

ОТЗЫВ

официального оппонента на диссертацию и автореферат Сташкова Дмитрия Викторовича «Системы автоматической группировки объектов на основе разделения смеси распределений», представленную на соискание ученой степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (космические и информационные технологии)

Актуальность темы работы

Задачи автоматической группировки объектов (иначе - задачи кластеризации) - один из наиболее распространенных классов задач машинного обучения без учителя. Эти задачи NP-трудны и методология их решения в основном развивается в сторону разработки алгоритмов поиска приближенного решения, что для динамических систем во многих случаях представляется вполне оправданным. Практика показывает, что эвристические и метаэвристические алгоритмы способны давать вполне удовлетворительные для практического применения результаты – с минимальными временными затратами и сравнительно небольшой ошибкой.

Для некоторых практических задач требуется строгая регламентация процесса решения, соответственно, результаты используемых алгоритмов должны быть не только близки к оптимуму, но и стабильны при многократных перезапусках алгоритма.

Одним из подходов к группировке объектов является рассмотрение совокупности параметров объектов одной группы как многомерной случайной величины, распределенной по некоторому известному закону распределения, и отделение объектов, предположительно порожденных одним из распределений в смеси, от других.

Диссертация посвящена разработке новых алгоритмов решения задач автоматической группировки объектов, основанных на разделении смеси распределений.

Структура работы

Введение диссертации посвящено обоснованию актуальности темы диссертации, формулировке цели и поставленных задач, а также положений, выносимых на защиту.

В первой главе проведен обзор постановок задач автоматической группировки и спектра методов их решения. Для задач автоматической группировки предложен подход из вероятностной диагностики, задача

автоматической группировки поставлена как задача разделения смеси распределений.

Вторая глава посвящена разработке алгоритмов решения задачи автоматической группировки на основе разделения смеси распределений и жадной агломеративной эвристики. Проводится анализ их эффективности по сравнению друг с другом и с известными алгоритмами.

В третьей главе проводится разработка модели разделения сборной партии промышленной продукции по результатам неразрушающих испытаний на примере партии электронных компонентов. Показана возможность применения модели разделения сборной партии изделий на основе разделения смеси гауссовых распределений.

В заключении диссертации подведены итоги, сформулированы основные результаты и сделаны выводы по работе.

Достоверность и обоснованность научных положений, выводов и рекомендаций

Автором проведен подробный анализ отечественной и зарубежной литературы по задачам кластеризации и методам разделения смеси распределений. Результаты и выводы подтверждены достаточным количеством численных экспериментов, которые подтверждают заявленные преимущества новых алгоритмов перед известными методами. Кроме того, результаты работы опубликованы в рецензируемых научных изданиях и апробированы на конференциях. Разработанные алгоритмы внедрены в опытную эксплуатацию на промышленном предприятии.

Новизна научных результатов и выводов

Новыми научными результатами диссертации Сташкова Д.В. являются:

- генетические алгоритмы, основанные на идеях метода жадных эвристик для задач разделения смеси распределений;
- алгоритмы поиска с чередующимися окрестностями для задач разделения смеси распределений;
- модель разделения сборных партий промышленных изделий на основе разделения смеси некоррелированных гауссовых распределений.

Результаты получены лично автором и полностью опубликованы.

Значимость полученных результатов для науки и практики

Работа дополняет арсенал методов решения задач автоматической группировки, а также дополняет область применения метода жадных эвристик задачами разделения смеси распределений.

Практическая ценность разработанных алгоритмов заключается, в том числе, в снижении процента ошибок при выявлении однородных партий промышленных изделий за счет применения разработанной модели разделения смеси распределений.

Общая оценка работы

Диссертация с приложениями изложена на 172 страницах состоит из введения, трех глав и заключения и включает 208 библиографических источников. Диссертация обладает внутренним единством, стиль изложения материала ясный. Содержание диссертации и полученные результаты соответствуют п.4 паспорта специальности 05.13.01 — системный анализ, управление и обработка информации.

Выявленные недостатки и замечания

1. Автор ничем не подкрепляет выбор количества запусков алгоритмов (стр. 15 автореферата): «Для большинства наборов данных было выполнено по 30 попыток запуска каждого из алгоритмов».

2. Результаты вычислительных экспериментов представлены в виде значений целевой функции для наборов данных при указанном числе групп (кластеров). В тексте не указывается истинное значение групп в обрабатываемых данных. Не понятно, чем обоснован выбор числа кластеров.

3. Не полностью раскрыта тема способа построения систем автоматической группировки. Не ясно, что понимается под *системой автоматической группировки*.

4. В постановке цели исследования и решения задач говорится о том, что в работе ставится задача построения алгоритмов на основе разделения смеси распределений в пространствах большой размерности (до сотен измерений). На сегодняшний день данный объем разделяемых данных нельзя назвать большим.

5. Не обоснован выбор типа генетического алгоритма — с вещественным алфавитом.

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

Имеющиеся замечания не снижают общей положительной оценки работы. Считаю, что диссертационная работа на соискание ученой степени кандидата технических наук Сташкова Д.В. является завершенной научно - квалификационной работой на актуальную тему, выполненной на высоком научном уровне. Полученные автором результаты достоверны, выводы и заключения являются обоснованными.

Автореферат соответствует основному содержанию диссертации, содержит описание основных этапов исследования, полученные результаты и выводы. Оформление автореферата и диссертации соответствует требованиям ВАК РФ.

Представленная диссертационная работа отвечает требованиям действующего Положения о порядке присуждения ученых степеней, предъявляемым к кандидатским диссертациям, а ее автор Сташков Дмитрий Викторович заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.01 — системный анализ, управление и обработка информации (космические и информационные технологии).

Официальный оппонент
доктор технических наук,
профессор, профессор кафедры
«Управление персоналом»
Красноярского института
Железнодорожного транспорта —
филиала ФГБОУ ВО ИрГУПС

Терсков Виталий Анатольевич

20.11.2017г.

КРИЖТ, филиал ФГБОУ ВО ИрГУПС
660028, Красноярский край, г. Красноярск, ул. Ладо Кецховели, 89
Тел. 8 (3912) 248-16-44, доб.20-37
E-mail:terskovva@mail.ru

Подпись
вер. спец. Б.Е.

