

На правах рукописи



**ЧЕПОВСКИЙ АЛЕКСАНДР АНДРЕЕВИЧ**

**МЕТОДЫ РАБОТЫ С НЕЯВНЫМИ СООБЩЕСТВАМИ НА  
ВЗВЕШЕННЫХ ГРАФАХ ВЗАИМОДЕЙСТВУЮЩИХ ОБЪЕКТОВ**

Специальность 1.2.2. – Математическое моделирование, численные методы и  
комплексы программ

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени

доктора физико-математических наук

Москва – 2025

Работа выполнена в Федеральном государственном автономном образовательном учреждении высшего образования «Национальный исследовательский университет «Высшая школа экономики».

Научный консультант: **Сигов Александр Сергеевич**, доктор физико-математических наук, профессор, академик РАН, президент ФГБОУ ВО «МИРЭА – Российский технологический университет».

Официальные оппоненты: **Галкин Валерий Алексеевич**  
доктор физико-математических наук, профессор,  
Бюджетное учреждение высшего образования Ханты-Мансийского автономного округа – Югры «Сургутский государственный университет» (СурГУ).

Профессор кафедры прикладной математики.

**Лаврентьев Михаил Михайлович**

доктор физико-математических наук, профессор,  
Федеральное государственное бюджетное учреждение науки  
Институт автоматики и электрометрии Сибирского отделения  
Российской академии наук (ИАиЭ СО РАН).

Заместитель директора по научной работе.

**Старков Сергей Олегович**

доктор физико-математических наук,  
Обнинский институт атомной энергетики (ИАТЭ)  
Национального исследовательского ядерного университета  
(НИЯУ) «МИФИ».

Начальник Отделения интеллектуальных кибернетических систем.

Ведущая организация: Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)».

Защита состоится «25» июня 2025 года в 12:00 часов на заседании на заседании диссертационного совета 24.2.332.02 при ФГБОУ ВО «Московский государственный технологический университет «СТАНКИН» по адресу: 127994, г. Москва, Вадковский пер., д. 1.

С диссертацией можно ознакомиться в библиотеке и на сайте ФГБОУ ВО «МГТУ «СТАНКИН», [www.stankin.ru](http://www.stankin.ru)

Автореферат разослан «\_\_\_» \_\_\_\_\_ 2025 г.

Ученый секретарь

диссертационного совета 24.2.332.02,  
к.т.н., доцент

Тюрбеева Татьяна Борисовна

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы исследования.** Системы, представляющие собой сетевые структуры, образованные взаимодействием между собой большого числа объектов, принято объединять под термином «сложные сети» (complex networks). Это, например, биологические, экологические, инфраструктурные, технологические, социальные сети. В качестве математической модели для таких сетей рассматриваются графы, в которых вершины соответствуют узлам сети, а ребра – связям между ними. При этом, как вершины, так и ребра могут обладать некоторой информацией, которая фиксируется как атрибуты соответствующих элементов множества вершин или множества ребер графа. Таким образом, получаются графы взаимодействующих объектов, анализ которых является существенной проблемой в области информационных технологий.

Актуальность рассматриваемой проблемы определяется «Стратегией национальной безопасности Российской Федерации», утвержденной Указом Президента РФ № 400 от 2 июля 2021 г., а именно задачами развития безопасного информационного пространства, защиты российского общества от деструктивного информационно-психологического воздействия. В частности, это задачи создания условий для эффективного предупреждения, выявления и пресечения преступлений и иных правонарушений, совершаемых с использованием информационно-коммуникационных технологий: задачи развития сил и средств информационного противоборства; задачи противодействия использованию информационной инфраструктуры Российской Федерации экстремистскими и террористическими организациями, специальными службами и пропагандистскими структурами иностранных государств для осуществления деструктивного информационного воздействия на граждан и общество.

Для информационно-аналитических систем важной составляющей является анализ графов взаимодействующих объектов, полученных из сетей передачи данных. Данный анализ есть существенная составляющая управления информационным пространством и аналитическими подсистемами,

применяемыми для обеспечения безопасности и контроля деятельности по распространению информации.

Решение данной проблемы имеет существенное значение в рамках современных разведывательных и контрразведывательных мероприятий, необходимости оценки источников размещения информации в социальных сетях, мессенджерах и выявления групп субъектов, использующих и активно поддерживающих данную информацию. Данные задачи актуальны для аналитических подразделений спецслужб и коммерческих структур, решающих как маркетинговые задачи, так и задачи борьбы с мошенничеством.

Анализ графов взаимодействующих объектов, включая построение методов выделения ключевой информации и разработку прикладного программного обеспечения для обработки данных, является важной составляющей для создания информационно-аналитических систем обеспечения безопасности. В частности, при работе с графами, полученными при импорте данных из социальных сетей и сетей мгновенного обмена сообщениями, особую ценность представляют следующие прикладные задачи: определение близости профилей пользователей, совпадения их интересов, степени (очного) знакомства; выявление наиболее активных единомышленников среди контактов заданного исходного пользователя, возможно напрямую с ним и не связанных; распознавание лидеров мнений; выявление каналов распространения и обмена информации между пользователями.

С точки зрения построения информационно-аналитических систем проблема анализа реальных графов взаимодействующих объектов влечет за собой необходимость решения следующих задач. Это разработка алгоритмов для выявления структуры графа; создание методик анализа сформированных данных, включая оценку корректности полученных результатов; программная реализация средств обработки графов больших размеров, включая создание специализированных эффективных хранилищ графов.

Последние 20 лет в области методов анализа структуры графа ведутся активные исследования по разработке алгоритмов выделения неявных сообществ на графах. Под выделением неявных сообществ на графе понимается разбиение

графа на подграфы, такое что плотность связей внутри этих подграфов значительно выше плотности связей между ними. Такое разбиение позволяет, в частности, переходить к выделению различных ролей у вершин графа. Один из самых интуитивно понятных и распространенных подходов к решению задачи выделения сообществ состоит в алгоритмах поиска разбиения графа на основе максимизации некоторого функционала, характеризующего качество разбиения и обычно называемого «модулярность». Различные аспекты выделения сообществ на графах рассматривали Newman M.E.J., Girvan M., Fortunato S., Blondel V. D. и другие авторы. Распространен также алгоритм на основе имитации условного динамического процесса на графе (Rosvall M., Bergstrom C. T.). При этом указанные наиболее развитые подходы не решают в полном объеме задачу выделения пересекающихся сообществ на графах взаимодействующих объектов, полученных из реальных данных о социальных коммуникациях, для которых характерны и играют существенную роль атрибуты ребер и вершин.

Наиболее спорной и практически открытой является проблема оценки корректности и эффективности работы алгоритмов и методов выделения сообществ на графах. Существует множество методов генерации случайных графов с последующим тестированием на них алгоритмов для получения оценки разбиения на сообщества (Lancichinetti A., Fortunato S., Radicchi F.). Есть иные методы, основанные на анализе разбиений анализируемого графа на основе оценки количества информации (Danon L., Díaz-Guilera A., Amelio A., Duch J., Arenas A., Pizzuti C.). Данные подходы не позволяют корректно оценить результаты работы с графами реальных сетей, особенно для графов, полученных при импорте данных из социальных сетей и сетей мгновенного обмена сообщениями.

Среди российских авторов работы, связанные с графами сложных сетей, в основном, носят либо обзорный характер (И.А. Евин, Н.Ф. Гусарова, Н.Г. Щербакова и др.), либо относятся к вопросам моделирования случайных графов и построения прогнозов их развития (А.М. Райгородский, В.Н. Задорожный, В.А. Бадрызлов и др.) и моделирования распространения информации в социальных сетях (Д.А. Губанов, Д.А. Новиков, А.Г. Чхартишвили).

Таким образом, актуальными являются разработка методов и алгоритмов выделения сообществ на реальных графах взаимодействующих объектов, формирование принципов тестирования алгоритмов, создание прикладного программного обеспечения, реализующего разработанные методы.

**Объект исследования** – графы взаимодействующих объектов, полученные из сетей передачи данных.

**Предмет исследования.** Предметом исследований являются разработка алгоритмов выделения сообществ на графах различной природы и моделирование методов формирования графов, полученных при импорте данных из социальных сетей и сетей мгновенного обмена сообщениями, включающих атрибутивные данные объектов и их взаимодействия.

**Целью** диссертационной работы является решение имеющей важное хозяйственное и социально-экономическое значение проблемы анализа коммуникационных данных, служащей существенным фактором в обеспечении технических и технологических подходов в сфере государственной и общественной безопасности, включая вопросы контроля информационного воздействия в социальных сетях и сетях мгновенного обмена сообщениями. Расширение средств и возможностей указанного контроля позволяет повысить безопасность информационного пространства, защиту общества от деструктивного информационно-психологического воздействия. Такое расширение включает в себя создание моделей, разработку 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.

**Научная проблема, имеющая важное хозяйственное значение,** состоит в создании методов, моделей и программного обеспечения для анализа структуры графов взаимодействующих объектов, полученных при импорте данных из социальных сетей и сетей мгновенного обмена сообщениями, с целью оценки информационного воздействия на субъектов через коммуникационные ресурсы. Решение данной проблемы должно предоставить средства для широкого класса

прикладных задач в сфере анализа данных коммуникаций, обеспеченных различными компьютерными технологиями. Указанные средства повышают возможности эффективного предупреждения, выявления и пресечения преступлений, предотвращению поддержки экстремистской и террористической деятельности, и иных правонарушений, совершаемых с использованием информационно-коммуникационных технологий.

Решение проблемы включает следующие задачи:

1. Построение для сетей передачи данных моделей графов взаимодействующих объектов, описывающих их информационное взаимодействие;
2. Построение и разработка итерационных численных методов и универсальных алгоритмов для выделения неявных сообществ и ключевых вершин графов с использованием эвристик;
3. Создание процедур по оценке качества выявленных сообществ на графе взаимодействующих объектов;
4. Разработка программного обеспечения, реализующего методы хранения, анализа и визуализации графов взаимодействующих объектов для сетей передачи данных.

**Методы исследования.** Для решения сформулированных проблем и поставленных задач использовалась методология информационного моделирования, аппарат решения экстремальных задач теории графов, и вычислительных методов оптимизации функционалов.

**На защиту выносятся следующие научные результаты:**

1. Модель формирования графа взаимодействующих объектов, полученного при импорте данных из социальных сетей и сетей мгновенного обмена сообщениями, описывающего информационное воздействие на объекты сети.
2. Итерационные численные методы для выделения неявных сообществ и ключевых вершин графов с использованием эвристик, а именно:
  - 2.1. «Комбинированный алгоритм» для выделения пересекающихся сообществ на графе, позволяющий убирать из рассмотрения малозначимые элементы

сети и предусматривающий параметрические модификации для формирования разнородных разбиений в зависимости от задач оператора.

2.2.«Метод ядра» для выделения непересекающихся сообществ на взвешенных графах, предусматривающий выделение ключевой компоненты на основании вычисляемых в явном виде характеристик графа.

2.3.«Метод Галактик» выделения пересекающихся сообществ на взвешенных графах, основанный на последовательном выделении сообществ и обработке исходного графа, переходе к мета-графу из мета-сообществ и последующем выделении итоговых пересекающихся сообществ.

3. Методика оценки корректности выделения пересекающихся сообществ на графах взаимодействующих объектов, основанная на анализе методами компьютерной лингвистики текстов – атрибутов вершин выделенных сообществ.

4. Программное обеспечение для построения графа взаимодействующих объектов (импорта данных из сетей коммуникации) и его последующего анализа, включающее:

- средства хранения графов;
- реализации алгоритмов выделения сообществ;
- средства графического отображения обнаруженной структуры графа.

Перечисленные научные результаты представляют основу нового научного направления «комплексный анализ структуры графов взаимодействующих объектов», сочетающего построение и анализ структуры соответствующих графов с целью описания информационного взаимодействия, в том числе в социальных сетях и сетях мгновенного обмена сообщениями.

**Научная новизна:**

- предложена модель формирования взвешенного графа взаимодействующих объектов при импорте данных из социальных сетей и сетей мгновенного обмена сообщениями, характеризующая информационное взаимодействие;
- построены итерационные числительные методы и алгоритмы для выделения неявных сообществ и ключевых вершин графов с использованием эвристик;

- предложен и реализован «Комбинированный алгоритм» выделения пересекающихся и вложенных сообществ на графе, позволяющий убирать из рассмотрения малозначимые элементы сети;
- предложен и реализован «Метод ядра» для выделения непересекающихся сообществ на взвешенных графах, предусматривающий выделение ключевой компоненты на основании вычисляемых в явном виде характеристиках графа;
- предложен и реализован «Метод Галактик» для выделения пересекающихся сообществ на взвешенных графах;
- разработана методика оценки эффективности выделения сообществ на графе с помощью алгоритмов компьютерной лингвистики для обработки текстовых метаданных – атрибутов вершин выделенных сообществ;
- разработана модель для эффективного хранения графов взаимодействующих объектов, основанная на алгоритмах сжатия и оптимизации операций с графами по памяти и по скоростным характеристикам.

**Теоретическая значимость результатов работы** состоит в том, что предложены новые методы и алгоритмы выделения неявных сообществ на графе взаимодействующих объектов, опирающиеся на структурные особенности графа, и в том, что предложены принципиально новые методы для оценки качества решения задачи по выделению сообществ.

**Практическая значимость** состоит в построении методов и алгоритмов для решения проблемы, имеющей важное хозяйственное значение, формировании новых подходов для анализа получаемых решений и реализации в комплексе программ. Получены свидетельства о регистрации программ для ЭВМ. Научные и практические результаты диссертации использованы в 2016 – 2021 годах в грантах РФФИ (участник грантов РФФИ): 16-07-00641 А. «Исследование и разработка математических моделей, методов и алгоритмов визуализации и анализа графов на примере социальных сетей». (2016-2018 г.г.); 16-29-09546 офим-м. «Разработка новых методов мониторинга и комплексного лингвистического и тематического анализа сообщений социальных медиа в целях противодействия экстремизму и

терроризму». (2016-2019 г.г.); 18-00-00233 КОМФИ. «Методы комплексного интеллектуального анализа информации различных типов для социо-гуманитарных исследований в социальных медиа». (2018-2020 г.г.); 19-07-00806 А. «Исследование и разработка методов и алгоритмов для создания и комплексного лингвистического анализа специализированных корпусов текстов». (2019-2021 г.г.). Результаты диссертации использованы в учебных пособиях. Результаты диссертации использованы в учебных пособиях [34 – 35].

**Достоверность результатов** и обоснованность научных положений диссертационной работы обеспечивается корректным использованием методов теории графов и соответствующего математического аппарата. Достоверность полученных выводов подтверждается согласованностью с результатами экспериментальных исследований и экспертной оценкой результатов. Достоверность результатов работы подтверждается работоспособностью предложенных методик в проведенных экспериментальных вычислениях с помощью разработанного программного обеспечения. Результаты прошли апробацию на международных и российских конференциях, принимались к публикации в рецензируемых журналах.

**Апробация работы.** Материалы отдельных разделов диссертации докладывались на международных конференциях:

- Международная конференция «Математика в созвездии наук» К юбилею ректора МГУ, академика В.А. Садовниченко, 1-2 апреля 2024 года. Секция «Дискретная математика, математическая кибернетика и теория интеллектуальных систем»;
- COMPLEX NETWORKS 2020. The 9th International Conference on Complex Networks and their Applications. December 1-3, 2020 - Madrid, Spain.;
- VII Международная научно-практическая конференция «Управление информационной безопасностью в современном обществе», Москва, 29 мая – 30 мая 2019 г.;

- The 6th International Conference on Complex Networks & Their Applications. Nov. 29 - Dec. 01, 2017, Lyon (France);
- V Международная научно-практическая конференция «Управление информационной безопасностью в современном обществе», Москва, 30 мая – 1 июня 2017 г.;
- IV Международная конференция «Управление информационной безопасностью в современном обществе», Москва, 31 мая – 02 июня 2016, Москва.
- 5 Международная конференция «Ситуационные центры и геоинформационные системы для задач мониторинга и безопасности (SCVRT 2016)», 21-25 ноября 2016, г. Пущино, Московская обл.;
- Международная научная конференция Resilience2014 Международного Центра по ядерной безопасности Института физико-технической информатики, Протвино, 2014 г.;
- Международная научная конференция по физико-технической информатике СРТ2014. Протвино, 2014;
- Международная научная конференция Международного центра по ядерной безопасности Института физико-технической информатики SCVRT2013, Протвино, 25-29 ноября 2013 г.

Материалы диссертационных исследований обсуждались на семинарах МИЭМ НИУ ВШЭ, научно-исследовательской лаборатории «Суперкомпьютерные технологии и машинное обучение» СПбПУ (май 2023 г.), Института перспективных технологий и индустриального программирования МИРЭА (март 2023 г.), отделения интеллектуальных кибернетических систем ИАТЭ НИЯУ «МИФИ» (апрель 2023 г.), кафедры Теоретической информатики механико-математического факультета МГУ имени М.В. Ломоносова (ноябрь 2023 г.).

**Публикации.** Основные результаты диссертационного исследования опубликованы в 37 работах [1 – 37]. Из них 22 статьи [3 – 24] 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) [17 – 24]. В том числе 12 статей в журналах категории K1 и K2 по распределению на 2022 и 2023 годы. Две рецензируемые монографии [1, 2], 13 публикаций [25 – 37] в трудах международных научных конференций. Два свидетельства о регистрации программ для ЭВМ [40, 41]. Результаты диссертации использованы в 2 учебных пособиях [38, 39].

**Личный вклад автора.** Основные научные результаты получены автором самостоятельно. Результаты, полученные соискателем лично, в работах, опубликованных в соавторстве:

модель формирования взвешенного графа информационного взаимодействия для разных сетей и методики применения модели в [6], [8], [17], [36], [37];

модификации классического алгоритма на основе оценки энтропии сети в [22], [23], [39];

модификации классического агломеративного иерархического алгоритма [10], [29], [33];

алгоритмы метода «Комбинированный алгоритм» и методика их применения, анализ результатов применения в [15], [28];

метод «Метод ядра», процедуры алгоритма, методики применения и анализ результатов применения в [9], [26];

алгоритмы метода «Метод Галактик» и методика их применения, анализ результатов применения в [7];

методики применения разработанных алгоритмов в [13], [20], [35];

применение методов анализа графов взаимодействующих объектов, полученных при импорте данных из социальных сетей, для формирования психологических показателей социального взаимодействия в [2], [19];

метод оценки корректности выделения сообществ на графе с помощью алгоритмов компьютерной лингвистики в [3], [5], [9], [11], [16], [34], [38];

структуры данных для хранения и анализа графов больших размеров в [12], [14], [21], [24], [31], [32], [36];

архитектура комплекса для анализа графов больших размеров и алгоритмы визуализации в [30], [35], [39].

Другим соавторам принадлежат программные реализации моделей и алгоритмов, анализ данных на сторонних программных продуктах, сопровождение разработанного программного обеспечения.

**Соответствие специальности.** Направление диссертационного исследования соответствует паспорту специальности 1.2.2. «Математическое моделирование, численные методы и комплексы программ», а именно, п. 1 «Разработка новых математических методов моделирования объектов и явлений» (физико-математические науки), п. 2 «Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий» и п. 8 «Комплексные исследования научных и технических проблем с применением современной технологии математического моделирования и вычислительного эксперимента».

**Структура и объем 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.

## СОДЕРЖАНИЕ РАБОТЫ

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

В **первой главе** «Неявные сообщества на графах взаимодействующих объектов» приведен обзор текущих и перспективных задач анализа графов, представляющих сложные сети взаимодействующих объектов. Описаны предметные области, для которых актуальны данные модели, рассматривается практика и проблемы анализа графов, в том числе полученных при импорте данных из социальных сетей и сетей мгновенного обмена сообщениями. Приведен обзор основных классических подходов по выделению неявных сообществ на графах. Освещена проблема тестирования получаемых разбиений и оценки их качества. Показана актуальность подходов, предложенных в диссертации.

Описаны проблемы создания хранилищ графов и программно-аналитические комплексы для визуализации и анализа графов взаимодействующих объектов.

*Для графов, представляющих реальные сети, часто можно выделить подграфы с высокой плотностью ребер внутри них и сравнительно низкой плотностью ребер между такими подграфами. Разбиение графа на такие подграфы и рассматривается как задача выделения неявных сообществ.*

Сделаны выводы о том, что в реальных сетях свойства объектов и связей между ними во многих случаях оказывают ключевое влияние на структуру сети. Поэтому одним из актуальных направлений исследований видится построение взвешенных графов взаимодействующих объектов на базе импорта данных из реальных сетей с последующим выделением на них неявных сообществ, в том числе с использованием механик, основанных на переходе от анализа всего исходного графа к его подграфам.

Во **второй главе** «Построение графа взаимодействующих объектов» изложены разработанные 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. В построенном графе вершинами являются пользователи, а ребра и вес на них описывают интенсивность или характер их взаимодействия.

Для семейства моделей вес  $w(e)$  на ребрах графа  $G(V, E)$  задается в общем случае взвешенной суммой  $N$  факторов взаимодействия объектов:

$$w(e) = \sum_k^N W_k \cdot \delta_e^k, \quad (2.1)$$

где  $W_k$  – весовой коэффициент для  $k$ -го фактора взаимодействия;

$\delta_e^k$  – функция, принимающая нулевое значение в случае отсутствия  $k$ -го фактора взаимодействия у ребра  $e$  и натуральные значения в иных случаях в зависимости от степени интенсивности взаимодействия, заложенной в конкретной модели.

Построенные по данной модели взвешенные графы взаимодействующих объектов обозначаются как  $G(V, \tilde{E})$ .

Рассматриваются методики формирования графов для трех типов сетей, примеры реальных данных из которых используются в работе для исследований: это социальная сеть *ВКонтакте*, сеть коротких сообщений *Twitter* и сеть каналов мессенджера *Telegram*.

Разработана и реализована  $(F, L, C, R)$ -модель информационного взаимодействия в сети *Twitter*, предусматривающая формирование взвешенного графа  $G(V, \tilde{E})$ , у которого на множестве ребер весовая функция  $w(e)$  с неотрицательными значениями задана следующим образом:

$$w(e) = F \cdot \delta_e^F + L \cdot \delta_e^L + C \cdot \delta_e^C + R \cdot \delta_e^R, \quad (2.2)$$

где:

$$\delta_e^F = \begin{cases} 2, & \text{если оба инцидентных пользователя подписаны друг на друга} \\ 1, & \text{если хотя бы один из инцидентных пользователей подписан на другого} \\ 0, & \text{иначе;} \end{cases}$$

$$\delta_e^L = \begin{cases} 2, & \text{если между инцидентными пользователями есть взаимные Like} \\ 1, & \text{если есть хотя бы один Like между инцидентными пользователями} \\ 0, & \text{иначе;} \end{cases}$$

$$\delta_e^C = \begin{cases} 2, & \text{если между инцидентными пользователями есть взаимные Reply} \\ 1, & \text{если есть хотя бы один Reply между инцидентными пользователями} \\ 0, & \text{иначе;} \end{cases}$$

$$\delta_e^R = \begin{cases} 2, & \text{если между инцидентными пользователями есть взаимные Retweet} \\ 1, & \text{если есть хотя бы один Retweet между инцидентными пользователями} \\ 0, & \text{иначе;} \end{cases}$$

и  $F, L, C, R$  – веса соответствующих факторов взаимодействия, фиксированные для всего графа.

В рамках модели (2.2) построены конкретные реализации посредством импорта данных из сети *Twitter*.

Дополнительно представлена модель построения хронологических подграфов из сети *Twitter*, где для каждого из заданных временных интервалов строится взвешенный граф  $G(V, \tilde{E})$ , характеризующий ситуацию в сети по параметру  $t$ . Для данной модели вес  $w(e_{ij})$  для каждого ребра  $e_{ij} \in E$ , инцидентного вершинам  $i$  и  $j$  определяется следующим образом:

$$w(e_{ij}) = 1 \cdot \delta_e^F + 2 \cdot likes_{ij} + 2 \cdot likes_{ji}, \quad (2.3)$$

где  $\delta_e^F$  определяется аналогично модели (2.2);

$likes_{ij} = 1$ , если за временной интервал  $t$  на постах пользователя, соответствующего вершине  $i$ , есть хотя бы один лайк от пользователя, соответствующего вершине  $j$ , иначе  $likes_{ij} = 0$ ;

$likes_{ji} = 1$ , если за временной интервал  $t$  на постах пользователя, соответствующего вершине  $j$ , есть хотя бы один лайк от пользователя, соответствующего вершине  $i$ , иначе  $likes_{ji} = 0$ .

В рамках модели (2.2) так же построены конкретные реализации посредством импорта данных из сети *Twitter*.

Разработана и реализована  $(U, M, R)$ -модель информационного воздействия в сети каналов мессенджера *Telegram*, предусматривающая формирование графа взаимодействующих объектов  $G(V, \tilde{E})$  из первичного графа  $G(V, E)$ , для которого  $V$  – множество вершин, представляющих каналы;  $E$  – множество всех возможных ребер, определяемых из взаимодействий между соответствующими парами каналов. Модель предусматривает выделение в мессенджере *Telegram* трех видов взаимодействий между каналами за временной промежуток  $T$ : наличие общих

внешних *URL*, упоминания одними каналами других и «репосты» между каналами. На множестве ребер  $E$  задается весовая функция  $w(e_{ij})$  ( $e_{ij} \in E; i, j \in V$ ) следующим образом:

$$w(e_{ij}) = U \cdot \delta_{e_{ij}}^U + M \cdot \delta_{e_{ij}}^M + R \cdot \delta_{e_{ij}}^R, \quad (2.4)$$

где  $\delta_{e_{ij}}^U$  – количество общих уникальных внешних ссылок в постах у соответствующих вершинам каналов за выбранный период  $T$ ;

$\delta_{e_{ij}}^M$  – суммарное количество за выбранный период  $T$  уникальных записей в одном канале, где упоминается другой канал и наоборот для соответствующих вершин;

$\delta_{e_{ij}}^R$  – суммарное количество за выбранный период  $T$  уникальных записей в одном канале, где цитируется (сделан «репост») другой канал и наоборот для соответствующих вершин;

$U, M, R$  – значения для  $W_k$  из формулы (2.1), фиксированные для всего графа, в данном случае значение  $N = 3$ .

В рамках модели (2.4) построены множественные примеры реализации посредством импорта данных из сети мгновенного обмена сообщениями *Telegram*.

Модели для построения взвешенных графов взаимодействующих объектов для социальной сети *ВКонтакте* основаны на обходе сети для импорта данных и применения формулы (2.1) в зависимости от выбранного набора факторов взаимодействия.

Разработаны и реализованы вариации модели для построения взвешенных графов взаимодействующих объектов, которые можно использовать при импорте данных из социальной сети *ВКонтакте*. Согласно этим моделям, строятся следующие типы графов: граф общего сходства пользователей, граф симпатии пользователей, граф информационного взаимодействия пользователей, граф чистого информационного взаимодействия пользователей, параметрический смешанный граф пользователей. Данные модели были апробированы на реальных данных из социальной сети *ВКонтакте*.

Приведены экспериментальные исследования предложенной в главе модели на примере оценки различных весовых функций. Показано, что данная модель позволяет строить взвешенные графы взаимодействующих объектов со степенным распределением веса вершин. Причем показатель степени находится в диапазоне характерном для сложных сетей. Разработанные модели позволяют из указанных сетей разного типа формировать графы взаимодействующих объектов для дальнейшего их анализа.

В третьей главе «Комбинированный алгоритм выделения сообществ» изложен разработанный метод, основанный на классических алгоритмах, в основу оценки результативности которых положено понятие «модулярность».

При рассмотрении графа, имеющего структуру сообществ, ожидается, что плотность ребер внутри этих сообществ будет выше, чем плотность ребер в остальном графе. Но также можно ожидать, и что плотность ребер внутри сообществ в таком графе должна быть выше, чем плотность ребер в графе, не обладающим структурой сообществ. Случайно выбранный граф, моделирующий некоторые свойства исходной сети, но не обладающий структурой сообществ, называется нулевой моделью. В идее модулярности используется не оценка, насколько внутри сообществ плотность ребер выше плотности ребер между ними, а насколько плотность выше по сравнению с нулевой моделью. Происходит сравнение с «нулевой гипотезой», заключающейся в том, что ребра распределены случайно, т.е. нет закономерностей в распределении плотности ребер внутри групп.

В нулевой модели Ньюмана-Гирвана, предполагающей сохранение степеней вершин графа вместе со случайным распределением ребер между ними, модулярность для разбиения на непересекающиеся сообщества определяется по следующей формуле:

$$Q = \frac{1}{2m} \sum_{ij} (A_{ij} - \frac{d_i \cdot d_j}{2m}) \cdot \delta(S_i, S_j), \quad (3.1)$$

где  $A_{ij}$  – элемент матрицы смежности;

$m$  – число ребер в графе;

$d_i$  – степень  $i$ -ой вершины;

$S_i$  – сообщество, к которому принадлежит  $i$ -ая вершина.

Для оценки качества разбиения графа на пересекающиеся сообщества применяется следующий вариант подсчета модулярности:

$$Q^{overlap} = \frac{1}{2m} \sum_i \sum_{v \in S_i, w \in S_i} \frac{1}{O_v O_w} \left[ A_{vw} - \frac{d_v \cdot d_w}{2m} \right], \quad (3.2)$$

где  $A_{vw}$  – элемент матрицы смежности;

$m$  – число ребер в графе;

$d_v$  – степень вершины  $v$ ;

$O_v$  – количество сообществ, которым принадлежит вершина  $v$ ;

$S = \{S_i\}$  – множество пересекающихся сообществ.

Разработанный Комбинированный алгоритм выделения пересекающихся сообществ строится как последовательное применение классических алгоритмов *Louvain* и *Clique Percolation Method (CPM)* с дополнительной параметризацией с целью удаления малозначимых элементов сети.

Агломеративный иерархический алгоритм *Louvain* состоит из двух этапов, которые повторяются итерационно. Изначально каждой вершине графа назначается свое сообщество. Затем, на первом этапе, для каждой вершины  $i$  рассматриваются соседние вершины  $j$  и вычисляется прирост модулярности, который может иметь место при удалении вершины  $i$  из своего сообщества и добавления ее в сообщество вершины  $j$ . Вершина  $i$  переносится в то сообщество, где достигается максимальный положительный прирост значения модулярности. Второй этап алгоритма заключается в построении нового графа, чьи вершины являются сообществами, найденными на первом этапе алгоритма, а вес ребер нового графа определяется как сумма весов ребер между всеми парами вершин, лежащих в разных сообществах.

В работе обоснованы особенности структуры сообществ, выделяемой алгоритмом *Louvain* на графе. Доказано отсутствие тривиальных сообществ в разбиении, основанном на максимизации модулярности. Описано формирование сообществ с листовыми вершинами, и следствие из этого – свойство для смежных

большому числу листов вершин попадать в сообщества с малым суммарным весом. Представлен известный «предел разрешения» – в максимальном по модулярности разбиении невзвешенного неориентированного графа не может существовать двух связанных сообществ, степень каждого из которых меньше, чем  $\sqrt{2m}$ . В совокупности эти свойства приводят к эффекту, названному автором «сбор мусора». Приведены подходы по борьбе с этим эффектом. В данной главе представлена параметрическая модулярность для корректировки «предела разрешения»:

$$Q(\alpha) = \frac{1}{2m} \sum_{ij} (A_{ij} - \alpha \frac{d_i \cdot d_j}{2m}) \cdot \delta(S_i, S_j), \quad (3.3)$$

где параметр  $0 < \alpha \leq 1$ .

Показано, что в использование параметрической модулярности позволяет снизить минимальный размер двух связанных сообществ до  $\sqrt{2m/\alpha}$ .

Приведены вычислительные и экспериментальные данные для выявленных особенностей структуры сообществ.

В алгоритме *CPM* нахождение пересекающейся структуры сообществ связано с задачей нахождения  $k$ -клик в графе. Для фиксированного  $k$  алгоритм сперва находит все  $k$ -клики графа. Затем, если две  $k$ -клики имеют  $k - 1$  общие вершины, то они считаются смежными друг с другом. Сообщество определяется максимальным набором смежных  $k$ -клик.

Идея Комбинированного алгоритма состоит в том, что на первом этапе граф разделяется на крупные непересекающиеся сообщества с помощью алгоритма *Louvain*, а затем на втором этапе в случае выполнения определенных условий внутри них находятся пересекающиеся сообщества меньшего размера с помощью алгоритма *CPM*. К каждому отдельному выделенному на первом этапе Комбинированного алгоритма сообществу применяется второй этап, только если для этого сообщества верно условие, что доля вершин с высокой центральностью по посредничеству не превосходит заранее заданного значения. На третьем этапе происходит перераспределение «мусора» (оставшихся отдельными вершин и

небольших сообществ) по входным параметрам алгоритма  $i_{opt}$  и  $m_{opt}$ , отвечающим за действия с оставшимися вершинами и маленькими сообществами соответственно.

Данный метод включает в себя как комбинацию двух классических алгоритмов (для выделения пересекающихся и не пересекающихся сообществ), так и дополнительную параметризацию на этапе обработки результатов между их применением. Что позволяет убирать из рассмотрения малозначимые элементы сети (провести процедуру «сбора мусора») или выделять их в отдельные сообщества в зависимости от задач оператора.

Таким образом, на первом этапе ищется  $S = \{S_0, S_1, S_2, \dots, S_r\}$  – такое множество непересекающихся сообществ на графе  $G$ , для которого достигается локальный максимум модулярности  $Q_{max}$ , где  $(r + 1)$  – число сообществ, выделенных на этом этапе. В результате любая вершина графа  $G$  лежит в каком-то одном сообществе  $C_i$ , а каждое такое сообщество входит в  $S$ :

$$Q_{max} = \max_{S: C_i \in S \forall i} Q \quad (3.4)$$

Далее, на втором шаге для всех  $S_k \in S$  проверяется условие:

$$\frac{\sum_{v \in S_k} \alpha(v)}{n_{S_k}} \leq \beta(n_{S_k}), \quad \alpha(v) = \begin{cases} c_B(v), & c_B(v) \geq \tilde{\alpha} \\ 0, & c_B(v) < \tilde{\alpha} \end{cases} \quad (3.5)$$

где  $n_{S_k} = |S_k|$ ;  $c_B(v)$  – центральность по посредничеству вершины  $v$ ;

$\tilde{\alpha}$  – пороговое значение, заданное исходно на полуинтервале  $[0; 1]$ ;

$\beta(n_{S_k})$  – задан изначально на полуинтервале  $(0; 1]$ .

Далее внутри каждого сообщества  $S_k$ , удовлетворяющего условию (3.5), ищется такое множество его внутренних сообществ  $S_k^{overlap} = \{S_k^0, S_k^1, S_k^2, \dots, S_k^{r_k}\}$ , при котором достигается локальный максимум модулярности:

$$Q_{S_k, max}^{overlap} = \max_{S_k^{overlap}} Q^{overlap} \quad (3.6)$$

На последнем шаге производится перераспределение «мусора» на основе двух параметров:  $i_{opt}$  и  $m_{opt}$ . Первый используется для распределения не вошедших на

втором этапе в клики вершин. Они могут быть выделены в индивидуальные сообщества или объединены в отдельное сообщество внутри рассматриваемого. Второй параметр аналогично определяет действия для маленьких (меньших размера клики) сообществ внутри  $S_k$ .

Были проведены вычислительные эксперименты применения Комбинированного алгоритма и классических алгоритмов на более чем 70 графах различных взаимодействующих объектов: социальные сети, сети цитирования, белок-белковые взаимодействия и т.п. Качественный экспертный анализ результатов применения Комбинированного алгоритма и классических алгоритмов на ряде графов, показал его эффективность, особенно в части работы с маленькими выделенными сообществами.

Алгоритмы применялись на графах социальной сети *ВКонтакте*, полученных из реальных данных методами, описанными в главе 2.

Разработан алгоритм выделения пересекающихся сообществ с несколькими степенями параметризации, среди которых можно подбирать наиболее подходящие значения для графов разной природы. В том числе для полученных при импорте данных из социальных сетей и сетей мгновенного обмена сообщениями.

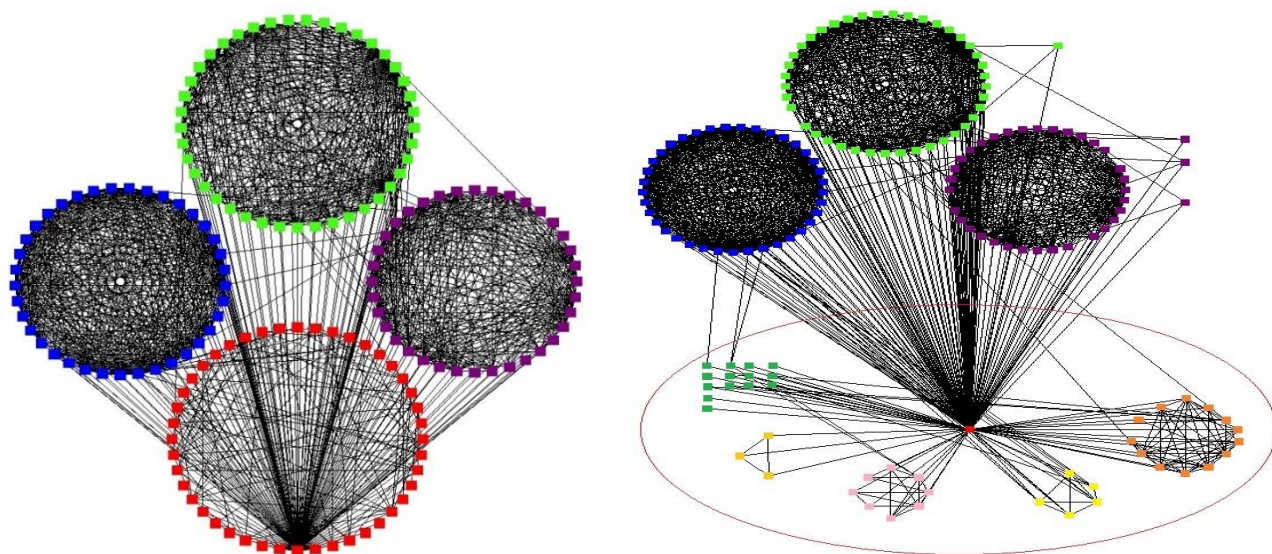


Рис 1. Пример применения Комбинированного алгоритма: выделение сообществ на первом этапе (слева) и итоговый результат выделения сообществ (справа)

Иллюстрация применения приведена на рисунке 1. На первом шаге выделяется 4 сообщества. На последующих шагах Комбинированного алгоритма выделяются новые сообщества внутри выделенных ранее.

В данной главе предложен итерационный алгоритм с модифицированным подсчетом весов на основании близости значений характеристик (атрибутов) у пар вершин. Он представляет собой модификацию алгоритма, в основе которого лежит идея сжатия информации о процессе случайного блуждания в графе за счет его кодирования.

В этой части работы также продемонстрирована методика анализа цифровых профилей социальных сетей на основе выделения сообществ на графах взаимодействующих объектов.

В **четвертой** главе «Метод ядра выделения сообществ» изложен разработанный метод, дающий комплексный подход к анализу взвешенного графа  $G = G(V, \tilde{E})$ .

Введены определения « $\alpha$ -звезды», « $\gamma$ -ядра», коэффициента взаимодействия, которые позволяют сформулировать через весовую функцию  $w$ , заданную на множествах вершин и ребер численное описание процедур, выполняемых для решения основных задач распознавания лидеров мнений и выявления каналов распространения и обмена информации между пользователями сети.

Назовем  $\alpha$ -звездой такие вершины  $v$  графа  $G$ , что  $v$  имеет вес, больший некоторого заданного для этого графа значения  $\alpha$ . Множество звезд  $Star_\alpha(G)$  тогда определяется следующим образом:

$$Star_\alpha(G) = \{v \in V \mid w(v) > \alpha\} \quad (4.1)$$

Для заданного разбиения  $S = \{S_0, S_1, S_2, \dots, S_r\}$  графа  $G$  на сообщества обозначим вес  $w(S_i)$  каждого из них как сумму весов входящих в это сообщество вершин. Сообщество с максимальным для заданного разбиения весом обозначим через  $Core(G)$ , его вес соответственно, как  $w(Core(G))$ . Если в исходном графе выделяются непересекающиеся сообщества  $S_i$ , вес которых очень близок или даже совпадает с максимальным весом  $w(Core(G))$ , то можно говорить о наличии

нескольких сообществ-ядер в графе  $G$ . Для заданной степени близости  $\gamma$  определим  $\gamma$ -ядро  $Core_\gamma(G)$ , как множество вершин из сообществ, которые удовлетворяет соотношению:

$$Core_\gamma(G) = \left\{ v \in V \mid v \in S_i : \frac{w(S_i)}{w(Core(G))} > \gamma \right\} \quad (4.2)$$

Исходя из определения  $\gamma \in [0; 1)$ , причем для выявления множества ключевых вершин графа требуется нахождение значения  $\gamma$ , близкого к 1.

Введено понятие внутреннего веса сообщества  $w^*(S_i)$  – суммы весов ребер, обе вершины которых лежат внутри сообщества  $S_i$ .

Упростить задачу анализа графа можно, проигнорировав не самые активные взаимодействия между объектами за счет удаления ребер, имеющих вес менее заданного  $\beta$ . После чего часть вершин станут изолированными и так же удаляются из анализируемого графа. Получим в итоге новый граф  $G'(V', E')$ . Метод ядра применяется после такой предобработки исходного графа уже на графе  $G'(V', E')$  активного информационного взаимодействия.

В работе вводится  $k_{S_i}(G')$  – коэффициент плотности сообщества  $S_i$  как отношение суммарного веса ребер внутри сообщества к суммарному весу ребер всего графа:

$$k_{S_i}(G') = \frac{w^*(S_i)}{\sum_{e \in E'} w(e)} \quad (4.3)$$

Введем коэффициент взаимодействия  $k_{int}(G')$  – отношение удвоенного числа ребер к квадрату числа вершин в графе  $G'$  активного информационного взаимодействия:

$$k_{int}(G') = \frac{2|E'|}{|V'|^2} \quad (4.4)$$

В работе показано, что  $k_{int}(G') \in (0; 1)$ , а для больших полных графов  $\lim_{n \rightarrow \infty} k_{int}(G') = 1$ . Коэффициент взаимодействия  $k_{int}(G)$  аналогично может быть подсчитан и для исходного графа  $G$ .

На основании введенных понятий и выявленных взаимосвязей построен алгоритм под названием «Метод ядра» для выделения непересекающихся

сообществ на взвешенных графах, предусматривающий выявление ключевой компоненты на основании вычисляемых в явном виде характеристик графа.

Алгоритм. Метод ядра для анализа взвешенных графов.

1. Удалить изолированные вершины. Наличие таковых может быть обусловлено особенностями исходной сети и процесса импорта данных. Получаем граф  $G$ .
2. Подсчитать исходный коэффициент взаимодействия графа  $k_{int}(G)$ . Он важен как ориентир для последующих шагов.
3. Убрать мусор. Необходимо определить значение  $\beta$  и выполнить операции по удалению ребер. Получаем граф  $G'$ . Если  $G' = G$  перейти к шагу 5.
4. Подсчитать обновленный коэффициент взаимодействия графа  $k_{int}(G')$ . Рекомендуемый диапазон изменения для графов из социальных сетей лежит в следующих пределах:  $0,8 < \frac{k_{int}(G')}{k_{int}(G)} < 0,9$ . В случае, если коэффициент изменился не в этом диапазоне, рекомендуется вернуться к шагу 3 и сделать его с иным значением  $\beta$ . Диапазон изменений может быть и иным, определяется опытным путем на массиве графов схожей природы.
5. Применить алгоритм разбиения на сообщества. Предполагается использование алгоритма, выделяющего непересекающиеся сообщества. Например, можно использовать вариации алгоритмов из главы 3.
6. Определить звезды. Выбрать значение  $\alpha$  и выделить множество вершин  $Star_\alpha(G')$ , состоящее из звезд графа. Проверить, что в наиболее крупных сообществах выделены звезды и такие сообщества имеют высокий показатель  $k_{S_i}(G')$ . Если нет, то поменять значение  $\alpha$ .
7. Выделить ядро. Определить  $\gamma$  и составить  $Core_\gamma(G')$ .

Данным алгоритмом в том числе выделяются вершины, осуществляющие наибольшее информационное воздействие на сеть, и наиболее активно взаимодействующие между собой вершины сети.

Таким образом, для графа  $G'$  активного информационного взаимодействия после нахождения локального максимума модулярности  $Q_{max}$  и соответствующего ему множества непересекающихся сообществ  $S$  Методом ядра находится значение параметра  $\alpha$ , для которого  $Star_\alpha(G)$  содержит вершины графа, осуществляющие основное информационное воздействие на остальных. После чего находится множество ключевых вершин графа  $Core_\gamma(G')$ :

$$\gamma \rightarrow \max_{Core_\gamma(G') \supset Star_\alpha(G')} \quad (4.5)$$

Рассмотрены конкретные примеры реализации метода на графах, полученных при импорте реальных данных из сети *Twitter*. Пример визуального представления результатов применения метода приведен на рисунке 2.

Апробация метода продемонстрирована на реальных данных из сети *Twitter*. Показано, в том числе с помощью методик, описанных в главе 6, что Метод ядра позволяет эффективно решать задачи распознавания лидеров мнений.

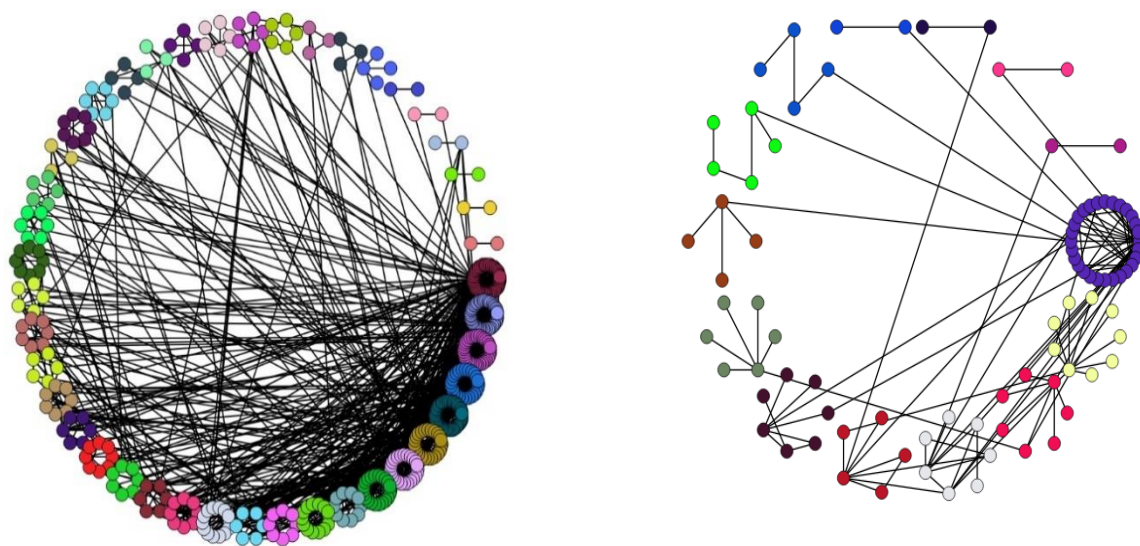


Рис.2 Пример применения метода ядра: разбиение графа на неявные сообщества (слева) и разбиение одного из полученных подграфов на внутренние сообщества (справа).

В пятой главе «Метод Галактик выделения сообществ» описан алгоритм работы со взвешенными графами  $G(V, \tilde{E})$ , полученными при импорте данных из сети мгновенного обмена сообщениями *Telegram* в соответствии с  $(U, M, R)$ -

моделью информационного воздействия (глава 2), где под  $\tilde{E}$  понимается множество ребер графа после удаления ребер с нулевым весом.

Предложен и реализован метод Галактик для выделения пересекающихся сообществ на взвешенных графах, основанный на конструировании алгоритма из базовых алгоритмов, обработке графа, переходам к мета-графу из мета-сообществ и последующем выделении пересекающихся сообществ.

Алгоритм «метод Галактик» для выделения неявных сообществ.

Шаг 1. На графе  $G(V, \tilde{E})$  выделяются пересекающиеся сообщества. Получаем набор сообществ  $S_{G(V, \tilde{E})}$ .

Шаг 2. Строится множество мета-вершин  $\tilde{S}_{G(V, \tilde{E})}$ , которое состоит из элементов  $S_{G(V, \tilde{E})}$ , содержащих более, чем одну исходную вершину.

Шаг 3. Для множества  $\tilde{S}_{G(V, \tilde{E})}$  вершин нового мета-графа  $\tilde{G}$  вычисляются веса его ребер и строится сам граф  $\tilde{G}$ .

Шаг 4. На графе  $\tilde{G}$  выделяются непересекающиеся сообщества.

Таким образом, на шаге 1 для графа  $G(V, \tilde{E})$  и модулярности  $Q_{G(V, \tilde{E})}^{overlap}$ , определенной по аналогии с (3.2), ищется набор сообществ  $S_{G(V, \tilde{E})}$  для локального максимума модулярности  $Q_{G(V, \tilde{E})}^{overlap}$ . Множество сообществ после «уборки мусора» на 2 шаге обозначим как  $\tilde{S}_{G(V, \tilde{E})} = \{\tilde{S}_0, \dots, \tilde{S}_{r_2}\}$ . Из этого множества и строится множество мета-вершин для нового графа  $\tilde{G}$ .

На 4 шаге для модулярности  $Q_{\tilde{G}}$ , определенной по аналогии с (3.1) ищется набор сообществ  $\tilde{\tilde{S}}_{\tilde{G}} = \{\tilde{\tilde{S}}_0, \dots, \tilde{\tilde{S}}_{r_3}\}$  для локального максимума:

$$Q_{\tilde{G}} \rightarrow \max_{\tilde{\tilde{S}}_{\tilde{G}}} \quad (5.1)$$

А так как  $\tilde{\tilde{S}}_i = \{\tilde{S}_{i_0}, \dots, \tilde{S}_{i_r}\}$  для всех  $i$ , и  $\tilde{S}_{i_k} = \{v_{i_{k1}}, \dots, v_{i_{kl}}\}$  для всех  $k$ , то получаем разбиение графа  $G(V, \tilde{E})$  на пересекающиеся сообщества.

В данной главе приведены примеры применения метода на реальных данных – графах, полученных при импорте данных из сети *Telegram*-каналов. На рисунке 3 приведен пример применения метода Галактик к графу *Telegram*-каналов, построенному при импорте данных по модели, описанной во второй главе; показано окончательное разбиение на 8 мета-сообществ (пронумерованы на рисунке).

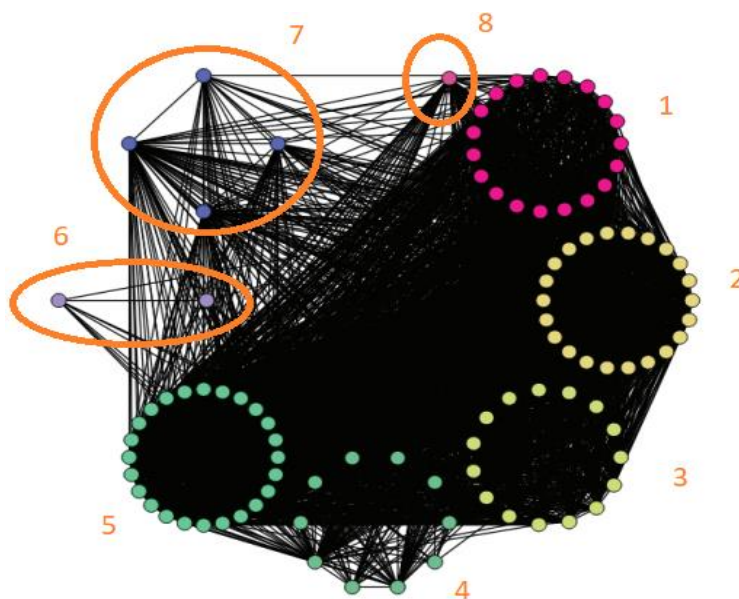


Рис.3. Пример применения метода Галактик: окончательное выделение мета-сообществ.

Каждое из 8 выделенных сообществ состоит из мета-вершин, которые представляют собой наборы вершин, отнесенных к различным исходно выделенным сообществам. Пример визуального представления внутри такой мета-вершины, содержащей в свою очередь 19 вершин, представлен на рисунке 4. В случае принадлежности вершины к нескольким сообществам, она разделена на сектора соответствующих разных цветов.

Для проверки эффективности выделения неявных сообществ «методом Галактик» на графах взаимодействующих объектов были проведены процедуры, описанные в главе 6. Результаты сравнения характеристик текстов выделенных сообществ в сопоставлении с экспертной оценкой их тематического содержания подтверждают эффективность «метода Галактик». Это свидетельствует о корректности выделения неявных сообществ описанным в данной главе методом.

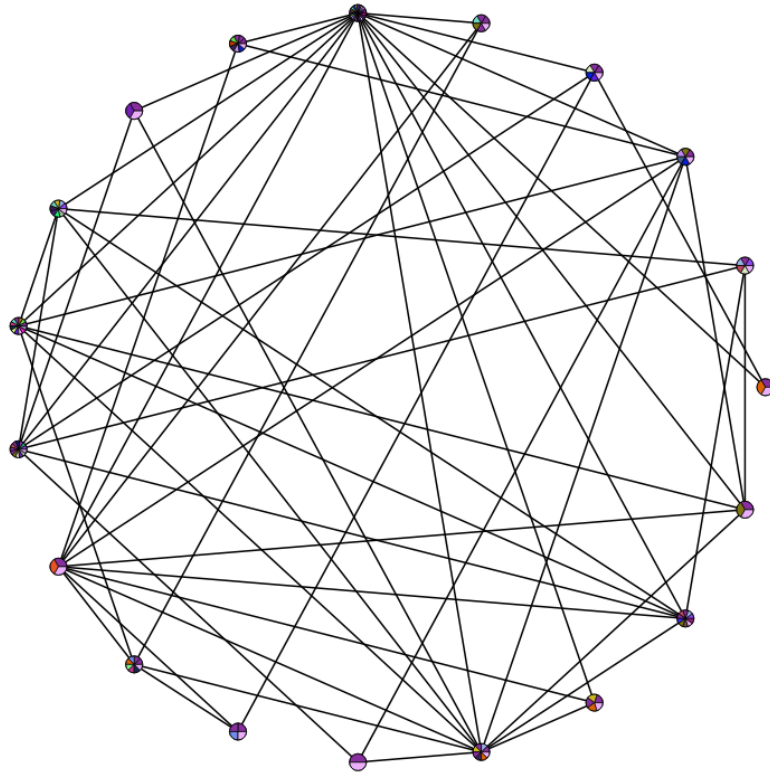


Рис.4. Представление внутри одной из мета-вершин с 19 вершинами

Предложенный метод Галактик позволяет для графов больших размеров посредством выделения мета-сообществ сокращать количество элементов визуализации, что повышает эффективность визуального анализа распространения и хранения информации в сетях взаимодействующих объектов больших объемов.

В **шестой главе** «Методики оценки качества выделения сообществ» представлен подход по оценке выделения сообществ на графе взаимодействующих объектов, имеющем атрибутивные текстовые данные для вершин. Методики основаны на алгоритмах компьютерной лингвистики и выделении психолингвистических факторов.

Результаты анализа объединенных массивов текстов выделенных сообществ позволяют оценить разделение на сообщества следующими методами компьютерной лингвистики:

- Парным сравнением частотных словарей различных лингвистических характеристик, составленных для наборов текстов каждого сообщества;

- Сравнением наборов статистических (психолингвистических) характеристик текстов, описывающих различное поведение участников неявных сообществ.

Тексты выделенных неявных сообществ сопоставляются путем попарного сравнения частотных словарей разных лингвистических характеристик, составленных для каждого из них. Проводится ранговый анализ частотных словарей. Записи словарей рассматриваются как случайные величины. Для каждой пары словарей вычисляются коэффициенты попарной ранговой корреляции, которые являются оценками близости словарей и, как следствие, близости текстов.

Считаем, что каждая пара словарей имеет одинаковый размер записей, который в реальных расчетах ограничивается сверху значением 10000. Это обеспечивается отбрасыванием характеристик с низкими частотами использования (как правило, единичными). Равные по значению элементы получают одинаковое значение усредненного ранга, не зависящее от их перестановки.

Рассматриваются словари текстов сообществ как выборки для двух случайных величин:  $X^n = \{X_i\}_{i=1}^n$ ,  $Y^n = \{Y_i\}_{i=1}^n$ . Для отсортированных по частотам словарей рассматриваются ранги элементов выборок  $rgX^n$  и  $rgY^n$ . Вычисляется коэффициент попарной ранговой корреляции для рассматриваемых выборок

$$r = r(rgX^n, rgY^n) = \frac{cov(rgX^n, rgY^n)}{\sigma_{rgX^n} \cdot \sigma_{rgY^n}} \quad (6.1)$$

где стандартные обозначения  $cov$  и  $\sigma$  – ковариация и дисперсия соответственно.

Пример сопоставления словарей глагольных групп текстов сообществ графа представлен в Таблице 1, где демонстрируется возможность разделения наборов текстов выделенных неявных сообществ. Граф из данного примера получен скачиванием *Telegram*-каналов начиная с наперед заданного канала.

Таблица 1. Коэффициенты корреляции словарей глагольных групп текстов сообществ

	S <sub>0</sub>	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>	S <sub>4</sub>	S <sub>5</sub>	S <sub>6</sub>	S <sub>7</sub>
S <sub>0</sub>	1							
S <sub>1</sub>	-0.694	1						
S <sub>2</sub>	-0.69	0.1	1					
S <sub>3</sub>	0.529	-0.153	-0.258	1				
S <sub>4</sub>	-0.741	-0.292	-0.409	-0.474	1			
S <sub>5</sub>	-0.69	0.71	0.065	-0.168	-0.198	1		
S <sub>6</sub>	-0.561	-0.055	-0.16	-0.213	-0.354	-0.029	1	
S <sub>7</sub>	-0.655	0.197	-0.064	-0.073	-0.122	0.305	-0.102	1

Методами компьютерной лингвистики для наборов текстов каждого из выделенных неявных сообществ вычислялись их статистические характеристики, которые рассматриваются как потенциальные психолингвистические параметры. Исходно были рассмотрены 23 показателя, но по результатам расчетов были выбраны только 5, наиболее явно изменяющихся в зависимости от рассматриваемого сообщества: Коэффициент лексического разнообразия 2 (ЛР2) – отношение числа уникальных псевдооснов к числу словоупотреблений; Коэффициент КД2 – отношение количества глаголов (с причастиями и деепричастными оборотами) к количеству прилагательных; Коэффициент опредмеченности действия (КОД) – соотношение количества глаголов к количеству существительных; Коэффициент логической связности 1 (ЛС1) – отношение общего количества служебных слов (союзов и предлогов) к общему количеству предложений; Отношение числа существительных и глаголов к количеству прилагательных и наречий.

Проведены так же исследования субъектности неявных выделенных сообществ графов взаимодействующих объектов на примере графов, полученных при импорте данных из социальной сети *Twitter*. Субъектность оценивалась экспертами Института психологии РАН. Общая субъектность оказалась значимо связанной с таким показателем как отношение числа вершин к числу ребер в графе: чем больше связей между аккаунтами, входящими в выделенное сообщество

выявлено, тем больше в контенте этих сообществ обнаруживается маркеров, связанных с определенными характеристиками субъектности. Таким образом, показана связь понятия субъектности с выделенными сообществами в графе социальных сетей.

По совокупности результатов сформулированы критерии оценки качества выделения неявных сообществ на графах, которые сводятся к сопоставлению текстов выделенных сообществ посредством рангового анализа словарей текстов и посредством сравнения психолингвистических характеристик текстов сообществ и оценке субъектности выделенных сообществ. Разделение по текстам характеризует корректность разделения на сообщества в условиях информационного взаимодействия в социальных сетях и менеджере *Telegram*.

В седьмой главе «Программное обеспечение анализа графов взаимодействующих объектов» представлены основные компоненты разработанного программного обеспечения для анализа графов взаимодействующих объектов. Реализованное программное обеспечение включает: процедуры импорта данных из социальных сетей и сетей мгновенного обмена сообщениями (*ВКонтакте*, *Twitter* и *Telegram*) и построение графов взаимодействующих объектов данных сетей для последующего анализа; хранилище графов; приложение *AVS (Analytics and Visualization System for graphs)* для анализа и визуализации графов.

В качестве базовой структуры хранилища *AVS-Storage* используется специализированная файловая система. Данная система реализуется внутри некоторого файла стандартной файловой системы компьютера, поэтому под логическим файлом будем понимать файл, принадлежащий данной специализированной системе. Предполагается, что каждый логический файл состоит из некоторых записей переменной длины. Поддерживаются следующие операции: Добавление записи в конец логического файла; Удаление логического файла целиком; Чтение всех записей логического файла с произвольной позиции.

Ключевыми элементами хранилищ данных являются индексы, формирование которых сопровождается сжатием сохраняемых записей о

вхождении индексируемого элемента. Сжатие производится при сохранении очередного буфера на диск. Для буферов обычного инвертированного индекса и индекса по подстрокам применяется один и тот же алгоритм сжатия.

Для более быстрого выполнения операций поиска вершин, по конкретным значениям атрибутов в хранилище присутствует специализированный индекс по атрибутам вершин. В хранилище графов реализованы следующие базовые операции с данными: Операции объединения и пересечения; Поиск путей на графах большого размера.

Для оценки эффективности реализации хранилища графов *AVS-Storage* было проведено сравнительное тестирование характеристик программного обеспечения, предназначенного для хранения и обработки больших графов. Сравнительное тестирование проводилось для разработанного хранилища *AVS-Storage* и следующих баз данных для хранения графов: *OrientDB*, *NEO4J* и *ArangoDB*. Для полноты сравнения эксперименты проводились в том числе и для наиболее популярных систем баз данных, позволяющих хранить графы в табличном представлении, – *Microsoft SQL Server* и *MySQL*, а также и для *MongoDB*.

Тестирование хранилища осуществлялось на графах взаимодействующих объектов, полученных при импорте данных из социальных сетей и взятых из репозитория на ресурсе *Network Repository*. Анализировались временные показатели загрузки тестовых графов в различные средства хранения и выполнения запросов к хранилищам с графами для получения атрибутов вершин, временные показатели и используемые объемы памяти при выполнении запросов по поиску соседей вершин графа. По результатам тестирования сделан вывод о том, что с учетом времени выполнения запросов, с точки зрения скорости, для обработки графов наиболее эффективным является разработанное хранилище *AVS-Storage*.

Программный комплекс *AVS* предусматривает загрузку графов больших размеров, которые могут быть получены из социальных сетей и сетей мгновенного обмена сообщениями. Модуль автоматического размещения графа программного комплекса *AVS* содержит реализованную на языке C++ библиотеку алгоритмов автоматического размещения графа на плоскости. Базовый метод для визуализации

результатов анализа – метод кругового размещения вершин, принадлежащих одному сообществу. Для визуального анализа графа взаимодействующих объектов предусмотрено представление сообщества как вершины, что дает возможность работать с мета-графом. Например, возможно еще раз выделить сообщества в полученном мета-графе.

Программный комплекс *AVS* реализован на языке *C++*. Реализация визуализации выполнена с использованием графической библиотеки *OpenGL*. Графические интерфейсы пользователя созданы на базе приложения *Qt Creator* на основе библиотеки *Qt*.

**В основных результатах работы** приведены основные научные результаты, полученные в диссертации, являющейся научно-квалификационной работой, в которой на основании выполненных исследований представлено решение поставленной в диссертационной работе научной проблемы.

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ, ПОЛУЧЕННЫЕ В ДИССЕРТАЦИОННОЙ РАБОТЕ

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.

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ ДИССЕРТАЦИИ ОПУБЛИКОВАНЫ В РАБОТАХ:

Монография:

1. Чеповский А.А. Анализ графов взаимодействующих объектов. – М.: Национальный открытый университет «ИНТУИТ», 2022. – 270 с.

Коллективная монография:

2. Субъектность и жизнеспособность сетевых сообществ в дискурсивном пространстве Интернета/ Алдашева А.А., Воронин А.Н., Гребенщикова Т.А., Китова Д.А., Ковалева Ю.В., Кубрак Т.А., Латынов В.В., Нестик Т.А., Павлова Н.Д., Рунец О.В., Смирнов И.В., Станкевич М.А., Чеповский А.А. –М.: Изд-во «Институт психологии РАН», 2021 – 373 с.

Публикации в журналах, входящих в перечень ВАК по специальности 1.2.2:

3. Фокина А. И., Чеповский А. А., Чеповский А. М. Использование платформы ТХМ корпусного анализа для анализа текстов сообществ социальных сетей // Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2023. – Т.21. – №2. С. 29–38. DOI: 10.25205/1818-7900-2022-21-2-29-38. [список ВАК (К2) 2023 № 583].
4. Чеповский А. А. О неявных сообществах на графе взаимодействующих объектов // Успехи кибернетики. – 2023. – Т.4. – № 1. – С. 56-64. DOI: 10.51790/2712-9942-2023-4-1-08. [список ВАК (К3) 2023 № 2719].
5. Аванесян Н.Л., Зенькова В.В., Чеповский А.А., Чеповский А.М. Анализ текстов сообществ социальных сетей // Успехи кибернетики. – 2023. – Т.4. – № 2. – С. 33-39. DOI: 10.51790/2712-9942-2023-4-2-05. [список ВАК (К3) 2022 № 2664].
6. Попов В. А., Чеповский А. А. Модели импорта данных из мессенджера Telegram // Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2022. – Т.20. – №2. С. 60–71. DOI: 10.25205/1818-7900-2022-20-2-60-71. [список ВАК (К1) 2022 № 518].
7. Попов В. А., Чеповский А. А. Выделение неявных сообществ на графе взаимодействия Telegram-каналов с помощью «метода Галактик» // Труды ИСА

- РАН. – 2022. – Т.72. – №4. С. 39–50. DOI: 10.14357/20790279220405. [список ВАК (К1) 2022 № 2348].
8. Попов В. А., Чеповский А. А. Модели импорта данных из Твиттера // Вестник НГУ. Серия: Информационные технологии. – 2021. – Т. 19. – № 2. – С. 76–91. DOI 10.25205/1818-7900-2021-19-2-76-91. [05.13.18 список ВАК (К1) 2021 № 486].
9. Аванесян Н. Л., Соловьев Ф. Н., Чеповский А. А. Характеристики текстов сообществ социальных сетей. Вестник НГУ. Серия: Информационные технологии. – 2021. – Т.19. – №1. – С. 5–14. DOI: 10.25205/1818-7900-2021-19-1-5-14 [05.13.18 список ВАК (К1) 2021 № 486].
10. Чеповский А. А., Орлов А. О. О свойствах модулярности и актуальных корректировках алгоритма Блонделя // Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2017. – Т. 15. – № 3. – С. 64-73. [05.13.00 список ВАК (К1) 2017 № 1834].
11. Поляков И. В., Соколова Т. В., Чеповский А. А., Чеповский А. М. Проблема классификации текстов и дифференцирующие признаки // Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2015. – Т. 13. – № 2. – С. 55-63. [список ВАК (К1) 2015 (до 30.06) № 365]
12. Поляков И. В., Чеповский А. А., Чеповский А. М. Хранение и обработка графа социальных сетей // Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2013. – Т. 11. – № 4. – С. 77-83. 14 [список ВАК (К1) 2013 № 365].
13. Соколова Т., Чеповский А. А. Анализ профилей сообществ социальных сетей // Системы высокой доступности. – 2018. – Т. 14. – № 3. – С. 82-86. [05.13.00 список ВАК (К2) 2018 № 1852].
14. Поляков И. В., Полякова В. И., Чеповский А. А. Особенности хранения графов социальной сети // Системы высокой доступности. – 2018. – Т. 14. – № 3. – С. 63-67. [05.13.00 список ВАК (К2) 2018 № 1852].

15. Чеповский А. А., Лобанова С. Ю. Комбинированный алгоритм выделения сообществ в графах взаимодействующих объектов // Бизнес-информатика. – 2017. – Т. 42. – № 4. – С. 64-73 (Cherovskiy A., Lobanova S. Combined method to detect communities in graphs of interacting objects / Пер. с рус. // Business Informatics. – 2017. – Vol. 42. – No. 4. – P. 64-73.) [05.13.00 список ВАК (К1) 2017 и 2018 № 90].
16. Михайлов А.С., Соколова Т.В., Чеповский А.А., Чеповский А.М. Выявление тематической направленности текстов на естественных языках // Искусственный интеллект и принятие решений. – 2016. – № 1. – С. 9-17. [05.13.00 список ВАК (К1) 2016 № 718].
- Публикации в научных изданиях, индексируемых в базах данных изданий, которые приравниваются к публикациям, в которых излагаются основные научные результаты диссертации:
17. Попов В. А., Чеповский А. А. О моделях построения графа взаимодействующих объектов в сети Telegram-каналов // Вопросы кибербезопасности. – 2024. – № 3 (61). – С. 105-112. DOI: 10.21681/2311-3456-2024-3-105-112. [RSCI].
18. Чеповский А. А. Об особенностях построения и анализа графов взаимодействующих объектов в сети Telegram-каналов // Вопросы кибербезопасности. – 2023. – № 1 (53). – С. 75-81. DOI:10.21681/2311-3456-2023-1-75-81 [RSCI].
19. Воронин А. Н., Ковалева Ю. В., Чеповский А. А. Взаимосвязь сетевых характеристик и субъектности сетевых сообществ в социальной сети Твиттер // Вопросы кибербезопасности. – 2020. – Т. 37. – № 3. – С. 40-57. [RSCI].
20. Лещёв Д. А., Сучков Д. В., Хайкова С. П., Чеповский А. А. Алгоритмы выделения групп общения // Вопросы кибербезопасности. – 2019. – Т. 32. – № 4. – С. 61-71. [RSCI].
21. Polyakov I. V., Cherovskiy A., Cherovskiy A. Data Compression in Big Graph Warehouse / Пер. с рус. // Journal of Mathematical Sciences. – 2020. – Vol. 245. – P. 197-201. DOI:10.1007/s10958-020-04686-4 (Поляков И. В., Чеповский А. А.,

- Чеповский А. М. Сжатие данных в хранилище больших графов // *Фундаментальная и прикладная математика*. – 2016. – Т. 21. – № 4. – С. 125-132) [Scopus, Q3].
22. Chervovskiy A., Chervovskiy A., Polyakov I. V., Kolomeychenko M. I. Detection of Communities in a Graph of Interactive Objects // *Journal of Mathematical Sciences*. – 2019. – Vol. 237. – No. 3. – P. 426-431 (Коломейченко М. И., Поляков И. В., Чеповский А. А., Чеповский А. М. Выделение сообществ в графе взаимодействующих объектов // *Фундаментальная и прикладная математика*. – 2016. – Т. 21. – № 3. – С. 131-139). [Scopus, Q3].
23. Kolomeychenko M. I., Chervovskiy A.A., Chervovskiy A.M. An Algorithm for Detecting Communities in Social Networks // *Journal of Mathematical Sciences*. – 2015. – Vol. 211. – No. 3. – P. 310-318 (Коломейченко М. И., Чеповский А. А., Чеповский А. М. Алгоритм выделения сообществ в социальных сетях // *Фундаментальная и прикладная математика*. – 2014. – Т. 19. – № 1. – С. 21-32) [Scopus, Q3].
24. Polyakov I.V., Chervovskiy A.A., Chervovskiy A.M. Algorithms for Searching Paths in Huge Graphs // *Journal of Mathematical Sciences*. – 2015. – Vol. 211. – No. 3. – P. 413-417. (Поляков И. В., Чеповский А. А., Чеповский А. М. Алгоритмы поиска путей на графах большого размера // *Фундаментальная и прикладная математика*. – 2014. – Т. 19. – № 1. – С. 165-172). [Scopus, Q3].
- Другие публикации в сборниках и трудах конференций:*
25. Чеповский А. А. Построение и анализ графов взаимодействующих объектов / В кн.: Международная конференция «Математика в созвездии наук». К юбилею ректора МГУ академика Виктора Антоновича Садовниченко: Тезисы докладов / Орг. комитет: В. А. Садовнический, А. И. Шафаревич, И. А. Соколов [и др.]. — Москва: Издательство Московского университета, 2024. — С. 355 - 357.
26. Chervovskiy A., Leshchev D., Khaykova S.P. Core Method for Community Detection, in: *Complex Networks & Their Applications IX. Volume 1: Proceedings of the Ninth International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2020*. – Springer, 2021. – P. 38-50. [Scopus].

27. Chervovskiy A. Methods to reveal communities without the property of "picking up junk" // In The 6 th International Conference on Complex Networks & Their Applications. Nov. 29 - Dec. 01, 2017. – Lyon (France). – P. 336-340. [Scopus].
28. Лобанова С. Ю., Чеповский А. А. О применении алгоритмов разбиения сети взаимодействующих объектов на сообщества / В кн.: Управление информационной безопасностью в современном обществе. Сборник научных трудов. М.: Изд. дом Высшей школы экономики, 2017. – С. 83-84.
29. Орлов А. О., Чеповский А. А. Особенности алгоритмов выделения сообществ в графах социальных сетей / В кн.: Управление информационной безопасностью в современном обществе. Сборник научных трудов. М.: Изд. дом Высшей школы экономики, 2017. – С. 108-109.
30. Коломейченко М. И., Поляков И. В., Чеповский А. А. Автоматическое размещение графа на основе метода физических аналогий / В кн.: Труды Международной научной конференции Московского физико-технического института (государственного университета) и Института физико-технической информатики (SCVRT1516). М., Протвино: Институт физико-технической информатики, 2016. – С. 93-97.
31. Поляков И. В., Чеповский А. А. Буферизация и сжатие данных при хранении мультиграфа / В кн.: Труды Международной научной конференции Московского физико-технического института (государственного университета) и Института физико-технической информатики (SCVRT1516). – М., Протвино: Институт физико-технической информатики, 2016. – С. 76-78.
32. Коломейченко М. И., Поляков И. В., Чеповский А. А., Чеповский А. М. О хранении графа социальной сети / В кн.: Труды Международной научной конференции по физико-технической информатике (СРТ2015). – М., Протвино: Институт физико-технической информатики, 2016. – С. 175-178.
33. Орлов А. О., Чеповский А. А. Особенности алгоритма Блонделя при выявлении сообществ в графе социальной сети / В кн.: Труды Международной научной конференции Московского физико-технического института (государственного

- университета) и Института физико-технической информатики (SCVRT1516). – М., Протвино: Институт физико-технической информатики, 2016. – С. 124-129.
34. Михайлов А. С., Соколова Т. В., Чеповский А. А., Чеповский А. М. Методика выявления нарушений в текстах Интернета / В кн.: Труды Международной научной конференции по физико-технической информатике (СРТ2014). – М., Протвино: Институт физико-технической информатики, 2015. – С. 115-119.
35. Золотых А. А., Коломейченко М. И., Чеповский А. А. О задаче анализа графа социальной сети / В кн.: Труды Международной научной конференции по физико-технической информатике (СРТ2014). – М., Протвино: Институт физико-технической информатики, 2015. – С. 131-134.
36. Коломейченко М. И., Поляков И. В., Чеповский А. А. Хранение и скачивание сетей больших размеров / В кн.: Труды Международной научной конференции Resilience2014 Международного Центра по ядерной безопасности Института физико-технической информатики. – М., Протвино: Институт физико-технической информатики, 2015. – С. 139-143.
37. Михайлов А. С., Чеповский А. А., Чеповский А. М. О моделях оценки информационного воздействия в социальных сетях. / В кн.: SCVRT2013-14 Труды Международной научной конференции Международного центра по ядерной безопасности Института физико-технической информатики. – Протвино: Изд-во ИФТИ, 2014. – С. 247-249.

Учебные пособия

38. Поляков И. В., Соловьев Ф. Н., Чеповский А. А., Чеповский А. М. Задача распознавания для текстов на естественных языках. – М.: Национальный открытый университет «ИНТУИТ», 2017. – 119 с.
39. Коломейченко М. И., Поляков И. В., Чеповский А. А., Чеповский А. М. Методы визуального анализа графов. «ИНТУИТ», 2016. – 276 с.

Список РИД - свидетельств о регистрации программы для ЭВМ:

40. Программа для импорта и построения графов на основе данных из сети мгновенного обмена сообщениями. Свидетельство о государственной

регистрации программы ЭВМ № 2024680205. Попов В.А., Чеповский А.А. Дата регистрации 27.08.2024.

41. Определение характеристик корпусов текстов и сравнения корпусов корреляционным анализом. Свидетельство о государственной регистрации программы ЭВМ № 2024680546. Аванесян Н.Л., Чеповский А.А., Чеповский А.М. Дата регистрации 30.08.2024.