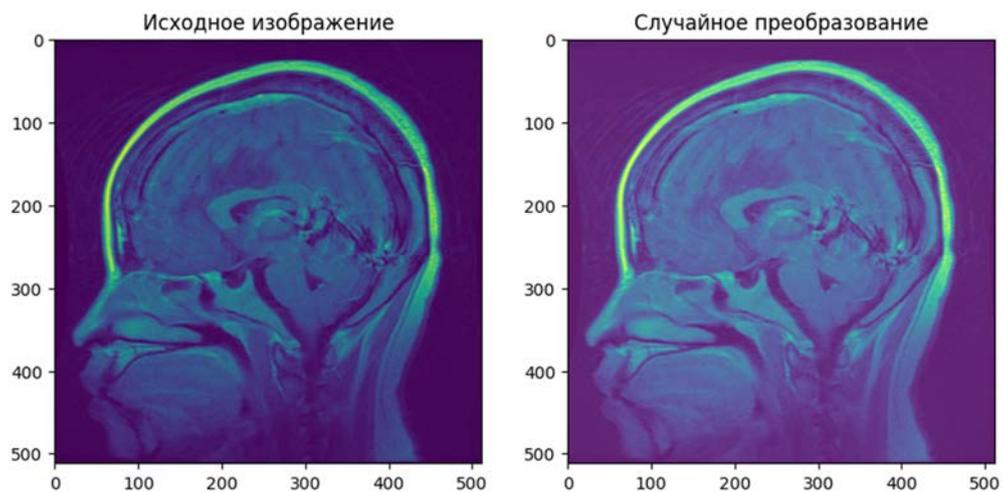


Рис. 7. Случайные изменения яркости изображения

Также было проверено, что обогащение объекта в некоторых случаях помогает при его распознавании. При распознавании по объекту было сгенерировано 50 псевдообъектов при обогащении. Каждый псевдообъект был классифицирован моделью и ответы агрегированы. Изначально модель классифицировала объект неверно, но при обогащении и агрегации модель классифицировала большее количество псевдообъектов (26 / 17 / 7) к верному классу.

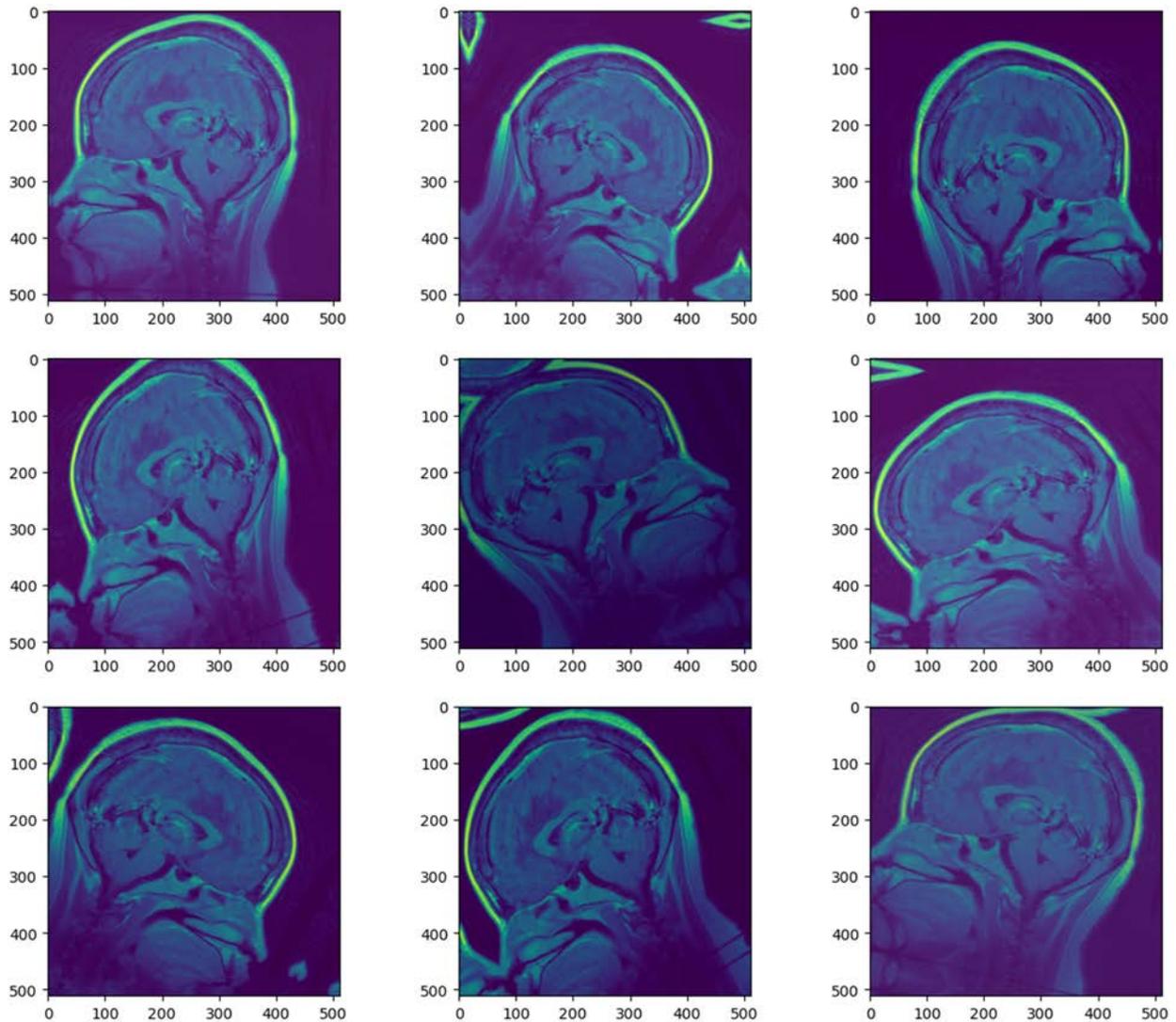


Рис. 8. Визуализация многократного совместного применения всех случайных преобразований для одного изображения

Разные режимы обучения

Для классификации применялись алгоритмы глубокого обучения. Глубокое обучение – это активно развивающаяся область машинного обучения, которая в настоящее время применяется для решения различных задач [4, 5].

Были рассмотрены несколько архитектур для совместного обучения этапов 2 и 3.

Первая архитектура (рис. 9) состоит из четырех последовательно расположенных сверточных слоев и операции взятия подвыборки (MaxPooling), за которыми идут два полносвязных слоя. Существенный недостаток такой архитектуры заключается в большом числе нейронов в полносвязном слое (262144).

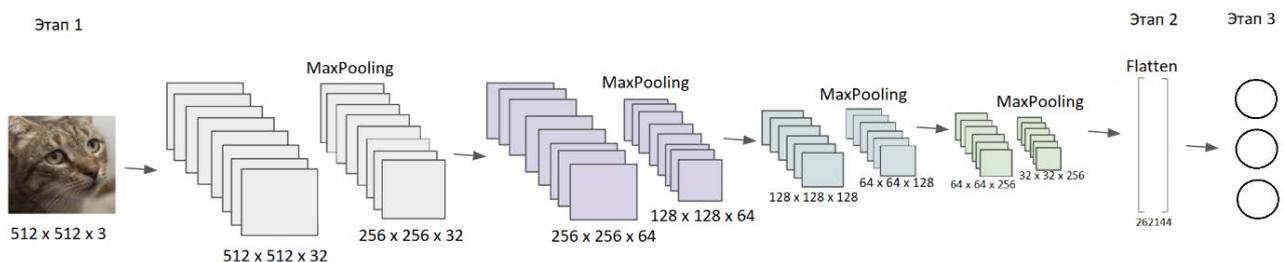


Рис. 9. Архитектура обучения № 1

Недостаток первой архитектуры может быть устранен добавлением еще двух сверточных слоев и уменьшением количества сверток на каждом сверточном слое (рис. 10).

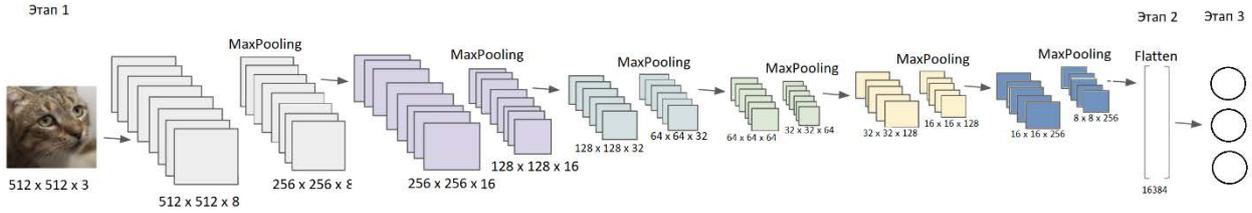


Рис. 10. Архитектура обучения № 2

В третьем виде архитектуры после каждого сверточного слоя и операции взятия подвыборки добавляется пакетная нормализация (Batch Normalization) для более быстрого и стабильного обучения нейронных сетей [6].

Также в данной архитектуре был уменьшен размер вектора основного векторного представления путем добавления дополнительного полносвязного слоя (рис. 11).

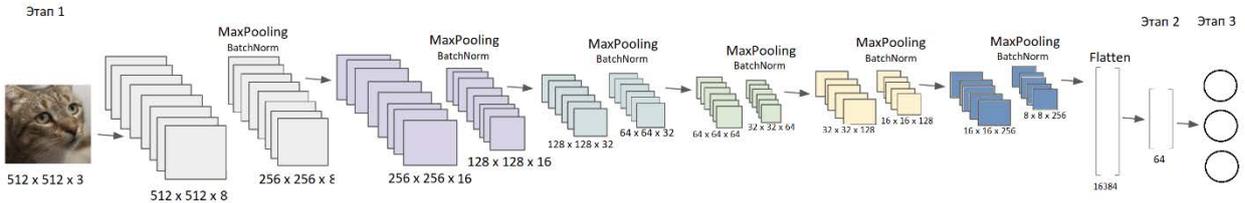


Рис. 11. Архитектура обучения № 3

На этапе 1 для всех трех архитектур имеется трехмерный тензор размерности [512, 512, 3], а на этапе 3 для всех трех архитектур имеется оценка апостериорного распределения на 3 класса (3 вида опухоли). На этапе 2 в зависимости от вида архитектуры были получены векторы различного размера:

- 1) вектор размера 262.144;
- 2) вектор размера 16.384;
- 3) вектор размера 64.

Следует отметить, что до получения основного представления было получено множество промежуточных представлений (рис. 9–11).

Обработка объектов осуществлялась индивидуально, т.е. была построена последовательная модель, в которой от основного представления выполнялся переход сразу к апостериорному распределению. Использовалась Cross-Entropy функция потерь:

$$H(p, q) = - \sum_x p(x) \cdot \log q(x),$$

$p(x)$ – истинное распределение на классы $y \in Y$; $q(x)$ – распределение на классы $y \in Y$, предсказанное моделью; $Y = \{0, 1, 2\}$ – множество видов опухолей.

1. Были рассмотрены несколько режимов обучения.

1) обучение производилось по жестким классам (использовалось One-Hot кодирование как истинное распределение на классы $p(x)$), т.е. для объекта (x, y) верное распределение на классы было следующим:

$$p_z(x) = \begin{cases} 1, & z = y \\ 0, & z \neq y \end{cases}, \quad z \in Y = \{0, 1, 2\};$$

2) обучение производилось по классам со сглаживанием меток (label smoothing) [8]:

$$p'_y(x) = (1 - \varepsilon) p_y(x) + \varepsilon u(y|x),$$

$u(y|x)$ – распределение шума; $\varepsilon \in [0; 1]$.

Также обработка объектов производилась наборами.

2. Обучение по парам объектов («сиамская архитектура», contrastive loss). Для получения основного векторного представления для каждого объекта пары (Этап 2) использовалась сверточная нейронная сеть из пяти последовательных сверточных слоев (включающих операции свертки, Max-Pooling и пакетную нормализацию) и двух полносвязных слоев (по 64 и 10 нейронов). Таким образом, эмбединг каждого объекта в данной архитектуре представляет собой вектор размера 10. Во время обучения для расчета расстояния между эмбедингами использовалось евклидово расстояние.

После обучения векторных представлений таким образом был отдельно обучен этап 3 (получение оценок апостериорных вероятностей).

3. Обучение по тройкам объектам (triplet loss) с отбором троек. Для обучения эмбединга по тройкам использовалась точно такая же архитектура нейронной сети, как в режиме 2. Таким образом, эмбединг каждого объекта представляет собой также вектор размера 10.

Следует отметить, что обучение во всех рассмотренных режимах производилось для архитектуры № 3, так как она продемонстрировала наилучшее качество при обучении. Другие архитектуры были обучены только в режиме обучения 1.1.

Также следует отметить, что архитектура модели при обучении и применении сильно отличается. В режимах обучения 1.1 и 1.2 обучение этапа 2 и 3 производится совместно. В режимах обучения 2, 3 обучение этапа 2 (построение векторного представления) происходит отдельно от этапа 3 (получение оценок апостериорных вероятностей).

Во время обучения этапа 3 веса нейронной сети, обученной для построения эмбединга, не изменялись.

Сравнение систем и окончательные выводы

Для корректности проводимого эксперимента обучение моделей всех архитектур производилось в течение 20 эпох. Для различных архитектур моделей и при наличии/отсутствии обогащения выборки было оценено качество моделей на обучающей и валидационной выборках (табл. 1). Из данной таблицы следует вывод, что для получения хорошего качества при обучении на обогащенной выборке требуется большее число эпох.

Таблица 1

Качество моделей различных архитектур при наличии/отсутствии обогащения выборки на обучающей и валидационной выборках

Модель	Наличие / отсутствие	Обучающая выборка	Валидационная выборка
№ 1	Без обогащения	1,00	0,90
	С обогащением	0,84	0,84
№ 2	Без обогащения	1,00	0,93
	С обогащением	0,90	0,85
№ 3	Без обогащения	1,00	0,96
	С обогащением	0,89	0,90

Были построены диаграммы качества моделей для различных архитектур моделей и при наличии/отсутствии обогащения выборки на обучающей и валидационной выборках (рис. 12, 13).

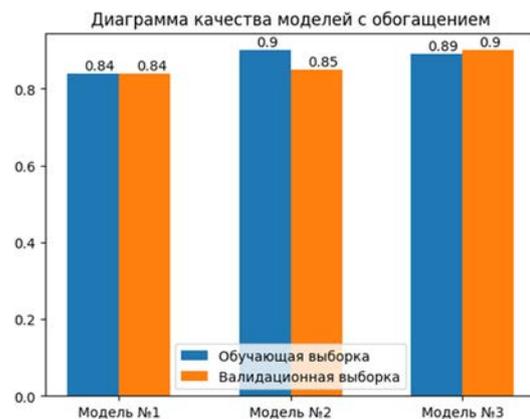
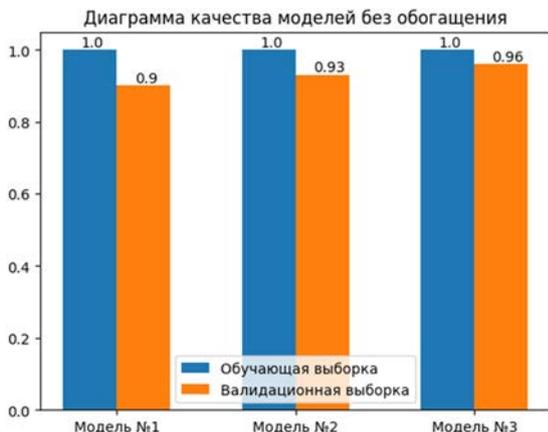


Рис. 12. Диаграмма качества моделей без обогащения

Рис. 13. Диаграмма качества моделей с обогащением

Было произведено сравнение времени обучения (всех модели также обучались ровно 20 эпох для чистоты эксперимента) и применения моделей всех рассмотренных архитектур при наличии/отсутствии обогащения выборки (табл. 2).

Таблица 2

Время обучения и применения моделей различных архитектур при наличии/отсутствии обогащения выборки

Модель	Наличие / отсутствие	Время обучения	Время применения
№ 1	Без обогащения	7 м 23 с	7 с
	С обогащением	11 м 10 с	6 с
№ 2	Без обогащения	4 м 47 с	7 с
	С обогащением	4 м 23 с	7 с
№ 3	Без обогащения	4 м 37 с	7 с
	С обогащением	5 м 1 с	7 с

Из данной таблицы видно, что время обучения моделей существенно отличается, а время применения практически одинаково.

Модель архитектуры № 3 была обучена в различных режимах обучения, было оценено качество (табл. 3) и время обучения/применения соответствующих моделей (табл. 4).

Таблица 3

Качество моделей при различных режимах обучения этапа 2 при наличии/отсутствии обогащения выборки на обучающей и валидационной выборках

Режим обучения	Наличие / отсутствие	Обучающая выборка	Валидационная выборка
Обучение по жестким классам	Без обогащения	1,00	0,96
	С обогащением	0,89	0,90
Обучение со сглаживанием меток	Без обогащения	1,00	0,96
	С обогащением	0,92	0,90
Обучение по парам объектов	Без обогащения	0,88	0,89
	С обогащением	0,88	0,90
Обучение по тройкам объектов	Без обогащения	0,63	0,63
	С обогащением	0,61	0,63

Таблица 4

Время обучения и применения моделей в различных режимах обучения при наличии/отсутствии обогащения выборки

Режим обучения	Наличие / отсутствие	Время обучения	Время применения
Обучение по жестким классам	Без обогащения	6 м 43 с	7 с
	С обогащением	5 м 32 с	7 с
Обучение со сглаживанием меток	Без обогащения	7 м 6 с	7 с
	С обогащением	4 м 40 с	9 с
Обучение по парам объектов	Без обогащения	7 м 10 с	1 с
	С обогащением	6 м 3 с	1 с
Обучение по тройкам объектов	Без обогащения	1 м 3 с	1 с
	С обогащением	1 м 10 с	1 с

Были построены диаграммы качества моделей в различных режимах обучения и при наличии/отсутствии обогащения выборки на обучающей и валидационной выборках (рис. 14, 15).

Модель-победитель выбиралась по качеству на валидационной выборке среди всех моделей. Моделью-победителем оказалась модель архитектуры № 3 без обогащения, достигшая качества 0,96 на валидационной выборке (и качества 1,0 на тренировочной выборке). Также для модели-победителя была получена точечная оценка качества на контрольной выборке: 0,989.

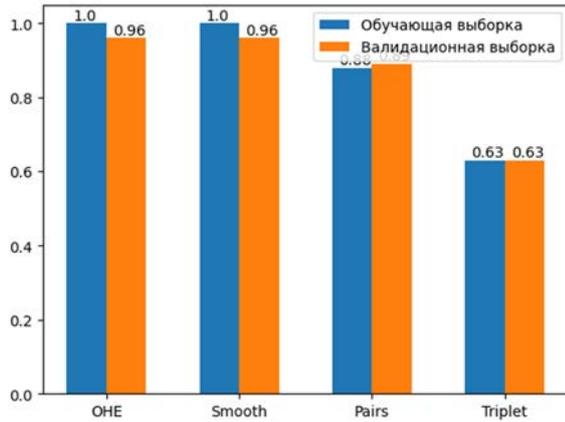


Рис. 14. Диаграмма качества режимов обучения без обогащения

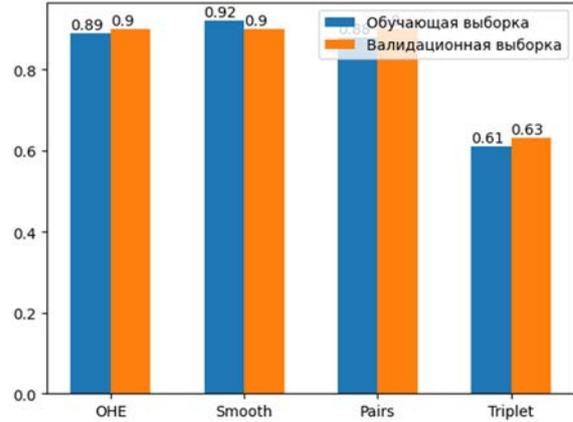


Рис. 15. Диаграмма качества режимов обучения с обогащением

Для получения доверительного интервала оценки качества модели-победителя использовался приближенный метод bootstrap, с помощью которого было сгенерировано 10 bootstrap-выборок (Было взято малое число выборок, так как обучение нейронных сетей требует достаточно времени). На каждой bootstrap-выборке была обучена модель архитектуры № 3 без обогащения и оценено качество на контрольной выборке. Был задан доверительный уровень $p = 0,95$ и рассчитан соответствующий доверительный интервал $I = [0,93; 0,98]$.

Интерпретируемость компонентов векторного представления

После выполнения этапа 2 было получено основное векторное представление, которое содержит меньше признаков, чем первичное представление, при этом признаки являются высокоуровневыми.

Для возможности интерпретации была выбрана модель архитектуры № 3, так как ее основное векторное представление представляет собой вектор размера 64.

Полученные основные представления достаточно тяжело проинтерпретировать, так как в процессе обучения нейронной сети этап 2 и 3 обучались совместно. Однако были получены основные векторные представления для объектов различных классов (рис. 16–18) и в них была выявлена закономерность: большинство компонент вектора основного представления оказались равны нулю, а лишь малая часть отлична от нуля; при этом для объектов одного класса индексы отличных от нуля компонент, как правило, одинаковы. Это свидетельствует о наличии некоторых высокоуровневых признаков в основном векторном представлении.

```

<tf.Tensor: shape=(64,), dtype=float32, numpy=
array([ 0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  51.27904,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  31.346498,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ], dtype=float32)>,
0.0,
<tf.Tensor: shape=(64,), dtype=float32, numpy=
array([ 0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  82.03193,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  32.462673,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ,
        0.      ,  0.      ,  0.      ,  0.      ,  0.      ,  0.      ], dtype=float32)>,
0.0,
    
```

Рис. 16. Основное представление модели для объектов класса 0

основного векторного представления объекты также будут обладать такой кластерной структурой, что и позволяет выполнять их верную классификацию.

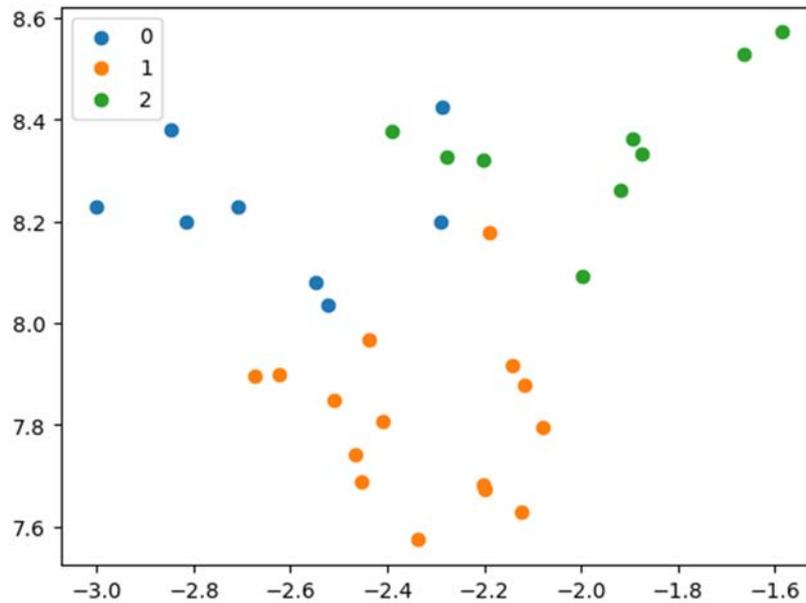


Рис. 19. Визуализация основного векторного представления пакета объектов тренировочной выборки на двумерной плоскости с помощью T-SNE

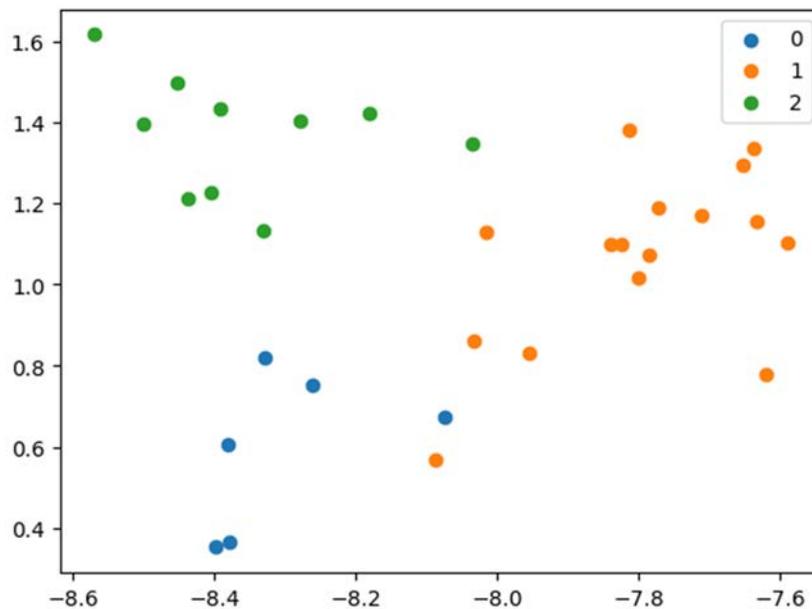


Рис. 20. Визуализация основного векторного представления пакета объектов валидационной выборки на двумерной плоскости с помощью T-SNE

Заключение

Таким образом, в ходе написания данной работы:

- была поставлена содержательная задача автоматизации процесса определения вида опухоли;
- были описаны генеральная совокупность, выборка, объекты и данные для поставленной задачи;
- была сформулирована формальная постановка задачи;
- произведено разбиение выборки с сохранением репрезентативности;
- был описан и программно реализован процесс распознавания объекта, т.е. процесс применения обученных моделей классификации различных архитектур;

- был описан и программно реализован переход от слабоструктурированного представления к структурированному представлению;
- было произведено обогащение выборки и произведено обучение моделей в различных режимах обучения и различных архитектур моделей при наличии и отсутствии обогащения;
- были описаны четыре режима обучения моделей для поставленной задачи;
- была описана попытка интерпретируемости основного векторного представления;
- произведено сравнение времени обучения и применения моделей; сравнение качества распознавания моделей. Сравнение качества по выборке было осуществлено не только точечной оценкой, но доверительным интервалом, полученным приближенно с помощью метода bootstrap. Окончательным победителем оказалась архитектура № 3.

Список литературы

1. Brain Tumor Image Dataset. URL: <https://www.kaggle.com/datasets/denizkavi1/brain-tumor> (дата обращения: 20.03.24).
2. Cheng J. Enhanced performance of brain tumor classification via tumor region augmentation and partition // *PloS one*. 2015. Т. 10, № 10. С. e0140381.
3. Оценка моделей. URL: http://www.machinelearning.ru/wiki/images/archive/5/54/20151010103437!Kitov-ML-05-Model_evaluation.pdf (дата обращения: 20.03.24).
4. Тимошкин М. О., Романова Е. Г. Применение метода Q-обучения для оптимизации производств // Аналитические и численные методы моделирования естественно-научных и социальных проблем : сб. ст. по материалам науч.-техн. конф. Пенза : Изд-во ПГУ, 2021. С. 100–107.
5. Тимошкин М. О., Романова Е. Г. Применение метода глубокого Q-обучения для обучения агентов в логических играх // Математическое и компьютерное моделирование естественно-научных и социальных проблем : материалы XVI Всерос. науч.-техн. конф. Пенза : Изд-во ПГУ, 2022. С. 199–206.
6. Генеральная совокупность – 2022. URL: https://ru.wikipedia.org/wiki/Генеральная_совокупность (дата обращения: 20.03.24).
7. Основы теории статистики – 2022. URL: <https://elar.ufrb.br/bitstream/10995/34746/1/978-5-7996-1520-8.pdf> (дата обращения: 20.03.24).
8. Label Smoothing. 2023. URL: <https://leimao.github.io/blog/Label-Smoothing/> (дата обращения: 20.03.24).
9. Image similarity estimation using a Siamese Network with a contrastive loss. 2023. URL: https://keras.io/examples/vision/siamese_contrastive/ (дата обращения: 20.03.24).
10. Batch normalization. URL: https://en.wikipedia.org/wiki/Batch_normalization (дата обращения: 20.03.24).
11. Репрезентативность. 2022. URL: <https://ru.wikipedia.org/wiki/Репрезентативность> (дата обращения: 20.03.24).
12. Звенигородский И. И., Зырянов Ю. Т., Ульшин Д. И. Использование нейросетевых регуляторов для повышения энергоэффективности и качества системы вентиляции и кондиционирования воздуха заглубленного сооружения // *Надежность и качество сложных систем*. 2023. № 1. С. 99–105.
13. Иванов А. И., Иванов А. П., Горбунов К. А. Нейросетевое преобразование биометрии в код аутентификации: дополнение энтропии хэмминга энтропией корреляционных связей между разрядами // *Надежность и качество сложных систем*. 2023. № 1. С. 91–98.
14. Иванов А. И., Кубасов И. А. Сильный искусственный интеллект: повышение качества нейросетевых решений с переходом к обработке входных данных большого объема // *Надежность и качество сложных систем*. 2021. № 1. С. 9–16.

References

1. *Brain Tumor Image Dataset*. Available at: <https://www.kaggle.com/datasets/denizkavi1/brain-tumor> (accessed 20.03.24).
2. Cheng J. Enhanced performance of brain tumor classification via tumor region augmentation and partition. *PloS one*. 2015;10(10):e0140381.
3. *Otsenka modeley = Evaluation of models*. (In Russ.). Available at: http://www.machinelearning.ru/wiki/images/archive/5/54/20151010103437!Kitov-ML-05-Model_evaluation.pdf (accessed 20.03.24).
4. Timoshkin M.O., Romanova E.G. Application of the Q-learning method for production optimization. *Analiticheskie i chislennye metody modelirovaniya estestvenno-nauchnykh i sotsial'nykh problem: sb. st. po materialam nauch.-tekhn. konf. = Analytical and numerical methods for modeling natural science and social problems : collection of articles based on the materials of scientific and technical conf.* Penza: Izd-vo PGU, 2021:100–107. (In Russ.)
5. Timoshkin M.O., Romanova E.G. Application of the deep Q-learning method for training agents in logic games. *Matematicheskoe i komp'yuternoe modelirovanie estestvenno-nauchnykh i sotsial'nykh problem: materialy XVI*

- Vseros. nauch.-tekhn. konf. = Mathematical and computer modeling of natural science and social problems : materials of the XVI All-Russian scientific and technical conf.* Penza: Izd-vo PGU, 2022:199–206. (In Russ.)
6. *General'naya sovokupnost' – 2022 = General combination – 2022.* (In Russ.). Available at: https://ru.wikipedia.org/wiki/General'naya_sovokupnost' (accessed 20.03.24).
 7. *Osnovy teorii statistiki – 2022 = Fundamentals of the theory of statistics – 2022.* (In Russ.). Available at: <https://elar.urfu.ru/bitstream/10995/34746/1/978-5-7996-1520-8.pdf> (accessed 20.03.24).
 8. *Label Smoothing.* 2023. Available at: <https://leimao.github.io/blog/Label-Smoothing/> (accessed 20.03.24).
 9. *Image similarity estimation using a Siamese Network with a contrastive loss.* 2023. Available at: https://keras.io/examples/vision/siamese_contrastive/ (accessed 20.03.24).
 10. *Batch normalization.* Available at: https://en.wikipedia.org/wiki/Batch_normalization (accessed 20.03.24).
 11. *Reprezentativnost' = Representativeness.* 2022. (In Russ.). Available at: <https://ru.wikipedia.org/wiki/Reprezentativnost'> (accessed 20.03.24).
 12. Zvenigorodskiy I.I., Zyryanov Yu.T., Ul'shin D.I. The use of neural network regulators to improve energy efficiency and quality of the ventilation and air conditioning system of a buried structure. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems.* 2023;(1):99–105. (In Russ.)
 13. Ivanov A.I., Ivanov A.P., Gorbunov K.A. Neural network transformation of biometrics into authentication code: addition of hamming entropy by entropy of correlations between bits. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems.* 2023;(1):91–98. (In Russ.)
 14. Ivanov A.I., Kubasov I.A. Strong artificial intelligence: improving the quality of neural network solutions with the transition to processing large-volume input data. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems.* 2021;(1):9–16. (In Russ.)

Информация об авторах / Information about the authors

Максим Олегович Тимошкин

студент,
Московский государственный университет
имени М. В. Ломоносова
(Россия, г. Москва, ул. Ленинские Горы, 1)
E-mail: max.timoshkin@inbox.ru

Maksim O. Timoshkin

Student,
Moscow State University
named after M.V. Lomonosov
(1 Leninskie Gory street, Moscow, Russia)

Елена Геннадьевна Романова

кандидат технических наук, доцент,
доцент кафедры высшей и прикладной математики,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: romanova.elenar2016@yandex.ru

Elena G. Romanova

Candidate of technical sciences, associate professor,
associate professor of the sub-department of higher
and applied mathematics,
Penza State University
(40 Krasnaya street, Penza, Russia)

Авторы заявляют об отсутствии конфликта интересов /

The authors declare no conflicts of interests.

Поступила в редакцию / Received 20.06.2024

Поступила после рецензирования / Revised 27.07.2024

Принята к публикации / Accepted 30.08.2024

ПРОЕКТИРОВАНИЕ И ТЕХНОЛОГИИ ПРИБОРОСТРОЕНИЯ И РАДИОЭЛЕКТРОННОЙ АППАРАТУРЫ

DESIGN AND TECHNOLOGY OF INSTRUMENTATION AND ELECTRONIC EQUIPMENT

УДК 681.586.2

doi: 10.21685/2307-4205-2024-3-7

ПРОЕКТИРОВАНИЕ ИМИТАЦИОННОЙ МОДЕЛИ ДАТЧИКА ТЕМПЕРАТУРЫ ПРИ ВОЗДЕЙСТВИИ КРИТИЧЕСКИХ ТЕМПЕРАТУР И СИНУСОИДАЛЬНОЙ ВИБРАЦИИ

Е. А. Фокина¹, А. А. Трофимов², В. Н. Пономарев³, С. А. Здобнов⁴

^{1, 2, 3, 4} Пензенский государственный университет, Пенза, Россия

¹ekaterina.isay1997@gmail.com, ²alex.t1978@mail.ru, ³revik2296@gmail.com, ⁴sa_zdobnov@mail.ru

Аннотация. *Актуальность и цели.* Измерение вибрации – это один из самых популярных и экономичных методов диагностики общего состояния оборудования и имеет неоспоримое значение в современной технике. Благодаря этому процессу производители и потребители могут быть уверены в надежности и качестве технических систем, а также предотвращать возможные поломки и сбои. А широкий объем вибрационных измерений в различных областях подтверждает их универсальность и важность в современном техническом прогрессе. Цель измерения воздействия критических температур на датчик температуры заключается в определении того, как изменение критической температуры может повлиять на точность и надежность работы преобразователя. Это позволит улучшить качество измерения температуры, исключить погрешности и ошибки в работе оборудования, а также повысить эффективность его использования в различных условиях эксплуатации. *Материалы и методы.* Методом исследования в данной статье является имитационная модель датчика, с помощью которой производится испытание на вибрацию и воздействие критических температур. Для создания модели используются программные обеспечения Ansys и SolidWorks Simulation, позволяющие избежать дорогостоящих и длительных циклов испытаний. *Результаты.* В результате имитационного моделирования были построены эпюры распределения механических напряжений при воздействии синусоидальной вибрации и эпюры распределения температур при воздействии критических температур.

Ключевые слова: датчик, имитационная модель, вибрация, чувствительный элемент

Для цитирования: Фокина Е. А., Трофимов А. А., Пономарев В. Н., Здобнов С. А. Проектирование имитационной модели датчика температуры при воздействии критических температур и синусоидальной вибрации // Надежность и качество сложных систем. 2024. № 3. С. 66–72. doi: 10.21685/2307-4205-2024-3-7

DESIGNING A SIMULATION MODEL OF A TEMPERATURE SENSOR UNDER CRITICAL TEMPERATURE AND SINUSOIDAL VIBRATION CONDITIONS

E.A. Fokina¹, A.A. Trofimov², V.N. Ponomarev³, S.A. Zdobnov⁴

^{1, 2, 3, 4} Penza State University, Penza, Russia

¹ekaterina.isay1997@gmail.com, ²alex.t1978@mail.ru, ³revik2296@gmail.com, ⁴sa_zdobnov@mail.ru

Abstract. *Background.* Vibration measurement is one of the most popular and cost-effective methods of diagnosing the overall condition of equipment and is of undeniable importance in modern technology. Thanks to this process, manufacturers and consumers can be sure of the reliability and quality of technical systems, as well as prevent possible breakdowns and failures. And the wide scope of vibration measurements in various fields confirms their versatility and importance in modern technical progress. The purpose of measuring the effects of critical temperatures on a temperature transducer is to determine how a change in critical temperature can affect the accuracy and reliability of the transducer. This will improve the quality of temperature measurement, eliminate inaccuracies and errors in equipment performance, and increase the efficiency of the equipment in various operating conditions. *Materials and methods.* The research method in this paper is a simulation model of the sensor, which is used to test for vibration and exposure to critical temperatures. Ansys and SolidWorks Simulation software are used to create the model to avoid costly and time-consuming test cycles. *Results.* As a result of the simulation modeling, the epuples of mechanical stress distribution under sinusoidal vibration and the epuples of temperature distribution under critical temperatures were constructed.

Keywords: sensor, imitation model, displacement, sensing element

For citation: Fokina E.A., Trofimov A.A., Ponomarev V.N., Zdobnov S.A. Designing a simulation model of a temperature sensor under critical temperature and sinusoidal vibration conditions. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):66–72. (In Russ.). doi: 10.21685/2307-4205-2024-3-7

Имитационной моделью датчика является логико-математическое описание объекта, используемое для анализа и оценки функционирования объекта на основе компьютерной модели, которая воспроизводит процессы функционирования реальной структуры изучаемого объекта.

Имитационное моделирование играет ключевую роль в разработке и тестировании датчиков температуры в условиях критических температур и воздействия синусоидальной вибрации [1, 2]. Имитационной моделью датчика является логико-математическое описание объекта, используемое для анализа и оценки функционирования объекта на основе компьютерной модели, которая воспроизводит процессы функционирования реальной структуры изучаемого объекта. Это позволяет инженерам и разработчикам предсказать поведение датчика в различных ситуациях, оптимизировать его работу и повысить надежность.

Проектирование имитационной модели датчика температуры при воздействии критических температур и синусоидальной вибрации требует учета множества факторов, таких как материалы, конструкция, теплопроводность и механическая устойчивость. Особое внимание уделяется точности и стабильности измерений при изменяющихся условиях окружающей среды [3, 4].

В данной статье мы рассмотрим основные принципы проектирования имитационной модели датчика температуры, а также рассмотрим влияние критических температур и синусоидальной вибрации на его работу.

Синусоидальная вибрация – это периодическое движение, закон движения которого может быть выражен выражением функции синуса. Например, все вибрации, создаваемые вращением, пульсацией и колебанием, являются синусоидальными колебаниями. Важно определять влияние синусоидальной вибрации на объекты эксперимента, так как действие вибрации может приводить к трансформированию внутренней структуры материалов и поверхностных слоев, изменению условий трения и износа на контактных поверхностях деталей, нагреву конструкций.

Для определения влияния синусоидальной вибрации на конструкцию датчика температуры в среде SolidWorks Simulation построена расчетная модель и схема граничных условий (ГУ), приведенная на рис. 1 [5, 6].

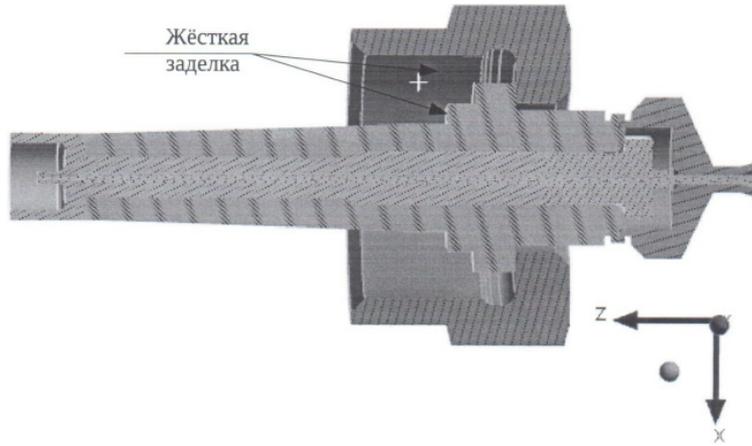


Рис. 1. Схема граничных условий

В ходе имитационного моделирования способность материалов рассеивать энергию определялась коэффициентом демпфирования (коэффициент демпфирования – это отношение внутреннего полного сопротивления усилителя к внутреннему полному сопротивлению системы колонок), который из-за отсутствия экспериментальных данных принимался равными 0,02–0,04. Также из-за отсутствия экспериментальных данных по усталостному разрушению материалов, входящих в состав сборки, не учитывалось время воздействия синусоидальной вибрации.

В результате имитационного моделирования была получена форма колебания резонансной частоты, попадающей в диапазон частот от 15 до 8000 Гц, представленная на рис. 2.

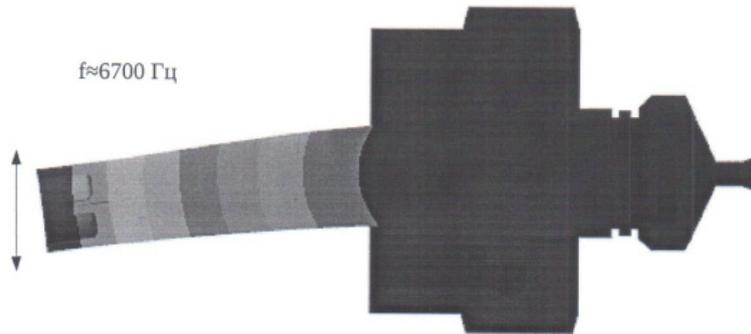


Рис. 2. Форма колебаний

Также были получены эпюры распределения механических напряжений при воздействии синусоидальной вибрации.

По данным рис. 3 видно, что максимальная величина механических напряжений варьируется в диапазоне от 406 до 780 МПа.

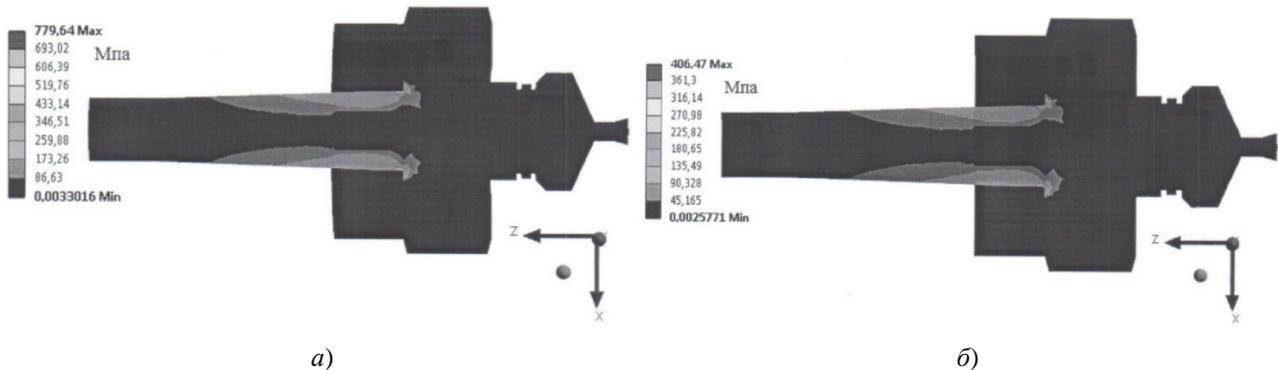


Рис. 3. Эпюры распределения механических напряжений:

а – при коэффициенте демпфирования 0,02; б – при коэффициенте демпфирования 0,04

Для определения теплопередачи при воздействии на датчик температуры критических температур и определения времени, за которое чувствительный элемент достигает измеряемой температуры, была построена 3D-модель датчика температуры в среде SolidWorks Simulation. 3D-модель и схема граничных условий для расчета представлены на рис. 4 [7].



Рис. 4. 3D модель и схема граничных условий для датчика температуры

В результате моделирования воздействия температур 1200 °C, 600 °C и минус 196 °C в течение 500 с были получены эпюры распределения температур, представленные на рис. 5–7.



Рис. 5. Эпюра распределения температуры от 600 °C до плюс 50 °C в течение 500 с

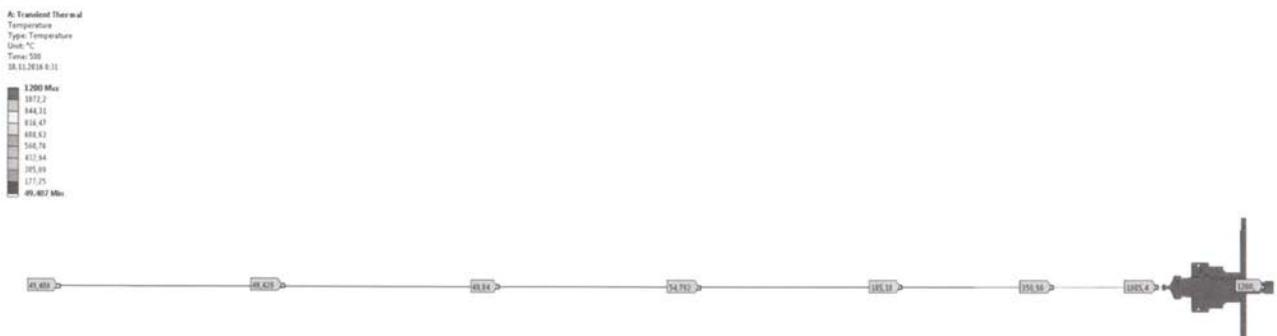


Рис. 6. Эпюра распределения температуры от 1200 °C до плюс 50 °C в течение 500 с



Рис. 7. Эпюра распределения температуры от минус 196 °C до плюс 50 °C в течение 500 с

Также был проведен анализ поведения модели при температуре измеряемой среды 300 °С. Температура окружающей среды составляет $t_{окр} = 22$ °С. Тип передачи тепла между окружающей средой и датчиком – конвекция. Данный тип передачи реализован соответствующим граничным условием. Его главной характеристикой является коэффициент конвективной теплоотдачи (рис. 8). Значения коэффициента конвективной отдачи взяты из стандартной базы данных программы Ansys для окружающей среды – воздух.

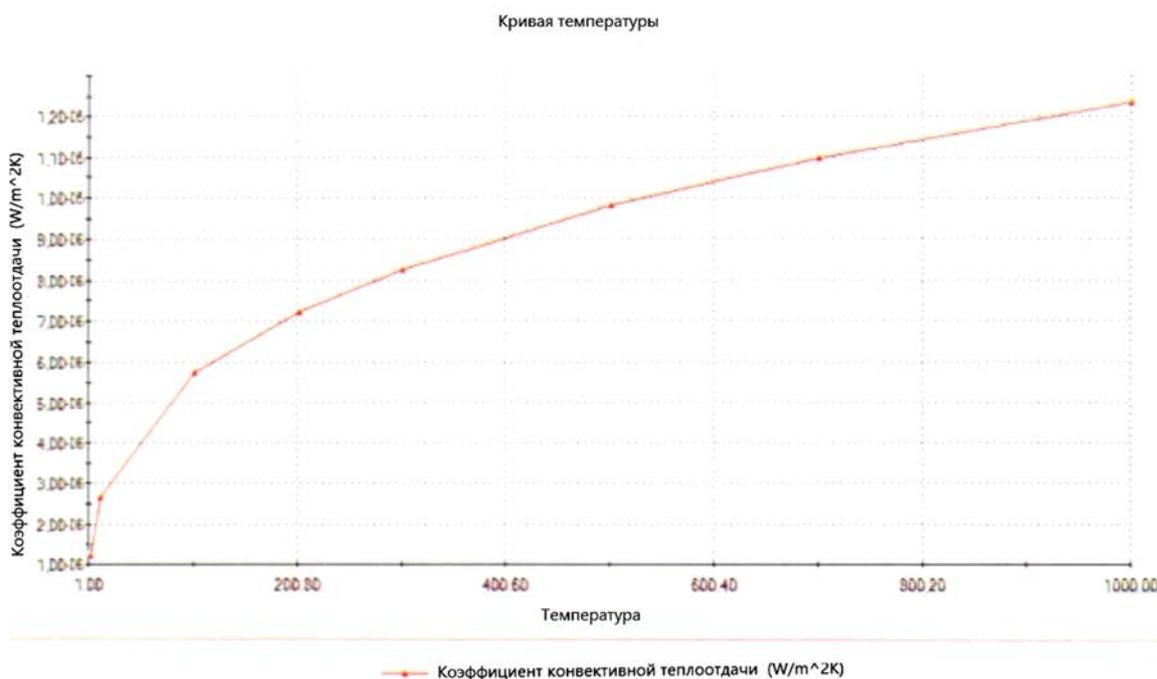


Рис. 8. Значения коэффициента конвективной теплоотдачи

Передача тепла осуществляется за счет конвекционных потоков от нагретой части датчика в окружающую среду и по средствам теплопроводности между деталями датчика. В момент времени 22 000 с наблюдается тепловое равновесие между датчиком и окружающей средой (выравнивание температурных полей).

Результаты моделирования показаны на эпюре распределения температурных полей в момент времени 22 000 (рис. 9).

В рамках данной работы была создана имитационная модель датчика температуры при воздействии критических температур и синусоидальной вибрации. Разработанная модель позволяет исследовать изменение характеристик датчика в условиях вибрации и воздействия критических температур.

В результате разработки имитационной модели воздействия синусоидальной вибрации на датчик температуры определена собственная резонансная частота конструкции датчика, входящая в диапазон 15–8000 Гц и равная 6700 Гц. Также из построенных эпюр выявлено максимальное механическое напряжение, которое образуется в конструкции датчика под воздействием синусоидальной вибрации от 0,8 до 1200 g в диапазоне частот от 15 до 8000 Гц, варьирующееся в диапазоне от 406 до 780 МПа.

По результатам моделирования имитационной модели воздействия критических температур на датчик температуры выявлено, что температура 1200 °С, 600 °С и минус 196 °С за время 500 с не оказывает влияния на соединитель. Основное влияние при воздействии критических температур на соединитель оказывает температура окружающей среды, которая за 500 с достигает соответственно плюс 50 °С и минус 50 °С.

Дополнительно по результатам моделирования имитационной модели воздействия критических температур на датчик температуры была определена максимальная температура, образующаяся на кабеле удлинителе и соединителе при воздействии на датчик температуры $t = 300$ °С. Температура на разъеме датчика составила 26 °С; на кабеле в месте крепления с печатной платой 180 °С.

Полученные результаты могут быть использованы для повышения метрологических характеристик датчика температуры в условиях вибрации и воздействия критических температур.

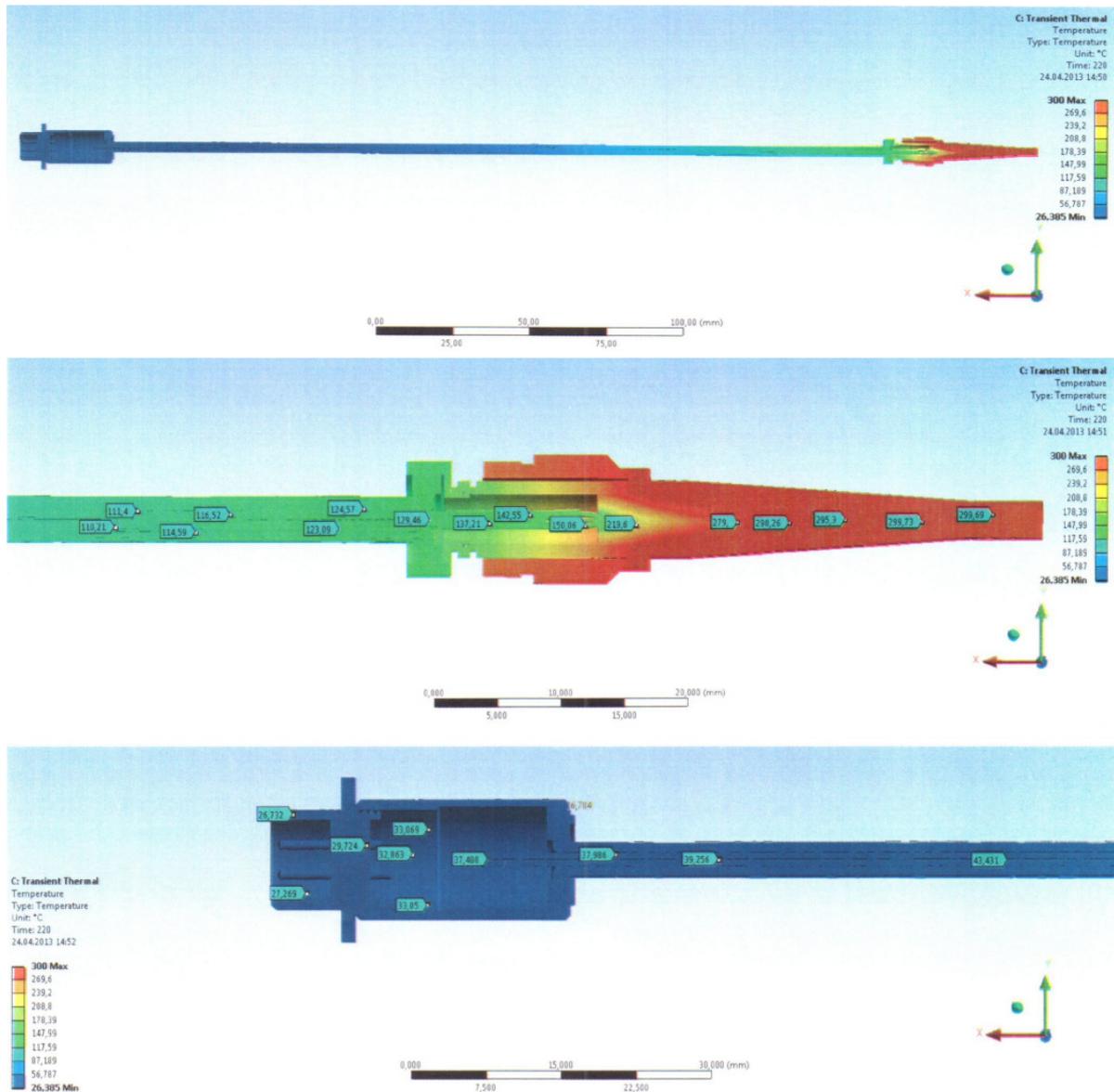


Рис. 9. Эпюра распределения температурных полей в момент времени $T = 22\,000$ с

Список литературы

1. Трофимов А. А., Пономарев В. Н., Дерябин Д. В. [и др.]. Проектирование датчика температуры методом имитационного моделирования // Надежность и качество сложных систем. 2023. № 3. С. 54–61.
2. Бастрыгин К. И., Трофимов А. А., Баранов А. С. [и др.]. Имитационное моделирование пьезоэлектрического датчика давления // Измерение. Мониторинг. Управление. Контроль. 2017. № 1. С. 20–28.
3. Мусаев Р. Ш., Трофимов А. А., Фролов М. А. Имитационное моделирование чувствительного элемента датчика давления струнного типа // Датчики и системы. 2014. № 7. С. 22–25.
4. Мусаев Р. Ш., Трофимов А. А., Фролов М. А. Имитационное моделирование тензорезисторного датчика избыточных давлений // Измерение. Мониторинг. Управление. Контроль. 2022. № 1. С. 56–63.
5. Рыбаков И. М. Модели и методы измерения электрических параметров в сложных системах электроники // Надежность и качество сложных систем. 2023. № 2. С. 44–50.
6. Рязанцев Д. А., Тимонин Р. М., Трофимов А. А. Имитационное моделирование датчика температуры // Датчики и системы. 2019. № 6. С. 11–14.
7. Алямовский А. А. Инженерные расчеты в SolidWorks Simulation. М. : ДМК Пресс, 2010. 464 с.

References

1. Trofimov A.A., Ponomarev V.N., Deryabin D.V., et al. Designing a temperature sensor by simulation modeling. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems*. 2023;(3):54–61. (In Russ.)

2. Bastrygin K.I., Trofimov A.A., Baranov A.S. et al. Simulation modeling of a piezoelectric pressure sensor. *Izmerenie. Monitoring. Upravlenie. Kontrol' = Measurement. Monitoring. Management. Control.* 2017;(1):20–28. (In Russ.)
3. Musaev R.Sh., Trofimov A.A., Frolov M.A. Simulation modeling of a sensitive element of a string-type pressure sensor. *Datchiki i sistemy = Sensors and systems.* 2014;(7):22–25. (In Russ.)
4. Musaev R.Sh., Trofimov A.A., Frolov M.A. Simulation modeling of a tensoresistor overpressure sensor. *Izmerenie. Monitoring. Upravlenie. Kontrol' = Measurement. Monitoring. Management. Control.* 2022;(1):56–63. (In Russ.)
5. Rybakov I.M. Models and methods of measuring electrical parameters in complex electronics systems. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2023;(2):44–50. (In Russ.)
6. Ryazantsev D.A., Timonin R.M., Trofimov A.A. Simulation modeling of a temperature sensor. *Datchiki i sistemy = Sensors and systems.* 2019;(6):11–14. (In Russ.)
7. Alyamovskiy A.A. *Inzhenernye raschety v SolidWorks Simulation = Engineering calculations in SolidWorks Simulation.* Moscow: DMK Press, 2010:464. (In Russ.)

Информация об авторах / Information about the authors

Екатерина Александровна Фокина

аспирант,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: ekaterina.isay1997@gmail.com

Ekaterina A. Fokina

Postgraduate student,
Penza State University
(40 Krasnaya street, Penza, Russia)

Алексей Анатольевич Трофимов

доктор технических наук, доцент,
профессор кафедры информационно-измерительной
техники и метрологии,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: alex.t1978@mail.ru

Aleksei A. Trofimov

Doctor of technical sciences, associate professor,
professor of the sub-department of information
and measuring equipment and metrology,
Penza State University
(40 Krasnaya street, Penza, Russia)

Владислав Николаевич Пономарев

аспирант,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: revik2296@gmail.com

Vladislav N. Ponomarev

Postgraduate student,
Penza State University
(40 Krasnaya street, Penza, Russia)

Сергей Александрович Здобнов

аспирант,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: sa_zdobnov@mail.ru

Sergey A. Zdobnov

Postgraduate student,
Penza State University
(40 Krasnaya street, Penza, Russia)

Авторы заявляют об отсутствии конфликта интересов /

The authors declare no conflicts of interests.

Поступила в редакцию/Received 11.06.2024

Поступила после рецензирования/Revised 02.08.2024

Принята к публикации/Accepted 03.09.2024

КОМПОНОВКА БЕЗРЕЗОНАНСНЫХ БЛОКОВ БОРТОВЫХ РАДИОЭЛЕКТРОННЫХ СИСТЕМ С ИСПОЛЬЗОВАНИЕМ ДВУХСЛОЙНЫХ ПЕЧАТНЫХ ПЛАТ С ЖЕСТКОЙ ОСНОВОЙ

С. И. Фролов¹, Е. А. Данилова², И. И. Кочегаров³

^{2,3} Пензенский государственный университет, Пенза, Россия
¹ sfrolov46@mail.ru, ² siori@list.ru, ³ kipra@mail.ru

Аннотация. *Актуальность и цели.* Целью работы является разработка инженерной методики расчета печатных узлов с жесткой подложкой. Проведен анализ методов и способов расчета безрезонансных (в рабочем диапазоне частот) электронных блоков бортовых радиоэлектронных средств (БРЭС) и в целом проектирования безрезонансных электронных блоков. Проведено моделирование собственных частот колебаний для различных способов закрепления печатных узлов. *Материалы и методы.* В применены математические положения теории механических колебаний для получения аналитических выражений собственной частоты колебаний конструкций. Для тестового сравнения полученных результатов использовался программный пакет ANSYS. *Результаты.* Проведено моделирование собственных частот колебаний для различных способов закрепления печатных узлов. Получены численные решения для различных вариантов многослойных печатных плат, отличающихся толщиной и материалами подложек. Проведена оценка точности расчетов собственной частоты колебаний. Предложена методика создания безрезонансных электронных блоков БРЭС и даны рекомендации по созданию технологичных безрезонансных конструкций блоков БРЭС на основе использования современных материалов подложки печатных плат. *Выводы.* Результаты работы в виде рекомендаций по повышению жесткости конструкций цилиндрических конструкций бортовых радиоэлектронных средств позволят создавать более надежные конструкции за счет улучшения стойкости к вибрации. Оперативный обоснованный выбор схем крепления печатных узлов на ранних этапах проектирования позволит сократить сроки работ.

Ключевые слова: печатные узлы, математическое моделирование, собственные частоты, электронные блоки, резонанс

Для цитирования: Фролов С. И., Данилова Е. А., Кочегаров И. И. Компоновка безрезонансных блоков БРЭС с использованием двухслойных печатных плат с жесткой основой // Надежность и качество сложных систем. 2024. № 3. С. 73–81. doi: 10.21685/2307-4205-2024-3-8

THE LAYOUT OF ON-BOARD ELECTRONIC SYSTEMS WITHOUT RESONANT BLOCKS USING DOUBLE-LAYER PRINTED CIRCUIT BOARDS WITH A RIGID BASE

S.I. Frolov¹, E.A. Danilova², I.I. Kochegarov³

^{2,3} Penza State University, Penza, Russia
¹ sfrolov46@mail.ru, ² siori@list.ru, ³ kipra@mail.ru

Abstract. *Background.* The aim of the work is to develop an engineering methodology for calculating printed circuit board assemblies with a rigid substrate. The analysis of methods and techniques for calculating resonance-free (in the operating frequency range) electronic units of on-board radio-electronic equipment (OREEq) and, in general, the design of resonance-free electronic units is carried out. The modeling of natural oscillation frequencies for various methods of fixing printed circuit board assemblies is carried out. *Materials and methods.* The mathematical principles of the theory of mechanical oscillations are applied to obtain analytical expressions for the natural oscillation frequency of structures. The ANSYS software package was used for test comparison of the obtained results. *Results.* The modeling of natural oscillation frequencies for various methods of fixing printed circuit board assemblies is carried out. Numerical solutions are obtained for various variants of multilayer printed circuit boards differing in thickness and substrate materials. The accuracy of calculations of the natural oscillation frequency is assessed. A method for

creating resonance-free electronic units of OREEq is proposed and recommendations are given for creating technological resonance-free designs of OREEq units based on the use of modern substrate materials for printed circuit boards. *Conclusions.* The results of the work in the form of recommendations for increasing the rigidity of the cylindrical structures of onboard electronic equipment will allow creating more reliable structures due to improved vibration resistance. Prompt, well-founded selection of printed circuit board mounting schemes at early design stages will reduce the work time.

Keywords: printed circuit assemblies, mathematical modeling, natural frequencies, electronic units, resonance

For citation: Frolov S.I., Danilova E.A., Kochegarov I.I. The layout of on-board electronic systems without resonant blocks using double-layer printed circuit boards with a rigid base. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):73–81. (In Russ.). doi: 10.21685/2307-4205-2024-3-8

Введение

В настоящее время перед разработчиками ставятся все более сложные задачи по созданию в сжатые сроки высоконадежных бортовых радиоэлектронных систем (БРЭС), эксплуатируемых в жестких условиях внешних механических воздействий. Наиболее опасными являются вибрации электронных блоков БРЭС из-за резонансных колебаний печатных плат с большим количеством электрорадиоизделий (ЭРИ).

Из-за сжатых сроков проектирования в качестве защиты от вибрации применяются главным образом виброизоляция электронных блоков БРЭС с использованием амортизаторов и/или демпфирование их печатных узлов (ПУ) [1]. Однако эти меры ухудшают массогабаритные характеристики БРЭС, а эффективность их недостаточна, так как поток отказов блоков аппаратуры от воздействия вибрации все еще остается значительным.

Исходя из вышеизложенного, по мнению авторов, для дальнейшего повышения уровня надежности блоков БРЭС практически безальтернативным методом обеспечения их вибропрочности и виброустойчивости является отстройка резонансных частот плоских элементов конструкции ПУ электронных блоков за пределы диапазона частот внешних вибрационных воздействий.

Помимо трудностей создания безрезонансных конструкций с требуемыми массогабаритными характеристиками, серьезной проблемой при конструировании безрезонансных (ПУ) является сложность реализации сквозного цикла проектирования печатных плат, так как с использованием существующих САПР на практике часто возникают трудоемкие итерационные петли работ, связанных с необходимостью перетрассировки плат для исключения их резонансов.

Решению вышеназванных проблем создания безрезонансных ПУ блоков БРЭС на ранних стадиях проектирования до трассировки печатных плат посвящена данная статья.

Разработка методики создания безрезонансных печатных узлов

В максимально упрощенном виде задача синтеза безрезонансных ПУ блоков БРЭС сформулирована в работе [2].

Исходные данные:

1. Диапазон частот, в котором у проектируемых пакетных ПУ не должно быть собственных частот колебаний (СЧК). Этот диапазон должен включать в себя диапазон частот возмущающих внешних вибрационных воздействий с коэффициентом запаса 1,5, т.е. $f_{СЧК} > 1,5 f_B$.

2. Размер монтажной площади S_m для размещения ЭРИ электрической схемы ПУ. Находится как сумма площадей ЭРИ, умноженная на коэффициент, учитывающий возможности трассировки платы с охранными зонами.

3. Механические характеристики материалов, используемых для изготовления ПУ.

4. На ранних стадиях проектирования влиянием на резонансную частоту ПУ массы малогабаритных ЭРИ можно пренебречь, так как оно в значительной мере компенсируется увеличением жесткости плат от закрепления ЭРИ на их поверхности жестким клеем и лаком.

5. Требуется найти варианты компоновки пакетных ПУ, у которых СЧК находятся вне диапазона частот вибрационных воздействий.

Для решения задач создания безрезонансных ПУ строятся необходимые для этого математические модели и алгоритмы синтеза на их основе [4].

Как известно, СЧК тонких пластин рассчитывается по формуле [3, 5]:

$$f_{СЧК} = A C h / S, \quad (1)$$

где f СЧК – первая СЧК платы, Гц; $A = \frac{1}{2\pi} \sqrt{\frac{E}{12(1-\nu^2)\rho}}$ – постоянный коэффициент, зависящий

от свойств материала платы; C – частотный коэффициент, зависящий от способа крепления и соотношения сторон платы, для нахождения которого можно использовать методику [2]; h – толщина платы, м; $S = a^2$ – площадь квадратной платы, где a – длина большей стороны, м; E – модуль упругости материала платы, Н/м²; ν – коэффициент Пуассона; ρ – плотность материала платы.

Ниже представлены виды крепления сторон печатных плат ПУ, расположенные в порядке увеличения массогабаритных характеристик блоков [5] (рис. 1–3).

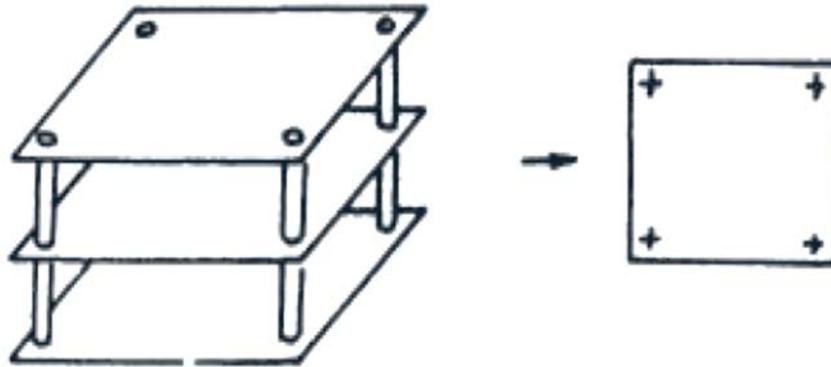


Рис. 1. Жесткое крепление в точках (обозначено крестиками) в случае сборки пакета ПУ на стойках. Края платы свободны

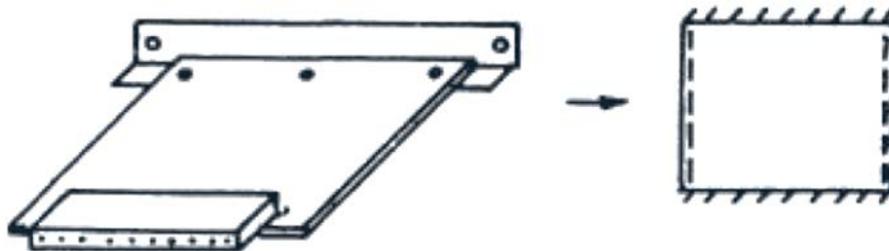


Рис. 2. Опираение сторон, обозначенное штриховой линией

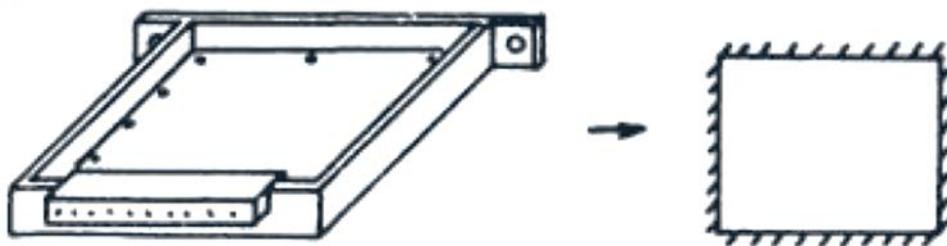


Рис. 3. Защемление края платы, обозначенное наклонной штриховкой

Обеспечивается, если край платы скользит вдоль направляющей в каркасе, а также, если край платы прикреплен к нежесткой планке или соединителю, когда край платы не может смещаться перпендикулярно плоскости платы, но может поворачиваться.

Обеспечивается, если край приклеен, приклепан или привинчен по длине в нескольких точках к жесткой рамке или если на краю платы закреплен разъемный соединитель с большим числом контактов, который при установке платы в блок соединяется с жестко закрепленной ответной частью соединителя.

Относительные значения частотных коэффициентов $C/C_{4т}$ для представленных выше видов крепления сторон плат приведены на графиках рис. 4. Здесь $C_{4т}$ – значение C для квадратной платы с креплением в четырех точках по углам.

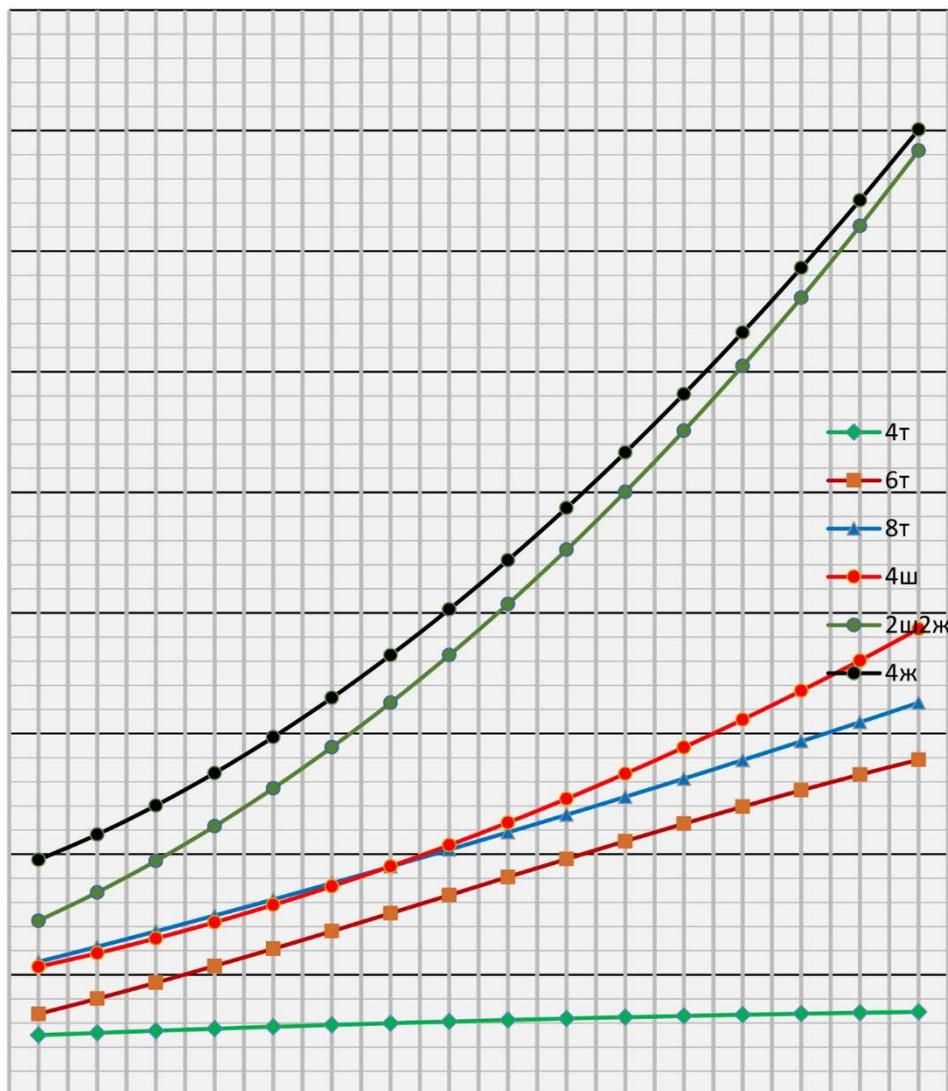


Рис. 4. Графики относительных коэффициентов C в зависимости от соотношения сторон платы a/b .

На графиках сделаны обозначения: 4т – 4 точки крепления по углам платы; 6т – 6 точек крепления, из них 4 по углам и 2 по середине длинных сторон; 8т – 8 точек крепления, из них 4 по углам и 4 по середине сторон; 4ш – шарнирное крепление всех сторон; 2ш2ж – шарнирное крепление коротких и упруго-жесткое крепление длинных сторон; 4ж – жесткое крепление всех четырех сторон

Для квадратных плат с жестким креплением четырех сторон (4ж) $C = 36$.

Графики рис. 4 позволяют сделать некоторые качественные выводы для облегчения принятий решения по синтезу компоновки безрезонансных ПУ в блоках БРЭС.

1. Самые низкие значения частотного коэффициента C и мало зависящие от соотношения сторон имеют платы, закрепленные в четырех точках по углам.

2. Шарнирное крепление всех сторон и крепление в восьми точках по периметру плат имеют практически одинаковый частотный коэффициент для всех соотношений сторон. При этом шарнирное крепление конструктивно сложнее и имеет массогабаритные характеристики блока выше, чем при креплении в точках.

3. Самые высокие значения частотного коэффициента C дает жесткое крепление сторон, но это обеспечивается за счет значительного увеличения массы и габаритов блоков БРЭС, что отрицательно влияет на характеристики их носителей.

4. При выборе конструктивного исполнения ПУ необходимо учитывать, что шарнирное крепление сторон имеет меньшую надежность, чем жесткое крепление сторон и в точках.

В принципе на ранних стадиях проектирования новых БРЭС простота формулы (1) и наглядность графиков рис. 4 позволяют опытному конструктору достаточно быстро определиться с конструктивными параметрами безрезонансного ПУ.

Так, можно сразу определить, что из рекомендованных к применению размеров плат 200×170, 170×150 и 170×75 безрезонансную конструкцию в диапазоне частот до 1000 Гц при максимальной рекомендованной толщине равной 2 мм обеспечивает только плата 170×75 как имеющая самую маленькую площадь. Остальные можно даже не рассматривать (табл. 1).

Таблица 1

Показатели	4т	6т	8т	4ш	2ш2ж	4ж
<i>C</i>	12,5	46	53	61	121	124
<i>f</i> , Гц	153	563	653	742	1479	1513

Как показывает приведенный пример и опыт конструирования БРЭС, точечное крепление плат по периметру при всей своей технологичности не позволяет обеспечить отсутствие резонансов в диапазоне частот вибрации выше 500 Гц. А размещение дополнительных точек крепления на поле платы для расширения диапазона до 1000 Гц приводит к уменьшению монтажной площади плат и усложнению сборки блоков. Варианты с жестким креплением по периметру плат обеспечивают более жесткую конструкцию, но это достигается за счет существенного увеличения массогабаритных характеристик блоков БРЭС, что существенно ограничивает возможность использования такого крепления для разработки безрезонансных ПУ. Таким образом, актуальным остается поиск конструктивных решений повышения жесткости ПУ и прежде всего с точечным креплением как превосходящим все другие способы крепления по другим важным показателям.

На практике в отдельных случаях для повышения жесткости ПУ используются двухслойные печатные платы с металлической подложкой, которая помимо всего прочего улучшает теплоотвод от ЭРИ [6]. Это позволяет добиться хороших результатов, но затруднения при сжатых сроках разработки заключаются в том, что сегодня нет удобной инженерной методики расчета параметров такой конструкции ПУ при дефиците времени на разработку новой БРЭС.

Рассмотрим один из возможных вариантов расчета основной СЧК квадратной двухслойной печатной платы с металлической подложкой с поперечным сечением, представленным на рис. 5.

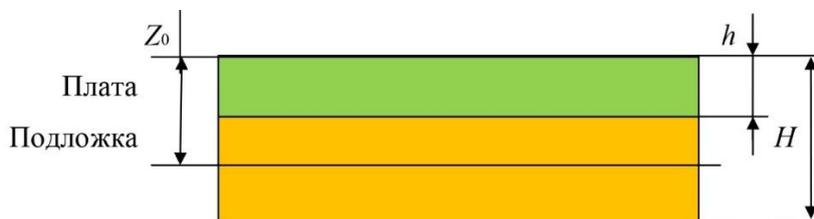


Рис. 5. Поперечное сечение печатной платы с металлической подложкой: *h* – толщина печатной платы; *H* – толщина ПУ; *Z0* – расстояние нейтральной поверхности ПУ от верхней граничной поверхности

В работе [3] приведена методика расчета СЧК многослойных печатных плат, которая на практике не очень удобна, но которой можно частично воспользоваться.

В нашем случае при расчете СЧК двухслойной печатной платы можно будет использовать расчетную толщину ПУ по формуле $H_{расч.} = (H - Z0) \cdot 2$, где *Z0* – расстояние до нейтральной поверхности в процессе изгибных колебаний двухслойного ПУ. Эта поверхность будет находиться в более жестком слое ПУ (слое подложки) и несколько смещена от оси подложки в сторону печатной платы.

Для многослойных плат приведены формулы [4] для определения *Z0*:

$$Z_0 = \frac{\sum_{i=1}^n \bar{E}_i h_i \left(2 \sum_{k=1}^i h_k - h_i \right)}{2 \sum_{i=1}^n \bar{E}_i h_i};$$

$$\bar{E}_i = E_i (1 - \nu_i^2), \nu = \frac{\sum_{i=1}^n \nu_i \bar{E}_i h_i}{\sum_{i=1}^n \bar{E}_i h_i},$$

где *E_i*, *ν_i*, *h_i* – модуль упругости, коэффициент Пуассона и толщина *i*-го слоя.

При расчете СЧК плотность пластины определяется по формуле

$$\bar{\rho} = \sum \rho_i h_i / \sum h_i,$$

где ρ_i – плотность материала i -го слоя.

Для проверки точности расчетов по предлагаемой методике выполнены расчеты СЧК по формуле (1) для ПУ в виде заземленных по всему периметру двухслойных квадратных пластин со сторонами равными 200 мм при $h = 1$ мм и различными значениями суммарной толщины $H = 2; 3; 4$ и 5 мм для вариантов подложки из дюралюминия и стали.

Механические характеристики материала слоев:

- плата (стеклотекстолит): $E = 0,33 \cdot 10^8$ кПа; $\nu = 0,27$; $\rho = 2,4$ т/м³;
- подложка (дюралюминий): $E = 0,72 \cdot 10^8$ кПа; $\nu = 0,34$; $\rho = 2,7$ т/м³;
- подложка (сталь): $E = 2,2 \cdot 10^8$ кПа; $\nu = 0,25$; $\rho = 7,8$ т/м³.

Для тех же ПУ выполнены расчеты СЧК с использованием комплекса программ ANSYS.

Результаты выполненных расчетов приведены в табл. 2.

Таблица 2

H, мм	Дюралюминий				Сталь			
	Формула	ANSYS	%	Вес, кг	Формула	ANSYS	%	Вес, кг
2	357,8	372,6	-3,9	0,204	330,2	327,7	0,8	0,408
3	560,7	568,4	-1,3	0,312	532,6	515,9	3,2	0,720
4	772,3	776,8	-0,6	0,420	741,9	726,3	2,1	1,032
5	986,5	991,1	-0,5	0,528	952,1	943,1	1,0	1,344

Из данных табл. 2 видно, что точность расчетов СЧК по предлагаемой методике находится в пределах 5 % по сравнению с результатами расчетов с использованием комплекса программ ANSYS и такая методика может быть рекомендована для инженерных расчетов.

Кроме того, из результатов расчетов видно, что стальная подложка ничего не дает и даже проигрывает дюралюминиевой по СЧК, а вес двухслойной пластины со стальной подложкой в разы больше. Поэтому варианты плат со стальной подложкой далее не рассматриваются.

Эффект от добавления дюралеминивой подложки можно оценить по табл. 3.

Таблица 3

H, мм	2	3	4	5
$f, \text{Гц}$	357,8	560,7	772,3	986,5
$f/f_{\text{пл}}$	2,2	3,5	4,8	6,2

Как видно из табл. 3, эффект от использования дюралеминивой подложки для повышения СЧК незначительный, что и объясняет, почему она используется на практике редко. Совершенно другая картина получается при использовании в качестве подложки более жесткого материала с меньшей плотностью, например, новых материалов, разработанных в ВИАМ [7], таких как углепластик с характеристиками $E = 1,25 \cdot 10^8$ кПа; $\rho = 1,5$ т/м³ и полиамид с углеродными волокнами, у которого $E = 2,38 \cdot 10^8$ кПа; $\rho = 1,34$ т/м³.

Расчеты для этих материалов приведены в табл. 4 и 5.

Таблица 4

H, мм	Дюралюминий		Углепластик		Полиамид с волокном	
	$f, \text{Гц}$	Вес, кг	$f, \text{Гц}$	Вес, кг	$f, \text{Гц}$	Вес, кг
2	357,8	0,204	452,9	0,156	559,4	0,150
3	560,7	0,312	780,2	0,216	1036,4	0,204
4	772,3	0,420	1126,7	0,276	1540,0	0,257
5	986,5	0,528	1479,9	0,336	2053,7	0,312

Таблица 5

H, мм	2	3	4	5
$f, \text{Гц}$	559,4	1036,4	1540,0	2053,7
$f/f_{\text{пл}}$	3,5	6,5	9,6	12,9

Сравнение полученных результатов табл. 5 с табл. 3 показывает, что использование новых материалов в качестве подложки дает эффект, позволяющий существенно расширить возможности создания безрезонансных ПУ БРЭС. Этот эффект можно оценить, заметив например, что при одинаковом весе (0,312 кг) полиамидная подложка при толщине ПУ 4 мм имеет СЧК почти в 4 раза больше, чем ПУ толщиной 3 мм с подложкой из дюралюминия.

Для ПУ рекомендованных минимальных размеров 170×75 при толщине печатной платы $h = 1$ мм и общей толщине ПУ с полиамидной подложкой $H = 4$ мм расчеты СЧК дают результаты, приведенные в табл. 6.

Таблица 6

a	0,17	b	0,075	a/b	2,27	
	4г	6г	8г	4ш	2ш2ж	4ж
C	12,5	46	53	61	121	124
$f, \text{Гц}$	819	3019	3503	3978	7934	8116

Здесь важно отметить, что даже для самых жестких условий эксплуатации, при верхней границе диапазона внешних возмущающих воздействий $f \text{ В} = 2000$ Гц, использование полиамидной подложки позволяет создавать безрезонансные конструкции блоков БРЭС пакетной компоновки с точечным креплением ПУ, например таких, как представлено на рис. 6.

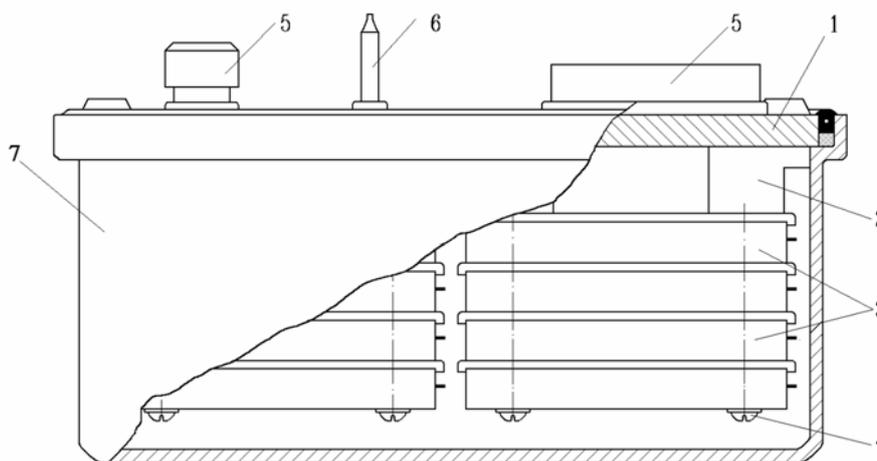


Рис. 6. Герметичный электронный блок БРЭС пакетной компоновки

Таким образом, изготовление подложек ПУ из современных материалов высокой жесткости и малой плотности, таких как углепластики, позволяет во многих случаях решить проблему вибропрочности и виброустойчивости блоков БРЭС, а значит, снизить поток отказов в процессе их эксплуатации. Кроме того, пакетная компоновка с точечным креплением таких ПУ обеспечивает высокую технологичность и возможность эксплуатации блоков БРЭС без амортизаторов и демпфирования даже в условиях жестких вибрационных воздействий. Соответственно, при прочих равных условиях более высокая надежность таких блоков достигается при наилучших массогабаритных характеристиках.

Заключение

На основе проведенных исследований предложены аналитические выражения и графики для определения на ранних стадиях проектирования основной СЧК печатных плат с жесткой подложкой для различных вариантов крепления их в электронных блоках БРЭС.

Выполненные тестовые расчеты с использованием пакета программ ANSYS подтвердили достоверность и необходимую точность результатов расчета СЧК печатных плат с подложкой для практического применения предложенных инструментов.

Предложена методика создания безрезонансных конструкций печатных узлов электронных блоков БРЭС с улучшенной технологичностью и массогабаритными характеристиками на основе использования современных материалов.

Возможность оперативного определения схем крепления и параметров безрезонансных ПУ на схемотехническом этапе и ранних этапах конструирования помимо обеспечения высокой надежности ПУ может позволить существенно уменьшить трудоемкость и сократить сроки работ за счет исключения лишних итерационных петель работ при конструировании.

Список литературы

1. Токарев М. Ф., Талицкий Е. Н., Фролов В. А. Механические воздействия и защита радиоэлектронной аппаратуры. М. : Радио и связь, 1984.
2. Фролов С. И., Кочегаров И. И., Данилова Е. А. [и др.]. Разработка методики компоновки безрезонансных блоков БРЭС на ранних стадиях проектирования // Труды Международного симпозиума Надежность и качество. 2020. Т. 1.
3. Карпушин В. Б. Вибрации и удары в радиоаппаратуре. М. : Сов. радио, 1971.
4. Юрков Н. К. Современное состояние исследований в области создания высоконадежной бортовой радиоэлектронной аппаратуры // Надежность и качество сложных систем. 2021. № 4. С. 5–12. doi: 10.21685/2307-4205-2021-4-1. EDN VBJPMK
5. Талицкий Е. Н. Защита электронных средств от механических воздействий. Владимир : ВГУ, 2001.
6. Борисов В. Ф. Проектирование РЭС. М. : МАИ, 2007.
7. Молчанов Б. И., Гудимов М. М. Свойства углепластиков и области их применения // Авиационная промышленность. 1997. № 3-4.

References

1. Tokarev M.F., Talitskiy E.N., Frolov V.A. *Mekhanicheskie vozdeystviya i zashchita radioelektronnoy apparatury = Mechanical effects and protection of radioelectronic equipment*. Moscow: Radio i svyaz', 1984. (In Russ.)
2. Frolov S.I., Kochegarov I.I., Danilova E.A. et al. Development of a methodology for the layout of non-resonant power plant units at the early stages of design. *Trudy Mezhdunarodnogo simpoziuma Nadezhnost' i kachestvo = Proceedings of the International Symposium Reliability and Quality*. 2020;1. (In Russ.)
3. Karpushin V.B. *Vibratsii i udary v radioapparature = Vibrations and shocks in radio equipment*. Moscow: Sov. radio, 1971. (In Russ.)
4. Yurkov N.K. The current state of research in the field of creating highly reliable on-board electronic equipment. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems*. 2021;(4):5–12. (In Russ.). doi: 10.21685/2307-4205-2021-4-1. EDN VBJPMK
5. Talitskiy E.N. *Zashchita elektronnykh sredstv ot mekhanicheskikh vozdeystviy = Protection of electronic devices from mechanical influences*. Vladimir: VGU, 2001. (In Russ.)
6. Borisov V.F. *Proektirovanie RES = Designing RES*. Moscow: MAI, 2007. (In Russ.)
7. Molchanov B.I., Gudimov M.M. Properties of carbon fiber plastics and their applications. *Aviatsionnaya promyshlennost' = Aviation industry*. 1997;(3-4). (In Russ.)

Информация об авторах / Information about the authors

Сергей Иванович Фролов

кандидат технических наук,
независимый исследователь
E-mail: sfrolov46@mail.ru

Евгения Анатольевна Данилова

кандидат технических наук,
доцент кафедры конструирования
и производства радиоаппаратуры,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: siori@list.ru

Игорь Иванович Кочегаров

кандидат технических наук, доцент,
доцент кафедры конструирования
и производства радиоаппаратуры,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: kipra@mail.ru

Sergey I. Frolov

Candidate of technical sciences,
independent researcher

Evgeniya A. Danilova

Candidate of technical sciences,
associate professor of the sub-department
of radio equipment design and production,
Penza State University
(40 Krasnaya street, Penza, Russia)

Igor I. Kochegarov

Candidate of technical sciences, associate professor,
associate professor of the sub-department
of radio equipment design and production,
Penza State University
(40 Krasnaya street, Penza, Russia)

Авторы заявляют об отсутствии конфликта интересов /

The authors declare no conflicts of interests.

Поступила в редакцию/Received 26.05.2024

Поступила после рецензирования/Revised 23.06.2024

Принята к публикации/Accepted 25.07.2024

МОДЕЛЬНОЕ ИССЛЕДОВАНИЕ ВЛИЯНИЯ ФОРМЫ ПЕРЕМЕННО-ФАЗНЫХ УЧАСТКОВ РАСПРЕДЕЛЕНИЯ ПОЛЯ НА ИЗЛУЧЕНИЕ АНТЕННЫ

А. Н. Якимов

Санкт-Петербургский государственный университет
аэрокосмического приборостроения, Санкт-Петербург, Россия
y_alder@mail.ru

Аннотация. *Актуальность и цели.* В радиотехнических системах различного назначения при решении задач на максимальных дальностях в широком секторе углов используются микроволновые антенны с диаграммой направленности секторного вида. Для формирования такой диаграммы направленности в антенне требуется создать переменнo-фазное распределение источников возбуждения (токов или полей). Модельное исследование влияния формы отдельных участков переменнo-фазным распределением поля в апертуре микроволновой антенны на параметры ее квазисекторной диаграммы направленности может быть проведено путем представления этого распределения комбинацией математических функций с использованием логики предикатов. *Материалы и методы.* Предложен подход к решению этой задачи, основанный на использовании свойства аддитивности определенного интеграла, позволяющего заменить его суммой определенных интегралов, взятых по частичным промежуткам, и независимо управлять распределением поля на этих промежутках. При этом появляется возможность дополнительного управления формой отдельных участков переменнo-фазного распределения поля в апертуре микроволновой антенны за счет изменения степени функций, описывающих эти участки. Показана возможность программной реализации предложенной математической модели в системе MATLAB с использованием логики предикатов. *Результаты и выводы.* Приведены результаты модельного исследования влияния формы отдельных участков переменнo-фазного распределения поля в апертуре микроволновой антенны на основные параметры ее диаграммы направленности. Полученные результаты указывают на перспективность модельного представления переменнo-фазного распределения поля в апертуре микроволновой антенны с квазисекторной ДН комбинацией тригонометрических функций произвольной степени с использованием логики предикатов для исследования влияния формы его отдельных участков на диаграмму направленности этой антенны и улучшения ее параметров.

Ключевые слова: микроволновая антенна, математическая модель, распределение поля, переменнo-фазные участки, диаграмма направленности, логика предикатов

Для цитирования: Якимов А. Н. Модельное исследование влияния формы переменнo-фазных участков распределения поля на излучение антенны // Надежность и качество сложных систем. 2024. № 3. С. 82–89. doi: 10.21685/2307-4205-2024-3-9

A MODEL STUDY OF THE INFLUENCE OF THE SHAPE OF VARIABLE-PHASE FIELD DISTRIBUTION AREAS ON ANTENNA RADIATION

A.N. Yakimov

Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia
y_alder@mail.ru

Abstract. *Background.* In radio engineering systems for various purposes, microwave antennas with a radiation pattern of a sector type are used to solve problems at maximum ranges in a wide range of angles. To form such a directional pattern in an antenna, it is necessary to create a variable-phase distribution of excitation sources (currents or fields). A model study of the influence of the shape of individual areas by the variable-phase field distribution in the aperture of a microwave antenna on the parameters of its quasi-sector radiation pattern can be carried out by representing this distribution by a combination of mathematical functions using predicate logic. *Materials and methods.* An approach to solving this problem is proposed based on the use of the additivity property of a definite integral, which allows replacing it with the sum of definite integrals taken over partial intervals and independently controlling the

distribution of the field over these intervals. At the same time, it becomes possible to additionally control the shape of individual areas of the variable-phase field distribution in the aperture of the microwave antenna by changing the degree of functions describing these areas. The possibility of software implementation of the proposed mathematical model in the MATLAB system using predicate logic is shown. *Results and conclusions.* The results of a model study of the influence of the shape of individual sections of the variable-phase field distribution in the aperture of a microwave antenna on the main parameters of its radiation pattern are presented. The results obtained indicate the prospects of a model representation of the variable-phase field distribution in the aperture of a microwave antenna with a quasi-sector radiation pattern by combination of trigonometric functions of arbitrary degree using predicate logic to study the influence of the shape of its individual areas on the radiation pattern of this antenna and improve its parameters.

Keywords: microwave antenna, mathematical model, field distribution, variable-phase areas, radiation pattern, predicate logic

For citation: Yakimov A.N. A model study of the influence of the shape of variable-phase field distribution areas on antenna radiation. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):82–89. (In Russ.). doi: 10.21685/2307-4205-2024-3-9

Введение

В радиотехнических системах различного назначения при решении задач на максимальных дальностях в широком секторе углов используются микроволновые антенны с диаграммой направленности (ДН) секторного вида. Задачу математического синтеза такой ДН относят к некорректным задачам математической физики, решение которых может быть получено лишь с некоторым приближением. Для формирования такой квазисекторной ДН в антенне требуется создать переменнo-фазное распределение источников возбуждения (токов или полей) [1, 2].

В большинстве известных математических решений, моделирующих излучение микроволновой антенны с квазисекторной ДН, параметры распределения источников возбуждения оказываются взаимозависимыми, что не позволяет провести детальное исследование связи параметров распределения источников возбуждения с параметрам ДН и получить оптимальное решение при проектировании антенн такого типа [3, 4].

Влияние основных параметров переменнo-фазного распределения источников возбуждения (максимального уровня противофазных участков (УПУ) и коэффициента, характеризующего долю их протяженности в главном сечении антенны) на параметры ДН в значительной мере уже исследовано, но влияние формы его отдельных участков на ДН антенны изучено недостаточно.

Таким образом, актуальной является задача модельного представления переменнo-фазного распределения поля в апертуре микроволновой антенны комбинацией математических функций с использованием логики предикатов для исследования влияния формы его отдельных участков на параметры антенны с квазисекторной ДН и их улучшения.

Постановка задачи

Рассмотрим апертурную микроволновую антенну с остронаправленной квазисекторной ДН, которую можно считать разделяющейся. У таких ДН горизонтальное сечение $F(\varphi)$ имеет одинаковую форму для разных фиксированных значений угла θ в вертикальной плоскости, и наоборот, вертикальное сечение $F(\theta)$ имеет одинаковую форму для разных фиксированных значений угла φ в горизонтальной плоскости. При этом возникает возможность управления формой ДН антенны за счет изменения амплитудно-фазового распределения электрической составляющей электромагнитного поля в каждой из плоскостей декартовой системы координат отдельно [5].

Математическая модель излучения такой микроволновой антенны в интегральной форме может быть представлена определенным интегралом следующим образом [3, 4]:

$$R(z) = \frac{1}{2\pi} \int_{-1}^1 f(y) e^{jzy} dy, \quad (1)$$

где $R(z)$ – функция, описывающая ДН антенны по полю в горизонтальной плоскости; $z = \pi(l/\lambda)\sin\varphi$; l – линейный размер главного сечения апертуры антенны в горизонтальной плоскости; λ – длина электромагнитной волны в свободном пространстве; φ – угол в горизонтальной плоскости между направлением на точку наблюдения и нормалью к антенне; $y = 2x/l$; x – расстояние до текущей точ-

ки главного сечения апертуры антенны относительно ее центра в горизонтальной плоскости; $f(y)$ – распределение поля в горизонтальном сечении апертуры антенны.

Свойство аддитивности определенного интеграла позволяет заменить его суммой определенных интегралов, взятых по частичным промежуткам, и независимо управлять распределением поля $f(y)$ на этих промежутках [6].

Таким образом, появляется возможность дополнительного управления формой отдельных участков переменного-фазного распределения поля в апертуре микроволновой антенны и исследовать влияние этой формы на ДН антенны. Перспективным математическим аппаратом для программной реализации решения этой задачи является логика предикатов [7, 8].

Построение математической модели

Для расчета ДН антенны в MatLAB удобно представить распределение поля $f(y)$ в смещенном виде, изменив в нем пределы интегрирования от -1 до 1 на интервал от 0 до 2 . При этом выражение (1) преобразуется к виду [9]

$$R(z) = \frac{1}{2\pi} \int_0^2 f(y) e^{jzy} dy. \quad (2)$$

Выражение (2) при этом может быть представлено как

$$R(z) = \int_0^2 f(y) e^{jyz} dy = \int_0^2 f(y) \cos yz dy + j \int_0^2 f(y) \sin yz dy. \quad (3)$$

Приняв обозначения

$$F_1(y) = f(y) \cos yz, \quad F_2(y) = f(y) \sin yz,$$

получим

$$R(z) = \int_0^2 F_1(y) dy + j \int_0^2 F_2(y) dy = A + iB = |R_f(z)| e^{i\Phi(z)}, \quad (4)$$

где $A = \int_0^2 F_1(y) dy$, $B = \int_0^2 F_2(y) dy$, $|R(z)| = \sqrt{A^2 + B^2}$ – искомая амплитудная ДН; $\Phi(z)$ – фазовая ДН.

Распределение поля $f(y)$ представим в виде трех независимых частей, описываемых тригонометрическими функциями. При этом центральный участок распределения поля от $y_{01} = 1 - (1/\alpha_0)$ до $y_{02} = 2 - y_{01}$ опишется выражением

$$f_1(y) = A_1 \cdot \cos^m \left[\frac{\pi \alpha_0}{2} (y - 1) \right], \quad (5)$$

где A_1 – коэффициент амплитуды главного лепестка распределения; m – произвольная (целая или дробная) степень; $\alpha_0 = 1/y_0$ – параметр, характеризующий долю противофазных участков в распределении поля; y_0 – координата точки смены фазы в сечении амплитудного распределения.

Первый противофазный участок распределения $f(y)$ на интервале от 0 до y_{01} опишется выражением

$$f_{21}(y) = -A_2 \cdot |\sin [\pi \alpha_0 (1 - y)]|^p, \quad (6)$$

где A_2 – коэффициент амплитуды противофазных участков распределения, задающий максимальный уровень противофазного участка (УПУ); p – произвольная степень.

Второй противофазный участок распределения $f(y)$, расположенный на интервале от y_{02} до 2 , опишется выражением

$$f_{22}(y) = -A_2 |\sin [\pi \alpha_0 (y_{02} - y)]|^p. \quad (7)$$

Для алгоритмического объединения выражений (5)–(7) в единую математическую модель удобно использовать логику предикатов. При этом обязательным является выполнение условия комплементарности, в соответствии с которым [7, 8]

$$a_1 + a_2 + \dots + a_n = 1, \quad a_i \in \{0,1\}, \quad i = 1, 2, \dots, n. \quad (8)$$

Здесь весовые коэффициенты a_i являются двузначными предикатами, а элементарными функциями, воспроизводящими операции выбора одной из двух переменных, являются предикатные конъюнкция и дизъюнкция.

Тогда операция суперпозиции элементов вектора предметных переменных $U = (u_1, u_2, \dots, u_n)$ с этими весовыми коэффициентами позволит подключать их независимо и получить результат G :

$$G = a_1 u_1 + a_2 u_2 + \dots + a_n u_n. \quad (9)$$

Предметными переменными в решаемой задаче являются функции $f_1(y)$, $f_{21}(y)$ и $f_{22}(y)$, описывающие отдельные участки переменного-фазного распределения поля, т.е. $n = 3$. Условия комплементарности при этом описываются как

$$a_1 + a_2 + a_3 = 1, \quad (10)$$

а весовые коэффициенты в виде удобном для программной реализации в системе MATLAB определены следующим образом:

$$\begin{cases} a_1 = (y \geq 0) \& (y < y_{01}), & a_2 = (y \geq y_{01}) \& (y \leq y_{02}), \\ a_3 = (y > y_{02}) \& (y \leq 2). \end{cases} \quad (11)$$

В результате суперпозиции в соответствии с формулой (9) получим

$$f(y) = a_1 f_{21}(y) + a_2 f_1(y) + a_3 f_{22}(y). \quad (12)$$

Таким образом, выражения (4)–(7), (11), (12) представляют собой математическую модель излучения микроволновой антенны с квазисекторной ДН и позволяют провести необходимые исследования.

Результаты моделирования

Расчеты, проведенные в MatLAB с использованием предложенной математической модели для $l/\lambda = 27$, $\alpha_0 = 1,9$, $A_1 = 1$ и $A_2 = 0,1$, позволили оценить влияния формы отдельных участков переменного-фазного распределения поля $f(y)$ в апертуре микроволновой антенны на основные параметры ее ДН [10].

Управление формой центрального участка распределения поля $f(y)$ за счет изменения степени m дает следующие результаты. Исходное распределение поля $f(y)$ в апертуре антенны при $p = 1$ и $m = 1$ (рис. 1, кривая 1) формирует ДН $R(\varphi)$ (рис. 2, кривая 1) шириной на уровне половинной мощности $2\varphi_{0,5} \approx 6,3$ град с малой пульсацией (провалом) вершины главного лепестка $|\Delta| \approx 0,03$ дБ и максимальным уровнем боковых лепестков (УБЛ) $R_{MSL} \approx -24,8$ дБ.

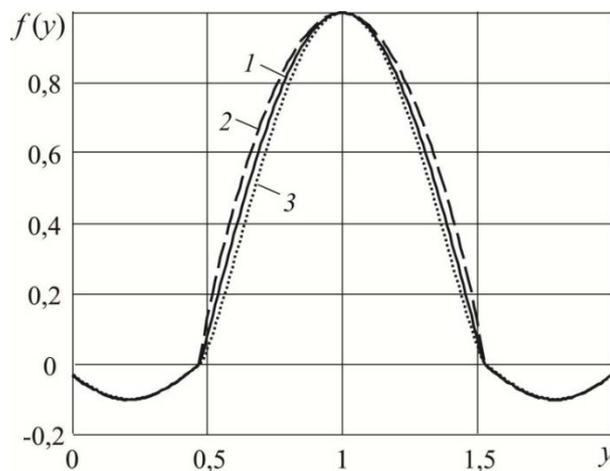


Рис. 1. Распределение поля в апертуре антенны $f(y)$ при $p = 1$:
 $1 - m = 1$; $2 - m = 0,8$; $3 - m = 1,2$

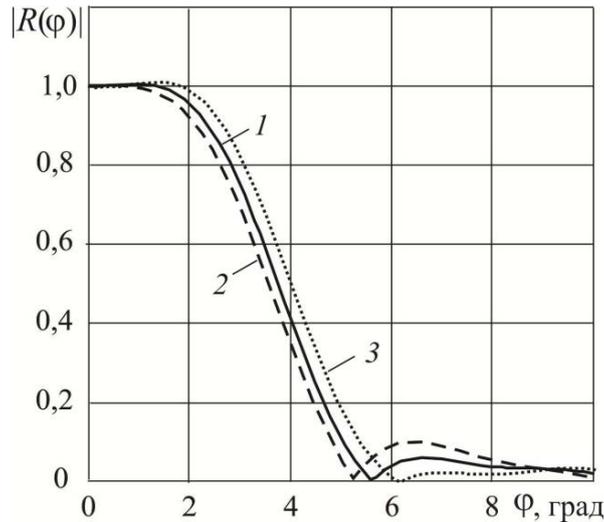


Рис. 2. Диаграммы направленности антенны $|R(\varphi)|$ при $p = 1$:
 1 – $m = 1$; 2 – $m = 0,8$; 3 – $m = 1,2$

Уменьшение степени m функции, описывающей центральный участок распределения поля $f(y)$, до $m = 0,8$ (рис. 1, кривая 2) приводит (рис. 2, кривая 2) к уменьшению пульсации до $|\Delta| = 0$ дБ, сужению ДН $R(\varphi)$ до $2\varphi_{0,5} \approx 6,0$ град и росту УБЛ до $R_{MSL} \approx -20,6$ дБ. Увеличение степени m функции, описывающей центральный участок распределения поля $f(y)$, до $m = 1,2$ (рис. 1, кривая 3) приводит (рис. 2, кривая 3) к увеличению пульсации до $|\Delta| \approx 0,11$ дБ, расширению ДН $R(\varphi)$ до $2\varphi_{0,5} \approx 6,5$ град и снижению УБЛ до $R_{MSL} \approx -29,2$ дБ.

Управление формой противофазных участков распределения поля $f(y)$ относительно исходного распределения поля $f(y)$ при $m = 1$ и $p = 1$ (рис. 3, кривая 1) и соответствующей ДН $R(\varphi)$ (рис. 4, кривая 1) за счет изменения степени p дает следующие результаты.

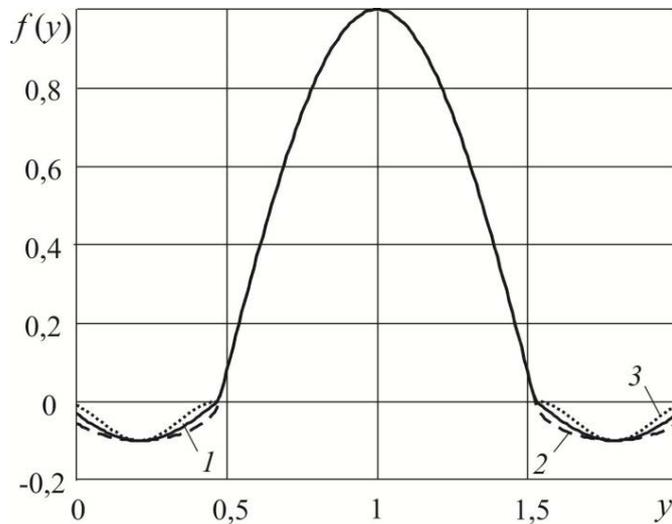


Рис. 3. Распределение поля в апертуре антенны $f(y)$ при $m = 1$:
 1 – $p = 1$; 2 – $p = 0,5$; 3 – $p = 2$

Лишь значительное уменьшение степени p функций, описывающих противофазные участки распределения поля $f(y)$, до $p = 0,5$ (рис. 3, кривая 2) приводит (рис. 4, кривая 2) к заметному увеличению пульсации вершины до $|\Delta| \approx 0,12$ дБ, расширению ДН $R(\varphi)$ до $2\varphi_{0,5} \approx 6,5$ град и незначительному росту УБЛ до $R_{MSL} \approx -24,4$ дБ. Увеличение степени p функций, описывающей противофазные участки распределения поля $f(y)$, до $p = 2$ (рис. 3, кривая 3) приводит (рис. 4, кривая 3) к уменьшению пульсации до $|\Delta| = 0$ дБ, сужению ДН $R(\varphi)$ до $2\varphi_{0,5} \approx 6,1$ град и снижению УБЛ до $R_{MSL} \approx -25,3$ дБ.

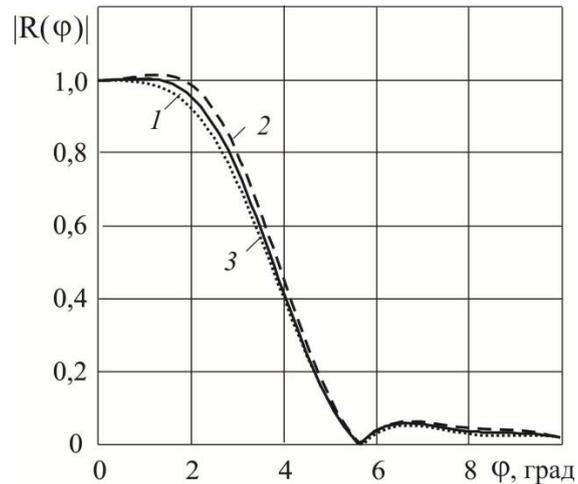


Рис. 4. Диаграммы направленности антенны $|R(\varphi)|$ при $m = 1$:
 $1 - p = 1$; $2 - p = 0,5$; $3 - p = 2$

Результаты модельного исследования указывают на то, что изменения формы центрального участка распределения поля $f(y)$ за счет изменения степени m оказывают более существенное влияние на параметры ДН микроволновой антенны, чем изменение формой противофазных участков этого распределения поля. Выявлены следующие тенденции изменения значений параметров ДН: с уменьшением m относительно значения $m = 1$ ширина $2\varphi_{0,5}$ и величина пульсации вершины ДН $|\Delta|$ уменьшаются, а УБЛ R_{MSL} растет; с увеличением m относительно значения $m = 1$ ширина $2\varphi_{0,5}$ и величина пульсаций вершины ДН $|\Delta|$ увеличиваются, а УБЛ R_{MSL} уменьшается. Изменение формы противофазных участков распределения поля выявило противоположную тенденцию: с уменьшением p относительно значения $p = 1$ ширина $2\varphi_{0,5}$ и величина пульсации вершины ДН $|\Delta|$ увеличиваются более интенсивно, чем УБЛ R_{MSL} ; с увеличением m относительно значения $m = 1$ ширина $2\varphi_{0,5}$ и величина пульсации вершины ДН $|\Delta|$ и УБЛ R_{MSL} уменьшаются.

Дополнительное модельное исследование позволило выявить следующие зависимости параметров квазисекторной ДН $R(\varphi)$ от степени m , оказывающей значительное большее влияние на ДН, чем степень p (рис. 5). В заданном интервале изменения m от 0,6 до 1,6 отличия расчетной ДН от секторной оказываются допустимыми.

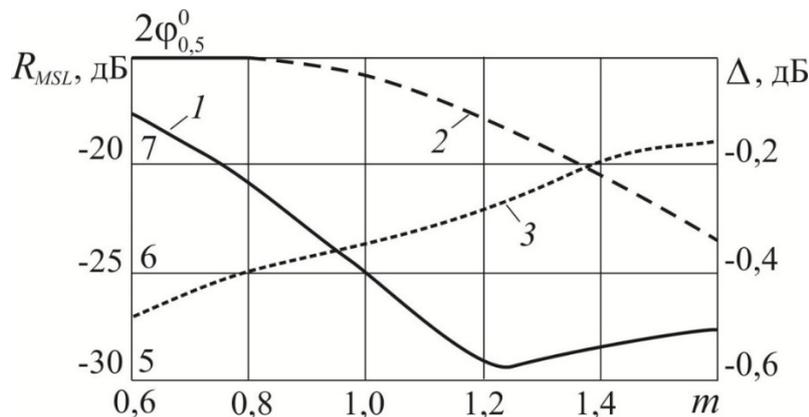


Рис. 5. Зависимости параметров квазисекторной ДН от m при $p = 1$:
 $1 - R_{MSL}$; $2 - \Delta$; $3 - 2\varphi_{0,5}$

За пределами указанного интервала отличия расчетной ДН от секторной становятся уже значительными: при $m < 0,6$ вершина главного лепестка ДН становится острой, а при $m > 1,6$ растет пульсация этого лепестка $|\Delta|$ и искажается его форма.

При этом видно, что УБЛ R_{MSL} (рис. 5, кривая 1) в заданном интервале изменения m имеет минимум, а величина пульсации $|\Delta|$ (рис. 5, кривая 2) и ширина ДН $2\varphi_{0,5}$ (рис. 5, кривая 3) увеличива-

ются, что указывает на возможность оптимизации такой микроволновой антенны с использованием предложенной математической модели.

Таким образом, полученные результаты указывают на перспективность модельного представления переменного-фазного распределения поля в апертуре микроволновой антенны с квазисекторной ДН комбинацией математических функций с использованием логики предикатов для исследования влияния формы его отдельных участков на характеристику направленности этой антенны и улучшения ее параметров. Они подтверждают возможность использования построенной математической модели в проектировании микроволновых антенн с переменного-фазным распределением поля совместно с другими моделями, разработанными ранее [4, 10, 11].

Заключение

Предложенный подход к решению задачи математического моделирования излучения микроволновой антенны с квазисекторной ДН, основанный на сочетании аналитических и численных методов, позволяет описать отдельные участки рассматриваемого переменного-фазного распределения поля тригонометрическими функциями произвольной степени, изменение которой дает возможность управлять формой этих участков. Использование логики предикатов в программной реализации предложенной математической модели значительно облегчает решение задачи моделирования и весьма перспективно.

Полученные результаты модельного исследования влияния формы отдельных участков переменного-фазного распределения поля в апертуре микроволновой антенны на основные параметры ее ДН указывают на возможность дополнительного улучшения этих параметров. Независимое управление формой отдельных участков распределения поля целесообразно использовать при проектировании микроволновых антенн для радиотехнических систем различного назначения.

Список литературы

1. Воскресенский Д. И., Гостюхин В. Л., Максимов В. М., Пономарев Л. И. Устройства СВЧ и антенны / под ред. Д. И. Воскресенского. М. : Радиотехника, 2006. 376 с.
2. Кюн Р. Микроволновые антенны. Л. : Судостроение, 1967. 518 с.
3. Минкович Б. М., Яковлев В. П. Теория синтеза антенн. М. : Сов. радио, 1969. 296 с.
4. Якимов А. Н. Условия формирования секторной диаграммы направленности с минимальным уровнем боковых лепестков // Вопросы радиоэлектроники. Сер.: ОТ. 1980. Вып. 4. С. 79–80.
5. Дабкин А. Л., Зузенко В. Л., Кислов А. Г. Антенно-фидерные устройства. М. : Сов. радио, 1974. 536 с.
6. Дмитриев В. И., Захаров Е. В. Метод интегральных уравнений в вычислительной электродинамике. М. : МАКС Пресс, 2008. 316 с.
7. Волгин Л. И. Элементарный базис комплементарной алгебры: комплементарный релятор // Проектирование и технология электронных средств. 2001. № 1. С. 10–11.
8. Якимов А. Н. Предикатная алгебра выбора в моделировании антенн сложной конфигурации // Надежность и качество сложных систем. 2013. № 1. С. 78–83.
9. Дьяконов В. П., Абраменкова И. В., Круглов В. В. MatLAB 5.3.1 с пакетами расширений / под ред. В. П. Дьяконова. М. : Нолидж, 2001. 880 с.
10. Якимов А. Н. Методы определения параметров при расчете диаграммы направленности антенны // Труды Международного симпозиума Надежность и качество. 2014. Т. 1. С. 74–75.
11. Якимов А. Н. Универсальная модель излучения микроволновой антенны // Надежность и качество сложных систем. 2024. № 1. С. 5–12.

References

1. Voskresenskiy D.I., Gostyukhin V.L., Maksimov V.M., Ponomarev L.I. *Ustroystva SVCh i anteny = Microwave devices and antennas*. Moscow: Radiotekhnika, 2006:376. (In Russ.)
2. Kyun R. *Mikrovolnovye anteny = Microwave antennas*. Leningrad: Sudostroenie, 1967:518. (In Russ.)
3. Minkovich B.M., Yakovlev V.P. *Teoriya sinteza antenn = Theory of antenna synthesis*. Moscow: Sov. radio, 1969:296. (In Russ.)
4. Yakimov A.N. Conditions for the formation of a sectoral directional pattern with a minimum level of side lobes. *Voprosy radioelektroniki. Ser.: OT = Radio electronics issues. Series: OT*. 1980;(4):79–80. (In Russ.)
5. Drabkin A.L., Zuzenko V.L., Kislov A.G. *Antenno-fidernye ustroystva = Antenna-feeder devices*. Moscow: Sov. radio, 1974:536. (In Russ.)
6. Dmitriev V.I., Zakharov E.V. *Metod integral'nykh uravneniy v vychislitel'noy elektrodinamike = Method of integral equations in computational electrodynamics*. Moscow: MAKS Press, 2008:316. (In Russ.)

7. Volgin L.I. Elementary basis of complementary algebra: complementary relator. *Proektirovaniye i tekhnologiya elektronnykh sredstv = Design and technology of electronic means*. 2001;(1):10–11. (In Russ.)
8. Yakimov A.N. Predicate algebra of choice in modeling antennas of complex configuration. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems*. 2013;(1):78–83. (In Russ.)
9. D'yakonov V.P., Abramenkova I.V., Kruglov V.V. *MatLAB 5.3.1 s paketami rasshireniy = MatLAB 5.3.1 with extension packages*. Moscow: Nolidzh, 2001:880. (In Russ.)
10. Yakimov A.N. Methods for determining parameters when calculating the antenna radiation pattern. *Trudy Mezhdunarodnogo simpoziuma Nadezhnost' i kachestvo = Proceedings of the International Symposium Reliability and Quality*. 2014;1:74–75. (In Russ.)
11. Yakimov A.N. Universal model of microwave antenna radiation. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems*. 2024;(1):5–12. (In Russ.)

Информация об авторах / Information about the authors

Александр Николаевич Якимов

доктор технических наук, профессор,
профессор кафедры конструирования
и технологий электронных и лазерных средств,
Санкт-Петербургский государственный университет
аэрокосмического приборостроения
(Россия, Санкт-Петербург, ул. Большая Морская, 67)
E-mail: y_alder@mail.ru

Aleksandr N. Yakimov

Doctor of technical sciences, professor,
professor of the sub-department of construction
and technology of electronic and laser equipment,
Saint Petersburg State University
of Aerospace Instrumentation
(67 Bolshaya Morskaya street, Saint Petersburg, Russia)

Автор заявляет об отсутствии конфликта интересов /

The author declares no conflicts of interests.

Поступила в редакцию / Received 30.05.2024

Поступила после рецензирования / Revised 25.06.2024

Принята к публикации / Accepted 18.08.2024

ДИАГНОСТИЧЕСКИЕ МЕТОДЫ ОБЕСПЕЧЕНИЯ НАДЕЖНОСТИ И КАЧЕСТВА СЛОЖНЫХ СИСТЕМ

DIAGNOSTIC METHODS FOR ENSURING RELIABILITY AND QUALITY OF COMPLEX SYSTEMS

УДК 621.371.334:537.874.6

doi: 10.21685/2307-4205-2024-3-10

ЭЛЕКТРОННО-УПРАВЛЯЕМЫЕ НЕЛИНЕЙНЫЕ УСТРОЙСТВА ДАЛЬНОГО И СРЕДНЕГО ИНФРАКРАСНОГО ДИАПАЗОНА НА ПЛАТФОРМЕ ГРАФЕНОВОЙ ПЛАЗМОНИКИ. ОБЗОР

Г. С. Макеева

Пензенский государственный университет, Пенза, Россия
radiotech@pnzgu.ru

Аннотация. *Актуальность и цели.* Актуальность тематики обусловлена необходимостью освоения инфракрасного (ИК) диапазона для оптических телекоммуникационных систем, ИК-спектроскопии и приложений биофотоники и диктуется потребностями нелинейной ИК-оптики в реализации активных оптических функций на основе графена, таких как генерация ИК-излучения, преобразование частоты для применений в оптической связи, обработке материалов, точных измерениях, спектроскопическом зондировании и биологических сенсорах. Цель данной работы – представить принципы построения и функционирования электрически перестраиваемых нелинейных устройств дальнего и среднего ИК-диапазона: генераторов гармоник и умножителей, преобразователей частоты и смесителей на основе графеновых емкостей, основываясь на новейших достижениях графеновой плазмоники. *Материалы и методы.* Графеновая плазмоника. *Результаты и выводы.* Уникальные свойства делают электронно-управляемые нелинейные устройства дальнего и среднего ИК-диапазона перспективным выбором для будущего нелинейного преобразователя частоты, оптического управления и обработки информации на платформе графеновой плазмоники.

Ключевые слова: инфракрасный диапазон, генераторы гармоник, преобразователи частоты, метаповерхность, графеновая плазмоника

Для цитирования: Макеева Г. С. Электронно-управляемые нелинейные устройства дальнего и среднего инфракрасного диапазона на платформе графеновой плазмоники. Обзор // Надежность и качество сложных систем. 2024. № 3. С. 90–103. doi: 10.21685/2307-4205-2024-3-10

ELECTRONICALLY CONTROLLED NONLINEAR DEVICES BASED ON THE GRAPHENE PLASMONICS PLATFORM IN THE FAR AND MID-INFRARED RANGE. REVIEW

G.S. Makeeva

Penza State University, Penza, Russia
radiotech@pnzgu.ru

Abstract. *Background.* The relevance of the topic is due to the need to master the infrared (IR) range for optical telecommunication systems, IR spectroscopy and biophotonics applications and is dictated by the needs of nonlinear IR optics for the implementation of active optical functions based on graphene, such as generation of IR radiation, frequency conversion for applications in optical communications, materials processing, precision measurements, spectroscopic sensing and biological sensors. The purpose of this work is to present the principles of construction and operation of electrically tunable nonlinear devices in the far and mid-IR range: harmonic generators and multipliers, frequency converters and mixers based on graphene metasurfaces using the latest achievements of graphene plasmonics. *Materials and methods.* Graphene plasmonics. *Results and conclusions.* The unique properties make electronically controlled nonlinear devices in the far and medium IR range a promising choice for the future nonlinear frequency converter, optical control and information processing on a graphene plasmonics board.

Keywords: infrared range, harmonic generators, frequency converters, metasurface, graphene plasmonics

For citation: Makeeva G.S. Electronically controlled nonlinear devices based on the graphene plasmonics platform in the far and mid-infrared range. Review. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):90–103. (In Russ.). doi: 10.21685/2307-4205-2024-3-10

Введение

Освоение инфракрасного (ИК) диапазона актуально для оптических телекоммуникационных систем, ИК-спектроскопии и приложений биофотоники [1]. В том числе ИК-визуализация является важным методом во множестве приложений, включая ночное видение, автономную навигацию транспортных средств, оптическую томографию и контроль качества пищевых продуктов [1, 2]. Исследования в ИК-диапазоне имеют большой потенциал для многих областей применения, таких как оптическая беспроводная связь, химическое и биомолекулярное зондирование [3].

ИК-излучение находит применение в ряде областей, включая спектроскопию, медицину и безопасность. Кроме того, недавние разработки в ИК-технологиях привели к появлению нелинейной ИК-оптики, где интенсивные электромагнитные поля перспективны для исследования сверхбыстродействующих электронных явлений и для изучения новых оптических нелинейностей [4].

Нелинейная оптика является важной областью исследований в ИК-диапазоне для реализации активных оптических функций, таких как генерация ИК-излучения, преобразование частоты и сверхбыстродействующее оптическое переключение для применений в оптической связи, обработке материалов, точных измерениях, спектроскопическом зондировании и биологических сенсорах [5].

С момента первого получения в 2004 г. графен (одноатомный слой углерода, упакованный в гексагональную решетку) привлек внимание благодаря своим замечательным электрическим, оптическим и физическим свойствам [1]. Исследован необычайно большой нелинейно-оптический отклик графена. Нелинейные процессы третьего порядка, такие как оптический эффект Керра, генерация третьей гармоники и четырехволновое смешение (ЧВС) [6], экспериментально проверены на тонких пленках графена с одним или несколькими слоями [7, 8]. Нелинейные процессы второго порядка, такие как генерация разностной частоты, обычно считаются запрещенными в изотропной среде, но возможны в графене из-за явно нелокального пространственного характера взаимодействия [1].

Графен является выдающимся плазмонным материалом из-за его низких потерь и способности поддерживать перестраиваемые поверхностные плазмон-поляритоны (ППП), в особенности в графеновых нанолентах (ГНЛ) и других ограниченных морфологиях в ТГц и среднем ИК-диапазонах [4]. Благодаря значительному увеличению поля ППП в графене могут усиливать присущую графену высокую нелинейность. Поскольку графен обладает уникальным свойством сверхвысокой нелинейности, особенно нелинейности третьего порядка, в широком спектральном диапазоне графен естественным образом адаптируется к нелинейным процессам преобразования частоты [9–13]. Графеновая плазмоника обеспечивает уникальную и превосходную платформу для типичных функций нелинейной оптической обработки сигналов в дальнем, среднем и ближнем ИК-диапазонах благодаря высокой нелинейной проводимости графена и жесткой локализации поля ППП [10].

Графеновая плазмоника возникла как платформа для реализации сильного взаимодействия излучения с материалами на сверхмалых масштабах длины (наноуровне) [11]. Возбуждение, детектирование и манипулирование ППП в графене с помощью зондов ближнего поля продемонстрированы на инфракрасных и терагерцевых частотах [11]. Высокая концентрация электрического поля как ключевой фактор нелинейных оптических явлений достигнута за счет возбуждения ППП с длиной волны всего несколько сотен нанометров, что примерно на два порядка короче длины электромагнитной волны в вакууме на этих же частотах [11].

ППП в графене обеспечивают стратегию для передовых ИК-технологий благодаря жесткому ограничению, сверхбыстродействующему отклику и возможности перестройки. Последние дости-

жения в области полностью оптической генерации ППП в графеновых планарных волноводах предлагают многообещающий метод высокоскоростной обработки сигналов в наноразмерных интегрированных оптоэлектронных устройствах [12].

Графен стал материалом, способным расширить область нелинейной плазмоники не только до ТГц, но и ИК-диапазона длин волн благодаря плазмонному отклику, исключительно высокой нелинейности третьего порядка, низким потерям в этих диапазонах, однослойной 2D-структуре, которая уменьшает объем взаимодействия излучения с веществом до чрезвычайно малых субволновых масштабов [14, 15]. ППП, возбуждаемые в графене, обеспечивают беспрецедентные уровни локализации излучения. Плазмонные резонансы в графене могут динамически перестраиваться с помощью электростатического смещения и позволяют создавать новое поколение реконфигурируемых плазмонных устройств, в частности ИК-биосенсоров, модуляторов, метаповерхностей и идеальных поглотителей [14]. Благодаря уникальному ограничению и частотной перестраиваемости графеновая плазмоника подходит для многих применений, включая ИК-детектирование, высокоэффективное ИК-поглощение, метаматериалы и химическое зондирование. В частности, графеновая плазмоника продемонстрировала большой потенциал в среднем ИК-диапазоне, где позволяет с высокой чувствительностью и избирательностью обнаруживать, например, молекулярные частицы по их колебательным спектрам и тепловому излучению [16].

Графен является самым нелинейным оптоэлектронным материалом из известных. Недавние эксперименты по терагерцевой, инфракрасной генерации высших гармоник в графене обеспечили непосредственное измерение коэффициентов нелинейности вплоть до седьмого порядка [17]. Обнаружено, что эти нелинейные коэффициенты на много порядков превышают соответствующие коэффициенты всех других известных электронных материалов [17]. Показано, что нелинейность графена как для сверхкоротких однопериодных, так и для квазимонохроматических многопериодных входных ТГц-, ИК-сигналов может эффективно управляться и контролироваться с помощью электрического стробирования с напряжением стробирования всего несколько вольт [17]. Показано, что нелинейность третьего порядка графена исключительно велика и может изменяться на порядки с помощью контролируемого легирования или сдвига химического потенциала [18]. Возникающая в результате гигантская нелинейность графена может найти большое применение в будущих оптоэлектронных устройствах. Действительно, в графене наблюдаются такие нелинейные эффекты, как генерация третьей гармоники, ЧВС, оптический эффект Керра, фазовая самомодуляция, двухчастотная когерентная оптическая инжекция тока и даже генерация высших гармоник [18].

В настоящее время продолжающиеся исследования нелинейных явлений в 2D материалах в основном сосредоточены на генерации гармоник. ЧВС в ближнем ИК-диапазоне осуществлено в монослое графена [19]. Недавно еще большее значение нелинейной восприимчивости третьего порядка получено в ГНЛ на частотах среднего ИК-диапазона, близких к поперечному плазмонному резонансу [5, 20].

Новой темой в исследованиях нелинейной оптики является реализация высокоэффективных оптических функций в сверхмалых субволновых структурах путем использования интересных оптических резонансов в метаповерхностях [21]. Усовершенствованные оптические элементы, состоящие из графеновых метаповерхностей, функционирующих в телекоммуникационном ИК-диапазоне, обещают уменьшить занимаемую площадь и раскрыть новые функциональные возможности оптических ИК-устройств [22, 23]. Результаты эксперимента, в котором ППП сильно влияют на нелинейные свойства графеновых метаповерхностей, подчеркивают потенциал графеновой плазмоники для разработки модуляторов, мультиплексоров, световой логики, которые можно использовать в оптических межсоединениях [20].

Метаповерхности открыли беспрецедентные возможности для необычных процессов волновой инженерии [24]. В нелинейных поляритонных метаповерхностях – субволновых структурах – эффективность, спектральный диапазон и локальная фаза при нелинейном смещении волн не зависят от фазового согласования, а только от нелинейного отклика составляющих метаатомов [25].

Визуализация на основе конверсии с повышением частоты – это успешный инновационный метод ИК-визуализации с точки зрения чувствительности, скорости и шумовых характеристик [26–31]. В этом подходе ИК-изображение преобразуется с повышением частоты для формирования видимого/ближнего ИК-изображения посредством параметрического трехволнового смещения с последующим обнаружением с помощью детектора или камеры на основе кремния [27]. Продemonстрировано усиление ЧВС в компактном микрорезонаторе из кремния и графена в резонаторе [28]. Представлены два нелинейных многоканальных компонента на основе графена [32], включающие усиленный графеном кремниевый дисковый резонатор для полностью оптической маршрутизации

в ближнем ИК вместе с графеновым микрополосковым резонатором для ЧВС в дальнем ИК-диапазоне [29, 30]. Показано, что значительная эффективность квантового преобразования СВЧ в оптический диапазон в многослойном графене достигается для микровольтовых СВЧ-возбуждающих напряжений [31, 33].

Принципы построения нелинейных устройств преобразования частоты среднего ИК-диапазона

Нелинейные оптические материалы используются для преобразования когерентного монохроматического излучения в другие требуемые длины волн в ультрафиолетовом, видимом, ИК- и ТГц-диапазонах. Преобразование происходит посредством ряда нелинейных процессов преобразования с повышением и понижением частоты, таких как генерация гармоник и генерация комбинационных частот [34].

Преобразование частоты с помощью нелинейных электронных компонентов – распространенный подход к обработке сигналов, необходимый в различных приложениях связи, выявил узкое место в рабочей полосе частот в ИК-диапазоне из-за ограниченной подвижности носителей заряда в традиционных материалах. Между тем волоконно-оптическая связь играет значительную роль в услугах связи благодаря своим превосходным характеристикам передачи сигналов. Однако передаваемые оптические сигналы должны быть преобразованы в электрические сигналы с помощью фотодетекторов до того, как преобразование частоты выполнено с помощью обычных электронных устройств, что делает эту систему преобразования очень сложной и дорогостоящей. Следовательно, разработка компактных оптических устройств, которые могут одновременно выполнять функции и преобразования частоты и фотодетектирования, имеет большое значение [35].

Полностью оптические преобразователи длин волн, использующие сверхбыстрые нелинейные эффекты, являются важными компонентами оптических сетей связи из-за их преимуществ высокой скорости, прозрачности, скорости передачи данных, низкого чирпа по сравнению с современными электронными системами обработки сигналов [3].

Разработка универсальных смесителей частоты для оптических частот остается сложной задачей: такие устройства обычно основаны на слабых нелинейных оптических процессах и, следовательно, должны удовлетворять условиям фазового синхронизма [36, 37.]. Хотя эффективность преобразования для оптических метасмесителей низкая, показано, что эффективность может быть значительно улучшена путем наложения нескольких слоев метаматериала [38].

Преобразователи частоты среднего ИК-диапазона на основе графеновых метаповерхностей

Нелокальная оптическая нелинейность, контролируемая ППП в структурированной графеновой среде, где наблюдается эффект электропоглощения, генерирует эффективное нелинейное смешение волн в среднем ИК-диапазоне, которое зависит от легирования графена, геометрии структуры, температуры и может быть на порядок больше, чем на обычной локальной оптической нелинейности графена. На рис. 1 схематически иллюстрируется принцип нелинейного смешения волн на основе графеновой метаповерхности в среднем ИК-диапазоне [20]. Волна сигнала, падающая на графеновую структуру (рис. 1,а), возбуждает колебания электронов в графене и поверхностные электрические токи (рис. 1,б), достаточно сильные для поля, находящегося в резонансе с «локализованным» ППП в ГНЛ [39].

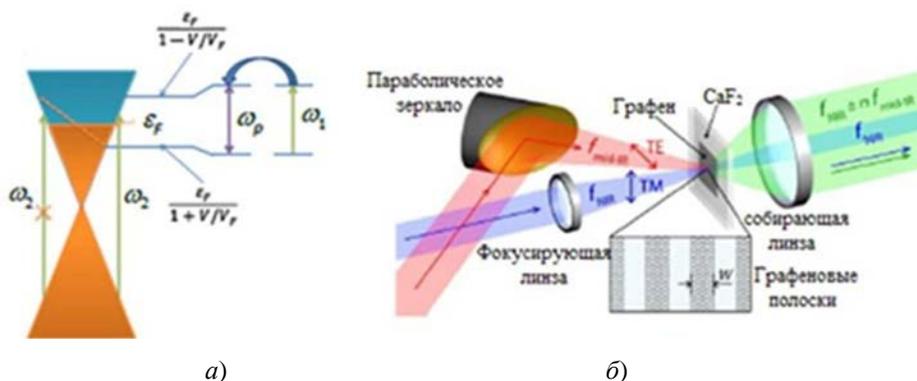


Рис. 1. Изменение поглощения волн графеном вследствие возбуждения ППП в ГНЛ (а) и принцип нелинейного смешения волн на основе графеновой метаповерхности в среднем ИК-диапазоне (б) [20]

Когда волна сигнала с частотой ω_2 проходит через ГНЛ, она частично поглощается при условии, что энергия фотона с частотой ω_2 более чем в два раза больше энергии Ферми (рис. 1,а). Если на ГНЛ падает волна накачки с частотой ω_1 (рис. 1,б), близкой к плазмонной частоте ω_p , токи в графене изменяют уровень Ферми (пунктирная желтая линия), при этом некоторые квантовые переходы становятся запрещенными (согласно принципу запрета Паули) (рис. 1,а).

В спектральной области квантовые переходы могут быть заблокированы или разблокированы, при этом рабочая частота определяется следующим образом [20]:

$$\frac{2E_F/h}{1+v_y/v_F} \ll \omega \ll \frac{2E_F/h}{1-v_y/v_F},$$

где v_y – скорость дрейфа, пропорциональная индуцированному току; E_F – уровень Ферми.

Поскольку ток и, следовательно, дрейфовая скорость периодически изменяются волной накачки, то поглощение волны сигнала модулируется с частотой волны накачки ω_1 , что приводит к эффективному нелинейному нелокальному смещению волн в среднем ИК-диапазоне [20].

Модель, позволяющая получить улучшенное нелинейное преобразование волн среднего ИК-диапазона через генерацию разностной частоты на основе нелинейности графенового листа, расположенного на подложке с металлической дифракционной решеткой (МДР) из полосок золота [1], представлена на рис. 2. Модель позволяет объединить ППП в благородных металлах и ППП в графене и предназначена для преобразования частоты волн среднего ИК-диапазона, улучшенного плазмонными резонансами высших мод ППП [1]. Слой графена наносится на подложку с МДР (периодом p) из полосок золота (толщиной h и шириной d) на диэлектрической подложке (рис. 2).

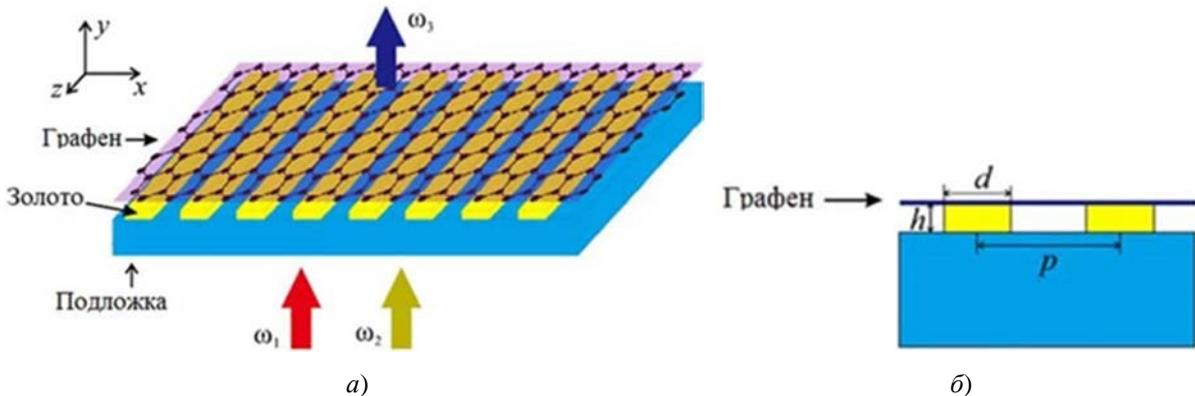


Рис. 2. Модель нелинейного преобразования частоты волн среднего ИК-диапазона (а) и проекция метаповерхности (б) на основе графенового листа на подложке с МДР из полосок золота [1]

В процессе генерации разностной частоты две волны с частотами ω_1 и ω_2 , нормально падающие на структуру (рис. 2), взаимодействуют в графеновом листе, генерируя волну на комбинационной частоте $\omega_3 = \omega_1 - \omega_2$ [8]. В данной структуре волна накачки и волна сигнала дальнего ИК-диапазона формируют плазмонные резонансы на поверхности золота, в то время как генерация волны на комбинационной частоте в среднем ИК-диапазоне формирует плазмонный резонанс на поверхности графена. В результате локальные электрические поля всех трех волн резонансно возрастают, что приводит к резкому усилению процесса генерации волны на комбинационной частоте [1].

Модуляторы среднего ИК-диапазона на основе метаповерхностей из легированных ГНЛ

За последние годы предприняты значительные усилия по созданию модуляторов на основе ППП в графене, функционирующих в диапазонах от ТГц до среднего ИК [40]. Эти устройства используют уникальное дисперсионное соотношение для ППП в графене, и основаны на геометриях, в которых используется резонансное поглощение в графеновых структурах. ППП моды могут быть электрически перестроены на контролируемую интенсивность и частоту либо путем оптического поглощения, либо излучения [41].

Однослойная атомная толщина и низкая плотность свободных носителей в графене ограничивают эффективность ТГц-, ИК-модуляторов, особенно на частотах среднего ИК-диапазона, где амплитуда ППП мод в графене низка [14, 42]. Для преодоления этих трудностей принят ряд стратегий, включая использование ионного геля или химического легирования для увеличения плотности носителей заряда в графене, а также включение металлического отражателя [43].

Конструкция модуляторов на основе метаповерхностей с легированными ГНЛ основана на тройной резонансной структуре, где плазмонные резонансы в ГНЛ согласованы с резонансами Фабри – Перо в диэлектрической подложке, а также с оптическими резонансами в субволновой щелевой МДР, которая предназначена для проявления экстраординарного оптического прохождения в среднем ИК-диапазоне [41]. В связанной структуре ГНЛ расположены внутри металлических щелей для модуляции связи между ППП модами на верхней и нижней металлических поверхностях. Ряд параметров обуславливают конструкцию щелевой МДР [41]: материал и толщина металла, ширина щели и период МДР.

Модель и принцип работы модулятора на основе графеновой метаповерхности с щелевой МДР показаны на рис. 3, конструкция метаповерхности – на рис. 4.

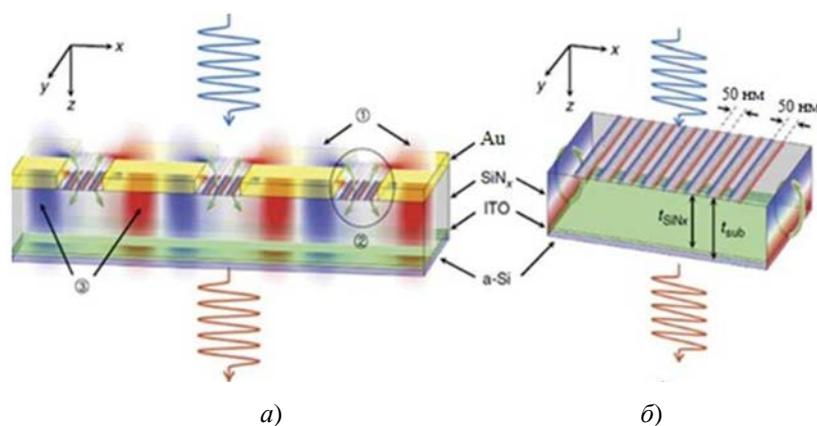


Рис. 3. Модель и принцип работы модулятора ИК-диапазона на основе метаповерхности с легированными ГНЛ [42]

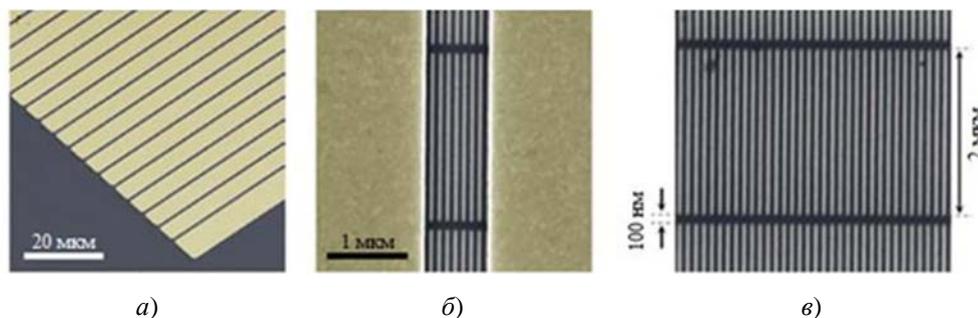


Рис. 4. Конструкция метаповерхности с легированными ГНЛ [41]

Падающая волна дифрагирует на периодической структуре, возбуждая ППП на верхней поверхности металла. Затем ППП через щели в металле субволновой длины возбуждают ППП на нижней поверхности металла, которые переизлучаются, в результате чего возникает дифракционный пик прохождения с сильной интенсивностью на резонансной частоте. ГНЛ помещаются в щели в металле для модуляции резонансной связи, что достигается электрической перестройкой плазмонных резонансов в ГНЛ в соответствии с частотой резонанса прохождения [41]. При совпадении электрические поля в щелях в металле порождают плазмонный резонанс в ГНЛ, что приводит к блокировке канала связи (4) для резонанса прохождения.

На рис. 5 представлены зависимости коэффициентов поглощения P , прохождения T , отражения R в зависимости от толщины подложки для уровня Ферми графена $E_F = 0,465$ эВ. Распределение электрического поля вокруг ГНЛ при $E_F = 0,465$ эВ для толщин подложки (б) 2,02 мкм и (в) 0,87 мкм; (г, д) зависимости – P и T от частоты.

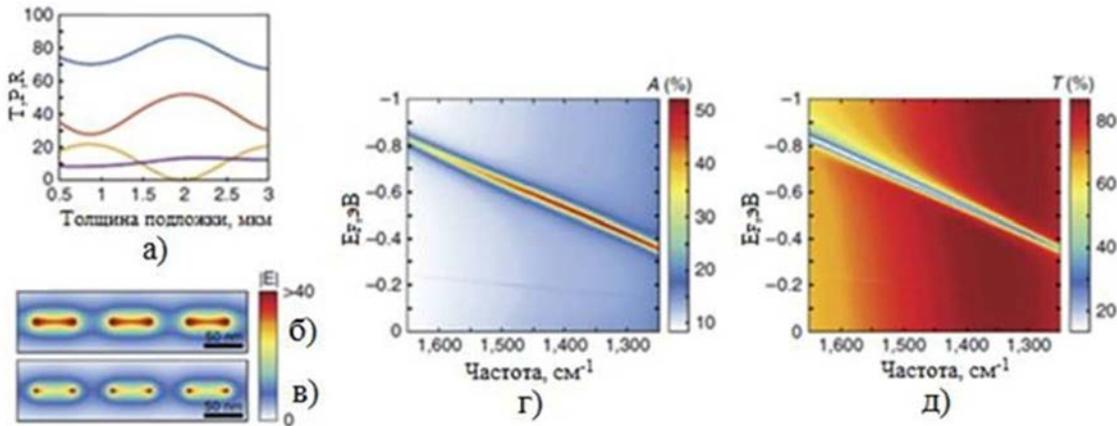


Рис. 5. Характеристики модулятора ИК-диапазона на основе метаповерхности с легированными ГНЛ [41]

Когда резонансные частоты ГНЛ и системы экстраординарного оптического прохождения сближаются, плазмонные резонансы могут изменять локальную диэлектрическую среду. Для значений E_F намного выше или ниже $E_F = 0,465$ эВ резонанс поглощения в ГНЛ находится далеко от резонанса прохождения щелевой МДР, и эти два эффекта ведут себя независимо. Однако при приближении E_F к значению $E_F = 0,465$ эВ плазмонный резонанс в ГНЛ создает расщепление, или антипересекающееся поведение между резонансной ППП модой графена и резонансной модой экстраординарного оптического прохождения [44–49].

Перестраиваемые генераторы гармоник и умножители частоты ИК-диапазона на основе плазмонных графеновых метаповерхностей

Генерация когерентного и перестраиваемого ТГц-излучения остается серьезной проблемой [50]. Фотомикширование с его преимуществами в виде сверхширокой полосы пропускания и точной перестройки стало важным методом генерации ТГц-, ИК-волн [50–52].

Генераторы гармоник среднего и дальнего ИК-диапазона на основе графеновых листов на кремниевой дифракционной решетке

Нелинейная восприимчивость графена третьего порядка и значительно увеличенная напряженность поля ППП в графене продемонстрировали большие возможности для повышения эффективности генерации третьей гармоники [53–55]. В работе [53] предложена кремниевая ДР на графене для генерации ППП с частотой третьей гармоники [53]. Графеновые ППП частоты ω_0 эффективно возбуждаются нормально падающей плоской волной (рис. 6) из-за резонанса основной моды ППП, а затем преобразуются в ППП с частотой третьей гармоники $3\omega_0$ с большой эффективностью, возникающей из-за высокой напряженности поля ППП частоты ω_0 благодаря кремниевой ДР под листом графена (p – период решетки, а w – ширина кремния) [53].

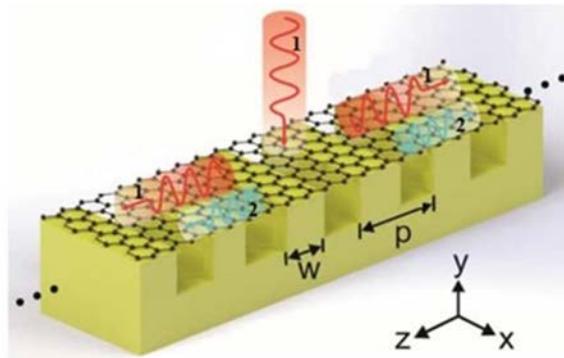


Рис. 6. Генерация третьей гармоники на основе графена на кремниевой ДР [54]

Рассмотренный принцип может стимулировать создание источников излучения на основе графена для фотоники среднего и дальнего ИК-диапазона на кремниевой подложке и, следовательно, расширить функциональные возможности кремниевой фотоники, такие как обработка сигналов, спектроскопия и зондирование.

Перестраиваемые генераторы второй и третьей гармоник ИК-диапазона на основе многослойных метаповерхностей из графеновых дифракционных решеток

Рассмотрим схему (рис. 7), позволяющую усилить генерацию второй и третьей гармоники многослойной метаповерхностью графен-диэлектрик-графен (ГДГ). В работе [54] разработана ГДГ-структура с уникальными оптическими свойствами, которые не могут быть достигнуты с помощью плазмонных наноструктур на основе металлов.

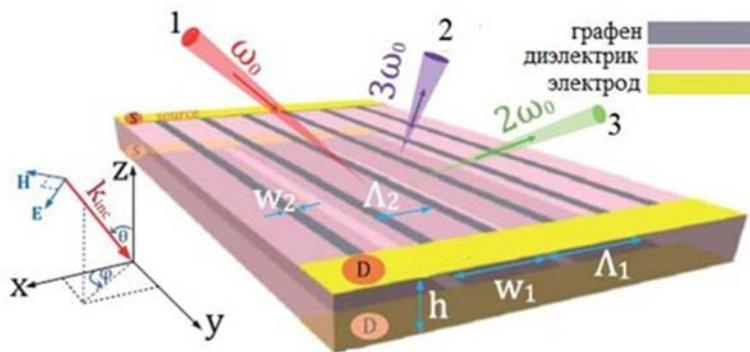


Рис. 7. Принцип функционирования перестраиваемого генератора ИК-диапазона на основе ГДГ-нанорезонатора, состоящего из двух ДР ГНЛ различной ширины, размещенных на противоположных гранях слоя диэлектрика:
 1 – излучение первой гармоники частоты ω_0 ; 2 – третьей гармоники частоты $3\omega_0$;
 3 – второй гармоники частоты $2\omega_0$) [54]

Конструкция генератора ИК-диапазона на основе многослойной метаповерхности – периодической ГДГ-структуры [54] – состоит из двух ДР ГНЛ, расположенных на противоположных гранях диэлектрической подложки (рис. 7). Электроды находятся в контакте с ГНЛ, что позволяет перестроить их уровень Ферми E_F . ТМ-поляризованная плоская волна с частотой ω_0 падает на ГДГ-структуру (рис. 7). Поскольку ППП в ГНЛ зависят от геометрии, ширина ГНЛ выбирается таким образом, чтобы ГНЛ в двух ДР имели резонансы основной моды ППП как на частоте первой гармоники $\omega_{\text{ПН}} = \omega_0$, так и высших гармоник ($\omega_{\text{ВГ}} = 2\omega_0$ для генерации второй гармоники и $\omega_{\text{ВГ}} = 3\omega_0$ для генерации третьей гармоники), как показано на рис. 7. Кроме того, нелинейно-оптический отклик ГДГ-структуры может быть дополнительно оптимизирован, требуя, чтобы ГНЛ в первой (нижней) ДР также имели резонансы мод ППП более высокого порядка.

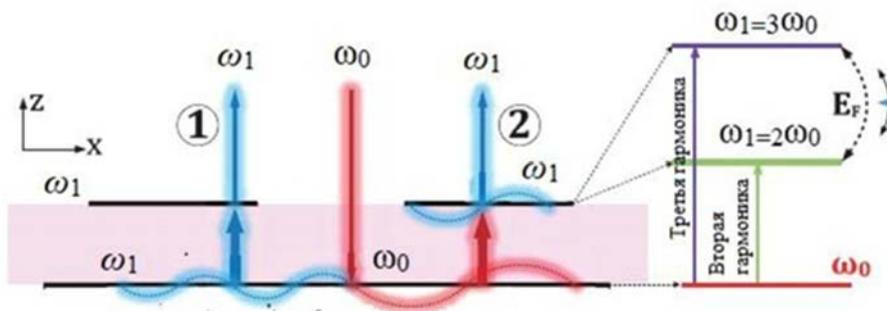


Рис. 8. Физический механизм увеличения генерации второй и третьей гармоники ГДГ метаповерхностью: ω_1 – частота ППП мод более высокого порядка; ω_0 – частота первой гармоники [54]

Существуют два ключевых механизма, которые способствуют заметному увеличению нелинейно-оптического отклика ГДГ структуры на несколько порядков по сравнению с графеновым слоем. Первый из них, обозначенный 1 на рис. 8, требует, чтобы существовали основная мода ППП в ГНЛ в первой (нижней) ДР при ω_0 и ППП моды более высокого порядка при ω_1 . Поле на частоте ω_0 , падающее на ГНЛ в первой ДР, генерирует сильное поле на этих ГНЛ при ω_0 через резонансное возбуждение основной моды ППП, а затем ППП мод более высокого порядка резонансно генерируются нелинейной поляризацией в этих же ГНЛ в первой решетке.

Гораздо более эффективный механизм, способствующий увеличению нелинейного резонанса ГДГ структуры, схематично обозначен 2 на рис. 8. ГНЛ во второй (верхней) ДР имеют основную моду ППП на высших гармониках. Усиленное оптическое поле за счет возбуждения основной моды ППП на ГНЛ в первой ДР индуцирует на ГНЛ во второй ДР сильную нелинейную поляризацию через взаимодействие ближнего поля. Это, в свою очередь, резонансно возбуждает основную моду ППП на ГНЛ во второй ДР. Кроме того, основная мода ППП на высших гармониках также генерируется через оптическую связь ближнего поля с модами ППП более высокого порядка ГНЛ в первой ДР [54].

На заключительном этапе моды более высокого порядка на нижних ГНЛ в первой ДР основная мода ППП ГНЛ во второй ДР связывается с излучательными модами для генерации сильного сигнала второй гармоники. Фактически эта система ГНЛ действует как нелинейная наноантенна Яги-Уда: нижняя и верхняя графеновые ДР являются рефлектором при ω_0 и директором при ω_{2g} соответственно.

Важным свойством ГДГ метаповерхности является усиление второй гармоники на широкополосной нелинейности, достигаемое путем перестройки энергии Ферми E_F в двух ДР. Причина этого уникального свойства раскрывается дисперсионной зависимостью поглощения оптимизированной ГДГ структуры, представленной на рис. 9,а.

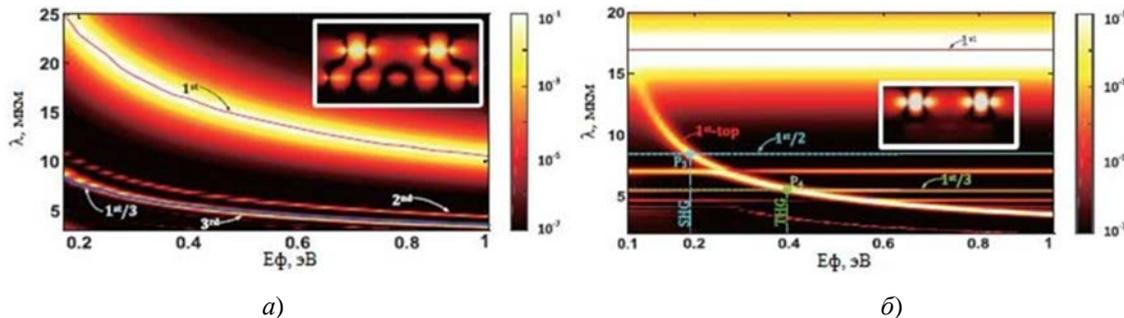


Рис. 9. Спектры поглощения оптимизированной ГДГ структуры в зависимости от энергии Ферми E_F двух графеновых ДР (а); спектры поглощения оптимизированной ГДГ структуры в зависимости от энергии Ферми E_F двух графеновых ДР, для случая когда изменяется только E_F во второй (верхней) ДР и $E_F = 0,4$ эВ в первой (нижней) ДР (б)

Соотношение между длинами волн основной моды ППП и моды третьего порядка в ГНЛ в первой ДР и основной моды ППП в ГНЛ во второй ДР остается постоянным при изменении E_F . Следовательно, свойство двойного резонанса сохраняется при изменении E_F . Как показано на рис. 9,а, при изменении E_F от 0,2 до 1 эВ резонансная длина волны моды ППП первого порядка ГНЛ в первой ДР и неявно рабочая длина волны на второй гармонике изменяются от 25 до 10 мкм [54].

Еще одним важным свойством ГДГ структуры является то, что она может увеличивать генерацию как третьей гармоники, так и второй. В частности, эта функциональность может быть реализована путем перестройки энергии Ферми E_F только во второй ДР так, что резонансная длина волны основной моды ППП ГНЛ в этой решетке смещается от третьей гармоники к второй. Об этом свидетельствует график поглощения ГДГ структуры, представленный на рис. 9,б, где показаны два типа плазмонных полос, а именно плоские полосы, соответствующие ППП в первой ДР, которые не зависят от E_F во второй ДР, и плазмонные полосы, связанные со второй ДР, резонансная длина волны которых зависит от E_F , в то время как резонансная длина волны основной моды ППП первой ДР остается постоянной, $\lambda_{FF} = 15,75$ мкм, резонансная длина волны основной моды ППП на второй ДР варьируется от $\lambda(P4) = \lambda_{FF} / 3 = 5,25$ мкм до $\lambda(P3) = \lambda_{FF} / 2 = 7,875$ мкм при перестройке E_F от 0,4 до 0,2 эВ соответственно.

Метаповерхность на основе ГДГ может быть использована для достижения перестраиваемого широкополосного увеличения как нелинейно-оптических, так и линейно-оптических взаимодействий; свойство, следующее из того факта, что система обладает перестраиваемыми двойными резонансами. Новый тип графеновых метаповерхностей открывает новое направление исследований в области разработки новых ультракомпактных и многочастотных активных фотонных наноструктур.

Заключение

Недавно появившийся графен стал материалом, способным расширить область нелинейной плазмоники не только до ТГц-, но и ИК-диапазона благодаря плазмонному отклику, исключительно высокой нелинейности третьего порядка, низким потерям в этих диапазонах, однослойной 2D-структуре, которая уменьшает объем взаимодействия света с веществом до чрезвычайно субволновых масштабов [14, 15]. ППП, возбуждаемые в графене, обеспечивают беспрецедентные уровни ограничения излучения. Плазмонные резонансы в графене могут динамически перестраиваться с помощью электростатического смещения и позволяют создавать новое поколение реконфигурируемых плазмонных устройств, в частности ИК-модуляторов, биосенсоров, метаповерхностей и идеальных поглотителей [14]. Благодаря уникальному ограничению и частотной перестраиваемости графеновая плазмоника подходит для многих применений, включая ИК-детектирование, высокоэффективное ИК-поглощение, метаматериалы и химическое зондирование. Графеновая плазмоника продемонстрировала большой потенциал в среднем ИК-диапазоне, где она позволяет, например, с высокой чувствительностью и избирательностью обнаруживать молекулярные частицы по их колебательным спектрам и тепловому излучению [47].

Нелинейно-оптические свойства многослойного графена, экспериментально изученные в широком спектре длин волн ИК-диапазона [20], определяют графен как перспективную среду для визуализации со сверхразрешением [56]. Плазмонный ИК-отклик многослойных графеновых структур, в частности двухслойного графена, обеспечивает более широкие спектральные диапазоны перестройки по сравнению с однослойным, в дополнение к поддержанию резонансов ППП мод на более высоких частотах вплоть до среднего ИК-диапазона и с более высоким поглощением [14]. Многослойный метаматериал из структурированного гиперболического графена может значительно улучшить нелинейные эффекты на ИК-частотах [15]. Исследованные спектральные, ближнепольные и свойства перестройки многослойных плазмонных ГНЛ показали, что большее плазмонное усиление связано с многослойной геометрией [14]. Экспериментально обнаружено улучшенное нелинейное смещение ближнего и среднего ИК-излучения в решетках ГНЛ, изучен отклик широкополосного ЧВС в графене с нанометровым и фемтосекундным пространственно-временным разрешением и обнаружено нелинейное усиление сигнала на краях ГНЛ в зависимости от количества слоев [6].

Это наряду с другими уникальными свойствами, такими как сильное нелинейное взаимодействие волн среднего ИК-диапазона с многослойными решетками ГНЛ, перестраиваемость эффективности нелинейного смещения волн и угловая нечувствительность, делают электронно-управляемые нелинейные устройства дальнего и среднего ИК-диапазона перспективным выбором для будущего нелинейного преобразования частоты, оптического управления и обработки информации на платформе графеновой плазмоники [6].

Список литературы

1. Cao Y., Kong, Gao S. [et al.]. Plasmon resonance enhanced mid-infrared generation by graphene on gold gratings through difference frequency mixing // *Optics Communications*. 2018. Vol. 406. P. 183–187. doi: 10.1016/j.optcom.2017.04.023.7
2. Barh A., Rodrigo P. J., Meng L., [et al.]. Parametric upconversion imaging and its applications // *Adv. Opt. Photon.* 2019. Vol. 11. P. 952–1019.
3. Du L., Ding X., Han D. [et al.]. 1.9 μm all-optical wavelength converter based on a graphene oxide coated micro-fiber // *Opt. Express*. 2021. Vol. 29. P. 40286–40293.
4. Houver S., Cavalie P., Meng L. [et al.]. Terahertz nonlinear optics with a compact semiconductor device // *SPIE news*. 2017. URL: <https://www.spie.org/news/6803-terahertz-nonlinear-optics-with-a-compact-semiconductor-device>
5. Raghunathan V., Deka J., Menon S. [et al.]. Nonlinear Optics in Dielectric Guided-Mode Resonant Structures and Resonant Metasurfaces // *Micromachines*. 2020. Vol. 11. P. 449. doi: 10.3390/mi11040449
6. Jiang T., Kravtsov V., Tokman M., [et al.]. Ultrafast coherent nonlinear nanooptics and nanoimaging of graphene // *Nature Nanotechnology*. 2019. Vol. 14. P. 838–843.

7. Головачева Е. В., Грибникова Е. И., Лерер А. М. [и др.]. Расчет электродинамических характеристик графеновых решеток // *Инженерный вестник Дона*. 2017. № 3.
8. Романенко П. М., Микаева С. А. Оптоэлектронные модуляторы излучения терагерцового диапазона на основе однослойного графена как перспективного компонента наноструктурированных метаповерхностей // *Информатика и технологии. Инновационные технологии в промышленности и информатике : сб. докл. конф. МИРЭА – Российский технологический университет. М., 2019. Т. 2. С. 268–274.*
9. Jiang Z., Wang Y., Chen L. [et al.]. Antenna-integrated silicon-plasmonic graphene sub-terahertz emitter // *APL Photonics*. 2021. Vol. 6. P. 066102.
10. Manabe S., Otsuji T., Satou A. Optical-to-THz Frequency Down-Conversion Utilizing Two-Dimensional Plasmons // *2019 44th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)*. 2019. P. 1–2.
11. Sun Z., Basov D. N., Fogler M. M. Graphene as a source of entangled plasmons // *Phys. Rev. Research*. 2022. Vol. 4. P. 023208.
12. Li Y., An N., Lu Z. [et al.]. Nonlinear co-generation of graphene plasmons for optoelectronic logic operations // *Nature Communications*. 2022. Vol. 3. P. 3138.
13. Mikhailov S. A. Quantum theory of the third-order nonlinear electrodynamic effects of graphene // *Phys. Rev*. 2016. Vol. 93, № 8. P. 085403.
14. Rodrigo D., Tittl A., Limaj O. [et al.]. Double-layer graphene for enhanced tunable infrared plasmonics // *Light: Science & Applications*. 2017. Vol. 6. P. e16277.
15. Jin B., Argyropoulos C. Nonlinear Graphene Metasurfaces with Advanced Electromagnetic Functionalities // *Proc. of SPIE – the International Society for Optical Engineering 16, Design, Materials Fabrication, Characterization, and Applications*. 2018. P. 107221R.
16. Gopalan K. K., Paulillo B., Mackenzie D. [et al.]. Scalable and Tunable Periodic Graphene Nanohole Arrays for Mid-Infrared Plasmonics // *Nano Lett*. 2018. Vol. 18, № 9. P. 5913–5918.
17. Kovalev S., Hafez H. A., Tielrooij K. J. [et al.]. Electrical tunability of terahertz nonlinearity in graphene // *Sci. Adv*. 2021. Vol. 7. eabf9809. doi: 10.1126/sciadv.abf9809
18. Jiang T., Huang D., Cheng J. [et al.]. Gate-tunable third-order nonlinear optical response of massless Dirac fermions in graphene // *Nature Photon*. 2018. Vol. 12. P. 430–436. doi: 10.1038/s41566-018-0175-7
19. Hendry E., Hale P. J., Moger J. [et al.]. Coherent Nonlinear Optical Response of Graphene // *Phys. Rev. Lett*. 2010. Vol. 105. P. 097401.
20. Kundys D., Van Duppen B., Marshall O. P. [et al.]. Nonlinear light mixing by graphene plasmons // *Nano Lett*. 2018. Vol. 18, № 1. P. 282–287.
21. Kislyakov M., Nunzi J.-M., Zhang X. [et al.]. Stimulated Brillouin scattering in dispersed graphene // *Opt. Express*. 2018. Vol. 26. P. 34346–34365.
22. Paterova A. V., Kalashnikov D. A., Khaidarov E. [et al.]. Non-linear interferometry with infrared metasurfaces // *Nanophotonics*. 2021. Vol. 10, № 6. P. 1775–1784.
23. Morales M. C., Rocco D., Xu L. [et al.]. Infrared imaging in nonlinear As metasurfaces // *SPIE Proceedings*. 2019. Vol. 11201: SPIE Micro + Nano Materials, Devices, and Applications. P. 112011S. doi: 10.1117/12.2541224/
24. Wu Y., Yao B. C., Feng Q. Y. [et al.]. Generation of cascaded four-wave-mixing with graphene-coated microfiber // *Photon. Res*. 2015. Vol. 3. P. A64–A68.
25. Deng H., Huang C., He Y. [et al.]. Quantum plasmon enhanced nonlinear wave mixing in graphene nanoflakes // *Chinese Physics B*. 2021. Vol. 30, № 4.
26. Guo T., Argyropoulos C. Recent advances in terahertz photonic technologies based on graphene and their applications // *Adv. Photonics Res*. 2021. Vol. 2. P. 2000168.
27. Li Y., An N., Lu Z. [et al.]. Nonlinear co-generation of graphene plasmons for optoelectronic logic operations // *Nature communications*. 2022. Vol. 3. P. 3138.
28. Pfeiffer T., Kutas M., Haase B. [et al.]. Terahertz detection by upconversion to the near-infrared using picosecond pulses // *Opt. Express*. 2020. Vol. 28. P. 29419–29429.
29. Salmon A., Bouchon P. Rapid prototyping of a bispectral terahertz-to-infrared converter // *Opt. Express*. 2021. Vol. 29. P. 18437–18445.
30. Pei S., Duan S., Yang N. [et al.]. Second-order nonlinear optical response of graphene irradiated by two-color lights: ellipticity and phase modulation // *J. Opt. Soc. Am*. 2021. Vol. 38. P. 2594–2602.
31. Sun Z., Basov D. N., Fogler M. M. Graphene as a source of entangled plasmons // *Phys. Rev. Research*. 2022. Vol. 4. P. 023208.
32. Christopoulos T., Tsilipakos O., Ataloglou V. G., Kriezis E. E. A Theoretical Study of Nonlinear Multi-Channel Graphene Components for the Near and Far-Infrared // *Frontiers in Optics. Laser Science / ed. by B. Lee, C. Mazzali, K. Corwin, R. Jason Jones. OSA Technical Digest (Optica Publishing Group)*. 2020. Paper JTh4B.17.
33. Qasymeh M., Eleuch H. Quantum microwave-to-optical conversion in electrically driven multilayer graphene // *Opt. Express*. 2019. Vol. 27. P. 5945–5960.
34. <https://www.materialstoday.com/metals-alloys/news/journal-of-alloys-and-compounds-call-for-papers>

35. Cheng C., Huang B., Mao X. [et al.]. Frequency conversion with nonlinear graphene photodetectors // *Nanoscale*. 2017. Vol. 9. P. 4082–4089.
36. Ji M., Cai H., Deng L. [et al.]. Enhanced parametric frequency conversion in a compact silicon-graphene microring resonator // *Opt. Express*. 2015. Vol. 23. P. 18679–18685.
37. Sheng L., Vaskin A., Reno J. L. [et al.]. An optical metamixer. arXiv:1711.0009010. doi.1038/s41467-018-04944-9
38. Liu S., Vaskin A., Reno J. L. [et al.]. An all-dielectric metasurface as a broadband optical frequency mixer // *Nat. Commun.* 2018. Vol. 9. P. 2507. doi: 10.1038/s41467-018-04944-9
39. Ooi K. J. A., Tan D. T. H. Nonlinear graphene plasmonics // *Proc. R. Soc.* 2017. Vol. 473. P. 20170433.
40. Ginis V., Tassin P., Koschny T. [et al.]. Tunable terahertz frequency comb generation using time-dependent graphene sheets // *Physical Review*. 2015. Vol. 91, № 16.
41. Kim S., Jang M., Brar V. [et al.]. Electronically tunable extraordinary optical transmission in graphene plasmonic ribbons coupled to subwavelength metallic slit arrays // *Nat Commun.* 2016. Vol. 7. P. 12323. doi: 10.1038/ncomms12323
42. Ullah Z., Witjaksono G., Nawi I. [et al.]. A Review on the Development of Tunable Graphene Nanoantennas for Terahertz Optoelectronic and Plasmonic // *Sensors*. 2020. Vol. 20, № 5. P. 1401.
43. Gopalan K. K., Paulillo B., Mackenzie D. M. A. [et al.]. Tunable Periodic Graphene Nanohole Arrays for Mid-Infrared Plasmonics // *Nano Lett.* 2018. Vol. 18, № 9. P. 5913–5918.
44. Cox J. D., Garcia de Abajo F. J. Plasmon-enhanced nonlinear wave mixing in nanostructured graphene // *ACS Photonics*. 2015. Vol. 2. P. 306–312.
45. Deinert J. C., Iranzo D. A., Perez R. [et al.]. Grating-graphene metamaterial as a platform for terahertz nonlinear photonic // *ACS Nano*. 2021. Vol. 15, № 1. P. 1145–1154.
46. Menendez G. A., Maes B. Frequency comb generation in a time-dependent graphene ribbon array // *Phys. Rev.* Vol. 95. P. 144307.
47. Cox J. D., Garcia de Abajo F. J. Nonlinear Graphene Nanoplasmonics // *Acc. Chem. Res.* 2019. Vol. 52, № 9. P. 2536–2547.
48. Sherrott M. C., Hon P. W. C., Fountaine K. T. [et al.]. Experimental Demonstration of $>230^\circ$ Phase Modulation in Gate-Tunable Graphene-Gold Reconfigurable Mid-Infrared Metasurfaces // *Nano Lett.* 2017. Vol. 17, № 5. P. 3027–3034.
49. An N., Tan T., Peng Z. [et al.]. Electrically Tunable Four-Wave-Mixing in Graphene Heterogeneous Fiber for Individual Gas Molecule Detection // *Nano Lett.* 2020. Vol. 20, № 9. P. 6473–6480.
50. Verma N., Govindan A. Nonlinear Mixing of Two Laser Beams and THz plasmons generation in graphene coated optical Fibre : preprint from Research Square. 2022. doi: 10.21203/rs.3.rs-1079347/v1 PPR: PPR438413
51. Kowligy A. S., Hickstein D. D., Lind A. [et al.]. Tunable mid-infrared generation via wide-band four-wave mixing in silicon nitride waveguides // *Optics Letters*. 2018. Vol. 43. P. 4220.
52. Menendez G. A., Maes B. Frequency comb generation in a time-dependent graphene ribbon lattice // *Phys. Rev.* 2017. Vol. 95. P. 144307.
53. Li J., Zhang T., Chen L. High-Efficiency Plasmonic Third-Harmonic Generation with Graphene on a Silicon Diffractive Grating in Mid-Infrared Region // *Nanoscale Research Letters*. 2018. Vol. 13. P. 338.
54. You J. W., Panoiu N. C. Tunable and dual-broadband giant enhancement of second-harmonic and third-harmonic generation in an optimized graphene-insulator-graphene metasurface // *Phys. Rev.* 2020. Vol. 102. P. 121403(R).
55. Ullah Z., Witjaksono G., Nawi I. [et al.]. A Review on the Development of Tunable Graphene Nanoantennas for Terahertz Optoelectronic and Plasmonic // *Sensors*. 2020. Vol. 20, № 5. P. 1401.
56. Demetriou G., Bookey H. T., Biancalana F. [et al.]. Nonlinear optical properties of multilayer graphene in the infrared // *Opt. Exp.* 2016. Vol. 24. P. 13033–13043.

References

1. Cao Y. Kong, Gao S. et al. Plasmon resonance enhanced mid-infrared generation by graphene on gold gratings through difference frequency mixing. *Optics Communications*. 2018;406:183–187. doi: 10.1016/j.optcom.2017.04.023.7
2. Barh A., Rodrigo P. J., Meng L. et al. Parametric upconversion imaging and its applications. *Adv. Opt. Photon.* 2019;11:952–1019.
3. Du L., Ding X., Han D. et al. 1.9 μm all-optical wavelength converter based on a graphene oxide coated microfiber. *Opt. Express*. 2021;29:40286–40293.
4. Houver S., Cavalie P., Meng L. et al. Terahertz nonlinear optics with a compact semiconductor device. *SPIE news*. 2017. Available at: <https://www.spie.org/news/6803-terahertz-nonlinear-optics-with-a-compact-semiconductor-device>
5. Raghunathan V., Deka J., Menon S. et al. Nonlinear Optics in Dielectric Guided-Mode Resonant Structures and Resonant Metasurfaces. *Micromachines*. 2020;11:449. doi: 10.3390/mi11040449

6. Jiang T., Kravtsov V., Tokman M. et al. Ultrafast coherent nonlinear nanooptics and nanoimaging of graphene. *Nature Nanotechnology*. 2019;14:838–843.
7. Golovacheva E.V., Gribnikova E.I., Lerer A.M. et al. Calculation of electrodynamic characteristics of graphene gratings. *Inzhenernyy vestnik Dona = Engineering Bulletin of the Don*. 2017;(3). (In Russ.)
8. Romanenko P.M., Mikaeva S.A. Optoelectronic modulators of terahertz radiation based on single-layer graphene as a promising component of nanostructured metasurfaces. *Informatika i tekhnologii. Innovatsionnye tekhnologii v promyshlennosti i informatike: sb. dokl. konf. MIREA – Rossiyskiy tekhnologicheskii universitet = Informatics and Technologies. Innovative technologies in industry and computer science : sat. dokl. konf. MIREA – Russian Technological University*. Moscow, 2019;2:268–274. (In Russ.)
9. Jiang Z., Wang Y., Chen L. et al. Antenna-integrated silicon–plasmonic graphene sub-terahertz emitter. *APL Photonics*. 2021;6:066102.
10. Manabe S., Otsuji T., Satou A. Optical-to-THz Frequency Down-Conversion Utilizing Two-Dimensional Plasmons. *2019 44th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)*. 2019:1–2.
11. Sun Z., Basov D.N., Fogler M.M. Graphene as a source of entangled plasmons. *Phys. Rev. Research*. 2022;4:023208.
12. Li Y., An N., Lu Z. et al. Nonlinear co-generation of graphene plasmons for optoelectronic logic operations. *Nature Communications*. 2022;3:3138.
13. Mikhailov S.A. Quantum theory of the third-order nonlinear electrodynamic effects of graphene. *Phys. Rev.* 2016;93(8):085403.
14. Rodrigo D., Tittl A., Limaj O. et al. Double-layer graphene for enhanced tunable infrared plasmonics. *Light: Science & Applications*. 2017;6:e16277.
15. Jin B., Argyropoulos C. Nonlinear Graphene Metasurfaces with Advanced Electromagnetic Functionalities. *Proc. of SPIE – the International Society for Optical Engineering 16, Design, Materials Fabrication, Characterization, and Applications*. 2018:107221R.
16. Gopalan K.K., Paulillo B., Mackenzie D. et al. Scalable and Tunable Periodic Graphene Nanohole Arrays for Mid-Infrared Plasmonics. *Nano Lett.* 2018;18(9):5913–5918.
17. Kovalev S., Hafez H.A., Tielrooij K.J. et al. Electrical tunability of terahertz nonlinearity in graphene. *Sci. Adv.* 2021;7:eabf9809. doi: 10.1126/sciadv.abf9809
18. Jiang T., Huang D., Cheng J. et al. Gate-tunable third-order nonlinear optical response of massless Dirac fermions in graphene. *Nature Photon.* 2018;12:430–436. doi: 10.1038/s41566-018-0175-7
19. Hendry E., Hale P.J., Moger J. et al. Coherent Nonlinear Optical Response of Graphene. *Phys. Rev. Lett.* 2010;105:097401.
20. Kundys D., Van Duppen B., Marshall O.P. et al. Nonlinear light mixing by graphene plasmons. *Nano Lett.* 2018;18(1):282–287.
21. Kislyakov M., Nunzi J.-M., Zhang X. et al. Stimulated Brillouin scattering in dispersed graphene. *Opt. Express*. 2018;26:34346–34365.
22. Paterova A.V., Kalashnikov D.A., Khaidarov E. et al. Non-linear interferometry with infrared metasurfaces. *Nanophotonics*. 2021;10(6):1775–1784.
23. Morales M.C., Rocco D., Xu L. et al. Infrared imaging in nonlinear As metasurfaces. *SPIE Proceedings*. 2019;11201(SPIE Micro + Nano Materials, Devices, and Applications):112011S. doi: 10.1117/12.2541224/
24. Wu Y., Yao B. C., Feng Q.Y. et al. Generation of cascaded four-wave-mixing with graphene-coated microfiber. *Photon. Res.* 2015;3:A64–A68.
25. Deng H., Huang C., He Y. et al. Quantum plasmon enhanced nonlinear wave mixing in graphene nanoflakes. *Chinese Physics B*. 2021;30(4).
26. Guo T., Argyropoulos C. Recent advances in terahertz photonic technologies based on graphene and their applications. *Adv. Photonics Res.* 2021;2:2000168.
27. Li Y., An N., Lu Z. et al. Nonlinear co-generation of graphene plasmons for optoelectronic logic operations. *Nature communications*. 2022;3:3138.
28. Pfeiffer T., Kutas M., Haase B. et al. Terahertz detection by upconversion to the near-infrared using picosecond pulses. *Opt. Express*. 2020;28:29419–29429.
29. Salmon A., Bouchon P. Rapid prototyping of a bispectral terahertz-to-infrared converter. *Opt. Express*. 2021;29:18437–18445.
30. Pei S., Duan S., Yang N. et al. Second-order nonlinear optical response of graphene irradiated by two-color lights: ellipticity and phase modulation. *J. Opt. Soc. Am.* 2021;38:2594–2602.
31. Sun Z., Basov D.N., Fogler M.M. Graphene as a source of entangled plasmons. *Phys. Rev. Research*. 2022;4:023208.
32. Christopoulos T., Tsilipakos O., Ataloglou V.G., Kriezis E.E. A Theoretical Study of Nonlinear Multi-Channel Graphene Components for the Near and Far-Infrared. *Frontiers in Optics. Laser Science*. Ed. by B. Lee, C. Mazzali, K. Corwin, R. Jason Jones. OSA Technical Digest (Optica Publishing Group, 2020), paper JTh4B.17.

33. Qasymeh M., Eleuch H. Quantum microwave-to-optical conversion in electrically driven multilayer graphene. *Opt. Express*. 2019;27:5945–5960.
34. Available at: <https://www.materialstoday.com/metals-alloys/news/journal-of-alloys-and-compounds-call-for-papers>
35. Cheng C., Huang B., Mao X. et al. Frequency conversion with nonlinear graphene photodetectors. *Nanoscale*. 2017;9:4082–4089.
36. Ji M., Cai H., Deng L. et al. Enhanced parametric frequency conversion in a compact silicon-graphene microring resonator. *Opt. Express*. 2015;23:18679–18685.
37. Sheng L., Vaskin A., Reno J.L. et al. *An optical metamixer*. arXiv:1711.0009010. doi:10.38/s41467-018-04944-9
38. Liu S., Vaskin A., Reno J.L. et al. An all-dielectric metasurface as a broadband optical frequency mixer. *Nat. Commun*. 2018;9:2507. doi: 10.1038/s41467-018-04944-9
39. Ooi K.J.A., Tan D.T.H. Nonlinear graphene plasmonics. *Proc. R. Soc.* 2017;473:20170433.
40. Ginis V., Tassin P., Koschny T. et al. Tunable terahertz frequency comb generation using time-dependent graphene sheets. *Physical Review*. 2015;91:16.
41. Kim S., Jang M., Brar V. et al. Electronically tunable extraordinary optical transmission in graphene plasmonic ribbons coupled to subwavelength metallic slit arrays. *Nat Commun*. 2016;7:12323. doi: 10.1038/ncomms12323
42. Ullah Z., Witjaksono G., Nawi I. et al. A Review on the Development of Tunable Graphene Nanoantennas for Terahertz Optoelectronic and Plasmonic. *Sensors*. 2020;20(5):1401.
43. Gopalan K.K., Paulillo B., Mackenzie D.M.A. et al. Tunable Periodic Graphene Nanohole Arrays for Mid-Infrared Plasmonics. *Nano Lett*. 2018;18(9):5913–5918.
44. Cox J.D., Garcia de Abajo F. J. Plasmon-enhanced nonlinear wave mixing in nanostructured graphene. *ACS Photonics*. 2015;2:306–312.
45. Deinert J.C., Iranzo D.A., Perez R. et al. Gratings-graphene metamaterial as a platform for terahertz nonlinear photonic. *ACS Nano*. 2021;15(1):1145–1154.
46. Menendez G.A., Maes B. Frequency comb generation in a time-dependent graphene ribbon array. *Phys. Rev.* 95:144307.
47. Cox J.D., Garcia de Abajo F. J. Nonlinear Graphene Nanoplasmonics. *Acc. Chem. Res.* 2019;52(9):2536–2547.
48. Sherrott M.C., Hon P.W.C., Fountaine K.T. et al. Experimental Demonstration of >230° Phase Modulation in Gate-Tunable Graphene-Gold Reconfigurable Mid-Infrared Metasurfaces. *Nano Lett*. 2017;17(5):3027–3034.
49. An N., Tan T., Peng Z. et al. Electrically Tunable Four-Wave-Mixing in Graphene Heterogeneous Fiber for Individual Gas Molecule Detection. *Nano Lett*. 2020;20(9):6473–6480.
50. Verma N., Govindan A. *Nonlinear Mixing of Two Laser Beams and THz plasmons generation in graphene coated optical Fibre: preprint from Research Square*. 2022. doi: 10.21203/rs.3.rs-1079347/v1 PPR: PPR438413
51. Kowligy A.S., Hickstein D.D., Lind A. et al. Tunable mid-infrared generation via wide-band four-wave mixing in silicon nitride waveguides. *Optics Letters*. 2018;43:4220.
52. Menendez G.A., Maes B. Frequency comb generation in a time-dependent graphene ribbon lattice. *Phys. Rev.* 2017;95:144307.
53. Li J., Zhang T., Chen L. High-Efficiency Plasmonic Third-Harmonic Generation with Graphene on a Silicon Diffractive Grating in Mid-Infrared Region. *Nanoscale Research Letters*. 2018;13:338.
54. You J.W., Panoiu N.C. Tunable and dual-broadband giant enhancement of second-harmonic and third-harmonic generation in an optimized graphene-insulator-graphene metasurface. *Phys. Rev.* 2020;102:121403(R).
55. Ullah Z., Witjaksono G., Nawi I. et al. A Review on the Development of Tunable Graphene Nanoantennas for Terahertz Optoelectronic and Plasmonic. *Sensors*. 2020;20(5):1401.
56. Demetriou G., Bookey H.T., Biancalana F. et al. Nonlinear optical properties of multilayer graphene in the infrared. *Opt. Exp.* 2016;24:13033–13043.

Информация об авторах / Information about the authors

Галина Степановна Макеева

доктор технических наук, профессор, профессор
кафедры радиотехники и радиоэлектронных систем,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: radiotech@pnzgu.ru

Galina S. Makeeva

Doctor of technical sciences, professor,
professor of the sub-department
of radio engineering and radio-electronic systems,
Penza State University
(40 Krasnaya street, Penza, Russia)

Автор заявляет об отсутствии конфликта интересов /

The author declares no conflicts of interests.

Поступила в редакцию/Received 15.06.2024

Поступила после рецензирования/Revised 26.07.2024

Принята к публикации/Accepted 31.08.2024

НАВИГАЦИОННЫЕ ЭЛЕМЕНТЫ ПОЛЕТА БЕСПИЛОТНОГО ЛЕТАТЕЛЬНОГО АППАРАТА

А. И. Годунов¹, С. А. Куканов², П. С. Суздальцев³

¹ Пензенский государственный университет, Пенза, Россия

^{2,3} Филиал Военной академии материально-технического обеспечения
имени генерала армии А. В. Хрулева в г. Пензе, Пенза, Россия

¹ avitelpgu@mail.ru, ² Kuk_@mail.ru, ³ Suzdal.1990@bk.ru

Аннотация. *Актуальность и цели.* Рассматривается проблема навигации и управления беспилотными летательными аппаратами (БЛА). Для обеспечения эффективного выполнения полетного задания предлагается учитывать в расчетах навигационные элементы полета беспилотного летательного аппарата. Особенностью предложенного метода является возможность оценки характеристик и параметров элементов, влияющих на траекторию и точность полета. Учитываются различные факторы, такие как аэродинамические свойства БЛА, условия окружающей среды, системы навигации и управления. *Материалы и методы.* В статье представлены математические модели и алгоритмы расчета оптимальных навигационных элементов, которые позволяют достичь наилучшей эффективности выполнения полетного задания. Предложенный подход отличается от существующих методов возможностью более точного определения траектории и маневров БЛА. *Результаты и выводы.* Результаты исследования могут найти практическое применение в разработке систем управления и навигации БЛА различного назначения, повышая их точность, надежность и эффективность выполнения задач. Представленные методы могут быть использованы в гражданской и военной авиации, а также в других областях, где применяются беспилотные летательные аппараты.

Ключевые слова: беспилотный летательный аппарат, квадрокоптер, система навигации, навигационные элементы полета

Для цитирования: Годунов А. И., Куканов С. А., Суздальцев П. С. Навигационные элементы полета беспилотного летательного аппарата // Надежность и качество сложных систем. 2024. № 3. С. 104–111. doi: 10.21685/2307-4205-2024-3-11

FLIGHT NAVIGATION ELEMENTS AN UNMANNED AERIAL VEHICLE

A.I. Godunov¹, S.A. Kukanov², P.S. Suzdaltsev³

¹ Penza State University, Penza, Russia

^{2,3} Branch of the Military Academy of Logistics named after Army General A.V. Khrulev in Penza, Penza, Russia

¹ avitelpgu@mail.ru, ² Kuk_@mail.ru, ³ Suzdal.1990@bk.ru

Abstract. *Background.* This article discusses the problem of navigation and control of unmanned aerial vehicles (UAVs). To ensure the effective performance of the flight task, it is proposed to take into account the calculations of the navigation elements of the flight of an unmanned aerial vehicle, the feature of the proposed method is the ability to evaluate the characteristics and parameters of elements that affect the trajectory and accuracy of flight. Various factors are taken into account, such as the aerodynamic properties of the UAV, environmental conditions, navigation and control systems. *Materials and methods.* The article presents mathematical models and algorithms for calculating optimal navigation elements, which allow to achieve the best efficiency of the flight task. The proposed approach differs from existing methods in the possibility of more accurate determination of the trajectory and maneuvers of the UAV. *Results and conclusions.* The results of the study can find practical application in the development of UAV control and navigation systems for various purposes, increasing their accuracy, reliability and efficiency of tasks. The presented methods can be used in civil and military aviation, as well as in other areas where unmanned aerial vehicles are used.

Keywords: unmanned aerial vehicle, quadcopter, navigation system, flight navigation elements

For citation: Godunov A.I., Kukanov S.A., Suzdaltsev P.S. Flight navigation elements an unmanned aerial vehicle. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):104–111. (In Russ.). doi: 10.21685/2307-4205-2024-3-11

Во время полета центр масс летательного аппарата движется в пространстве по некоторой траектории. При решении различных задач обычно определяют линии положения летательного аппарата (рис. 1). Она характеризуется расстоянием до одной или нескольких точек на земной поверхности, направлением на точку земной поверхности или небесной сферы, высотой светила. В качестве линий положения используются ортодромия, локсодромия, линия равных азимутов, линия равных радиопеленгов, линии равных расстояний и разностей расстояний.

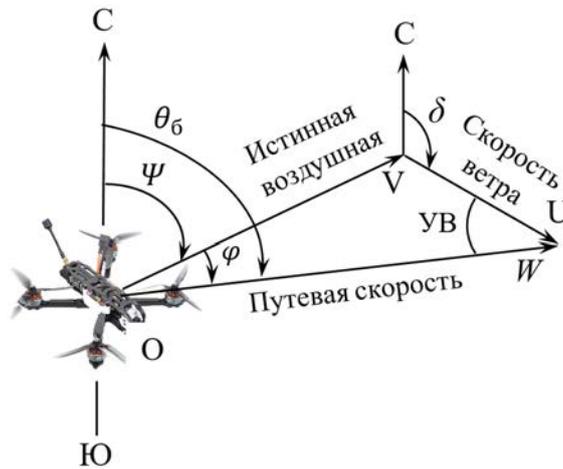


Рис. 1. Навигационный треугольник скоростей

Величины, определяющие направление и траекторию полета, называются навигационными элементами. Основные навигационные элементы полета: курс Ψ ; заданный и фактический путевые углы; истинная воздушная скорость V ; индикаторная воздушная скорость $V_{\text{инд}}$; путевая скорость W ; угол сноса φ ; высота полета H . Обычно считается, что вектор воздушной скорости V совпадает с продольной осью летательного аппарата. Поэтому угол между северным направлением меридиана и вектором V равен курсу самолета Ψ . Заданный путевой угол определяется как угол между северным направлением меридиана и прямолинейным отрезком линии заданного пути, определяя направление полета. Фактический путевой угол определяется как угол между северным направлением меридиана и линией фактического пути. Направление вектора скорости ветра определяется углом δ между северным направлением меридиана и вектором ветра. Угол сноса φ представляет собой угол между векторами истинной воздушной и путевой скоростей. Путевой угол $\theta_б$ отсчитывается по ходу часовой стрелки от северного направления меридиана до направления вектора путевой скорости. Величина путевого угла равна алгебраической сумме курса и угла сноса:

$$\theta_б = \Psi \pm \varphi. \quad (1)$$

Угол ветра отсчитывается по ходу часовой стрелки от вектора путевой скорости до вектора ветра и может быть представлен в виде

$$\text{УВ} = \delta - \theta_б. \quad (2)$$

Угол сноса и угол ветра связаны зависимостью

$$\sin \varphi = \frac{U}{V} \sin \text{УВ}. \quad (3)$$

Путевая скорость летательного аппарата

$$W = V \cos \varphi + U \cos \text{УВ}. \quad (4)$$

Так как движение центра масс летательного аппарата по траектории полета характеризуется истинной воздушной и путевой скоростями, отклонением от заданной траектории по высоте H , боковым линейным отклонением z , величиной пройденного пути S и составляющими S_x, S_y, S_z , то траекторию полета летательного аппарата можно записать следующими системами уравнений:

$$\frac{dH}{dt} = V_y = V \sin \theta, \quad (5)$$

$$\frac{dz}{dt} = V_z = V \sin \theta_\delta, \quad (6)$$

$$\frac{dV}{dt} = \frac{1}{m} \sum F_x, \quad (7)$$

$$W = V \cos \varphi + U \cos \gamma_B, \quad (8)$$

$$\frac{dS}{dt} = W, \quad (9)$$

$$\sin \theta = \sin \vartheta - (\alpha \cos \gamma + \beta \sin \gamma) \cos \vartheta, \quad (10)$$

$$\theta_\delta = \psi + \varphi. \quad (11)$$

Интегрирование дифференциальных уравнений (5), (7), (9) позволяет определить соответственно высоту полета H , боковое линейное отклонение от заданной траектории z , истинную воздушную скорость V ; пройденный путь S . Решение уравнений (8)–(11) дает возможность вычислить путевую скорость W . Скорость ветра U при решении системы уравнений (5)–(11) может быть представлена постоянной величиной или величиной, изменяющейся по случайному закону.

Курс летательного аппарата вычисляется в системе моделирования динамики полета в результате интегрирования уравнения

$$\frac{d\psi}{dt} = \frac{1}{\cos \vartheta} (\omega_y \cos \gamma - \omega_z \sin \gamma). \quad (12)$$

Однако величина ψ из уравнения (12) определяется в предположении, что полет совершается на небольшое расстояние, а также, что при полете летательный аппарат находится достаточно близко к ортодромии, проходящей через центральную точку района полета (место взлета). На основании этих предположений сферичность Земли не учитывается, и задача определения местоположения летательного аппарата решается на горизонтальной плоскости.

Местоположение самолета на поверхности Земли может быть определено в географических и ортодромических системах координат. Ортодромическая система – это сферическая система координат с началом в произвольной точке земной поверхности. При определении местоположения летательного аппарата на горизонтальной плоскости используются прямоугольные условные системы координат. Обычно применяются земная прямоугольная система $O_0x_0y_0z_0$ и аэродромная $O_ax_ay_az_a$. В земной системе начало координат находится в центральной точке района полета, ось O_0x_0 направлена на север, ось O_0z_0 на восток, ось O_0y_0 совпадает с вертикалью в точке начала координат (рис. 2).

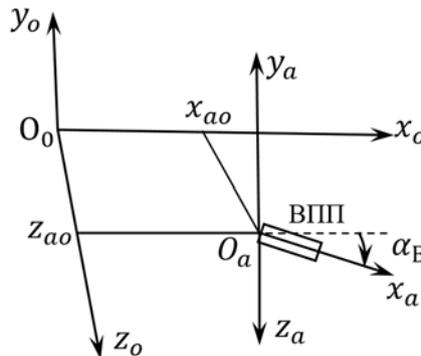


Рис. 2. Земная и аэродромная система координат

Аэродромная система координат ориентируется так, что ось $O_a x_a$ совпадает с направлением оси ВПП. Начало координат выбирается обычно в начале ВПП. Положение начала аэродромной системы в земной системе определяется координатами $x_{a0} z_{a0}$. В частном случае начала обеих систем могут совпадать, тогда $x_{a0} = 0, z_{a0} = 0$. Но во всех случаях оси аэродромной системы координат повернуты вокруг вертикальной оси на угол курса взлетно-посадочной полосы или на угол ВПП α_B .

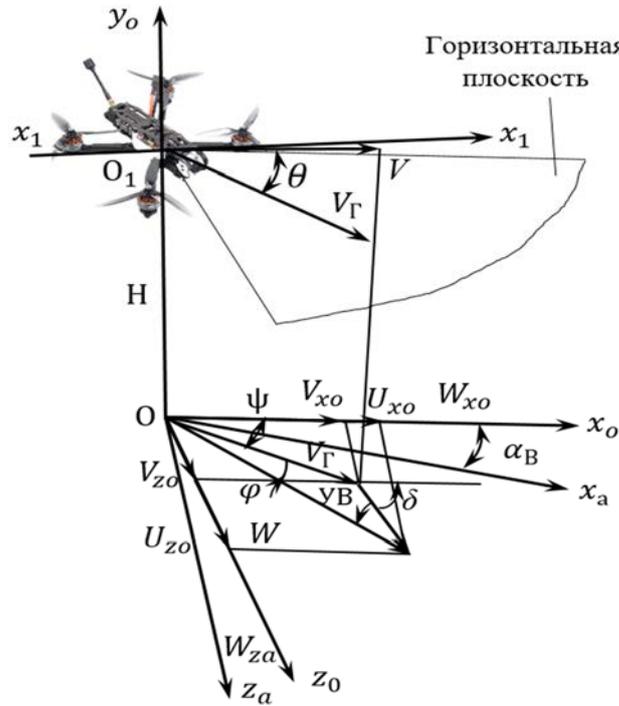


Рис. 3. Определение навигационных элементов

Летательный аппарат в общем случае может совершать негоризонтальный полет, тогда вектор истинной воздушной скорости V будет наклонен под некоторым углом θ к горизонтальной плоскости (рис. 3). При этом вектор V может не совпадать с направлением продольной оси летательного аппарата $x_1 x_1$.

Проекция вектора V на горизонтальную плоскость можно определить как

$$V_\Gamma = V \cos \theta. \quad (13)$$

Если считать, что полет совершается без скольжения, то вектор скорости V совпадает с направлением продольной оси беспилотного летательного аппарата. Поэтому курс ψ может быть определен углом в горизонтальной плоскости между направлением оси $O_0 x_0$ и направлением вектора V_Γ . Теперь проекции вектора V_Γ на оси земной системы координат есть произведения:

$$V_{x0} = V_\Gamma \cos \psi = V \cos \theta \cos \psi; \quad (14)$$

$$V_{z0} = V_\Gamma \sin \psi = V \cos \theta \sin \psi. \quad (15)$$

Как правило, полет летательного аппарата совершается в условиях ветра. В частном, простейшем, случае предполагается, что горизонтальная составляющая скорости ветра U – величина постоянная. Также неизменным считается направление ветра, которое характеризуется углом δ . Поэтому проекции вектора U на оси системы координат составляют

$$U_{x0} = U \cos \delta; \quad (16)$$

$$U_{z0} = U \sin \delta. \quad (17)$$

Проекция вектора путевой скорости W на оси земной системы координат определяется следующими суммами:

$$W_{x_0} = V_{x_0} + U_{x_0} = V \cos \theta \cos \psi + U \cos \delta; \quad (18)$$

$$W_{z_0} = V_{z_0} + U_{z_0} = V \cos \theta \sin \psi + U \sin \delta. \quad (19)$$

Для получения составляющих путевой скорости в аэродромной системе координат необходимо учитывать смещение начала этой системы относительно начала земных координат, т.е. координаты точки O_a . Кроме того, должен быть учтен угол поворота α_B аэродромных координат относительно земной системы. В частном случае, когда начало аэродромной системы совпадает с началом земной системы, учитывается только угол α_B (рис. 4). Проекции вектора путевой скорости на оси аэродромной системы координат составляют

$$W_{x_a} = W_{x_{0a}} + W_{z_{0x_a}}; \quad (20)$$

$$W_{z_a} = W_{z_{0a}} + W_{x_{0z_a}}. \quad (21)$$

Проекции составляющих вектора путевой скорости в земной системе координат W_{x_0} , W_{z_0} на оси аэродромной системы равны

$$W_{x_{0a}} = W_{x_0} \cos \alpha_B; \quad (22)$$

$$W_{x_{0z_a}} = -W_{x_0} \sin \alpha_B; \quad (23)$$

$$W_{z_{0x_a}} = W_{z_0} \sin \alpha_B; \quad (24)$$

$$W_{z_{0a}} = W_{z_0} \cos \alpha_B. \quad (25)$$

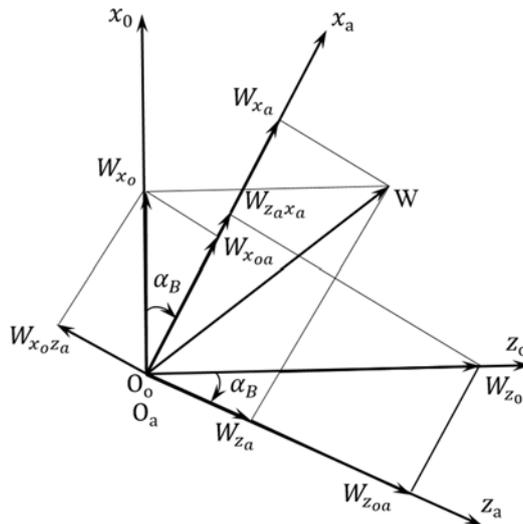


Рис. 4. Земная и аэродромная системы координат на горизонтальной плоскости

После подстановки значений, составляющих вектора W из уравнений (22)–(24) в уравнения (20) и (21) составляющие вектора путевой скорости будут определяются следующими равенствами:

$$W_{x_a} = W_{x_0} \cos \alpha_B + W_{z_0} \sin \alpha_B; \quad (26)$$

$$W_{z_a} = W_{z_0} \cos \alpha_B + W_{x_0} \sin \alpha_B. \quad (27)$$

Интегрирование составляющих W_{x_a} , W_{z_a} позволяет вычислить координаты беспилотного летательного аппарата в аэродромной системе, расстояние S_{x_a} начала координат и боковое отклонение S_{z_a} , от оси взлетно-посадочной полосы:

$$S_{xa} = \int_0^t W_{xa} dt ; \quad (28)$$

$$S_{za} = \int_0^t W_{za} dt . \quad (29)$$

Координаты S_{xa} , S_{za} необходимы для записи маршрута имитации работы навигационных автоматов.

В зависимости от направления, относительно которого ориентируется продольная ось беспилотного летательного аппарата, курс может быть истинным, магнитным, ортодромическим. Ортодромический курс условный и определяется углом между направлением условного меридиана и проекцией на горизонтальную плоскость продольной оси самолета. За направление условного меридиана можно принять истинный меридиан аэродрома взлета или любой точки маршрута, направление ВПП и ортодромии. В ортодромической системе координат в качестве условного меридиана принимается направление главной ортодромии для района полета. Условный курс, измеряемый в ортодромической системе координат, называется ортодромическим курсом (ОК или ψ_0) (рис. 5). При решении задачи управления беспилотного летательного аппарата на плоскости ортодромические координаты являются условными прямоугольными координатами $O_0x_0y_0z_0$. Значение ортодромического курса получается непрерывным интегрированием угловой скорости разворота беспилотного летательного аппарата $\dot{\psi} = \frac{d\psi}{dt}$, т.е.

$$\psi_0 = \psi_n + \int_0^t \dot{\psi} dt , \quad (30)$$

где ψ_n – начальный ортодромический курс.

Истинный курс:

$$\psi_n = \psi_0 - \sigma , \quad (31)$$

где σ – угол схождения между текущим географическим меридианом места самолета и направлением оси O_0x_0 системы координат.

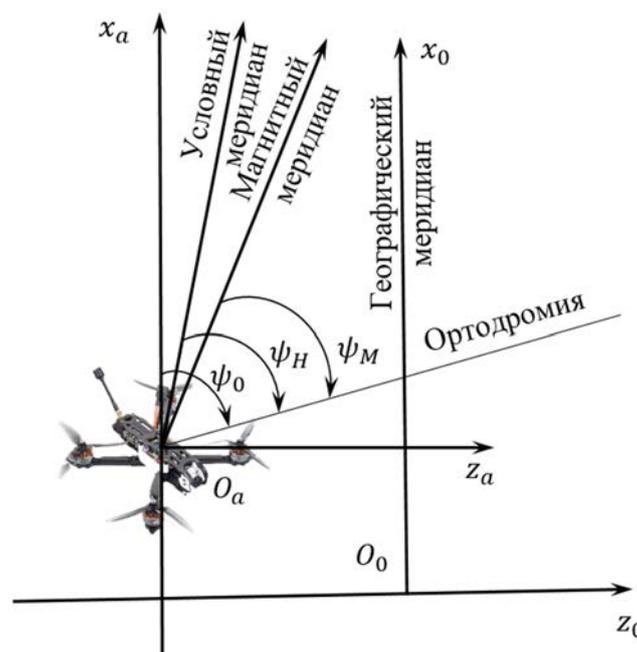


Рис. 5. Курсы БЛА и зависимости между ними

Для определения магнитного курса ψ_m должна учитываться величина магнитного склонения Δ_m , поэтому

$$\psi_m = \psi_0 - \sigma - \Delta_m. \quad (32)$$

Ортодромический курс может быть также определен из соотношения

$$\psi_m = \psi_0 + \sigma + \Delta_m. \quad (33)$$

Текущие значения угла схождения меридианов σ и географические координаты (широта φ и долгота λ) вычисляются по формулам связи географических и ортодромических координат. Географические координаты точки O_0 (широта φ_0 и долгота λ_0) определяются по карте, поэтому углы схождения меридианов можно рассчитать как

$$\sigma = \Delta\lambda \sin \varphi_0. \quad (34)$$

Заключение

Математическая модель полета беспилотного летательного аппарата [6], математическая модель системы управления беспилотного летательного аппарата [7] и приведенные выше уравнения навигации позволяют решить задачу управления полетом и задачи ближней и дальней навигации.

Список литературы

1. Остославский И. В., Стражева И. В. Динамика полета: траектории летательных аппаратов. М. : Машиностроение, 1969. 505 с.
2. Алхаддад Мухаммад. Моделирование и управление ориентацией квадрокоптера с использованием линейного квадратического регулятора // Актуальные проблемы авиации и космонавтики. 2016. Т. 1, № 12. С. 883–886.
3. Горбатенко С. А., Макашов Э. М., Полушкин Ю. Ф. [и др.]. Механика полета. Общие сведения. Уравнения движения. М. : Машиностроение, 1969. 419 с.
4. Лебедев А. А., Чернобровкин Л. С. Динамика полета беспилотных летательных аппаратов. М. : Машиностроение, 1973. 615 с.
5. Красовский Н. Н., Летов А. М. К теории аналитического конструирования регуляторов // Автоматика и телемеханика. 1962. № 6. С. 713–720.
6. Годунов А. И., Суздальцев П. С., Жежук А. А. [и др.]. Математическая модель полета беспилотного летательного аппарата // Надежность и качество сложных систем. 2024. № 1 (45). С. 21–31.
7. Годунов А. И., Суздальцев П. С., Куканов С. А., Мухамбетов А. М. Математическая модель системы управления беспилотного летательного аппарата // Надежность и качество сложных систем. 2024. № 2 (46). С. 25–31.

References

1. Ostoslavskiy I.V., Strazheva I.V. *Dinamika poleta: traektorii letatel'nykh apparatov = Flight dynamics: trajectories of aircraft*. Moscow: Mashinostroenie, 1969:505. (In Russ.)
2. Alkhaddad Mukhammad. Modeling and control of the orientation of a quadcopter using a linear quadratic controller. *Aktual'nye problemy aviatsii i kosmonavтики = Actual problems of aviation and cosmonautics*. 2016;1(12):883–886. (In Russ.)
3. Gorbatenko S.A., Makashov E.M., Polushkin Yu.F. et al. *Mekhanika poleta. Obshchie svedeniya. Uravneniya dvizheniya = Mechanics of flight. General information. Equations of motion*. Moscow: Mashinostroenie, 1969:419. (In Russ.)
4. Lebedev A.A., Chernobrovkin L.S. *Dinamika poleta bespilotnykh letatel'nykh apparatov = Flight dynamics of unmanned aerial vehicles*. Moscow: Mashinostroenie, 1973:615. (In Russ.)
5. Krasovskiy N.N., Letov A.M. On the theory of analytical design of regulators. *Avtomatika i telemekhanika = Automation and telemechanics*. 1962;(6):713–720. (In Russ.)
6. Godunov A.I., Suzdal'tsev P.S., Zhezhuk A.A. et al. Mathematical model of flight of an unmanned aerial vehicle. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems*. 2024;(1):21–31. (In Russ.)
7. Godunov A.I., Suzdal'tsev P.S., Kukanov S.A., Mukhambetov A.M. Mathematical model of the control system of an unmanned aerial vehicle. *Nadezhnost' i kachestvo slozhnykh system = Reliability and quality of complex systems*. 2024;(2):25–31. (In Russ.)

Информация об авторах / Information about the authors

Анатолий Иванович Годунов

доктор технических наук, профессор,
заслуженный деятель науки РФ,
профессор кафедры автоматики и телемеханики,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: avitelpgu@mail.ru

Сергей Анатольевич Куканов

кандидат технических наук,
доцент кафедры средств ближнего боя,
Филиал Военной академии
материально-технического
обеспечения имени генерала армии
А. В. Хрулева в г. Пензе
(Россия, г. Пенза, Военный городок)
E-mail: Kuk_@mail.ru

Павел Сергеевич Суздальцев

адъюнкт,
Филиал Военной академии
материально-технического
обеспечения имени генерала армии
А. В. Хрулева в г. Пензе
(Россия, г. Пенза, Военный городок)
E-mail: suzdal.1990@bk.ru

Anatoly I. Godunov

Doctor of technical sciences, professor,
honored worker of Russia,
professor of the sub-department
of automation and telemechanics,
Penza State University
(40 Krasnaya street, Penza, Russia)

Sergei A. Kukanov

Candidate of technical sciences, associate professor
of the sub-department of melee weapons,
Branch of the Military Academy
of Logistics named after Army General
A.V. Khrulev in Penza
(Military town, Penza, Russia)

Pavel S. Suzdaltsev

Adjunct,
Branch of the Military Academy
of Logistics named after Army General
A.V. Khrulev in Penza
(Military town, Penza, Russia)

Авторы заявляют об отсутствии конфликта интересов /

The authors declare no conflicts of interests.

Поступила в редакцию / Received 27.05.2024

Поступила после рецензирования / Revised 06.07.2024

Принята к публикации / Accepted 18.08.2024

ИССЛЕДОВАНИЕ АЛГОРИТМОВ ИЗМЕРЕНИЯ СОСТАВЛЯЮЩИХ ВЕКТОРА СКОРОСТИ РАДИОВЫСОТОМЕРНЫХ СИСТЕМ, АДАПТИВНЫХ К ПОДСТИЛАЮЩЕЙ ПОВЕРХНОСТИ

А. В. Васильева

Уральское проектно-конструкторское бюро «Деталь», Каменск-Уральский, Россия
niokr@upkb.ru

Аннотация. *Актуальность и цели.* Современный этап развития авиации характеризуется устойчивой тенденцией к повышению требований к тактико-техническим характеристикам летательных аппаратов, в том числе к системам маршрутной навигации. Естественное следствие такой тенденции – повышение роли и значения радиолокационных датчиков навигационной информации. К таковым относятся радиовысотомерные системы (РВС), обеспечивающие измерение высоты и составляющих вектора скорости в связанной системе координат. В настоящее время актуальной задачей является повышение точности измерения параметров движения летательных аппаратов над неоднородными сложными поверхностями. *Материалы и методы.* Рассмотрены следующие способы решения поставленной задачи: адаптация алгоритмов оценки составляющих вектора скорости РВС к подстилающей поверхности путем изменения частоты следования зондирующих импульсов в зависимости от значения высоты полета летательного аппарата; адаптация алгоритмов оценки составляющих вектора скорости РВС к подстилающей поверхности путем изменения конфигурации антенной системы; адаптация алгоритмов оценки составляющих вектора скорости РВС к подстилающей поверхности путем изменения несущей частоты. *Результаты и выводы.* Применение рассмотренных способов позволит существенно повысить тактико-технические характеристики как существующих РВС, так и новых разработок, сократит сроки их отработки.

Ключевые слова: радиовысотомерная система; радионавигация; точностные характеристики

Для цитирования: Васильева А. В. Исследование алгоритмов измерения составляющих вектора скорости радиовысотомерных систем, адаптивных к подстилающей поверхности // Надежность и качество сложных систем. 2024. № 3. С. 112–117. doi: 10.21685/2307-4205-2024-3-12

INVESTIGATION OF ALGORITHMS FOR MEASURING THE COMPONENTS OF THE VELOCITY VECTOR OF RADIO ALTIMETER SYSTEMS ADAPTIVE TO THE UNDERLYING SURFACE

A.V. Vasilieva

Ural design bureau "Detal", Kamensk-Uralsky, Russia
niokr@upkb.ru

Abstract. *Background.* The current stage of aviation development is characterized by a steady tendency to increase requirements for the tactical and technical characteristics of aircraft, including navigation systems. A natural consequence of this tendency is the increasing role and importance of radar sensors of navigation information. These include radio altimeter systems (RAS), which measure the altitude and components of the velocity vector in a linked coordinate system. At present, an urgent task is to improve the accuracy of RAS characteristics over heterogeneous complex surfaces. *Materials and methods.* Considers the following methods for solving the problem: adaptation of algorithms for estimating the components of the RAS velocity vector to the underlying surface by changing the pulses time depending on the value of the aircraft flight altitude; adaptation of algorithms for estimating the components of the RAS velocity vector to the underlying surface by changing the antenna system configuration; adaptation of algorithms for estimating the components of the RAS velocity vector to the underlying surface by switching the RAS frequency. *Results and conclusions.* Methods usage will significantly improve the tactical and technical characteristics of both existing RAS and new developments, and will reduce the time of development.

Keywords: radio altimeter system; radio navigation; accuracy characteristics

For citation: Vasilieva A.V. Investigation of algorithms for measuring the components of the velocity vector of radio altimeter systems adaptive to the underlying surface. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):112–117. (In Russ.). doi: 10.21685/2307-4205-2024-3-12

Достигнутый уровень развития авиации отличается высокими показателями эффективности решения целевых задач, эксплуатационными характеристиками, в частности, предъявляются жесткие требования по внешним воздействующим факторам в ожидаемых условиях эксплуатации, требования к точности, надежности и отказобезопасности систем, а также требования к массогабаритным характеристикам и энергопотреблению.

Возникающие требования оказывают непосредственное влияние на всю совокупность тактико-технических характеристик (ТТХ), предъявляемых в том числе к навигационным системам; вызывают необходимость переоценки существующих возможностей и создания новых методов решения существующих задач. Естественное следствие такой тенденции – повышение роли и значения радионавигационных средств.

В данной статье приведены результаты исследования алгоритмов измерения составляющих вектора скорости радиовысотомерных систем (РВС), адаптивных к подстилающей поверхности. Рассмотренные далее способы направлены на обеспечение точностных характеристик РВС над сложными поверхностями.

Первый способ – адаптация алгоритма оценки составляющих вектора скорости РВС к подстилающей поверхности путем изменения частоты следования зондирующих импульсов в зависимости от значения высоты полета летательного аппарата (ЛА).

При полете ЛА над неоднородной поверхностью (поверхность, сочетающая в себе участки с разной рассеивающей способностью) происходит искажение формы взаимно корреляционных функций (ВКФ). При этом в аппроксимируемую область попадает как узкая (соответствует пролету над поверхностью с широкой диаграммой обратного рассеяния (ДОР)), так и широкая часть (соответствует пролету над поверхностью с узкой ДОР) ВКФ.

Моделирование работы скоростного канала над неоднородными поверхностями подтвердило (рис. 1), что полет над поверхностью с чередующимися участками суши и водной поверхности на интервале накопления ВКФ приводит к искажению формы ВКФ и некорректной их аппроксимации.

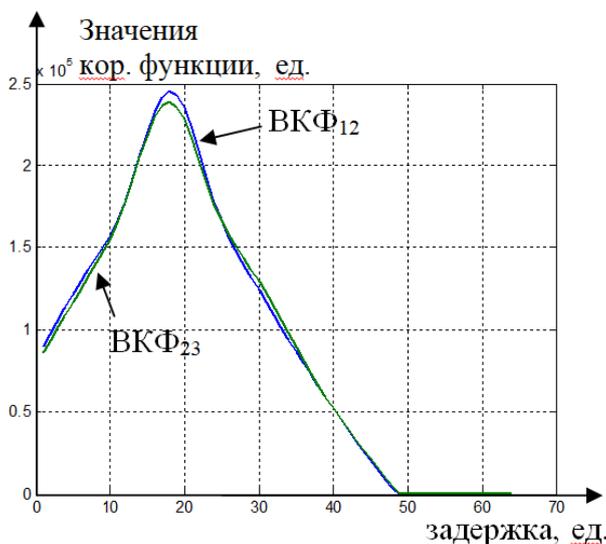


Рис. 1. Результаты моделирования работы РВС над неоднородной поверхностью

Для устранения вышеописанного явления необходимо осуществлять аппроксимацию искаженной ВКФ только по узкой части. Это возможно путем увеличения частоты следования импульсов зондирующего сигнала, в связи с чем предложен алгоритм обработки сигнала с изменением дискрета ВКФ в зависимости от высоты полета ЛА (блок-схема приведена на рис. 2). Следует отметить, что дискрет ВКФ определяется средним периодом повторения импульса T_{cp} и количеством импульсов с фазами 0 и $\pi/2$, принятых 1-й, 2-й и 3-й антеннами [1].

Для подтверждения эффективности предложенного решения проведен анализ алгоритма обработки сигнала с изменением периода повторения зондирующего сигнала в зависимости от высоты полета ЛА над неоднородной поверхностью посредством моделирования работы РВС. Разработанный алгоритм позволяет уменьшить флуктуации измеряемых параметров V_x , V_z в 2,5 раза, а также исключить их постоянное смещение над неоднородной поверхностью.

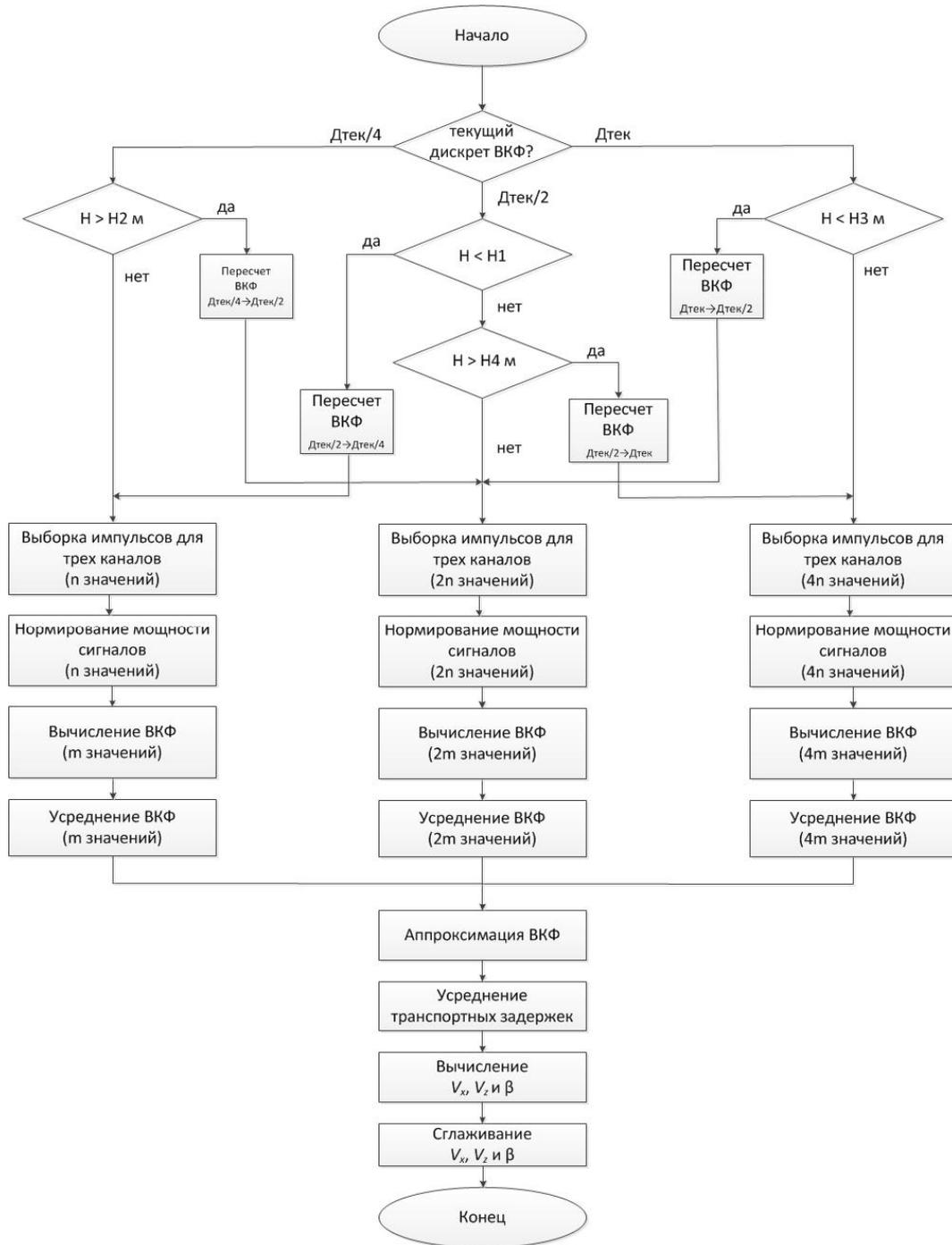


Рис. 2. Блок-схема алгоритма обработки сигнала с изменением дискрета ВКФ

Второй способ – адаптация алгоритма оценки составляющих вектора скорости РВС к подстилающей поверхности путем изменения конфигурации антенной системы (АС).

Максимальные погрешности измерения поперечной составляющей вектора скорости V_z возникают при полете над поверхностями с узкой ДОР, поскольку ВКФ при этом расширяется, а крутизна дискриминационной характеристики снижается [2]. При использовании АС, представленной на рис. 3, появляется возможность снижения погрешности поперечной составляющей вектора скорости над такими поверхностями путем увеличения поперечного разноса антенн, сигналы которых сравниваются. Действительно, при построении ВКФ пар антенн № 1, 5 и № 2, 3 поперечный параметр антенны Z_0 увеличивается в три раза, во столько же раз уменьшается погрешность определения V_z .

Следует отметить, что при полете над поверхностью с широкой ДОР для получения оценок составляющих вектора скорости целесообразно использовать одну из троек, образующих равнобед-

ренный треугольник (геометрия идентична классическому построению АС, представленному в работе [3]), в основании которого две рядом расположенные антенны, например, антенны № 1, 3, 4.

Поскольку предлагаемая АС содержит дополнительные приемные антенны, то появляется возможность уменьшить погрешность измерения составляющих вектора скорости: при построении ВКФ двух комбинаций антенн № 1, 3, 4 и № 2, 4, 5 получаем две независимые оценки.

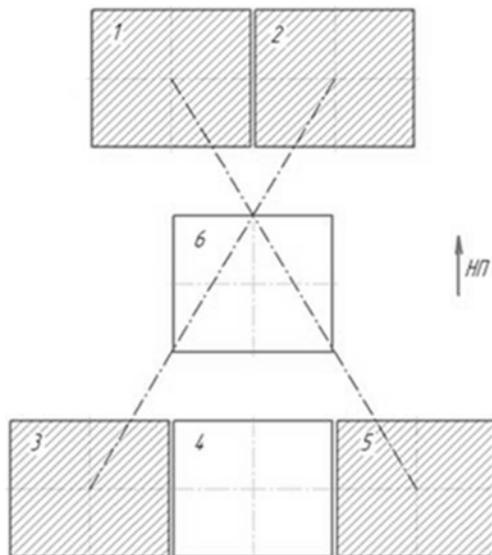


Рис. 3. Геометрия АС

В силу симметричного расположения названных выше троек антенн возможные медленно меняющиеся погрешности боковой составляющей скорости, вызванные взаимным влиянием антенн (отклонение оси ДНА, увеличение расстояния между фазовыми центрами), противоположны по знаку и компенсируются при усреднении полученных двух оценок [4].

Для оценки эффективности применения в РВС предлагаемой конфигурации АС было проведено моделирование РВС над поверхностями с различной отражающей способностью. В связи с необходимостью использования разных комбинаций антенн АС разработан алгоритм переключения в зависимости от ширины ДОР подстилающей поверхности [5].

По результатам моделирования:

– подтверждены теоретические постулаты: над поверхностью с широкой ДОР эффективно использовать АС с обычной поперечной базой (рис. 3: антенны № 1, 3 и 4), над поверхностью с узкой ДОР – АС с расширенной поперечной базой (антенны № 1, 2, 3 и 5). Увеличение поперечной базы АС в три раза дает возможность уменьшения погрешности измерения параметра V_z над поверхностями с узкой ДОР также в 3 раза;

– подтверждена эффективность использования двух независимых оценок составляющих вектора скорости для двух комбинаций антенн (рис. 3: антенны № 1, 3, 4 и № 2, 4, 5) над поверхностью с широкой ДОР – погрешность измерения поперечной составляющей вектора скорости V_z уменьшается в $\sqrt{2}$ раз;

– разработан и проверен алгоритм переключения комбинации антенн в предлагаемой АС в зависимости от ширины ДОР поверхности.

На данную АС получен патент № 2601441 на изобретение «Антенное устройство радиовысотомерной системы с повышенной точностью измерения поперечной составляющей скорости», зарегистрирован в Государственном реестре изобретений Российской Федерации 12 октября 2016 г.

Третий способ – адаптация алгоритмов оценки составляющих вектора скорости РВС к подстилающей поверхности путем переключения несущей частоты

Как акцентировано ранее, максимальные погрешности измерения поперечной составляющей вектора скорости V_z возникают при полете ЛА над поверхностью с узкой ДОР:

$$\varphi(\beta) = \frac{8\pi\Delta_z^2 X_0 Z_0}{\lambda^2} \sin(2\beta). \quad (1)$$

В свою очередь, крутизна дискриминационной характеристики S является первой производной по углу сноса:

$$S = \frac{d\varphi}{d\beta} = \frac{16\pi\Delta_3^2 X_0 Z_0}{\lambda^2} \cos(2\beta). \quad (2)$$

При малых углах сноса $\cos(2\beta) \approx 1$, поэтому в выражении (2) можно опустить $\cos(2\beta)$:

$$S = \frac{16\pi\Delta_3^2 X_0 Z_0}{\lambda^2}. \quad (3)$$

Из выражения (3) следует, что крутизна дискриминационной характеристики S падает пропорционально квадрату эффективной ширины ДНА Δ_3 , что и является причиной увеличения погрешности измерения поперечной составляющей вектора скорости V_z над поверхностью с узкой ДОР.

Крутизна дискриминационной характеристики оказывает влияние на погрешность измерения угла сноса: так, при малой крутизне малейшее расхождение максимумов ВКФ, вызванное, например, их искажением, приводит к значительному изменению угла сноса. При большой крутизне то же самое изменение угла сноса требует существенного расхождения максимумов ВКФ, которое уже не может быть вызвано их искажением [6, 7].

Следовательно, для снижения погрешности измерения параметра V_z необходимо повышать крутизну дискриминационной характеристики. В соответствии с выражением (3) крутизна дискриминационной характеристики растет:

- прямо пропорционально: продольному параметру АС X_0 ; поперечному параметру АС Z_0 ;
- обратно пропорционально: квадрату длины волны λ .

Следует отметить, что уменьшение длины волны λ в два раза для снижения погрешности измерения параметра V_z над гладкой поверхностью эквивалентно расширению поперечной базы АС Z_0 в 4 раза. Уменьшение длины волны над поверхностью с узкой ДОР приводит к сужению ВКФ и, в свою очередь, повышает точность определения максимума ВКФ, что благоприятно и для измерения продольной составляющей вектора скорости V_x . Таким образом, для повышения точности измерения составляющих вектора скорости РВС над поверхностью с узкой ДОР целесообразно уменьшать длину волны.

Для подтверждения эффективности предложенного решения было проведено моделирование РВС над поверхностью с узкой ДОР в разных частотных диапазонах.

Результаты моделирования подтвердили, что уменьшение длины волны в 2(3) раза обеспечивает погрешности измерения составляющих вектора скорости над поверхностями с узкой ДОР соответственно в 2(3) раза меньше по сравнению со штатной длиной волны λ .

Предложенный способ адаптации к подстилающей поверхности путем переключения несущей частоты (λ и $\lambda/3$) реализован в программном обеспечении РВС. Эффективность подтверждена в ходе проведения автономных летных испытаний (АЛИ) РВС.

По результатам АЛИ, сравнение корреляционных характеристик над различными типами подстилающих поверхностей показало, что:

- над поверхностью с узкой ДОР (водная поверхность, ширина ДОР порядка 7 град.) корректно осуществлялось измерение параметров с длиной волны $\lambda/3$;
- над поверхностью с широкой ДОР (лесной массив, ширина ДОР порядка 30 град.) корректно осуществлялось измерение параметров с длиной волны λ .

На данный способ адаптации алгоритма измерения скорости РВС к подстилающей поверхности путем переключения несущей частоты получен патент № 2672098 на изобретение «Радиовысомерная система с адаптацией к гладкой водной поверхности».

Использование приведенных результатов позволило существенно повысить ТТХ РВС, сократить сроки их отработки и приступить к разработке нового поколения бортовых измерителей параметров движения, удовлетворяющих требованиям систем маршрутной навигации, обусловленным современными реалиями.

Список литературы

1. Васильева А. В., Калмыков Н. Н., Мельников С. А., Соловьев В. В. Модель коррелированных сигналов для имитатора скорости корреляционного измерителя // Вестник Концерна ВКО «Алмаз-Антей». 2016. № 4. С. 33–38.

2. Васильева А. В. Повышение точностных характеристик скоростного канала радиовысотометрической системы конфигурированием антенной системы // Радиолокация, навигация и связь : сб. тр. XXIII Междунар. конф. Воронеж, 2017. С. 657–663.
3. Боркус М. К., Черный А. Е. Корреляционные измерители путевой скорости и угла сноса летательных аппаратов. М. : Сов. радио, 1973. 169 с.
4. Рабинович С. Г. Погрешности измерений. М. : Энергия, 1978.
5. Васильева А. В., Седов Д. П. Малогабаритный радиовысотометр с расширенными функциональными возможностями. М., 2021. 22 с.
6. Пономарев Л. И., Калмыков Н. Н., Кац М. И. [и др.]. Результаты отработки алгоритмов функционирования и перспективы применения комбинированных радиолокационных измерителей в составе бортового комплекса управления полетом летательного аппарата // Вестник Южно-Уральского государственного университета. Сер.: Компьютерные технологии, управление, радиоэлектроника. 2011. Вып. 14. С. 116–119.
7. Маковецкий П. В., Васильев В. Г. Отражения радиолокационных сигналов. Л., 1975. 50 с.

References

1. Vasil'eva A.V., Kalmykov N.N., Mel'nikov S.A., Solov'ev V.V. A model of correlated signals for a speed simulator of a correlation meter. *Vestnik Kontserna VKO «Almaz-Antey» = Bulletin of the East Kazakhstan Region Almaz-Antey Concern*. 2016;(4):33–38. (In Russ.)
2. Vasil'eva A.V. Improving the accuracy characteristics of the high-speed channel of the radio altimeter system by configuring the antenna system. *Radiolokatsiya, navigatsiya i svyaz': sb. tr. XXIII Mezhdunar. konf. = Radar, navigation and communications : proceedings of the XXIII International conference*. Voronezh, 2017:657–663. (In Russ.)
3. Borkus M.K., Chernyy A.E. *Korrelyatsionnye izmeriteli putevoy skorosti i ugla snosa letatel'nykh apparatov = Correlation meters of ground speed and angle of demolition of aircraft*. Moscow: Sov. radio, 1973:169. (In Russ.)
4. Rabinovich S.G. *Pogreshnosti izmereniy = Measurement errors*. Moscow: Energiya, 1978. (In Russ.)
5. Vasil'eva A.V., Sedov D.P. *Malogabaritnyy radiovysotomer s rasshirennymi funktsional'nymi vozmozhnostyami = Small-sized radio altimeter with extended functionality*. Moscow, 2021:22. (In Russ.)
6. Ponomarev L.I., Kalmykov N.N., Kats M.I. et al. Results of working out algorithms for the functioning and prospects for the use of combined radar meters as part of the on-board flight control system of an aircraft. *Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Ser.: Komp'yuternye tekhnologii, upravlenie, radioelektronika = Bulletin of the South Ural State University. Ser.: Computer technologies, management, radio electronics*. 2011;(14):116–119. (In Russ.)
7. Makovetskiy P.V., Vasil'ev V.G. *Otrazheniya radiolokatsionnykh signalov = Reflections of radar signals*. Leningrad, 1975:50. (In Russ.)

Информация об авторах / Information about the authors

Анна Валерьевна Васильева

ведущий инженер,
Уральское проектно-конструкторское
бюро «Деталь»
(Россия, г. Каменск-Уральский, ул. Пионерская, 8)
E-mail: niokr@upkb.ru

Anna V. Vasilieva

Leading engineer,
Ural design bureau "Detal"
(8 Pionerskaya street, Kamensk-Uralsky, Russia)

Автор заявляет об отсутствии конфликта интересов /

The author declares no conflicts of interests.

Поступила в редакцию / Received 23.05.2024

Поступила после рецензирования / Revised 25.06.2024

Принята к публикации / Accepted 13.08.2024

РАЗРАБОТКА ПРОТОТИПА АРХИТЕКТУРЫ ИНТЕЛЛЕКТУАЛЬНОГО МИКРОКОНТРОЛЛЕРА ПРИ РЕАЛИЗАЦИИ ПРИЛОЖЕНИЙ ИНТЕРНЕТА ВЕЩЕЙ

Ю. Т. Зырянов¹, Д. Ш. Калхиташвили², С. П. Хрипунов³

¹ Тамбовский государственный технический университет, Тамбов, Россия

² Российская академия народного хозяйства и государственной службы
при Президенте Российской Федерации, Москва, Россия

³ Институт проблем управления имени В. А. Трапезникова РАН, Москва, Россия

¹ zut-tmb@mail.ru, ² davidkalkhitashvili@gmail.com, ³ hsp61@yandex.ru

Аннотация. *Актуальность и цели.* В рамках современной парадигмы цифровой трансформации общества и бурного роста количества цифровых технологий внедрение Интернета вещей (IoT) происходит во все сферы жизни человека. Данные технологии применяются в промышленном секторе экономики и в частной жизни. Разработка и внедрение Интернета вещей – это парадигма, заключающаяся в разработке связанных между собой технических объектов при помощи различных методов и протоколов связи, с передачей данных в облачный центр. Данная технология имеет ряд ограничений: в виде объема памяти, электрического заряда и низкой эффективности процессора. Второе ограничение – отсутствие алгоритмов машинного обучения. В связи с этим требуются новые методы и инженерные подходы для решения данной проблемы. Цель статьи – разработка архитектуры прототипа микроконтроллера для контроля технического состояния информационно-измерительных и управляющих систем на основе концепции Интернета вещей и алгоритмов машинного обучения для обработки многопоточных источников данных. *Материалы и методы.* В качестве объекта исследования выступает архитектура микроконтроллера Интернета вещей. Для решения поставленной задачи применен метод системного подхода. *Результаты.* Разработана архитектура интеллектуального микроконтроллера с использованием шины для шлюзов, а также с применением технологии MEMS и алгоритмов машинного обучения. *Выводы.* Данный подход позволяет обрабатывать многопоточные данные в динамическом режиме с применением алгоритмов машинного обучения, используемые в парадигме Интернета вещей, для разработки алгоритмов контроля технического состояния информационно-измерительных и управляющих систем.

Ключевые слова: приложения Интернета вещей, информационно-измерительные и управляющие системы, алгоритмы машинного обучения, архитектура интеллектуального микроконтроллера

Для цитирования: Зырянов Ю. Т., Калхиташвили Д. Ш., Хрипунов С. П. Разработка прототипа архитектуры интеллектуального микроконтроллера при реализации приложений интернета вещей // Надежность и качество сложных систем. 2024. № 3. С. 118–123. doi: 10.21685/2307-4205-2024-3-13

DEVELOPMENT OF A PROTOTYPE OF INTELLIGENT MICROCONTROLLER ARCHITECTURE WHEN IMPLEMENTING INTERNET OF THINGS APPLICATIONS

Yu.T. Zyryanov¹, D.Sh. Kalkhitashvili², S.P. Khripunov³

¹ Tambov State Technical University, Tambov, Russia

² Russian Academy of National Economy and Public Administration
under the President of the Russian Federation, Moscow, Russia

³ Institute of Management Problems named after V.A. Trapeznikov RAS, Moscow, Russia

¹ zut-tmb@mail.ru, ² davidkalkhitashvili@gmail.com, ³ hsp61@yandex.ru

Abstract. *Background.* Within the framework of the modern paradigm of the digital transformation of society and the rapid growth of the number of digital technologies, the introduction of the Internet of things (IoT) occurs in all spheres of human life. These technologies are used in the industrial sector of the economy and in private life. The development and implementation of the Internet of Things is a paradigm that consists of the development of interconnected technical objects using various communication methods and protocols, with data transferred to a cloud center.

This technology has a number of limitations: in the form of memory capacity, electrical charge and low processor efficiency. Second limitation: lack of machine learning algorithms. In this regard, new methods and engineering approaches are required to solve this problem. Purpose of the article – development of a microcontroller prototype architecture for monitoring the technical condition of information, measuring and control systems (IMS) based on the concept of the Internet of Things and machine learning algorithms for processing multi-threaded data sources. *Materials and methods*: the object of research is the architecture of the Internet of Things microcontroller. To solve the problem, a systematic approach method was used. *Results*. An intelligent microcontroller architecture has been developed using a gateway bus, as well as using MEMS technology and machine learning algorithms. *Conclusions*. This approach allows you to process multi-threaded data in a dynamic mode using machine learning algorithms, used in the Internet of Things paradigm to develop algorithms for monitoring the technical condition of automated control systems.

Keywords: Internet of Things applications, information-measuring and control systems, machine learning algorithms, intelligent microcontroller architecture

For citation: Zyryanov Yu.T., Kalkhitashvili D.Sh., Khripunov S.P. Development of a prototype of intelligent microcontroller architecture when implementing Internet of things applications. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems*. 2024;(3):118–123. (In Russ.). doi: 10.21685/2307-4205-2024-3-13

Введение

В последнее время ведется активная работа по разработке концептуальных идей, связанных с интернет-вещами, а также их воплощения, а также созданию различных подходов с их использованием. В научной литературе представлены инженерные решения технических объектов и их элементов на уровне аппаратной среды [1–6]. Разрабатываются различные компоненты данных технических объектов и усвершенствуются готовые решения. Необходимо отметить, что основной задачей интернет-вещей является процесс обмена данных с облачными центрами. Иными словами, Интернет вещи выполняет роль маршрутизатора: от сенсорных устройств передают данные в облачный центр, после чего получают обратно сигнал о выполнении той или иной функции, после чего микроконтроллер данного устройства подает электромагнитный сигнал на движущее устройство [7]. Такой принцип работы не позволяет в полной мере называть данную категорию технических объектов – «умными устройствами», так как они не выполняют интеллектуальных функций внутри себя, а выполняют роль ретранслятора сигнала для исполнительного устройства (системы). Предлагается внедрить алгоритмы машинного обучения (технология MEMS) [8–10] как один из приемов, позволяющих увеличить производительность данного микроконтроллера, а также внедрения шлюзов, объединенных сетевой шиной для динамической обработки данных.

Результаты исследования

В качестве объекта исследования использована архитектура программной среды интернет-вещей, представленная на рис. 1.



Рис. 1. Архитектура ASIS программной среды микроконтроллера

Типовая архитектура программной среды микроконтроллера имеет три уровня: восприятия, связи и пользовательский.

Уровень восприятия получает данные с датчиков и передает их пользователю, уровень связи осуществляет процесс передачи данных с помощью протоколов связи, а пользовательский уровень представлен для анализа полученных данных и принятия решения о выборе выполняемой функции (рис. 2).



Рис. 2. Упрощенная схема AS-IS принципа работы «интеллектуального» микроконтроллера

С учетом ограничений вычислительных мощностей и объема электрического заряда батареи были проанализированы разработки микроконтроллеров с использованием шлюзов, позволяющих оптимизировать вычислительные мощности на процесс обмена данными с другими микропроцессами и облачным центром. Основываясь на данных анализа по процессу «бесшовного» обмена данными, был выделен подход к применению общей сетевой и локальных шин, позволяющий снизить нагрузку на процессор (рис. 3).

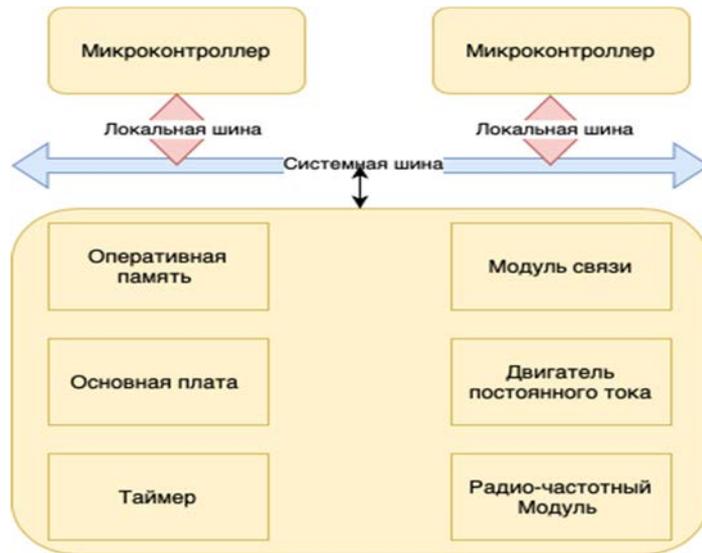


Рис. 3. Архитектура «интеллектуального» микроконтроллера

Использование шин необходимо также для обеспечения работы с динамическими потоками данных. Однако применение шины в свою очередь повлияло на архитектуру программного уровня микроконтроллера. После изменений, вызванных применением шины, архитектура интеллектуального микроконтроллера включает в себя [11, 12]: основную плату, модуль связи, двигатель постоянного тока, оперативную память, радиочастотный модуль, таймер.

Для добавления функции применения алгоритмов машинного обучения в работу микроконтроллера и для повышения эффективности обработки потока данных использовалась технология МЭМС (microelectromechanical systems MEMS).

После того, как данные подходы были применены, в архитектуру программной среды добавляется уровень управления (рис. 4).

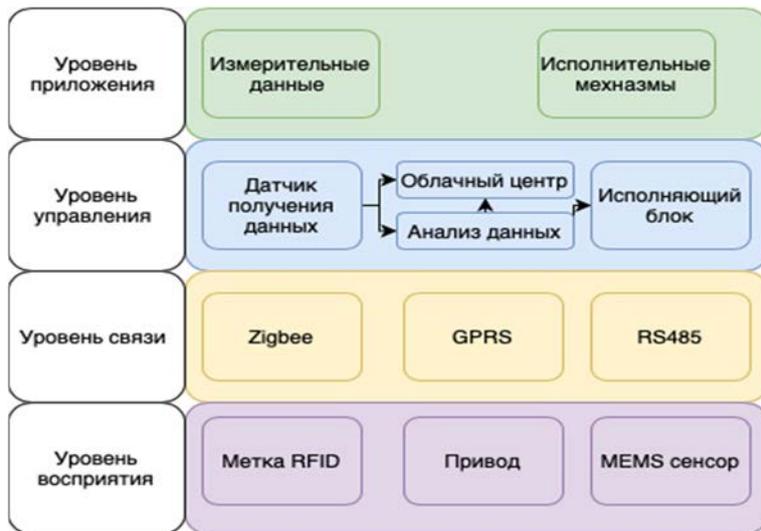


Рис. 4. Архитектура TOBE программной среды «интеллектуального» микроконтроллера

Таким образом, полученная архитектура программной среды с использованием алгоритмов машинного обучения и применение меток RFID [13] придают техническому объекту атрибуты интеллектуального информационно-измерительного устройства (рис. 5).

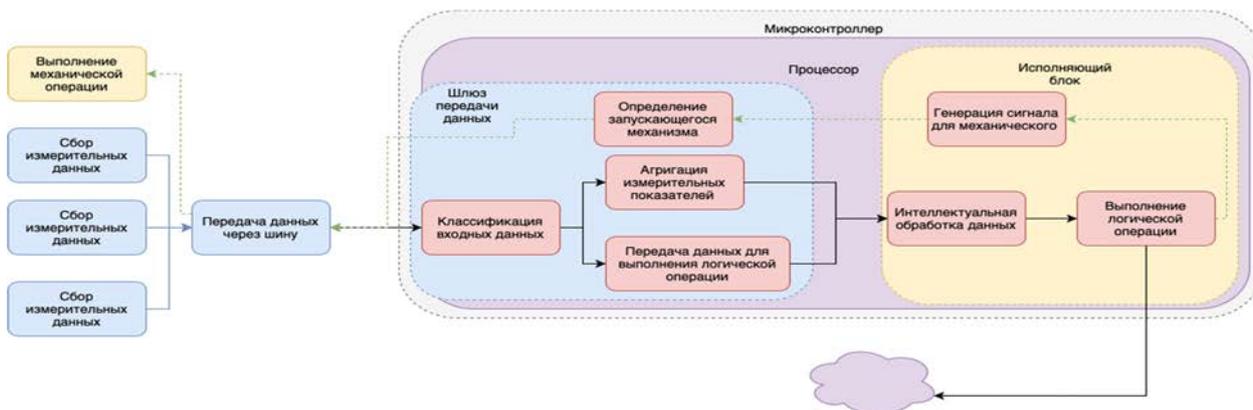


Рис. 5. Упрощенная схема То-Ве принципа работы «интеллектуального» микроконтроллера

В связи с тем, что данное устройство собирает данные из внешней среды, обрабатывает данные и выполняет логические операции, а после посылает сигналы исполнительным системам – это позволяет действовать данной системе автономно.

В табл. 1 представлена сравнительная характеристика функциональных возможностей рассматриваемых архитектур.

Таблица 1

Сравнительная характеристика функциональных возможностей рассматриваемых архитектур

Архитектура	Интеллектуальный микроконтроллер	Стандартный микроконтроллер
Обработка динамических потоков данных	Да	Только их передача и получение готового результата
Наличие алгоритмов машинного обучения	Да	Полное отсутствие
Способ взаимодействия с иными микроконтроллерами	Анализ входных данных и передача результата	Передача входных данных в облачный центр
Участие человека в принятии решения	Не требуется	Требуется активное участие человека
Наличие человеческого фактора	Отсутствует	Присутствует

Заключение

В результате исследования проведен анализ литературы, связанный с разработкой элементов для Интернета вещей, были проанализированы архитектуры Интернета вещей. Модифицирована существующая архитектура, которая позволяет производить обработку многопоточных данных с помощью интеллектуального шлюза в динамическом режиме. Предложенная архитектура позволяет применять алгоритмы машинного обучения для обработки многопоточных источников данных, полученных от сенсорных сетей, что в свою очередь за счет автоматизации процесса позволяет уменьшить расходы на содержание оператора управляющей системы, а также снижает уровень опасности при обслуживании такого оборудования, как электрогенераторы, атомные реакторы, теплоэлектростанции и др.

Список литературы

1. Li Deyi, Du Yi. *Artificial Intelligence with Uncertainty*. New York : Chapman and Hall, 2007. 376 p. doi: 10.1201/9781584889991
2. Наим М. Выбор распределенного шлюза для связи M2M в когнитивных сетях 5G // *IEEE*. 2017. № 31. С. 94–100.
3. Чен Ч. Х. Беспроводной биомеметический датчик для обнаружения активных белков на основе наномеханики // *IEEE Transactions on Biological Engineering*. 2009. № 56. С. 462–470.
4. Чен Ч. Х. Шлюз периферийных вычислений промышленного Интернета вещей с использованием нескольких совместных микроконтроллеров // *IEEE*. 2018. № 32. С. 24–32.
5. Банайе Ф. Анализ производительности многопоточного шлюза Интернета вещей // *IEEE Internet Things J*. 2018. № 6. С. 3143–3155.
6. Кан Б. Интернет всего: крупномасштабный автономный шлюз Интернета вещей // *IEEE Transactions on computing systems*. 2017. № 3. С. 206–214.
7. Интеллектуальный контроллер АТС. URL: <https://energy-utilities.com>
8. Буржуа Дж., Гольдштейн С. С. Распределенные интеллектуальные МЭМС: достижения и перспективы // *Материалы Международной конференции по инновациям в области ИКТ (Охрид, Македония, 12–15 сентября 2012 г.)*. Берлин ; Гейдельберг : Springer, 2012. С. 15–25.
9. Белокопытов М. Л., Ушанев К. В. Модель интеллектуализированной системы контроля и диагностирования бортовой аппаратуры космических аппаратов на этапе их испытаний и применения // *Измерение. Мониторинг. Управление. Контроль*. 2023. № 1. С. 62–69. doi: 10.21685/2307-5538-2023-1-8
10. Вийкари В., Сеппа Х. Концепция RFID Mems-датчика, основанная на интермодуляционных искажениях // *IEEE Sens. J*. 2009. № 9. С. 1918–1923.
11. Лазареску М. Т. Разработка платформы WSN для долгосрочного мониторинга окружающей среды для приложений Интернета вещей // *IEEE J. Emerging and Selected Topics in Circuits and Systems*. 2013. № 3. С. 45–54.
12. Григорьян Л. Р., Богатов Н. М., Григорьян Р. Л. Применение микроконтроллеров в программно-конфигурируемом измерительном тракте коммутационно-тактовой фазоизмерительной аппаратуре // *Измерение. Мониторинг. Управление. Контроль*. 2023. № 1. С. 45–55. doi: 10.21685/2307-5538-2023-1-6
13. Уэлбурн Э. Создание Интернета вещей с использованием RFID: опыт экосистемы RFID // *IEEE Internet Computing*. 2009. № 13. С. 48–55.

References

1. Li Deyi, Du Yi. *Artificial Intelligence with Uncertainty*. New York: Chapman and Hall, 2007:376. doi: 10.1201/9781584889991
2. Naim M. Choosing a distributed gateway for M2M communication in 5G cognitive networks. *IEEE*. 2017;(31):94–100. (In Russ.)
3. Chen Ch.Kh. Wireless biomemetic sensor for detecting active proteins based on nanomechanics. *IEEE Transactions on Biological Engineering*. 2009;(56):462–470. (In Russ.)
4. Chen Ch.Kh. Gateway of peripheral computing of the industrial Internet of Things using several joint microcontrollers. *IEEE*. 2018;(32):24–32. (In Russ.)
5. Banaye F. Performance analysis of the multithreaded gateway of the Internet of Things. *IEEE Internet Things J*. 2018;(6):3143–3155. (In Russ.)
6. Kan B. Internet of Everything: a large-scale autonomous gateway of the Internet of Things. *IEEE Transactions on computing systems*. 2017;(3):206–214. (In Russ.)
7. *Intellektual'nyy kontroller ATS = Intelligent PBX controller*. (In Russ.). Available at: <https://energy-utilities.com>
8. Burzhua Dzh., Gol'dshteyn S.S. Distributed intelligent MEMS: achievements and prospects. *Materialy Mezhdunarodnoy konferentsii po innovatsiyam v oblasti IKT (Okhrid, Makedoniya, 12–15 sentyabrya 2012 g.)* =

Proceedings of the International Conference on Innovations in the field of ICT (Ohrid, Macedonia, September 12-15, 2012). Berlin; Geydel'berg: Springer, 2012:15–25. (In Russ.)

9. Belokopytov M.L., Ushanev K.V. Model of intellectualized systems monitoring and diagnostics of on-board spacecraft equipment at the stage of their testing and application. *Izmerenie. Monitoring. Upravlenie. Kontrol' = Measurement. Monitoring. Management. Control*. 2023;(1):62–69. (In Russ.). doi: 10.21685/2307-5538-2023-1-8
10. Viykari V., Seppa Kh. The concept of an RFID Mems sensor based on intermodulation distortions. *IEEE Sens. J.* 2009;(9):1918–1923. (In Russ.)
11. Lazaresku M.T. Development of the WSN platform for long-term environmental monitoring for Internet of Things applications. *IEEE J. Emerging and Selected Topics in Circuits and Systems*. 2013;(3):45–54. (In Russ.)
12. Grigor'yan L.R., Bogatov N.M., Grigor'yan R.L. The use of microcontrollers in a software-configurable measuring path of switching-contact phase measuring equipment. *Izmerenie. Monitoring. Upravlenie. Kontrol' = Measurement. Monitoring. Management. Control*. 2023;(1):45–55. (In Russ.). doi: 10.21685/2307-5538-2023-1-6
13. Uelburn E. Creation of the Internet of Things using RFID: the experience of the RFID ecosystem. *IEEE Internet Computing*. 2009;(13):48–55. (In Russ.)

Информация об авторах / Information about the authors

Юрий Трифионович Зырянов

доктор технических наук, профессор, профессор кафедры конструирования радиоэлектронных и микропроцессорных систем, Тамбовский государственный технический университет (Россия, г. Тамбов, ул. Советская, 106)
E-mail: zut-tmb@mail.ru

Давид Шалвович Калхиташвили

старший аналитик Дирекции данных, Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации (Россия, г. Москва, пр-кт Вернадского, 82/2)
E-mail: davidkalkhitashvili@gmail.com

Сергей Петрович Хрипунов

доктор технических наук, профессор, ведущий научный сотрудник, Институт проблем управления имени В. А. Трапезникова Российской академии наук (Россия, Москва ул. Профсоюзная, 65)
E-mail: hsp61@yandex.ru

Yuri T. Zyryanov

Doctor of technical sciences, professor, professor of the sub-department of design of radioelectronic and microprocessor systems, Tambov State Technical University (106 Sovetskaya street, Tambov, Russia)

David Sh. Kalkhitashvili

Senior analyst of the Data Directorate, Russian Academy of National Economy and Public Administration under the President of the Russian Federation (82/2 Vernadsky avenue, Moscow, Russia)

Sergey P. Khripunov

Doctor of technical sciences, professor, leading researcher, Institute of Management Problems named after V.A. Trapeznikov of Russian Academy of Sciences (65 Profsoyuznaya street, Moscow, Russia)

Авторы заявляют об отсутствии конфликта интересов /

The authors declare no conflicts of interests.

Поступила в редакцию / Received 27.05.2024

Поступила после рецензирования / Revised 30.06.2024

Принята к публикации / Accepted 15.08.2024

СИСТЕМА ПОДДЕРЖКИ АВИАПАССАЖИРОВ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ МОНИТОРИНГА ИХ МЕСТОПОЛОЖЕНИЯ

Д. В. Егжов¹, Л. А. Надейкина², П. О. Терешко³

^{1, 2, 3} Московский государственный технический университет гражданской авиации, Москва, Россия
¹ mr.danil1509@mail.ru, ² ludmila.nadejkina@yandex.ru, ³ pav.tereshko@yandex.ru

Аннотация. *Актуальность и цели.* Представлена система мониторинга местоположения авиапассажиров в здании аэропорта, предоставляющая сервис оповещения опаздывающих на посадку, а также предоставляющая эффективный способ скорейшего следования к месту посадки. *Материалы и методы.* Предлагается карта построенного кратчайшего пути в мобильном приложении. Система также предоставляет возможность навигации по зданию аэропорта. Пассажир фиксирует на карте конечные точки маршрута, получает в приложении выстроенный маршрут. *Результаты и выводы.* Предлагаемая система обеспечивает повышение качества обслуживания и комфорта пассажиров, а также решение ряда посадочных проблем, что повышает лояльность авиапассажиров.

Ключевые слова: архитектура приложения, паттерны проектирования, БД SQLite, поддержка авиапассажиров, идентификация, мониторинг местонахождения пассажиров, навигация

Для цитирования: Егжов Д. В., Надейкина Л. А., Терешко П. О. Система поддержки авиапассажиров с использованием технологий мониторинга их местоположения // Надежность и качество сложных систем. 2024. № 3. С. 124–129. doi: 10.21685/2307-4205-2024-3-14

AIR PASSENGER SUPPORT SYSTEM USING LOCATION MONITORING TECHNOLOGIES

D.V. Egzhov¹, L.A. Nadeykina², P.O. Tereshko³

^{1, 2, 3} Moscow State Technical University of Civil Aviation, Moscow, Russia
¹ mr.danil1509@mail.ru, ² ludmila.nadejkina@yandex.ru, ³ pav.tereshko@yandex.ru

Abstract. *Background.* The article presents a system for monitoring the location of air passengers in the airport building, providing a service for notifying those who are late for boarding, as well as providing an effective way to quickly get to the boarding point. *Materials and methods.* Offering a map of the shortest route built in a mobile application. The system also provides the ability to navigate around the airport building. The passenger records the end points of the route on the map, receives the built route in the application. *Results and conclusions.* The proposed system improves the quality of service and passenger comfort, as well as solves a number of boarding problems, which increases the loyalty of air passengers.

Keywords: application architecture, design patterns, SQLite database, air passenger support, identification, monitoring of passenger location, navigation

For citation: Egzhov D.V., Nadeykina L.A., Tereshko P.O. Air passenger support system using location monitoring technologies. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):124–129. (In Russ.). doi: 10.21685/2307-4205-2024-3-14

Введение

В связи с ежегодным ростом пассажирского трафика повышение качества обслуживания становится одной из первостепенных задач, которые стоят перед современными аэропортами [1]. Пассажиры, прошедшие регистрацию и получившие посадочный талон, проходят в так называемую «чистую зону», в которой находится большое количество магазинов разного назначения, кафе и ресторанов. Порою пассажиры переоценивают временной и физический ресурс и, как результат, опаздывают на посадку. Еще одним недостатком, который особенно ощущается в крупных аэропортах, является отсутствие способа эффективного ориентирования по зданию аэропорта.

Для решения описанных выше проблем была разработана система мониторинга местоположения авиапассажиров, предоставляющая сервис оповещения опаздывающих на посадку авиапассажиров, а также предоставляющая эффективный способ скорейшего следования их к месту посадки. В мобильное приложение такого авиапассажира будет поступать соответствующее push-уведомление и карта кратчайшего маршрута до места посадки.

Представленная многопоточная программная система поддержки авиапассажиров Росавиации, предназначенная для одновременной работы с большим числом пассажиров, способна при необходимости масштабироваться. При этом приложение обладает следующими критериями качества: быть надежным, масштабируемым и удобным в сопровождении.

Магазины, кафе и прочие организации зачастую уже используют для своей работы web-приложения, а значит, имеют подготовленную для их разворачивания инфраструктуру. Поэтому оптимальным решением является построение системы на базе web-приложения.

Структура модульной системы на базе фундаментального архитектурного паттерна MVC (Model-View-Controller) представлена на рис. 1 и состоит из внешней и внутренней частей.

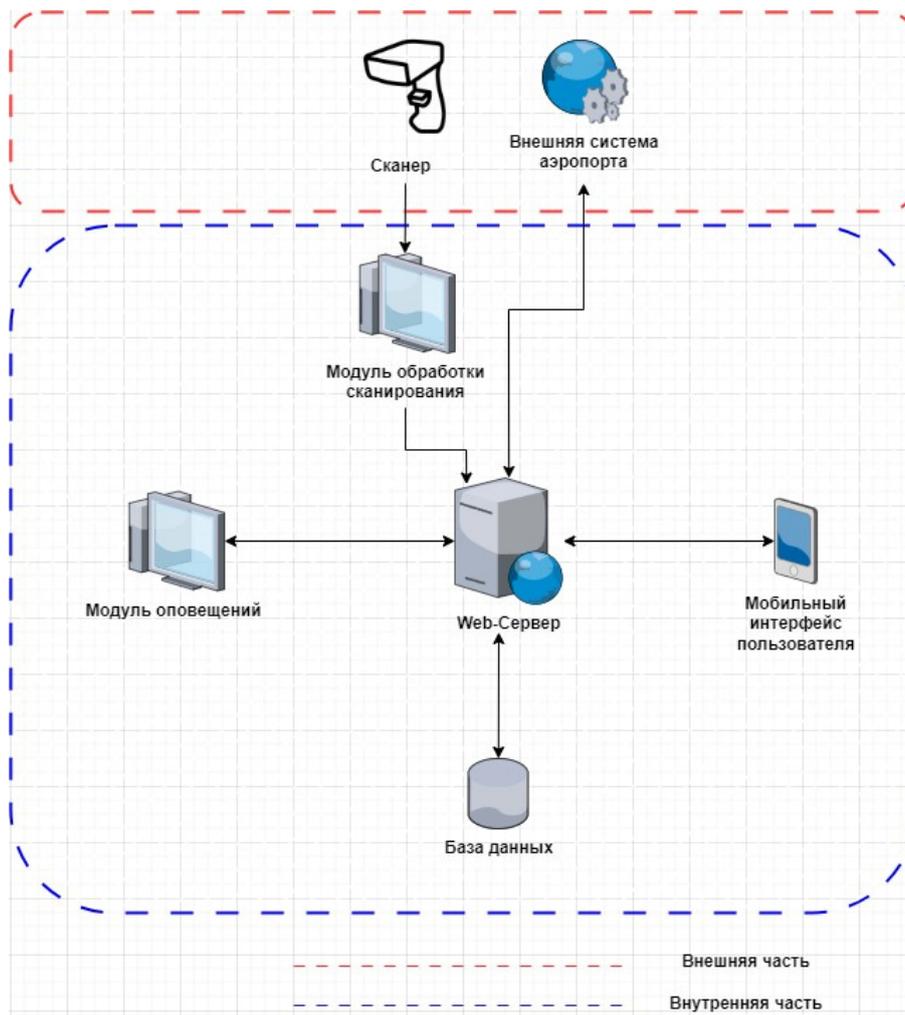


Рис. 1. Структурная схема системы оповещения

Модульная архитектура системы обеспечивает удобство внесения изменений и их интеграцию в компоненты системы либо полную их замены в случае необходимости.

К внешнему сегменту системы относится информационная система аэропорта.

К внутреннему сегменту относятся web-сервер, сервер базы данных, модули сбора данных, оповещения и интерфейс клиентского мобильного приложения.

Web-сервер выполняет роль связующего звена системы [2]. Он опрашивает модуль сбора данных, получая новые данные о посетителе заведения, записывает их в базу данных. Модуль оповещений извлекает данные о пассажире в нужном формате через Web-сервер из базы данных для даль-

нейшей обработки. Web-сервер запрашивает у модуля оповещения информацию: кому из опаздывающих пассажиров нужно направить в мобильное приложение оповещение и при необходимости маршрут до места посадки.

Система оповещения обладает следующим функционалом:

- фиксация местоположения авиапассажира, извлечение данных о рейсе, времени и терминале посадки, определение резерва времени и вероятности опоздания на посадку;
- отправление оповещения в мобильное приложение пассажира о необходимости следовать на посадку и предоставление карты кратчайшего маршрута;
- построение маршрутов на карте аэропорта.

Приложение включает представленные ниже функциональные модули.

Модуль обработки сканирования

Назначение – обработка поступающих сведений от сканера и отправка их в базу данных. Модуль написан на языке программирования Python.

Обеспечивает связь со сканером, обработку данных сканирования посадочного талона, фиксацию координат точки локации авиапассажира и терминала вылета, данные о пассажире, а также идентификатор рейса, которым данный пассажир летит, проводит верификацию данных с помощью запроса к внешней информационной системе аэропорта. Также осуществляет обработку данных к нужному формату и отправку данных в базу данных. В самой базе данных запись о каждом пассажире, зарегистрированном на рейс, содержит сведения о точке последнего сканирования, благодаря этому определяется местоположение авиапассажира. При получении новых данных от модуля обработки сканирования сведения о точке последнего сканирования обновляются.

База данных

Назначение – хранение, прием на хранение и предоставление данных по запросам.

Так как итогом работы будет являться система, с которой будет взаимодействовать большое количество пользователей, для хранения всех данных было решено использовать базу данных. Логическая модель базы данных представлена на рис. 2, включающая следующие таблицы: рейсы, аэропорты вылета/прилета, города и страны вылета/прилета, наполнение которых очевидно.

Таблица «посадочные талоны» хранит информацию о посадочных талонах, зарегистрированных в приложении пассажиров, такую как номер рейса, точка последнего сканирования.

Таблица «пассажиры» хранит информацию о пассажирах, зарегистрированных в приложении.

Таблица «точки маршрута» хранит информацию о списке точек местоположений авиапассажира в предпосадочной зоне.

Таблица «маршруты» – сведения, необходимые для генерации маршрута из точки *A* в точку *B*.

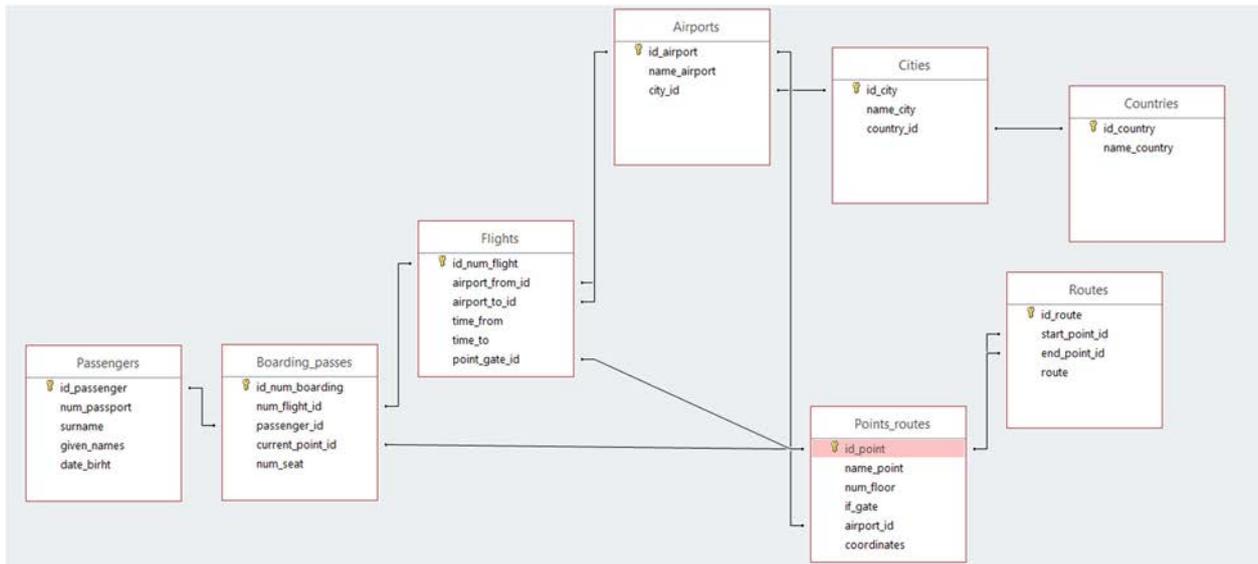


Рис. 2. Логическая модель базы данных

После создания логической модели база данных была реализована с использованием СУБД SQLite. Выбор был обусловлен мультиплатформенностью, а также небольшой загруженностью данной системы управления базами данных, при этом отвечающей всем нужным требованиям создаваемой системы. SQLite эффективно работает с Python, так как его библиотеки позволяют с легкостью работать с созданной базой данных напрямую из кода, без задействования менеджеров СУБД и каких-либо других средств.

Модуль оповещений

Основной модуль системы. Назначения:

- проверка необходимости отправки уведомлений;
- формирование запросов к базе данных;
- запрос данных из внешней информационной системы аэропорта;
- обработка данных, поступающих от мобильного приложения;
- передача данных в приложение.

Для реализации данного модуля применялся язык программирования Python.

Для проверки необходимости вывода уведомления реализуется поведенческий паттерн проектирования «Наблюдатель» [3]. Паттерн «Наблюдатель» определяет зависимость между объектами «один-ко-многим» так, что при изменении состояния одного объекта все зависящие от него объекты уведомляются и обновляются автоматически.

В нашем случае наблюдатель подписывается на уведомления от источников событий и реагирует на изменения состояния в этих источниках. Когда состояние источника изменяется, наблюдатель получает уведомление и выполняет какие-то действия в соответствии с изменением. Обработывается следующее событие.

Событие – обнаружен опаздывающий пассажир. Источником события является модуль оповещений. Наблюдателем – мобильное приложение. Модуль оповещений получает от мобильного приложения текущее время, от базы данных время отправления рейса, а также расстояние от местоположения пассажира до нужного выхода на посадку. Затем производит вычисление критического времени, когда надо отправлять оповещение. Оно определяется как частное от деления расстояния от местоположения пассажира на среднюю скорость ходьбы здорового человека в произвольном темпе, равную 1,43 м/с, согласно расчетам, проведенным в статье [4]. Если разница между текущим временем и временем отправки рейса равна критическому, пользователю отправляется уведомление о необходимости срочно проследовать к выходу на посадку.

Событие – пассажир прошел через выход на посадку. Источником события является модуль обработки сканирования. Наблюдателем – данный модуль оповещений. При проходе через выход на посадку пассажир сканирует свой посадочный талон, модуль обработки сканирования определяет точку сканирования как выход на посадку и уведомляет об этом модуль оповещений. Модуль оповещений со своей стороны удаляет все записи о пассажире из базы данных и завершает сеанс наблюдения за пассажиром.

Для работы системы необходимы предоставляемые аэропортом данные, такие как: сведения о рейсах и списки зарегистрированных на них пассажиров, которые выполнили авторизацию в нашем приложении, сведения об общественных объектах аэропорта, карта-схема доступной посетителям части аэропорта. Получаемые из внешней системы сведения проходят обработку, после чего помещаются в базу данных.

Кроме того, при поступлении запроса о предоставлении данных от мобильного приложения модуль запрашивает необходимые данные из базы данных, затем преобразует в необходимый формат, формирует json-файл и направляет его мобильному приложению как ответ на запрос.

Мобильный интерфейс пользователя

Мобильный интерфейс пользователя выполнен в виде мобильного приложения, написанного на языке Python с использованием одного из самых популярных фреймворков для мобильной разработки на языке Python – Kivy [5]. Служит для взаимодействия авиапассажира с разработанной системой.

В момент регистрации в мобильном приложении пассажир выражает свое согласие на обработку его персональных данных системой оповещения с момента входа в систему до отправки рейса. При авторизации производится сканирование посадочного талона.

Далее приложение отправляет запрос данных о пользователе и рейсе, после получения которых выполняется вход и данные отображаются на главном экране мобильного приложения.

Как только модуль оповещения определил, что пассажир опаздывает, пассажиру посылается соответствующее push-уведомление, в котором также предлагается отобразить маршрут следования к выходу на посадку.

Вместе с этим в приложении имеется карта аэропорта. Кроме основного сервиса, приложение предоставляет возможность проложить маршрут между любыми двумя точками доступной зоны аэропорта. Для этого пользователю требуется перейти в предназначенное для этого окно и, используя предложенную карту аэропорта, зафиксировать на карте две точки – начало и конец маршрута. В этот момент срабатывает паттерн «Наблюдатель», реализованный также и в мобильном приложении, событием для которого является ввод точек начала и конца маршрута, ответом «Наблюдателя» является прорисовка маршрута и отображение его на карте.

Для реализации модулей использовалось следующее программное обеспечение:

- Visual Studio Code (VS Code) – редактор исходного кода, разработанный компанией Microsoft для Windows, Linux и macOS. Включает в себя отладчик, инструменты для работы с Git, подсветку синтаксиса, IntelliSense и средства для рефакторинга;

- SQLiteStudio – свободный мультиплатформенный русскоязычный менеджер для систем управления базами данных SQLite с открытым исходным кодом.

Также использовались следующие технологии:

- Kivy – мультиплатформенный фреймворк с открытым исходным кодом для разработки приложений на языке Python. Используется для создания сложных многофункциональных мобильных и десктопных приложений с технологией мультитач, которая позволяет сенсорной панели или сенсорному экрану распознавать более одной точек соприкосновения с поверхностью или с естественным пользовательским интерфейсом, который не активен до взаимодействия с пользователем;

- JSON (JavaScript Object Notation) – стандартный текстовый формат для хранения и передачи структурированных данных. Он основан на синтаксисе объекта в JavaScript, но не привязан к нему.

Заключение

Разработана система, предоставляющая услугу оповещения авиапассажиров, опаздывающих на рейс, результатом внедрения которой должно быть повышение качества обслуживания и комфорта пассажиров, а также решение ряда посадочных проблем, что должно повысить лояльность авиапассажиров.

Основные функции системы:

- 1) фиксация местоположения авиапассажира, определение резерва времени до посадки и вероятности опоздания;

- 2) отправление оповещения в мобильное приложение пассажира о необходимости следовать на посадку и предоставление ему карты кратчайшего маршрута;

- 3) построение кратчайших маршрутов на карте для более эффективного передвижения по зданию аэропорта.

Преимуществом системы является используемая при проектировании системы плагиновая архитектура [3], которая дает возможность изменения старых модулей и подключения новых модулей, не меняя существенно остальные модули системы.

Наравне с этим преимуществом системы является ее автоматизированная работа без задействования больших людских ресурсов и наличие мобильного приложения для удобства пользователей.

Целесообразно интегрировать в приложение модуль биометрического контроля выхода на посадку: возможность выхода на посадку с электронным посадочным талоном с дескриптором лица. На пассажира будет затрачиваться не более 1–2 с.

Список литературы

1. Большие данные в большой гражданской авиации: обзор мировых и российских трендов. URL: <https://integral-russia.ru/2020/12/09/bolshie-dannye-v-bolshoj-grazhdanskoj-aviatsii-obzor-mirovyh-i-rossijskih-dostizhenij/> (дата обращения: 19.02.2024).
2. Шалин К. А., Черновская В. В. Разработка принципиальной схемы и конструирование высоконадежного сервера // Труды Международного симпозиума Надежность и качество. 2021. Т. 1. С. 264–266.

3. Надейкина Л. А., Черкасова Н. И. Использование архитектурных паттернов и функциональной декомпозиции для повышения качества и надежности программного обеспечения // Труды Международного симпозиума Надежность и качество. 2016. Т. 1. С. 148–151.
4. Витензон А. С. Зависимость биомеханических параметров от скорости ходьбы // Протезирование и протезостроение. М. : ЦНИИПП, 1974. С. 53–65
5. Основы Kivy. URL: <https://kivy.org/doc/stable/guide/basic.html/> (дата обращения: 12.03.2024)

References

1. *Bol'shie dannye v bol'shoy grazhdanskoj aviatsii: obzor mirovykh i rossijskikh trendov = Big Data in Big Civil aviation: an overview of global and Russian trends.* (In Russ.). Available at: <https://integral-russia.ru/2020/12/09/bolshie-dannye-v-bolshoj-grazhdanskoj-aviatsii-obzor-mirovykh-i-rossijskikh-dostizhenij/> (accessed 19.02.2024).
2. Shalin K.A., Chernoverskaya V.V. Development of a schematic diagram and design of a high-performance server. *Trudy Mezhdunarodnogo simpoziuma Nadezhnost' i kachestvo = Proceedings of the International Symposium Reliability and Quality.* 2021;1:264–266. (In Russ.)
3. Nadeykina L.A., Cherkasova N.I. The use of architectural patterns and functional decomposition to improve the quality and reliability of software. *Trudy Mezhdunarodnogo simpoziuma Nadezhnost' i kachestvo = Proceedings of the International Symposium Reliability and Quality.* 2016;1:148–151. (In Russ.)
4. Vitenzon A.S. Dependence of biomechanical parameters on walking speed. *Protezirovaniye i protezostroeniye = Prosthetics and prosthetics.* Moscow: TsNIIPP, 1974:53–65. (In Russ.)
5. *Osnovy Kivy = The basics of Kiwi.* (In Russ.). Available at: <https://kivy.org/doc/stable/guide/basic.html/> (accessed 12.03.2024)

Информация об авторах / Information about the authors

Данил Вячеславович Егзов

студент,
Московский государственный технический
университет гражданской авиации
(Россия, г. Москва, Кронштадтский б-р, 20)
E-mail: mr.danil1509@mail.ru

Людмила Анатольевна Надейкина

кандидат физико-математических наук, доцент,
доцент кафедры вычислительных машин,
комплексов, систем и сетей,
Московский государственный технический
университет гражданской авиации
(Россия, г. Москва, Кронштадтский б-р, 20)
E-mail: ludmila.nadejkina@yandex.ru

Павел Олегович Терешко

студент,
Московский государственный технический
университет гражданской авиации
(Россия, г. Москва, Кронштадтский б-р, 20)
E-mail: pavel120304@yandex.ru

Danil V. Egzhov

Student,
Moscow State Technical
University of Civil Aviation
(20 Kronshtadtskiy boulevard, Moscow, Russia)

Ludmila A. Nadeikina

Candidate of physical and mathematical sciences,
associate professor, associate professor
of the sub-department of computers,
complexes, systems and networks,
Moscow State Technical
University of Civil Aviation
(20 Kronshtadtskiy boulevard, Moscow, Russia)

Pavel O. Tereshko

Student,
Moscow State Technical
University of Civil Aviation
(20 Kronshtadtskiy boulevard, Moscow, Russia)

Авторы заявляют об отсутствии конфликта интересов /

The authors declare no conflicts of interests.

Поступила в редакцию / Received 17.06.2024

Поступила после рецензирования / Revised 18.07.2024

Принята к публикации / Accepted 30.08.2024

БЕЗОПАСНОСТЬ В ЧРЕЗВЫЧАЙНЫХ СИТУАЦИЯХ

SAFETY IN EMERGENCY SITUATIONS

УДК 004.05, 006.015.8, 519.718, 519.876.2
doi: 10.21685/2307-4205-2024-3-15

AN INDEX-BASED METHOD FOR INTEGRAL ESTIMATION OF REGIONAL CRITICAL INFRASTRUCTURE RESILIENCE USING FUZZY CALCULATIONS (PART 2. RESILIENCE CAPACITY MODELS AND BACKBONE CAPABILITIES)

A.V. Masloboev

Putilov Institute for Informatics and Mathematical Modeling of the Federal Research Centre
"Kola Science Centre of the Russian Academy of Sciences", Apatity, Russia
Institute of North Industrial Ecology Problems of the Federal Research Centre "Kola Science Centre
of the Russian Academy of Sciences", Apatity, Russia
masloboev@iimm.ru

Abstract. *Background.* The study is aimed at developing well-known and designing novel models and methods for decision support in the field of security and resilient operation management of critical infrastructures and socio-economic facilities in the Arctic region of Russian Federation. This urgent problem is especially relevant at the regional level in terms of the need to protectability heightening of critical facilities/infrastructures, cascading effects restricting of the multiple threats of various nature on higher-level systems and favorable conditions providing for mitigation of the negative consequences of influencing factors on the performance of the elements of these systems. *Materials and methods.* For easy understanding, the work structurally is decomposed in two parts. In the first part, a formal problem statement is given. The substantiation of mathematical apparatus for problem-solving is carried out. The generic framework of the developed method for assessment and analysis of the regional critical infrastructures resilience based on a fuzzy-set approach and expert judgements is proposed. In the second part, the efficiency Q-function computational models of the organizational and technical systems resilience, such as anticipation ability, responsiveness, recoverability and adaptability, which are the central elements of the optimization model of critical infrastructures resilience integral index, are examined. *Results and conclusions.* An index-based method for the integral estimation and analysis of the regional critical infrastructures resilience, based on fuzzy calculations of the level and ratio of aggregated reliability, security and robustness indices, has been developed. The method allows on basis of incomplete data to quantify systemic risks affecting the critical infrastructure resilience, performances, savings and possible losses under sampling and implementing the anti-crisis measures at all stages of the resilience management life-cycle. A distinctive feature of the method is its universality, i.e., applicability to all types of critical infrastructures. The method can be practically used by operators and analysts of regional situational centers to train and generate design decisions for counteracting the actual threats and local failures in the operation of regional critical infrastructures under uncertainty.

Keywords: system analysis, resilience, security, critical infrastructure, integral performance index, expert judgement, fuzzy calculations

Financing: the work was carried out within the framework of the State Research Program of the Putilov Institute for Informatics and Mathematical Modeling KSC RAS (project No. FMEZ-2022-0023).

For citation: Masloboev A.V. An index-based method for integral estimation of regional critical infrastructure resilience using fuzzy calculations (Part 2. Resilience capacity models and backbone capabilities). *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems.* 2024;(3):130–156. (In Russ.). doi: 10.21685/2307-4205-2024-3-15

ИНДИКАТОРНЫЙ МЕТОД ИНТЕГРАЛЬНОЙ ОЦЕНКИ ЖИЗНЕСПОСОБНОСТИ РЕГИОНАЛЬНЫХ КРИТИЧЕСКИХ ИНФРАСТРУКТУР НА ОСНОВЕ НЕЧЕТКИХ ВЫЧИСЛЕНИЙ (ЧАСТЬ 2. МОДЕЛИ ПОКАЗАТЕЛЕЙ КАЧЕСТВА ЖИЗНЕСПОСОБНОСТИ)

А. В. Маслобоев

Институт информатики и математического моделирования имени В. А. Путилова
Кольского научного центра Российской академии наук, Апатиты, Россия
Институт проблем промышленной экологии Севера
Кольского научного центра Российской академии наук, Апатиты, Россия
masloboev@imm.ru

Аннотация. *Актуальность и цели.* Исследование направлено на развитие известных и разработку новых моделей и методов поддержки принятия решений в области управления безопасностью и устойчивым функционированием критических инфраструктур и социально-экономических объектов Арктической зоны Российской Федерации. Эта задача особенно актуальна на региональном уровне с точки зрения необходимости повышения защищенности критически важных объектов/инфраструктур, сдерживания каскадных эффектов воздействия множественных угроз различной природы на системы более высокого уровня и обеспечения благоприятных условий для смягчения негативных последствий влияющих факторов на состояние работоспособности элементов этих систем. *Материалы и методы.* Работа состоит из двух частей. В первой части дана формальная постановка задачи, приводится обоснование математического аппарата для ее решения и представлена общая структура разработанного метода оценки и анализа жизнеспособности региональных критических инфраструктур на основе нечетко-множественного подхода и экспертных оценок. Во второй части исследуются вычислительные модели целевых функций качества устойчивости организационных и технических систем таких, как упреждаемость, реактивность, восстанавливаемость и адаптируемость, являющихся центральными компонентами оптимизационной модели интегрального показателя жизнеспособности критических инфраструктур. *Результаты и выводы.* Разработан индикаторный метод интегральной оценки и анализа жизнеспособности региональных критических инфраструктур, основанный на нечетких вычислениях уровня и соотношения агрегированных показателей надежности, безопасности и устойчивости. Метод позволяет на основе неполных данных количественно оценить системные риски, влияющие на жизнеспособность критической инфраструктуры, полезные эффекты и возможные потери при выборе и реализации антикризисных мер на всех стадиях жизненного цикла управления устойчивостью. Отличительной особенностью метода является его универсальность, т.е. применимость ко всем типам критических инфраструктур. Метод может быть использован операторами и аналитиками региональных ситуационных центров для подготовки проектных решений по противодействию актуальным угрозам и локальным сбоям в работе критических инфраструктур региона в условиях неопределенности.

Ключевые слова: системный анализ, жизнеспособность, безопасность, критическая инфраструктура, интегральный показатель, экспертная оценка, нечеткие вычисления

Финансирование: работа выполнена в рамках государственного задания ИИММ КНЦ РАН (НИР № FMEZ-2022-0023).

Для цитирования: Маслобоев А. В. Индикаторный метод интегральной оценки жизнеспособности региональных критических инфраструктур на основе нечетких вычислений (Часть 2. Модели показателей качества жизнеспособности) // Надежность и качество сложных систем. 2024. № 3. С. 130–156. doi: 10.21685/2307-4205-2024-3-15

Introduction

Nowadays, risk reduction, security ensuring and the resilience improvement of the critical entities and infrastructures are still major problems in management of regional socio-economic and organizational systems. This is confirmed by a number of legislations and state protection programs adopted at the highest official level both in our country and abroad, e.g.¹, etc. In the last fifteen years, foreign security policies

¹ Федеральный закон от 26.07.2017 г. № 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации». URL: <http://static.kremlin.ru/media/acts/files/0001201707260023.pdf> ; Приказ ФСТЭК России от 06.12.2017 г. № 227 «Об утверждении Порядка ведения реестра значимых объектов критической информационной инфраструктуры Российской Федерации». URL: <http://publication.pravo.gov.ru/document/0001201802090010> ; Директива Совета Европейского Союза 2008/114/EC от 8 декабря 2008 г. «О Европейских критических инфраструктурах и мерах по их защите». URL: <https://base.garant.ru/70333008/> ; Директива Европейского Парламента и Совета Европейского Союза 2022/2557 от 14 декабря 2022 г. «Об устойчивости критически важных организаций». URL: <https://base.garant.ru/407633886/>

have shown an onrush shift from the protection of critical infrastructures towards the resilience of critical entities. In Russia such a conversion is more regular and step-by-step in nature, and the focus shifting in the field of safety practices is still in progress, but deems very challenging. Global rethinking protection of critical infrastructures in the context of system resilience at the technological and political levels suspects concentrating the activities more on maintaining the essential functions which the critical infrastructures provide by adding improved absorptive, restorative and adaptive capacities or other control features, along with preventing and reducing threat, vulnerability and impact of numerous hazards by traditional management facilities. Thus, the resilience concept is a refocus from protection (security) to adaptation and recovery of the critical infrastructure systems. Reputable experts define the resilience concept as an extension of modern safety studies, namely the risk analysis and assessment, and position it as a new era of risk management, even though this concept contested and ambiguous in some cases is. Consequently, critical infrastructure resilience is a recent trend of the safety sciences conditioned by the current worldwide geopolitical situation, and its popularity has increasingly exploded in both academic and policy discourses.

From the system of systems approach perspective the critical infrastructures is commonly understood as distributed, multi-level, highly dynamic complex systems that are comprised of the interdependent sub-systems, which themselves may be large-scale, compounding and multifaceted, and operate in an emergent or synergistic manner. This means that considered class of systems have unique properties, such as large number of interacting components, emergent properties difficult to anticipate from the knowledge of single components, adaptability to absorb random disruptions, and highly vulnerability to widespread failure under adverse conditions. Accounting of these capacities is important when examining overall resilience of critical infrastructures. In accordance with¹, a critical infrastructure is defined as an "asset, system or part thereof located Member States which is essential for the maintenance of vital societal functions, health, safety, security, economic or social well-being of people, and the disruption or destruction of which would have a significant impact in a Member State as a result of the failure to maintain those functions". Considering the complexity and interconnectivity, it is obvious that critical infrastructures are highly vulnerable systems to and may be threatened by multiple hazards and disruptions of various natures.

Resilience is characterized as an immanent and relevant, but abstract system property of its self-preservation, because of the exponential growth (in number and dependence) of the internal and external threats and hazards that directly or indirectly may affect critical infrastructure performance. In turn, the loss of essential functionality of critical infrastructures due to adverse events may hurt the well-being of the society in toto. More formally, resilience is defined as the ability of a critical infrastructure system, facility or asset to anticipate/prevent, absorb/withstand, respond to, recover from and adapt to a potentially disruptive event, shock, threat or a changing environment within acceptable losses of functionality, cost and time, which should be as minimum as possible. In other words, resilience is the ability to reduce the magnitude and/or duration of disruptive events, and to cope with future risks. Broadly speaking, a resilient critical infrastructure refers to a system capable to subsist under anticipated and unpredictable events by efficiently planning, reducing vulnerability, absorbing and minimizing the consequences of multiple threats, quickly recovering and adapting all its elementary functions and structures. However, resilience concept is used in different contexts and composed of several dimensions that are related to a specific resilience management strategy each, which addresses to areas of control and actions that can be implemented in order to increase the various aspects of critical infrastructure resilience.

Despite the resilience concept has become well studied and better understood among system engineers, risk managers and owners/operators of critical infrastructures, there is still a lack of consensus regarding its formal unambiguous definition, as well as consistency and accuracy in its measurement by "one number". The absence of a common framework and standardized metrics for measuring the critical infrastructure resilience undermines the effectiveness of decision-making in the field of resilience management and situational control in the face of potential threats and uncertainties caused and triggered by disruptive events or dynamically changing environment. For the purpose of providing adequate and efficient situational management, the critical infrastructure resilience should be assessed all-round before, during and after the occurrence of disruptive events. Implementation of the proper preventive actions and protective measures on the basis of these assessments will improve system resilience, lead to useful effects and savings, as well as optimize system performance and functionality level. Thus, considering these issues, a holistic, transparent and easy-to-use methodology for comprehensive assessment and analysis of critical infrastructures resilience – from withstanding specific threats and mitigating negative impacts to eliminating post-event consequences and returning to normal operation conditions, as well as to support decision-making for risk management, is imperatively needed.

¹ Директива Совета Европейского Союза 2008/114/ЕС от 8 декабря 2008 г. «О Европейских критических инфраструктурах и мерах по их защите». URL: <https://base.garant.ru/70333008/>

Therefore, the aim of this study is to develop computable methods for integral estimation of the critical infrastructure resilience and to perform an analysis of resilience backbone capabilities, as well as to select appropriate resilience capacity models relevant and suitable for combined use within the proposed assessment procedures. Background is based on a systematic literature survey of current methodologies for evaluation of resilience concept, which enable its operationalization to critical infrastructures, and summarizing benefits and drawbacks of the existing approaches for the assessment and control of critical infrastructure resilience. Most of the state-of-the-art frameworks and methodologies reviewed in the first part of this study [1] are based on indicators (quantitative, semi-quantitative or qualitative criteria), simulation, expert judgments and fuzzy calculations. Four resilience capacities, i.e. resistive, absorptive, restorative and adaptive, are the target objectives of these approaches and are closely related with the different stages of typical resilience cycle [2]. All these resilient system capabilities (resistivity, absorbability, recoverability, adaptability) are poorly formalizeable, quantifiable and manageable, and, thus, require detailed analysis and consideration.

This article being a logical continuation of the study [1], where a generic framework of the proposed index-based method for integral estimation of the critical infrastructure resilience based on fuzzy calculations has been developed, is organized as follows. Section 1 outlines briefly related work and the background of the study. In section 2 the backbone resilience capabilities and dynamic characteristics of critical infrastructures are systematized and analyzed. Section 3 encompasses the applicable computational models of the resilience capacities, which are the central components (Q-functions) of the general estimation model of critical infrastructure resilience integral index. Finally, conclusions are drawn and the future research directions are highlighted.

Background and Related work

The resilience concept in the context of critical infrastructures has evolved from existing disciplines in other fields and is related to the foundations of risk, reliability and security. For a system to be characterized as resilient, it is important to be able to bring the system back to its original state or an adjusted state, as well as to provide a minimum service level while undergoing changes or facing disturbances [3]. According to [3], resilience is defined as the overarching goal of a system to continue to function to the fullest possible extent in the face of stress to achieve its purpose, where resilience is a function of both the vulnerability of the system and its adaptive capacity. Disruptive events and crises that start in one critical infrastructure can spread through a network of critical infrastructures, affecting them also and other sectors of socio-economic systems. According to [3], two resilience types are distinguished: internal resilience (the resilience level of the critical infrastructure, where the triggering event occurs) and external resilience (the resilience level of the rest of the external involved critical entities).

The majority of the available approaches for studying resilience are only resilience analysis methodologies. The subsequent stage of resilience evaluation is often missing, and where it is present then it is only in the form of a comparison of the resilience of the organization, asset, or system in question with other comparable objects. Thus, the evaluation is reduced to a simple comparison with ones peers. The implementation of resilience concepts to critical infrastructure on this basis seems to be rather arbitrary and this points towards the need for a framework for assessing resilience which includes some sort of evaluation process based on the needs and requirements of stakeholders of the critical infrastructure, including dependent entities, governments and the society which the critical infrastructures serve. The elaboration of this framework is one of the objectives of the current study, however, the intention is that it will be able to incorporate the results from all of the analysis methodologies reviewed.

Resilience assessment is a process for knowing its value or level by applying appropriate steps [4]. To evaluate the resilience of critical infrastructures, different metrics and definitions are discussed in up-to-date academic literature. Thereto, the commonly used approaches are qualitative, quantitative, hypothetical and empirical methods based on diverse data. However, these methods are limited to the availability of information, subjectivity of the responses provided, to a specific critical infrastructure or scenario and lack in generalization [5]. While conducting academic literature review, it is found that there are several models and tools exist for evaluating and measuring resilience. However, there are a rather limited number of freely available frameworks, and only limited information about them is publicly available. Moreover, they tend to cover specific domains/dimensions of resilience, and are sectorally limited to a specific type/class of critical infrastructure or territorially limited to a region/country.

Findings reported in [2, 6] give the following definitions for various stages in a resilience assessment framework, which are based on the similar definitions for risk assessment¹:

¹ ISO 31000:2018 Risk management – Guidelines. 2nd Edition. Switzerland, International Organization for Standardization, 2018. 24 p.

- *Resilience analysis* is the process to comprehend and to determine the level of resilience based on selected resilience indicators.
- *Resilience evaluation* is the process of comparing the results of resilience analysis with criteria or objectives to determine whether resilience level is acceptable and identify areas for improvement.
- *Resilience assessment* is the overall process of resilience analysis and evaluation.

There are many proposed methods for resilience assessment and analysis, some of which are directly targeted to critical infrastructures and few others long-listed in [1], which may apply to critical infrastructures. These estimation methods differ considerably in their background, focus and application. While a few of them are already in operational use, others exist only as theoretical and methodological models. The output of all of the methods is also expressed differently and the question remains what should be done with the calculated resilience of critical infrastructure. The following frameworks for evaluating resilience of critical infrastructures have been considered in [1]: Critical Infrastructure Resilience Indicator (CIRI), Resilience Management Index (RMI), Benchmark Resilience Tool (BRT), Guidelines for Critical Infrastructures Resilience Evaluation ("Guidelines"), Organisational Resilience Health Check (ORHC), Resilience Analysis Grid (RAG) and the "Swiss approach". A generic conceptual framework for analysis and assessment of critical infrastructure resilience is schematically represented in Fig. 1.

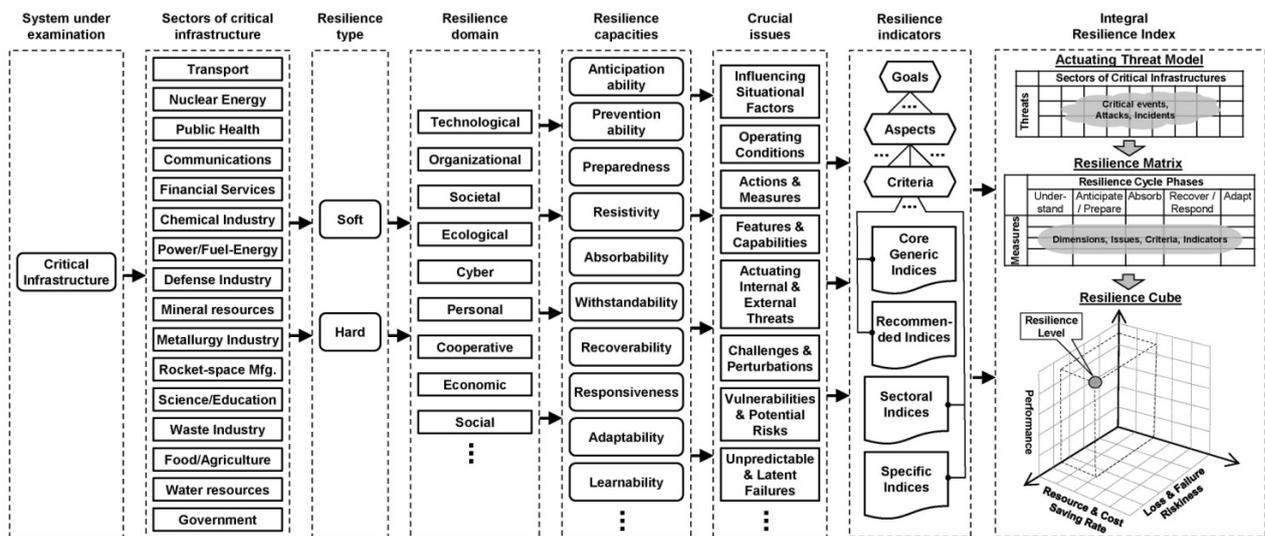


Fig. 1. The conceptual framework for analysis and assessment of critical infrastructure resilience

The broadly used and named above methods in science and practice of resilience assessment of critical infrastructures are mostly based on indicators [4]. Indicator, being a less abstract concept than resilience, can be used to show positive or negative changes in resilience. Therefore, the index-based resilience assessment could help stakeholders to analyze critical infrastructures on a practical and situational basis and to make efficient decisions. The identification of indicators is considered key before assessing resilience. These indicators exist already as safety or risk indicators, and are mainly taken from official statistics, reports and standards, current guidelines and practices, etc. They are based on historical and on-line data or expert judgements that are produced under strict quality assurance. Values of the indicators from any of the above sources can be numerical, fuzzy or non-numerical and in a general case need to be transferred into the single crisp score on a common relative or interval scale when applying resilience assessment procedure (selecting, measuring, weighting and aggregating the indicators). Therefore, the main challenge of resilience assessment is to transform expert knowledge and data into actionable measures by the means of indicators [4].

Resilience indicators are related to measurable variables that can be used, either alone or in combination, as a formal representation of resilience. Qualitative, semi-quantitative or quantitative indicators are analyzed and, when sufficient, aggregated to a measure of resilience. The resilience indicators should be clearly defined, in order to ensure objectivity and a proper balance between generality and specificity. To monitor resilience over time or comparing to similar critical infrastructures, the indicators must also provide reproducibility and repeatability. Measurement scales for the indicators and their possible weight factors should ideally be benchmarked at a sectoral level. Based on literature and defined requirements from critical infrastructure operators associated with regional situational centers, the resilience indicators to be included in the overall resilience assessment need concerted selection and optimization actions, because they

relate to the different resilience domains and issues. Indicators and criteria are an important part of various analysis methodologies used for resilience assessment.

Obviously, the more indicators are chosen to measuring resilience, the better the coverage of an issue (anything important in order to be resilient against severe threats) is going to be, but it is also obvious that the larger the number of indicators, the more complex their handling is going to be [7]. The way out has two possible directions suggested in [7]:

- finding the right number of indicators relevant to the resilience problem tackled (in practice, the more critical the situation, the smaller the number of indicators recognized and managed by operators, i.e. in absolute emergency situations operators can hardly look at more than 5–7 indicators, and in preplanned situations – 120–150 indicators are usually a maximum);

- allowing to drill-down in cases when one or more indicators need further explanation.

Resilience assessment has become convenient and common tool for resilience management, as assessment results provide useful information to critical infrastructure managers for reasoned decision-making. However, resilience assessment of critical infrastructures is facing challenges of being practical to use on the operational level of risk management [8], where there is often no or minimum time to respond to the disruptions, impacts and perturbations. Most existing resilience assessment methodologies make both general and specific criteria generalization quite complicated. Although these methodologies are diverse and multidisciplinary, they have some several common limitations. Besides, these methodologies are not comprehensive enough.

As substantiated in [8], the current lack of thinking about spatial and temporal interactions across the network of critical infrastructures prevents designing beneficial actions and suppressing dangerous ones. A critical event often causes cascading effects while optimization measures could lead to side effects. In addition, the vagueness existing recently in critical infrastructure resilience definition makes it difficult to develop generalizable indicators or criteria for resilience assessment. At once, each critical infrastructure adverse event has uniqueness, but only few existing criteria are specific enough to fully correspond to concrete situations aimed by different critical infrastructure stakeholders. It results that most resilience assessments for critical infrastructures cannot make the resilience concept usefulness at the operational level of risk and emergency management.

Some review studies on critical infrastructure resilience assessment [2, 4–6, 8–13] assign the different criteria, dimensions and aspects of resilience that existing estimation methods are currently focused on. However, most of the state-of-the-art studies for resilience analysis of critical infrastructures do not discuss assessment criteria, even though they are focused on dimensions or perspectives, such as capacities, capabilities and characteristics, could be further developed and translated to criteria. Therein, as declared in [8], without assessment criteria critical infrastructure operators have practically no envisaged positive outcomes of estimation results. During assessment processes, a target criterion is the desired direction of selected objective information, i.e. an index that is used to monitor the evolution of a specific aspect of the issue dealt with. Estimates consisting of criteria and indices provide a commonly agreed framework for articulating and defining targets and expectations, developing management methodologies, best practices and performance elements, and are then used in monitoring and evaluating attainability of those expectations and targets [8].

Generally, the weighted aggregation process for resilience criteria and indicators within the adjusted assessment method [1] rely on a conceptual hierarchical structure shown in Fig. 2, which is traditionally used for analyzing and modeling of complex systems.

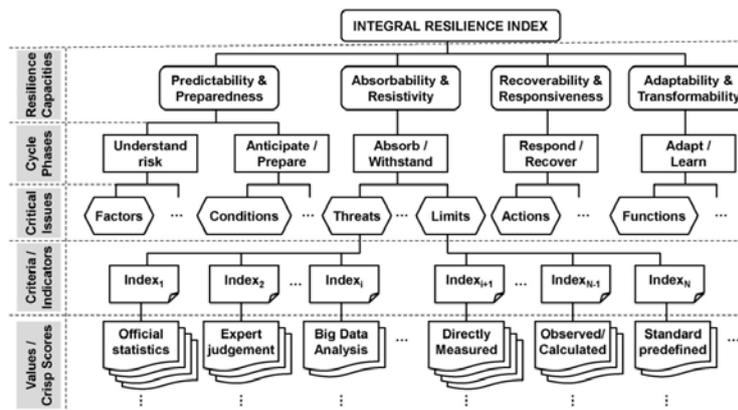


Fig. 2. The general structure of the typical index-based methodology for measuring the overall resilience of critical infrastructures based on the bottom-up weighted average aggregation

Developing generalizable criteria for resilience assessment is a current challenge to turn resilience into operational tools, i.e. resilience operationalization, because the existing formulations and theoretical models of resilience are multitude and different, but, nevertheless, are very valuable. Several studies, like [8, 14], insist that for resilience theory to become practical, it is necessary to consider not only the cost-effectiveness and negative effects of the critical infrastructure operation, but also the uniqueness of each situation. According to researches [8, 14], the operationalization of resilience concept to critical infrastructures refers to making a theory have practical and operational significance, transforming a theory into an object of practical value, regarding in the broader sense of using a theory for different purposes. Therefore, the proposed method [1] allows a wide margin of autonomy for managers and policymakers, who have the responsibility for maintaining critical infrastructure resilience and need support and guidance to operationalize the resilience-maintaining process. The adjusted method [1] is based on the multi-criteria evaluation/optimization framework similar to [7, 8, 14] and provides a regular step-by-step multidimensional aggregated assessment of positive and negative aspects, including influencing situational factors, which can better help critical infrastructure operators to make ad-hoc decisions that are better informed and profitable. It is worth noting that the usefulness and effectiveness of multi-criteria assessment approach to safety and resilience management problem-solving, as well as for the other multidisciplinary applications and issues [3, 15] have been already proofed by reputable researchers all over the world.

Thus, it is necessary to design a more complete methodology to cover the various aspects relevant to critical infrastructure resilience for the practical issues of its in-depth understanding and management. While resilience maintenance of critical infrastructures is very time and resource consuming, regular assessment and gap analysis of the functionality level of critical infrastructures exposed to disruptive events is a best practice of reacting to urgent problems as they arise, as well to planning and implementing protective measures for the future risks, and at the expense of this provide critical infrastructure system performance improvement or adaptation.

Critical infrastructure resilience backbone capabilities

In the first part of this study [1], a systematic view on resilience backbone capabilities of critical infrastructures as its target indicators used at different levels of the index-based hierarchical estimation model of the overall system resilience, has been proposed. Now, let's focus closely at the physical meaning of these key elements of the multi-level metrics system for aggregated assessment of the critical infrastructures resilience. Based on the detailed analysis of state-of-the-art literature surveys of the resilience measurement methods and frameworks [2, 3, 6, 9–11, 15, 16], the following main resilience capabilities inherent both to soft (socio-economic systems) or hard (engineering systems) resilience types and the most of resilience domains (technological, organizational, ecological, cyber, etc.) can be conditionally distinguished:

– **Reliability** is "the ability of the system to maintain its required capacity and performance during a given period of time (or mission time) under stated operating conditions" [17]. In other words, for critical infrastructures this means the capability to implement the needed performance under certain conditions and over some time without loss of performance. When the critical infrastructure is in a normal state (before a disruptive event), reliability provides its essential function. The aim of *absorptive, adaptive, and restorative capabilities* is to enhance the critical infrastructure reliability degradation due to disruptive events. *Reliability* focuses on avoiding disruptions, while resilience also counts the critical infrastructure recovery. Therefore, *reliability* and *recoverability* are complement and greatly related to the critical infrastructures resilience.

– **Maintainability** is "the ability of an item under given conditions of use, to be retained in, or restored to, a state in which it can perform a required function, when maintenance is performed under given conditions and using stated procedures and resources" [18]. *Maintainability* is a measure of how easily the critical infrastructures are repaired to a specified condition. In practice, *recovery speed* or *recovery time* is mostly used to quantify critical infrastructures maintainability. Therefore, if the time required to recover the critical infrastructure is short, it indicates proper critical infrastructure maintainability. The aim of *absorptive, adaptive, and restorative capabilities* is to increase the ease of critical infrastructure recovery by reducing the critical infrastructure damages caused by disruption or adverse events.

– **Supportability** is the critical infrastructure "ability to be supported to sustain the required availability with a defined operational profile and given logistic and maintenance resources"¹. This capability refers to the intrinsic features of the critical infrastructures that facilitate efficient and effective support of the

¹ ГОСТ IEC 60050-191 International Electrotechnical Vocabulary (IEV). 2017. 149 p.

critical infrastructures throughout its life cycle [19]. *Resourcefulness* and *mean time to support (service)* are often used as a measure of system supportability. *Supportability* is heavily influenced by logistics considerations, such as *spare parts*, *personnel availability*, strategic resources, test equipment and tools [20]. *Supportability* can be characterized as *planned* (preventive) or *unplanned* (corrective) *maintenance* activities. At once, according to study [21], the system ability to support the mission objectives includes passive and active supportabilities. Passive supportability refers to the *resource provision* (e.g., *spare parts*) at the system design phase. On the other hand, active supportability refers to the *resource allocation* at the system operational phase (e.g., *spare parts transportation speed*). Thus, passive and active supportabilities affect the critical infrastructure supportability in toto. Supportability is a characteristic that influences the availability.

– **Availability** is the critical infrastructure "ability to be in the state to perform a required function under given conditions at a given instant of time or over a given time interval, assuming that the required external resources are provided"¹. This critical infrastructure ability depends on the combined aspects of reliability performance, maintainability performance and maintenance support performance. Operational availability of critical infrastructures is formally defined as the critical infrastructure uptime ratio to the total critical infrastructure uptime and downtime. Thus, the critical infrastructure availability refers to the portion of time that the critical infrastructure can be used. The operational availability is generally used as a performance measure for a given system.

Reliability, maintainability, supportability and availability refer to the technical aspect of critical infrastructures resilience. In addition, managers need another measure to evaluate resilience from the organizational aspect.

– **Organizational resilience** considers the resilience of the critical infrastructure owner. It plays an important role in the critical infrastructure resilience. Applying this measure helps organizations to be able to deal effectively with hazards, especially when the situation is very uncertain and unstable [22]. Organizational resilience includes all actors involved in resilience management of critical infrastructures, such as resilience analysts, experts, personnel, managers and operators of situational centers. The general purpose of organizational resilience is to enhance organizational management performance in the face of irregular conditions and to provide an efficient problem-solving mentality at the organizational level of resilience control hierarchy. In [22] the organizational resilience is estimated using internal processes of an organization, including risk management, innovation, learning and design processes, which provide the proper conditions for critical infrastructures to adapt to disruptions.

– **Prevention ability (predictability)** refers to the early warning, anticipation and detection ability of disruptions and adverse events in the critical infrastructures and directly affects the critical infrastructures *recoverability*. In [14], the Prognostic and Health Management (PHM) system is used as a useful tool for prediction multiple threats and pre-event early warning. The PHM system assesses the critical infrastructures current state by monitoring facilities, anticipates potential defects by analyzing the monitoring data and assists in the proper management of critical infrastructures throughout their life cycle [14]. Early warning and predictability provide timely information to implement efficient response measures against disruptive events. Therefore, it can positively affect the dedicated costs and time for the critical infrastructures recovery process. Resilience can be described as a function of reliability and restoration, where *restoration* is defined as "the ability of an engineered system to restore its capacity and performance by detecting, predicting, and mitigating or recovering from the system-wide effects of adverse events" [17]. *Restoration* or *recoverability* can be considered as the degree of reliability of the restoration, formulated as the joint probability of a system failure event, a correct diagnosis event, and a correct prognosis event, and a mitigation/recovery action success event [17]. Hence, by knowing the actual condition of the system (diagnosis), one can estimate the maintenance and support that is needed (prognosis), and thus, the *repair/recovery time* can be optimized.

– **Absorbability (absorptive capacity)** is the degree that the critical infrastructure can absorb the negative impact of the disruptive event automatically. This capability is often considered as an immanent critical infrastructure characteristic to minimize the disruptive effects of the adverse events. Absorptive capacity includes a set of proactive actions that should be implemented in the critical infrastructure preparedness phase. Robustness is commonly used to quantify the adsorptive capacity of critical infrastructures.

– **Redundancy** refers to the degree to which critical infrastructure or its elements exist that are interchangeable and can meet functional needs in the presence of adverse events, degradation or inoperability.

¹ GOCT IEC 60050-191 International Electrotechnical Vocabulary (IEV). 2017. 149 p.

Redundancy creates alternative functions for the critical infrastructure items operation under disruption and its goal is to achieve a robust critical infrastructure. *Redundancy* increases the absorptive capacity of critical infrastructures. In addition, the redundancy is also related to backup resource and asset diversity. To provide backup for the replacement of failing functionality both *internal* and *external redundancy* can be used.

Thus, discussed key system resilience capabilities influence the recoverability (restorative capacity) and responsiveness of critical infrastructures.

– **Recoverability** is the ability of a system or critical infrastructure to restore its capacity and performance promptly by recovering from the negative effects of adverse events during a period of time under certain conditions using the available resources required to perform the adequate recovery actions. *Recoverability* is formally defined as the probability that a failed critical infrastructure element or system as a whole recovers quickly to perform the required functions at given time.

– **Responsiveness** is the ability of critical infrastructure to understand and carry out its tasks in a timely manner. Responsiveness refers to the way the system reacts quickly and effectively to a wide range of disruption events within possible modes of system operation as they occur.

– **Restorative capacity** is the degree to which the critical infrastructure can effectively restore its damaged performance and is typically affected by available budget and resources. Therefore, this capacity is affected by the critical infrastructure supportability. Restorative capacity provides permanent solutions to damages caused by the disruptions. Rapidity is commonly used to quantify the restorative capacity of critical infrastructures. The cost of restorative capacity is much more than an adaptive capacity.

– **Adaptability (adaptive capacity)** is the self-organization degree to the new conditions and changes, to which the critical infrastructure can arrange itself and use temporary and often non-standard actions to prevent critical infrastructure downtime during and after the disruption events. This capacity can prevent sudden collapses in the critical infrastructure performance level, but these actions have a temporary nature and for the critical infrastructure performance recovery permanent actions should be taken as soon as possible.

– **Learnability (learning capacity)** is the degree to which the critical infrastructure can learn from the occurred disruptions to prevent similar future events. The obtained experience and knowledge from past events can be incorporated for future iterations.

A graphical interpretation of the physical meaning of system resilience capabilities at different phases of the critical infrastructures resilience management cycle is shown in Fig. 3.

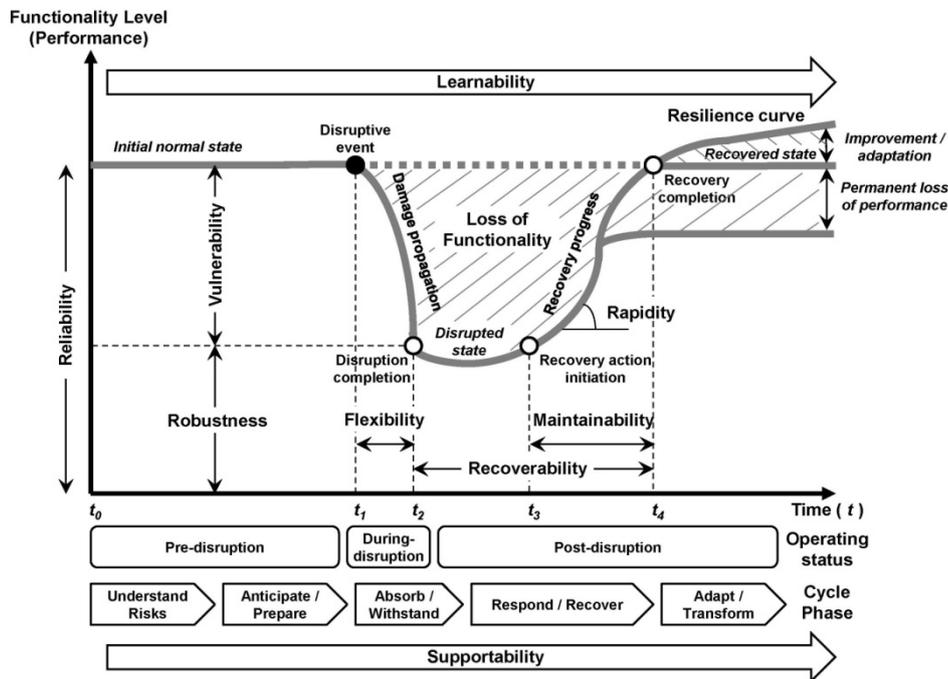


Fig. 3. The physical meaning of critical infrastructure resilience capabilities at different phases of the resilience management cycle (adapted from [9]).

Specification of the given resilience capabilities is structurally proposed in Tables 1–4. It should be noted that all the resilience capabilities are much interconnected and complementary.

Table 1

Characteristics determining the preventive capacity of critical infrastructure resilience

Resilience capacity	Characteristics / Indicators	Description & Definition of capabilities
Anticipation ability & Resistance	Preparedness degree	An extent that characterizes the available set of facilities and assets, as well as on-the-shelf strategies, response plans and actions for implementing and executing the precautionary measures relevant to the system in the face of potential disruptive events and emergency situations influencing critical infrastructure resilience. In other words, the state of critical infrastructure of being ready for the occurrence and impact of disruptive events (e.g., failure, error, critical situation, crisis, etc.). Preparedness is defined as "a continuous cycle of planning, organizing, training, equipping, exercising, evaluating, and taking corrective active in an effort to ensure effective coordination during incident response" and includes a checklist of protective measures relevant to implementation in the current situation, i.e. a set of the analytical-planning documents to increase the preparedness of critical infrastructure elements for adverse events. As a measure of preparedness can be the resource intensity of forces and means to reduce the effects of disruptions
	Reliability	This capability was described above
	Detection ability	Probability and/or time of identification of the disruptive events
	Prognostic & Health Management (PHM)	This capability was described above
	Planned Maintenance	The percentage of time that the critical infrastructure was dedicated to planned maintenance activities relative to the total operating time. This general metric provides insights into the efficiency and effectiveness of critical infrastructure maintenance program in ensuring optimal performance and reliability. It refers to as scheduled maintenance and focuses on minimizing the downtime and costs associated with breakdowns, as well as fault tolerance improvement
	Joint activity cooperation plan	Refers to a degree of concordance and relevance of the planned preventive measures and actions, decentralized local decisions made and implemented to the operational context and specificity of disruptive events occurred
	Protectability	Ability of critical infrastructure to be safe and preserve its essential functions under the impact of negative situational factors of various natures. The risk impact and risk probability factors determine the system protection level. A higher protectability score indicates a lower overall risk level and better protection in the face of multiple potential threats or vulnerabilities
	Operability	Ability of critical infrastructure that has the capacity and flexibility to achieve a range of operating conditions safely, reliably, profitably and with positive dynamic performance and quality, i.e. the <i>operability</i> ensures to keep the system, its units or service in a safe and reliable functioning condition that is pre-defined operational requirements. The <i>operability</i> is a measure of the resilience support of critical infrastructure through the ability to adapt and transform to restore system critical functions at the operational level, which can be achieved before all system repairs are made. In other words, the <i>operability</i> can be described as the fitness, capacity, or system ability to use to provide critical services allowing stakeholders and management entities to receive required, or near required, essential functions from a potentially impaired critical infrastructure following a disruptive event
	Sensitivity	A degree to which a critical infrastructure is vulnerable or susceptible to disruptions and threats that could impact its operation, performance, or ability to perform essential functions. As well, it describes the extent to which the dynamics of a critical infrastructure change in response to perturbations or parameter variations. Unlike robustness and flexibility, which are measured in terms of behavioral fitness, sensitivity refers only to the responsiveness of a system to perturbations. Sensitivity to perturbations can be either beneficial, detrimental, or neutral with respect to task performance, and sensitivity refers only to the magnitude of the response to a perturbation, rather than its effect on fitness [23]
Resistance / Resistivity	The ability of a critical infrastructure and its functional elements to prevent and withstand the occurrence of the disruptive or undesirable events	

Table 2

Characteristics determining the absorptive capacity of critical infrastructure resilience

Resilience capacity	Characteristics / Indicators	Description & Definition of capabilities
1	2	3
Absorbability & Robustness	Robustness	The strength or ability of critical infrastructure to resist a certain disruption level (withstand a given level of stress or demand) and absorb its primary effects without suffering degradation or significantly reducing performance (losses of functionality). A critical infrastructure with high robustness maintains its central function in a disruptive event. <i>Robustness</i> is measured by the critical infrastructure amount of residual performance after a disruption. Furthermore, <i>survivability</i> , <i>resistant ability</i> and <i>stability</i> of critical infrastructures have a similar formal definition to its robustness
	Fragility	A hazard specific indicator of the critical infrastructure performance loss function within the absorption and response phase of the system resilience life-cycle as discussed in [15]. <i>Fragility</i> of a critical infrastructure refers to the conditional probability of failure or a given level of damage conditioned on the response parameter (intensity measure), i.e. the probability of reaching or exceeding a given damage level as a function of the hazard intensity. The greater degree of damage is associated with higher fragility, and a lower robustness or reliability. <i>Fragility</i> is often used as a specific description of vulnerability
	Vulnerability	The degree of the critical infrastructures sensitivity to disruption. There is no consensus about the relationship between vulnerability and resilience, but it seems that higher vulnerability of the critical infrastructures leads to lower resiliency and vice versa.. Vulnerability is an inherent feature of the critical infrastructures, even before any disruptive event. Analyzing the vulnerabilities can help to identify the possible weak points of the critical infrastructure operation, which cause the most damage during disruption, and generate proper control strategies to fallback. The key parameters of vulnerability are: stress to which a critical infrastructure is exposed, its sensitivity and its adaptive capacity. <i>Exposure</i> is perceived as the degree to which a critical infrastructure is exposed to a given stressor. <i>Sensitivity</i> is the degree to which a stressor impacts the critical infrastructure. <i>Adaptive capacity</i> is perceived as the potential for the critical infrastructure to adjust or cope with impact
	Stress rate	A resilience measure used to determine the stability of a given item, system, critical infrastructure or any other entity when deliberately intense or thorough testing. Stress rate measurement involves testing beyond normal operational capacity, often to a breaking point, in order to observe the outcomes and consequences for the purpose of further training and enhancing the system performance in different operational conditions. Such an examination and fitness of the system resistance provide damage level and limitation exercise, as well as identifying ultimate stress limits under disruptions, deviations or disturbances
	Independency	Refers to the ability of a critical infrastructure to function autonomously and self-sufficiently without being overly reliant on external resources, dependencies, or vulnerabilities. An independent critical infrastructure is characterized by its capacity to operate independently and sustainably, even in the face of disruptions, failures, or external threats. This characteristic describes the level of autonomy and isolation of components within the critical infrastructure from each other, i.e. the degree of interdependence among these components. <i>Independency</i> of critical infrastructures is crucial for ensuring their reliability, security and resilience in performing essential functions. By identifying potential points of failure and bottlenecks, as well as reducing risks, vulnerabilities and dependencies, an independent critical infrastructure can enhance its ability to adapt to dynamically changing environment, mitigate threats and maintain the continuity of essential functions for society and the economy

1	2	3
	Resourcefulness	The ability of a system to direct resources to critical infrastructure support by using, mobilizing and supplying the required resources (spare parts, finances, information, laborers, technology, etc.) to identify and solve problems under adverse event of a disturbance or shock in a prioritized manner. It describes a level of system preparedness to effectively combat an adverse event. The main destination of resourcefulness is to enhance the critical infrastructure rapidity and increase the restorative capacity of critical infrastructure
	Facilitation ability	Refers to the capacity of a critical infrastructure to support, enable and enhance the efficient and effective performing of essential functions and operational services. A critical infrastructure system with strong <i>facilitation ability</i> can streamline control processes and provide coordination of decision-making, foster creativity and innovation of resilience management, and improve collaboration and communication among its components, operators or stakeholders to address complex challenges and to ensure the reliable and continued delivery of critical services. Critical infrastructures with high <i>facilitation ability</i> are more efficient in withstanding impacts of multiple potential threats and in achieving their goals and objectives as well. Facilitation plays a key role in optimizing the performance of critical infrastructures intended for resilient economic growth, public safety and national security support
	Internal redundancy	The <i>internal redundancy</i> is provided by a part of the critical infrastructure which is always online. With no redundancy the impact of a disruptive event on the performance of critical infrastructure results in a considerable drop. The presence of internal redundancy with additional capacity within the critical infrastructure results in the lesser or minor drop in its performance under unexpected circumstances. If the critical infrastructure has a sufficient <i>internal redundancy</i> , the system performance can be restored using alternative pathways. The effect can be indirectly measured in recovery time or backup cost. A redundant critical infrastructure is expected to have lesser recovery time, but the initial backup expenses may be considerable. <i>Internal redundancy</i> contributes to robustness (insensitivity to local failure) and could be described as the means to decrease the dependence of a critical infrastructure to its components
	Safe failure	The ability of a critical infrastructure to absorb shocks and the cumulative effects of slow-onset challenges in ways that avoid catastrophic failure or irretrievable losses
	Situational awareness	The ability of decision makers and the operators of critical infrastructures to perceive, comprehend, and project relevant information in a given context, as well as their capacity to maintain a clear understanding of the current situation and respond effectively to changing conditions or risks. It enables critical infrastructure managers to aware/anticipate the information on potential threats, perturbations and adverse events in agreed manner, analyze/interpret that information within a unified context, using it to make informed decisions and implement appropriate control actions (preventive, mitigating or proactive measures), i.e. respond effectively to impact of the changing environment and deviating operating characteristics of critical infrastructures. As a measure of situational awareness can be put forward the completeness of understanding current situation and anticipation of risks before and after the disruptive event occurred, or the entropy of the situational control data, or the response time which refers to the time it takes for critical infrastructure operators to recognize changes in the situation (risk identification) and make appropriate corrective control actions

Table 3

Characteristics determining the restorative capacity of critical infrastructure resilience

Resilience capacity	Characteristics/ Indicators	Description & Definition of capabilities
1	2	3
Recoverability & Responsiveness	Maintainability	This capability was described above. In terms of hard resilience (engineering systems) maintainability of critical infrastructures is usually associated with technological repairability of the system
	Supportability	This capability was described above
	Restoration index	A measure of how quickly critical infrastructure system can be restored to full functionality after a failure or disruption. This index is an important metric for evaluating the resilience and reliability of a critical infrastructures
	Downtime	Refers to the period during which a critical infrastructure or its components are not operational or available for use. Consideration of downtime is urgent for assessing the reliability and performance of critical infrastructures, as well as for identifying opportunities to improve uptime and minimize disruptions
	Rapidity	The ability of a critical infrastructure to return to normal operating capacity in a timely manner. It is also a rate at which a critical infrastructure can recover a satisfactory performance level. <i>Rapidity</i> refers to the critical infrastructures performance curve slope during the recovery process and is often known as the <i>recovery rate</i> of system or its elements functionality in a timely manner. <i>Rapidity</i> reflects also how quickly the spare parts can be accessed and applied to improve critical infrastructure resilience
	Safety margin	A measure of how much extra capacity or capability a critical infrastructure has beyond its normal operating requirements to ensure safe and reliable operation. It is often expressed as a percentage to indicate the level of safety buffer built into the system (critical infrastructure)
	External redundancy	The ability of a critical infrastructure to carry on providing a service in the case of failure enabled by external means. As is known from real practice, external resources are not immediately available to reduce the impact of an adverse event and they contribute to the quicker recovery of the critical infrastructure functionality. The <i>external redundancy</i> depends on availability of external reserves (services) and can be ensured by its sufficient number and capacity subject to supply rapidity with minimal cost and delay
	Modularity	Refers to the means to modular organization of critical infrastructures based on various combinations of operational units, each contributing to system performance and performing a specific system function. This characteristic of critical infrastructures is perceived as the system capacity for proper re-engineering to accommodate increasing failure or damage under contingency situations, as well as to provide flexible pathways and options for system performance enhancement by replacing the interacting components composed each other if one, or even more, fail, and by integrating new functional elements (services) if necessary. Generally, <i>modularity</i> refers to the degree to which a system can be divided into separate, independent modules or components that can be developed, maintained, and modified independently. <i>Modularity</i> of critical infrastructure system is crucial for ensuring its flexibility, scalability, and ease of maintenance
Segregability	Refers to the degree to which a critical infrastructure system can be segregated into separate, isolated components, parts or units that can function independently without affecting each other. Segregability is important for system security, fault tolerance and scalability	

1	2	3
	Decomposability	Is similar to segregability of a critical infrastructure and refers to the ease with which a critical infrastructure system can be broken down into smaller, more manageable components, items or units. Decomposability is useful for critical infrastructure effective design, analysis and maintenance
	Unplanned maintenance	Refers to the corrective maintenance activities that are carried out in response to unexpected system failures, malfunctions or other crisis events that disrupt normal operation of critical infrastructures, i.e. where there is a sudden failure which was unpredicted. Unplanned maintenance is typically reactive in nature and occurs outside of scheduled maintenance plans or preventive maintenance routines. It is often necessary to address emergency situations that require immediate attention to restore the items or assets of critical infrastructures to operational status and minimize downtime. <i>Mean Time Between Failures</i> and <i>Mean Time to Repair</i> are common measures of the unplanned maintenance, which represent the average time between two consecutive failures of a system or its elements and the average time it takes to repair a system or its elements after it has failed, respectively. Both metrics can be used to estimate the general costs associated with unplanned maintenance of the functionality and operability of critical infrastructures
	Functionality	A capacity or the state of a critical infrastructure operating properly to provide a regular reliable service at, or as close as possible to, what the critical infrastructure provided prior to an adverse event. The <i>functionality</i> is a measure of the critical infrastructure resilience and is not fully restored until all system repairs are completed and operational restrictions removed
	Feasibility	Refers to ability of a critical infrastructure to be successfully implemented, operated, and maintained within the constraints of available resources, technology and time. Feasibility assessments are conducted to determine whether a proposed critical infrastructure is viable and achievable, taking into account factors such as technical feasibility, economic feasibility, operational feasibility, legal and regulatory feasibility. Technical feasibility assesses whether the necessary technology and expertise are available to develop and implement the critical infrastructure. Economic feasibility evaluates whether the benefits of the critical infrastructure outweigh the costs and if the project is financially viable. Operational feasibility examines whether the critical infrastructure can be effectively integrated into existing systems, processes and operations. Legal and regulatory feasibility considers compliance with established laws, regulations and standards. A critical infrastructure that is deemed feasible is more likely to be successfully implemented and deliver the intended benefits. Feasibility analysis is essential for identifying potential challenges, risks and opportunities early in the planning phase to ensure the successful development and deployment of critical infrastructures
	Autonomy	Refers to the ability of a critical infrastructure to operate independently or with minimal human-aided. The measurement of critical infrastructure autonomy is based on different factors such as the degree of automation, decision-making capabilities, self-sufficiency and adaptive capacity of the system, etc
	Insurance rate	A characteristic used to assess the risk level associated with the critical infrastructure and determine the premium that needs to be paid to insure its operational units against potential losses of functionality or damages. The insurance rate is typically based on such factors as the value of a system, its susceptibility to risks and the desired level of coverage
	Restart ability	Refers to the capability of a critical infrastructure system and its components to recover and resume normal operation after a failure or disruption. The <i>restart ability</i> depends on such factors as recovery time, reliability of restart procedures and the effectiveness of fault detection mechanisms

Table 4

Characteristics determining the adaptive capacity of critical infrastructure resilience

Resilience capacity	Characteristics/ Indicators	Description & Definition of capabilities
1	2	3
Adaptability & Learnability	Flexibility	The ability of a critical infrastructure to perform essential tasks under a wide range of conditions, and to convert assets or modify structures to introduce new ways of achieving essential goals, as well as to react to disruptions and adjust its internal mechanisms with the help of adaptive capacity without the consideration of any prior responses
	Technological transformability	The capability of a critical infrastructure to effect transformational change. System transformability depends on the following attributes: getting beyond the state of denial (acknowledging the need for transformational change); creating options for transformational change; having the capacity for transformative change. Such a change suspects the transition to an entirely new stability system configuration defined by new state variables, or the old state variables supplemented by new ones. To a wide extent, transformability is the ability of a critical infrastructure to create a new stability state space (configuration) for all its functional units and the new system functioning pathways under the impact of multiple internal and external threats when the operating system being unstable. The changes introduced by the transformability cascade through and may transform the whole existing system with all its constituent adaptive cycles
	Technological upgradability	The ability of a critical infrastructure to restore system functionality quickly and to adjust it to increased demands by means of replacement some system components by new or similar ones, but with different (improved) characteristics. The upgraded critical infrastructure provides a higher system performance level leading to an improved resilience. <i>Upgradability</i> regards also to data acquisition on performance and expected changes in demands. The collected data can be effectively used for upgrading the system by removing or reducing any critical weaknesses that lead to higher demands on service and maintenance. The data can also be used to make prognoses on future maintenance and support needs, and to predict when to upgrade, modify or replace the critical infrastructure components and assets
	Integrability	The ability of a critical infrastructure system to integrate external heterogeneous elements inbye and provide their communication with existing items and each other on the basis of compatible technical, organizational and normative regulations, protocols and standards
	Interoperability	The ability of critical infrastructure elements to interact (data/control exchange) with external entities and with each other based on common conceptual models and context interpretation of information for the purpose of providing completeness of situational awareness and formation of the unified information field for decision-making under joint activities
	Composability	The ability of critical infrastructure elements to interact with any other elements in a recombinant manner to satisfy requirements based on the expectation of the behaviors of the interacting parties, as well as to form a steady composition for improving system resilience
	Reconfiguration ability	Refers to the capability of a critical infrastructure to adapt, modify or reorganize its structure, components or configuration in response to changing requirements, conditions or failures. This ability allows the system to maintain functionality, performance and reliability even in dynamic or uncertain environments. Reconfiguration involves adding, modifying and removing components, changing relationships between components, adjusting parameters or switching between different operating modes. Critical infrastructures with a high reconfiguration ability are usually more resilient, flexible, and efficient

1	2	3
	Personnel availability	Refers to the degree of readiness and presence of the qualified staff, operators and other essential personnel required to operate, maintain, and manage the critical infrastructure effectively. Needs of personnel skilled and trained to respond to critical events, handle disruptions, implement preventive measures, conduct repairs and risk elimination, and ensure the continuous operation of critical infrastructure assets and components. Personnel competences directly impact the ability of a critical infrastructure to respond to and recover from disruptions or unexpected events. Adequate staffing levels, proper training, clear communication protocols, effective coordination among personnel, response time, adequacy of shift schedules and coverage and skill set diversity are essential factors in ensuring the personnel availability to support the uninterrupted operation of critical infrastructures
	Spare parts availability	Refers to the accessibility and supply of essential resources (assets, materials, components, equipment, etc.) that are necessary for the operation, maintenance and recovery of critical infrastructures when needed to quickly address failures, breakdowns, or disruptions within the system. Common measures used to assess spare parts availability are resource availability rate and the mean time to repair. Additional metrics such as stock-out rates, inventory turnover, lead times for spare parts delivery, and percentage of critical spare parts in stock can also provide the estimation of spare parts availability for critical infrastructures
	Long-term/short-term reconstruction ability	Refers to the capacity of a critical infrastructure to recover and rebuild after a disruptive event or disaster, i.e. the ability of a system to bounce back from a crisis. Short-term reconstruction refers to the immediate response and recovery efforts following a disruption. It includes activities such as restoring essential services/units, repairing damaged elements of a critical infrastructure and ensuring its safety. Short-term reconstruction focuses on rapid and effective response to minimize the impact of the event and restore basic functionality to the system. Long-term reconstruction pertains to the ability of a critical infrastructure to fully recover and rebuild over an extended period of time. It involves strategic planning, policy changes, system upgrades, and more comprehensive efforts to address the underlying vulnerabilities and improve the capacity of a critical infrastructure to withstand future disruptions. Common measure used to assess the long-term/short-term reconstruction ability of a critical infrastructure is the recovery time objective and recovery point objective. Recovery time objective refers to the targeted duration within which a critical infrastructure should be restored to full operational capacity after a disruptive event. It measures the time it takes for the system to recover and resume normal operations. Recovery point objective measures the acceptable performance loss of a critical infrastructure in the event of a disruption. It defines the maximum performance that can be lost without causing significant harm to the functional units of critical infrastructure
	Self-organization ability	Refers to the capacity of a critical infrastructure to adapt, evolve and organize itself without external actions. Self-organization ability is often influenced by such factors as system complexity, diversity, coherence and feedback mechanisms. There is still no universally accepted measure for estimating self-organization ability of complex systems
	Creativity & improvisation ability	Refers to the management system capacity of a critical infrastructure to generate novel ideas and ways to solve new and existing control problems using cumulative knowledge, and utilise innovative and creative approaches to developing solutions for the purpose of adaptation of critical infrastructure components and assets to changing circumstances. The degree of creativity and improvisation of risk analysis procedures within the resilience management of critical infrastructures directly affects the effectiveness of decision-making and situational control in times of crisis and disruptive events

Critical infrastructures resilience is currently determined by capabilities represented above and their combinations that characterize different life-cycle phases of system resilience (understand risks, anticipate/prepare, absorb/withstand, respond/recover, adapt/learn) and appropriate resilience components (capacities), specifically such as anticipation and prevention ability, absorbability and responsiveness, recoverability and adaptability. In the last decade, comprehensive analysis of these resilience components has been carried out by a great number of reputable studies, but the lion's share of them was promoted abroad. In our homeland, the resilience management support of critical infrastructures is a quite new and challenging field of research, intersecting with pioneering safety, reliability and situational control fundamentals.

Generic indicators and dimensions of system resilience listed in Tables 1–4 are commonly used within the state-of-the-art estimation models and assessment methodologies for measuring the overall resilience of critical infrastructures considered in [5–7]. The choice of the specific indicators and their measure of influence on resilience under its assessment and management depends on the types of critical infrastructures, scope, context and resilience domains, as well as the subjective preferences of experts, and remains with risk-analysts or decision makers. It is worth noting that the selection of resilience metrics is typically made in relation to the class of system under study and the nature of influencing situational factors being the sources of system disturbance or shock and clearly defined.

As is declared in [24], resilience indicators can be applied in a-priori manner when assessing resilience of critical infrastructure before a disruptive event, and post-hoc manner when giving an absolute measure of the indicator that is directly benchmarked against a predetermined baseline, and estimated following some system perturbation. According to report [24], a-priori resilience indices include failure probability, critical infrastructure quality, pre-event functionality, substitutability, interdependence, extent of mitigating features; quality of planning/response under disturbance, quality of crisis communications/information sharing, security of critical infrastructure, etc. Ad-hoc resilience indices include systems failure, severity of failure, post-event functionality, post-event damage assessment, cost of reinstating functionality post-event, recovery time post-event, recovery or loss ratio, etc. The permanent increasing of complexity and uncertainty in operation of existing critical infrastructures requires regular review, updating and improvement of resilience metrics for adequate valuation and efficient management of critical infrastructure resilience. Moreover, since acting critical infrastructures are usually connected to each other and interdependencies between them exist, the quantification of critical infrastructure resilience becomes even more complex.

Next, after discussing the conceptualization of the resilience phenomena and its capabilities, let's move to the formal representation of the main resilience capacities mathematically formalized by well-known reputable resilience researchers in specific manner.

Critical infrastructure resilience capacity models

The existing index-based methods for critical infrastructure resilience assessment found and reviewed in this study are generally intended for modeling and scenario analysis of the critical infrastructure behavior during a particular disruptive event, i.e. over scenario time. Defining the critical functionality of a critical infrastructure enables to precisely and quantitatively define and construct the system resilience curve in scenario time and analyze the main characteristic points of its performance level in discrete or continuous time. The resilience curve can be used to monitor the critical infrastructure functionality level dynamics and to define resilience dynamic characteristics (capabilities), such as reliability, robustness, vulnerability, recoverability, rapidity, maintainability, supportability, etc., mentioned in the previous section. Thus, the resilience capacity models, which correspond to the resulting macro-indicators of critical infrastructure resilience under consideration, selected and used within the framework of the designed estimation method [1] are mainly based on the mathematical formulations given in [25] and resilience curve analysis notionally illustrated in Fig. 3 and the Fig. 4. As noted in [7], these resulting macro-indicators are not the same as the input resilience and functional indices defined at the lower level of the assessment hierarchy (see Fig. 2) and then bottom-up aggregated to the macro level of the overall resilience estimates. In practice, combinations of these macro-indicators are well suitable for stress-testing of critical infrastructures by comparing their values measured or computed with the allowed critical thresholds defined for the specific operating conditions.

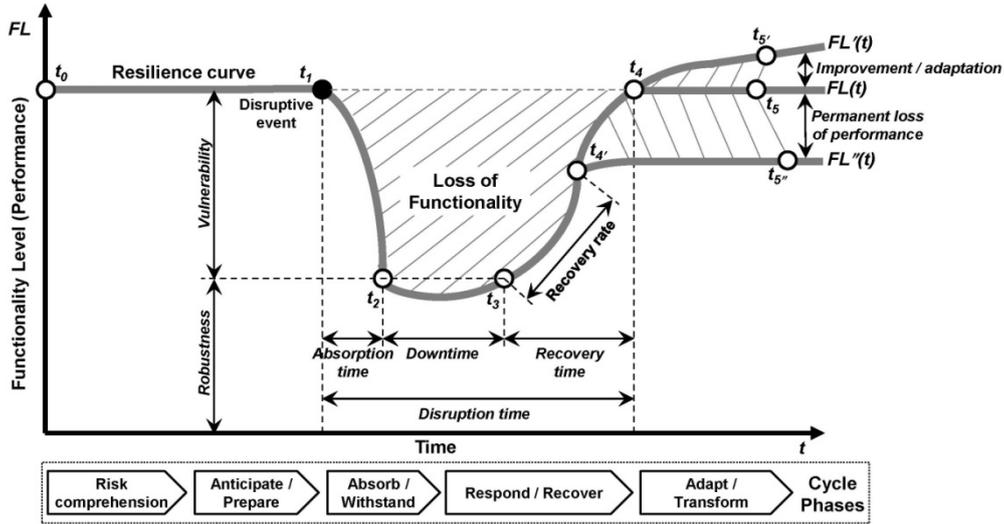


Fig. 4. Resilience curve general view: The dynamics of the resilience level of critical infrastructure over time expressed by the performance loss and recovery function (adopted from [25])

The notations used in Fig. 4 are as follows [25]: t_0 is a time before the disruptive event or a starting point of the simulating scenario; t_1 is a time at which the adverse event occurs; t_2 is a time at which the critical infrastructure reaches the minimum performance level, i.e. a starting point of its functionality loss; t_3 is a time at which the critical infrastructure starts to recover; t_4 is a time at which the critical infrastructure reaches the initial functionality level or a starting point of a new steady-state level, but with lesser performance ($t_4 = t_{4'}$); t_5 is a time at which the scenario ends or at which the critical infrastructure increases its functionality via adapting, transforming and learning ($t_5 = t_{5'}$), or, in the worst case, the system shows a permanent loss of functionality ($t_5 = t_{5''}$).

An accident or disruptive event E occurred at time t_1 within the critical infrastructure, which is initiated by some actuating threats as a root cause of its emergence, refers to an incident formally interpreted as a process of parametric variation (system change). Meanwhile, generally speaking, an incident is any parametric or structural change in a critical infrastructure system that is associated with various failures in operation of its components and accompanied by a loss of functionality and irreversible transition process of the system state from a normal operation to an emergency one. In this context, the process of changing the system operation conditions can be formalized as follows:

$$CI(\pi^{no}, \sigma^{no}) \rightarrow CI(\pi^{de}, \sigma^{de}) \rightarrow CI(\pi^{eo}, \sigma^{eo}), \quad (1)$$

where (π^{no}, σ^{no}) , (π^{de}, σ^{de}) , (π^{eo}, σ^{eo}) are the values of the parametric and structural state variables of the critical infrastructure system under different operating conditions: normal operation, failure-caused disruption, emergency operation.

A sudden failure of critical infrastructure system is understood as a rapid (stepwise) change in the values of system state variables that determine its quality (reliability, safety, resilience, etc.), which leads to a complete loss of its functionality at an arbitrary point of time. For the formalized representation of a sudden failure, the Heaviside unit function $1(t)$ [26] is used. The operation of the critical infrastructure system from the point of time $t = t_0$ until the loss of its functionality when $t = t_1 = (t_0 + T)$ is shown in Fig. 4. Then, in compliance with such a way of formal definition of system failure, the following mathematical formulation can be written:

$$x(t) = FL(t) [1(t - t_0) - 1(t - t_0 - T)], \quad (2)$$

where $FL(t)$ is a system performance function of the critical infrastructure resilience curve; $x(t) \equiv FL(t)$ is a signal actuating at the output of the system in normal operating conditions, and as a result of a failure, $x(t) = 0$; T is a critical period of time when a failure or disruptive event occurs.

In addition to sudden failures, there is also a possible case of stepless degradation of the critical infrastructure system (gradual failures), characterized by the accumulation of hazards within the system and, consequently, a slow (gradual) change in the operating characteristics of the system. Let ρ be a certain variable that expresses an internal danger to the critical infrastructure system. Let us introduce a function $J(x_i(\rho, t))$ that reflects by x_i the damage (fatigue) accumulation within the system at time t :

$$J(x_i(\rho, t)) = \int_{t_0}^t x_i(\rho, \tau) d\tau. \quad (3)$$

It is obvious, when the $t_{i+1} > t_i$, an inequality meets $J(x_i(\rho, t_{i+1})) > J(x_i(\rho, t_i))$, and if the proposition $(x_i(t_{i+1}) > x_i(t_i)) \wedge (x_i(t_{i+1}) > x^{\max})$ is true, then from the point of time t_{i+1} the system experiences stepless degradation due to the accumulation of structural changes in it (e.g., variation of constraints, interconnections or control coefficients), where x^{\max} is a maximum permissible value (upper bound) of the system state variables in the normal operating conditions.

Next, the measures for modeling the impact on critical infrastructure system are considered.

Robustness ($Rob, \%$) characterizes the absorption capacity of the critical infrastructure [25]. It is measured as the ratio of the percentage of the lowest post-disruption functionality level, i.e. at point of time t_2 , to the pre-disruption functionality level, i.e. at point of time t_0 during normal operation. The appropriate formulation can be written as follows:

$$Rob = \frac{FL_2(t)}{FL_0(t)} \cdot 100\%. \quad (4)$$

Absorption time (AT, t), measured in hours, is defined as the time during which the critical infrastructure absorbs a disruptive event while the critical infrastructure undergoes a decrease in its functionality level. It is measured as the difference between points of time t_2 and t_1 . The following formulation is given:

$$AT = t_2 - t_1. \quad (5)$$

Loss of functionality ($LoF, \% * t$) is the critical infrastructure performance lost in a given adverse situation [25]. It is measured by the area of the curve (an approximation) between the time when the critical infrastructure starts to lose its functionality (t_1) to the time when it reaches the initial state (t_4) (Fig. 4). The approximation is done for the area above the curve to a well-defined shape (e.g., a triangle) [25]. The output is measured in percentage loss of functionality over time (the time is measured in hours).

$$LoF = \int_{t_1}^{t_4} [FL_1(t) - FL(t)] dt. \quad (6)$$

The value of the functionality level $FL(t)$ of the critical infrastructure system at a particular time is calculated by aggregating the relevant indicators scores (in a particular case of $FL(t)$, the specific functionality indices) starting from t_0 and makes up $FL(t) = 100\%$.

Downtime (DT, t), measured in hours, is defined as the time duration for which the critical infrastructure is not functional. In respect to critical infrastructures, this could apply if the critical infrastructure stops functioning. In this case, the functionality level of the critical infrastructure remains below the threshold level of functionality [25]. It can be measured as the difference in time between points of time t_3 and t_2 , as illustrated in Fig. 4 and the following formulation is assumed:

$$DT = t_3 - t_2. \quad (7)$$

This calculation is conducted when the threshold level of functionality is defined (in [25], it is assumed that the threshold level is $FL_{t_2} (= FL_{t_3})$).

Recovery refers to the ability to not only return to acceptable operating levels, but also to recover fully from the effects of a disruptive event in the maximum allowable/acceptable recovery time [25]. Recovery time (RT, t) , measured in hours, is defined as the time at which the critical infrastructure recovers from the disruptive event and gains its initial or desired functionality level [25]. It can be measured as the time taken to recover the functionality level, i.e. the time between points of time t_3 and t_4 . The following formulation can be written:

$$RT = t_4 - t_3. \quad (8)$$

Since the functionality level at the end of the scenario time may be different from at the start of the scenario, the recovery time may have to be measured at a new steady-state level [25].

Recovery rate $(RR, \% / t)$, measured in percentage, is defined as the rate at which the critical infrastructure recovers from a disruptive event and gets back to its initial functionality level [25]. It characterizes the recovery trajectories of the critical infrastructure system from the point it starts recovering from the given scenario to the final recovery. Recovery rate is measured as the ratio of change in functionality level between points of time t_3 and t_4 , as shown in Fig. 4. The following formulation is given:

$$RR = \frac{FL_4(t) - FL_3(t)}{t_4 - t_3}. \quad (9)$$

Disruption time (DT, t) , measured in hours, characterizes the recover capacity of the critical infrastructure to return to the desired functionality level and is defined as the total time taken by the critical infrastructure to recover [25]. In the functionality level over time FL/t curve, it is a time between points of time t_1 and t_4 when the disruptive event occurs and the critical infrastructure has fully recovered, respectively. This situation is formally represented in Fig. 4 and formulated as:

$$DT = t_4 - t_1. \quad (10)$$

Final recovery of the functionality level of a critical infrastructure could be equal to, better than $(FL'(t))$, or worse than $(FL''(t))$ the original system performance $(FL(t))$. Hence, the model schematically illustrated in Fig. 4 allows for the calculation of the system "improvement/adaptation/transformation" capacity $(IAT, \%)$ measured in percentage [25]. This is the capacity of the critical infrastructure to learn from a disruptive event (e.g. a revision of plans, modification of procedures, introduction of new tools and technologies) [25]. It is measured as the ratio of change in functionality level during and after the disruptive event over the initial functionality level:

$$IAT = \frac{FL_5(t) - FL_0(t)}{FL_0(t)} \cdot 100\%. \quad (11)$$

According to [25], such resilience macro-indicators are ideal for comparing the functionality level responses for multiple case studies, critical infrastructures, entities, facilities and assets, etc. They allow an objective evaluation of not only how the functionality level of a system might react to a disruptive event, but also how and when it can recover. Using a theoretical acceptance level, a stress-test can also be performed.

Other important factors to take into consideration when quantifying the resilience of critical infrastructures are: reliability and recoverability of the critical infrastructure components, as well as maintainability and supportability of the disrupted system elements, the prognostics and health management efficiency of the critical infrastructure system in the case of disruption.

The reliability function of critical infrastructure R_{CI} is formally defined as the probability that the system will not fail during the specified period of time t under given conditions.

$$R^{CI}(t) = \Pr(\text{the system doesn't fail during } [0, t]) = 1 - F(t), \quad (12)$$

where reliability $R^{CI}(t)$ is a decreasing function with time t , i.e. for $t_1 < t_2$, $R^{CI}(t_1) \geq R^{CI}(t_2)$, and it is usually assumed that $R^{CI}(0) = 1$; $F(t)$ is a failure function and is a basic (logistic) reliability measure which is defined as the probability that an element of critical infrastructure will fail before or at the moment of op-

erating time t ; t is a system operation time which is used in a generic sense and can have units such as hours, number of cycles, etc.

$$F(t) = \Pr(\text{failure will occur before or at the time } t) = \Pr(TTF \leq t), \quad (13)$$

$$F(t) = \int_0^t f(u) du, \quad (14)$$

where $f(t)$ is the probability density function of the time-to-failure random variable (TTF) in the case of an absolutely continuous distribution function.

Recoverability can be expressed as a non-linear function of system reliability, indicating that the performance of recovery actions is affected by the health of the critical infrastructure system. Special cases of the hybrid and gamma families of recoverability functions expressed in terms of a measure of system functionality (performance) level FL are proposed in study [27]:

$$1) FL(t) = 1 - \exp[-c\tilde{t}], \text{ when } TP_0 = 1, a = 1, b = 1, g(t) = \tilde{t}; \quad (15)$$

$$2) FL(t) = TP_0 - (TP_0 - FL_{\min}) \cdot \exp[-c\tilde{t}], \text{ when } a = TP_0 - FL_{\min}, b = 0, g(t) = \tilde{t}; \quad (16)$$

$$3) FL(t)_{Exp} = 1 - \alpha \cdot \exp\left[-c \frac{\tilde{t}}{\tilde{T}_{rec}}\right], \text{ when } TP_0 = 1, b = 0, g(t) = \frac{\tilde{t}}{\tilde{T}_{rec}}; \quad (17)$$

$$4) FL(t)_{lin} = 1 - \frac{\tilde{T}_{rec} (TP_0 - FL_{\min})}{2TP_0 t} \text{ and } FL(t)_{step} = 1 - \frac{\tilde{T}_{rec} (TP_0 - FL_{\min})}{TP_0 t}, \text{ when } TP_0 = 1, a = TP_0 - FL_{\min}$$

$$, b = 0, c = 0, g(t)_{lin} = \frac{\tilde{T}_{rec}}{2TP_0 t}, g(t)_{step} = \frac{\tilde{T}_{rec}}{TP_0 t}; \quad (18)$$

$$5) FL(t)_{crit_damp} = \frac{d}{dt} FL(t)|_{t=0} t \exp[-\omega t], \text{ when } TP_0 = 0, a = \frac{d}{dt} FL(t)|_{t=0}, b = 1, c = \omega; \quad (19)$$

where $FL(t)$ is the measure of system functionality (performance) which is quantifiable and time-dependent and is a composite function of time; TP_0 is the target functionality level before the disruption; $[TP_0 - \eta, TP_0 + \eta]$ is the system robustness range; η is the robustness parameter which characterizes how much system performance level can deviate from the target without affecting its core functionalities; a, b, c are parameters to be estimated; $a, a \geq 0$ is a location parameter that is associated with the magnitude of the maximum incurred functionality loss, which depends upon the severity of the disruption and the extent to which the system absorbs the shock; $b, b \geq 0$ is a shape parameter that is associated with the rates of functionality loss and restoration; $c, c \geq 0$ is a scale parameter which indicates the magnitude of the functionality loss over time for fixed loss and restoration rates, has a constant effect on the recovery process during the entire period and, therefore, is associated with the degree of absorptive capability, which is intrinsic to the system, and persists over time; parameters b, c characterize the time to recovery; ω is a parameter that describes the natural frequency with which the system would oscillate if no damping occurred; $g(t)$ is a non-decreasing function such that $g(0) = 0$ and describes system performance monotonic time-domain transformations and contributes into recovery function; $\tilde{t} = t - t_{\min}$ and $\tilde{T}_{rec} = T_{rec} - t_{\min}$; T_{rec} is the time to recovery.

On the other hand, the study [28] provides a basis for estimating the recoverability of critical infrastructures using the following formulations:

$$REC^{CI} = D(t) \times RA(t) \times RC(t), \quad (20)$$

where $D(t)$ is the diagnosis capability which is the ability of a critical infrastructure system to identify true failure elements or failure modes when disruption occurs; $RA(t)$ is the resource accessibility which is the ability of a critical infrastructure system to access enough resources after correct diagnosis; $RC(t)$ is the

repair capability which is the ability of a critical infrastructure system to accomplish the repair process after receiving enough resources.

Diagnosis capability of a critical infrastructure system is formulated as follows:

$$D(t) = \mu_D \cdot \frac{1}{e^{\alpha t_D}}, \quad (21)$$

where μ_D is the diagnosis accuracy; t_D is the diagnosis time; α is the coefficient of an exponential utility function used to consider the time effect.

To quantify the resource accessibility of a critical infrastructure system the following formulations are used:

$$RA(t) = \Omega(ava, req) \cdot u(t_{RA}), \quad u(t_{RA}) = \frac{1}{e^{\beta t_{RA}}}, \quad (22)$$

where ava is the available amount of resources; req is the required amount of resources; t_{RA} is the time to obtain the resources which is affected by the design of critical infrastructure, resource allocation, amount of required extra resources, etc.; $u(t_{RA})$ is the time utility function; β is the coefficient of the utility function; $\Omega(ava, req)$ is the resource index function.

Quantification of the repair capability is provided by the formulae:

$$RC(t) = L \cdot u(t_w) \cdot u(t_{R\gamma}) \cdot \kappa_\gamma, \quad u(t_w) = \frac{1}{e^{\omega t_w}}, \quad u(t_{R\gamma}) = \frac{1}{e^{\gamma t_{R\gamma}}}, \quad (23)$$

where L is the labor availability; t_w is the required time to retrieve the labor; t_{RC} is the repair time related with available technology, structural design of the element or system, and retest process after the repair; κ_γ is the successful repair rate; $u(t_w)$ and $u(t_{R\gamma})$ are the utility functions of required time and repair time, respectively; ω and γ are the utility coefficients, respectively.

Multiplication of the expressed formulations denotes that the failed critical infrastructure elements or system as a whole can only be recovered with successful operation of all three actions.

The efficiency of the prognostic and health management (PHM) system before and after disruption can be defined as the performance of a critical infrastructure system to failure detection, diagnosis and prediction the effects of possible adverse events [29]. This index is used to maintain and increase the backbone resilience capacities described above. According to researches [29, 30], PHM efficiency is mainly determined by the probability of the correct failure diagnosis event and the probability of the correct failure prognosis event by applying Fuzzy Fault Tree Analysis. At the same time, the efficiency of system PHM depends on the accuracy of defect detection and failure prediction by the critical infrastructure operators and maintenance personnel. Thus, the probability of failure of the PHM system efficiency can be estimated using the following formulation proposed in [29]:

$$FP(\Lambda_{PHM}) = \prod_{i=1}^m FP(BE_i), \quad (24)$$

$$P(\Lambda_{PHM}) = 1 - FP(\Lambda_{PHM}), \quad (25)$$

where $FP(BE_i)$ is the failure possibility of i -th basic event; m is the number of basic events; $P(\Lambda_{PHM})$ is the efficiency index of the PHM system which is equal to the complement of the failure possibility of this system.

Another formal expression of the critical infrastructure system PHM efficiency also mostly applied in hard resilience studies can be given as follows:

$$PHM^{CI} = \left(1 - \frac{n}{m}\right) \cdot \left(1 - \frac{t_{dp}}{t_m}\right) \cdot \left(1 - \frac{k}{m}\right) \cdot 100\%, \quad (26)$$

where n is the number of detected defects (failures); m is the total number of supervisions (predictions); t_{dp} is the time between failure detection and its prevention or elimination; t_m is the total observation (prognosis) period; k is the number of false alarms of the PHM system.

The higher the value of this indicator, the more efficient the PHM system is in the context of the critical infrastructure resilience management.

For any critical infrastructure, the system maintainability (M^{CI}) can be calculated using the following equations:

$$M^{CI}(t) = 1 - e^{(-\mu t)}, \quad \mu = \frac{1}{MTTR}, \quad MTTR^{CI} = \frac{\sum_{i=1}^n \frac{MTTR_i}{MTBF_i}}{\sum_{i=1}^n \frac{1}{MTBF_i}}, \quad (27)$$

where μ is the repair rate; $MTTR_{CI}$ is a mean time to repair of the critical infrastructure and is calculated as a function in mean time to repair ($MTTR$) and mean operating time between failure ($MTBF$) of critical infrastructure element i ; n is a number of critical infrastructure elements.

$MTTR$ represents the expectation of the time to system restoration. $MTBF$ is extremely difficult to predict for fairly reliable system elements. However, it can be estimated if the appropriate failure data are available. In fact, it is very rarely predicted with an acceptable accuracy.

Consequently, the value of the operational availability of critical infrastructure A^{CI} can be determined by the following formula:

$$A^{CI} = \frac{MTTF}{MTTF + MTTR + MTTT}, \quad (28)$$

where $MTTF$, $MTTR$ and $MTTS$ represent the mean time to failure, mean time to repair and mean time to support, and are measures of the system reliability, maintainability and supportability characteristics, respectively.

Mean time to failure ($MTTF$) represents the expectation of the time to failure and is used as a measure of reliability for non-repairable system elements. Mathematically, $MTTF$ can be defined as follows:

$$MTTF = \int_0^{\infty} rf(t) dt = \int_0^{\infty} R^{CI}(t) dt. \quad (29)$$

$MTTS$ can be defined as a term that represents the expectation of the time to support and is a measure of the critical infrastructure supportability characteristics. $MTTS$ is a measure of an on-product maintainability characteristic related to servicing that is calculated by dividing the total scheduled crew/operator/driver servicing time by the number of times the item was serviced.

The discussed resilience capacity models can be adapted in various ways and applied to all types of critical infrastructures and resilience domains for the overall resilience index assessment and analysis of the given class of complex dynamic systems.

Conclusion

Through the last decades, critical infrastructures have progressively begun to be the most essential complex systems influencing the socio-economic development and public welfare as well. In this connection, concerns about the protection and maintenance of critical infrastructures result into a series of state and sector-specific programs targeted to improve security and lately the resilience of this class of systems for withstanding multiple threats and hazards. The high level objective of the most of these programs is development of standards and guidelines for identifying risk factors and interdependencies, evaluating threats and impact, preparing measures to reduce vulnerabilities and to mitigate the consequences of disruptive events, as well as establishing best practices to increase resilience, validating and operationalizing methodologies and tools for system resilience management support in practice. To achieve this goal, the multi-disciplinary integrated studies in the line of critical infrastructure resilience assessment and analysis should be first of all carried out.

The main difference between foreign and Russian studies and practices in the field of critical infrastructure resilience management consist in the fact that Russian approaches are mostly focused on pre-event and during disruption measures (prevention and absorption phases, respectively) for the resilience maintenance, while the foreign methodologies concentrate on the post-event measures along with that, and enclose the coping of recovery and adaptation phases as well. At the same time, both ways are complementary and accompanying within the specific case studies of infrastructure resilience issues, notably, resilience estimation and control problems of critical entities or assets.

Nevertheless, it is worth noting that some shortcomings and contradictions exist between the science and practice of resilience management that should be eliminated. In particular, the current well-developed methods of crisis and risk management require modification and adaptation in the face of new challenges brought by the real practice of situational control of critical entities, as well as approaches known from theory can be rather inefficient for protection and resilience maintenance of critical infrastructures when certain theoretical scenarios are irrelevant and mismatched to current threats which may be more complicated, compounding, diverse or unexpected in reality. Moreover, theoretical models for resilience management seems to be ideal and verified, but in real applications can meet complications due to the uncertainties, restrictions, resource limits, changing operation conditions or other influencing factors that are not fully accounted within these models. In addition, theoretical and empirical methods cannot cope all of interdependencies between situational factors, resilience aspects and dynamic characteristics of critical infrastructures when assessing the overall system resilience in real practice. The manner of system behavior and latent nature of dependences between the interconnected critical entities may differ also on conceptual (research) and operational (applied) level of resilience management framework operationalization to critical infrastructures.

To level these bottlenecks, the enhancing of critical infrastructure resilience requires regular evaluation and strengthening the capabilities of critical infrastructures to anticipate and prevent (preparedness, predictability), to resist and absorb (withstandability, absorbability), to react and recover (recoverability, responsiveness), to adapt and transform (adaptability, transformability) in the face of context-dependent disruptive events, adverse circumstances and perturbations. Thereto, an effort to develop a more complete assessment and analysis procedure of the critical infrastructure resilience has been made. It is index-based and applicable to all types of critical infrastructures of the regional scale. The proposed method uses selected estimation models of the resilience capacities and provides quantification of the level and ratio of aggregated reliability, robustness and security indices, as well as the calculation of performance level, savings and losses rate, and control risks for obtaining end-to-end resilience assessment within the all stages of the resilience management cycle.

Combining the developed method with other models of resilience capabilities and indicators allows for a comprehensive assessment of systemic risks that can support decision-making about protection, emergency and situational management of regional critical infrastructures, and thus, in conjugation with other resilience measurement tools and frameworks, allows critical infrastructure operating conditions to be compared in terms of performance characteristics, vulnerabilities, threat impacts, possible consequences, effectiveness of the preventive/mitigation measures and ultimately resilience control strategies.

This research outputs can be practically used as reliable guidance for operators and analysts of regional situational centers to train and generate design decisions about counteracting the current threats, actuating hazards and local failures in the operation of sector-specific critical infrastructure systems under uncertain situations. It is urgent and imperative to get a relevant, holistic comparative picture on the respective functionality level of critical entities and infrastructures based on adequate assessments to control and improve their resilience efficiently. In this case, the proposed method is supposed to be implemented and introduced within the decision support systems of regional and sector situational centers controlling critical entities, or in wider scale applications. In the reality, however, critical infrastructure managers are reluctant to become compared by the auditing services or security authorities, and, naturally, are uninterested to reveal the detailed resilience level across indicators and the current points of system vulnerability as well. Therefore, with a view to this fact, the proposed method can be positioned as a self-assessment tool rather than a regulative and control mechanism of the public authorities.

The future research will be aimed at analysis of the existing normative documents and legal acts adopted in the field of resilience management of critical infrastructures that are regulating and reasoning the assessment criteria and procedure of the critical entities resilience and protection. As a result, findings will be used as a basis for precision adjustment of the proposed estimation method and its further implementa-

tion as a software tool suitable and easily tailorable to specific managerial and information support needs of the resilience maintenance, assessment and control.

References

1. Masloboev A.V. An index-based method for integral estimation of regional critical infrastructure resilience using fuzzy calculations. Part 1. Problem statement and method generic structure. *Reliability and quality of complex systems*. 2024;(1):124–141.
2. Reitan N.K. et al. Evaluation of resilience concepts applied to critical infrastructure using existing methodologies. *IMPROVER Project Report: Deliverable 2.3*. 2016:97.
3. Melkunaite L. et al. International Survey. *IMPROVER Project Report: Deliverable 1.1*. 2016:343.
4. Yang Zh. et al. Indicator-based resilience assessment for critical infrastructures – A review. *Safety Science*. 2023;160:106049.
5. Osei-Kyei R., Almeida L.M., Ampratwum G., Tam V. Systematic review of critical infrastructure resilience indicators. *Construction Innovation*. 2023;23(5):1210–1231.
6. Dan G., Shan M., Owusu E.K. Resilience Assessment Frameworks of Critical Infrastructures: State-of-the-Art Review. *Buildings*. 2021;11(10):464.
7. Jovanović A., Klimek P., Renn O. et al. Assessing resilience of healthcare infrastructure exposed to COVID-19: emerging risks, resilience indicators, interdependencies and international standards. *Environment Systems and Decisions*. 2020;40(2):252–286.
8. Yang Zh. et al. A multi-criteria framework for critical infrastructure systems resilience. *International Journal of Critical Infrastructure Protection*. 2023;42:100616.
9. Mottahedi A., Sereshki F., Ataei M., Nouri Qarahasanlou A., Barabadi A. The Resilience of Critical Infrastructure Systems: A Systematic Literature Review. *Energies*. 2021;14(6):1571.
10. Almaleh A. Measuring Resilience in Smart Infrastructures: A Comprehensive Review of Metrics and Methods. *Applied Sciences*. 2023;13(11):6452.
11. Sathurshan M., Saja A., Thamboo J., Haraguchi M., Navaratnam S. Resilience of Critical Infrastructure Systems: A Systematic Literature Review of Measurement Frameworks. *Infrastructures*. 2022;7(5):67.
12. Wells E.M., Boden M., Tseytlin I., Linkov I. Modeling critical infrastructure resilience under compounding threats: A systematic literature review. *Progress in Disaster Science*. 2022;15:100244.
13. Rehak D., Senovsky P., Hromada M., Lovecek T. Complex Approach to Assessing Resilience of Critical Infrastructure Elements. *International Journal of Critical Infrastructure Protection*. 2019;25:125–138.
14. Rød B., Barabadi A., Gudmestad O.T. Characteristics of Arctic infrastructure resilience: Application of expert judgement. *Proceedings of the 26th International Ocean and Polar Engineering Conference (Rhodes, Greece, June 26 – July 1, 2016)*. Rhodes, Greece, 2016:1226–1233.
15. Honfi D. et al. Technological resilience concepts applied to critical infrastructure. *IMPROVER Project Report: Deliverable 3.2*. 2017:67.
16. Davis C.A. Understanding Functionality and Operability for Infrastructure System Resilience. *Natural Hazards Review*. 2021;22(1):0000431.
17. Youn B.D., Hu C., Wang P.F. Resilience-driven system design of complex engineered systems. *Journal of Mechanical Design*. 2011;133(10):101011(15).
18. Barabadi A., Markeset T. Reliability and maintainability performance under Arctic conditions. *International Journal of System Assurance Engineering and Management*. 2011;(2):205–217.
19. Saraswat S., Yadava G. S. An overview on reliability, availability, maintainability and supportability (RAMS) engineering. *International Journal of Quality and Reliability Management*. 2008;25(3):330–344.
20. Smith, C., Knezevic, J. Achieving quality through supportability: Part 1. Concepts and Principles. *Journal of Quality in Maintenance Engineering*. 1996;2(2):21–29.
21. Tortorella M. *Reliability, Maintainability, and Supportability: Best Practices for Systems Engineers*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2015:464.
22. Rehak D. Assessing and strengthening organisational resilience in a critical infrastructure system: case study of the Slovak Republic. *Safety Science*. 2020;123:104573.
23. Lyttle D.N., Gill J.P., Shaw K.M. et al. Robustness, flexibility, and sensitivity in a multifunctional motor control model. *Biological Cybernetics*. 2017;111:25–47.
24. Prior T. *Measuring Critical Infrastructure Resilience: Possible Indicators. Risk and Resilience Report 9*. Zurich, Switzerland, Center for Security Studies, ETH Zurich. 2015:13.
25. Jovanović A., Øien K., Jelic M. et al. Modeling the impact of an adverse event on the "absorb" and "recover" capacity of a smart critical infrastructure, based on resilience indicators. *H2020 Project: Smart Resilience Indicators for Smart Critical Infrastructure. Report Deliverable No: D3.3*. Stuttgart, 2018:53.
26. Davies B. *Integral Transforms and their Applications. Third Edition. Applied Mathematics*. Springer Science & Business Media, 2012;41:370.

27. Cassottana B., Shen L., Tang L. Ch. Modeling the recovery process: A key dimension of resilience. *Reliability Engineering and System Safety*. 2019;190:106528.
28. Li J., Xi Z. Engineering Recoverability: A New Indicator of Design for Engineering Resilience. *Proceedings of the ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. 40th Design Automation Conference. August 17–20, 2014. Buffalo, New York, USA. 2014;2A:V02AT03A044.*
29. Rahimazar A., Nouri Qarahasanlou A., Khanzadeh D., Tavaghi M. Assessing resilience in mechanical systems: an industrial perspective. *International Journal of Quality & Reliability Management*. 2024.
30. Yurkov N.K., Mikhaylov V.S. *Integral'nye otsenki v teorii nadezhnosti. Vvedenie i osnovnye rezul'taty = Integral estimates in reliability theory. Introduction and main results.* Moscow: Tekhnosfera, 2020:152. (In Russ.)

Список литературы

1. Masloboev A. V. An index-based method for integral estimation of regional critical infrastructure resilience using fuzzy calculations. Part 1. Problem statement and method generic structure // *Reliability and quality of complex systems*. 2024. № 1. P. 124–141.
2. Reitan N. K. et al. Evaluation of resilience concepts applied to critical infrastructure using existing methodologies // *IMPROVER Project Report: Deliverable 2.3*. 2016. 97 p.
3. Melkunaite L. et al. International Survey // *IMPROVER Project Report: Deliverable 1.1*. 2016. 343 p.
4. Yang Zh. et al. Indicator-based resilience assessment for critical infrastructures – A review // *Safety Science*. 2023. Vol. 160. P. 106049.
5. Osei-Kyei R., Almeida L.M., Ampratwum G., Tam V. Systematic review of critical infrastructure resilience indicators // *Construction Innovation*. 2023. Vol. 23, № 5. P. 1210–1231.
6. Dan G., Shan M., Owusu E.K. Resilience Assessment Frameworks of Critical Infrastructures: State-of-the-Art Review // *Buildings*. 2021. Vol. 11, № 10. P. 464.
7. Jovanović A., Klimek P., Renn O. et al. Assessing resilience of healthcare infrastructure exposed to COVID-19: emerging risks, resilience indicators, interdependencies and international standards // *Environment Systems and Decisions*. 2020. Vol. 40, № 2. P. 252–286.
8. Yang Zh. et al. A multi-criteria framework for critical infrastructure systems resilience // *International Journal of Critical Infrastructure Protection*. 2023. Vol. 42. P. 100616.
9. Mottahedi A., Sereshki F., Ataei M., Nouri Qarahasanlou A., Barabadi A. The Resilience of Critical Infrastructure Systems: A Systematic Literature Review // *Energies*. 2021. Vol. 14, № 6. P. 1571.
10. Almaleh A. Measuring Resilience in Smart Infrastructures: A Comprehensive Review of Metrics and Methods // *Applied Sciences*. 2023. Vol. 13, № 11. P. 6452.
11. Sathurshan M., Saja A., Thamboo J., Haraguchi M., Navaratnam S. Resilience of Critical Infrastructure Systems: A Systematic Literature Review of Measurement Frameworks // *Infrastructures*. 2022. Vol. 7, № 5. P. 67.
12. Wells E. M., Boden M., Tseytlin I., Linkov I. Modeling critical infrastructure resilience under compounding threats: A systematic literature review // *Progress in Disaster Science*. 2022. Vol. 15. P. 100244.
13. Rehak D., Senovsky P., Hromada M., Lovecek T. Complex Approach to Assessing Resilience of Critical Infrastructure Elements // *International Journal of Critical Infrastructure Protection*. 2019. Vol. 25. P. 125–138.
14. Rød B., Barabadi A., Gudmestad O. T. Characteristics of Arctic infrastructure resilience: Application of expert judgement // *Proceedings of the 26th International Ocean and Polar Engineering Conference (Rhodes, Greece, June 26 – July 1, 2016)*. Rhodes, Greece, 2016. P. 1226–1233.
15. Honfi D. et al. Technological resilience concepts applied to critical infrastructure // *IMPROVER Project Report: Deliverable 3.2*. 2017. 67 p.
16. Davis C. A. Understanding Functionality and Operability for Infrastructure System Resilience // *Natural Hazards Review*. 2021. Vol. 22, iss. 1. P. 0000431.
17. Youn B. D., Hu C., Wang P. F. Resilience-driven system design of complex engineered systems // *Journal of Mechanical Design*. 2011. Vol. 133. P. 10. P. 101011.
18. Barabadi A., Markeset T. Reliability and maintainability performance under Arctic conditions // *International Journal of System Assurance Engineering and Management*. 2011. № 2. P. 205–217.
19. Saraswat S., Yadava G. S. An overview on reliability, availability, maintainability and supportability (RAMS) engineering // *International Journal of Quality and Reliability Management*. 2008. Vol. 25, № 3. P. 330–344.
20. Smith, C., Knezevic, J. Achieving quality through supportability: Part 1. Concepts and Principles // *Journal of Quality in Maintenance Engineering*. 1996. Vol. 2, № 2. P. 21–29.
21. Tortorella M. *Reliability, Maintainability, and Supportability: Best Practices for Systems Engineers*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2015. 464 p.
22. Rehak D. Assessing and strengthening organisational resilience in a critical infrastructure system: case study of the Slovak Republic // *Safety Science*. 2020. Vol. 123. P. 104573.
23. Lyttle D. N., Gill J. P., Shaw K. M. et al. Robustness, flexibility, and sensitivity in a multifunctional motor control model // *Biological Cybernetics*. 2017. Vol. 111. P. 25–47.

24. Prior T. Measuring Critical Infrastructure Resilience: Possible Indicators. Risk and Resilience Report 9. Zurich, Switzerland, Center for Security Studies, ETH Zurich. 2015. 13 p.
25. Jovanović A. Øien K., Jelic M. et al. Modeling the impact of an adverse event on the "absorb" and "recover" capacity of a smart critical infrastructure, based on resilience indicators // H2020 Project: Smart Resilience Indicators for Smart Critical Infrastructure. Report Deliverable No: D3.3. Stuttgart, 2018. 53 p.
26. Davies B. Integral Transforms and their Applications. Third Edition. Applied Mathematics. Springer Science & Business Media, 2012. Vol. 41. 370 p.
27. Cassottana B., Shen L., Tang L. Ch. Modeling the recovery process: A key dimension of resilience // Reliability Engineering and System Safety. 2019. Vol. 190. P. 106528.
28. Li J., Xi Z. Engineering Recoverability: A New Indicator of Design for Engineering Resilience // Proceedings of the ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. 40th Design Automation Conference. August 17-20, 2014, Buffalo, New York, USA. 2014. Vol. 2A. P. V02AT03A044.
29. Rahimazar A., Nouri Qarahasanlou A., Khanzadeh D., Tavaghi M. Assessing resilience in mechanical systems: an industrial perspective // International Journal of Quality & Reliability Management. 2024.
30. Юрков Н. К., Михайлов В. С. Интегральные оценки в теории надежности. Введение и основные результаты. М. : Техносфера, 2020. 152 с.

Информация об авторах / Information about the authors

Андрей Владимирович Маслобоев

доктор технических наук, доцент,
ведущий научный сотрудник лаборатории
информационных технологий управления
техногенно-природными системами,
Институт информатики и математического
моделирования имени В. А. Путилова
Кольского научного центра Российской академии наук;
главный научный сотрудник,
Институт проблем промышленной экологии Севера
Кольского научного центра Российской академии наук
(Россия, Мурманская область, г. Апатиты,
ул. Ферсмана, 14)
E-mail: masloboev@iimm.ru

Andrey V. Masloboev

Doctor of technical sciences, associate professor,
leading researcher of the laboratory of information
technologies for industrial-natural system management,
Putilov Institute for Informatics
and Mathematical Modeling of the Kola Science
Centre of the Russian Academy of Sciences;
chief researcher,
Institute of North Industrial Ecology Problems
of the Kola Science Centre
of the Russian Academy of Sciences
(14 Fersmana street, Apatity, Murmansk region, Russia)

Автор заявляет об отсутствии конфликта интересов /

The author declares no conflicts of interests.

Поступила в редакцию/Received 28.05.2024

Поступила после рецензирования/Revised 25.06.2024

Принята к публикации/Accepted 16.08.2024