

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

Актуальность темы исследования, представленного в автореферате В.Л. Казаковцева, обусловлена стремительным ростом объемов неструктурированных данных и необходимостью совершенствования механизмов их обработки в рамках векторных баз данных. В условиях работы с массивами информации, исчисляемыми сотнями миллионов и миллиардов объектов, задача повышения быстродействия алгоритмов приближенного поиска ближайших соседей (ANN) без потери точности становится критически важной для современных высокопроизводительных систем.

В работе предложен оригинальный подход к классификации запросов по степени сложности с использованием IVF-индекса. Научный интерес представляет разработанный автором метод оценки доли эффективных (результативных) кластеров на начальных этапах поиска, что позволяет динамически определять параметры процесса и избегать избыточных вычислений. Особо следует отметить отказ от использования дистанционных статистик, доказавших свою малоинформативность в пространствах высокой размерности, и переход к адаптивной стратегии на основе обучаемого классификатора. Предложенный алгоритм характеризуется высокой эффективностью: при простоте реализации классификатор обеспечивает точность на уровне 0,81, достигая баланса между скоростью и точностью поиска.

Теоретическая значимость работы заключается в развитии методов системного анализа применительно к индексированию и поиску в крупномасштабных векторных пространствах. Автором предложены новые подходы к определению оптимального объема поискового пространства на основе минимальной выборки обработанных данных. Практическая ценность результатов подтверждается возможностью внедрения разработанных алгоритмов в современные СУБД для ускорения обработки сложных запросов. О высокой востребованности темы также свидетельствуют опубликованные работы, участие в конференциях, а также поддержка исследований грантами и их апробация в рамках взаимодействия с технологическими компаниями.

В качестве замечания к автореферату следует указать следующее:

1. Отсутствие детального анализа влияния размерности векторов на стабильность работы обученного классификатора сложности запросов.

2. В тексте недостаточно полно освещен вопрос влияния распределения данных (плотности кластеров) на точность прогнозирования параметра n_{res}

3. Не приводится прямого сравнения с алгоритмами ANN поиска, помимо стандартного IVF.

В целом, судя по автореферату, диссертация Казаковцева В.Л. представляет собой завершенное научное исследование, содержащее новые научные результаты, имеющие существенное значение для теории и практики системного анализа, управления и обработки информации, полностью удовлетворяет требованиям ВАК, предъявляемым к диссертациям на соискание ученой степени кандидата технических наук по специальности 2.3.1 – Системный анализ, управление и обработка информации, статистика, а ее автор заслуживает присуждения указанной степени.

Кандидат технических наук,
доцент института
прикладных компьютерных
наук Университета ИТМО
19.05.2026

Муравьев Сергей Борисович

Адрес организации:
Кронверкский проспект, 49
литер А.
г. Санкт-Петербург, 197101,
Россия
e-mail: smuravyov@itmo.ru
Телефон: +7(900)6442206