

УТВЕРЖДАЮ

Проректор по науке и инновациям
ФГБОУ ВО «Воронежский
государственный технический
университет»
доктор технических наук, доцент

Башкиров А.В.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ ФГБОУ ВО «Воронежский государственный технический университет»

на диссертацию и автореферат Казаковцева Владимира Львовича на тему «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

1. Актуальность исследования

Современные информационные системы сталкиваются с экспоненциальным ростом объемов и сложности данных, особенно в виде высокоразмерных векторных представлений (эмбедингов). Точные алгоритмы поиска ближайших соседей становятся неприменимыми из-за вычислительной сложности, что создаёт острую потребность в эффективных методах приближенного поиска (ANN). Эти методы являются критически важным компонентом для семантического поиска, рекомендательных систем, компьютерного зрения, обработки естественного языка, обнаружения аномалий и анализа научных данных. Только приближенные подходы позволяют выполнять подобные запросы в реальном времени, иногда сокращая время отклика на несколько порядков при допустимом снижении точности.

Особое место среди методов ANN занимают алгоритмы на основе инвертированного файлового индекса (IVF). Их эффективность напрямую зависит от качества решения задачи кластерного анализа данных и от стратегии поиска внутри индекса. Существующие подходы часто используют

субоптимальные методы кластеризации и неадаптивные стратегии поиска, что открывает значительное поле для их совершенствования.

Параллельно усложняется природа самих данных. Все чаще требуется обработка мультимодальных данных - комплексных объектов, представленных в нескольких разнородных форматах одновременно. Существует явный дефицит теоретических моделей и практических инструментов для прямой кластеризации и поиска в мультимодальном пространстве, что сдерживает развитие целого класса интеллектуальных систем.

Таким образом, актуальность диссертационного исследования обусловлена критической зависимостью современных систем от эффективных алгоритмов приближенного поиска ближайших соседей, наличием существенных резервов для повышения эффективности популярных IVF-методов, острой практической потребностью в новых подходах для работы с мультимодальными данными и междисциплинарным характером проблематики, находящейся на стыке теории алгоритмов, машинного обучения и анализа данных. Работа, направленная на развитие эволюционных алгоритмов для IVF-индексов, создание адаптивных методов поиска и предложение новой модели для мультимодальной кластеризации, является своевременным вкладом в решение этих задач.

2. Достоверность и научная новизна результатов работы

Обоснованность и достоверность полученных результатов подтверждаются проведением автором многократных численных экспериментов с использованием репрезентативного набора тестовых задач и методов статистической обработки данных. Валидность выводов обеспечивается корректным применением методов кластерного анализа, приближенного поиска ближайших соседей, оптимизации, математической статистики и результатами решения прикладных задач.

Научная новизна определяется разработанными новыми подходами приближенному поиску ближайших соседей на основе индекса обратного файла, основу которых составляют: классификатор сложности запросов, алгоритм определения глубины поиска на основе классификации сложности запроса, алгоритмы кластеризации на основе жадной агломеративной процедуры, а также новой моделью кластеризации мультимодальных данных.

3. Наиболее существенные результаты исследований и ценность для практического использования полученных соискателем результатов

1. Разработан новый алгоритм классификации запросов по уровню сложности для приближенного поиска ближайших соседей с использованием IVF-индекса, позволяющий определять требуемую для достижения целевого показателя полноты область поиска на основе числа результативных кластеров после начального этапа поиска;

2. Разработан новый адаптивный алгоритм поиска ближайших соседей в

векторной базе данных на основе IVF-индекса, отличающийся от известных использованием классификатора сложности запросов на основе результатов предварительного поиска, что позволяет повысить среднюю эффективность поиска;

3. Разработаны новые эволюционные алгоритмы решения задачи k -средних, отличающиеся от известных оператором мутации, основанном на ускоренной жадной агломеративной процедуре, и позволяющих повысить точность решения задачи кластеризации и качество построения IVF-индекса;

4. Предложена новая модель кластеризации мультимодальных данных, которая, в отличие от известных моделей, позволяет напрямую, без приведения к единому векторному виду, применять алгоритмы кластеризации к мультимодальным данным.

Практическая ценность диссертационной работы заключается в возможности применения разработанных моделей и алгоритмов для решения задач кластеризации объектов и поиска данных в информационных системах, работающих с разнородными наборами данных в различных прикладных областях. Разработанный эволюционный алгоритм и адаптивный алгоритм поиска могут эффективно использоваться в системах управления векторными базами данных, при этом применение эволюционных алгоритмов типа $(1+\lambda)$ и упрощенного жадного агломеративного алгоритма позволяет снизить долю неверно определенных ближайших соседей в базах данных, хранящих эмбединги. Адаптивный алгоритм приближенного поиска ближайших соседей и классификатор сложности запросов находят применение в векторных базах данных с сотнями миллионов объектов для эффективного приближенного поиска. Разработанная мера расстояния в пространстве мультимодальных данных обеспечивает возможность вычисления расстояний между мультимодальными объектами, что позволяет не только адаптировать существующие алгоритмы кластеризации для такого пространства, но и сравнивать расстояния между собой, делая возможным и эффективным приближенный поиск ближайших соседей в пространстве мультимодальных данных.

4. Достоверность и научная новизна результатов работы

Диссертация Казаковцева В.Л. представлена на 142 страницах, включая приложения, основной текст состоит из введения, четырех глав и заключения.

Во введении представлено обоснование актуальности проведенного диссертационного исследования, сформулированы его цель и задачи, определены элементы научной новизны и практической значимости полученных результатов. Также в данном разделе описаны использованные методы исследования и изложены основные положения, выносимые на защиту.

В первой главе представлен обзор современных алгоритмов приближенного поиска ближайших соседей, включая методы, основанные на квантизации. Описан подход, использующий инвертированный файл (IVF),

приведены способы его построения, а также рассмотрены современные методы обработки мультимодальных данных.

Вторая глава содержит описание разработанного адаптивного алгоритма поиска ближайших соседей с использованием IVF-индекса. Основная идея метода заключается в динамической оценке сложности входящего запроса на основе анализа количества кластеров, содержащих релевантные векторы на начальном этапе поиска. Для этого разработан классификатор, определяющий «результативность» кластеров относительно конкретного запроса. В зависимости от полученной оценки алгоритм автоматически корректирует глубину поиска (количество просматриваемых векторов) для достижения заданного целевого уровня полноты (Recall). Экспериментальные исследования демонстрируют, что предложенный подход позволяет значительно сократить время обработки запросов по сравнению с неадаптивными методами при сохранении высокого качества результатов по метрике полноты.

В третьей главе рассматривается задача повышения качества кластеризации данных для последующего построения более эффективного IVF-индекса. В рамках исследования предложены новые эволюционные алгоритмы решения задачи k -средних. Используемый оператор мутации на основе жадной агломеративной процедуры обеспечивает эффективное исследование пространства решений и позволяет избегать попадания в локальные оптимумы, характерные для классического алгоритма Ллойда. Специализированный алгоритм кластеризации эмбедингов позволяет использовать жадную агломеративную эвристику для данных высокой размерности, используя особенность таких данных для ускорения жадной агломеративной процедуры. Проведенные эксперименты показывают, что предложенные алгоритмы формируют более качественное разбиение данных на кластеры, что в конечном итоге способствует повышению точности и скорости поиска по IVF-индексу, построенному на их основе.

В четвертой главе разработан новый метод кластеризации мультимодальных данных, включая, в частности, объекты вида «текст–изображение». В отличие от подходов, опирающихся на сложные нейросетевые модели для отображения различных модальностей в общее векторное пространство, здесь предложена модель кластеризации, основанная на специально сконструированной агрегированной мере расстояния. Указанная мера определяется как взвешенное сочетание расстояний между объектами в каждом из модальных пространств по отдельности. Благодаря этому становится возможным напрямую применять традиционные алгоритмы кластеризации, такие как k -средних и иерархическая кластеризация, к исходным мультимодальным данным. Результаты экспериментов на сгенерированных и реальных наборах данных подтверждают эффективность предложенной модели, демонстрируя ее способность формировать интерпретируемые и содержательные группы объектов без привлечения методов глубокого обучения.

В заключении представлены выводы и результаты проведенного диссертационного исследования, указывающие на выполнение поставленных задач и достижение цели диссертационной работы.

По результатам диссертационного исследования опубликованы 24 научные работы, включая 6 статей в журналах, входящих в Перечень рецензируемых изданий ВАК РФ. Кроме того, в Федеральной службе по интеллектуальной собственности зарегистрированы 2 программных продукта, созданных в рамках выполненного исследования.

Содержание диссертации изложено логично и последовательно. Стиль изложения отличается ясностью и структурной целостностью. Оформление работы соответствует установленным требованиям ВАК. Автореферат в полной мере отражает ключевые положения диссертации, ее основные выводы и рекомендации.

5. Рекомендации по использованию результатов диссертационной работы

Результаты диссертационной работы Казаковцева В. Л. целесообразно использовать в научно-исследовательской и практической деятельности организаций, разрабатывающих и эксплуатирующих системы обработки больших массивов векторных и мультимодальных данных. В первую очередь разработанные алгоритмы адаптивного приближенного поиска ближайших соседей на основе IVF-индекса могут быть рекомендованы к внедрению в векторные базы данных, поисковые платформы, рекомендательные системы, системы интеллектуального анализа данных, компьютерного зрения, обработки текстов и видео-контента, где предъявляются повышенные требования к скорости отклика при сохранении высокой полноты поиска.

Предложенные методы построения IVF-индекса, включая эволюционные алгоритмы кластеризации и ускоренную жадную агломеративную процедуру, рекомендуется использовать при проектировании и модернизации высокопроизводительных хранилищ эмбедингов, а также при создании специализированных программных модулей индексации для информационно-аналитических систем, функционирующих в условиях больших объемов данных и ограничений по вычислительным ресурсам.

Разработанную модель кластеризации мультимодальных данных и агрегированную меру расстояния представляется целесообразным применять в задачах интеллектуальной обработки разнородной информации без обязательного построения единого векторного пространства и без использования трудоемких процедур обучения глубоких моделей. Указанные результаты могут быть использованы в системах анализа коллекций документов, мультимедийных архивах, цифровых библиотеках, медицинских и промышленных информационных системах.

6. Замечания и недостатки

1. Ограниченно освещены вычислительные затраты на обучение классификатора сложности запросов и его последующее сопровождение при изменении состава или распределения данных.

2. Предложенная ускоренная агломеративная процедура основана на предположении о зависимости прироста целевой функции преимущественно от мощности удаляемого кластера, однако в диссертационной работе недостаточно подробно исследованы границы применимости данного предположения для наборов данных с различной структурой, плотностью и степенью неоднородности.

3. В главе, посвященной кластеризации мультимодальных данных, автором заявлено отсутствие необходимости в формировании единого векторного пространства, однако выбор весовых коэффициентов модальностей, как следует из автореферата, во многом остается эвристическим, что может затруднять практическое применение предложенной модели в предметных областях с различной информативностью модальностей без использования экспертной оценки.

4. Экспериментальная часть производит благоприятное впечатление, однако включение сравнений с более широким кругом современных методов приближенного поиска и мультимодальной обработки данных позволило бы сделать выводы о преимуществах предложенных решений еще более аргументированными.

5. Во второй главе 2 используются обозначения переменных (статистических признаков результата запроса), не вполне привычные для данной области, например, много обозначений с использованием угловых скобок, что затрудняет восприятие.

Отмеченные замечания носят частный характер и направлены, в первую очередь, на дальнейшее развитие представленного исследования.

7. Выводы

Работа Казаковцева Владимира Львовича на тему «Алгоритмы ускоренного поиска в векторных базах данных», представленная на соискание ученой степени кандидата технических наук по специальности 2.3.1. Системный анализ, управление и обработка информации, статистика, является самостоятельной, завершенной научно-исследовательской работой, выполненной на высоком научном уровне и имеющей теоретическую значимость и практическую ценность. Работа является актуальной и имеет существенное значение для теории и практики системного анализа, управления и обработки информации. Диссертация соответствует паспорту специальности 2.3.1. Системный анализ, управление и обработка информации, статистика.

Представленная диссертационная работа отвечает требованиям пп. 9–14 «Положения о присуждении ученых степеней», утвержденного постановлением Правительства РФ от 24.09.2013 № 842 в редакции от 11 сентября 2021 г., предъявляемым ВАК РФ к диссертациям на соискание ученой степени

кандидата технических наук, а ее автор – Казаковцев Владимир Львович – заслуживает присуждения ему ученой степени кандидата технических наук по специальности 2.3.1. Системный анализ, управление и обработка информации, статистика.

Отзыв на диссертацию и автореферат обсужден и утвержден на заседании кафедры автоматизированных и вычислительных систем, протокол № 12 от «24» апреля 2026 г.

Заведующий кафедрой автоматизированных и вычислительных систем ФГБОУ ВО «Воронежский государственный технический университет»,
д-р техн. наук, профессор

Барабанов Владимир Федорович
24.04.2026

Профессор кафедры автоматизированных и вычислительных систем ФГБОУ ВО «Воронежский государственный технический университет»,
д-р техн. наук, профессор

Кравец Олег Яковлевич
24.04.2026

Сведения о составителях отзыва:

Барабанов Владимир Федорович, д.т.н., профессор, защитил докторскую диссертацию по специальности 05.13.12.

Контактный телефон: +7473 2072220

Адрес электронной почты: bvf@list.ru

Кравец Олег Яковлевич, д.т.н., профессор, защитил докторскую диссертацию по специальности 05.13.09.

Контактный телефон: +7473 2072220

Адрес электронной почты: csit@bk.ru

Сведения о ведущей организации:

Федеральное государственное бюджетное образовательное учреждение высшего образования "Воронежский государственный технический университет"

Почтовый адрес: ул. 20-летия Октября, д. 84, г. Воронеж, 394006

Телефон: +7(473) 271-59-05

E-mail: rector@vorstu.ru Сайт: <http://cchgeu.ru>