

На правах рукописи



Токарева Виктория Андреевна

Математические модели и алгоритмы для формирования расписания в распределённых системах обработки данных с агрегированным доступом к информационным ресурсам

Специальность 05.13.18

«Математическое моделирование, численные методы и комплексы программ»

АВТОРЕФЕРАТ

диссертации на соискание учёной степени
кандидата физико-математических наук

Воронеж — 2022

Работа выполнена на кафедре математических методов исследования операций факультета прикладной математики, информатики и механики ФГБОУ ВО «Воронежский государственный университет»

Научный руководитель: кандидат физико-математических наук,
доктор технических наук, доцент,
Бондаренко Юлия Валентиновна

Официальные оппоненты: **Лазарев Александр Алексеевич**,
доктор физико-математических наук, профессор,
ФГБУН «Институт проблем управления им. В.А. Трапезникова РАН», лаборатория «Теории расписаний и дискретной оптимизации», и.о. заведующего, главный научный сотрудник

Киселева Екатерина Игоревна,
кандидат физико-математических наук,
ФГБОУ ВО «Воронежский государственный педагогический университет», кафедра педагогики и методики дошкольного и начального образования, доцент

Ведущая организация: Федеральное государственное автономное образовательное учреждение высшего образования «Южный федеральный университет»

Защита состоится «29» июня 2022 г. в 15.10 часов на заседании диссертационного совета Д 212.038.20 на базе Воронежского государственного университета по адресу: 394018, Россия, г. Воронеж, Университетская площадь 1, ауд. 333.

С диссертацией и авторефератом можно ознакомиться в библиотеке Воронежского государственного университета, а также на сайте http://www.science.vsu.ru/dissertations/10330/Диссертация_Токарева_В.А..pdf

Отзывы на автореферат в двух экземплярах, заверенные печатью учреждения, просьба направлять по адресу: 394018, Россия, г. Воронеж, Университетская площадь 1, учёному секретарю диссертационного совета Д 212.038.20

Автореферат разослан «__» ____ 20__ г.

Учёный секретарь диссертационного совета Д 212.038.20,
доктор физико-математических наук



С.А. Шабров

Общая характеристика работы

Актуальность темы. Усиление цифровизации общества в представлении широкого круга исследователей связывается с социально-экономическим феноменом больших данных и вызванного им интенсивного развития прикладных информационных систем различного назначения, таких как системы электронной коммерции, автоматизированные системы управления различными процессами, системы компьютерного моделирования, системы автоматизации (научных исследований, проектирования, производства) и др.

Сопутствующая этому развитию эволюция информационных технологий приводит, с одной стороны, к увеличению количества данных, обрабатываемых такими системами, и, как следствие, к ужесточению требований, предъявляемых к их работе, увеличению вычислительной нагрузки как в плане количества обрабатываемых информационных единиц, так и в плане роста числа выполняемых внутренних задач. Работа таких систем требует эффективной диспетчеризации, которая может быть достигнута за счёт использования моделей и алгоритмов такой области исследования операций, как теории расписаний. С другой стороны, наблюдается тренд на дифференциацию систем обработки больших данных, что позволяет выделять новые классы таких систем и разрабатывать для них более эффективные специализированные подходы к обработке информации.

Таким образом, актуальным становится развитие математического аппарата теории расписаний, создание эффективных алгоритмов диспетчеризации и их реализации в виде комплексов проблемно-ориентированных программ.

Степень разработанности темы. Научный поиск моделей и алгоритмов составления расписаний опирается прежде всего на формальное описание свойств ресурсов, их типологию и ограничений на количество доступных ресурсов в системе, раскрываемых в работах таких отечественных учёных, как В.Н. Бурков, Е.Р. Гафаров, В.С. Гордон, А.А. Лазарев, В.С. Танаев, Я.М. Шафранский и др., в частности, в работах воронежских учёных: Т.В. Азарновой, А.Я. Асниной, С.А. Баркалова, Ю.В. Бондаренко, Т.М. Леденевой, М.Г. Матвеева и др. Среди зарубежных учёных можно выделить таких как: Я. Блажевич, П. Брукер, Г. Валигора, Д.С. Джонсон, Г. Кендалл и др.

Теоретической основой исследования послужили модели и алгоритмы составления расписаний для приборов и сетевого планирования. Выдающиеся результаты в области составления расписаний для приборов с дискретными ограничениями на ресурсы (таких как обслуживание на одном приборе, и различных постановок задач цеха для нескольких приборов) получены таким авторами, как М.Р. Гэри, А.Х.Г. Ринной Кан, Э.Г. Кофман-мл, Р.Дж. Деннинг и др. Составление расписаний для приборов в задачах с непрерывными ресурсными ограничениями рассматривалось в работах Я. Юзефовски, Р.М. Карпа, А. Яняка, Е. Новицки и др. Различные модификации модели минимизации времени выполнения проекта с учетом ограничения на ресурсы (обозначаемой в литературе как RCPSP), а также модели выравнивания по ресурсам (RLP), поиска компромисса между временем и затратами и вероятностные модели были разработаны такими авторами как В. де Рейк, С. Барум, А. Шпрехер и др.

В основу разработки алгоритмов в данном исследовании положен анализ генетических алгоритмов, алгоритмов муравьиных колоний, методов динамического перебора, а так же широко применяемых в системных планировщиках ИС робастных эвристик, таких как простые правила диспетчеризации (SPT, STF и др.) или алгоритм Round-robin (А.К. Гупта, Н. Арора, Ф. Алаа, М.М. Зулиха и др.).

Однако, несмотря на несомненную значимость проанализированных работ, изложенные в них результаты, хотя и являются достаточными для решения широкого круга про-

блем, требуют расширения математического аппарата моделями и алгоритмами, обеспечивающими построение расписаний с учётом специфики информационных ресурсов. Существенным отличием таких моделей и алгоритмов должно стать такое современное требование работы с большими данными, как обеспечение баланса между эффективностью расписания и временем выполнения расчета. В этой связи разработка математических моделей, формирование эффективных алгоритмов и проблемно-ориентированных комплексов программ для построения расписаний в системах с ограничениями доступа к распределенным информационным ресурсам является важной научной и прикладной проблемой.

Объект исследования. Процессы формирования расписаний в системах с несколькими приборами и дополнительными возобновляемыми ресурсами, ограниченными в терминах качественной доступности применительно к классу распределённых систем хранения и обработки данных, использующих удалённые информационных ресурсы.

Предмет исследования. Дискретные и дискретно-непрерывные математические модели, методы, вычислительные алгоритмы, предназначенные для составления расписаний в системах обработки данных с несколькими приборами и дополнительными ресурсами.

Целью исследования является разработка и теоретическое обоснование новых математических моделей, быстрых вычислительных алгоритмов для составления расписаний в распределенных системах обработки данных с агрегированным поиском, основанных на учёте ограничений на дополнительные информационные ресурсы, а также создание комплексов параллельных программ для оперативного решения задач. Для достижения поставленной цели необходимо решить следующие **задачи**:

- провести анализ процессов формирования расписания в информационных системах с агрегированным доступом к данным, обзор существующих математических моделей и алгоритмов теории расписаний с целью выявления основным проблем применения аппарата теории расписаний в прикладных информационных системах и обоснование актуальности построения моделей, предусматривающих ограничения на дополнительные информационные ресурсы;
- разработать подход к математическому моделированию распределённых систем обработки данных с несколькими приборами и дополнительными информационными ресурсами;
- сформулировать постановку математической задачи поиска оптимального расписания в исследуемом классе систем;
- построить семейство моделей составления расписаний для нескольких приборов с ограничениями на доступность ресурсов;
- разработать алгоритмы и численные методы решения задач составления расписаний в многоприборных системах с ресурсами, обладающими ограниченной качественной доступностью;
- реализовать разработанные математические модели и алгоритмы в виде комплексов проблемно-ориентированных программ для проведения комплексного исследования проблем составления расписаний в распределённых системах обработки данных с несколькими приборами и дополнительными информационными ресурсами.

Соответствие диссертации Паспорту научной специальности. Исследование соответствует формуле специальности 05.13.18 «Математическое моделирование, численные методы и комплексы программ» по следующим пунктам: п. 4 «Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента», п. 5 «Комплексные исследова-

ния научных и технических проблем с применением современной технологии математического моделирования и вычислительного эксперимента», п. 8 «Разработка систем компьютерного и имитационного моделирования».

Материалы и методы исследования. Для решения задач, поставленных в работе, применялись фундаментальные положения ряда научных направлений: теории расписаний, теории массового обслуживания, математических методов исследования операций, статистики, теории вероятностей, теории алгоритмов.

Комплексы программ для создания имитационных моделей, проведения численных экспериментов и решения задач диспетчеризации в системе доступа и распределённой обработки данных GRADLCI реализованы на языке Python 3 для работы в дистрибутивах операционной системы Linux или MacOS с использованием библиотек plotly и matplotlib для визуализации результатов и библиотек subprocess и multiprocessing для параллельных асинхронных вызовов в многопроцессорной вычислительной системе для параллельной реализации разработанных алгоритмов.

В качестве **методологической основы** использован системно-аналитический подход с фокусом на структурном, функциональном и динамическом аспектах, численные методы, методы математического моделирования и программирования. В качестве вспомогательных методологических средств использованы контент-анализ, методы интеллектуальной схемотехники, сравнение, систематизация. Работа основывается на принципах системности, комплексности, детерминированности.

В диссертации получены следующие основные результаты, характеризующиеся **научной новизной**:

1. Предложено формальное описание класса распределённых систем обработки данных с агрегированным поиском, отличающихся наличием дополнительных информационных ресурсов ограниченных в терминах качественной доступности. Сформулирована математическая задача поиска оптимального расписания в рассмотренном классе систем.
2. Разработаны оптимизационные дискретная и дискретно-непрерывная математические модели составления расписаний для нескольких приборов, отличающиеся учетом дополнительных ограничений на качественную доступность ресурсов и функцией цели, заключающейся в минимизации времени выполнения агрегированной работы. Проведено исследование существования аналитического решения для случаев дискретных ограничений на ресурсы и различных видов функции расхода ресурса для модели с непрерывными ограничениями на ресурсы.
3. Сформированы новые алгоритмы решения задач составления расписаний в многоприборных системах с ресурсами, обладающими ограниченной качественной доступностью, основанные на приоритето-порождающих функционалах для различных соотношений числа приборов к числу ресурсов. Для предложенных алгоритмов исследована асимптотическая сходимость для худшего случая выполнения.
4. Разработаны комплексы программ для направленного численного эксперимента по исследованию свойств разработанных алгоритмов, имитационного моделирования поведения систем агрегации при различных заданных параметрах.
5. Разработанные численные методы и алгоритмы реализованы в виде комплекса программ для работы с пользовательскими заявками и диспетчеризации задач в центре сбора и анализа данных экспериментальной астрофизики частиц GRADLCI.

Практическая значимость заключается в построении комплекса дискретных и дискретно-непрерывных математических моделей и алгоритмов, позволяющих сформировать

ровать расписание обслуживания информационных запросов в широком классе систем с дополнительными информационными ресурсами за практически приемлемое время.

Разработаны комплексы программ на языке Python 3 для построения имитационных моделей и проведения численных экспериментов, использующие асинхронные распаралеленные вызовы, реализованные с использованием библиотек `subprocess` и `multiprocessing`, что определило такие отличительные свойства разработанных программ, как оптимизированное выполнение высоконагруженных вычислений, улучшенная масштабируемость, ускоренный по сравнению с последовательным выполнением отклик пользовательских интерфейсов и более эффективное использование системных ресурсов. Разработанные программные комплексы позволяют: создавать и конфигурировать имитационные модели распределённых систем обработки данных с агрегированным доступом к удалённым информационным ресурсам; осуществлять моделирование различных сценариев составления расписаний в многоприборных распределённых системах с ограниченными в терминах качественной доступности информационными ресурсами; осуществлять за практически приемлемое (полиномиальное либо линейно-логарифмическое) время составление расписаний обработки пользовательских запросов на среднесрочном временном интервале для пользовательских запросов, распределённых согласно экспоненциальному распределению с периодически изменяющимся параметром интенсивности потока заявок. Получены 3 свидетельства о регистрации ПО ЭВМ [12—14].

Разработанные алгоритмы и модели нашли своё применение при создании автором системы доступа к данным экспериментов астрофизики частиц GRADLCI, используемой профессиональными исследователями в области астрофизики частиц. Полученные результаты включены в программу учебных курсов «Теория игр и исследование операций», «Имитационное моделирование», «Имитационное моделирование в задачах машинного обучения», «Методы оптимизации», преподаваемых на факультете прикладной математики, информатики и механики ФГБОУ ВО «Воронежский государственный университет».

Основные результаты и положения, выносимые на защиту:

1. Формальный аппарат описания класса распределённых систем обработки данных с агрегированным поиском, отличающихся наличием дополнительных информационных ресурсов, позволяет выделить характерные особенности процесса составления расписаний и сформировать базовые составляющие для построения комплекса моделей для построения эффективных расписаний в данном классе систем.
2. Комплекс математических моделей составления расписаний для нескольких приборов, включающих дискретную и дискретно-непрерывную оптимизационные модели составления расписаний в многоприборных системах с ограничениями на дополнительные информационные ресурсы. В моделях учтены ограничения на количество соединений и скорость доступа к удалённым информационным ресурсам. Комплекс предназначен для формирования оптимальных расписаний по одному из следующих критериев: минимизации длины расписания, минимизации суммы времён обслуживания задач и минимизации суммарного времени выполнения агрегированных работ. Математическое исследование существования аналитического решения для случаев дискретных ограничений на ресурсы и различных видов функции расхода ресурса для модели с непрерывными ограничениями на ресурсы, позволяет обосновать использование численных алгоритмов решения.
3. Алгоритмы численного решения задач составления расписаний в многоприборных системах с ресурсами, обладающими ограниченной качественной доступностью, основанные на приоритето-порождающих функционалах для случаев числа приборов равного чис-

лу ресурсов и числа приборов меньшего числа ресурсов, делают возможным отыскание эффективного решения за практически приемлемое время.

4. Комплексы программ, включающие в себя: систему имитационного моделирования, предназначенную для создания имитационных моделей процессов формирования расписаний в распределённых системах обработки данных с агрегированным доступом к информационным ресурсам для различных комбинаций параметров; комплекс программ для реализации численных экспериментов, предназначенный для моделирования различных сценариев составления расписаний в многоприборных распределённых системах с агрегированным поиском; для работы с пользовательскими заявками и диспетчеризации задач в центре сбора и анализа данных экспериментальной астрофизики частиц GRADLCI.

Достоверность полученных результатов основывается на строгих формулировках и доказательствах и подтверждается проведёнными вычислительными экспериментами.

Апробация работы. Основные выводы и теоретические положения диссертационного исследования были представлены в докладах на научных мероприятиях международного и всероссийского уровня, в числе которых: The 8th International Conference “Distributed Computing and Grid-technologies in Science and Education” (Дубна, 2018); School for Astroparticle Physics Obertrubach-Bärnfels (Obertrubach, Germany, 2018); Matter and the Universe Days (Hamburg, Germany, 2019); Workshop “Big Data Science in Astroparticle Research” (Aachen, Germany, 2019, 2020); 6th KSETA Plenary workshop, (Durbach, Germany, 2019, 2020); 19th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (Saas Fee, Switzerland, 2019); DPG conference (Aachen-Berlin, Germany, 2019, 2021); International workshop “Data life cycle in physics” (Карлсруэ-Иркутск-Москва, 2018, 2019, 2020); Sparse Digital Radio Arrays online workshop (Москва, 2019); International Conference on Computer Simulation in Physics and beyond (Москва, 2020); The XXIV International Scientific Conference of Young Scientists and Specialists (Дубна, 2020); 13-я Международная конференция «Интеллектуализация обработки информации» (Москва, 2020); 5th International Workshop on Deep Learning in Computational Physics (Москва, 2021); 37th International Cosmic Ray Conference-2021 (Berlin, Germany, 2021); Международная научная конференция «Актуальные проблемы прикладной математики, информатики и механики» (Воронеж, 2021); семинар лаборатории №68 «Теории расписаний и дискретной оптимизации» ИПУ РАН; а также рабочих совещаниях коллаборации GRADLCI/APPDS и семинарах Института Астрофизики Частиц Технологического Института Карлсруэ.

Работа была выполнена при поддержке грантов РФФИ № 18-41-06003 и фонда Гельмгольца № HRSF-0027.

Публикации. Основные результаты по теме диссертации изложены в 15 статьях [1—15], 11 из которых проиндексированы базами Scopus и Web of Science [1—11] и 3 являются свидетельствами о регистрации ПО ЭВМ [12—14].

Личный вклад. Все результаты, изложенные в диссертации, получены автором самостоятельно. Из совместных работ в диссертацию включены только результаты, полученные лично автором.

Объём и структура работы. Диссертация состоит из введения, четырёх глав, заключения, библиографического списка и приложений. Полный объём диссертации составляет 175 страниц, включая 44 рисунка и 22 таблицы. Список литературы содержит 159 наименований.

Содержание работы

Во **Введении** обоснована актуальность темы диссертационного исследования, определена степень её разработанности, сформулированы объект, предмет, цель и задачи работы, выделена научная новизна, обоснована практическая значимость, приведены материалы и методы исследования, а также основные положения, выносимые на защиту, представлены сведения об апробации и соответствии паспорту научной специальности.

В **первой главе** представлен обзор существующих исследований в теории расписаний, подходы к записи и классификации задач, разделение задач теории расписаний на классы разрешимости. Выполнен анализ дискретных и дискретно-непрерывных моделей составления расписаний в системах с ограничениями на дополнительные ресурсы.

Анализ существующих классификаций ресурсных ограничений показал, что в существующих моделях не учитываются случаи, когда ограниченная доступность ресурсов связана не непосредственно с его имеющимся количеством, а с совокупностью внешних факторов, ограничивающих как одновременный доступ к ресурсу несколькими приборами — обозначаемый в работе как ϕ , так и объём ресурса, получаемый системой в единицу времени $\hat{\mu}$ (далее определяемый как «скорость получения ресурса»). Предложена форма записи таких ограничений $\text{res}\lambda\sigma\rho\phi\hat{\mu}$, которая является расширением принятой в литературе нотации $\text{res}\lambda\sigma\rho$. Ресурсы, ограниченные в данном смысле, определены в категорию «качественно ограниченных».

Выделен и описан класс распределённых информационно-вычислительных систем «системы агрегации» (СА), отличительной чертой которых является наличие доступа к ряду изолированных распределённых информационных ресурсов, обладающих свойством качественной ограниченности. Произведено построение концептуальной Функциональной модели составления адаптивных расписаний в СА. Для уточнение свойств элементов системы и связей между ними разработана и проанализирована логическая модель Сущность-Связь для СА. Определен теоретико-множественный описатель СА.

Делаются выводы о необходимости адаптации и расширения существующих дискретных и дискретно-непрерывных моделей составления расписаний в системах с несколькими приборами и ограничениями на дополнительные ресурсы для учёта характеристик и потребностей СА.

Во **второй главе** исследуется влияние факторов качественной доступности ресурсов при составлении расписаний в многолинейных СА с информационными ресурсами, ограниченными в терминах качественной доступности. Рассматриваются случаи дискретных и дискретно-непрерывных ограничений на дополнительные ресурсы в системе.

Для случая дискретных ограничений приводится следующая модель:

$$\left\{ \begin{array}{l} C_a = \sum_{j=1}^n \max_{k=1, \dots, s} C_{jk} \rightarrow \min; C_{ji} = s_{ji} + p_{ji}; \\ p_{ji} = \mu_i V_{ji}; s_{ji} \geq r_{ji}; s_j = \min_{i=1, \dots, k} s_{ji}; r_{ji} = r_j, \forall i \\ PR_{js}(t) = \begin{cases} 1, & \text{если } T_{js} \text{ выполняется в момент времени } t; \\ 0, & \text{иначе;} \end{cases} \\ B_s(t) = \sum_{j=1}^n PR_{js} \leq L; \sum_{s=1}^S B_s(t) \leq m; B_s(t) \in 0, \dots, m; \\ j \in \overline{1, n}, k \in \overline{1, S}, i \in \overline{1, k}; L \in \mathbb{N}, m \in \mathbb{N}, \mu_s \in \mathbb{N}^s, V \in \mathbb{N}^{n \times s}, r_j \in \mathbb{N}. \end{array} \right. \quad (1)$$

Здесь J_j — работы, $j = 1, \dots, n$; T_{ji} — задачи, $j = 1, \dots, n, i \in i_1, \dots, i_k$; R_s — ресурсы, $s = 1, \dots, S$; P_l — приборы, $l = 1, \dots, m$; s_j, s_{ji} — времена начала работ и задач; r_j, r_{ji} — времена готовности задачи; C_j, C_{ji} — времена завершения работ и задач; $p_{ji} = C_{ji} - s_{ji}$ — оценка времени выполнения задачи; V_{ji} — ресурсоёмкость задачи; μ_s — скорость доступа к ресурсу; $PR_{j_s}(t)$ — выполнение задачи T_{j_s} в момент времени t ; $B_s(t)$ — количество процессоров, одновременно обращающихся к ресурсу R_s ; L — ограничение одновременного доступа к ресурсу.

Работы J_j , для которых возможна декомпозиция на задачи T_{ji} , называются агрегированными. Время начала выполнения агрегированной работы определено как $s_j = \min_i s_{ji}$, время её завершения как $C_j = \max_i C_{ji}$. Тогда нелинейный критерий оптимальности $C_a = \sum_{ij} \max_i C_{ji} \rightarrow \min$ имеет смысл минимизации суммарного времени завершения агрегированных работ. Наряду с данным критерием оптимальности в исследовании рассматривается также использование следующих критериев оптимальности: C_{\max} — минимальная длина расписания; $\sum_j \sum_k C_{jk}$ — сумма времён завершения работ.

Таким образом, приходим к следующим формулировкам задач поиска оптимального расписания в СА в зависимости от выбранного критерия оптимальности C_O : $\langle P_m \mid p_{ij} = \mu_{i_k} V_{j_{i_k}}, \text{res}S..L\mu_s \mid C_O \rangle$.

Для практически значимого случая, когда один ресурс может быть одновременно доступен не более, чем одному прибору, т. е. $L = 1$, исследована область определения допустимого расписания в следующей теореме.

Теорема 1. Пусть для модели (1) $L = 1, m = S$ и выполнено

$$\nexists x \in N : \left(\frac{\mu_s}{\mu_{s'}} \right)^\kappa, \quad s \neq s', \quad s, s' \in 1, \dots, S.$$

где $\kappa \in \{-1, 1\}$, времена готовности $r_{ij} = 0, \forall i, j$, а ресурсоёмкость задач задана бинарной матрицей V размерности $n \times s$. Тогда допустимыми будут являться только те расписания, в которых на любой прибор P_l могут быть назначены только задачи из множества $T_l = \{T_{j_s} : s = l, \forall j = 1, \dots, n\}, l = 1, \dots, m$.

Из доказанной теоремы непосредственно вытекает

Следствие. Пусть выполнены все условия теоремы, но $m > S$. Тогда при $L = 1$ будут простаивать $m - S$ приборов, которое показывает, что невозможно построить эффективное (т. е., обеспечивающее использование приборов без простоя) расписание в случае $m > S$. В связи с этим далее в работе поиск алгоритмов составления эффективных расписаний ведётся для $m \leq S$.

В информационных системах, помимо ресурсов с ограничением одновременного доступа встречаются также ресурсы с практически неограниченным параллельным доступом и возможностью разделения ресурса между несколькими приборами (т. е. клиентами) в произвольной пропорции. Таким образом, в дополнение к дискретной модели (1) вводятся обозначения: t — непрерывное время модели, $t \in [0; +\infty)$; X_{ji} — объёмы задач; $x_{ji}(t)$ — объёмы задач, выполненные к моменту времени t ; $v_{ji}(t)$ — скорость выполнения задачи; $u_{ji}(t)$ — объём ресурса, используемый для выполнения задачи в момент времени t ; U — доступный объём ресурса (полностью возобновляемый ресурс). Предполагается, что $\tilde{v}_i(u_i(t))$ — строго возрастающая функция, т. е. увеличение объёма используемого ресурса всегда ускоряет выполнение заданий. Считается, что $\forall i : v_i(0) = 0$ — т. е. без нужного ресурса обслуживание задачи невозможно.

Таким образом, формулируется дискретно-непрерывная модель многолинейной системы с ограниченными восполнимыми ресурсами для СА:

$$\left\{ \begin{array}{l} C_a = \sum_{j=1}^n \max_{k=1, \dots, s} C_{jk} \rightarrow \min; \quad C_{ji} = s_{ji} + p_{ji}; \\ s_{ji} \geq r_{ji}; \quad s_j = \min_{i=1, \dots, k} s_{ji}; \quad r_{ji} = r_j, \forall i; \\ v_{ji}(t) = \frac{dx_{ji}(t)}{dt} = \tilde{v}_{ji}(u_{ji}(t)); \quad X_{ji} = \int_{s_{ji}}^{C_{ji}} \tilde{v}_{ji}(u_{ji}(t)) dt; \\ \sum_{j=1}^n \sum_{i=1}^S u_{ji}(t) \leq U; \quad m \in \mathbb{N}, j \in \overline{1, n}, k \in \overline{1, S}, i \in \overline{1, k}. \end{array} \right. \quad (2)$$

Анализ модели потребовал сравнить время выполнения задач при различных вариантах выделения ресурса с течением времени. Условие такого сравнения показано в **Лемме 1**: если объёмы задач равны

$$\int_{C_0}^C \tilde{v}(u(t)) dt = \int_{C_0}^C \tilde{v}(u'(t)) dt, \quad u \in [0, +\infty),$$

а зависимость скорости выполнения задачи от объёма выделенного ресурса $\tilde{v}(u)$ неотрицательна и монотонно возрастает, то для выделения ресурса $u(t) \leq u'(t)$ верно соотношение времён завершения задач $C > C'$.

На основе Леммы 1 доказана Лемма 2, утверждающая, что выгодно выбирать такое использование ресурсов, что $u_i(t) = U$.

Лемма 2. Для расписания, минимизирующего значения критериев $\sum C_i$ и C_{\max} , в любой момент времени t выполняется $\sum_i u_i(t) = U$.

Леммы 1–2 используются в доказательстве теорем 2–4.

В рамках модели (2) рассмотрен случай параллельной многозадачной системы с неограниченным числом приборов, единственным ресурсом и неограниченным параллельным доступом к нему. Полученные результаты расширены на случаи ограниченного числа приборов, более чем одного ресурса и декомпозиции работ на задачи.

Рассмотрены три случая для зависимости скорости выполнения задачи от объёма выделенного ресурса: когда $\tilde{v}_i(u_i(t))$ — линейная функция, т. е. $v_i(t) = a_i i(t)$, когда $\tilde{v}_i(u_i(t))$ — выпуклая, и когда $\tilde{v}_i(u_i(t))$ — вогнутая функция.

Для линейного случая доказана **Теорема 2**: минимальные значения критериев оптимальности

$$C_{\max} = X_i/U; \quad \sum_i C_i = (n + 1 - i)X_i/U$$

достигаются при строго последовательном выполнении задач в порядке $X_1 < X_2 < \dots < X_n$, т. е. при использовании правила диспетчеризации SPT (кратчайшее время выполнения).

Для вогнутого случая доказана **Теорема 3**: минимальные значения критериев оптимальности достигаются при строго последовательном выполнении задач в порядке возрастания времени выполнения.

Для случая выпуклых функций в **Теореме 4** показано, что наилучшие значения критериев оптимальности C_{\max} и $\sum_i C_i$ достигаются при резервировании ресурсов, зависящем от формы функций $\tilde{v}_i(u_i)$ для конкретной задачи.

Полученные результаты обобщены на случай ограниченного числа приборов, более чем одного ресурса и декомпозиции работ на задачи, в результате дискретно-непрерывная модель (2) для линейного и вогнутого случая сводится к рассмотренной ранее модели (1)

с дискретными ресурсами, и оптимальный порядок выполнения задач и обращения к ресурсам определяется на основе данной модели.

В **третьей главе** описываются и исследуются предлагаемые алгоритмы построения расписаний обслуживания задач в многоприборных системах с ограниченными воспользуемыми ресурсами.

Принимая во внимание декомпозицию на задачи, приоритето-порождающие функционалы и эвристическую догадку о том, что упорядочивание по времени возможно не только между работами, но и между задачами, строится следующая совокупность алгоритмов, адаптированных для следующих случаев:

- Алгоритм численного решения задач, основанный на функционалах 1-приоритета, для случая выделенных приборов (АППФ1-ВП);
 - Эвристический алгоритм двухуровневой диспетчеризации для случая выделенных приборов (АДД-ВП);
 - Алгоритм численного решения задач, основанный на функционалах 1-приоритета, для случая независимых приборов, количество которых меньше числа ресурсов (АППФ1-НП);
 - Эвристический алгоритм двухуровневой диспетчеризации для случая независимых приборов, количество которых меньше числа ресурсов (АДД-НП);
- и приводятся примеры их использования. Производится асимптотический анализ предела худшего времени выполнения алгоритмов, доказываются теоремы о значениях данного предела, представленных в таблице 1.

Таблица 1 — Асимптотический анализ времени выполнения для алгоритмов, разработанных в рамках диссертационного исследования

Алгоритмы		Разновидности архитектуры СА	
		Приборы	
		Выделенные	Независимые
		Соотношение приборов к ресурсам	
Принцип	Характеристики	$m = S$	$m < S$
ППФ 1-го порядка	Название	АППФ1-ВП	АППФ1-НП
	$T(n, S)$	$O(nS \log nS)$	$O(n^2 S \log S)$
Двухуровневой диспетчеризации	Название	АДД-ВП	АДД-НП
	$T(n, S)$	$O(nS \log nS)$	$O(nS \log nS)$

Асимптотическая сложность $T(n, S)$ времени выполнения разработанных алгоритмов была доказана в четырёх теоремах, приведенных в диссертационной работе.

В **четвёртой главе** приводится описание разработанных комплексов программ и результаты численных экспериментов, выполненных с их помощью.

В рамках данной главы была выполнена разработка следующего программного обеспечения (ПО):

1. Система имитационного моделирования (СИМ) СА, предназначенная для моделирования поведения СА с различной архитектурой и параметрами работы. Описаны разработка логической структуры СИМ на основе моделей «Чёрный ящик» и «Белый ящик»; построение моделирующего алгоритма для проведения направленных численных экспериментов, модульная и файловая структура СИМ, показана диаграмма потока данных и интерфейс запуска, приведены сведения об эксплуатации программы. Программа реализована на языке программирования Python3 с использованием библиотек statistics, pandas, logging, numpy, weakref, json, copy, locale, random, subprocess, multiprocessing, argparse, csv,

matplotlib в виде консольного приложения для запуска в UNIX подобных операционных системах. Диаграмма потока данных при взаимодействии модулей СИМ СА приведена на рисунке 1.

2. Программный комплекс для проведения направленных численных экспериментов по составлению расписаний в СА. Описаны разработка структуры и программной архитектуры комплекса программ, моделирующие алгоритмы, интерфейсы взаимодействия и эксплуатация, а также порядок проведения численных экспериментов.

Данное ПО написано на Python3 с использованием библиотек math, copy, weakref, locale, random, multiprocessing, subprocess, csv, matplotlib и numpy. Эксплуатация возможна в ОС Linux, Windows 10, 8 либо 7, MacOS 10.0.4 и установленным интерпретатором языка Python 3.6 или новее.

3. Комплекс программ для агрегированной обработки распределенных смешанных данных в центре анализа и обработки данных астрофизики частиц GRADLCI (German-Russian Data Life Cycle Initiative). Дата-центр GRADLCI является практическим примером СА, в котором под дополнительными ресурсами, необходимыми для выполнения пользовательских задач понимаются данные экспериментов астрофизики частиц KASCADE, KASCADE-GRANDE], LOPES, Tunka-133, Tunka-Rex, Maked-Ani, находящиеся в удалённых хранилищах данных. Оптимизация обслуживания пользовательских запросов выполнена через агрегированный поиск с использованием базы метаданных MDDB.

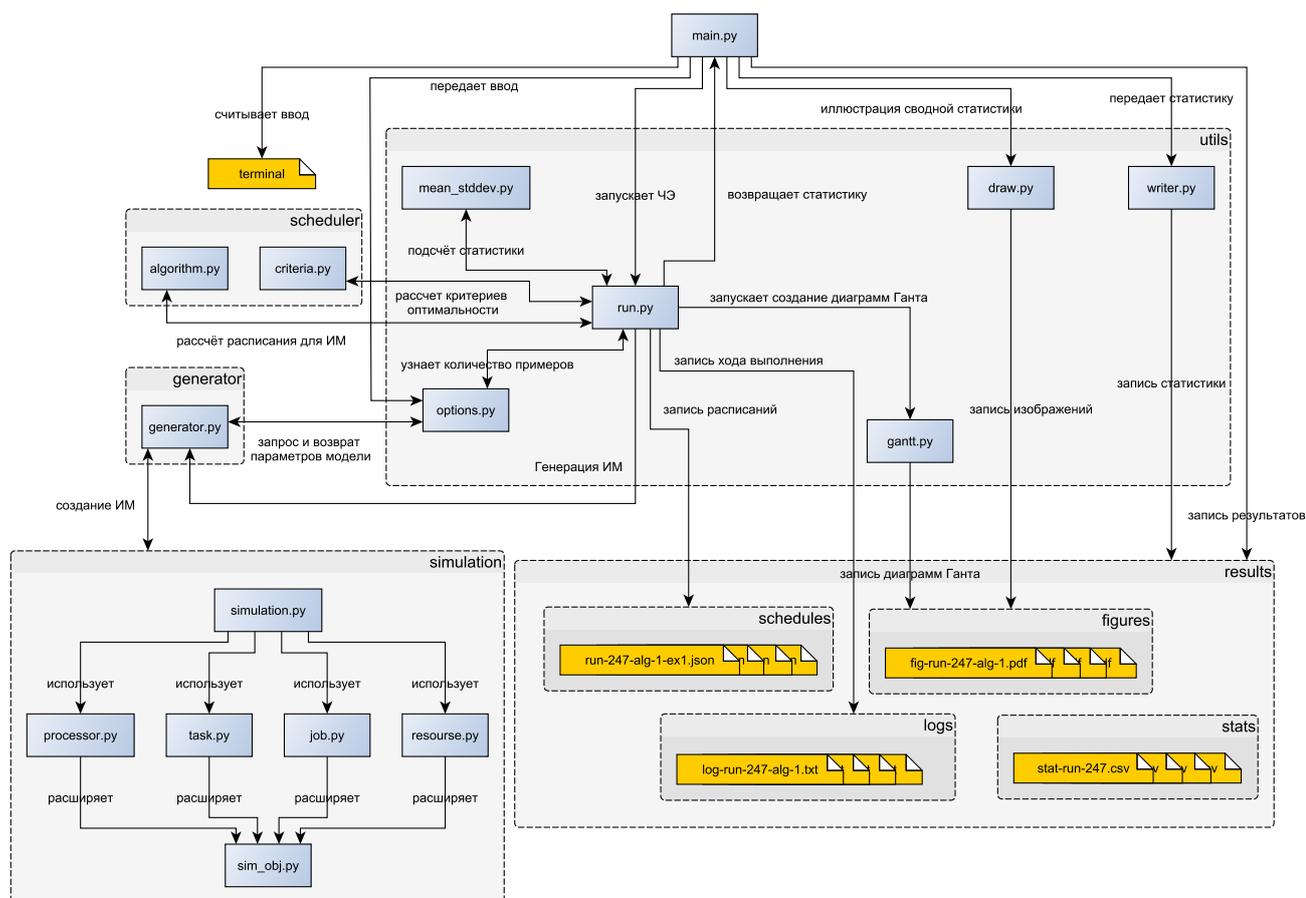


Рисунок 1 — Диаграмма потока данных при взаимодействии модулей системы имитационного моделирования.

Эффективность работы данных программных комплексов определена использованием параллельных асинхронных программных вызовов, реализованных с использованием библиотек subprocess и multiprocessing, а также за счёт использования паттернов парал-

лельного программирования. Использование параллельного программирования позволило оптимизированное выполнение высоконагруженных вычислений до 73% по сравнению с однопоточной версией и добиться более эффективного использования оперативной памяти системы.

С помощью разработанных программ были проведены численные эксперименты:

1. Численный эксперимент по применению приоритетно-порождающих функционалов первого порядка для построения расписаний в многолинейной СА с выделенными приборами, количество которых равно количеству ресурсов.
2. Численный эксперимент по созданию имитационной модели СА, реализующей эвристические алгоритмы двухуровневых перестановок для случая выделенных и независимых приборов и произвольного соотношения между числом приборов и числом ресурсов, и позволяющей сравнить их с алгоритмами, основанными на функционалах 1-приоритета.

Результаты первого численного эксперимента показывают, что на то, какое правило диспетчеризации окажется наиболее выгодной для распределения задач в системе, влияет как выбор критерия оптимальности, так и выбор количества ресурсов и загрузки системы. При этом, эвристика STPT оказалась среди «лучших» в 0.93% случаев (для некоторых из этих случаев близкий к ней результат показывают FIFO и SPT). Это может служить рекомендацией к её использованию для решения данных классов задач. Примеры результатов первого численного эксперимента приведены на рисунке 2.

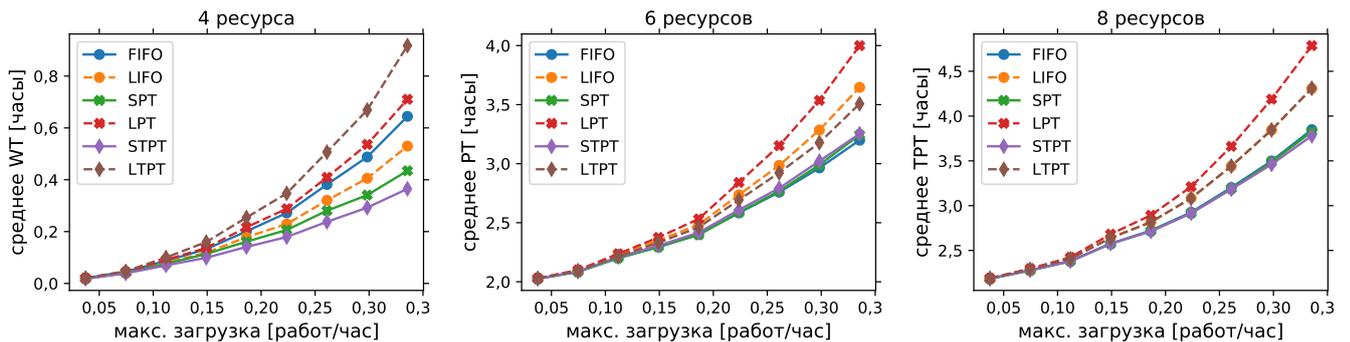


Рисунок 2 — Примеры результатов моделирования времён-критериев оптимальности расписания для различных правил определения приоритета, количества ресурсов и средних частот запросов.

Результаты второго численного эксперимента, сравнивающие эффективность алгоритма двухуровневых перестановок с алгоритмами, основанными на функционалах 1-приоритета, показывают, что из функционалов 1-приоритета по критерию $\sum_{ji} C_{ji}$ стабильно худшим является LPT, лучшим — SPT. При этом лучший по критерию C_a — STPT, худший — LTPT. Критерий C_{\max} оказывается практически нечувствительным к выбору функционала. Что касается алгоритма двухуровневых перестановок, он превосходит алгоритмы, основанные на функционалах 1-приоритета, почти при всех параметрах модели. При этом особенно ярко его преимущества проявляются при условии, что число приборов m меньше числа ресурсов S . Примеры результатов второго численного эксперимента приведены на рисунке 3.

Если $m = S$, то с ростом числа работ алгоритмы SPT и STPT начинают работать так же хорошо, как алгоритм двухуровневых перестановок, а SPT даже может превосходить его по критерию $\sum_{ji} C_{ji}$ на 2–3%.

Пример расписания, построенного с использованием алгоритма двухуровневых перестановок, приведён на рисунке 4.

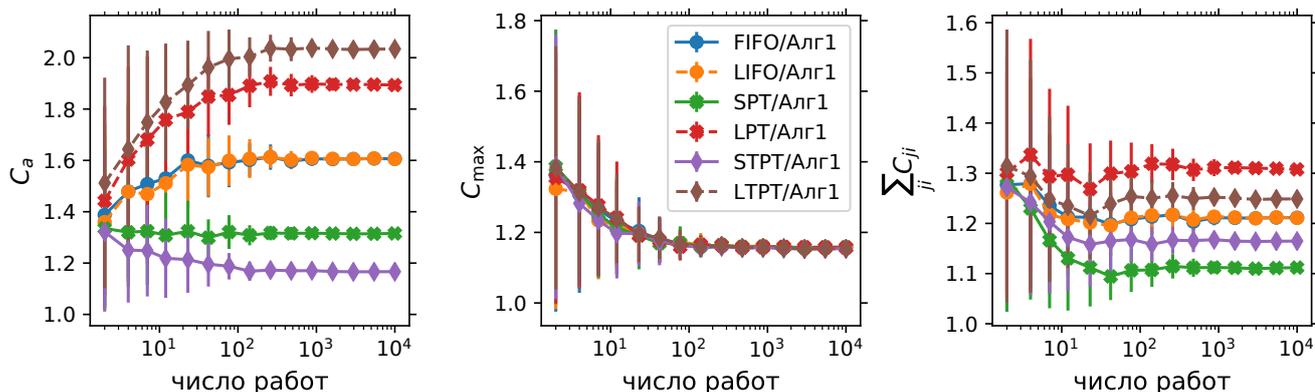


Рисунок 3 — Примеры отношений критериев оптимальности для различных правил диспетчеризации из класса приоритето-порождающих функционалов 1-го порядка к критериям оптимальности для алгоритма двухуровневых перестановок, для 3 процессоров, 4 ресурсов, до 4 задач в работе.

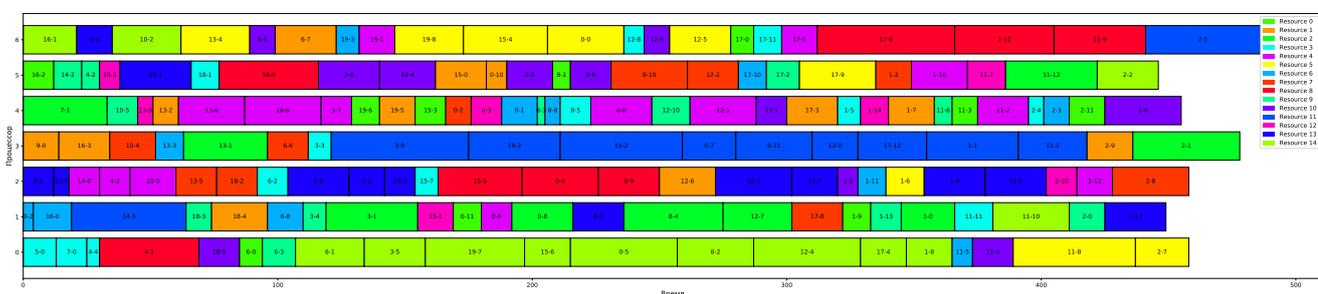


Рисунок 4 — Пример диаграммы Ганта, полученной программой для случая решения примера с параметрами $n = 20$, $m = 7$, $S = 15$ алгоритмом двухуровневых перестановок.

В рамках второго численного эксперимента была также проведена оценка временной сложности алгоритма двухуровневых перестановок для среднего случая выполнения.

В Заключение диссертации сформулированы следующие выводы и результаты:

1. Осуществлен анализ проблематики моделирования ресурсных ограничений в современных информационно-вычислительных системах. Выделено понятие информационного ресурса и категория «качественная доступность» ресурсных ограничений. Выделен и описан класс распределённых информационно-вычислительных систем «системы агрегации», отличительной чертой которых является наличие доступа к ряду распределённых информационных ресурсов, используемых для решения задачи агрегированного поиска. Сформулированы определения допустимости, эффективности и оптимальности расписаний, поставлена оптимизационная задача построения расписаний в системах агрегации с ресурсными ограничениями качественной доступности.
2. Разработан комплекс математических моделей составления расписаний для нескольких приборов для случаев дискретных и дискретно-непрерывных ограничений качественного доступа к удалённым информационным ресурсам.
3. Проведено математическое исследование существования аналитического решения для случаев дискретных ограничений на ресурсы и различных видов функции расхода ресурса для модели с непрерывными ограничениями на ресурсы, использовано для обоснования необходимости разработки численных алгоритмов решения.
4. Разработаны новые алгоритмы численного решения задач составления расписаний в многоприборных системах с ресурсами, обладающими ограниченной качественной доступностью, основанные на приоритето-порождающих функционалах для случаев числа приборов равного числу ресурсов и числа приборов меньшего числа ресурсов;

5. Разработаны комплексы программ для направленного численного эксперимента по исследованию свойств разработанных алгоритмов, имитационного моделирования поведения СА при различных заданных параметрах;
6. Разработанные численные методы и алгоритмы реализованы в виде комплекса программ для работы с пользовательскими заявками и диспетчеризации задач в центре сбора и анализа данных экспериментальной астрофизики частиц GRADLCI.

Полученные в диссертационной работе результаты были **приняты к использованию** в центре анализа и обработки данных GRADLCI для составления расписаний обработки пользовательских запросов и подтвердили свою практическую целесообразность.

Публикации автора по теме диссертации

1. *Tokareva V.* Optimization of request processing times for a heterogeneous data aggregation platform / V. Tokareva // Journal of Physics: Conference Series. — 2021. — Vol. 1740, no. 1. — P. 012058. — DOI: [10.1088/1742-6596/1740/1/012058](https://doi.org/10.1088/1742-6596/1740/1/012058).
2. *Tokareva V.* Towards a coherent Data Life Cycle in Astroparticle Physics / V. Tokareva, A. Haungs, D. Kang, D. Kostunin, F. Polgart, D. Wochele, J. Wochele // Journal of Physics: Conference Series. — 2020. — Vol. 1525, no. 1. — P. 012070. — DOI: [10.1088/1742-6596/1525/1/012070](https://doi.org/10.1088/1742-6596/1525/1/012070).
3. *Tokareva V.* Data Aggregation Platform for Experiments of Astroparticle Physics / V. Tokareva, A. Haungs, D. Kang, F. Polgart, D. Wochele, J. Wochele // CEUR Workshop Proceedings. — 2020. — Vol. 2679. — P. 134–142.
4. *Wochele D.* Data structure adaption from large-scale experiment for public re-use / D. Wochele, J. Wochele, F. Polgart, V. Tokareva, D. Kang, A. Haungs // CEUR Workshop Proceedings. — 2019. — Vol. 2406. — P. 114–121.
5. *Tokareva V.* Development of a data infrastructure for a global data and analysis center in astroparticle physics / V. Tokareva, A. Haungs, D. Kang, D. Kostunin, F. Polgart, D. Wochele, J. Wochele // CEUR Workshop Proceedings. — 2019. — Vol. 2406. — P. 106–113.
6. *Haungs A.* German-Russian Astroparticle Data Life Cycle Initiative / A. Haungs, I. Bychkov, J. Dubenskaya, O. Fedorov, A. Heiss, D. Kang, Y. Kazarina, E. E. Korosteleva, D. Kostunin, A. Kryukov, A. Mikhailov, M.-D. Nguyen, F. Polgart, S. Polyakov, E. Postnikov, A. Shigarov, D. Shipilov, A. Streit, V. Tokareva, D. Wochele, J. Wochele, D. Zhurov // Proceedings of Science. — 2019. — Vol. 358. — P. 284. — DOI: [10.22323/1.358.0284](https://doi.org/10.22323/1.358.0284).
7. *Bychkov I.* Russian–German astroparticle data life cycle initiative / I. Bychkov, A. Demichev, J. Dubenskaya, O. Fedorov, A. Haungs, A. Heiss, D. Kang, Y. Kazarina, E. Korosteleva, D. Kostunin, A. Kryukov, A. Mikhailov, M.-D. Nguyen, S. Polyakov, E. Postnikov, A. Shigarov, D. Shipilov, A. Streit, V. Tokareva, D. Wochele, J. Wochele, D. Zhurov // Data. — 2018. — Vol. 3, no. 4. — P. 56. — DOI: [10.3390/data3040056](https://doi.org/10.3390/data3040056).
8. *Tokareva V. A.* Current status of data center for cosmic rays based on KCDC / V. A. Tokareva, D. G. Kostunin, A. Haungs // CEUR Workshop Proceedings. — 2018. — Vol. 2267. — P. 405–409.
9. *Tokareva V.* German-Russian Astroparticle Data Life Cycle Initiative to foster Big Data Infrastructure for Multi-Messenger Astronomy / V. Tokareva, I. Bychkov, A. Demichev, J. Dubenskaya, O. Fedorov, A. Haungs, D. Kang, Y. Kazarina, E. Korosteleva, D. Kostunin, A. Kryukov, A. Mikhailov, M.-D. Nguyen, F. Polgart, S. Polyakov, E. Postnikov, A. Shigarov, D. Shipilov, A. Streit, D. Wochele, J. Wochele, D. Zhurov // Proceedings of Science. — 2021. — Vol. 395. — P. 938. — DOI: [10.22323/1.395.0938](https://doi.org/10.22323/1.395.0938).

10. Tokareva V. Optimization of aggregated requests scheduling in a system with non-separable resources and parallel data processing / V. Tokareva // AIP Conference Proceedings. Vol. 2377. — AIP Publishing LLC. 2021. — P. 040009. — DOI: [10.1063/5.0063574](https://doi.org/10.1063/5.0063574).
11. Tokareva V. A. Schedules with Priorities for Online Resource Management Problems in Aggregated Data Access Systems / V. A. Tokareva // Automation and Remote Control. — 2021. — Vol. 82, no. 11. — P. 1939–1948. — DOI: [10.1134/S0005117921110096](https://doi.org/10.1134/S0005117921110096).

Свидетельства о государственной регистрации программ для ЭВМ

12. Многопроцессорное клиент-серверное приложение с механизмом адаптивного составления расписаний для обработки информационных запросов в распределенных хранилищах данных : Свидетельство о государственной регистрации программы для ЭВМ № 2021680467 / В. А. Токарева, Ю. В. Бондаренко. — № 2021680026 ; заявл. 06.12.2021 ; опублик. 10.12.2021 (Рос. Федерация).
13. Программный комплекс для моделирования процесса составления адаптивных расписаний для многолинейной системы с ограниченными восполнимыми ресурсами : Свидетельство о государственной регистрации программы для ЭВМ № 2021681570 / В. А. Токарева. — № 2021668937 ; заявл. 23.11.2021 ; опублик. 23.12.2021 (Рос. Федерация).
14. Система имитационного моделирования для исследования проблем составления расписаний в распределенных системах с ограниченными ресурсами : Свидетельство о государственной регистрации программы для ЭВМ № 2022610178 / В. А. Токарева. — № 2021681892 ; заявл. 24.12.2021 ; опублик. 10.01.2022 (Рос. Федерация).

Статьи и материалы конференций

15. Токарева В. А. Расписания с приоритетами для задач онлайн-управления ресурсами в системе агрегированного доступа к данным / В. А. Токарева // Автоматика и телемеханика. — 2021. — № 11. — С. 135—147. — DOI: [10.31857/S000523102111009X](https://doi.org/10.31857/S000523102111009X).