

На правах рукописи



Мезенцев Юрий Анатольевич

**ЭФФЕКТИВНЫЕ ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ РЕШЕНИЯ
ДИСКРЕТНЫХ ЗАДАЧ ОПТИМИЗАЦИИ УПРАВЛЕНИЯ
ПРОИЗВОДСТВЕННЫМИ ПРОЦЕССАМИ**

Специальность: 05.13.01 – Системный анализ, управление и обработка
информации (промышленность)

А В Т О Р Е Ф Е Р А Т
диссертации на соискание ученой степени
доктора технических наук

Новосибирск – 2013

Работа выполнена в Федеральном государственном бюджетном образовательном учреждении высшего профессионального образования «Новосибирский государственный технический университет»

Научный консультант: доктор технических наук, профессор,
Авдеенко Татьяна Владимировна

Официальные оппоненты: Фроловский Владимир Дмитриевич, доктор технических наук, профессор, Новосибирский государственный технический университет, профессор кафедры автоматизированных систем управления

Окольнишников Виктор Васильевич, доктор технических наук, Конструкторско-технологический институт вычислительной техники СО РАН, ведущий научный сотрудник лаборатории информационных систем

Пятковский Олег Иванович, доктор технических наук, профессор, Алтайский государственный технический университет им. И.И. Ползунова, заведующий кафедрой информационных систем в экономике

Ведущая организация: Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Национальный исследовательский Томский политехнический университет»

Защита состоится « 5 » ноября 2013 г. в 10.00 часов на заседании совета по защите докторских и кандидатских диссертаций Д 212.173.05 при Новосибирском государственном техническом университете по адресу 630073, г. Новосибирск, пр. К. Маркса, д. 20.

С диссертацией можно ознакомиться в научно-технической библиотеке Новосибирского государственного технического университета.

Автореферат разослан « _____ » октября 2013 г.

Ученый секретарь
диссертационного совета,
д.т.н., профессор



Юркевич Валерий
Дмитриевич

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы и цель работы

История развития современного аппарата дискретной оптимизации (ДО) началась с работ Балаша, Беллмана, Бендерса, Данцига, Гомори, Канторовича, Ху, ряда других авторов (конец 50-х, начало 60-х годов XX века). Позднее появилась теория вычислительной сложности задач дискретной оптимизации, и выделились основные направления развития алгоритмов их решения. Определелись, ставшие классическими, сферы приложений: дискретные задачи управления производством (календарное планирование), дискретные задачи оптимального проектирования (размещение, комплектация), дискретные задачи управления логистикой (раскрой и упаковка, транспортировка, выбор поставщиков, ассортимента, определение маршрутов, управление запасами). Список можно существенно расширить. В настоящее время теоретические исследования в большинстве своем направлены на создание эффективных алгоритмов приближенного решения отдельных подклассов задач ДО с хорошими априорными либо апостериорными оценками точности. Под вычислительной эффективностью (или просто эффективностью) алгоритма понимается полиномиальная трудоемкость поиска решения задач ДО, что для практических применений зачастую важнее точности.

Достигнутые к настоящему времени, как теоретические, так и практические результаты едва ли можно считать удовлетворительными. Несмотря на взрывной рост доступности и возможностей вычислительной техники, на практике методы дискретной оптимизации в перечисленных областях применяются крайне недостаточно. Основной причиной является вычислительная трудоемкость задач дискретного программирования, к каковым сводится большинство дискретных задач управления. Значительная их часть NP-трудна в сильном смысле, а реализации для практических нужд имеют большие размерности.

Поэтому для науки актуальны исследования в области *дискретной* оптимизации, а для производства - разработки систем поддержки принятия решений на этой основе.

Теоретическая часть диссертации посвящена разработке эффективных алгоритмов решения общей задачи ДО. Прикладная часть ориентирована на традиционные и наиболее актуальные сферы применения моделей и методов дискретной оптимизации. Это важнейшие для практики разделы теории расписаний, используемые в управлении производственно-технологическими процессами. В данном случае в качестве абстрактных объектов исследования рассматриваются параллельные обслуживающие системы с возможными задержками начала работы приборов и поступления заявок, а также многостадийные параллельно-последовательные системы. Другой актуальной сферой применения аппарата дискретной оптимизации, выбранной в качестве объекта исследований и разработок в диссертации, явились процессы управления взаимодействием предприятия с внешней средой. Наконец, еще один прикладной раздел посвящен задачам выбора оборудования и комплектации технических и технологических систем составляющими подсистемами.

Следует отметить не только самостоятельную значимость, но и явную системную связь перечисленных направлений исследований. Действительно, исследуемые модели теории расписаний описывают технологические (внутренние) процессы промышленного производства, строительства, либо любой другой продуктивной деятельности. Модели, которые в диссертации достаточно условно (по преобладающей предметной области) именуется математическими моделями логистики, отображают входные и выходные (внешние) материальные потоки. Таким образом, на основе выбранных направлений, объекты управления рассматриваются системно, во взаимосвязи внешних и внутренних факторов производственной деятельности.

В работе над диссертацией использованы как классические труды в области дискретной оптимизации и смежных разделах, так и работы современных авторов, теоретиков и практиков. В частности, главы, связанные с теорией расписаний, опираются на труды Гимади Э.Х., Гордона В.С., Jonson S.M., Загидуллина Р.Р., Конвей Р.В., Коффмана Э.Г., Левина В.И., Максвелла В.Л., Миллера Л.В., Мирносецкого Н.Б., Севастьянова С.В., Сотскова Ю.Н., Струевича В.А., Танаева В.С., Фролова Е.Б., Шафранского Я.М., Шкурбы В.В. При разработке алгоритмов решения нелинейных оценочных задач использованы результаты Boyd S., Евтушенко Ю.Г., Немировского А.С., Нестерова Ю.Е., Vandenberghe L., Vanderbei, R. J. Алгоритмы, основанные на лагранжевой декомпозиции, либо на неполной декомпозиции опираются на работы Бахтина А.Е., Гилл Ф., Гольштейна Е.Г., Gondzio J., Лэдсона Л.С., Малкова У.Х., Муртаф Б., Мюррей У., Райт М. При создании новых методов и алгоритмов решения общей задачи линейного программирования с булевыми переменными учтены достижения в этой области Алексева О.Г., Гомори Р.Е., Емеличева В.А., Забиняко Г.И., Заславского А.А., Колоколова А.А., Комлик В.И., Korte В., Корбут А.А., Кочетова Ю.А., Лебедева С.С., Леонтьева В. К., Лихтенштейна В.Е., Михалевича В.С., Пападимитриу Х., Сергиенко И.В., Стайглиц К., Ху Т., Vygen J. Работы Дикина И.И., Frisch R.A., Gonzio J., Евтушенко Ю.Г., Жадан В.Г., Зоркальцева В.И., Jansen В., Karmarkar N. К., Немировского А.С. Нестерова Ю.Е., Ye Y., Roos С., Terlaky T., Todd M.J., Vial J.- P. послужили основой для разработки эффективной модификации алгоритма следования центральному пути, применяемого для решения последовательности оценочных подзадач в алгоритмах дискретной оптимизации.

Фундаментальная научная проблема на решение которой направлены исследования и разработки, осуществленные в рамках диссертационной работы, - создание вычислительно эффективных методов решения дискретных задач оптимизации. Экспертные оценки улучшения значений критериев эффективности при достижении оптимумов в сравнении с имеющимися на сегодняшний день результатами разнятся для разных сфер приложения методов дискретной оптимизации. Однако в среднем эти значения оцениваются величиной не менее чем 20%. Фактически это оценка скрытого резерва повышения эффективности управления при разрешении упомянутой фундаментальной научной проблемы. Полученные в прикладных разделах диссертации при решении тестовых задач апостериорные оценки отклонения решений от оптимумов в большинстве сво-

ем попадают в интервал 0-5%, что может служить предварительной количественной оценкой степени достижения поставленной цели.

Актуальность работы обусловливается отсутствием вычислительно эффективных методов *точного* решения общей задачи *дискретной* оптимизации, отсутствием эффективных методов *точного* решения наиболее ценных с точки зрения практики частных *дискретных* задач управления, недостаточным развитием средств быстрого приближенного решения с приемлемыми апостериорными оценками близости к оптимумам.

Опираясь на вышеизложенное, можно сформулировать цель работы и задачи исследований, степень успешности решения которых определяет её теоретическую и практическую ценность.

Цель работы состоит в разработке и исследовании вычислительно эффективных методов решения дискретных задач оптимизации управления производственными процессами промышленных предприятий.

Задачи исследований

Поставленная цель достигается решением следующих основных задач.

1. Формализация и разработка алгоритмов решения важнейших задач теории расписаний для промышленных предприятий, включая общую задачу оптимизации расписаний параллельно-последовательных систем.

2. Постановка задач оптимизации входных и выходных материальных потоков предприятия. Сведение к задачам дискретного программирования и разработка эффективных алгоритмов их решения.

3. Разработка вычислительно эффективного алгоритма решения задачи оптимальной комплектации технических систем подсистемами на основе линеаризации условий и лагранжевой декомпозиции.

4. Развитие методов отсечений решения общей задачи дискретной оптимизации. Разработка метода бинарных отсечений решения задачи линейного программирования с булевыми переменными и распространение результатов на случай общей задачи частично-целочисленного программирования.

5. Разработка алгоритмов и программ решения последовательностей оценочных задач линейного и полуопределенного программирования, порождаемых методами дискретной оптимизации при решении задач исследуемых классов. Модификация компонент алгоритма следования центральному пути для повышения его быстродействия.

Область исследования

Содержание диссертации соответствует положениям п.п. 1,2,3,4,5 областей исследований паспорта специальности 05.13.01.

Объект и предмет исследования

Объектом исследования являются производственно-технологические системы и процессы, математическое описание проблем управления которыми может быть представлено в виде задач дискретной оптимизации. При этом

ограничивающие и целевые условия задаются в виде алгебраических уравнений и неравенств. *Предмет* исследования – дискретные задачи оптимального управления этими системами и процессами.

Методы исследования

Основные теоретические результаты работы базируются на точных и приближённых методах системного анализа и исследования операций, непрерывной и дискретной оптимизации, включая барьерно-ньютоновские методы, методы декомпозиции, линеаризации, ветвлений и отсечений, динамическое программирование, а также ряд эвристических алгоритмов.

Эффективность всех разработанных вычислительных методов решения поставленных задач дискретной оптимизации проверена экспериментально.

Научная новизна

Научная новизна работы состоит в том, что впервые были получены следующие результаты.

1. Осуществлены содержательная постановка и формальное описание совокупности задач оптимизации управления производством, отображающих, как внутренние производственно-технологические процессы, так и процессы взаимодействия с внешней средой. Все сформулированные задачи принадлежат классу NP-трудных задач ДО.

2. Разработаны вычислительно эффективные методы решения поставленных задач ДО, использующие упрощающую редукцию и декомпозиционные схемы: неполную декомпозицию, совмещающую разбиение на подзадачи меньшей размерности с адаптивным сужением области поиска и отсевом бесперспективных вариантов; сочетание лагранжевой декомпозиции с линеаризацией и параметрическим анализом последовательностей подзадач; динамическое программирование в сочетании с декомпозицией и жадным алгоритмом. При решении последовательностей подзадач использованы вычислительные схемы ветвей и границ, ветвей и отсечений.

3. Предложен новый вычислительный метод и созданы алгоритмы решения общей задачи булева программирования, основанные на процедуре конструирования бинарных отсечений. Разработаны эвристические процедуры построения бинарных отсечений, позволившие реализовать быстрые алгоритмы решения задач линейного программирования с булевыми переменными.

4. Разработан вычислительный метод решения последовательностей подзадач линейного и полуопределенного программирования (порождаемых методами решения задач ДО), являющийся эффективной модификацией алгоритма следования центральному пути в недопустимой области с длинным шагом. Достигнута универсальность метода относительно используемых барьерных функций, процедур определения длины шага и других параметров.

5. Разработаны быстродействующие варианты прямого и итерационного алгоритмов решения систем линейных алгебраических уравнений специального вида, порождаемых на каждом шаге модифицированного алгоритма следования

центральному пути. Определены параметры, обеспечивающие его наибольшее быстроедействие среди всех возможных реализаций.

Практическая ценность и значимость

Заключаются в возможности использования разработанных методов, алгоритмов и прототипов программного обеспечения для решения прикладных задач управления производственными процессами промышленных предприятий.

Результаты диссертационной работы были использованы производственными компаниями ООО «СофтМастер» и ООО ТК «Маркет Регион», а также ИЭОПП СО РАН при выполнении договора № 2010/08/0134 от 01.08.2010 на тему «Разработка перспективной программы оптимизации издержек ООО «Газпром добыча Надым» на базе инновационных технологий» по заказу ООО «Газпром добыча Надым» (г. Надым), что подтверждается соответствующими актами внедрения разработок. Часть практических результатов получена при поддержке грантов: Минобрнауки, «ТП-8.536.2011, Разработка интеллектуальных технологий, средств компьютерного моделирования и эффективных методов оптимизации, как функционального наполнения информационно-аналитических систем поддержки принятия решений», РГНФ «09-02-00093а, Разработка и исследование моделей планирования производства изделий, основанных на нанотехнологиях».

Получены свидетельства о государственной регистрации программного обеспечения (ПО), разработанного в рамках диссертационных исследований: «Программная реализация прямо-двойственного барьерного алгоритма следования центральному пути для решения задач линейного программирования большой размерности», «Программный комплекс быстрого решения задач оптимальной комплектации систем», «Программный комплекс оптимального планирования закупок и сбыта продукции», «Программный комплекс оптимального управления поставками сырья и комплектующих на предприятии».

Достоверность

Обоснованность и достоверность научных положений и выводов, содержащихся в диссертационной работе, подтверждается вычислительными экспериментами. Получены апостериорные оценки точности и быстрогодействия всех разработанных в рамках диссертационной работы алгоритмов оптимизации. Сгенерированы и просчитаны тестовые задачи в широком диапазоне размерностей. Результаты тестирования программных модулей, реализующих разработанные алгоритмы, проверены с использованием IBM ILOG CPLEX Studio (v:12.1-12.5).

Апробация работы

Основные положения и результаты диссертационной работы докладывались и обсуждались на следующих научно-технических конференциях и симпозиумах: Международной научной конференции «Моделирование 2010» (Киев, Украина, 2010); на международном индо-российском совместном семинаре по

компьютерному интеллекту и современным эвристикам в автоматизации и робототехнике «DST-RFBR Sponsored Indo-Russian Joint Workshop on Computational Intelligence and Modern Heuristics in Automation and Robotics» (Сурат, Индия, 2010); на международном российско-индийском совместном семинаре по компьютерному интеллекту и современным эвристикам в автоматизации и робототехнике «RFBR-DST Sponsored Russian-Indian Joint Workshop on Computational Intelligence and Modern Heuristics in Automation and Robotics» (Новосибирск, Россия, 2011); XV международной открытой конференции «Современные проблемы информатизации в экономике и обеспечении безопасности». (Воронеж, Россия, 2010); Международной научно-практической конференции «Агропромышленный комплекс, как фактор развития национальной экономики республики Казахстан» (Семипалатинск, Казахстан, 2004); VI международной конференции «Актуальные проблемы электронного приборостроения» (Новосибирск, Россия, 2002); Международной научно-технической конференции «Информационные системы и технологии» (Новосибирск, Россия, 2003); X Международной научно-практической конференции «Методы и алгоритмы прикладной математики в технике, медицине и экономике» (Новочеркасск, Россия, 2010).

Публикации

По результатам выполненных в диссертации исследований опубликовано 35 печатных работ, в том числе 2 монографии, 18 статей в журналах из перечня ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание учёной степени доктора наук, 4 статьи в научных периодических журналах и тематических сборниках, 7 докладов в материалах и трудах международных и всероссийских конференций, 4 свидетельства о государственной регистрации программного обеспечения.

Научные положения, выносимые на защиту

1. Постановка задачи синтеза оптимальных по быстродействию расписаний параллельной обслуживающей системы с задержками начала обслуживания заявок, ее редукция в задачу линейного программирования с булевыми переменными больших потенциальных размерностей реализаций и эффективные декомпозиционные алгоритмы поиска оптимума.

2. Формальные постановки задач синтеза оптимальных по быстродействию расписаний многостадийных параллельно-последовательных обслуживающих систем и эффективные декомпозиционные алгоритмы решения.

3. Формальные постановки задач оптимального управления входными и выходными материальными потоками предприятия, содержащие непрерывные и целочисленные переменные, отягощенные квадратичными ограничениями с альтернативными условиями и большими размерностями для реализаций. Итерационные вычислительно эффективные алгоритмы приближенного решения сформулированных задач оптимизации.

4. Эффективный асимптотически точный по размерности алгоритм решения задач оптимальной комплектации систем составляющими элементами, основанный на линеаризации и лагранжевой декомпозиции.

5. Метод бинарных отсечений решения общей задачи линейного программирования с булевыми переменными, эвристические процедуры построения бинарных отсечений, позволившие реализовать быстрые алгоритмы оптимизации.

6. Быстродействующая модификация барьерно-ньютоновского алгоритма следования центральному пути при решении (непрерывных) линейных и квадратичных оценочных подзадач, позволившая обеспечить максимальное быстродействие при реализации алгоритмов дискретной оптимизации.

Личный вклад автора в проведённых исследованиях

Все представленные в диссертации результаты исследований получены лично автором или при его непосредственном участии. В работах, опубликованных в соавторстве, автор внёс определяющий вклад в постановку задач, выбор направлений и методов исследований, аналитические выкладки, анализ и интерпретацию полученных результатов, написание работы.

Структура и объём диссертации

Диссертация состоит из введения, шести глав и заключения, изложенных на 307 страницах основного текста, включая 37 рисунков, 59 таблиц, список использованных источников отечественных и зарубежных авторов из 120 наименований, дополненного приложениями.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Текст диссертации разбит на две части. Первая, «Специальные прикладные задачи дискретной оптимизации, методы моделирования и синтеза управляющих воздействий», состоящая из четырех глав, посвящена исследованиям и разработке ряда прикладных оптимизационных моделей, алгоритмов оптимизации и программных компонент, применяемых для решения дискретных задач оптимизации управления производственными процессами. Общим свойством всех рассматриваемых в первой части задач управления производственными процессами является их принадлежность (либо сводимость) к классу NP-трудных задач ДО. Единым подходом, использованным во всех четырех главах при постановке задач и разработке эффективных алгоритмов их приближенного решения, является применение методов декомпозиции и релаксации. В большинстве случаев разработаны и применены алгоритмы неполной декомпозиции. Это позволило, с одной стороны получить апостериорные оценки быстродействия алгоритмов, подтверждающие их вычислительную эффективность, с другой стороны, обеспечить приемлемую точность решений, что также подтверждается вычислительными экспериментами.

Вторая часть «Быстрые методы решения общей задачи целочисленного программирования» посвящена разработке алгоритмов решения общей задачи дискретной оптимизации и состоит из двух глав (главы 5 и 6). Ее тесная взаи-

мосвязь с материалами первой части диссертации объясняется следующим обстоятельством. Использование методов неполной декомпозиции ограничено потенциальными потерями точности при увеличении глубины декомпозиции. Иными словами разбиение исходных задач на подзадачи меньшей размерности возможно только в разумных пределах и, таким образом, общая проблема эффективности алгоритмов решения некоторых частных задач ДО, рассмотренных в рамках первой части диссертации, ослаблена, но отнюдь не снята. Снята она может быть только в случае построения *эффективных алгоритмов решения общей задачи целочисленного программирования*. Именно разработке этой темы посвящена вторая часть работы.

В первой главе даётся обзор и обобщение работ, посвящённых теоретическим и прикладным аспектам теории расписаний. Сформулированы общие задачи теории расписаний, рассматриваемые в рамках диссертационной работы.

Осуществлена аналитическая постановка задачи оптимизации расписаний параллельной системы с задержками поступления заявок, как задачи дискретной оптимизации. Рассматриваемые объекты (разновидности параллельной обслуживающей системы, или ПОС).

Для каждого прибора i ПОС определено время обслуживания каждой заявки j ($0 \leq t_{i,j} \leq \infty$). Прерываний обслуживания одних заявок в пользу других не допускается, приоритеты заявок и приборов в общем случае не определены, дисциплина обслуживания очередей – произвольная. Показано, что за полиномиальное время ПОС вида «сборка» и «разузлование» приводятся к виду «линейное обслуживание». Поэтому основные исследования в рамках главы сосредоточились на ПОС данной разновидности.

Тогда в общем случае модель, описывающая процесс функционирования ПОС, представима в виде:

$$\sum_{i=1}^I x_{i,j} = 1, \quad j = \overline{1, J}, \quad \underline{b}_i \leq \sum_{j=1}^J x_{i,j} \leq \overline{b}_i, \quad i = \overline{1, I}, \quad (1)$$

$$x_{i,j} = \begin{cases} 1, & \text{если заявка } j \text{ назначается на прибор } i, \\ 0 & \text{в противном случае,} \end{cases} \quad y_{i,j} \geq 0, \quad i = \overline{1, I}, \quad j = \overline{1, J},$$

$$\tau_{i,j} = \tau_j^0 - \sum_{l=1}^{j-1} (\tau_{i,l} + t_{i,l}) x_{i,l}, \quad i = \overline{1, I}, \quad j = \overline{1, J}, \quad (2)$$

$$\widehat{\tau}_{i,j} = \tau_{i,j} + y_{i,j} \geq 0, \quad i = \overline{1, I}, \quad j = \overline{1, J}, \quad (3)$$

$$\sum_{j=1}^J \widehat{\tau}_{i,j} x_{i,j} + \sum_{j=1}^J t_{i,j} x_{i,j} \leq \lambda, \quad i = \overline{1, I}, \quad (4)$$

Здесь τ_j^0 - задержка поступления j -й заявки в ПОС ($T^0 = \|\tau_j^0\|$ - расписание на входе системы), $t_{i,j}$ - время обслуживания заявки j прибором i ($T = \|t_{i,j}\|$), $y_{i,j}$ - компенсирующие переменные, которые вводятся для того, чтобы избежать

появления отрицательных задержек $(\tau_{i,j})$. $\widehat{\tau}_{i,j}$ имеет смысл фактической задержки начала выполнения j -й заявки i -м прибором после завершения обслуживания им предшествующей заявки. Пусть также $\Gamma^1 = \|\tau_j^1\|$ - расписание на выходе системы. Система ограничений может дополняться ресурсными ограничениями вида:

$$\sum_{i=1}^I \sum_{j=1}^J a_{i,j}^r (\widehat{\tau}_{i,j} + t_{i,j}) x_{i,j} \leq b^r, \quad r = \overline{1, R}, \quad (5)$$

где $a_{i,j}^r$ норматив расхода ресурса r -го вида при обслуживании i -м прибором j -й заявки, b^r - лимит r -го ресурса в рассматриваемом временном интервале. Если при этом определен критерий качества (эффективности) расписания, например

$$Z = \lambda \rightarrow \min, \quad (6)$$

то (1)-(6) определяют задачу синтеза оптимальных по быстродействию расписаний параллельной системы с задержками на входе. (4) и (6) реализуют минимаксный критерий, имеющий смысл критерия максимального быстродействия. Кроме (6) определены и некоторые альтернативные критерии.

Условия (2)-(6) содержат рекурсивные функции фактических задержек $\tau_{i,j}$.

Для решения поставленных задач в общем случае требуется раскрытие рекурсий и линеаризации получаемых при этом условий. В четвертом разделе главы эта операция выполнена. В результате получена модель ПОС с задержками поступления заявок следующего вида:

$$\begin{aligned} \sum_{i=1}^I x_{i,j} &= 1, \quad j = \overline{1, J}, & \underline{b}_i &\leq \sum_{j=1}^J x_{i,j} \leq \overline{b}_i, \quad i = \overline{1, I}, \\ x_{i,j} &= \begin{cases} 1, & \text{если заявка } j \text{ закрепляется за прибором } i; \\ 0 & \text{в противном случае;} \end{cases} \\ u_{i,j,k} &= \begin{cases} 1, & \text{при истинности выражения } f_{i,j,k}(x_{i,k}, x_{i,k+1}, \dots, x_{i,j-1}) = x_{i,j} x_{i,k} \prod_{l=k+1}^{j-1} x_{i,l}, \\ 0 & \text{в противном случае;} \end{cases} \\ y_{i,j} &\geq 0, \quad i = \overline{1, I}, \quad j = \overline{1, J}, \quad i = \overline{1, I}, \quad j = \overline{1, J}, \end{aligned} \quad (7)$$

$$-K + 2 \leq x_{i,j} + x_{i,k} - \sum_{l=k+1}^{j-1} x_{i,l} - K u_{i,j,k} \leq 1, \quad K = j - k + 1, \quad i = \overline{1, I}, \quad j = \overline{1, J}, \quad k = \overline{1, J}, \quad (8)$$

$$\begin{aligned} \widehat{\tau}_{i,j} &= y_{i,j} + \tau_j^0 - \sum_{k=1}^{j-1} (\tau_k^0 + t_{i,k}) u_{i,j,k} \geq 0, \quad \text{или} \\ -y_{i,j} + \sum_{k=1}^{j-1} (\tau_k^0 + t_{i,k}) u_{i,j,k} &\leq \tau_j^0, \quad i = \overline{1, I}, \quad j = \overline{1, J}, \end{aligned} \quad (9)$$

$$\sum_{j=1}^J (\tau_j^0 + t_{i,j}) x_{i,j} + \sum_{j=1}^J y_{i,j} - \sum_{j=1}^J \sum_{k=1}^{j-1} (\tau_k^0 + t_{i,k}) u_{i,j,k} \leq \lambda, \quad i = \overline{1, I}. \quad (10)$$

При использовании критерия (6): $\lambda \rightarrow \min$ получаем задачу частично-целочисленного программирования с булевыми переменными. Представленная задача NP-трудна и относительно исходной (1)-(6) имеет размерность, большую в $J/2$ раз и по числу булевых переменных (за счет $u_{i,j,k}$), и по числу ограничений (8), (9). Ее вычислительная сложность для сколько-нибудь реальных объектов такова, что попытки прямого (точного) решения существующими средствами вычислений не имеют перспектив.

Поэтому в первой главе были разработаны три модификации вычислительно эффективного алгоритма приближенного решения задачи, отличающиеся между собой трудоемкостью и апостериорными оценками близости получаемых решений к оптимальным. В простейшем варианте алгоритм предусматривает редукцию задачи дискретной оптимизации (6)-(10) в бикритериальную задачу того же класса относительно малой размерности следующего вида:

$$\sum_{i=1}^I x_{i,j} = 1, \quad j = \overline{1, J}, \quad \underline{b}_i \leq \sum_{j=1}^J x_{i,j} \leq \overline{b}_i, \quad i = \overline{1, I},$$

$$x_{i,j} = \begin{cases} 1, & \text{если заявка } j \text{ назначается на прибор } i, \\ 0 & \text{в противном случае,} \end{cases} \quad i = \overline{1, I}, \quad j = \overline{1, J} \quad (11)$$

$$\sum_{j=1}^J \tau_j^0 x_{i,j} \leq \beta, \quad i = \overline{1, I}, \quad \sum_{j=1}^J t_{i,j} x_{i,j} \leq \lambda, \quad i = \overline{1, I}, \quad \lambda \rightarrow \min, \quad \beta \rightarrow \min.$$

Из контекста понятно, что данный подход к распределению заявок между приборами предполагает компромиссное решение по «чистому» быстродействию ОС без учета задержек (λ) и по равномерности распределения задержек между приборами (β).

Упрощенная задача (11) NP – трудна в сильном смысле, поскольку кроме ограничений задачи о назначениях, содержит также $2 \cdot I$ ранцевых ограничений. Разработан алгоритм ее решения, включающий построение Парето-оптимального множества расписаний (параметрического решения, при котором один из критериев вводится в состав ограничений с параметром в правой части). Далее получаемое множество Парето - оптимальных решений упрощенной задачи (11) оценивается в терминах задачи (6)-(10) непосредственной постановкой в условия последней. Таким образом, из Парето – оптимального множества решений упрощенной задачи выбирается наилучшее решение исходной задачи (1)-(6). Тестирование алгоритма определило среднее 5.5% отклонение от оптимума значений целевой функции, что в большинстве случаев вполне приемлемо.

Обеспечение большей точности решения потребовало модификаций алгоритма (разделы 7-9). Для этого применены разбиения расписания на входе T^0 на ряд подмножеств, а также разбиения соответствующих задач на ряд подзадач с решением получаемой последовательности. Декомпозиция осуществлена на основе метода динамического программирования. Разработанный при этом алгоритм синтеза расписаний на основе точной (6)-(10) и упрощенной (11) постановок является асимптотически точным, позволяет снизить размерности решаемой

емых подзадач, но существенно увеличивает трудоемкость решения по сравнению с простым алгоритмом. Для снижения трудоемкости применен жадный алгоритм, являющийся обобщением простого алгоритма. Применение разработанного аппарата в совокупности позволило снизить апостериорную оценку отклонения от оптимума до 2% при обеспечении вычислительной эффективности алгоритмов.

В качестве комплексной иллюстрации применения разработанных моделей и алгоритмов к реальному объекту, в десятом разделе главы приведен пример формализации и решения задачи оптимизации расписания строительства кустов скважин нефтегазоконденсатного месторождения. Использование для этого ПОС с задержками поступления заявок определяет поэтапный подход, или подход с неявным заданием последовательностей бурения. В этом случае в качестве заявок выступают кусты скважин, в качестве приборов – буровые установки.

Для описания процессов бурения и строительства скважин НГКМ потребовалось дополнение моделей (1)-(5) и (7)-(10) условиями:

$$\sum_{i=1}^I x_{i,j} \leq 1, \quad j = \overline{1, J}, \quad \sum_{j=1}^J d_j \sum_{i=1}^I x_{i,j} \geq D, \quad (12)$$

где d_j - дебит j - й скважины (куста скважин), D - совокупный спрос на продукты НГКМ, и расширение понятия задержек поступления заявок в системе:

$$\text{му: } T^0 = \|\tau_{i,j}^0\|, \quad \tau_{i,j}^0 = \begin{cases} \tau_i^{\bar{0}}, & \text{если } \tau_i^{\bar{0}} > \tau_j^{c0}, \\ \tau_j^{c0}, & \text{если } \tau_j^{c0} \geq \tau_i^{\bar{0}} \end{cases} - \text{исходная задержка начала бурения}$$

куста скважин j на БУ i . ПОС, которым соответствуют такие условия, названы системами с задержками начала обслуживания заявок.

Кроме этого расширен состав критериев эффективности показателями затрат на бурение $\sum_{i=1}^I \sum_{j=1}^J c_{i,j} x_{i,j} \rightarrow \min$ и прибыли

$$r \sum_{j=1}^J d_j \sum_{i=1}^I x_{i,j} - \sum_{i=1}^I \sum_{j=1}^J c_{i,j} x_{i,j} \rightarrow \max, \quad \text{где } c_{i,j} \text{ затраты, связанные с бурением}$$

скважин j -го куста i -й буровой, определяемые как эксплуатационные расходы и амортизационные отчисления за время бурения, и r - оптовая цена реализации 1 м^3 газа.

Описания всех алгоритмов оптимизации снабжены численными примерами, иллюстрирующими их работу. Приведены результаты решения сгенерированных тестовых задач в широком диапазоне размерностей с оценками быстродействия и точности алгоритмов.

В последнем разделе главы дано краткое описание программной реализации разработанных алгоритмов оптимизации расписаний параллельных систем с задержками начала обслуживания заявок.

Вторая глава посвящена исследованию и разработке моделей и алгоритмов синтеза расписаний многостадийных систем. Рассматриваемые объекты

(многостадийные параллельно-последовательные обслуживающие системы, или ППОС) представимы в виде схемы (рис. 1). ППОС состоит из совокупности блоков, или подсистем, содержащих взаимозаменяемые, в общем случае, неидентичные приборы.

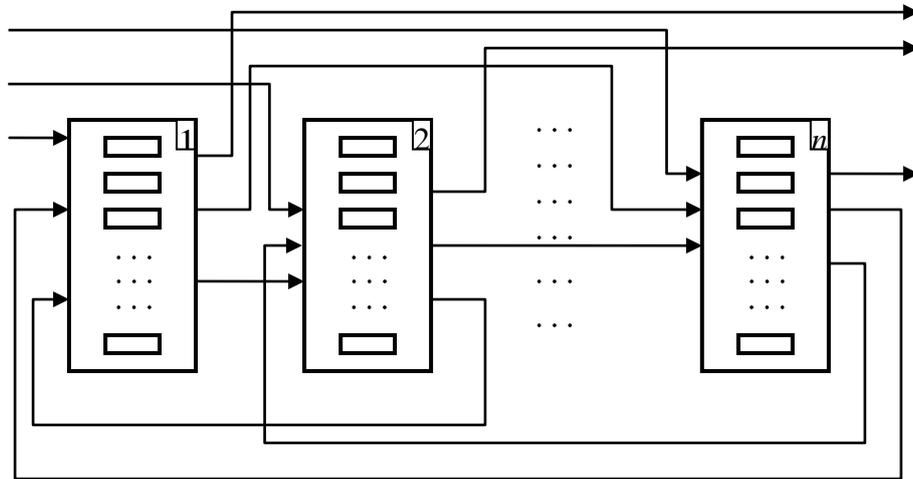


Рис. 1. Параллельно-последовательная обслуживающая система

На рис. 1. они пронумерованы от 1 до n . Заданы технологические последовательности обслуживания всех заявок с точностью до блока (подсистемы). Технологические маршруты фиксированы, но различны для различных заявок. Известно время обслуживания каждой заявки каждым прибором из каждого блока $t_{i,j}$. Время детерминировано, но в общем случае различно для каждого прибора. Прерываний обслуживания одних заявок в пользу других не допускается.

Для представленной системы естественно определить задачу (задачи) синтеза расписаний ее работы, т.е. расписаний обслуживания всех заявок всеми приборами по всем технологическим маршрутам, оптимальные по какому-либо критерию эффективности.

Подсистемы ППОС являются параллельными системами с задержками поступления заявок, модели функционирования которых, как и алгоритмы решения соответствующих задач теории расписаний рассмотрены в первой главе. Следовательно алгоритмы оптимизации расписаний ППОС могут являться синтезом алгоритмов оптимизации расписаний последовательных многостадийных обслуживающих систем (ПМОС) с алгоритмами оптимизации расписаний ПОС, разработанными в рамках первой главы.

Поэтому первым объектом изучения явилась ПМОС, схематично отображенная на рис. 2. ПМОС состоит из k различных приборов, которые обслуживают J заявок. Маршруты обслуживания заявок приборами фиксированы, но различны. Длительности обслуживания каждой заявки каждым прибором также различны. Одновременно одним прибором может обслуживаться только одна заявка. Прерываний обслуживания не допускается. Дисциплина обслуживания очередей - произвольная. Данный тип систем относительно хорошо изучен.

Задача синтеза расписаний для подобных систем является NP-трудной, может интерпретироваться, например, как задача календарного планирования деятельности производственного участка (цеха) и известна в литературе, как

job-shop (JS). Ее исследования начались еще с работ Джонсона, но до сих пор проблема построения полиномиальных (приближенных с оцениваемой точностью) алгоритмов оптимизации расписаний ПМОС для произвольного числа приборов и заявок не потеряла актуальности.

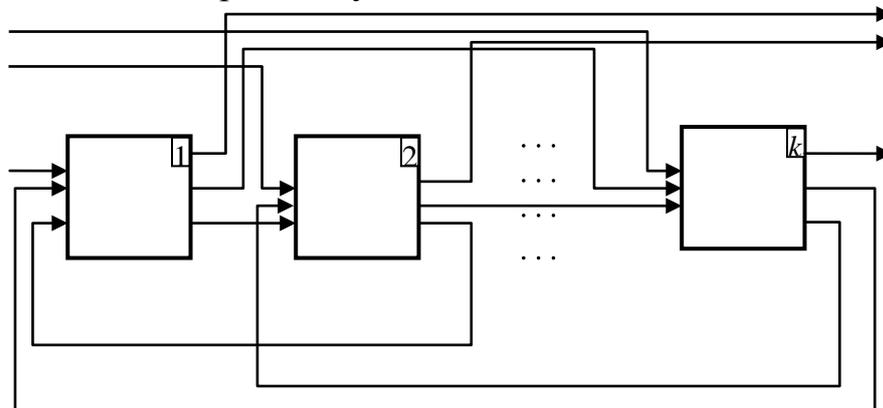


Рис. 2. Последовательная многостадийная обслуживающая система

В главе 2 предложено несколько модификаций алгоритмов решения задачи JS для графического и аналитического ее представлений. В частности, в разделе 2 приведен алгоритм ветвей и границ с использованием в качестве решателя оценочных подзадач алгоритма критического пути. Там же приведена модификация алгоритма ветвей и отсечений, формирующая на каждом шаге подзадачи ЛП из условий и использующая альтернативные (бинарные) отсечения с исключением дополнительных булевых переменных и ограничений. Идея бинарных отсечений развита и обобщена в главе 5.

В разделах 3-5 второй главы разработан и исследован алгоритм неполной декомпозиции задачи оптимизации расписаний ПМОС, сочетающий метод декомпозиции условий с сужением области поиска, позволивший фактически снять проблему размерности общей задачи JS. Алгоритм приведен для графического и аналитического представлений задачи. Декомпозиция осуществляется за счет разбиения множества конфликтов, (ребра сети в графическом представлении) на ряд подмножеств. Из всего многообразия возможных разбиений эмпирическим путем выбраны два способа, дающих хорошие оценки быстродействия и точности.

При первом способе на любом шаге алгоритма в очередное подмножество включаются ребра в порядке следования операций в матрице маршрутов (в приемлемом по размерности количестве, называемом толщиной активного слоя). При этом хорошие результаты дает принцип равномерности потоков операций относительно заявок. При втором способе ребра в текущий слой включаются в соответствии с их рангами.

Сужение области поиска достигается за счет того, что разрешенные конфликты, вышедшие за пределы активного слоя, фиксируются. Тем самым трудоемкость решаемой на любом шаге алгоритма подзадачи дискретной оптимизации определяется только толщиной текущего слоя. Поэтому для толщины слоя существует предел, определяемый возможностями общего алгоритма дискретной оптимизации, который применяется для решения оценочных подзадач.

Поскольку число слоев конечно, алгоритм при названных обстоятельствах (соблюдении регулируемых ограничений на размерности порождаемых подзадач) фактически является эффективным. В пятом разделе приведены результаты численных экспериментов с программной реализацией алгоритма неполной декомпозиции, определившие среднее отклонение решений от оптимума около 2% в широком диапазоне размерностей задачи JS. Показано также, что при его использовании необходимо находить разумный компромисс между точностью и трудоемкостью, которые регулируются параметрами алгоритма.

Далее в разделах (6-10) главы на основе алгоритмов оптимизации расписаний одностадийных параллельных систем с задержками на входе (результаты первой главы) и алгоритма неполной декомпозиции задач оптимизации расписаний последовательных систем разработан общий декомпозиционный алгоритм синтеза оптимальных расписаний параллельно-последовательных систем.

Формальная запись условий общей задачи оптимизации расписаний ППОС приведена ниже.

$$x_{i,j}^{q,p,k} = \begin{cases} 1, & \text{если заявка } j \text{ закрепляется за прибором } i \\ & \text{подсистемы } p \text{ при обращении } k \text{ на шаге } q, \\ 0 & \text{в противном случае;} \end{cases} \quad (13)$$

ограничения на назначения заявок на приборы:

$$\sum_{i \in I_p} x_{i,j}^{q,p,k} = 1, \quad j = \overline{1, J}, q = 1, 2, \dots, k = 1, 2, \dots, p = \overline{1, n},$$

$$\underline{b}_i^{q,p,k} \leq \sum_{j=1}^J x_{i,j}^{q,p,k} \leq \overline{b}_i^{q,p,k}, \quad \forall i \in I_p, q = 1, 2, \dots, k = 1, 2, \dots, p = \overline{1, n},$$

переменные, компенсирующие возможные отрицательные значения задержек $\tau_{i,j}^{q,p,k}$:

$$y_{i,j}^{q,p,k} \geq 0, \quad \forall i \in I_p, q = 1, 2, \dots, k = 1, 2, \dots, p = \overline{1, n}, j = \overline{1, J}.$$

Фактические задержки (расписания) поступления заявок в ОС:

$$\widehat{\tau}_{i,j}^{q,p,k} = \begin{cases} \tau_{i,j}^{q,p,k}, & \text{если } \tau_{i,j}^{q,p,k} \geq 0 \\ 0 & \text{в противном случае} \end{cases}, \quad \forall i \in I_p, q = 1, 2, \dots, k = 1, 2, \dots, p = \overline{1, n}, j = \overline{1, J}.$$

Фактические задержки учитываются посредством использования компенсирующих переменных $y_{j,i}^{q,p,k}$ и нижеследующих ограничений:

- зависимости текущих задержек начала обслуживания заявок от предшествующих задержек

$$\tau_{i,j}^{q,p,k} = \tau_j^{q-1,p} - \sum_{l=1}^{j-1} (\tau_{i,l}^{q,p,k} + t_{i,l}^{p,k}) x_{i,l}^{q,p,k}, \quad \forall i \in I_p, q = 1, 2, \dots, k = 1, 2, \dots, p = \overline{1, n}, j = \overline{1, J},$$

$$\tau_j^{q,p,k} = \sum_{i \in I_p} (\widehat{\tau}_{i,j}^{q,p,k} + t_{i,j}^{p,k}) x_{i,j}^{q,p,k}, \quad \forall q = 1, 2, \dots, k = 1, 2, \dots, p = \overline{1, n}, j = \overline{1, J},$$

- компенсированные задержки

$$\widehat{\tau}_{i,j}^{q,p,k} = \tau_{i,j}^{q,p,k} + y_{i,j}^{q,p,k} \geq 0, \quad \forall i \in I_p, \quad q=1,2,\dots, k=1,2,\dots, p=\overline{1,n}, \quad j=\overline{1,J}, \quad (14)$$

- отношения предшествования-следования технологических операций при соблюдении временных ограничений на межоперационные задержки:

$$\begin{aligned} \widehat{\tau}_{i,j}^{q,p,k} - \widehat{\tau}_{i',j'}^{q,p',k'} + t_{i,j}^{p,k} x_{i,j}^{q,p,k} &\leq 0, \quad \forall (i, j, p, k; i', j, p', k') \in I_{(A,U,\emptyset)}, \\ \widehat{\tau}_{i,j}^{q,p,k} - \widehat{\tau}_{i',j'}^{q,p',k'} + (t_{i,j}^{p,k} + \widehat{t}_{i,j}^{p,k}) x_{i,j}^{q,p,k} &\geq 0, \quad \forall (i, j, p, k; i', j, p', k') \in I_{(A,U,\emptyset)}, \end{aligned} \quad (15)$$

- условия неодновременности выполнения некоторых пар операций

$$Bw_{i,j,j'}^{p,k,k'} - t_{i,j'}^{p,k} x_{i,j'}^{q,p,k'} \geq \widehat{\tau}_{i,j'}^{q,p,k'} - \widehat{\tau}_{i,j}^{q,p,k} \geq t_{i,j}^{p,k} x_{i,j}^{q,p,k} - B(1 - w_{i,j,j'}^{p,k,k'}),$$

$$\forall (i, j, p, k; i, j', p, k') \in I_{(A,\emptyset,V)},$$

$$w_{i,j,j'}^{p,k,k'} = \begin{cases} 1, & \text{если операция } i, j, p, k \text{ предшествует } i, j', p, k' \\ 0, & \text{если операция } i, j', p, k' \text{ предшествует } i, j, p, k \end{cases}$$

$$\forall (i, j, p, k; i, j', p, k') \in I_{(A,\emptyset,V)},$$

- минимаксный критерий быстродействия ППОС

$$\sum_{j=1}^J \widehat{\tau}_{i,j}^q x_{i,j}^q + \sum_{j=1}^J t_{i,j} x_{i,j}^q \leq \lambda, \quad \forall i \in I_p, \quad \lambda \rightarrow \min. \quad (16)$$

Условия общей задачи оптимизации расписаний (ОЗОР) ППОС (13)-(16) содержат рекурсивные функции (формально идентичные (2)), при раскрытии которых представленная задача редуцируется в задачу частично целочисленного линейного программирования с реализациями сверхбольших (для реальных объектов) размерностей. Упрощенным примером может служить задача (6)-(10) на практике неразрешимая точными методами. Поэтому подобная редукция не имеет практических перспектив, а общий алгоритм решения задачи (13)-(16) может быть реализован без этой процедуры. Для решения ОЗОР ППОС разработаны две модификации декомпозиционного алгоритма: 1) прямой, являющийся обобщением декомпозиционного алгоритма оптимизации расписаний ПМОС; 2) циклический алгоритм оптимизации расписаний многостадийных ППОС. Последний основан на поэтапном решении составляющих подзадач: оптимизации расписаний одностадийных параллельных систем с задержками на входе и подзадач оптимизации расписаний ПМОС.

Прямой алгоритм опирается на представление задачи синтеза расписаний ППОС в виде обобщенной смешанной сети специального вида (рис. 3) и основывается на алгоритмах оптимизации расписаний последовательных систем.

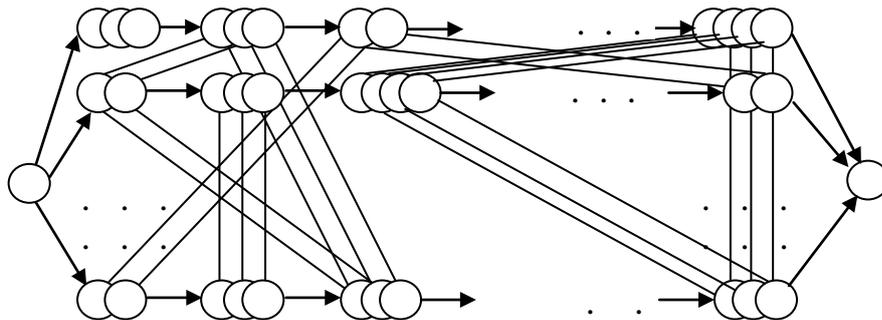


Рис. 3. Графическое представление задачи синтеза расписаний ППОС

Назначения заявок на параллельные приборы в этом случае определяются прямым перебором вариантов. Прямой алгоритм применим при наличии небольшого числа (нескольких единиц) параллельных приборов в подсистемах ППОС.

При выполнении операций, отображаемых вершинами сети на рис. 3, должен осуществляться выбор единственного варианта из альтернатив, определяемых множествами приборов. Этот факт отражен на рисунке с помощью сцепленных вершин. Каждой подобной сцепке соответствует наличие альтернативных вариантов выполнения операции.

В разделе 7 описаны модификации алгоритмов, основанные на вычислительной схеме метода ветвей и отсечений для решения представленной задачи, являющиеся обобщением алгоритмов оптимизации расписаний ПМОС (с неполной декомпозицией в том числе). Настоящий подход распространен и применен при решении прикладной задачи оптимизации расписаний строительства скважин нефтегазоконденсатного месторождения (разделы 11, 12).

Аналитическое представление задачи приведено ниже.

Пусть τ_j время начала бурения скважины (куста скважин) под номером j

$$\tau_j \geq 0, x_{i,j} = \begin{cases} 1, & \text{если буровая } i \text{ назначается на куст скважин } j, \\ 0 & \text{в противном случае, } i = \overline{1, I}, j = \overline{1, J} \end{cases} \quad (17)$$

$$\sum_{i=1}^I x_{i,j} = 1, j = \overline{1, J}, \quad \underline{b}_i \leq \sum_{j=1}^J x_{i,j} \leq \overline{b}_i, i = \overline{1, I},$$

$$\tau_{j'} - \tau_j - \sum_{i=1}^I t_{i,j} x_{i,j} \geq 0, j', j = \overline{1, J}, \text{ если строительство куста } j \text{ предшествует } j'$$

$$\tau_{j^k} + \sum_{i=1}^I t_{i,j^k} x_{i,j^k} \leq \lambda, \quad \forall j^k \in J^k \text{ (множество замыкающих кустов),}$$

$$w_{j,j'} = \begin{cases} 1, & \text{если операция (строительство куста) } j \text{ предшествует } j', \\ 0 & \text{в противном случае,} \end{cases} j', j = \overline{1, J}$$

$$Bw_{j,j'} - \sum_{i=1}^I t_{i,j'} x_{i,j'} \geq \tau_{j'} - \tau_j \geq \sum_{i=1}^I t_{i,j} x_{i,j} - B(1 - w_{j,j'}), j', j = \overline{1, J}, \quad (18)$$

(18) означают неодновременность начала выполнения операций на кустах j', j и являются частным случаем общих, условий характерных для ППОС:

$$Bw_{j,j'} - t_{i,j'} x_{i,j'} \geq \tau_{j'} - \tau_j \geq t_{i,j} x_{i,j} - B(1 - w_{j,j'}), i = \overline{1, I}, j', j = \overline{1, J}. \quad (19)$$

Условия (19) означают, что кусты скважин j' и j не могут быть пробурены одновременно одной буровой i .

$$\sum_{j=1}^J t_{i,j} x_{i,j} \leq \lambda, \quad i = \overline{1, I}, \quad \lambda \rightarrow \min. \quad (20)$$

Задачу (17)-(20) можно рассматривать как обобщение JS и частный случай ОЗОР ППОС (13)-(16). Фактически (17)-(20) является формализацией ЗТР, известной как flexible job-shop (FJS). В отличие от ОЗОР ППОС, она не содержит

рекурсивных функций и для ее решения могут быть применены все модификации прямого алгоритма решения. Необходимо отметить также существование взаимно однозначного соответствия между графическим представлением задачи синтеза расписаний ППОС (рис. 3) и условиями (17)-(20).

В качестве иллюстрации применения представленных в главе моделей и алгоритмов к реальному объекту, приведен пример формализации и решения задачи оптимизации расписания строительства кустов скважин нефтегазоконденсатного месторождения. Используемый при этом подход предусматривает явное задание технологических последовательностей бурения скважин в форме (17)-(20). Показана взаимодополняемость поэтапного подхода (задача (7)-(10), (12)), более применимого при оперативном планировании, и подхода на основе обобщенных смешанных сетей (в аналитическом представлении (17)-(20)), целесообразного при среднесрочном планировании. Результаты решения (17)-(20) служат исходными условиями (расписанием на входе) для поэтапного подхода.

В разделах 9 и 10 представлен циклический декомпозиционный алгоритм решения общей задачи (13)-(16). Регулируя параметры декомпозиционного алгоритма, определяющие размерности решаемых подзадач частично целочисленного линейного программирования с булевыми переменными на каждом шаге, можно добиться фактического полиномиального быстродействия. Единственным ограничением, выявленным эмпирическим путем, является ограничение на число булевых переменных в подзадачах, порождаемых алгоритмом, которое не должно превышать тысячи.

Анализ полученных результатов экспериментов позволяет охарактеризовать разработанный алгоритм как эффективный алгоритм решения общей задачи теории расписаний для параллельно-последовательных систем. Вычислительные эксперименты с разработанными моделями и алгоритмами позволили сделать вывод о наличии условий, при которых гарантируется получение оптимальных по быстродействию расписаний параллельно-последовательных систем.

В последнем разделе главы дано краткое описание программной реализации разработанных алгоритмов оптимизации расписаний ППОС.

В третьей главе осуществлена формальная постановка двух задач оптимизации внешних материальных потоков предприятия.

Решением первой задачи (разделы 2, 3) является: оптимальный выбор параметров входных и выходных материальных потоков.

Формальная запись задачи приведена ниже.

Определить $x_{i,l,k}(t)$, $y_{i,j}(t)$, $w_{j,k}(t)$ при условиях:

$$\sum_{i=1}^I C_{i,j}(t) y_{i,j}(t) = d_j(t), \quad j = \overline{1, J}, t = \overline{1, T}, \quad (21)$$

$$d_j(t) - h_{j,k}(t) w_{j,k}(t) \geq 0, \quad j = \overline{1, J}, t = \overline{1, T}$$

$$0 \leq w_{j,k}(t) \leq 1, \quad w_{j,k}(t) - \text{целые}, \quad y_{i,j}(t) \geq 0, \quad i = \overline{1, I}, j = \overline{1, J}, t = \overline{1, T}, k = \overline{1, K},$$

$$\sum_{j=1}^J [d_j(t) - d_j(t) \sum_{k=1}^K g_{j,k}(t) w_{j,k}(t)] + \Delta Q(t) - Q(t) = 0, \quad (22)$$

$$\Delta Q(t) \geq 0, Q(t) \geq 0 \quad t = \overline{1, T},$$

$$g_j(t) = \begin{cases} g_{j,1}(t), & \text{если } d_j(t) \leq h_{j,1}(t), \\ g_{j,2}(t), & \text{если } h_{j,1}(t) < d_j(t) \leq h_{j,2}(t), \\ \dots \\ g_{j,K}(t), & \text{если } d_j(t) > h_{j,K}(t), \end{cases} \quad j = \overline{1, J}, t = \overline{1, T}$$

$$x_{i,l,1}(t) \leq s_{i,l,1}(t), \quad i = \overline{1, I}, l = \overline{1, L}, t = \overline{1, T},$$

$$x_{i,l,k}(t) \leq s_{i,l,k}(t) - \sum_{k'=1}^{k-1} x_{i,l,k'}(t), \quad i = \overline{1, I}, l = \overline{1, L}, k = \overline{1, K}, t = \overline{1, T},$$

$$\sum_{j=1}^J y_{i,j}(t) + O_i(t-1) \geq \sum_{l=1}^L \sum_{k=1}^K x_{i,l,k}(t), \quad i = \overline{1, I}, t = \overline{1, T},$$

$$O_i(t) = \sum_{j=1}^J y_{i,j}(t) + O_i(t-1) - \sum_{l=1}^L \sum_{k=1}^K x_{i,l,k}(t), \quad i = \overline{1, I}, t = \overline{1, T},$$

$$\Delta Q(t) = \Delta Q(t-1) + \sum_{i=1}^I \sum_{l=1}^L \sum_{k=1}^K p_{i,l,k}(t-1) x_{i,l,k}(t-1) - N(t-1) -$$

$$- \sum_{j=1}^J [d_j(t) - d_j(t) \sum_{k=1}^K g_{j,k}(t) w_{j,k}(t)], \quad t = \overline{1, T}, \quad (23)$$

$$\sum_{t=2}^T \alpha(t) Q(t) \rightarrow \max, \quad \text{при условиях } 0 \leq \alpha(t) \leq 1, \sum_{t=2}^T \alpha(t) = 1,$$

$$\Delta Q(T+1) \rightarrow \max. \quad (24)$$

Здесь i - индекс продукта, j - поставщика, l - типа потребителя, k - интервала шкалы объемов скидок, t - номер месяца.

$y_{i,j}(t)$ и $x_{i,l,k}(t)$ - входной и выходной потоки, $w_{j,k}(t)$ - булева переменная выбора интервала шкалы,

$C_{i,j}(t)$ и $p_{i,l,k}(t)$ - заданные коэффициенты (цены, оценки);

$h_{j,k}(t)$ - значение правой границы интервала объема потока $y_{i,j}(t)$ (рис. 4),

$g_{j,k}(t)$ - соответствующая интервалу k шкалы объемов скидка (в долях единицы),

$s_{i,l,k}(t)$ - значение правой границы интервала k шкалы функции $g_j(t)$.

$Q(t)$, $\Delta Q(t)$ и $N(t)$ - заданные значения, $O_i(t)$ - остатки запасов.

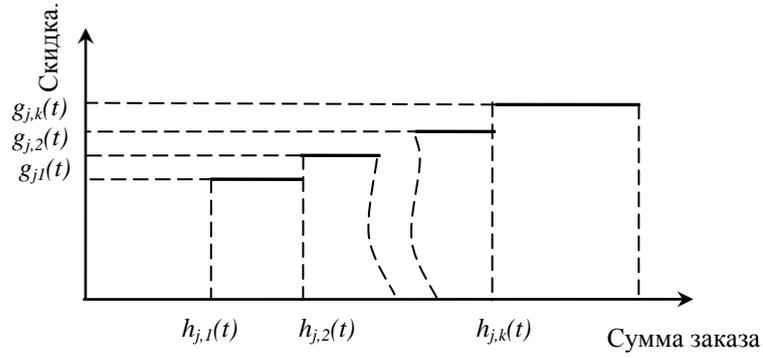


Рис 5. Условия поставок

Показано, что сформулированная задача не имеет точного прямого алгоритма решения, поскольку содержит ограничения (22) и (23), нелинейно зависящие от непрерывных $y_{i,j}(t)$ и булевых $w_{j,k}(t)$ переменных. Вместе с тем, разработан эффективный итерационный алгоритм, позволяющий получать близкие к оптимальным ее решения (разделы 6, 7). Суть алгоритма заключается в итерационном решении релаксированной по отношению (21)-(24) линейной задачи. Релаксация осуществляется по переменным $w_{j,k}(t)$ заменой условий (22) и (23) на ограничения:

$$\sum_{j=1}^J \sum_{i=1}^I [1 - \tilde{g}_j(t)] C_{i,j}(t) y_{i,j}(t) + \Delta Q(t) - Q(t) = 0, \quad t = \overline{1, T}. \quad \text{и} \quad \Delta Q(t+1) = \Delta Q(t) + \sum_{i=1}^I \sum_{l=1}^L \sum_{k=1}^K p_{i,l,k}(t) x_{i,l,k}(t) - N(t) - \sum_{i=1}^I \sum_{j=1}^J [1 - \tilde{g}_j(t+1)] C_{i,j}(t+1) y_{i,j}(t+1) \quad (25)$$

Очевидно, если $\tilde{g}_j(t)$ являются оптимальным выбором для задачи (21)-(24), то оптимальное решение ее релаксации является также оптимальным решением самой задачи (21)-(24). Быстрая сходимость обеспечивается благодаря тому, что функции скидочных коэффициентов $\tilde{g}_j(t)$ являются неубывающими кусочно-линейными зависимостями (рис. 4). Тестирование алгоритма на задачах реальных размеров подтвердило эффективность предложенного подхода.

В разделе 8 представлена квадратичная модификация задачи, более точно учитывающая условия формирования выходных потоков. Для этого ограничения, определяющие условия формирования выходных потоков, дополнены явно задаваемыми зависимостями $p_{i,l}(x_{i,l}(t))$ следующего вида (рис. 5).

Тогда зависимость $p_{i,l}(x_{i,l}(t))$ определяется в соответствии с:

$$p_{i,l}(x_{i,l}(t)) = \begin{cases} p_{i,l,1}(t), & \text{если } x_{i,l}(t) \leq s_{i,l,1}(t), \\ a_{i,l}^0(t) + a_{i,l}^1(t)x_{i,l}(t), & \text{если } x_{i,l}(t) > s_{i,l,1}(t), \end{cases} \quad i = \overline{1, I}, \quad l = \overline{1, L}, \quad t = \overline{1, T},$$

где $a_{i,l}^0(t) = \frac{p_{i,l,2}(t)s_{i,l,1}(t) - p_{i,l,1}(t)s_{i,l,2}(t)}{s_{i,l,1}(t) - s_{i,l,2}(t)}$; $a_{i,l}^1(t) = -\frac{p_{i,l,2}(t) - p_{i,l,1}(t)}{s_{i,l,1}(t) - s_{i,l,2}(t)}$, и

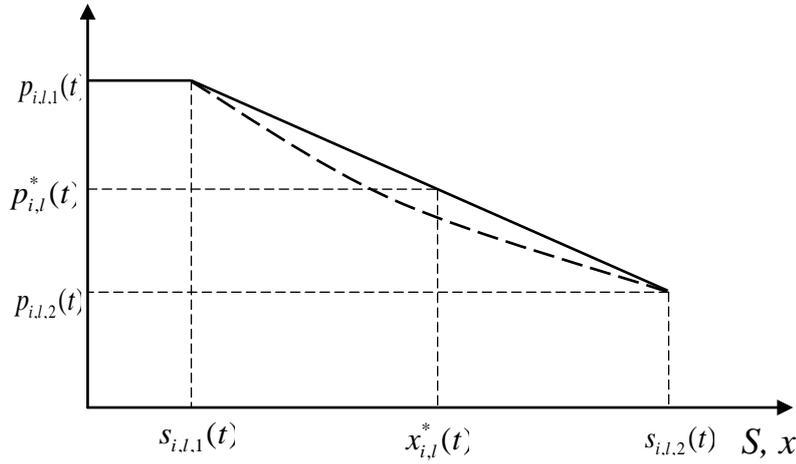


Рис. 5. Зависимости цены от объема продаж продукта

уравнения (25) принимают вид:

$$\begin{aligned}
 & -\Delta Q(t+1) + \Delta Q(t) + \sum_{i=1}^I \sum_{l=1}^L x_{i,l}(t) p_{i,l}(x_{i,l}(t)) - \\
 & - \sum_{i=1}^I \sum_{j=1}^J [1 - \tilde{g}_j(t+1)] C_{i,j}(t+1) y_{i,j}(t+1) = N(t), \quad t = \overline{1, T}, \\
 & \sum_{i=1}^I \sum_{l=1}^L x_{i,l}(t) p_{i,l}(x_{i,l}(t)) \geq \begin{cases} \sum_{i=1}^I \sum_{l=1}^L [a_{i,l}^0(t) x_{i,l}(t) + a_{i,l}^1(t) (x_{i,l}(t))^2], & \text{если } x_{i,l}(t) > s_{i,l,1}(t), \\ \sum_{i=1}^I \sum_{l=1}^L p_{i,l,1}(t) x_{i,l}(t), & \text{если } x_{i,l}(t) \leq s_{i,l,1}(t). \end{cases} \quad (26)
 \end{aligned}$$

Таким образом, модифицированная задача оптимизации дополняется квадратичными ограничениями-неравенствами с альтернативными условиями (26).

В работе предложен эффективный итерационный алгоритм, позволяющий находить близкие к оптимальным решения данной задачи. Эффективность алгоритма обусловлена тем, что исходная квадратичная задача с альтернативными условиями сводится к решению последовательности подзадач полуопределенного программирования. Для подзадач этого класса существуют полиномиальные алгоритмы, относящиеся к группе методов внутренних точек (материалы главы 6). Показано, что максимальное число таких подзадач линейно зависит от числа переменных. Тестирование алгоритма подтвердило его эффективность.

Другая формальная постановка, представленная в третьей главе (разделы 4,5), определена как динамическая задача управления входным потоком промышленного предприятия. Формальная запись представлена ниже.

Определить $x_{i,j}(t)$, $y_{i,l}(t)$, $t = \overline{1, T}$, $i \in I$, $j \in J$, $l \in L$, при условиях:

$$O_j(t-1) + \sum_{i \in I} x_{i,j}(t) \geq P_j(t) + R_j(t), \quad t = \overline{1, T}, \quad j \in J,$$

$$O_j(t-1) + \sum_{i \in I} x_{i,j}(t) - O_j(t) = P_j(t) + R_j(t), \quad t = \overline{1, T}, \quad j \in J,$$

$$O_j(t) \geq R_j(t), \quad t = \overline{1, T}, \quad j \in J,$$

$$\begin{aligned}
& \sum_{j \in J} a_j O_j(t-1) + \sum_{j \in J} a_j \sum_{i \in I} x_{i,j}(t) \leq E(t), \quad t = \overline{1, T}, \\
& \sum_{j \in J} x_{i,j}(t) \leq \sum_{j \in J} \sum_{l \in L} e_{l,j} y_{i,l}(t), \quad t = \overline{1, T}, \quad i \in I, \\
& x_{i,j}(t) \geq 0, \quad t = \overline{1, T}, \quad i \in I, \quad j \in J, \\
& y_{i,l}(t) \geq 0, \quad \text{целые}, \quad t = \overline{1, T}, \quad i \in I, \quad l \in L,
\end{aligned} \tag{27}$$

$$Z = \sum_{t=1}^T \left\{ \sum_{i \in I} \sum_{j \in J} c_{i,j} x_{i,j}(t) + \sum_{i \in I} \sum_{l \in L} s_{i,l} y_{i,l}(t) + d(t) \sum_{j \in J} h_j(t) O_j(t) + N(t) \right\} \rightarrow \min, \tag{28}$$

$$Q = \sum_{i \in I} \sum_{j \in J} q_{i,j} x_{i,j}(t) \rightarrow \max, \tag{29}$$

где i - индекс поставщика, j - материала, l - вида транспорта, t - номер временного интервала, $x_{i,j}(t)$ - размер потока, $y_{i,l}(t)$ - варианты доставки;

$c_{i,j}$, $q_{i,j}$ - цена и оценка качества; $s_{i,l}$ - затраты на единицу вида l ;

$e_{l,j}$ - вместимость единицы вида l для элемента потока j ;

$P_j(t)$, $O_j(t)$, $R_j(t)$ и $h_j(t)$ - потребность, остаток, страховой резерв и штраф за хранение запаса; a_j - площадь на единицу элемента j ;

$E(t)$, $d(t)$, $N(t)$ - площадь складских помещений, средняя цена единицы входного потока, издержки;

Ограничения (27) с критерием эффективности (28), либо (29) составляют NP - трудную задачу частично-целочисленного линейного программирования.

Ее разбиение на подзадачи меньших размерностей и решение генерируемой последовательности подзадач осуществляется с помощью специального алгоритма (раздел 9). Алгоритм решения основан на схеме неполной декомпозиции. Декомпозиция осуществляется по целочисленным переменным $y_{i,l}(t)$, $t = \overline{1, T}$, $i \in I$, $l \in L$, по индексу интервала времени t (время дискретно).

С учетом потенциальных размерностей реализаций сформулированной задачи управления, алгоритм фактически является эффективным. Данный вывод подтвержден при тестировании алгоритма. (раздел 10).

В разделе 11 описаны программные средства, реализующие разработанные алгоритмы решения поставленных в рамках третьей главы задач дискретной оптимизации. На сгенерированных тестовых примерах больших размерностей продемонстрированы высокие точность и быстродействие программных реализаций алгоритмов оптимизации, возможность решения поставленных задач ДО для реальных объектов и фактически в реальном масштабе времени.

В четвертой главе осуществлена формальная постановка общей задачи оптимальной комплектации (ЗОК) технической системы, относящейся к классу задач оптимального проектирования. В общем случае ЗОК представима в виде:

$$Q_{об} = \sum_{v \in V_j^L} \sum_{j \in J^L} r_j y_{j,v}(u) x_{j,v} \rightarrow \max, \quad (30)$$

$$\sum_{v \in V_j^L} \sum_{j \in J^L} c_{j,v}(u) x_{j,v} \leq P, c_{j,v}(u) \in C^j, \quad (31)$$

$$G(x_{j,v}, u_i) = 0, u_i, x_{j,v} = \{0 \vee 1\}, \forall v \in V_j^L, j \in J^L, i \in I, \quad (32)$$

где r_j значимость j -го показателя качества комплекта (подсистемы), $y_{j,v}(u)$ значение j -го показателя качества при v - м варианте реализации, $u = \|u_i\|_I$ - набор дискретных количественных характеристик, от которых зависит качество комплектации, $x_{j,v}$ переменные выбора варианта комплектации,

$$x_{j,v} = \begin{cases} 1, & \text{если выбирается } v\text{-вариант комплектации } j\text{-й подсистемы,} \\ 0, & \text{в противном случае,} \end{cases}$$

$c_{j,v}(u)$ затраты на достижение j -го показателя качества при v - м варианте реализации, P - общие затраты на комплектующие.

(32) задают все возможные логические ограничения, связывающие переменные ЗОК. В простейшем случае переменные выбора $x_{j,v}$ удовлетворяют условию $\sum_{v \in V_j^L} x_{j,v} = 1, \forall j \in J^L$.

Во втором и третьем разделах главы определены основные особенности базовой ЗОК (30)-(32). Во-первых, это единственность ресурсного ограничения (31). Если положить $c_{j,v}(u)$ константами, то (31) - ранцевое ограничение. Во-вторых, посредством эквивалентных преобразований множество логических ограничений (32) можно привести к линейному относительно переменных виду (материалы раздела 3) и показать, что в большинстве случаев коэффициенты при переменных такой линейризованной подсистемы образуют фрагмент абсолютно унимодулярной базисной матрицы системы (31)-(32), в связи с чем ограничения (32) не оказывают влияния на вычислительную сложность ЗОК (30)-(32). И, в-третьих, ограничения ЗОК системы образуют блочную матрицу, что позволяет применить известные методы лагранжевой декомпозиции при решении. Также в третьем разделе приведены правила, используемые для линейризации ЗОК относительно булевых переменных $x_{j,v}$.

В четвертом разделе осуществлена декомпозиция линейризованной задачи оптимальной комплектации (ЛЗОК) и изложены основные постулаты, позволившие заменить булевы переменные $x_{j,v}$ непрерывными. В частности использованы положения теоремы Эверетта о целочисленности решений задач рассматриваемого класса в седловых точках функции Лагранжа при замене дискретных переменных непрерывными. Там же предложен алгоритм параметрического решения выделенных подзадач ЛЗОК для всех седловых точек функции Лагранжа. В качестве параметра рассматривается значение P_j (правая

часть (31) для подсистемы j). Параметру P_j на каждом интервале устойчивости соответствует множитель Лагранжа (двойственная оценка, или предельная полезность) $\lambda_j(P_j)$. Результаты параметрического анализа подзадач выбора подсистем ЛЗОК представимы, например, в виде таблиц:

Таблица 1.

ФПП затрат и функция качества комплектации подсистемы

P_j	$P_j^{\min} < P_j \leq P_j^2$...	$P_j^s < P_j \leq P_j^{s+1}$...	$P_j > P_j^{\max}$
$\lambda_j(P_j)$	λ_j^{\max} ...	λ_j^s ...	0
$x_{j,v}^s$	$x_{j,v}^1(P_j)$...	$x_{j,v}^s(P_j)$...	$x_{j,v}^m(P_j)$
$Q_j(P_j)$	$Q_j^1(P_j)$...	$Q_j^s(P_j)$...	$Q_j^m(P_j)$

В табл. 1 $x_{j,v}^s$, λ_j^s , $Q_j^s(P_j)$ значения переменной выбора, двойственной оценки затрат и критерия качества для подсистемы j на интервале устойчивости s , $s = \overline{1, m}$; $\lambda_j(P_j)$ - невозрастающая функция предельной полезности (ФПП), т.е. $\lambda_j^{\max} \geq \dots \geq \lambda_j^s \geq \dots > 0$, и λ_j^s - константы; $x_{j,v}^s(P_j)$ - параметрические решения на интервале s .

Разработанный алгоритм, основой которого является двойственный симплекс-метод, позволяет построить локальные ФПП всех подсистем комплектуемой системы. Тем самым реализуется декомпозиционная схема решения ЛЗОК, завершает которую алгоритм построения общей ФПП затрат на комплектацию системы и оптимизации по критерию качества, т.е. решения линеаризованной задачи (30)-(32).

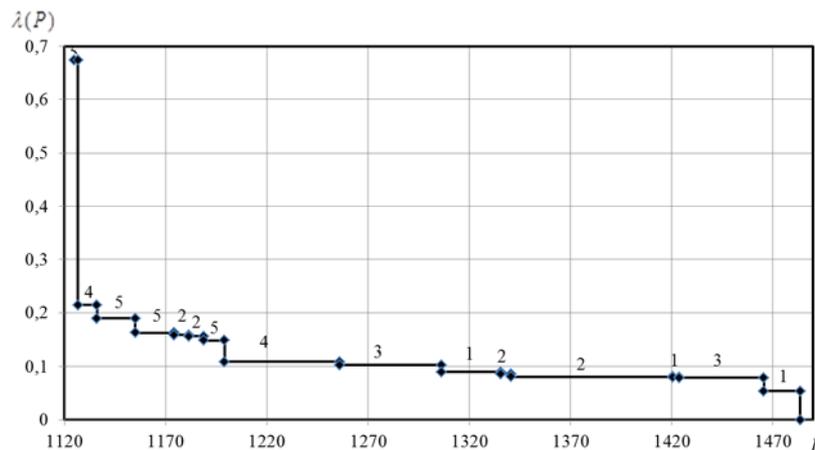


Рис. 6. Пример ФПП ЛЗОК

На рис. 6 изображен график ФПП приведенного в разделе 4 примера задачи комплектации системы пятью подсистемами (от 15 до 22 вариантов комплектации каждой). Маркеры на рисунке соответствуют седловым точкам функции Лагранжа. В этих точках алгоритм определяет точные оптимальные решения ЗОК. Цифры над интервалами устойчивости ФПП отображают номера комплектуемых подсистем.

Алгоритм является асимптотически точным по размерности задач оптимальной комплектации. Этот факт иллюстрируется рис. 7, отображающим соотношение качество-цена $Q(P)$ для одной из тестовых ЛЗОК, содержащей 10^4 булевых переменных. Маркеры на рисунке также соответствуют седловым точкам функции Лагранжа, наглядно демонстрируя точность разработанного алгоритма оптимизации для ЛЗОК больших размерностей.

Программная реализация созданного алгоритма оптимизации ЛЗОК позволила получить также апостериорные оценки точности и быстродействия для задач оптимальной комплектации в широком диапазоне размерностей (разделы 5-7). Табл. 2 содержит результаты решения 30 тестовых примеров, сгенерированных программой-генератором исходных данных ЛЗОК.

Для проведения вычислительных экспериментов были сгенерированы три группы тестовых задач разной размерности (по 10 тестов в группе с числом булевых переменных 1000, 10000 и 100000 соответственно).

Таблица 2.

Результаты тестирования алгоритма

Группа	Число тестов	Число подсистем	max число переменных ПС	Всего булевых переменных	δ_{\max}	Δ_{\max} (%)	Время решения (сек.)
1	10	10	123	1000	0.096337	0.0252	0.000001
2	10	10	1230	10000	0.236492	0.0386	0.000191
3	10	10	12300	100000	0.024639	0.0040	0.187240

Графа «max число переменных ПС» отображает максимальное число переменных выбора альтернативных вариантов комплектации одной из подсистем. Общее же число таких вариантов определяется количеством булевых переменных ЛЗОК (следующая графа). В колонках δ_{\max} и Δ_{\max} отображены максимальные возможные абсолютные и относительные отклонения от оптимума по всем десяти тестам каждой группы, в колонке «время решения» указано максимальное время вычислений также по всем тестам группы.

Из таблицы видно, что возможные отклонения от теоретических оптимумов крайне незначительны и уже при 1000 переменных пренебрежимо малы. С ростом числа булевых переменных растет и точность алгоритма, что подтверждает вывод об асимптотической точности последнего. Данный вывод хорошо иллюстрируется графиком $Q(P)$ для одного из тестов второй группы (рис. 7).

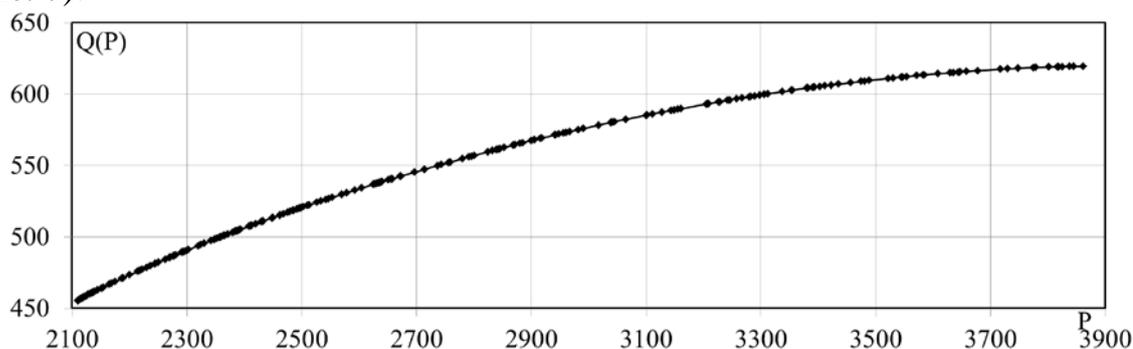


Рис. 7. Соотношение качество-цена для ЛЗОК (10^4 булевых переменных)

Маркеры на рисунке соответствуют седловым точка функции Лагранжа. С ростом размерности число седловых точек пропорционально возрастает.

При этом быстроедействие программной реализации алгоритма решения ЗОК таково, что на практике позволяет решать рассматриваемые задачи любой размерности в реальном масштабе времени.

Пятая глава посвящена описанию алгоритмов нового метода решения задач целочисленного линейного программирования, основанного на использовании бинарных отсечений. Теоретические материалы предваряет краткий обзор и обобщение работ, посвящённых теоретическим и прикладным аспектам теории дискретной оптимизации. Далее, во втором разделе дано теоретическое описание основных положений метода бинарных отсечений (МБО), сформулированы и доказаны ряд утверждений, на которые опираются алгоритмы МБО.

Объект исследования – задача вида

$$\gamma(x) = (c^T, x) \rightarrow \max, \quad (33)$$

$$Ax \leq b, \bar{0} \leq x \leq \bar{1}, \quad (34)$$

$$x \in I_2^n, \quad (35)$$

являющаяся задачей линейного программирования с булевыми переменными (ЗЛПБП). Условия $x \in I_2^n$ определяют принадлежность решения одной из вершин единичного гиперкуба размерности n , $c, x, \bar{0}, \bar{1}$ – векторы той же размерности ($\bar{0}$ – нулевой вектор, $\bar{1}$ – вектор, состоящий из единиц).

Это не уменьшает общности метода, поскольку к ЗЛПБП может быть компактно сведена любая задача целочисленного линейного программирования.

Рассматриваются дополняющие задачу (33)-(34) неравенства следующего вида: $(\hat{\alpha}^T, x) \leq \hat{\beta}_0, \hat{\alpha}_j \in \{-1, 0, 1\}, j = \overline{1, n}, \beta_0 = (\hat{\alpha}^T, x^0),$ (36)

где $\hat{\beta}_0 = [\beta_0]$ и x^0 – оптимальное решение задачи (33)-(34). Если вводимое ограничение является неравенством – следствием базисной системы неравенств (34), то оно именуется правильным сильным отсечением. Такие отсечения являются частным случаем отсечений Гомори и не всегда могут быть построены в силу условия $\hat{\alpha}_j \in \{-1, 0, 1\}$.

Введены ключевые понятия правильного слабого отсечения и соответствующего порождающего неравенства. Первое определяется как любое неравенство (36), отсекающее от ОДР (34) область, не содержащую целых точек. В отличие от сильных, всегда существует множество слабых отсечений (36). Порождающими могут являться любые активные ограничения (34), т.е. неравенства образующие базисную систему при оптимальном решении текущей задачи, а также неравенства – следствия базисной системы неравенств. На этой же основе определены бинарные отсечения (БО), являющиеся частным случаем слабых отсечений вида (36). Введены они для упрощения изложения и обоснования алгоритмов оптимизации посредством нижеследующих преобразований. Слабому БО не соответствует никакое отсечение Гомори. Теоретически, любое

такое отсечение можно построить только с помощью некоторой конечной последовательности отсечений Гомори.

Нормализуем задачу (33)-(35) таким образом, чтобы коэффициенты целевой функции (33) были неотрицательны $c_j \geq 0, j = \overline{1, n}$. Тогда в выражении (36) можно положить $\hat{\alpha}_j \in \{0, 1\}, j = \overline{1, n}$. Это преобразует (36) в неравенства, известные, как отсечения покрытия (обычно применяемые при решении задачи о покрытии).

Редуцированная подобным образом задача (33)-(35):

$$\gamma(\tilde{x}) = (\tilde{c}^T, \tilde{x}) + C \rightarrow \max, \quad (37)$$

$$\tilde{A}\tilde{x} \leq \tilde{b}, \bar{0} \leq \tilde{x} \leq \bar{1}, \quad (38)$$

$$\tilde{x} \in I_2^n, \quad (39)$$

$$\text{где } \tilde{c}_j \geq 0, j = \overline{1, n}, \tilde{x}_j = \begin{cases} x_j, & \text{если } c_j \geq 0 \\ 1 - x_j, & \text{если } c_j < 0 \end{cases}, \tilde{c}_j = \begin{cases} c_j, & \text{если } c_j \geq 0 \\ -c_j, & \text{если } c_j < 0 \end{cases},$$

$$C = - \sum_{j|\forall c_j < 0} c_j, \tilde{b}_i = b_i - \sum_{j|\forall c_j < 0} a_{i,j}; \tilde{a}_{i,j} = \begin{cases} a_{i,j}, & \text{если } c_j \geq 0 \\ -a_{i,j}, & \text{если } c_j < 0 \end{cases}.$$

В этом случае коэффициенты левой части любого i -го отсечения (36) обладают свойством: $\hat{\alpha}_j^i \in \{0, 1\}, j = \overline{1, n}$. Обозначив через \tilde{x}^0 решение задачи (37)-(39), определим БО по отношению к системе (38):

$$(\hat{\alpha}^T, \tilde{x}) \leq \hat{\beta}_0, \hat{\alpha}_j \in \{0, 1\}, j = \overline{1, n}, \hat{\beta}_0 = [\beta_0], \beta_0 = (\hat{\alpha}^T, \tilde{x}^0), \quad (40)$$

а также дополняющую систему БО к системе ограничений (38)

$$A^D \tilde{x} \leq \beta, \quad (41)$$

где $A^D = \|\hat{\alpha}_j^i\|_{m^D \times n}$, $\hat{\alpha}_j^i \in \{0, 1\}, j = \overline{1, n}$ - матрица коэффициентов дополняющей системы, и вектор правых частей отсечений β определяется в соответствии с (40).

Если A^D обладает свойством абсолютной унимодулярности и все отсечения (41) правильные (A^D - часть базисной матрицы системы (38),(41)), то при $m^D \geq n$, A^D включает базисную матрицу и оптимальное решение задачи ЛП (37),(38),(41) является оптимальным решением ЗЛПБП (37)-(39).

Определим условия, которым правильные БО должны удовлетворять.

Рассмотрим вспомогательную задачу:

$$\beta(z) = (\hat{\alpha}^T, z) \rightarrow \max, \quad (42)$$

$$\tilde{A}z \leq \tilde{b}, \bar{0} \leq z \leq \bar{1}, \quad (43)$$

$$\hat{\beta} = [\beta(z^0)]. \quad (44)$$

где z^0 оптимальное решение задачи (42) - (43) и $\hat{\beta}$ - правая часть (40).

Возможны три исхода.

1. $z^0 = \tilde{x}^0$. В этом случае (40) является сильным отсечением.

2. $z^0 \neq \tilde{x}^0$ и $\left[(\hat{\alpha}^T, \tilde{x}^0) \right] = \left[\beta(z^0) \right]$. Тогда (40) - соответственно слабое правильное отсечение.

3. $z^0 \neq x^0$ и $\left[(\hat{\alpha}^T, \tilde{x}^0) \right] < \left[\beta(z^0) \right]$. При данных обстоятельствах имеет место положительная целочисленная невязка $\mu = \left[\hat{\beta}(z^0) \right] - \left[(\hat{\alpha}^T, \tilde{x}^0) \right] > 0$ и неравенство $(\hat{\alpha}^T, \tilde{x}) \leq \left[(\hat{\alpha}^T, \tilde{x}^0) \right]$ - неправильное отсечение. Тогда $(\hat{\alpha}^T, \tilde{x}) \leq \hat{\beta}$ с правой частью (44) не имеет смысла, поскольку не будет активным при дополнении им задачи (37)-(38).

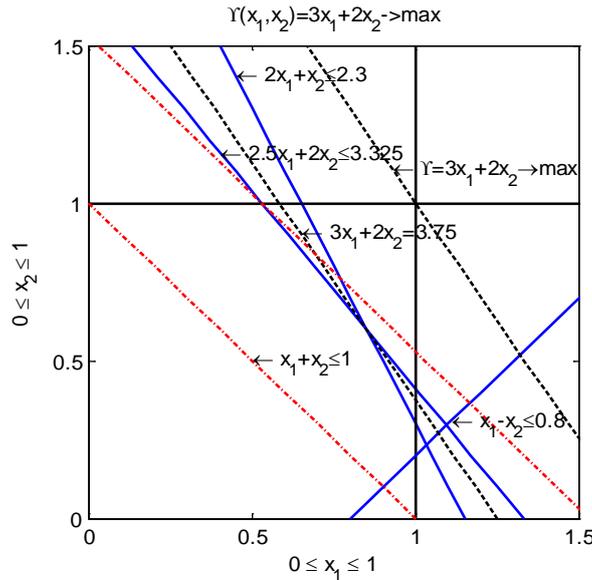


Рис. 8. Пример «слабого» правильного бинарного отсечения

Графическая иллюстрация построения слабого БО на примере: $\gamma(x) = 3x_1 + 2x_2 \rightarrow \max$, $2x_1 + x_2 \leq 2.3$, $2.5x_1 + 3x_2 \leq 3.325$, $x_1 - x_2 \leq 0.8$, $x_1, x_2 \in I_2^2$ представлена на рис. 8.

В релаксированной задаче последнее условие заменено на: $0 \leq x_j \leq 1, j=1,2$. Ее оптимальное решение: $(x^0)^T = (0.85, 0.6)$ и $\gamma_{\max} = 3.75$. Целевая функция задачи (42)-(44) $\beta(z) = z_1 + z_1 \rightarrow \max$.

Решение вспомогательной задачи $(z^0)^T = (0.53, 1)$ не совпадает с x^0 , однако отсечение $x_1 + x_2 \leq 1$ является правильным (выполнено условие 2).

На рис. 8 границы ограничений (38) изображены сплошными линиями, линии уровня целевой функции – пунктиром, границы порождающего неравенства и слабого БО – штрихпунктирными линиями. Для этого примера не существует сильного БО.

Для определения правильности БО кроме (42)-(44) можно воспользоваться альтернативным признаком. Определим:

$$(\hat{\alpha}^T, \tilde{x}) \geq \beta(\tilde{x}^0) + 1, \tag{45}$$

где $\beta(\tilde{x}^0) = \left[(\hat{\alpha}^T, \tilde{x}^0) \right]$ и \tilde{x}^0 - решение задачи (37)-(38).

Если задача (37)-(38), (45) имеет решение, то отсечение $(\hat{\alpha}^T, \tilde{x}) \leq \beta(\tilde{x}^0)$ правильное. И напротив, если (37)-(38), (45) решения не имеет из-за противоречивости системы (38), (45), то $(\hat{\alpha}^T, \tilde{x}) \leq \beta(\tilde{x}^0)$ - правильное БО.

Так, для рассмотренного примера (рис. 8) легко убедиться в том, что дополнительное ограничение $x_1 + x_2 \geq 2$ делает систему неравенств несовместной и $x_1 + x_2 \leq 1$ - правильное отсечение.

Полиномиально трудоемкого алгоритма синтеза правильных БО пока построить не удалось. В работе предложены эвристические процедуры решения задачи синтеза в некотором смысле хороших бинарных отсечений.

В третьем разделе главы предложены четыре подобные процедуры, из которых только первая представляется относительно простым правилом. Вторая эвристика описывается простым быстрым алгоритмом. Последующие (третья и четвертая эвристики), производные от первых двух являются кратно более «тяжелыми» в вычислительно смысле процедурами.

1) *Правило ненулевых координат предшествующей релаксации.* Коэффициенты левой части БО вычисляются по правилу:

$$\hat{\alpha}_j = \begin{cases} 1, & \text{если } \tilde{x}_j^0 > 0 \\ 0, & \text{если } \tilde{x}_j^0 = 0 \end{cases}$$

Правая часть вычисляется в соответствии с (40).

2) *Правило ближайшего отсечения к порождающему неравенству.* Сводится к решению следующей задачи:

найти коэффициенты $\hat{\alpha}_j \in \{0, 1\}$, $j = \overline{1, n}$, максимизирующие выражение

$$k(\hat{\alpha}) = \frac{(\bar{\zeta}^T, \hat{\alpha})}{|\zeta|_2 |\hat{\alpha}|_2} \quad (46)$$

Для точного решения этой задачи разработан алгоритм трудоемкости $O(n \log n)$, основой которого является сортировка.

3) *Выбор на множестве оценок переменных текущей релаксации.*

Пусть λ_j^0 - оценки переменных \tilde{x}_j задачи (37)-(38), определяемые как оптимальные двойственные оценки ограничений $\tilde{x}_j \leq 1$, $j = \overline{1, n}$. ν_j^0 - оценки ограничений задачи двойственной к (37)-(38). Из условий дополняющей нежесткости следует: $\lambda_j^0 = 0$, если $\tilde{x}_j^0 < 1$ и $\lambda_j^0 \geq 0$, если $\tilde{x}_j^0 = 1$. Причем случай $\lambda_j^0 = 0$ и $\tilde{x}_j^0 = 1$ может иметь место, только при наличии альтернативных оптимумов в текущей релаксации. В качестве оценок базисных переменных \tilde{x}_j для случая $0 < \tilde{x}_j^0 < 1$ используем штрафы за увеличение \tilde{x}_j до единицы (\bar{h}_j) и за уменьшение до нуля (\underline{h}_j).

Если обозначить через $\tilde{a}_{l,j}^0$ коэффициенты строки разложения l - й базисной переменной в оптимальной таблице, а соответствующий коэффициент

столбца правых частей через $\tilde{x}_l^0 (< 1)$, то для точной оценки штрафа за снижение \tilde{x}_l до нуля, в преобразованную задачу необходимо добавить ограничение $\sum_{j=1}^n -\tilde{a}_{l,j}^0 \xi_j \leq -\tilde{x}_l^0$, где ξ_j - небазисные переменные в оптимальной таблице. Для

точной оценки штрафа за увеличение \tilde{x}_l до единицы, в преобразованную задачу необходимо добавить ограничение $\sum_{j=1}^n \tilde{a}_{l,j}^0 \xi_j \leq \tilde{x}_l^0 - 1$. Обе такие оценки требуют

серьезных затрат времени на вычисления из-за потребности в решении дополнительных задач ЛП. Поэтому ограничимся минимальными оценками, которые можно определить следующим образом:

$$\underline{h}_l = -\max_{<0} \left\{ \frac{\tilde{c}_j^0}{-\tilde{a}_{l,j}^0} \right\} \cdot \tilde{x}_l^0, \quad \bar{h}_l = -\max_{<0} \left\{ \frac{\tilde{c}_j^0}{\tilde{a}_{l,j}^0} \right\} \cdot (1 - \tilde{x}_l^0).$$

В общем случае оба штрафа вычислим в соответствии с выражениями:

$$\underline{h}_l = \begin{cases} \lambda_l^0, & \text{если } x_l^0 = 1, \\ -\max_{<0} \left\{ \frac{\tilde{c}_j^0}{-\tilde{a}_{l,j}^0} \right\} \cdot \tilde{x}_l^0, & \text{если } 0 < x_l^0 < 1, \\ 0, & \text{если } x_l^0 = 0. \end{cases}$$

$$\bar{h}_l = \begin{cases} 0, & \text{если } x_l^0 = 1, \\ -\max_{<0} \left\{ \frac{\tilde{c}_j^0}{\tilde{a}_{l,j}^0} \right\} \cdot (1 - \tilde{x}_l^0), & \text{если } 0 < x_l^0 < 1, \\ \nu_l^0, & \text{если } x_l^0 = 0. \end{cases}$$

Тогда приоритет координаты \tilde{x}_l в отсечении можно определить, например, как показано ниже (47). Посредством (47) можно оценить все координаты и ввести в соответствии с $h_j, j = \overline{1, n}$ отношение нестрогого порядка на всем множестве $\tilde{x}_j, j = \overline{1, n}$. Благодаря этому возможны построение и оценка целого ряда альтернативных отсечений (40), а не единственного, как при использовании первой и второй эвристик.

$$h_l = \begin{cases} \lambda_l^0, & \text{если } x_l^0 = 1, \\ -\max_{<0} \left\{ \frac{\tilde{c}_j^0}{-\tilde{a}_{l,j}^0} \right\} \cdot \tilde{x}_l^0 + \max_{<0} \left\{ \frac{\tilde{c}_j^0}{\tilde{a}_{l,j}^0} \right\} \cdot (1 - \tilde{x}_l^0), & \text{если } 0 < x_l^0 < 1, \\ -\nu_l^0, & \text{если } x_l^0 = 0. \end{cases} \quad (47)$$

В соответствии с (47) разобьем все множество переменных на три непересекающихся подмножества и упорядочим по убыванию. Пусть n_1, n_2, n_3 - число

ции дополняющая матрица A^D коэффициентов, составляющих БО, всегда является частью базисной матрицы системы ограничений (38), (41).

В этом случае построение A^D полного ранга гарантирует оптимальность решения ЗЛПБП (37)-(39). Рис. 9 иллюстрирует результаты работы циклического алгоритма при решении следующей нормализованной ЗЛПБП.

$\gamma(x) = 3x_1 + 2x_2 \rightarrow \max$, $2x_1 + x_2 \leq 2.3$, $2x_1 + 3x_2 \leq 3.5$, $x_1 - x_2 \leq 0.8$, $x_1, x_2 \in I_2^2$.
В релаксированной задаче последнее условие заменяется на $0 \leq x_j \leq 1$, $j = 1, 2$.

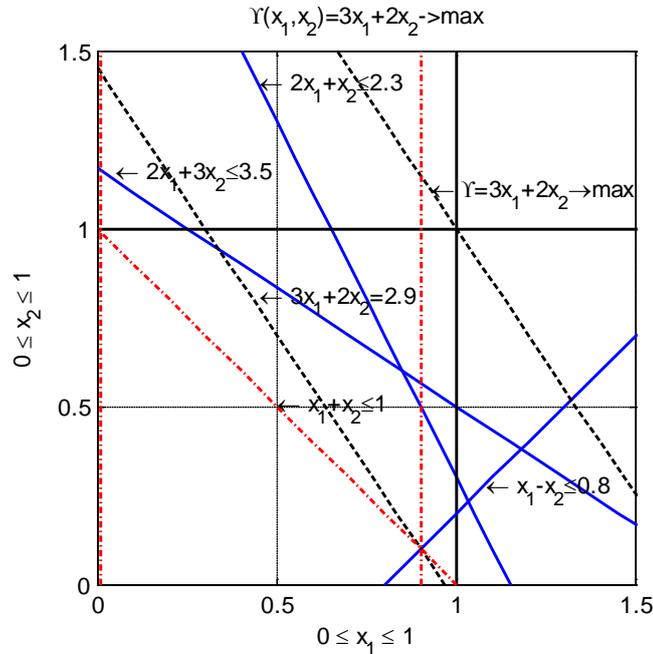


Рис. 9. Построение полной системы бинарных отсечений

Схема формирования подсистемы (41) в данном случае выглядит следующим образом (t - номер шага алгоритма):

$$t := 1 \rightarrow (\hat{a}^{tT}, \tilde{x}) := x_1 + x_2, \hat{b}_1 = 1 \rightarrow x_1 + x_2 \leq 1; \quad t := 2 \rightarrow (\hat{a}^{tT}, \tilde{x}) := x_1, \hat{b}_2 = 0 \rightarrow x_1 \leq 0. \text{ Тогда } A^D = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, b^D = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, x^* = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \gamma_{\max} = 2.$$

Оптимальное решение найдено за два шага с построением A^D полного ранга. Схема и результат полностью совпадает с аналогичной схемой и результатом решения рассмотренного выше примера (рис. 8).

Вместе с тем циклический алгоритм обнаружил ряд отрицательных свойств. Во-первых, это необходимость на каждом его шаге обеспечивать активность подсистемы введенных ранее отсечений. Во-вторых, циклический алгоритм приводит к вырожденности формируемой задачи (37)-(38), (41) до построения дополняющей подсистемы полного ранга, что, в свою очередь, приводит к существенному усложнению алгоритма. В-третьих, при фиксации величины μ для обеспечения эффективности алгоритма, существует опасность исключения из рассмотрения оптимумов ЗЛПБП (37)-(39).

Перечисленные обстоятельства привели к необходимости построения более простого и технологичного алгоритма, основанного на приведенных выше эвристиках, который является гибридным алгоритмом бинарных отсечений и ветвлений (АБОВ), сочетающим идею метода ветвей и границ с построением отсекающих плоскостей рассматриваемого типа. АБОВ (седьмой раздел главы) не требует построения иерархической системы отсечений и, следовательно, обеспечения активности всех вводимых БО на каждом шаге. Проблема возможности порождения неправильного БО решается с помощью схемы ветвлений с использованием свойства БО, определяемого правилом (45). При этом, чем лучше применяемая процедура конструирования отсечений, тем меньше ветвей в дереве решений, и тем эффективней алгоритм оптимизации.

Свойства программной реализации АБОВ исследованы в разделах 7 и 8. В табл. 3 приведены апостериорные оценки, полученные в результате расчетов 30 сгенерированных тестовых примеров различной размерности. Апробированы простейшие варианты процедуры синтеза бинарных отсечений. Эффективность предложенных производных процедур (3 и 4) требует дополнительного анализа, что отмечено в табл. 3 звездочками.

Экспериментально подтверждены надежность алгоритма бинарных отсечений и ветвлений и его применимость для решения смешанных частично-целочисленных задач оптимизации. Последнее свойство также является преимуществом АБОВ перед циклическим алгоритмом.

Сравнение предложенных алгоритмов метода бинарных отсечений возможно только по верхним оценкам трудоемкости.

Верхняя оценка трудоемкости решения ЗЛПБП, общая для всех переборных алгоритмов, составляет: $N = 2^n$, где N - число релаксированных подзадач, которые нужно решить для гарантированного определения оптимума, n - число булевых переменных в задаче. Такую оценку трудоемкости обеспечивают, например, метод полного перебора и динамического программирования.

Таблица 3.

Эффективность применения эвристических процедур синтеза БО

Наименование процедуры	Число правильных отсечений в общем количестве (%)		
	минимум	максимум	в среднем
1) По ненулевым координатам предшественника	31	63	50
2) выбор ближайшего отсечения	15	40	30
3) выбор на множестве оценок переменных	-*	91*	-*
4) выбор на множестве ближайших отсечений	-*	75*	-*

В МВГ N соответствует максимальному числу концевых вершин дерева решений. Поскольку длина любой ветви древа от начальной вершины до концевой составляет $O(n)$, то верхняя оценка трудоемкости МВГ выше, чем у полного перебора: $N_{bb} = O(n)2^n$.

Определим N_{bc} - верхнюю оценку трудоемкости АБОВ. Верхняя оценка числа бинарных отсечений, необходимых для нахождения оптимума ЗЛПБП, при условии, что все отсечения правильные, равна $O(n^2)$. Доля правильных БО

в общем количестве является мерой эффективности эвристики, применяемой в рассматриваемых алгоритмах. Обозначим эту величину через δ , $0 \leq \delta \leq 1$. Тогда $N_{bc} = O(n^2)2^{(1-\delta)n}$ - верхняя оценка трудоемкости АБОВ. Обратимся вновь к табл. 3, из которой следует апостериорная оценка гарантированного числа правильных БО 30% от общего числа отсечений. Таким образом, имеем оценку эффективности первой эвристики $\delta = 0,3$, откуда общая оценка $N_{bc} = O(n^2)2^{0,7n}$.

Третья и четвертая эвристики позволят повысить эффективность отсечений. Полученные результаты могут влиять и на нижние оценки, и потенциально означают многократное преимущество представленного алгоритма перед всеми известными точными алгоритмами целочисленной оптимизации. Например, сравним N_{bb} и N_{bc} : $N_{bb} / N_{bc} = \frac{O(n)2^n}{O(n^2)2^{0,7n}} = \frac{2^{0,3n}}{O(n)}$. Если положить $n = 1000$, получим трудоемкость N_{bc} на 86 порядков ниже чем N_{bb} .

Заметим также, что при стремлении δ к единице, АБОВ становится вычислительно эффективным алгоритмом целочисленного программирования. Этот вывод справедлив и для циклического алгоритма.

Материалы шестой главы обеспечивают необходимый фундамент решения оценочных задач линейного и полуопределенного программирования, используемых в вычислительных методах, разработанных в рамках прикладного раздела (первая часть работы, главы 1-4), а также в рамках пятой главы. Глава посвящена разработке эффективной модификации барьерного алгоритма следования центральному пути (АСЦП), входящего в группу методов внутренних точек. Алгоритм имеет полиномиальную сходимость при решении задач линейного и полуопределенного программирования. Разработанная модификация включает ряд усовершенствований, содержащих выбор наилучших комбинаций процедур выбора длины шага, вида барьерной функции, модификаций прямого и итерационного алгоритмов решения порождаемых последовательностей систем линейных алгебраических уравнений (СЛАУ).

Пара задач ЛП, решаемых АСЦП с учетом ослабляющих переменных:

P^s : найти $\max\{c^T x : Ax + y = b, x \geq 0, y \geq 0\}$, и

D^s : найти $\min\{b^T u : A^T u - v = c, u \geq 0, v \geq 0\}$, задается в виде однородной самодвойственной задачи (с учетом обозначений $\bar{b} = b - Ax^0$, $\bar{c} = c - A^T u^0 - s^0$, $\bar{z} = c^T x^0 + 1 - b^T u^0$):

найти $\min_{(x,s,u,\theta)} ((x^0)^T s^0 + 1)\theta$ при ограничениях

$$\begin{array}{rcll}
-c^T x & +b^T u & +\bar{z}\theta & \geq 0, \\
c\tau & & -A^T u & -\bar{c}\theta \geq 0, \\
-b\tau & +Ax & & +\bar{b}\theta = 0, \\
\bar{z}\tau & +\bar{c}^T x & -\bar{b}^T u & = -((x^0)^T s^0 + 1), \\
& & (\tau, x) & \geq 0, \\
& & (u, \theta) & \text{свободные.}
\end{array}$$

Обозначим также:

$$M = \begin{bmatrix} 0 & A^T & -c \\ -A & 0 & +b \\ +c^T & -b^T & 0 \end{bmatrix}, \quad \hat{x} = \begin{bmatrix} x \\ u \\ \tau \end{bmatrix}, \quad \hat{s} = s(\hat{x}) = \begin{bmatrix} s(x) \\ s(u) \\ s(\tau) \end{bmatrix} = \begin{bmatrix} v \\ y \\ \rho \end{bmatrix} \Rightarrow M\hat{x} \geq 0; \quad \hat{x} \geq 0;$$

$$s(\hat{x}) = Mx.$$

Решение исходной пары задач P^s и D^s определяется как $\left(\frac{x}{\tau}, \frac{u}{\tau}\right)$.

Если определить вектор $r = e - Me$, матрицу $\begin{bmatrix} M & r \\ -r^T & 0 \end{bmatrix}$, число $\lambda = n + 1$ и

дополнить переменные задачи парой \mathcal{G} и w , то система, определяющая центральный путь, запишется следующим образом:

$$-\begin{bmatrix} M & r \\ -r^T & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \mathcal{G} \end{bmatrix} + \begin{bmatrix} \hat{s} \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda \end{bmatrix}; \quad \begin{bmatrix} \hat{x} \\ \mathcal{G} \end{bmatrix}, \begin{bmatrix} \hat{s} \\ w \end{bmatrix} \geq 0, \quad \begin{bmatrix} \hat{x} \\ \mathcal{G} \end{bmatrix} \begin{bmatrix} \hat{s} \\ w \end{bmatrix} = \begin{bmatrix} \mu e \\ \mu \end{bmatrix} \quad (48)$$

В качестве стартовой точки можно выбрать единичный вектор $(\hat{x}^0, \mathcal{G}^0, \hat{s}^0, w^0) = (e, 1, e, 1)$, который всегда является допустимым решением (48), т.к. принадлежит центральному пути. Начальное значение параметра $\mu = 1$.

Шаг метода Ньютона состоит в поиске приращений $(\Delta\hat{x}, \Delta\mathcal{G}, \Delta\hat{s}, \Delta w)$ после редукции параметра μ , таких, что:

$$-\begin{bmatrix} M & r \\ -r^T & 0 \end{bmatrix} \begin{bmatrix} \hat{x} + \Delta\hat{x} \\ \mathcal{G} + \Delta\mathcal{G} \end{bmatrix} + \begin{bmatrix} \hat{s} + \Delta\hat{s} \\ w + \Delta w \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda \end{bmatrix}; \quad \begin{bmatrix} \hat{x} + \Delta\hat{x} \\ \mathcal{G} + \Delta\mathcal{G} \end{bmatrix}, \begin{bmatrix} \hat{s} + \Delta\hat{s} \\ w + \Delta w \end{bmatrix} \geq 0, \\
\begin{bmatrix} \hat{x} + \Delta\hat{x} \\ \mathcal{G} + \Delta\mathcal{G} \end{bmatrix} \begin{bmatrix} \hat{s} + \Delta\hat{s} \\ w + \Delta w \end{bmatrix} = \begin{bmatrix} \mu e \\ \mu \end{bmatrix}.$$

Пренебрегая квадратичными слагаемыми $\Delta\hat{X}\Delta\hat{S}$ и $\Delta\mathcal{G}\Delta w$ системы уравнений, получаем ее линейную аппроксимацию:

$$\begin{array}{rcll}
-M\Delta\hat{x} & -r\Delta\mathcal{G} & +\Delta\hat{s} & = 0 \\
r^T\Delta\hat{x} & & +\Delta w & = 0 \\
\hat{S}\Delta\hat{x} & +\hat{X}\Delta\hat{s} & & = \mu e - \hat{X}\hat{s} \\
w\Delta\mathcal{G} & +\mathcal{G}\Delta w & & = \mu - \mathcal{G}w
\end{array} \quad (49)$$

После определения $(\Delta\hat{x}, \Delta\mathcal{G}, \Delta\hat{s}, \Delta w)$, следует скорректировать текущее решение, используя при необходимости ослабляющий фактор α :

$$\hat{x} := \hat{x} + \alpha\Delta\hat{x}, \mathcal{G} := \mathcal{G} + \alpha\Delta\mathcal{G}, \hat{s} := \hat{s} + \alpha\Delta\hat{s}, w := w + \alpha\Delta w.$$

Основные параметры, с которыми оперирует модифицированный АСЦП в рамках диссертационной работы, определены таким образом, чтобы добиться универсальности относительно используемых барьерных функций (БФ):

1. В качестве меры близости к центральному пути используется любое из выражений:

$$\Psi(v) = \sum_{j=1}^n \psi(v_j) \text{ и } \delta(v) = \frac{1}{2} \|\nabla\Psi(v)\|_2, \text{ где } v_j = \left(\frac{x_j s_j}{\mu} \right)^{\frac{1}{2}}, \text{ а } \psi(t) - \text{любая са-}$$

мосогласующаяся одномерная БФ, $\nabla\Psi(v)$ - градиент $\Psi(v)$ ($\Psi(v)$ и $\delta(v)$ по определению являются самосогласующимися многомерными БФ).

2. Трансформируется система (49):

$$\begin{aligned} -M\Delta\hat{x} \quad -r\Delta\mathcal{G} \quad +\Delta\hat{s} &= 0 \\ r^T\Delta\hat{x} &+ \Delta w = 0 \\ \hat{S}\Delta\hat{x} \quad +\hat{X}\Delta\hat{s} &= -\mu w \nabla\Psi(v) \\ w\Delta\mathcal{G} \quad +\mathcal{G}\Delta w &= -\mu \left(\frac{\mathcal{G}w}{\mu} \right)^{\frac{1}{2}} \psi' \left(\sqrt{\frac{\mathcal{G}w}{\mu}} \right) \end{aligned} \quad (50)$$

3. Определяется окрестность центрального пути для всех v_j , посредством решения уравнения $\psi(t) = \beta$, β - предельная мера близости к центральному пути. Вычисленные корни уравнения $0 < \gamma_l < 1$, $1 < \gamma_u < \infty$ определяют допустимую окрестность центрального пути. Тем самым задаются условия $\gamma_l \leq v_j \leq \gamma_u, \forall j = \overline{1, k+1}$.

4. Критерии останова алгоритма.

Критерий останова классического АСЦП $(n+1)\mu \leq \varepsilon$ имеет скорее теоретическое, чем прикладное значение. Данное условие при больших n на конечных итерациях опосредует столь малые значения барьерного параметра μ , что на фоне вычислительных возмущений алгоритм теряет устойчивость. Это может привести к нулевой длине шага с заикливанием и плохой обусловленности матрицы \tilde{A} . Причиной является погрешность метода, заключающаяся в образовании на завершающих шагах как близких нулю, так и машинной бесконечности элементов диагональной матрицы $\hat{X}^{-1}\hat{S}$ (при оптимальном решении x и s ортогональны). На этот естественный процесс накладываются ошибки округления, из чего вытекает потребность в двух вспомогательных критериях останова:

$$\alpha < \varepsilon_1, \text{ и } \Delta\tau < \varepsilon_2, \varepsilon_2 \geq \varepsilon_1 \geq \varepsilon. \quad (51)$$

Обычно в практике применения МВТ используется совокупность критериев – невязок решений пары P^s и D^s , либо свертка невязок. В частности, в Matlab в качестве критерия останова модуля *lipsol* применяется

$$\frac{\|r_b\|_2}{\max(1, \|b\|_2)} + \frac{\|r_c\|_2}{\max(1, \|c\|_2)} + \frac{|c^T x - b^T u|}{\max(1, |c^T x|, |b^T u|)} \leq \varepsilon, \quad (52)$$

где, $r_b = Ax + y - b$, $r_c = A^T u - v - c$.

Условия Куна – Таккера для P^s и D^s опосредуют:

для активных ограничений P^s : $U(b - Ax) = \mu e \Rightarrow Ur_b = \mu e \Rightarrow r_b = U^{-1} \mu e$;

для активных ограничений D^s : $X(A^T u - c) = \mu e \Rightarrow Xr_c = \mu e \Rightarrow r_c = X^{-1} \mu e$.

Поэтому $\|r_b\|_2 = \frac{\mu}{\|u\|_2}$, $\|r_c\|_2 = \frac{\mu}{\|x\|_2}$, $|c^T x - b^T u| = \rho$. Тогда критерий останова

(52) применительно к АСЦП будет выглядеть следующим образом:

$$\frac{\mu}{\max(1, \|u\|_2, \|b\|_2 \|u\|_2)} + \frac{\mu}{\max(1, \|x\|_2, \|c\|_2 \|x\|_2)} + \frac{\rho}{\max(1, |c^T x|, |b^T u|)} \leq \varepsilon. \quad (53)$$

Очевидно:

$$\frac{\mu}{\max(1, \|u\|_2, \|b\|_2 \|u\|_2)} \leq \mu; \quad \frac{\mu}{\max(1, \|x\|_2, \|c\|_2 \|x\|_2)} \leq \mu; \quad \frac{\rho}{\max(1, |c^T x|, |b^T u|)} \leq \rho.$$

Поскольку ρ также пропорционально μ , то, по сути дела, критерий останова (53) эквивалентен $\mu \kappa \leq \varepsilon$, где $\kappa > 1$ - некоторая константа. Несмотря на то, что (53) не зависит от размерности и много мягче $(n+1)\mu \leq \varepsilon$, он не исключает употребление дополнительных условий останова (51).

В разделе 6 представлена процедура определения оптимальной длины шага АСЦП. Показано, что ослабляющий параметр длины шага α определяется, как результат совместного решения в общем случае $m+n+1$ следующих одномерных задач оптимизации.

Найти $\alpha = \min(\max_j(\alpha_{xj}), \max_j(\alpha_{sj}))$ при условиях

$$\alpha_{xj}^2 + \frac{\mu v_j \psi'(v_j)}{\Delta x_j s_j} \alpha_{xj} + \left(\frac{\mu v_j \psi'(v_j)}{\Delta x_j s_j} \frac{x_j}{\Delta x_j} - \left(\frac{x_j}{\Delta x_j} \right)^2 + \frac{\mu \gamma_l^2}{\Delta x_j s_j} \frac{x_j}{\Delta x_j} \right) \leq 0, \quad \forall j = \overline{1, n},$$

$$\alpha_{sj}^2 + \frac{\mu v_j \psi'(v_j)}{\Delta x_j s_j} \alpha_{sj} + \left(\frac{\mu v_j \psi'(v_j)}{\Delta x_j s_j} \frac{x_j}{\Delta x_j} - \left(\frac{x_j}{\Delta x_j} \right)^2 + \frac{\mu \gamma_u^2}{\Delta x_j s_j} \frac{x_j}{\Delta x_j} \right) \geq 0, \quad \forall j = \overline{1, n}.$$

При этом доказано, что БФ должны обладать свойством: $v_j^2 - v_j \psi'(v_j) \geq \gamma_l^2$

Тогда:

$$\alpha_{xj}^* = -\frac{\mu v_j \psi'(v_j)}{2s_j \Delta x_j} + \sqrt{\left(\frac{\mu v_j \psi'(v_j)}{2s_j \Delta x_j} - \frac{x_j}{\Delta x_j} \right)^2 - \frac{\mu \gamma_l^2}{s_j \Delta x_j} \frac{x_j}{\Delta x_j}}, \text{ и}$$

$$\alpha_{sj}^* = -\frac{s_j \Delta x_j}{\Delta s_j x_j} \alpha_{xj}^* - \frac{\mu v_j \psi'(v_j)}{\Delta s_j x_j}, \quad \forall j = \overline{1, n}.$$

Для рассматриваемой модификации АСЦП по результатам анализа (разделы 6 и 7) определены две наилучшие БФ (из 9 рассмотренных разновидностей, принадлежащих классам саморегулируемых и самосогласующихся барьерных функций): $\psi_9(t)$, включающая, как частный случай $\psi_4(t)$ и являющаяся, по сути, семейством функций, и классическая логарифмическая $\psi_1(t)$

$$(\psi_1(t) = \frac{t^2 - 1}{2} - \ln t, \psi_9(t) = \frac{t^{p+1} - 1}{p(p+1)} + \frac{t^{1-q} - 1}{q(q-1)} + \frac{p-q}{pq}(t-1), \quad 0 < p \leq 1, q \geq 1).$$

В разделах 8 и 9 разработаны модификации прямого и итерационного алгоритмов решения последовательности СЛАУ специального вида (50), порождаемой АСЦП. Прямой алгоритм основан на разновидности треугольного разложения, максимально учитывающей специфику матрицы системы и имеющую минимальную для прямых методов трудоемкость. При этом система (50) приводится к квазисимметричному виду, что и позволяет получить совокупную трудоемкость прямого алгоритма решения рассматриваемых СЛАУ, не превышающую трудоемкость метода квадратного корня (для симметричных систем).

Модификация итерационного алгоритма является разновидностью предобусловленного алгоритма бисопряженных градиентов, также максимально учитывающей особенности матрицы системы (50). Предобусловливатель естественным образом определяется посредством применения прямого алгоритма на ранних стадиях АСЦП.

В совокупности разработанная модификация АСЦП включает все перечисленные усовершенствования и выглядит следующим образом:

1. Выбор: 1) самосогласующейся БФ $\psi(t)$ и ее параметров, удовлетворяющих условиям процедуры определения оптимальной длины шага; 2) собственно процедуры вычисления длины шага α ; 3) предельной меры близости к центральному пути β ; 4) параметров точности решения $\varepsilon, \varepsilon_1, \varepsilon_2 > 0, \varepsilon_2 < \varepsilon_1 < \varepsilon$; 5) декремента барьерного параметра $\theta, 0 < \theta < 1$;

2. Задание начальной точки (\hat{x}^0, \hat{s}^0) , и $\mu^0 \leq 1$, таких, что $\Psi(\hat{x}^0, \hat{s}^0, \mu^0) \leq \beta$;

3. $\hat{x} := \hat{x}^0$; $\hat{s} := \hat{s}^0$; $\mu := \mu^0$;

4. Если $\frac{\mu}{\max(1, \|u\|_2, \|b\|_2 \|u\|_2)} + \frac{\mu}{\max(1, \|x\|_2, \|c\|_2 \|x\|_2)} + \frac{\rho}{\max(1, |c^T x|, |b^T u|)} < \varepsilon$,

или $\alpha < \varepsilon_1$, или $\Delta\tau < \varepsilon_2$, конец работы алгоритма, иначе:

5. $\mu := (1 - \theta) \mu$;

6. Если $\Psi(\hat{x}, \hat{s}, \mu) < \beta$, возврат к п. 4, иначе:

7. Решение системы (50), (определение $\Delta\hat{x}, \Delta\hat{s}$);

8. Вычисление параметра длины шага α ;

9. $\hat{x} := \hat{x} + \alpha\Delta\hat{x}$; $\hat{s} := \hat{s} + \alpha\Delta\hat{s}$; возврат к п. 6. ■

Программная реализация и тестирование алгоритма подтвердили его высокую эффективность. Так быстродействием программная реализация многократно превзошла модуль Matlab *lipsol* и почти в 1.5 раза реализацию барьерного алгоритма в CPLEX для однопроцессорного (без распараллеливания вычислений) варианта. При этом алгоритм, протестированный на множестве сгенериро-

ванных задач больших размерностей, обнаружил высокую численную устойчивость. Одним из результатов проведенного эмпирического анализа является вывод о фактической независимости числа шагов АСЦП от размерности решаемых задач, которое принадлежит интервалу $[16 \div 24]$ (размерности тестов менялись в диапазоне $[400 \times 300 \div 10000 \times 5000]$ при плотной заполненности матрицы ограничений).

В заключении отмечено, что все предложенные в работе новые математические модели, методы и алгоритмы направлены на решение глобальной проблемы: создания вычислительно эффективных инструментов решения дискретных задач оптимизации управления производством. При этом достигнута главная цель работы и решены все поставленные задачи, разработаны программные средства, реализующие созданные в рамках диссертационной работы вычислительные методы. С помощью этого ПО оценены быстродействие и точность алгоритмов оптимизации. Во всех случаях получены ожидаемые, вполне отвечающие практическим нуждам оценки быстродействия и точности.

В рамках решения поставленных задач автором впервые получены следующие научные и практические результаты.

1. Формализована задача оптимизации расписаний параллельной ОС с задержками начала работы приборов и поступления заявок на обслуживание. Ценой увеличения размерности осуществлена редукция задачи оптимизации расписаний параллельной системы, содержащей нелинейные ограничения и булевы переменные, в задачу частично целочисленного линейного программирования.

2. В целях обеспечения эффективного решения задачи оптимизации расписаний параллельных ОС с задержками начала обслуживания использован принцип упрощающей редукции и сужения области поиска. В результате сформулирована упрощенная бикритериальная задача линейного программирования с булевыми переменными, и разработаны эффективные алгоритмы ее решения, отличающиеся апостериорными оценками трудоемкости и близости решений к оптимумам.

3. Разработанные алгоритмы синтеза расписаний параллельных систем непосредственно применены в рамках поэтапного подхода для формализации и решения крупной народнохозяйственной проблемы: разработки оптимальных по быстродействию календарных графиков обустройства нефтегазоконденсатных месторождений с учетом ресурсных и временных ограничений.

4. Разработаны алгоритмы оптимизации расписаний классической задачи job shop: точный алгоритм (на основе метода ветвей и границ) и приближенные алгоритмы, основанные на неполной декомпозиции, фактически являющиеся эффективными при соблюдении регулируемых ограничений на размерности подзадач, порождаемых в процессе счета.

5. Дана общая формальная постановка задачи теории расписаний для многостадийных обслуживающих систем. На базе алгоритмов оптимизации расписаний одностадийных параллельных систем с задержками на входе и алгоритма неполной декомпозиции задач оптимизации расписаний

последовательных систем, разработан эффективный алгоритм приближенного решения задачи оптимизации расписаний ППОС.

6. В рамках концепции параллельно-последовательных систем рассмотрен подход к формализации и решению задачи календарного планирования строительства скважин НГКМ с явным заданием последовательности бурения. Результирующая формальная постановка позволяет эффективно применить разработанные прямые алгоритмы оптимизации расписаний ППОС для нефтегазоконденсатных месторождений.

7. Осуществлена формальная постановка двух типичных задач оптимизации управления входными и выходными материальными потоками предприятия. Рассмотрен ряд модификаций, содержащих целочисленные переменные и квадратичные ограничения, дополненные альтернативными условиями. Созданы эффективные итерационные алгоритмы их решения.

8. Формализована общая задача оптимальной комплектации технической системы подсистемами, являющаяся NP-трудной задачей дискретной оптимизации. На основе методов линеаризации и лагранжевой декомпозиции разработан эффективный алгоритм поиска ее решений. Алгоритм является асимптотически точным по размерности задачи оптимальной комплектации.

9. Предложен новый подход к решению общей задачи целочисленного линейного программирования, основанный на процедуре конструирования бинарных отсечений. Разработаны эвристические процедуры построения бинарных отсечений, позволившие реализовать быстрые алгоритмы решения задач линейного программирования с булевыми переменными. Программно реализованы два варианта процедуры синтеза отсечений.

10. Разработаны два алгоритма метода бинарных отсечений: циклический алгоритм и алгоритм бинарных отсечений и ветвлений. Экспериментально подтверждена применимость метода для решения смешанных частично-целочисленных задач оптимизации.

11. Усовершенствован алгоритм следования центральному пути в недопустимой области с длинным шагом. Алгоритм используется в качестве вычислителя при решении оценочных задач линейного и полуопределенного программирования, порождаемых алгоритмами ДО. Удалось добиться его универсальности относительно используемых барьерных функций, процедур определения длины шага и других параметров. Применительно к АСЦП разработаны эффективные модификации прямого и итерационного алгоритмов решения последовательностей СЛАУ специального вида.

Основные результаты диссертации опубликованы:

Статьи в журналах из перечня ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертаций:

1. *Мезенцев Ю.А.* Алгоритмы синтеза расписаний многостадийных обслуживающих систем в календарном планировании. Омский научный вестник. Омск, Изд-во ОГТУ 2006. №3(36) С. 97–102.

2. *Мезенцев Ю.А.* Декомпозиционный метод решения одного класса задач оптимального проектирования. Научный вестник НГТУ. Новосибирск, Изд-во НГТУ 2006. №3(24), С. 67–100.
3. *Мезенцев Ю.А.* Практические аспекты эффективности алгоритма следования центральному пути метода внутренних точек. Научный вестник НГТУ. Новосибирск, Изд-во НГТУ 2006. №4(25) С. 67–104.
4. *Иванов Л.Н., Мезенцев Ю.А.* Модели синтеза расписаний параллельных обслуживающих систем. Омский научный вестник. Омск, Изд-во ОГТУ 2006. №9(46) С. 164–167.
5. *Иванов Л.Н., Мезенцев Ю.А.* Методы оптимизации расписаний параллельных обслуживающих систем. Программные продукты и системы. Тверь, Изд-во МНИИПУ и НИИ «Центрпрограммсистем» 2008. №1 С. 72-74.
6. *Мезенцев Ю.А.* Оптимизация расписаний параллельных динамических систем в календарном планировании. Информационные технологии. М., Изд-во «Новые технологии» 2008. №2 С. 24-33.
7. *Мезенцев Ю.А.* Математические модели управления подсистемами логистики на предприятиях. Журнал «Автоматизация и современные технологии», М, Изд-во «Машиностроение» 2008. №8. С. 46-55.
8. *Мезенцев Ю.А.* Оптимизация расписаний последовательно- параллельных обслуживающих систем. Программные продукты и системы. Тверь, Изд-во МНИИПУ и НИИ «Центрпрограммсистем» 2009. №1 С. 22-26.
9. *Мезенцев Ю.А.* Неполная факторизация и вопросы эффективности алгоритма центрального пути. Научный вестник НГТУ. Новосибирск, Изд-во НГТУ 2009. №1(34), С. 69–85.
10. *Мезенцев Ю.А.* Эффективный алгоритм целочисленного программирования. Научный вестник НГТУ. Новосибирск, Изд-во НГТУ 2009. №2(35), С. 91–114.
11. *Мезенцев Ю.А.* Оптимизация расписаний параллельно-последовательных систем в календарном планировании // Информационные технологии. М., Изд-во «Новые технологии» 2009. №6 С. 35-41.
12. *Авдеенко Т.В., Кравченко А.В., Мезенцев Ю.А.* Модели планирования производства изделий, основанных на нанотехнологиях // Программные продукты и системы. Тверь, Изд-во МНИИПУ и НИИ «Центрпрограммсистем» 2009. №4 С. 22-25.
13. *Мезенцев Ю.А.* Метод бинарных отсечений и ветвлений целочисленного программирования // Доклады академии наук высшей школы РФ. Новосибирск: Изд-во НГТУ 2011. № 1(16) С. 12-25.
14. *Мезенцев Ю.А., Павлов П.С.* Реализация алгоритма решения специальных задач полуопределенного программирования с использованием IBM PLOG CPLEX // Научный вестник НГТУ. Новосибирск, Изд-во НГТУ 2011. №4(45), С. 25-34.
15. *Мезенцев Ю.А., Павлов П.С.* К программной реализации декомпозиционного алгоритма решения одного класса задач дискретной оптимизации с полуопределенной релаксацией // Информационные технологии. М., Изд-во «Новые технологии» 2012. №2 С. 54-59.

16. *Мезенцев Ю.А.* Эффективный алгоритм решения одного класса задач целочисленного программирования. // Доклады академии наук высшей школы РФ. Новосибирск: Изд-во НГТУ 2012. № 2(19) С.42- 53.
17. *Мезенцев Ю.А., Павлов П.С.* Практические аспекты решения одной задачи оптимального планирования обустройства нефтегазоконденсатных месторождений // Научный вестник НГТУ. Новосибирск, Изд-во НГТУ 2012. №4(49), С. 48-55.
18. *Мезенцев Ю.А.* Практические аспекты реализации эффективного алгоритма решения задач оптимальной комплектации. // Доклады академии наук высшей школы РФ. Новосибирск: Изд-во НГТУ 2013. № 1(20) С. 26-34.

Свидетельства об официальной регистрации программных систем разработанных на основе результатов диссертации:

19. Программная реализация прямо-двойственного барьерного алгоритма следования центральному пути для решения задач линейного программирования большой размерности / *Мезенцев Ю.А.* // Свидетельство о государственной регистрации № 2012611589 (Россия). Зарегистрировано 10.02.2012. Приоритет от 14.12.2011. Заявка № 2011619552.
20. Программный комплекс быстрого решения задач оптимальной комплектации систем / *Мезенцев Ю.А.* // Свидетельство о государственной регистрации № 2012617398 (Россия). Зарегистрировано 16.08.2012. Приоритет от 19.06.2012. Заявка № 2012615085.
21. Программный комплекс оптимального планирования закупок и сбыта продукции / *Мезенцев Ю.А., Павлов П.С.* // Свидетельство о государственной регистрации № 2012611279 (Россия). Зарегистрировано 31.01.2012. Приоритет от 02.12.2011. Заявка № 2011619206.
22. Программный комплекс оптимального управления поставками сырья и комплектующих на предприятии / *Мезенцев Ю.А., Павлов П.С.* // Свидетельство о государственной регистрации № 2012617400 (Россия). Зарегистрировано 16.08.2012. Приоритет от 19.06.2012. Заявка № 2012615087.

Монографии:

23. *Наумов А.А., Мезенцев Ю.А.* Оптимальное управление инвестиционным портфелем. Новосибирск: Издательская компания Лада, 2002, 192 с.
24. *Мезенцев Ю.А.* Математические задачи оптимального управления реализацией проектов. Новосибирск: Изд-во НГТУ, 2013, 142 с.

Доклады на международных конференциях:

25. *Avdeenro T.V., Mezentsev Y.A.* Time-table optimization for parallel-serial production systems on the basis of heuristics methods // Proceedings of DST-RFBR Sponsored Indo-Russian Joint Workshop on Computational Intelligence and Modern Heuristics in Automation and Robotics, 20-22 September 2010, Surat, India, pp. 25-29.

26. *Авдеенко Т.В., Мезенцев Ю.А.* Задачи синтеза оптимальных расписаний нанотехнологических производств // Материалы международной научной конференции «Моделирование 2010» Том 1, Изд-во Института проблем моделирования в энергетике, Киев, 2010., С. 88-96.
27. *Мезенцев Ю.А.* Задачи и алгоритмы оптимизации расписаний параллельных обслуживающих систем с динамическим входом // Материалы X Международной научно-практической конференции «Методы и алгоритмы прикладной математики в технике, медицине и экономике». Новочеркасск, Изд-во ЮРГТУ 2010., С. 13-17.
28. *Мезенцев Ю.А.* Об одной задаче математической логистики // Сборник трудов XV международной открытой конференции «Современные проблемы информатизации в экономике и обеспечении безопасности». Воронеж, Изд-во «Научная книга» 2010. вып.15, С. 22-29.
29. *Avdeenko T.V., Mezentsev Y.A.* Models and algorithms for calendar production planning // Proceedings of RFBR and DST Sponsored “The 2-nd Russian-Indian Joint Workshop on Computational Intelligence and Modern Heuristics in Automation and Robotics”, 10 – 13 September, 2011, Additional vol., pp.11-15.
30. *Мезенцев Ю.А.* Об одном алгоритме решения задач теории расписаний в приложении к календарному производственному планированию // Материалы VI международной конференции «Актуальные проблемы электронного приборостроения» (АПЭП-2002), т.7. Новосибирск 2002, С. 184-190.
31. *Мезенцев Ю.А., Преображенская Т.В.* Модель оптимального выбора проекта по критерию цена/качество (на основе функционально-стоимостного анализа) // Материалы Международной научно-технической конференции «Информационные системы и технологии» ИСТ’2003 Том 1 НГТУ 2003, С. 95-96.

Отпечатано в типографии Новосибирского
государственного технического университета
630073, г. Новосибирск, пр. К. Маркса, 20,
тел./факс (383) 346-08-57

формат 60□84/16, объем 2.5 п. л., тираж 120 экз.,
заказ № , подписано в печать . 2013 г.