

УТВЕРЖДАЮ

Директор Федерального государственного
бюджетного учреждения науки
«Институт математики и механики
им. Н.Н. Красовского Уральского отделения
Российской академии наук» (ИММ УрО
РАН)

доктор физико-математических наук,
член корреспондент РАН,

Николай Юрьевич Лукоянов



« 7 » сентября

2018 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

на диссертацию Гайнанова Дамира Насибулловича «Математическое и программное обеспечение вычислительных комплексов для решения задач анализа несовместных систем с массивно параллельной обработкой данных», представленную на соискание ученой степени доктора физико-математических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей» и специальности 05.13.18 – «Математическое моделирование, численные методы и комплексы программ».

Диссертация Гайнанова Д. Н. посвящена исследованию актуальных оптимизационных проблем, связанных с моделированием производственных процессов, характеризующихся большой размерностью, противоречивостью и нестационарностью ограничений, построением и обоснованием соответствующего алгоритмического аппарата, а также проектированием и разработкой современного вычислительного комплекса.

Задачи логистики и маршрутизации в своих различных экстремальных постановках исследуются в многочисленных работах как отечественных (Э.Х. Гимади, А.И. Кибзуна, Ю.А. Кочетова, С.В. Севастьянова и др.), так и зарубежных авторов (P. Brucker, G. Gutin, N. Mladenovic, P.M. Pardalos, J. Pei, и др.)

Несовместные системы алгебраических ограничений и тесно связанные с ними несобственные задачи математического программирования традиционно являются одними из общепризнанных математических моделей,

ОБЩИЙ ОТДЕЛ МАИ
Вх. №
" 7 " 09 2018

используемых при описании производственных процессов, сопряженных с оптимальным планированием, распределением ограниченных ресурсов, эффективным распределением работ, выбором оптимальных маршрутов доставки продукции и т.п. Поскольку по ряду объективных причин возникающие на этом пути математические задачи оказываются противоречивыми, не обладая решениями в классическом понимании этого слова, одной из первоочередных трудностей, возникающих при анализе подобных моделей, является строгая формализация самого понятия решения.

Широко известен подход к коррекции несобственных экстремальных задач и несовместных систем ограничений, восходящий к работам П.Л.Чебышева, Р.Рокафеллара, И.И.Еремина, Е.Г.Гольштейна и др. в рамках которого часть ограничений несовместной системы (или даже вся система целиком) подвергаются ослаблению или заменяются подходящими штрафными слагаемыми в целевом функционале.

Наряду с классическим подходом с 70-х гг. прошлого века активно развивается альтернативный дискретный подход к обобщению понятия решения, связанный с поиском решений максимальных по включению совместных подсистем (МСП) исходной системы и комбинации их с использованием различных агрегирующих процедур, среди которых выделяется процедура голосования большинством голосов. Немалое влияние на развитие этого подхода оказали фундаментальные результаты С.Н.Черникова в области алгебраической теории линейных неравенств, метод комитетных обобщенных решений Вл.Д.Мазурова, алгебраический подход к проблеме распознавания Ю.И.Журавлева, К.В.Рудакова, Н.Н.Катериночкиной, результаты Д.Джонсона, И.Динур, Н.Мегиддо и др.

В диссертации Гайнанова Д.Н. исследуются важные классы таких задач, в которых множество ограничений на переменные характеризуется крайне высокой размерностью и несовместностью по совокупности. В этой связи одной из задач исследования явилась разработка эффективных методов и вычислительных технологий для реализации параллельной обработки реальных данных большого объема.

Таким образом, **актуальность** темы диссертации и научная ценность полученных результатов не вызывают сомнений.

Краткая характеристика работы

Диссертация состоит из введения с кратким обзором, шести глав, заключения и списка использованной литературы.

Во **введении** обоснована актуальность темы исследования, приводится обзор результатов, полученных другими авторами, включая результаты,

полученные в смежных областях, а также обозначены цель работы и задачи, решение которых влечет достижение поставленной цели. В рамках краткого содержания работы во введении также представлена общая структурная схема вычислительного комплекса и обозначены те его компоненты, разработке которых будут посвящены последующие главы диссертации.

В **первой** главе формализованы понятие монотонных несовместных систем условий и задачи анализа таких систем в терминах задачи распознавания образов в геометрической постановке и в терминах задачи расшифровки монотонных булевых функций (МБФ). Автор показывает, что перечисленные задачи тесно связаны друг с другом, допуская однотипное описание в терминах абстрактных симплициальных комплексов (АСК). Завершает главу краткий обзор известных результатов в области оценивания экстремального числа граней АСК различной размерности.

Вторая глава посвящена разработке теоретико-графовых методов математического моделирования в задачах анализа несовместных систем ограничений. Вводится новое понятие графа системы независимости, обобщающего известный объект - граф максимальных по включению совместных подсистем. В основе данного обобщения лежит свойство монотонности семейства подсистем исследуемого класса систем ограничений и близкое по природе свойство абстрактных симплициальных комплексов.

В рамках исследования структурных особенностей графов систем независимости установлено одно из важнейших их свойств – связность графа системы независимости для широкого класса исследуемых систем. В утверждении 2.4 показывается, что связность графа системы независимости порождается связностью соответствующего топологического пространства, что является достаточно слабым ограничением и свидетельствует тем самым о широком диапазоне класса графов систем независимости, обладающих свойством связности. Для частного случая графа системы независимости — графа максимальных совместных подсистем (МСП) системы однородных линейных неравенств данный результат удается существенно усилить. Так, показывается, что граф МСП обладает циклом нечетной длины, не превосходящей число неравенств (Теорема 2.3), обосновываются условия его k -связности (Теорема 2.2), нижние оценки степеней вершин и верхние оценки диаметра (Утв. 2.12).

Третья глава посвящена исследованию структуры семейств совместных и несовместных подсистем несовместной системы линейных неравенств в терминах комбинаторной теории выпуклых многогранников. Автором вводится новое понятие G -диагонали набора точек евклидова пространства. Показывается, что классификация выпуклых многогранников по типам G -

диагоналей эквивалентна классической классификации по комбинаторному типу решеток граней.

По-видимому, основным результатом данной главы является Теорема 3.1, устанавливающая взаимно однозначную взаимосвязь между индексами максимальных совместных подсистем (минимальных несовместных подсистем) системы однородных линейных неравенств и семействами диагоналей (гиперграней) подходящего выпуклого многогранника и открывающая возможность применения развитого аппарата комбинаторной теории многогранников.

Четвертая глава посвящена разработке вычислительных алгоритмов поиска и подсчета МСП, а также для решения взаимосвязанных задач расшифровки монотонных булевых функций (МБФ) и задач распознавания образов. Алгоритмы поиска и подсчета всех МСП представлены в своих различных вариантах, включая приближенные алгоритмы и их экономную реализацию с точки зрения сложности вычислений. В основе всех этих алгоритмов лежат свойства графов систем независимости такие, как связность графа МСП и существование в нем цикла нечетной длины.

Наиболее яркими результатами данной главы представляются алгоритмы расшифровки МБФ, в том числе, порождаемой несовместной системой ограничений общего вида. Так, заключительный раздел главы содержит описание алгоритмов, обладающих рекордно низкой трудоемкостью в терминах числа обращений к оракулу (Утв. 4.6 и 4.8).

Пятая глава посвящена исследованию прикладных задач анализа несовместных систем. Задача планирования грузовых железнодорожных перевозок в условиях противоречивости заданного множества допустимых расписаний приводится в своей известной постановке и сводится к задаче расшифровки МБФ, индуцированной независимыми множествами заданного неориентированного графа. Такой подход был впервые разработан в совместных работах автора и представлен в диссертации в качестве обзора результатов в этой области. Следующий этап решения задачи состоит в назначении локомотивов на исполнение перевозок, заданных некоторым допустимым расписанием, и характеризуется крайне высокой комбинаторной сложностью. Для реализации этого этапа в пятой главе разработана методология параллельной обработки данных на множестве путей ориентированного графа, а также приводятся результаты вычислительных экспериментов с использованием этой методологии на тестовых данных большой размерности.

Другой класс прикладных задач анализа несовместных систем условий исследуется в контексте задачи управления технологическими процессами.

Поиск (суб)оптимального технологического маршрута реализован в виде распределенной вычислительной системы, использующей среди прочего, обучение по исторической базе прецедентов. Лежащий в ее основе математический аппарат опирается на результаты предыдущих глав в области поиска максимальных совместных подсистем и построения комитетных обобщенных решений малой мощности. Для эффективной реализации этапа классификации входного вектора (текущего технологического маршрута) в пятой главе разработан метод параллельных (независимых) вычислений, позволяющий сократить время получения решения, что особенно важно в условиях большой размерности технологической базы производства и непрерывного ее пополнения новыми данными.

В **шестой** главе приведено описание структуры разработанного вычислительного комплекса. Комплекс состоит из нескольких взаимодействующих программных компонент, названных в работе «серверами»: серверов сбора и хранения данных, сервера управляющих программ, сервера математических моделей, серверов планирования и принятия решений. Математический аппарат, лежащий в основе основных компонент комплекса, всецело опирается на результаты представленные в предыдущих главах работы, являясь по сути их программной реализацией.

Глава посвящена описанию основных функций и схемы взаимодействия разработанных программных модулей в рамках общей архитектуры вычислительного комплекса.

В **заключении** приводятся основные результаты, полученные в диссертации и выносимые на защиту.

Научная новизна. В диссертации получены следующие новые результаты:

1. Разработаны математическое и программное обеспечение вычислительных комплексов в составе управляющих программ и проблемно-ориентированных комплексов программ для решения прикладных оптимизационных задач (управление технологическими и транспортными процессами) в условиях большой размерности и высокой комбинаторной сложности, реализующие разработанные методы параллельной обработки данных различной структуры, а также алгоритмы решения задач анализа несовместных систем в рамках взаимосвязи с расшифровкой МБФ и распознаванием образов в геометрической постановке.
2. Разработана методология параллельной обработки данных на сети задач распознавания образов в геометрической постановке. Для задачи оптимизации управления технологическим маршрутами на дискретном

производстве разработаны полиномиальные эвристические алгоритмы дихотомии обучающей выборки. Эффективность предложенного подхода подтверждается результатами вычислительных экспериментов на больших наборах случайных данных.

3. Разработана методология параллельной обработки данных с использованием алгоритма декомпозиции множества путей ориентированного графа на множестве пересекающихся сильно связанных подграфов. Эффективность подхода подтверждается на реальных задачах управления транспортными процессами большой размерности.
4. Разработан новый метод математического моделирования монотонных несовместных систем, в основе которого лежит понятие графа системы независимости. Доказана связность такого графа для некоторых классов систем, непосредственно вытекающая из связности топологического пространства. Для графа МСП системы линейных неравенств, частного случая графа системы независимости, доказана теорема о существовании нечетного цикла, обоснованы оценки степеней его вершин и диаметра, а также достаточные условия более сильных типов связности. На основе полученных результатов построены комбинаторные и комбинаторно-графовые алгоритмы построения графа МСП.
5. Впервые доказана теорема о взаимосвязи нечетного цикла в графе МСП с комитетом исходной несовместной системы линейных неравенств и разработан новый приближенный алгоритм синтеза минимального комитета. Для классификации комитетов равной мощности разработан новый критерий, в основе которого лежит понятие альтернативных покрытий.
6. Разработан новый метод математического моделирования монотонных несовместных систем линейных неравенств, в основе которого лежит понятие G -диагонали выпуклого многогранника. Доказана теорема о двойственной взаимосвязи между семейством G -диагоналей и семейством дополнений МСП соответствующей системы. На основе взаимосвязи, установленной в этой теореме, получены нижние оценки для максимального числа МСП.
7. Разработан новый критерий оптимальной расшифровки МБФ, связанной с несовместной системой условий общего вида, учитывающий объективную сложность расшифровки конкретной МБФ. Предложен алгоритм, оптимальный по этому критерию для МБФ, порождаемых несовместными системами линейных неравенств. Получены результаты вычислительных экспериментов с использованием эвристического алгоритма расшифровки МБФ,

порожденных неориентированным графом на базе примеров из открытой библиотеки DIMACS, подтверждающие качество получаемых решений и значительное сокращение времени расчетов в сравнении с известными алгоритмами.

Практическая ценность

Результаты, полученные в диссертации, могут быть использованы специалистами в области математического и программного обеспечения вычислительных комплексов, машинного обучения и распознавания образов, кластерного анализа, теории графов и комбинаторной оптимизации, математического программирования и исследования операций, а также при подготовке соответствующих специальных курсов для магистрантов и аспирантов. Разработанный в диссертации вычислительный комплекс с массивно параллельной обработкой данных может быть использован для решения актуальных оптимизационных задач в условиях противоречивости, например, в рамках Свердловской железной дороги и металлургических предприятий Урала и Южного Урала (Нижнетагильский металлургический комбинат, металлургические предприятия Свердловской и Челябинской областей).

Личный вклад. Все исследования, результаты которых изложены в диссертационной работе, выполнены лично соискателем. Из совместных публикаций в диссертацию вошли лишь те результаты, которые принадлежат лично соискателю.

Работа прошла всестороннюю апробацию и неоднократно докладывалась на международных и всероссийских конференциях. По результатам диссертации опубликованы две монографии, 20 статей в рецензируемых изданиях, индексируемых международными базами Scopus и Web of Science, в том числе 13 статей в журналах, входящих в перечень ВАК и 2 статьи в журналах, входящих только в перечень ВАК, 4 свидетельства о государственной регистрации программ для ЭВМ и 4 патента на изобретения.

Достоверность результатов, представленных в диссертационной работе, подтверждается адекватными математическими моделями, строгими математическими доказательствами, проведенными вычислительными экспериментами, подтверждающими эффективность предлагаемых алгоритмов.

Замечания

1. Высокая эффективность предложенной в главе 5 автоматической системы прогнозирования бракованных изделий, основанная на дообучении в реальном времени комитетного классификатора, представляется крайне любопытной. Подобного типа процедуры являются предметом изучения одного из активно развивающихся разделов современной теории анализа данных — обучения с подкреплением (reinforcement learning). Представляет интерес сравнение полученных в работе результатов с последними достижениями в данной области.
2. Приведенные в главе 5 эвристические алгоритмы дихотомии обучающей выборки для решения многоклассовой задачи распознавания образов представляются разновидностью другого широкого класса алгоритмов обучения — решающих деревьев. В этой связи представляется уместным проведение (теоретического или численного) сравнения данных алгоритмов с известными представителями этого класса: C4.5, C5.0, RF др.
3. Утверждение об оптимальности предложенного в главе 4 алгоритма расшифровки монотонной булевой функции, порожденной несовместной системой линейных неравенств представляется не вполне корректным. Традиционно трудоемкость подобных алгоритмов оценивается в терминах т. н. «оракульной» сложности - числа обращений в подпрограмме-оракулу, вычисляющей значение восстанавливаемой функции к конкретной вершине булева куба. Известна простая нижняя оценка такой сложности, совпадающая с числом верхних нулей и нижних единиц функции. Автором показано, что число обращений к оракулу предложенного им алгоритма совпадает с указанной нижней оценкой. Однако оракул, используемый в данном алгоритме, имеет существенно более сложную структуру и предположительно большую трудоемкость, что затрудняет объективную интерпретацию данного результата.
4. В этой же главе представлен быстрый эвристический алгоритм поиска максимального независимого множества графа. Показано, что алгоритм является точным в классе ациклических графов. Предположительно высокая точность в классе произвольных графов подтверждена результатами численных экспериментов на графах из библиотеки DIMACS. Кроме того, утверждается, что данный алгоритм обладает «абсолютной оценкой точности», поскольку для каждого конкретного графа наряду с полученным независимым множеством алгоритм

позволяет оценить отклонение его мощности от оптимума задачи. Последнее утверждение представляется терминологически не вполне обоснованным, поскольку термин «алгоритм с оценкой точности» предполагает равномерность данной оценки для некоторого нетривиального класса графов.

Отметим, что высказанные замечания в целом носят рекомендательный характер и не снижают общей высокой квалификационной оценки работы.

Заключение

Диссертация Гайнанова Д.Н. «Математическое и программное обеспечение вычислительных комплексов для решения задач анализа несовместных систем с массивно параллельной обработкой данных» представляет собой законченную и целостную научно-исследовательскую работу, выполненную на высоком научном уровне. В ней получены новые результаты, обладающие высокой научной и практической значимостью. Совокупность полученных теоретических результатов можно квалифицировать как решение крупной научной проблемы, а внедрение полученных практических результатов вносит существенный вклад в решение важных прикладных задач в области оптимального управления производственными и транспортными процессами в условиях противоречивости ограничений на использование ресурсов инфраструктуры.

Таким образом, **диссертация удовлетворяет** всем требованиям п.9 «Положения о порядке присуждения ученых с степеней», утвержденного Постановлением Правительства РФ от 24 сентября 2013 года №842 и 21 апреля 2016 года №335, предъявляемым к диссертациям на соискание ученой степени доктора физико-математических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей» и по специальности 05.13.18 – «Математическое моделирование, численные методы и комплексы программ», а ее автор, **Гайнанов Дамир Насибуллович, заслуживает присуждения искомой степени.**

Отзыв на диссертацию и автореферат обсужден на семинаре отдела математического программирования 24 августа 2018 г., протокол № 297.

Заведующий отделом математического
программирования ИММ УрО РАН,
доктор физико-математических наук,
профессор РАН

Подпись завершено
Ученый секретарь
ИММ УрО РАН



М. Ю. Хачай

Ульянов О.Н.