

ОТЗЫВ

на автореферат диссертации Ахматшина Фариды Галиулловича
«Модели и алгоритмы автоматической группировки объектов для систем
анализа и хранения данных на основе методов семейства k -средних»,
представленной на соискание ученой степени кандидата технических наук по
специальности 2.3.1 – Системный анализ, управление и обработка
информации, статистика

В рамках роста накапливаемых данных современное и будущее развитие технологий анализа и хранения информации связано, в том числе, с использованием методов автоматической группировки (кластеризации) данных и приближенного поиска ближайших соседей (ANN – Approximate Nearest Neighbors) как ценного источника новых знаний.

Исследования, выполненные автором, позволили получить важные результаты, обладающие научной новизной:

1. Подход к нормализации данных для предобработки входных данных, используемых в системах анализа, обеспечивает повышение точности решения задачи кластеризации на 10% по индексу Рэнда, на примере задачи автоматической группировки электрорадиоизделий.

2. Алгоритм кластеризации для системы анализа данных электрорадиоизделий, обеспечивает повышение точности (по индексу Рэнда) и скорости получения результатов автоматической группировки по сравнению с алгоритмом k -средних.

3. Алгоритм кластеризации для построения индекса векторной базы данных, обеспечивает компромисс между точностью и временем вычислений, существенно улучшает метрику полноты в задачах приближенного поиска ближайших соседей.

4. Алгоритм автоматической группировки повторяющихся фрагментов блоков данных, обеспечивает увеличение эффективности сжатия данных в системах хранения данных.

5. Процедура инициализации центров кластеров, снижает вычислительные затраты в сравнении с алгоритмом k -means++ без снижения качества получаемого начального решения.

Замечания:

1. Нет пояснений, что в формуле (2) обозначено p , каким образом определяются значения коэффициента t . Рассматривались ли другие значения этого коэффициента?

2. В таблицах 1 и 2 не все значения размаха и вариации соответствуют результатам расчетов этих показателей по приведенным в таблицах числовым данным.

Замечания не влияют на положительную оценку диссертационного исследования. Работа Ахматшина Ф.Г. «Модели и алгоритмы автоматической группировки объектов для систем анализа и хранения данных на основе методов семейства k -средних» соответствует критериям, установленным Положением о присуждении ученых степеней, предъявляемым к кандидатским диссертациям. Автор работы Ахматшин Фарид Галиуллович заслуживает присуждения ученой степени кандидата технических наук по специальности 2.3.1 – Системный анализ, управление и обработка информации, статистика.

Ведущий специалист по
АСУ ТП горнорудной
Промышленности АО «СИНЕТИК»,
Кандидат технических наук

11 марта 2025 г.



Сташков Дмитрий Викторович

Подпись Сташкова Д.В. заверяю:
Директор инженерного центра АО «СИНЕТИК»



А.А. Калашников

Телефон: (383)-266-51-40
e-mail: stashkov@sinetic.ru
Сайт в сети Интернет: www.sinetic.ru

