

ОТЗЫВ

на автореферат диссертации Казаковцева Владимира Львовича
«Алгоритмы ускоренного поиска в векторных базах данных»,
представленной на соискание ученой степени кандидата технических наук по
специальности 2.3.1 – Системный анализ, управление и обработка
информации, статистика

В соответствии с содержанием автореферата, диссертационная работа В.Л. Казаковцева посвящена решению значимой научно-практической задачи, связанной с созданием эффективных методов приближенного поиска ближайших соседей (ANN) в векторных базах данных, характеризующихся высокой размерностью признакового пространства и значительными объемами хранимой информации. Развитие ANN-алгоритмов представляет несомненный интерес для развития современных информационных технологий, в том числе рекомендательных систем, средств компьютерного зрения и инструментов обработки мультимодальных данных.

К числу основных результатов, определяющих научную новизну исследования, автор относит предложенный классификатор сложности запросов для IVF-индекса, основанный на анализе доли кластеров, оказывающихся результативными на ранних стадиях поиска, а также адаптивный алгоритм поиска, использующий результаты такой классификации. Кроме того, в работе получили дальнейшее развитие эволюционные подходы к решению задачи k -средних, необходимых для построения IVF-индекса, а также предложена новая модель кластеризации мультимодальных данных.

Для обоснования полученных выводов использованы методы математической статистики, включая t -критерий и критерий Манна-Уитни-Уилкоксона, методы компьютерного эксперимента на крупных массивах данных, а также подходы системного анализа и теории оптимизации. Представленные результаты подтверждают прикладную значимость предложенных решений, в том числе для векторных баз данных, содержащих сотни миллионов объектов.

Теоретическая и практическая ценность выполненного исследования подтверждается апробацией результатов в рамках государственных проектов и работ по хозяйственным договорам, а также наличием зарегистрированных программных продуктов. Существенным подтверждением степени проработанности темы является и публикационная активность автора: по материалам диссертации опубликовано 24 научные работы, включая свидетельства о регистрации ЭВМ и статьи в изданиях, рекомендованных ВАК, индексируемых в Web of Science и Scopus.

Несмотря на перечисленные достоинства, в работе можно найти некоторые недостатки:

1. Упоминается использование статистических критериев, но не указаны их конкретные значения.

2. В модели мультимодальной кластеризации веса модальностей α_i предлагается задавать экспертным путём. Отсутствует обсуждение методов автоматической настройки этих весов (например, по вкладу в разделимость кластеров), что снижает универсальность подхода для полностью неразмеченных данных.

3. Для построения классификатора неинформативность признаков, основанных на расстояниях, требует большего подтверждения.

Тем не менее, диссертационная работа В. Л. Казаковцева соответствует требованиям, предъявляемым к кандидатским диссертациям по специальности 2.3.1, а ее автор заслуживает присуждения ему ученой степени кандидата технических наук.

Доктор физико-математических наук,
доцент, главный научный сотрудник,
директор филиала,

14 05 2026

Омский филиал Федерального
государственного бюджетного
учреждения науки Института
математики им. С.Л. Соболева
Сибирского отделения Российской
академии наук

Еремеев Антон Валентинович

Адрес организации:
ул. Певцова, 13.
г. Омск, 644099, Россия
e-mail: eremeev@ofim.oscsbras.ru
Телефон: +7(3812)236567