

УТВЕРЖДАЮ

Проректор МГТУ им. Н.Э. Баумана по
науке и цифровому развитию, д.э.н.,
профессор

ДРОГОВОЗ П. А.

« 15 »

2025 г.



ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

**Федеральное государственное автономное образовательное учреждение
высшего образования «Московский государственный технический
университет имени Н.Э. Баумана»**

(национальный исследовательский университет)

**на диссертационную работу Александра Андреевича Чеповского
«Методы работы с неявными сообществами на взвешенных графах
взаимодействующих объектов», представленной на соискание ученой
степени доктора физико-математических наук по специальности
1.2.2 – «Математическое моделирование, численные методы и
комплексы программ»**

Актуальность темы исследования

Диссертационная работа посвящена актуальному направлению исследований, связанному с разработкой методов и алгоритмов выделения пересекающихся и непересекающихся сообществ на реальных графах

взаимодействующих объектов, что является сложной проблемой в исследованиях телекоммуникационного взаимодействия. Анализ графов взаимодействующих объектов необходим для разработки эффективных информационно-аналитических систем обеспечения государственной безопасности и организации деятельности по распространению достоверной информации в обществе.

Для создания информационно-аналитических систем проблема анализа реальных графов взаимодействующих объектов, включая сети мгновенного обмена сообщениями, приводит к необходимости разработки алгоритмов для выявления структуры графа; формирования методов анализа соответствующих данных; разработки методов оценки корректности результатов анализа; программной реализации средств обработки графов больших размеров, включая создание специализированных эффективных хранилищ графов. Для данных, полученных при импорте из социальных сетей и мессенджеров, возникают задачи формирования графов взаимодействия, корректных для различных источников, и выбора методов их анализа с целью выявления каналов распространения и обмена информацией между пользователями.

Исследования по разработке алгоритмов выделения неявных сообществ на графах ведутся два десятилетия, в основном за рубежом, но созданные методы не решают в полном объеме задачу выделения пересекающихся и непересекающихся сообществ на графах взаимодействующих объектов, полученных из реальных данных о социальных и телекоммуникационных взаимодействиях. Остается нерешенной задача оценки корректности и эффективности разрабатываемых алгоритмов и методов.

Вышеизложенное демонстрирует высокую актуальность диссертационных исследований для области анализа реальных графов взаимодействующих объектов, для которых характерны и играют существенную роль атрибуты ребер и вершин.

Содержание диссертационной работы

Диссертация состоит из введения, семи глав, выводов, библиографического списка, включающего в себя 180 наименований, и 2 приложений. Работа содержит 205 страниц машинописного текста основной части, включающей 78 рисунков, 51 таблицу и 19 страниц библиографии. Приложения содержат 12 страниц машинописного текста.

Во **введении** сформулирована актуальность исследуемой проблемы, определены цели и задачи диссертационной работы, перечислены полученные в диссертации новые результаты, их практическая ценность, представлены положения, выносимые на защиту.

В **первой главе** приведен аналитический обзор проблемы анализа графов, представляющих сложные сети взаимодействующих объектов. Приведено описание предметных областей, для которых актуальны модели графов взаимодействующих объектов. Приведен обзор основных классических подходов по выделению неявных сообществ на графах. Указаны в том числе недостатки стандартных подходов, не учитывающих атрибутивные данные вершин и ребер графов подобного рода, а также важность разработки методов для оценки качества разбиения графа. Показана актуальность подходов, предложенных в диссертации.

Во **второй главе** предложены модели построения графа взаимодействующих объектов при импорте реальных данных из социальных сетей и сетей мгновенного обмена сообщениями. Формирование графов взаимодействующих объектов обеспечивается разработанными процедурами импорта данных из соответствующих сетей с последующим построением взвешенного графа на основе выбранных атрибутов и весовых функций. Показано, что разработанные модели позволяют из сетей разного типа формировать графы взаимодействующих объектов для дальнейшего их анализа.

В **третьей главе** представлен разработанный «комбинированный алгоритм», основанный на классических алгоритмах, в оценке результативности которых положено в том числе понятие «модулярность».

Разработанный алгоритм выделения пересекающихся сообществ имеет несколько степеней параметризации, среди которых можно подбирать наиболее подходящие значения для графов разной природы, включая графы, полученные при импорте данных из социальных сетей и сетей мгновенного обмена сообщениями. Подробно изложена логика работы алгоритма и принцип его построения. Вычислительные эксперименты применения комбинированного алгоритма проведены на большом числе графов различных взаимодействующих объектов.

В четвертой главе описан разработанный «метод ядра», для выделения непересекающихся сообществ на взвешенных графах, предусматривающий выявление ключевой компоненты на основании вычисляемых в явном виде характеристик графа. Метод демонстрирует комплексный подход к анализу взвешенного графа. Дано подробное описание методики работы с графом, лежащей в основе алгоритма. Апробация метода продемонстрирована на реальных данных из сети Twitter.

В пятой главе предложен реализованный «метод Галактик» для выделения пересекающихся сообществ на взвешенных графах, полученных при импорте данных из сети мгновенного обмена сообщениями Telegram. Метод основан на конструировании алгоритма из базовых алгоритмов, переходам к мета-графу из мета-сообществ и выделении сообществ на мета-графе с последующим возвратом к исходному графу. Предложенный метод позволяет проводить эффективные процедуры анализа для графов больших размеров посредством выделения мета-сообществ, что сокращает количество элементов, исследуемых с точки зрения распространения и хранения информации в сетях взаимодействующих объектов больших объемов. Показаны результаты применения алгоритма на реальных данных.

Шестая глава посвящена методикам оценки качества выделения сообществ на графе взаимодействующих объектов, имеющем атрибутивные текстовые данные для вершин. Методики основаны на алгоритмах компьютерной лингвистики и выделении психолингвистических факторов для

объединенных массивов текстов выделенных сообществ. Предложены две методики оценки разделения на сообщества. Первая основана на попарном сравнении методами рангового анализа частотных словарей различных лингвистических характеристик, составленных для наборов текстов каждого сообщества. Вторая базируется на сравнении наборов статистических (психолингвистических) характеристик текстов, описывающих различную направленность неявных сообществ.

В **седьмой главе** представлены основные компоненты разработанного программного обеспечения для анализа графов взаимодействующих объектов. Реализованное программное обеспечение включает в себя важные части – хранилище графов и приложение для анализа и визуализации графов. Для оценки эффективности реализации хранилища графов проведено сравнительное тестирование характеристик программного обеспечения, предназначенного для хранения и обработки больших графов.

В **заключении** сформулированы основные научные результаты, полученные в ходе диссертационного исследования.

Содержание автореферата соответствует содержанию диссертации.

Оценка изложения результатов диссертационных исследований

Текст диссертации хорошо структурирован. Формулировки задач, разработанных методов и алгоритмов представлены в соответствии с общепринятыми нормами. Разработанные алгоритмы описаны в соответствии с принятыми для компьютерных наук и научных стандартов требованиями, соответствуют специальности 1.2.2 – «Математическое моделирование, численные методы и комплексы программ».

Научная новизна исследования и полученных результатов

В диссертации получены следующие научные результаты, являющиеся новыми.

1. Представлено решение научной проблемы, имеющей важное хозяйственное значение, заключающееся в создании моделей, разработки численных методов и программного обеспечения для анализа структуры

графов взаимодействующих объектов, полученных при импорте данных из социальных сетей и сетей мгновенного обмена сообщениями с целью описания информационного взаимодействия объектов.

2. Разработаны и реализованы в программном обеспечении вариации модели формирования взвешенного графа информационного взаимодействия для разных социальных сетей и сетей мгновенного обмена сообщениями. Приведены вариации данной модели для импорта данных из сети Twitter, сети Telegram-каналов и социальной сети ВКонтакте.
3. Построены итерационные численные методы и алгоритмы для выделения неявных сообществ и ключевых вершин графов с использованием эвристик, а именно.
 - 3.1. Предложен и реализован «Комбинированный алгоритм» для выделения пересекающихся и вложенных сообществ на графе, позволяющий убирать из рассмотрения малозначимые элементы сети и предусматривающий параметрические модификации для формирования разнородных разбиений в зависимости от задач оператора.
 - 3.2. Предложен и реализован «Метод ядра» для выделения непересекающихся сообществ на взвешенных графах, предусматривающий выделение ключевой компоненты на основании вычисляемых в явном виде характеристик графа. Апробация метода продемонстрирована на реальных данных из сети Twitter.
 - 3.3. Предложен и реализован «Метод Галактик» для выделения пересекающихся сообществ на взвешенных графах, основанный на последовательном применении других алгоритмов, обработке графа, переходам к мета-графу из мета-сообществ и последующем выделении пересекающихся сообществ. Показано применение алгоритма на

реальных данных из сети Telegram-каналов с последующим экспертным обоснованием качества полученного разбиения.

4. Предложена методика оценки эффективности выделения сообществ на графе с помощью алгоритмов компьютерной лингвистики для обработки текстовых метаданных (атрибутов вершин выделенных сообществ) и анализа психолингвистических факторов.
5. Разработана модель для эффективного хранения графов взаимодействующих объектов, основанная на алгоритмах сжатия и оптимизации операций с графами по памяти и по скоростным характеристикам. Создано программное обеспечение для анализа графов взаимодействующих объектов.

Степень достоверности результатов, изложенных в диссертации

Достоверность результатов и обоснованность научных положений диссертационной работы обеспечивается корректным использованием математических методов. Достоверность выводов по результатам исследований и разработок подтверждаются экспериментальными исследованиями и экспертной оценкой результатов. Выводы диссертации прошли апробацию на международных и российских конференциях, принимались к публикации в рецензируемых журналах.

Полнота изложения материалов диссертации в публикациях

Основные результаты диссертационного исследования опубликованы в 37 работах. Из них 22 статьи в ведущих рецензируемых научных журналах, которые входят в утвержденный ВАК Минобрнауки России «Перечень российских рецензируемых научных журналов, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней доктора и кандидата наук» по специальности 1.2.2 (05.13.18) и приравненных к ним зарубежных рецензируемых изданиях (Scopus) и входящих в базу RSCI (на платформе Web of Science). В том числе 12 статей в журналах категории К1 и К2 по распределению на 2022 и 2023 годы. Две рецензируемые монографии, 13 публикаций в трудах международных

научных конференций. Два свидетельства о регистрации программ для ЭВМ. Результаты диссертации использованы в 2 учебных пособиях. Основные результаты диссертации докладывались на семинарах и международных конференциях. Основные научные результаты получены автором самостоятельно. Содержание диссертации и основные положения, выносимые на защиту, отражают персональный вклад автора в опубликованные работы. Публикации полностью соответствуют теме диссертационного исследования и раскрывают ее основные положения.

Значимость полученных результатов для развития отрасли науки

Теоретическая значимость результатов работы состоит в том, что разработаны новые методы и алгоритмы выделения неявных пересекающихся и непересекающихся сообществ на графе взаимодействующих объектов, опирающиеся на структурные особенности графа, и в том, что предложены неприменяемые ранее методы для оценки качества решения задачи по выделению сообществ, основанные на методах компьютерной лингвистики.

Теоретическая и практическая значимость определяется разработкой и применением методов построения и анализа графов к различным социальным сетям и сетям мгновенного обмена сообщениями.

Большую теоретическую и практическую значимость представляют разработанные методы хранения графов с атрибутивной информацией, положенные в основу разработанного и реализованного эффективного хранилища графов.

Практическое значение заключается в создании прикладного программного обеспечения, реализующего разработанные алгоритмы и средства визуального анализа графов. Получены два свидетельства о регистрации программ для ЭВМ.

Научные и практические результаты диссертации использованы в 2016 – 2021 годах в четырех грантах РФФИ. Результаты диссертации использованы в 2 учебных пособиях.

Рекомендации по возможности использования результатов диссертационной работы

Разработанные в диссертации методики и результаты исследований рекомендуются для использования крупными интеграторами программных продуктов при создании программного обеспечения для операторов связи и государственных организаций, включая специальные службы и службы безопасности коммерческих структур. Разработанные методики рекомендуются для применения социологическим службам и службам, занимающимся противодействию информационному воздействию.

Разработанные в диссертации методики и программные продукты рекомендуются для использования в подразделениях Федеральной службы по финансовому мониторингу (Росфинмониторинг) и Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор).

Результаты исследований могут использоваться в учебном процессе при подготовке специалистов по информационным технологиям и информационной безопасности.

Замечания по работе.

При анализе содержания текста диссертации и автореферата отмечены следующие недостатки:

1. Во второй главе представлена модель построения графов взаимодействующих объектов для популярных сетей, актуальных для задач государственной безопасности. При этом описание модели для сети Telegram-каналов хоть и полно с точки зрения решаемой соискателем проблемы, но не содержит классификацию вершин (каналов).

2. В главе 7 для разработанного хранилища графов экспериментальные оценки приведены для сравнения с пятью программами, что оставляет вопрос сравнения с другими существующими решениями.

Отмеченные замечания не снижают научной и практической ценности результатов, полученных в диссертации.

Заключение

Диссертация Чеповского Александра Андреевича на тему: «Методы работы с неявными сообществами на взвешенных графах взаимодействующих объектов» отвечает требованиям, изложенным в пунктах 9 – 11, 13 и 14 «Положения о присуждении ученых степеней» Постановления Правительства РФ от 24.09.2013 г. № 842 (в ред. от 16.10.2024 №1382), предъявляемым к диссертациям на степень доктора наук.

Диссертация является самостоятельной законченной научно-квалификационной работой, в которой содержится решение имеющей важное хозяйственное и социально-экономическое значение проблемы анализа телекоммуникационных данных. Решение данной проблемы содержит новые научно обоснованные решения, внедрение которых вносит значительный вклад в развитие методов анализа данных в сфере государственной и общественной безопасности.

Диссертация выполнена в соответствии с паспортом научной специальности 1.2.2 – «Математическое моделирование, численные методы и комплексы программ». Содержание работы соответствует пункту 1 «Разработка новых математических методов моделирования объектов и явлений» (физико-математические науки), пункту 2 «Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий» и пункту 8 «Комплексные исследования научных и технических проблем с применением современной технологии математического моделирования и вычислительного эксперимента».

Автор диссертации Чеповский Александр Андреевич заслуживает присуждения ученой степени доктора физико-математических наук по специальности 1.2.2 – «Математическое моделирование, численные методы и комплексы программ».

Отзыв заслушан и единогласно одобрен на заседании кафедры «Защита информации» (ИУ-10) МГТУ им. Н.Э. Баумана (Протокол № 04.22-04/12 от 14 апреля 2025 г.) Присутствовало 20 чел.

Отзыв составили:

Профессор кафедры «Защита информации» (ИУ-10)
МГТУ им. Н.Э. Баумана, д.т.н.,
профессор
« 15 » апреля 2025 г.

Сычев Михаил Павлович

Профессор кафедры «Защита информации» (ИУ-10) МГТУ им. Н.Э. Баумана, академик Академии криптографии Российской Федерации, д.т.н., профессор
« 15 » апреля 2025 г.

Смирнов Сергей Николаевич

Личные подписи Сычева М.П. и Смирнова С.Н. заверяю
Проректор МГТУ им. Н.Э. Баумана по науке и цифровому развитию, д.э.н., профессор



Выходные данные организации:

Федеральное государственное автономное образовательное учреждение высшего образования «Московский государственный технический университет имени Н.Э. Баумана» (национальный исследовательский университет) (МГТУ им. Н.Э. Баумана).

Адрес: 105005, Москва, 2-я Бауманская, д.5, стр.1

Телефон: 8-499-263-63-91

Сайт: <https://bmstu.ru>. Адрес электронной почты: bauman@bmstu.ru