

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»**

Актуальность диссертационных исследований

Диссертационное исследование посвящено выявлению фундаментальных общих закономерностей в формировании и поведении сложных сетевых структур на основе математического моделирования взаимодействия их элементов при отображении на графы. Центральным вопросом рассматриваемого класса задач является выявление подмножеств (именуемых в диссертационной работе сообществами) на основе плотности распределения связей между вершинами графа и дополнительной информации, специфической для конкретных приложений.

Начало таких исследований восходит к работам Л. Эйлера (1736 г.) по топологической структуре графов, связанных с оптимизационными задачами построения путей между заданными точками (вершинами графа). Несомненно, сюда примыкают задачи теории электрических цепей, модели ферромагнетизма В. Ленца и Э. Изинга (1920–1925 г.), модели доменной структуры ферромагнетиков Л.Д. Ландау и Е.М. Лифшица (1935 г.), фундаментальные исследования по методам оптимизации и линейного программирования Л.В. Канторовича (1938 г.), диаграммы Р. Фейнмана в квантовой электродинамике (1949 г.), математические методы в теории игр Дж. Фон Неймана, О. Моргеншерна, Дж. Ф. Нэша (1944-1949 г.), алгоритмы генерации вычислительных сеток и параллельные алгоритмы сортировки больших объемов данных М.В. Якововского (2004 г.), множество IT разработок по фрагментации-дефрагментации дискового пространства.

Цифровизация современной общественной жизни вызывает необходимость разработки математических методов, способных эффективно выявлять и анализировать структурные единицы – пересекающиеся и непересекающиеся сообщества, являющиеся подмножествами на графах с дополнительной информационной структурой.

Диссертационная работа А.А. Чеповского посвящена выявлению существенных закономерностей взаимодействия в таких сообществах, включая коммуникационные, инфраструктурные, биологические сети и т.п., на основе разработанных им авторских алгоритмов.

Общепринятые алгоритмы для решения задачи выявления сообществ на графах, реализующих подобные сети, демонстрируют ограничения при работе в случаях наличия атрибутивной информации у вершин и ребер, ибо основаны на обработке исключительно топологии графа. Это затрудняет их применение в критически важных областях: обнаружение каналов дезинформации, анализ распространения информации, управление телекоммуникационными и рекламными сетями. Т. к. для рассматриваемой проблематики характерны взвешенные графы и такие атрибутивные свойства их вершин как текстовые метаданные, то диссертационная работа восполняет актуальный пробел в научной и хозяйственной сфере, предлагая методы, сочетающие структурный анализ графов с обработкой семантических данных, что расширяет возможности моделирования сложных сетевых систем и проведения на них вычислительных процедур.

Этой тематике посвящено значительное количество научных публикаций и международных научных конференций. Таким образом, актуальность темы диссертации несомненна.

Диссертационная работа является развитием идей анализа доменных структур в ферромагнитных материалах на основе оптимизации так называемых модулярностей – прямым аналогом «штрафных» функций в теории игр и энергии взаимодействия магнитных моментов «соседей» в магнитном домене. Структура сообществ анализируется на основе плотности распределения связей между вершинами графа.

Представленная диссертационная работа демонстрирует принципиальную новизну в области анализа коммуникационных сетей, выражающуюся в синтезе математических методов из теории графов, теории алгоритмов, компьютерной лингвистики. Разработана формальная модель интеграции атрибутивных данных в структурный анализ графов. Предложены новые авторские критерии оценки качества разбиения, объединяющие топологические характеристики и семантические параметры.

Приведенные в работе многочисленные примеры показывают результативность разработанных автором алгоритмов по выявлению «сообществ» и их актуальность для различных распространенных среди пользователей коммуникационных сетей. Такое представление ограничивается тестовыми и экспериментальными примерами, при этом не приводятся доказательные утверждения о достаточных условиях сходимости этих алгоритмов. Но это можно отнести к большинству известных из публикаций алгоритмов.

Научные результаты диссертационной работы А.А. Чеповского могут быть использованы для разработки стандартов в области анализа сетевых структур, ГОСТов и отраслевых стандартов в области обработки данных.

Структура и краткое содержание диссертационной работы

Объем диссертации составляет 217 страниц. Диссертация представляет собой системное исследование, организованное в семь взаимосвязанных глав и включает введение, заключение, список литературы из 180 наименований, два приложения.

Диссертационное исследование выполнено в строгом соответствии с методологией математического моделирования и организовано как последовательность взаимосвязанных этапов: построение формальных моделей в терминах теории графов с учетом атрибутивных свойств; разработка алгоритмов и численных методов; программная реализация посредством создания программных комплексов для обработки данных; вычислительные эксперименты как проверка адекватности моделей на реальных данных.

В первой главе дан методологический базис. Проведен критический анализ существующих подходов к выделению сообществ на графах с позиций математического моделирования. Установлены ключевые ограничения базовых алгоритмов, заключающиеся в игнорировании атрибутивных свойств вершин графов и отсутствии эффективных методов оценки качества разбиения.

Глава 2 посвящена процессу построения математической модели исследуемых коммуникационных сетей и формированию взвешенных графов, отражающих их структуру.

В главе 3 изложены численные методы, используемые в предложенных алгоритмах, в том числе итерационный поиск разбиения на сообщества за счет вычисления на каждой итерации показателей качества разбиения, учитывающих атрибутивный состав вершин графа. Показана линейная зависимость времени работы алгоритма от средней степени вершин в графе. В этой же главе представлен «комбинированный алгоритм» с итерационным применением процедур выделения пересекающихся и непересекающихся сообществ за счет учета специально введенной параметрической эвристики. Эти методы используются для максимизации модулярностей – прямым аналогом «штрафных» функций в теории игр, подсчитываемых для отдельных частей графа. Приведены результаты вычислительных экспериментов, подтверждающих эффективность разработанных алгоритмов.

В главе 4 изложен алгоритм выделения непересекающихся сообществ и выявления ключевых с точки зрения активной коммуникации вершин, названный «метод ядра». Введены нестандартные показатели топологических особенностей графа, с помощью вычисления которых выделяются структурно значимые вершины и строятся шаги алгоритма. Вычисления для «метода ядра» произведены на графах, построенных согласно введенным в первой главе моделям из данных Twitter.

В главе 5 описан алгоритм «Галактик» для выделения пересекающихся сообществ за счет редукции исходного графа в мета-граф, что обеспечивает масштабируемость его анализа. Приведены результаты вычислительных

экспериментов и оценки качества разбиения по авторской методике анализа атрибутивных свойств групп вершин – текстов Telegram-каналов.

В главе 6 формализуется подход по оценке качества выделения сообществ на графах взаимодействующих объектов с текстовыми метаданными. Для этих целей используется аппарат компьютерной лингвистики и ранговый анализ словарей, составленных для выделенных сообществ.

В главе 7 детализируется архитектура программного комплекса: оптимизированное графовое хранилище с механизмами сжатия данных и приложение для анализа графов. Продемонстрированы результаты экспериментальных вычислений в сравнении данного подхода и классических СУБД. Показана эффективность этой модели по скоростным характеристикам.

В заключении обобщены и приведены достигнутые основные результаты, полученные в диссертационной работе.

Автореферат полностью отражает содержание диссертации, подтверждая её научную достоверность.

Соответствие научной специальности

Диссертационная работа полностью соответствует научной специальности ВАК 1.2.2 – «Математическое моделирование, численные методы и комплексы программ» и содержит все три составляющие данной специальности.

1. Разработанные математические модели формирования взвешенных графов, отражающих структуру исследуемых коммуникационных сетей (глава 2) и модели текстов сообществ для сравнительного анализа на основе статистических характеристик (глава 6).
2. Разработанные численные методы: итерационный поиск разбиения на сообщества с вычислением показателей качества разбиения на каждой итерации, итерационные процедуры максимизации модулярностей при выделении сообществ на основе параметрической эвристики (глава 3).

3. Созданное программное обеспечение для хранения, визуализации и анализа графов (глава 7), реализующее разработанные в диссертации методы и алгоритмы.

Таким образом, диссертационная работа соответствует паспорту научной специальности, а именно: разработанные методы моделирования взаимодействующих объектов как модели коммуникационных сетей, пункту 1 «Разработка новых математических методов моделирования объектов и явлений» (физико-математические науки); методы и алгоритмы выделения сообществ пункту 2 «Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий» и комплекс алгоритмов и программ выделения сообществ с методами оценки качества разбиения пункту 8 «Комплексные исследования научных и технических проблем с применением современной технологии математического моделирования и вычислительного эксперимента».

Научная новизна исследования и полученных результатов

Научная ценность исследования заключается в преодолении ключевых ограничений классических подходов к выделению сообществ и разработке новых методов работы с ними, что подтверждается следующими аспектами:

1. Введено понятие графа взаимодействующих объектов как математической модели, дающей содержательное описание для коммуникационных сетей.
2. Предложена математическая модель построения графа взаимодействующих объектов, позволяющая совместить структурный анализ полученной топологии графа с контекстной интерпретацией объектов сети, что решает проблему игнорирования контекста в существующих классических методах выделения сообществ (таких как Louvain, Infomap).
3. Разработаны уникальные методы и алгоритмы работы с графами взаимодействующих объектов: «комбинированный алгоритм», «метод ядра» и «метод Галактик».

4. Сформулированы формальные характеристики для оценки семантического различия сообществ и качества полученного разбиения.
5. Предложен и реализован программный комплекс для хранения, обработки и визуализации графов взаимодействующих объектов.

Все это в совокупности представляет полноценную связную работу и устраняет пробел в оценке разбиения на сообщества для графов с текстовыми метаданными.

Работа вносит принципиально новый вклад в теорию анализа графов, преодолевая ключевые ограничения классических подходов к работе с такими сетями. Новизна работы носит фундаментальный характер, устанавливая новые стандарты в теории графов. Разработанный аппарат создает основу для решения задач анализа сетей коммуникации.

Практическая значимость подтверждена реализацией программного комплекса, успешно апробированного на данных Twitter и Telegram, а также интеграцией результатов в исследовательские проекты РФФИ и образовательные программы.

Обоснованность и достоверность научных результатов и выводов диссертационных исследований

Достоверность результатов и обоснованность научных положений диссертационной работы подтверждается корректным использованием математического аппарата теории графов и вычислительных методов. Достоверность выводов диссертационных исследований подтверждается экспертными оценками и их соответствием результатам вычислительных экспериментов. Представленные в диссертации результаты прошли апробацию на международных и российских конференциях, принимались к публикации в рецензируемых журналах.

Представленные в диссертации результаты исследований опубликованы в 37 работах. Из них 22 статьи в ведущих рецензируемых научных журналах, входящих в список ВАК (в том числе 12 статей в журналах категории К1 и К2) 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 соискателя ученой степени раскрывают содержание основных положений диссертации и отражают персональный вклад автора в диссертационные исследования.

Значимость результатов диссертационных исследований для развития отрасли науки

Результаты исследования вносят существенный вклад в развитие методов анализа сложных сетей. Теоретическая значимость заключается в расширении математического аппарата обработки графов за счет учета атрибутивных данных при построении весовых функций; сочетания вычислительных процедур теории графов и учета атрибутивных данных; разработки многокритериальных методов оценки качества, сочетающих топологические и семантические параметры.

Прикладная ценность проявляется в создании технологий, адаптированных к особенностям социальных сетей и сетей мгновенного обмена сообщениями. Работа вносит существенный вклад в развитие методов анализа сложных сетей. Расширены границы применимости алгоритмов за счет интеграции атрибутивных данных и семантического анализа. Предложены новые методы оценки качества, сочетающие структурные и лингвистические параметры.

Результаты своевременно опубликованы, доложены научному сообществу, подтверждаются участием в научных проектах РФФИ, связанных с анализом коммуникационных сетей.

Диссертационная работа А.А. Чеповского вносит существенный вклад в развитие математического моделирования, расширяя методологический аппарат анализа сетевых структур.

Замечания по работе.

1. Некоторые из рассматриваемых коммуникационных сетей имеют тенденцию к постоянным изменениям их функциональных возможностей. В частности, анализ нестационарных процессов

динамики распределения кластеров по размерам хорошо известен в теории уравнения Смолуховского кинетической теории коагуляции и в моделях динамики дефектов в конструкционных материалах ядерных реакторов под действием радиационных потоков. В тексте диссертации не сказано, насколько описанная модель построения графов может быть применима в случае изменения функционала информационного сервиса и продолжают ли в текущий момент исследования с ее использованием.

2. Следовало бы чётко определять в тексте использование термина «случайный граф», т.к. это предполагает определённый закон распределения, например, «равномерное» или иное распределение, используемое при генерации графа.
3. В главе 6 подробно описана методика обработки текстовых данных на русском языке, используемая для оценки качества выделенных сообществ. Но не приведено данных о такой работе с другими языками.

Приведенные замечания не влияют на положительную оценку результатов диссертационных исследований и их изложения в тексте 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.

Представленные диссертационные исследования является выполненной самостоятельно завершённой научно-квалификационной работой, в которой содержится решение имеющей важное хозяйственное и социально-экономическое значение проблемы анализа графов, возникающих в информационно-коммуникационных технологиях.

Содержание диссертации полностью удовлетворяет 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), 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».

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

Официальный оппонент:

Доктор физико-математических наук, профессор,
профессор кафедры прикладной математики БУ ВО
«Сургутский государственный университет»

Галкин В.А.

«15» мая 2025 г.

Подпись Галкина Валерия Алексеевича заверяю
Ученый секретарь УС СурГУ, д.б.н.

Козлова В.В.

Выходные данные организации:
Бюджетное учреждение высшего образования
Ханты-Мансийского автономного округа – Югры
«Сургутский государственный университет»

Адрес: Россия, 628412, Ханты.Мансийский
автономный округ- Югра, г. Сургут, пр. Ленина, д. 1.
Телефон: 8:9120870620.
Адрес электронной почты: val-gal@yandex.ru.

