

## **ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА**

**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»**

### **Оценка актуальности темы диссертационных исследований**

Диссертация посвящена актуальным вопросам научной тематики, относящейся к разделам алгоритмистики и численных методов, используемых для решения задач поиска пересекающихся и непересекающихся подграфов. В научной литературе рассматриваются графы сложных сетей и выделяемые в них подграфы, называемые сообществами и имеющие относительно большую внутреннюю плотность ребер по сравнению с плотностью ребер между ними. Такие подграфы в коммуникационных сетях часто свидетельствуют об активном взаимодействии между объектами той сети, которую моделирует соответствующий им граф.

Затронутая в работе проблематика является продолжением многих классических задач и активно изучается в научных кругах разных стран, особенно в последние десятилетия. Исследования, проводимые в этой сфере для коммуникационных сетей, имеют важное научное и прикладное значение. Их результаты могут быть использованы при создании программно-аппаратных комплексов, нацеленных на решение задач как государственной безопасности, так и маркетинговых исследований. Важным аспектом при построении информационных систем, нацеленных на решение подобных задач, является проектирование и реализация алгоритмов выявления таких скрытых структур, построение комплекса программ, предусматривающего

хранение и обработку графов больших размеров – графовых хранилищ. Моделирование таких объектов как графы коммуникационных сетей, рассматриваемые в диссертационной работе, характеризуется актуальными для современной науки и общества запросами.

Отдельно стоит отметить актуальную задачу оценки качества выделения сообществ. Классические методы анализа полученных сообществ не позволяют оценить результаты для графов, полученных в результате импорта реальных данных из коммуникационных сетей, ибо не предусматривают учет атрибутивных свойств объектов. Для коммуникационных сетей такие данные представляют собой крайне важный аспект этой задачи, мало освещенный в работах других авторов. Между тем, именно использование атрибутивных данных как дополнительных эвристик для алгоритмов и заложенных в них численных методов, а также для оценки качества выделенной структуры, представляют из себя неоспоримо важную научную задачу.

Вышеописанные особенности рассматриваемой в работе научной проблематики подтверждают научную новизну и практическую актуальность исследований, описанных в диссертационной работе. Автором предложены востребованные для современных целей новые алгоритмы и методы работы с графами, моделирующими коммуникационные сети.

### **Структура и 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** описана актуальность диссертационных исследований, сформулированы цели и задачи работы, представлены положения, выносимые на защиту.

В **первой главе** сделан подробный аналитический обзор о проблематике работ с графами взаимодействующих объектов. Проанализированы

классические методы по выделению неявных сообществ на графах. Сформулированы недостатки существующих подходов, не учитывающих атрибутивные данные вершин и ребер графов, а также важность разработки методов для оценки качества разбиения графа.

Во **второй главе** описаны модели, используемые при построении графа взаимодействующих объектов на основе данных из социальных сетей и сетей мгновенного обмена сообщениями. Разработаны процедуры импорта данных из рассматриваемых сетей с последующим формированием взвешенного графа на основе выбранных атрибутов и весовых функций.

В **третьей главе** сформулирован «комбинированный алгоритм», в основе которого лежат классические алгоритмы и параметризованные эвристики для их применения к различным фрагментам графа. Подробно изложены принципы построения алгоритма. Проведены вычислительные эксперименты на большом числе графов. В этой главе представлены модификации итерационных алгоритмов, учитывающие атрибутивные свойства вершин.

В **четвертой главе** описан «метод ядра», который позволяет выделять непересекающиеся сообщества на взвешенных графах. Метод основывается на выявлении ключевых вершин графа с помощью явно вычисляемых характеристик графа. Алгоритм корректно и подробно представлен. Работоспособность метода продемонстрирована на реальных данных из сети Twitter и показывает комплексный подход к анализу взвешенного графа.

В **пятой главе** изложен «метод Галактик» предназначенный для выделения пересекающихся сообществ на взвешенных графах, построенных по данным из Telegram. За счет перехода между исходным графом и метаграфом его сообществ данный метод дает возможность рассмотрения графов больших размеров с меньшим числом элементов визуализации для пользователей. Применение алгоритма продемонстрировано на реальных данных, приведено обоснование качества выделяемых этим методом сообществ.

В **шестой главе** предлагаются две методики для оценки качества выделения сообществ на рассматриваемых графах, у которых среди атрибутов вершин имеются содержательные текстовые данные. Одна из них состоит в построении частотных словарей по обобщенным текстам атрибутов вершин из выделенных сообществ и дальнейшем сравнении рангов их элементов. Другая основана на сравнении лингвистических характеристик текстов, построенных как объединение данных из атрибутов у выделенных сообществ. Приведены множественные примеры применения этих методик.

В **седьмой главе** описаны ключевые элементы разработанного программного обеспечения, предназначенного для анализа графов взаимодействующих объектов. Ключевыми составляющими данного программного обеспечения являются хранилище графов и приложение для анализа и визуализации графов. Представлено тестирование и сравнение его характеристик со стандартным программным обеспечением для хранения и работы с графами больших размеров.

В **заключении** сформулированы полученные научные результаты.

Содержание автореферата соответствует содержанию диссертации.

В целом текст диссертации представляет собой хорошее, логически четкое изложение материала. Отметим, что разработанные методы и алгоритмы описаны в тексте диссертации в духе классических трудов по алгоритмам и вычислительной математике.

#### **Научная новизна исследования и полученных результатов**

Все научные результаты, выносимые на защиту, являются новыми. В рамках диссертационного исследования созданы следующие составляющие. Предложена концепция моделей построения взвешенных сетевых структур на основе данных, импортируемых из социальных сетей и мессенджеров, отражающая специфику информационных взаимосвязей между объектами. Предложен и реализован комплекс алгоритмических решений для структурного анализа сетей: «комбинированный алгоритм» для выявления пересекающихся сообществ и исключения из анализа ее малозначимых

элементов; «метод ядра» для выделения непересекающихся сообществ и ключевого набора вершин за счет вычисляемых в явном виде характеристик; «метод Галактик» для выделения пересекающихся сообществ на взвешенных графах. Разработана инновационная система оценки качества полученного разбиения графа, интегрирующая лингвистические алгоритмы обработки текстовых данных из атрибутов вершин в сформированных сообществах. Предложена архитектура хранения графовых данных, сочетающая механизмы сжатия и алгоритмические оптимизации для повышения производительности операций с графами.

Практическим итогом работы стало создание программного комплекса с функционалом для анализа сетевых структур, включая инструменты интерактивной визуализации. Разработанные решения прошли апробацию в реальных условиях, подтвердив свою эффективность и соответствие заявленным характеристикам.

Достигнутые в диссертационном исследовании результаты формируют целостный методологический базис для решения проблемы сетевого анализа в условиях больших данных, сочетающий теоретическую глубину с прикладной ориентированностью.

#### **Обоснованность и достоверность научных результатов и выводов, изложенных в диссертации**

Достоверность результатов и обоснованность научных положений диссертационной работы обеспечивается корректным использованием математическим инструментарием из теории графов. Достоверность выводов, изложенных в диссертации, подтверждает их соответствие результатам вычислительных экспериментов и их экспертной оценкой. Результаты прошли апробацию на международных и российских конференциях, принимались к публикации в рецензируемых журналах.

Основные результаты диссертационного исследования опубликованы в 37 работах. Из них 22 статьи в ведущих рецензируемых научных журналах, входящих в список ВАК (из которых 12 статей в журналах категории K1 и K2)

и приравненных к ним зарубежных рецензируемых изданиях. По результатам диссертации издано две рецензируемые монографии, два свидетельства о регистрации программ для ЭВМ. Результаты диссертации использованы в 2 учебных пособиях.

Публикации автора полностью соответствуют содержанию диссертации и раскрывают ее основные положения и отражают персональный вклад автора в опубликованные работы.

### **Значимость полученных результатов для развития отрасли науки**

Теоретическая значимость диссертационного исследования заключается в создании инновационных подходов к обнаружению неявных сообществ на графах взаимодействующих объектов, как для случаев пересекающихся, так и для случаев непересекающихся сообществ. Эти подходы базируются на анализе топологических особенностей сетевых структур, а также дополнены оригинальными критериями для оценки эффективности решений, разработанными с привлечением методов лингвистического анализа данных.

Научная и прикладная значимость исследования подтверждается адаптацией предложенных методик к анализу разнородных сетевых данных, включая социальные сети и сервисы мгновенного обмена сообщениями. Особого внимания заслуживает концепция атрибутивного хранения сетевых структур, которая легла в основу специализированного программного комплекса с оптимизированными возможностями обработки больших массивов данных.

С практической точки зрения, работа демонстрирует высокую результативность за счет программной реализации авторских алгоритмов и интерактивных решений для визуализации графа. По разработкам получены два свидетельства о регистрации программ для ЭВМ.

Результаты исследования нашли применение в научно-исследовательской и образовательной сферах в виде использования в четырех грантах при поддержке РФФИ (2016–2021 гг.) и включения материалов в два учебных пособия.

Таким образом, диссертационная работа вносит вклад как в развитие теоретических основ сетевого анализа, так и в создание технологических решений, актуальных для задач управления данными в социально-ориентированных системах.

### **Замечания по работе.**

1. В третьей главе изложен «комбинированный алгоритм», объяснены его идеи и суть проводимых вычислений. При этом введено достаточно много специализированных обозначений для описания проводимых в алгоритме процедур, что усложняет восприятие.

2. В пятой главе при описании применения «метода Галактик» к реальным данным сказано про существенное сокращение количества элементов визуализации. При этом в явном виде не указано, каков порядок этого сокращения элементов.

Данные замечания не являются принципиальными и не влияют на общую положительную оценку основных результатов диссертации.

### **Заключение**

Диссертация Чеповского Александра Андреевича на тему: «Методы работы с неявными сообществами на взвешенных графах взаимодействующих объектов» является выполненной самостоятельно законченной научно-квалификационной работой, в которой содержится решение имеющей важное хозяйственное и социально-экономическое значение проблемы анализа графов, возникающих в информационно-коммуникационных технологиях.

Диссертация отвечает требованиям, изложенным в пунктах 9 – 11, 13 и 14 «Положения о присуждении ученых степеней» (Постановление правительства РФ от 24.09.2013 г. № 842 (в ред. от 16.10.2024 № 1382), предъявляемым к диссертациям на степень доктора наук.

Диссертация полностью соответствует паспорту научной специальности 1.2.2 – «Математическое моделирование, численные методы и комплексы программ», а именно: пункту 1 «Разработка новых математических методов моделирования объектов и явлений» (физико-математические науки),

