

ОТЗЫВ

официального оппонента

Левановой Татьяны Валентиновны

на диссертацию Рожнова Ивана Павловича

«Алгоритмы поиска с чередующимися рандомизированными окрестностями для задач автоматической группировки объектов», представленной на соискание ученой степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (космические и информационные технологии)

Актуальность темы диссертационной работы

Предметом исследования диссертационной работы Рожнова Ивана Павловича являются алгоритмы автоматической группировки (кластеризации) с комбинированным применением алгоритмов поиска с чередующимися рандомизированными окрестностями и жадных эвристических алгоритмов автоматической группировки. Указанные алгоритмы и их параллельные модификации позволяют существенно расширить рамки применения метода жадных эвристик на случаи задач большой размерности, содержащих до сотен тысяч векторов многомерных данных при дополнительных повышенных требованиях к точности получаемого результата.

Проведенный соискателем анализ существующего программного обеспечения для решения рассматриваемой задачи выявил дефицит алгоритмов, способных стабильно получать близкие к оптимальным решения за относительно небольшое время. Вопросы создания быстрых и при этом достаточно точных алгоритмов решения задач автоматической группировки объектов, применимых к решению задач с большим объемом входных данных, к которым предъявляются высокие требования по точности и стабильности результата (достигаемого значения целевой функции за заданное время), остаются недостаточно исследованными. Это дает основание утверждать, что научная проблема, сформулированная в диссертации Рожнова И.П., является актуальной.

Структура диссертационной работы

Диссертационная работа состоит из введения, четырех глав, заключения, списка литературы и приложений. Объем работы составляет 176 страниц. Список литературы содержит 255 наименований.

Во введении автор обосновывает актуальность, описывает степень разработанности темы, указывает цель и поставленные задачи, формулирует новые научные результаты и положения, выносимые на защиту.

Первая глава диссертации Рожнова И.П. посвящена анализу текущего состояния и развития методов и задач автоматической группировки во взаимосвязи с задачами теории размещения. Обозначены проблемы, возникающие при решении задач кластеризации объектов с повышенными требованиями к точности и стабильности получаемого результата.

Во второй главе описана разработка комбинированных алгоритмов метода жадных эвристик для задач автоматической группировки с совместным применением алгоритмов поиска с чередующимися рандомизированными окрестностями, а также параллельных жадных эвристических алгоритмов автоматической группировки для массивно-параллельных систем применительно к задаче k -средних. Предложены параллельные модификации алгоритмов метода жадных эвристик для архитектуры CUDA, которые показывают преимущество в скорости по сравнению с вычислениями на центральном процессоре. По результатам вычислительных экспериментов показано, что новые алгоритмы за фиксированное время позволяют получать более точный и стабильный результат (по достигаемому значению целевой функции) в сравнении с известными алгоритмами автоматической группировки.

В третьей главе описана разработка комбинированных алгоритмов метода жадных эвристик для задач автоматической группировки с повышенными требованиями к точности и стабильности результата с применением алгоритмов поиска с чередующимися рандомизированными окрестностями применительно к задаче k -медоид и максимизации функции правдоподобия математического ожидания. Сформулирован новый подход к разработке эффективных алгоритмов автоматической группировки, основанный на сочетании известных алгоритмов локального поиска, жадных агломеративных эвристических процедур и алгоритмов поиска с чередующимися рандомизированными окрестностями. Для его реализации предложен оригинальный способ построения окрестностей с использованием одной из жадных агломеративных эвристических процедур.

Четвертая глава посвящена построению процедуры составления оптимальных ансамблей алгоритмов автоматической группировки с комбинированным применением генетического алгоритма метода жадных эвристик и согласованной матрицы бинарных разбиений для практических задач. Предложенная процедура позволяет уменьшить число ошибок при разделении сборной партии промышленной продукции на однородные партии с использованием данных неразрушающих тестовых испытаний.

В заключении сформулированы основные выводы и результаты работы, показано, что цель диссертации достигнута путем решения поставленных задач.

В приложения помещены дополнительные данные о результатах вычислительных экспериментов, акты об использовании результатов исследования и Свидетельство о государственной регистрации программы для ЭВМ.

Степень достоверности и обоснованность научных положений, выводов и рекомендаций

В диссертации Рожнова Ивана Павловича довольно подробно проанализирована отечественная и зарубежная литература по задачам автоматической группировки, теории размещения и эволюционным алгоритмам. Автор достаточно корректно использует известные научные

методы обоснования полученных результатов, выводов и рекомендаций. Автором изучены и критически анализируются известные достижения и теоретические положения других авторов по исследуемым вопросам.

Достоверность экспериментальных данных обеспечивается использованием современных средств и методик проведения исследований. С целью подтверждения преимуществ новых алгоритмов перед известными разработками проведено достаточное количество экспериментов. Основные положения и результаты диссертационной работы апробированы на конференциях и опубликованы в 18 работах (7 работ в ведущих рецензируемых журналах, рекомендуемых в действующем перечне ВАК, 5 – в международных изданиях, индексируемых в системах цитирования Web of Science и Scopus). Имеется свидетельство о государственной регистрации программы для ЭВМ. Разработанные автором новые алгоритмы были внедрены в эксплуатацию на двух промышленных предприятиях.

Оценка новизны и достоверности

В качестве новых научных результатов диссертантом предложены:

1) новый подход к разработке алгоритмов автоматической группировки, основанных на параметрических оптимизационных моделях, с комбинированным применением алгоритмов поиска с чередующимися рандомизированными окрестностями и жадных агломеративных эвристических процедур. Показано, что применение данного подхода позволяет создавать эффективные алгоритмы автоматической группировки (по достигаемому значению целевой функции за фиксированное время);

2) параллельные модификации алгоритмов с жадной агломеративной эвристической процедурой для больших задач автоматической группировки, адаптированные к архитектуре CUDA. Было выявлено, что параллельная реализация алгоритма локального поиска, а также отдельных шагов жадной агломеративной эвристической процедуры, позволяет построить алгоритм автоматической группировки с высоким коэффициентом ускорения, сокращающим время расчетов в десятки раз без ухудшения достигаемого значения целевой функции;

3) процедура составления оптимальных ансамблей алгоритмов автоматической группировки с комбинированным применением генетического алгоритма метода жадных эвристик и согласованной матрицы бинарных разбиений для практических задач. Было выявлено, что точность разделения сборной партии промышленной продукции с особыми требованиями качества на однородные партии, выполненного с применением получаемых ансамблей, выше усредненной точности разделения с применением отдельных алгоритмов, отобранных для составления ансамбля.

Диссертантом был проведен значительный объем научной работы. Результаты получены лично автором и полностью опубликованы в рецензируемых научных изданиях, прошли апробацию на конференциях различного уровня и внедрены в производственный процесс.

Научная значимость диссертационных исследований состоит в разработке нового подхода к созданию алгоритмов автоматической группировки, основанных на параметрических оптимизационных моделях, с комбинированным применением алгоритмов поиска с чередующимися рандомизированными окрестностями и жадных агломеративных эвристических процедур, а также в построении новой процедуры составления оптимальных ансамблей алгоритмов кластеризации.

Диссертационная работа Рожнова И.П. изложена на 176 страницах, состоит из введения, четырех глав и заключения. Содержание диссертации и полученные результаты соответствуют п. 4 паспорта специальности 05.13.01 – Системный анализ, управление и обработка информации.

Выявленные недостатки и замечания

1. Математическая модель задачи k -средних (стр. 19, формула (1.1)) содержит опечатку: суммирование ведётся по i , но в формуле написаны индексы j .

2. Суть алгоритма с чередующимися окрестностями состоит в специальном переборе окрестностей. На стр. 31 правильнее было бы написать, что излагается суть алгоритма, основанного на поиске с переменными окрестностями, поскольку указан другой перебор.

3. На рис 1.1 (стр. 33) не отмечено, в каких случаях осуществляется переход по стрелкам, это затрудняет понимание. Нужно было описать схему более подробно.

4. Блок-схема алгоритма VNS приведена с ошибкой (рис. 2.2, стр. 43). При $k > k_{max}$ (идем по стрелке «нет») невозможно выбрать точку x' .

5. В алгоритмах 2.5, 3.1 и 3.4 решение S не может принимать числовое значение 1, 2 или 3.

6. В описании метрик (стр. 85) ошибочно присутствуют булевы переменные $x_1, \dots, x_N$.

7. Текст содержит некоторое количество опечаток. Например, стр. 30, первая строка: $k=1, \dots, k_{max}$ – пропущены точка и запятая; стр. 43, последний абзац: неправильные падежи (в окрестностях известного промежуточноГО решениЯ); стр. 16, второй абзац снизу: некорректное название института (правильное название – им. С.Л. Соболева СО РАН); и другие.

Заключение о соответствии диссертации требованиям и критериям, установленным Положением о порядке присуждения ученых степеней

Работа базируется на достаточном объеме исходных данных, примеров и расчетов. По каждой главе и работе в целом сделаны четкие выводы. Вышеуказанные замечания не снижают положительной оценки работы. Диссертация «Алгоритмы поиска с чередующимися рандомизированными окрестностями для задач автоматической группировки объектов» является завершенным научно-исследовательским трудом, выполненным автором самостоятельно и на высоком научном уровне.

Автореферат достаточно полно и правильно отражает основное содержание диссертации. Оформление автореферата и диссертации соответствует требованиям ВАК РФ.

Представленная диссертационная работа отвечает требованиям п. 9 «Положения о порядке присуждения ученых степеней» постановления Правительства Российской Федерации от 24.09.2013 г. № 842, предъявляемым к кандидатским диссертациям, а ее автор Рожнов Иван Павлович заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (космические и информационные технологии).

Официальный оппонент,
старший научный сотрудник,
Омский филиал Института математики
им. С.Л. Соболева СО РАН,
кандидат физ.-мат. наук, доцент

Леванова Татьяна Валентиновна

11.09.2019

Адрес организации:
644043, г. Омск, ул. Певцова, 13
e-mail: levanovat@mail.ru

11.09.2019

Подпись
Левановой Татьяны Валентиновны
заверяю
Ученый секретарь



Планкова В.А.