

Санкт-Петербургский государственный университет

на правах рукописи

Чи Чжао

**Моделирование динамики бинарных
мнений в социальных сетях сложных
конфигураций**

Научная специальность 2.3.1.

Системный анализ, управление и обработка информации, статистика

**Диссертация на соискание ученой степени
кандидата технических наук**

Перевод с английского языка

Научный руководитель:
доктор физико-математических наук,
Проф. Е. М. Парилина

Санкт-Петербург
2024

Содержание

| | |
|---|-----------|
| Введение | 5 |
| Глава 1. Динамика мнений в двухслойных сетях в присутствии лицемерия | 24 |
| 1.1 Базовая и скрытая модели избирателя | 24 |
| 1.1.1 Базовая модель избирателя (BVM) | 25 |
| 1.1.2 Скрытая модель избирателя (CVM) | 27 |
| 1.2 Общая скрытая модель избирателя (GCVM) | 31 |
| 1.2.1 Обоснование модели | 31 |
| 1.2.2 Симметричный случай: полная внутренняя сеть | 32 |
| 1.2.3 Асимметричные случаи: неполная внутренняя сеть | 34 |
| 1.2.4 Асимметричный случай: Неполная внешняя сеть | 40 |
| 1.3 Эксперименты и результаты | 44 |
| 1.4 Заключение к Главе 1 | 48 |
| Глава 2. Анализ времени достижения консенсуса и вероятности победы мнения в двухслойных сетях различных структур при наличии лицемерия | 50 |
| 2.1 Многослойная сеть с узлами-репликами | 50 |
| 2.2 Двухслойная сеть с узлами-репликами | 51 |
| 2.3 Модель | 52 |
| 2.3.1 Общая скрытая модель избирателя (макроверсия) | 52 |
| 2.3.2 Общая скрытая модель избирателя (микроверсия) | 52 |
| 2.4 Эксперименты и результаты | 56 |
| 2.4.1 Общее описание эксперимента | 56 |
| 2.4.2 Основные результаты и наблюдения | 58 |
| 2.5 Заключение ко второй главе | 76 |
| Глава 3. Меры центральности и динамика мнений в двухслой- ных сетях с узлами-репликами | 78 |
| 3.1 Метод построения однослойной взвешенной сети по двухслойной сети | 78 |
| 3.2 Сеть клуба каратэ Закари в двухслойной сетевой постановке | 80 |

| | | |
|--|---|------------|
| 3.3 | Меры центральности в однослойных и двухслойных сетях | 82 |
| 3.3.1 | Классические меры центральности | 83 |
| 3.3.2 | Меры центральности при случайном блуждании | 85 |
| 3.3.3 | Теоретико-игровые меры центральности | 87 |
| 3.4 | Эксперименты | 92 |
| 3.4.1 | Центральности, основанные на векторе Шепли и значении Майерсона | 94 |
| 3.4.2 | Эксперименты по изучению корреляции характеристик се- ти и динамики мнений | 102 |
| 3.5 | Заключение к главе 3 | 107 |
| | | |
| Глава 4. SharG: новый метод оценки значимости признаков на основе вектора Шепли | | 109 |
| 4.1 | Методы объяснимого искусственного интеллекта (ХАИ) | 109 |
| 4.2 | SharG: новый метод ХАИ | 111 |
| 4.2.1 | Вектор Шепли | 111 |
| 4.2.2 | SharG (модель на основе вектора Шепли для графов) | 112 |
| 4.3 | Эксперименты | 118 |
| 4.3.1 | Описание наборов данных | 118 |
| 4.3.2 | Предварительная обработка данных для SharG | 119 |
| 4.3.3 | Модели предсказания ИИ | 120 |
| 4.3.4 | Оценка методов ХАИ | 124 |
| 4.4 | Результаты и их анализ | 125 |
| 4.4.1 | Значимость признаков, вычисленная SharG | 125 |
| 4.4.2 | Оценка методов ХАИ | 128 |
| 4.4.3 | Объяснение сложных моделей | 135 |
| 4.5 | Заключение к главе 4 | 138 |
| | | |
| Заключение | | 139 |
| | | |
| Список сокращений и условных обозначений | | 142 |
| | | |
| Список иллюстраций | | 144 |
| | | |
| Список таблиц | | 146 |
| | | |
| Список литературы | | 147 |

Приложение 158

Введение

Актуальность темы диссертации

Динамика мнений является важной областью исследования, посвящённой формированию убеждений, взглядов и точек зрения, их распространению и изменению в рамках социальных сетей. Социальные сети, состоящие из индивидов или объектов, связанных различными отношениями, в настоящее время могут считаться основными каналами для обмена информацией и распространения мнений. Поскольку мир становится всё более взаимосвязанным, понимание процессов формирования, распространения и изменения мнений в сложных сетях становится ключевым для решения множества актуальных проблем, таких как борьба с дезинформацией, разработка эффективной общественной политики и др. Изучение динамики мнений объединяет знания из социологии, психологии, физики и компьютерных наук для моделирования и анализа взаимодействий между индивидами, приводящих к коллективному поведению и социальным феноменам.

Социальные сети по своей природе являются сложными системами, часто состоящими из нескольких взаимосвязанных слоёв, которые представляют собой различные типы отношений или каналов коммуникации между индивидами. В современную цифровую эпоху, когда онлайн-платформы и социальные медиа радикально изменили «природу» человеческого взаимодействия и распространения информации, они также добавили новые сложности в социальные сети, такие как возможность быстрого распространения информации, формирование эхо-камер и разнообразие структуры сети. Кроме того, наличие индивидов, которые могут публично выражать мнения, отличные от своих персональных, добавляет ещё один уровень сложности в эти системы. Это явление характерно для множества реальных сценариев, например, в рабочей среде, где сотрудники могут скрывать свои истинные мнения, чтобы соответствовать корпоративной культуре, или в политике, где индивиды могут изменять свою публичную позицию, чтобы соответствовать большинству. Понимание влияния такого поведения на общую динамику мнений важно для разработки более точных моделей социального влияния и создания стратегий, направленных наощрение подлинного дискурса.

Когда речь идёт о распространении мнений, часто имеют в виду процесс достижения консенсуса или общего согласия по конкретному вопросу. Интересно исследовать динамику формирования и распространения мнений в социальных сетях со сложными конфигурациями, где индивиды могут проявлять упрямство, лицемерие, изменчивость и т.д., а структура сети может быть многослойной, циклической, звездообразной или иной. Различные структуры сети характеризуются различными числовыми характеристиками и неоднородностью индивидов в сети. Например, структура сети может существенно изменить центральность индивидов, среднюю длину кратчайшего пути и траектории распространения информации, что влияет на процесс распространения мнений. Поведение индивидов также может влиять на этот процесс, например, упрямые индивиды никогда не меняют своё мнение, лицемерные могут иметь разные частные и публичные мнения, а изменчивые могут часто менять свои взгляды. Наличие таких индивидов создаёт новые вызовы для процесса формирования консенсуса и распространения мнений. Анализ времени достижения консенсуса и вероятности выигрыша определенного мнения для различных сетевых структур и начальных мнений индивидов позволяет лучше понять влияние топологии сети и поведения индивидов на скорость и исход процесса распространения мнений. Такие знания могут быть применены для разработки более эффективных стратегий коммуникации и понимания устойчивости различных социальных структур к изменению мнений.

Диссертация посвящена изучению динамики мнений в социальных сетях со сложными конфигурациями, в частности, с точки зрения структуры сети и поведения индивидов. Ключевым аспектом данного исследования является изучение двухслойных моделей сетей, которые различают публичные (внешние) и частные (внутренние) уровни коммуникации. Такой подход позволяет более тонко и правдоподобно описывать взаимодействия индивидов в реальных социальных системах. В то же время, в рамках данной работы исследуются меры центральности в сложных сетевых структурах, что даёт ценные знания для выявления влиятельных индивидов в социальной системе. Это имеет практическое применение в таких областях, как таргетированный маркетинг, политические кампании или общественные программы здравоохранения, где выявление ключевых индивидов или лидеров мнений является важным для эффективного распространения информации. Кроме того, в этой диссертации демонстрируется применение разработанных мер центральности для создания нового метода объясняемого искусственного интеллекта. Использование разработанных мер

центральности для оценки важности признаков в задачах регрессии и классификации полезно для исследователей в области искусственного интеллекта для нахождения оптимальной стратегии обучения моделей. С социальной точки зрения, оптимальная стратегия обучения может сократить потребление энергии, что приведёт к снижению выбросов углекислого газа и, таким образом, внесёт вклад в достижение углеродной нейтральности.

Обзор результатов в данной области

Новаторская работа Винера в области кибернетики [1] заложила основы для социокibernетики [2]. Социокibernетика фокусируется на внутренних законах самоорганизации и самоадаптации социальных систем, исследуя то, как социальные механизмы и структуры системы, представляющие общество, могут спонтанно выполнять специфические координационные и контрольные действия. Сочетание социологии, системного анализа и теории управления сместило акцент исследований социальных сетей с анализа на изучение эволюции восприятия, поведения и социальных взаимоотношений в социальных сетях с точки зрения динамических систем, что дало начало новой области исследований — динамике мнений.

В XXI веке развитие многоагентных систем и сложных сетей предоставило исследователям широкий спектр математических моделей и инструментов для количественного анализа или численного моделирования крупных социальных сетей.

Модели динамики мнений можно разделить на две основные группы: макроскопические и микроскопические. Макроскопические модели исследуют социальные сети с использованием статистико-физических методов, применяя теорию вероятностей и статистику для анализа того, как изменяется распределение мнений, например, модель Изинга [3] и модель избирателя [4]. Модель Изинга имеет долгую историю в статистической физике [5]. Модель Шнайда [6] является одной из известных модификаций модели Изинга. В каждом раунде модели Шнайда пара агентов a_i и a_{i+1} выбирается для воздействия на ближайших соседей, то есть агентов a_{i-1} и a_{i+2} . В модели избирателя [4] случайным образом выбирается агент a_i , а затем выбирается его сосед, и этот сосед принимает мнение агента a_i .

Микроскопические модели напрямую описывают, как изменяются мнения индивидов с социальной точки зрения, например, см. модель Де Гроота [7], модель Фридкина-Джонсона (F-J) [8] и модели ограниченного доверия [9, 10].

В модели Де Гроота каждый индивид обновляет своё мнение на основе своего собственного мнения и мнений соседей. Модель F-J является одной из основных обобщений модели Де Гроота, в которой присутствуют упрямые агенты, которые учитывают свои первоначальные убеждения на каждом этапе обновления мнений [11]. Возможность управления мнениями агентов посторонними участниками сети рассматривается в работах [12, 13]. Модифицированная модель F-J с пассивными и активными агентами представлена в [14]. Модель ограниченного доверия (VCM) — это модель, в которой агенты игнорируют мнения, которые слишком сильно отличаются от их собственных [5]. VCM включает две основные модели: модель Деффуана-Вейбуша (D-W), предложенную в работе [10], и модель Хегсельмана-Краузе (H-K), представленную в работе [9]. В модели D-W случайным образом выбираются два индивида a_i, a_j , и они решают, взаимодействовать ли, в зависимости от уровня доверия [15]. Модель H-K также является расширением модели Де Гроота, в которой предполагается, что у каждого индивида в сети есть порог доверия, и его мнение может быть изменено только другими индивидами, которые находятся в пределах этого порога. Таким образом, правило обновления мнений в модели H-K зависит от текущего состояния. Модель Краснощёкова (K), описывающая динамику мнений в обществе, представленном одним слоем, была введена в [16] и затем исследована в [17]. Эта динамика может быть сведена к динамике модели F-J. Фактически, если применить правило обновления мнений, представленное в модели K, соответствующая динамика приведёт к сближению мнений агентов до определённых конечных мнений, и, следовательно, достигается консенсус. Обзор [18] рассматривает различные модели в области динамики мнений с ограниченным доверием, выделяя ключевые механизмы, приводящие к достижению консенсуса, поляризации и фрагментации в группах.

Наличие группы агентов, способных манипулировать мнениями общества, исследуется в [19]. В этой модели предполагается, что агенты гетерогенны, причём группа так называемых лидеров знает начальные мнения всех агентов, в то время как остальные агенты не владеют этой информацией.

Согласно [20], модели динамики мнений обычно состоят из нескольких основных элементов: (i) форматы выражения мнений, определяющие, как математически представлять мнение, (ii) правило слияния, определяющее, как ин-

дивиды взаимодействуют друг с другом, (iii) среда динамики мнений, то есть структура такой социальной сети.

В социальной сети индивиды не полностью принимают и не полностью игнорируют мнения других. В некоторой степени они учитывают эти мнения при формировании новых мнений в процессе, определённом правилом слияния. В ходе группового взаимодействия индивиды непрерывно обновляют и интегрируют свои мнения по одному и тому же вопросу. В конечном итоге возможны три варианта стабильных результатов: консенсус, поляризация, фрагментация и один нестабильный результат — колебания [9].

Базовая модель избирателя, предложенная Ричардом и Томасом, называется *BVM* [4], а скрытая модель голосующего, предложенная Гастнером и др., называется *CVM* [21,22]. Обе модели относятся к макроскопическим. Структура *BVM* основана на полной сети, и основное предположение этой модели заключается в том, что индивиды всегда публично выражают своё мнение. Поэтому правило слияния в *BVM* довольно интуитивно: случайным образом выбирается индивид и его сосед, и индивид принимает мнение соседа. *CVM* предполагает, что социальная сеть разделена на внешний и внутренний слой, и индивиды могут свободно скрывать или публично выражать свои мнения. Внешний слой *CVM* представляет собой полную сеть, и каждая вершина во внешнем слое связана с вершиной во внутреннем слое. Кроме того, во внутреннем слое отсутствуют связи между вершинами, поэтому взаимодействие во внутреннем слое не допускается.

В данной диссертации предполагается, что индивиды могут взаимодействовать во внутреннем слое, что отличает нашу модель от *CVM*. Мы предлагаем эту идею, основываясь на том, что индивиды всегда делятся своими реальными мнениями с близкими друзьями. Многоуровневый подход к моделированию структуры общества, включая командные сети в компаниях, широко используется в моделях промышленной организации [23]. Поэтому мы также включаем этот многоуровневый подход к коммуникации в модель динамики мнений и называем *CVM* с возможными взаимодействиями во внутреннем слое — *GCVM* [24–27]. Главу 1 мы начинаем с сетевой структуры и используем статистико-физические методы и теорию вероятностей для формулирования и моделирования процесса динамики мнений (т.е. в моделировании не создаётся реальная сеть, а модель основана на формулах), и предлагаемую модель мы называем макроскопической *GCVM* [24]. В главе 2 мы создаём соответствующие сети для заданных внутренних и внешних структур вместо использования

только статистико-физических методов, и такая модель называется микроскопической GCVM [25–27].

Поскольку структура сети в GCVM является двухслойной, интересно исследовать, как эта структура в целом, и такие характеристики сети как, например, различные меры центральности [28], влияют на динамику мнений и достижимость консенсуса. Мы рассматриваем два ключевых показателя эффективности динамики мнений, а именно коэффициент выигрыша и время достижения консенсуса.

Социальная власть (влияние центральности) — это концепция, которая ранжирует важность вершин в сети. Меры центральности используются для выявления наиболее влиятельных вершин в сети. Знание влиятельных вершин очень важно для динамики мнений, что может помочь понять, какие вершины играют ключевую роль в распространении мнений. Наиболее распространённые меры центральности — степень посредничества [29], степень близости [30–32] и степень связности [33]. Существуют также некоторые меры центральности, основанные на случайных блужданиях в сети, такие как занятость при случайном блуждании [34], которая является частотой посещения вершины в сети во время случайного блуждания, и нагрузка узла при случайном блуждании [35], которая представляет собой долю путей через вершину среди всех путей во время случайного блуждания. Нагрузка узла при случайном блуждании не зависит от кратчайшего пути, поэтому она более общая, чем степень посредничества. Также стоит упомянуть близость при случайном блуждании, которая является вариантом степени близости [34], и вычисление этой центральности основано на времени первого прохождения (MFPT). Аналитические выражения центральностей, основанных на случайных блужданиях, можно найти в работе [34]. Теоретико-игровые меры центральности — гибкий и сложный подход к выявлению наиболее влиятельных вершин в сети, идея которого заимствована из кооперативной теории игр. Тарковски написал хороший обзор теоретико-игровых мер центральностей в работе [36]. Вектор Шепли [37] и значение Майерсона [38] — концепции решений в кооперативной теории игр, которые используются для справедливого распределения общего выигрыша среди игроков на основе их предельного вклада. В статье [39] авторы представляют, как использовать вектор Шепли для определения топ- k вершин в социальной сети. Мазалов и др. предложили теоретико-игровую меру центральности для взвешенного графа на основе значения Майерсона в работе [40]. В [41] Мазалов и Хитрая предложили модифицированное значение Майерсона для невзвешенных неори-

ентрированных графов. Характеристическая функция, использованная для этой модификации, учитывает не только простые пути, но и циклы. В следующей работе [42] Мазалов и Хитрая вводят концепцию интегральной центральности для невзвешенных ориентированных графов и приводят точное математическое доказательство того, что эта мера центральности удовлетворяет аксиомам Болди-Виньи [43].

В главе 3 мы исследуем связь между мерами центральности и динамикой мнений на основе сети клуба каратэ Закари. Здесь предложены два быстрых и точных алгоритма для приближённого вычисления теоретико-игровых мер центральности, которые проверяются на случайно сгенерированной сети и сети клуба каратэ Закари.

Если рассматривать признаки в модели машинного обучения как вершины в сети, а важность признаков как силу вершин, то меры центральности могут быть использованы для объяснения моделей машинного обучения. Важность признаков — это ключевое понятие в машинном обучении, которое помогает понять модель и принять правильное решение о том, какие признаки выбрать для создания точной модели. Арриета и др. рассмотрели концепции, связанные с объяснимым ИИ, и проанализировали объяснения, предоставляемые ХАИ, в основном, делящихся на два типа: глобальные и локальные [44]. Методы локальных объяснений в ХАИ ориентированы на предоставление объяснений для отдельных выборок или предсказаний [45], тогда как методы глобальных объяснений предоставляют объяснения для всей модели [46]. Эти методы крайне полезны для того, чтобы помочь пользователям понять процесс принятия решений и важность признаков модели, повышая доверие к модели и обеспечивая поддержку принятия решений на основе объяснений. В локальных объяснениях часто используются два метода: LIME (Локальные Интерпретируемые Модель-Агностические Объяснения) [47] и SHAP (Объяснения на основе вектора Шепли) [48]. В глобальных объяснениях FI и SHAP также могут использоваться для объяснений значимости переменных в модели [49]. В главе 4 предложен новый метод глобальных объяснений и проведено его сравнение с современными методами, такими как SHAP и Feature Importance.

Подводя итог, скажем, что важно и значимо исследовать динамику мнений в социальных сетях со сложными конфигурациями, в частности, различные структуры сетей и поведение индивидов, меры центральности и их связь с динамикой мнений, а также практическое применение разработанных мер центральности. Результаты этого исследования могут быть применены в таких

областях, как социология, психология, физика, компьютерные науки и искусственный интеллект.

Цели диссертации

Цель диссертации — изучение динамики мнений в социальных сетях со сложными конфигурациями с использованием различных подходов и их анализ с различных точек зрения, когда динамика мнений моделируется статистико-физическими (макроскопическими) методами или на основе реальной сети (микроскопическими методами). В моделировании также предполагается наличие лицемерных индивидов, а также учитываются различные характеристики сети, в частности, с точки зрения мер центральности для изучения процессов распространения и практического применения разработанных методов. Для достижения основной цели в четырёх главах диссертации рассматриваются и решаются несколько конкретных вопросов.

Диссертация представляет собой серию результатов, включающих моделирование, анализ динамики мнений в социальной сети со сложными конфигурациями, а также анализ свойств сети с их применениями.

Основные задачи

Для достижения целей диссертации были определены следующие ключевые задачи:

1. Моделирование динамики мнений в двухслойных сетях с лицемерными индивидами с использованием статистико-физических методов и теории вероятностей, предполагая, что взаимодействие во внутреннем слое допускается, что является более реалистичным, чем скрытая модель голосящего. Сравнение процессов распространения мнений при конкретных структурах внутренних сетей (т.е. пустая сеть, полная сеть, циклическая сеть, звездообразная сеть, сети - двойная звезда и двойная клика) и начальных мнениях индивидов. Ключевые показатели — время достижения консенсуса (время консенсуса) и вероятность выигрыша конкретного мнения (например, красного).
2. Использование микроскопического подхода для моделирования динамики мнений при тех же предположениях, что и в задаче 1. Предложение модели для двухслойной сети, которая должна быть достаточно общей, при-

менимой для любой заданной структуры двухслойной сети, соответствующей определению. Провести моделирование процесса динамики мнений в различных структурах сетей и поведеньях индивидов. Проанализировать, как свойства сети и поведение индивидов влияют на ключевые показатели (время консенсуса и коэффициент выигрыша).

3. Генерация более реалистичной двухслойной сети в соответствии с определением двухслойной сети и реального набора данных (например, сеть клуба каратэ Закари). Исследование микроскопической модели на сгенерированной сети. Меры центральности используются для выявления самых влиятельных вершин в сети, однако о мерах центральности в двухслойной сети известно мало. Поэтому необходимо выяснить, как упростить двухслойную сеть до однослойной и применить меры центральности к упрощённой сети. Кроме того, следует разработать и проверить новые меры центральности. Исследовать связь между мерами центральности и ключевыми показателями динамики мнений.
4. Поиск практических применений разработанных мер центральности. Важность признаков при построении моделей – это ключевое понятие в машинном обучении, которое помогает понять модель и принять правильное решение, какие признаки использовать в модели. Разработанные меры центральности могут быть использованы для определения важности признаков в моделях машинного обучения, если рассматривать признаки как вершины сети. Разработать новый алгоритм определения важности признаков на основе предложенных мер центральности и сравнить его с современными методами (SHAP, LIME и Feature Importance).

Научная новизна

В данной диссертации предложена общая модель скрытого избирателя (GCVM), в которой индивиды взаимодействуют на двух уровнях и могут обмениваться мнениями во внутреннем слое сети, которая их связывает. Такое взаимодействие не предусмотрено в модели скрытого избирателя (CVM). Под обменом мнениями во внутреннем слое подразумевается, что индивиды делятся своими реальными или внутренними мнениями с близкими друзьями. Существуют две версии GCVM: макроскопическая и микроскопическая. Макроскопическая GCVM основана на статистико-физических методах и теории веро-

ятностей, тогда как микроскопическая GCVM опирается на моделирование на основе реальной сетевой структуры.

Для макроскопической GCVM в диссертации представлен процесс формирования мнений в различных внутренних структурах GCVM. Мы проводим серию численных симуляций макроскопической GCVM с различными сетевыми структурами (как внешними, так и внутренними) и получаем несколько интересных выводов. Например, выясняется, что иногда при относительно простой структуре внешнего слоя консенсус среди мнений индивидов не достигается, а если индивиды в сети не склонны выражать свои мнения публично (во внешнем слое), обмен мнениями с близкими друзьями (во внутреннем слое) почти бесполезен.

Для микроскопической GCVM мы даём определение двухслойной сети для динамики мнений. Макроскопическая GCVM использует только статистико-физические методы, тогда как микроскопическая GCVM начинается с моделирования реальной сети. Мы проводим серию симуляций с различными сетевыми структурами и поведением индивидов. Более того, мы выдвигаем и подтверждаем гипотезу о существовании сильной линейной зависимости между временем консенсуса и средними попарными кратчайшими путями d в структуре сети. Мы применяем метод контролируемой переменной для оценки влияния каждого индивидуального параметра на ключевые показатели эффективности (KPI), включая время достижения консенсуса и вероятность выигрыша мнения. Кроме того, оценивается влияние комбинаций параметров (некоторых специфических поведений индивидов) на KPI, анализируя результаты с помощью алгоритма K-means. Заключаем, что определённые комбинации параметров могут существенно влиять на время достижения консенсуса.

Мы исследуем меры центральности, определённые на двухслойных сетях. Предложен подход к упрощению двухслойной сети, связанной с динамикой мнений, до однослойной взвешенной сети. Предложены два быстрых и точных алгоритма для однослойной взвешенной сети для приближённого вычисления теоретико-игровых мер центральности и исследована связь между мерами центральности и характеристиками процессов динамики мнений в таких сетях. В качестве примера рассматривается социальная сеть клуба каратэ Закари, расширенная за счёт добавления второго (внутреннего) слоя коммуникации. Структуры внешнего и внутреннего слоёв могут отличаться. Под ключевыми характеристиками процесса динамики мнений понимаются время кон-

сенсуса и вероятность выигрыша конкретного мнения. Обнаружена значимая положительная корреляция между плотностью внутреннего графа и временем достижения консенсуса, а также значительная отрицательная корреляция между центральностью влиятельных вершин и временем достижения консенсуса.

Разработан новый метод объяснимого искусственного интеллекта (ХАИ) под названием ShapG (объяснения на основе значения Шепли, посчитанного для графов) для измерения важности признаков. ShapG — модель-агностический метод глобальных объяснений. На первом этапе определяется неориентированный граф на основе набора данных, где вершины представляют признаки, а рёбра добавляются на основе расчёта коэффициентов корреляции между признаками. На втором этапе рассчитывается приближённый вектор Шепли с использованием выборки данных, принимая во внимание структуру графа. Выборочный подход ShapG позволяет эффективно рассчитывать важность признаков, то есть снижать вычислительную сложность. Сравнение ShapG с другими существующими методами ХАИ показывает, что он предоставляет более точные объяснения, что подтверждается на двух рассмотренных наборах данных. Мы также сравнили другие разработанные методы ХАИ, основанные на кооперативной теории игр, с ShapG по времени выполнения, и результаты показывают, что ShapG обладает очевидными преимуществами по времени выполнения, что дополнительно подтверждает эффективность ShapG. Кроме того, многочисленные эксперименты демонстрируют широкий спектр применений метода ShapG для объяснения сложных моделей. ShapG зарекомендовал себя как важный инструмент для повышения объяснимости и прозрачности систем ИИ, что позволяет надеяться на его широкое применение в различных областях.

Методы исследования

В диссертации использованы методы моделирования в статистической физике (метод Монте-Карло, моделирование молекулярной динамики), статистики (коэффициент корреляции, проверка гипотез), теории вероятностей (распределения случайных величин и стохастические процессы), теории игр (элементы кооперативной теории игр, включая вектор Шепли и значение Майерсона), машинного обучения (SHAP, LIME, важность признаков, LightGBM и многослойный перцептрон), теории графов (меры центральности и кратчайший путь), а также теории оптимизации и кластеризации (алгоритм K-means).

Теоретическая и практическая значимость

Результаты, представленные в данной диссертации, сосредоточены на моделировании динамики мнений, анализе сетей и их применении. Теоретическая значимость может быть кратко изложена следующим образом: (i) предложены математические формулировки статистико-физической модели и микроскопической модели динамики мнений в социальных сетях со сложными конфигурациями, что отражает новый взгляд на моделирование динамики мнений в таких сетях и модели могут быть использованы для изучения влияния различных сетевых структур и поведения индивидов на распространение мнений; (ii) предложено общее определение двухслойной сети с узлами-репликами, разработан новый подход к упрощению двухслойной сети до однослойной взвешенной сети на основе заданной динамики мнений; (iii) разработаны два быстрых и точных алгоритма для нахождения мер центральности, основанных на теории игр; (iv) предложен новый подход к определению важности признаков в моделях машинного обучения на основе мер центральности, и разработан новый метод объяснимого искусственного интеллекта (XAI) под названием ShapG.

Практическая значимость результатов проявляется в исследованиях по разработке эффективных стратегий коммуникации, таргетированного маркетинга, политических кампаний, общественных мероприятий здравоохранения и искусственного интеллекта. Разработанные меры центральности могут быть использованы для выявления самых влиятельных узлов в сети, что имеет важное значение не только для динамики мнений, но и для транспортных сетей, коммуникационных сетей и т.д. Разработанный алгоритм важности признаков (ShapG) может быть использован для объяснения моделей машинного обучения, разработки оптимальной стратегии обучения, экономии энергии при построении моделей машинного обучения и содействия углеродной нейтральности.

В главе 1 проведена серия численных симуляций на основе математической формы макроскопической модели GCVM. Обнаружено, что внешняя циклическая структура может значительно продлить время достижения консенсуса, а если индивиды в сети не склонны публично выражать свои мнения, обмен мнениями с близкими друзьями оказывается почти бесполезным. В главе 2 предложена и подтверждена гипотеза о существовании сильной линейной зависимости между временем достижения консенсуса и средними попарными кратчайшими путями d в структуре сети, что согласуется с результатами ис-

следования модели макроскопической GCVM. Глава 3 посвящена исследованию связи между мерами центральности и динамикой мнений в сети клуба каратэ Закари. Обнаружена значительная отрицательная корреляция между центральностью влиятельных вершин и временем достижения консенсуса. В Главе 4 разработан метод глобальных объяснений на основе вектора Шепли для графов (SharG), который является более точным и эффективным по сравнению с существующими методами (SHAP, LIME и Feature Importance).

Исследования, проведённые в диссертации, поддержаны стипендией Правительства Китая (CSC) № 202109010084 (2021–2025); грантом Российского научного фонда (РНФ) № 22-21-00346 «Игровые методы управления динамикой мнений в социальных сетях» (2022–2023).

Краткое описание структуры диссертации

Диссертация состоит из введения, четырёх глав, заключения, списка сокращений и символов, списка рисунков, списка таблиц, библиографии и аппендикса. Содержание каждой главы включает основные обозначения, описание модели и определения, подробное описание экспериментов, анализ результатов и краткое резюме главы. Диссертация содержит 158 страниц (149 страниц в английской версии), включая 38 рисунков и 34 таблиц. В библиографии цитируется 109 источников.

Первая глава диссертации посвящена исследованию динамики мнений в двухслойных сетях с лицемерными индивидами с использованием макроскопического подхода на основе методов статистической физики и теории вероятностей. В разделе 1.1 вводятся базовая модель избирателя (BSM) и модель скрытого избирателя (SVM). В разделе 1.2 приводится мотивация для общей модели скрытого избирателя (GCVM) (т.е. введение взаимодействия во внутреннем слое), и предложена математическая формулировка макроскопической GCVM с различными структурами сетей. В разделе 1.3 представлено описание экспериментов и результаты численных симуляций, которые обсуждаются в этом разделе. Раздел 1.4 содержит краткое заключение первой главы.

В отличие от первой главы, вторая глава сосредоточена на микроскопическом подходе к моделированию динамики мнений в двухслойных сетях с лицемерными индивидами. В разделах 2.1 и 2.2 приводятся определения многослойной и двухслойной сетей с узлами-репликами соответственно. В разделе 2.3 проводится сравнение микроскопической GCVM с макроскопической GCVM. В

разделе 2.4 представлены и обсуждаются численные эксперименты и их результаты. В разделе 2.5 кратко подводятся итоги второй главы.

Вместо использования классических сетевых структур, в третьей главе рассматривается реальная сетевая структура, а именно, сеть клуба каратэ Закари, которая расширена за счёт добавления второго (внутреннего) слоя коммуникации. Исследуется связь между центральностью влиятельных узлов и ключевыми показателями эффективности (временем достижения консенсуса и вероятности выигрыша мнения). В частности, в разделе 3.1 представлен подход к упрощению двухслойной сети с заданной динамикой мнений до однослойной взвешенной сети. В разделе 3.2 описывается сеть клуба каратэ Закари. В разделе 3.3 обсуждаются некоторые свойства сети (средние попарные кратчайшие пути и плотность графа), серия мер центральности (классические, теоретико-игровые и на основе случайных блужданий), а также два алгоритма для приближённого вычисления теоретико-игровых мер центральности. В разделе 3.4 проверяется эффективность подхода к упрощению двухслойной сети в однослойную, точность и производительность теоретико-игровых мер центральности, а также корреляция между свойствами сети и ключевыми показателями эффективности (временем достижения консенсуса и вероятностью выигрыша мнения). В разделе 3.5 делаются выводы по третьей главе.

Применяя идею приближённых теоретико-игровых мер центральности из третьей главы, четвёртая глава сосредоточена на применении разработанного алгоритма в области объяснимого искусственного интеллекта. В разделе 4.1 кратко рассматриваются существующие методы объяснимого искусственного интеллекта. В разделе 4.2 предложен новый метод глобальных объяснений. Подробное описание экспериментов приводится в разделе 4.3. В разделе 4.4 проводится сравнение результатов предложенного метода с существующими методами, такими, как SHAP и Feature Importance. В разделе 4.5 приводятся выводы по четвёртой главе.

Заключение диссертации содержит краткое описание полученных результатов.

Положения, выносимые на защиту

1. Предложена математическая формулировка динамики мнений под названием общая модель скрытого избирателя (GCVM). Модель обобщает мо-

дель скрытого избирателя (CVM) за счёт введения внутренних взаимодействий между агентами.

2. Предложена математическая формулировка статистико-физической модели (макроскопической GCVM) динамики мнений в различных структурах двухслойных сетей с лицемерием и внутренним взаимодействием. В рамках этой модели исследована взаимосвязь между характеристиками индивидов и процессами динамики мнений (время достижения консенсуса и вероятность выигрыша определенного мнения).
3. Предложено определение двухслойной сети с узлами-репликами. Представлена микроскопическая модель GCVM для любой структуры сети, удовлетворяющей данному определению. Разработана методология для проверки влияния каждого индивидуального параметра на время консенсуса и вероятность выигрыша. В рамках этой модели динамики мнений исследована взаимосвязь между свойствами сети и процессами динамики мнений.
4. Представлен подход к упрощению двухслойной сети с заданной динамикой мнений до однослойной взвешенной сети. Предложены новые подходы к приближённому вычислению теоретико-игровых мер центральности, которые значительно ускоряют оригинальные меры с высокой точностью. Исследована связь между мерами центральности и процессами динамики мнений.
5. Предложен новый метод определения важности признаков под названием ShapG для объяснения моделей машинного обучения. Методология ShapG основана на рассмотрении признаков в моделях машинного обучения как вершин в сети, при этом важность признаков рассматривается как центральность вершин. В диссертации разработан подробный алгоритм метода ShapG. Тестирование ShapG показывает, что он является более точным и эффективным, чем имеющиеся современные методы (SHAP, LIME и Feature Importance).
6. Разработана реализация с открытым исходным кодом глобального алгоритма объяснений ShapG [50].

7. Разработаны программы на Golang для моделирования микроскопической и макроскопической GCVM с различными структурами сетей и поведением индивидов.

Основные научные результаты

1. Предложена общая модель скрытого избирателя (GCVM), её математическая формулировка и соответствующие механизмы взаимодействия в процессах динамики мнений, см. статьи [24, 25, 27] в библиографии (личный вклад не менее 70%).
2. Проведено исследование взаимосвязей между характеристиками индивидов, свойствами сети и процессами динамики мнений (время достижения консенсуса и вероятность выигрыша мнения) с использованием численного моделирования, см. статьи [24, 25, 27] в библиографии (личный вклад не менее 70%).
3. Введено определение двухслойной сети с реплицированными узлами, подходящей для микроскопической модели GCVM для любой заданной сети, см. статью [51] в библиографии (личный вклад не менее 70%).
4. На основе численных экспериментов проведена проверка взаимосвязи центральности влиятельных узлов с процессами динамики мнений, см. статьи [26, 51] в библиографии (личный вклад не менее 70%).
5. Предложен метод упрощения двухслойной сети с заданной динамикой мнений до однослойной взвешенной сети, см. статьи [26, 51] в библиографии (личный вклад не менее 70%).
6. Предложены новые подходы аппроксимации теоретико-игровых мер центральности с высокой производительностью и точностью, см. статью [51] в библиографии (личный вклад не менее 70%).
7. Предложен новый метод SharG определения значимости признаков, применяемый для объяснения моделей машинного обучения, см. статью [52] в библиографии (личный вклад не менее 70%).
8. Разработана программа с открытым исходным кодом, реализующая работу метода SharG, см. [50] в библиографии (личный вклад – 100%).

9. Разработана программа на языке Golang, предназначенная для моделирования макроскопической модели GCVM с различной сетевой структурой и поведением индивидов, см. сертификат программы [53] в библиографии (личный вклад – 100%).

Апробация результатов

Основные результаты диссертации были представлены на Международной конференции «Stability and control processes» (Санкт-Петербург, 2022); Международной конференции «Game Theory and Management» (Санкт-Петербург, 2023, 2024); Научной школе «Dynamics of Complex Networks and their Applications» (2023); семинаре «Dynamic games and applications» GERAD (онлайн, 2024); Международной конференции «Mathematical Optimization Theory and Operations Research» (Омск, 2024).

Публикации

На основе результатов диссертации были опубликованы следующие работы: [24–27], все они индексируются в Scopus. Следующие статьи [24, 25] опубликованы в рецензируемых журналах из списка ВАК. Эти статьи [51, 52] поданы в журналы и в настоящее время находятся на рассмотрении. Получено свидетельство о регистрации программы для ЭВМ [53] с регистрационным номером № 2023661532. Разработана программа с открытым исходным кодом [50] и выпущена в индексе пакетов Python.

- [24] Zhao C., Parilina E. M. Opinion Dynamics in Two-Layer Networks with Hypocrisy // Journal of the Operations Research Society of China. – 2024. – Mar. – Vol. 12, no. 1. – P. 109-132. – Access mode: <https://doi.org/10.1007/s40305-023-00503-2>.
- [25] Zhao C., Parilina E. M. Analysis of consensus time and winning rate in two-layer networks with hypocrisy of different structures // Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes. – 2024. – Vol. 20, no. 2. – P. 170-192.
- [26] Zhao C., Parilina E. M. Network Structure Properties and Opinion Dynamics in Two-Layer Networks with Hypocrisy // Mathematical Optimization Theory

- and Operations Research. – Cham : Springer Nature Switzerland. – 2024. – P. 300-314.
- [27] Zhao C., Parilina E. M. Consensus time and winning rate based on simulations in two-layer networks with hypocrisy // 2023 7th Scientific School Dynamics of Complex Networks and their Applications (DCNA). – 2023. – P. 68-71.
- [50] Zhao C. shapG. – PyPI: <https://pypi.org/project/shapG/>; GitHub: <https://github.com/vectorsss/shapG>. – 2024. – July. – (Accessed on 11/16/2024).
- [51] Zhao C., Parilina E. M. Centrality measures and opinion dynamics in two-layer networks with replica nodes // arXiv preprint arXiv:2406.18780v2. – 2024. – 2406.18780v2.
- [52] Zhao C., Liu J., Parilina E. M. ShapG: new feature importance method based on the Shapley value // arXiv preprint arXiv:2407.00506. – 2024. – 2407.00506.
- [53] Свидетельство о государственной регистрации программы для ЭВМ № 2023661532 Российская Федерация : Программа для моделирования динамики распространения бинарных мнений в двухслойных сетях Чжао Ч. (CN) ; федеральное государственное бюджетное образовательное учреждение высшего образования “Санкт-Петербургский государственный университет” (СПбГУ) ; Чжао Чи. – № 2023660681 ; заявл. 24.05.2023 ; опубл. 01.06.2023.

Благодарности

Автор выражает глубокую признательность многим увлечённым и неравнодушным преподавателям, которые направляли его, сделав возможным написание этой диссертации. Искренние благодарности относятся доктору физико-математических наук, профессору Парилиной Елене Михайловне (Санкт-Петербургский государственный университет) за её бесценные рекомендации, поддержку и плодотворные обсуждения, которые существенно способствовали профессиональному росту автора как исследователя. Автор также выражает благодарность Ганкевичу Ивану Геннадьевичу, своему первому руководителю в Huawei, от которого он многому научился во время их совместной работы. Шаблон данной диссертации в основном основан на работе с открытым исходным

кодом, созданной Ганкевичем. Автор хотел бы выразить благодарность своей девушке Челноковой Анастасии Дмитриевне за ее неизменную любовь, понимание и постоянную поддержку в любых обстоятельствах. Особую благодарность автор выражает своему дедушке, своему первому учителю математики, за то, что он вдохновил его любовь к учёбе. Автор искренне благодарит свою семью за их неустанную поддержку и ободрение на протяжении всего этого сложного, но увлекательного пути. Кроме того, он хотел бы отметить поддержку и содействие, полученные от Китайского Совета по стипендиям в ходе обучения в аспирантуре.

Глава 1. Динамика мнений в двухслойных сетях в присутствии лицемерия

В этой главе механизмы внутреннего взаимодействия добавлены в скрытую модель избирателя (CVM) и формализуется макроскопическая модель GCVM для различных сетевых структур. Проводится серия численных экспериментов, результаты которых анализируются. Обсуждается влияние структуры сети на время достижения консенсуса и вероятность победы определенного мнения. Результаты, представленные в этой главе, опубликованы в статье [24].

1.1 Базовая и скрытая модели избирателя

Предполагается, что общество представлено агентами или индивидами, которые имеют мнения по обсуждаемой теме. Мнение бинарно, т.е. мнение каждого индивида может быть, скажем, либо красным, либо синим. Вводим модели динамики мнений с учетом сетевой структуры, которую формируют индивиды в социальной сети, называемой общей скрытой моделью избирателя (GCVM). Прежде чем вводить GCVM, кратко представим базовую модель избирателя (BVM) и скрытую модель избирателя (CVM) (см. [21, 22]).

Назовем классическую модель избирателя базовой моделью избирателя. В BVM предполагается, что каждый в сети может публично выражать свое мнение, поэтому в этой модели динамики мнений существует только один слой обмена информацией и единственное мнение индивида по обсуждаемой теме. В CVM, описанной в разделе 1.1.2, индивиды могут либо выражать, либо скрывать свое истинное (скрытое) мнение публично. У индивида есть как публично выраженное, так и личное мнение по теме, и личное мнение неизвестно другим. Если индивид скрывает свое истинное мнение, т.е. его публично известное мнение отличается от личного, мы называем это *лицемерием*: т.е. индивид с

когнитивным диссонансом [54]. В SVM существует два слоя (внешний и внутренний), но обмен информацией во внутреннем слое не допускается.

Сфокусируемся на времени, необходимом для достижения консенсуса группами в различных сетевых структурах, это время называется временем достижения консенсуса, а также на вероятности победы мнения.

1.1.1 Базовая модель избирателя (BVM)

Предположим, что имеется заданная сеть G , представляющая структуру коммуникации участников в обществе. Обозначим количество индивидов в сети G как N . Рассмотрим эволюцию сети во времени. Используем следующие обозначения:

- $\omega_{ext}(\alpha, t) \in \{0, 1\}$ — мнение индивида α в момент времени t (0 соответствует синему, 1 — красному мнению);
- c — коэффициент копирования того, что индивид примет мнение своего соседа;
- r — количество индивидов с красным мнением;
- ρ — доля индивидов с красным мнением в популяции (сила красного мнения), то есть $\rho = r/N$;
- T_{cons} — время достижения консенсуса, в BVM оно означает время, необходимое для того, чтобы все индивиды пришли к одному мнению ($\rho = 1$ или $\rho = 0$).

Доля красного мнения используется для представления состояния такой системы в BVM, и $\rho \in \{0, \frac{1}{N}, \dots, \frac{N-1}{N}, 1\}$. Предполагается, что любой индивид может изменить своё мнение в результате реализации случайного события. Полагаем, что интервал между такими событиями подчиняется экспоненциальному распределению с интенсивностью λ . Также определяем следующую стохастическую матрицу с элементами $P(\rho, \cdot)$ для каждого $\rho \notin \{0, 1\}$, где $P(\rho, \rho')$ — вероятность перехода системы BVM из состояния ρ в состояние ρ' после наступления события. Если $\rho \in \{0, 1\}$, то $P(0, 0) = 1$ и $P(1, 1) = 1$, что означает,

что если все индивиды в сети имеют одинаковое мнение, они никогда его не изменят. Доля красного мнения ρ может измениться на $\rho \pm \frac{1}{N}$ или остаться той же с положительными вероятностями:

$$\begin{aligned} P(\rho, \rho + \frac{1}{N}) &= \frac{c(N-r)r}{N(N-1)} = \frac{c(1-\rho)r}{N-1}, \\ P(\rho, \rho - \frac{1}{N}) &= \frac{cr(N-r)}{N(N-1)} = \frac{cr(1-\rho)}{N-1}, \\ P(\rho, \rho) &= 1 - 2\frac{cr(1-\rho)}{N-1}. \end{aligned} \quad (1)$$

Если изменение состояния системы BVM рассматривается как случайное событие, тогда скорость изменения системы за единицу времени может быть представлена как сумма скоростей изменения в положительном направлении λ_+ и в отрицательном направлении λ_- , где

$$\begin{aligned} \lambda_- &= cr\frac{N-r}{N} = \lambda_+ = c(N-r)\frac{r}{N}, \\ \lambda &= \lambda_+ + \lambda_- = 2c(N-r)\frac{r}{N}. \end{aligned} \quad (2)$$

Таким образом, имеется система с интенсивностью событий λ за единицу времени. Следующие шаги показывают, как индивиды взаимодействуют друг с другом в BVM:

1. **Инициализация:** Задана доля ρ индивидов с красным мнением, остальные индивиды придерживаются синего мнения. Инициализируем время $t \rightarrow 0$.
2. **Итерация:**
 - а. Выбрать «фокусного» индивида f случайным образом из всех N индивидов.
 - б. Выбрать соседа a фокусного индивида случайным образом из всех его соседей.
 - с. Сгенерировать временную переменную $\lambda_t = \lambda \cdot u$, где $u \sim U(0, 1)$, и индивид f принимает мнение соседа a следующим образом:
 - i. Состояние системы изменится с ρ на $\rho + \frac{1}{N}$, если $\lambda_t < \lambda_+$;
 - ii. Состояние системы изменится с ρ на $\rho - \frac{1}{N}$, если $\lambda_+ \leq \lambda_t \leq \lambda_+ + \lambda_- = \lambda$,

- d. Увеличить время на случайное число Δt , выбранное из экспоненциального распределения с математическим ожиданием $\frac{1}{\lambda}$, $t \rightarrow t + \Delta t$.
- e. Если группа достигла консенсуса, установить $t \rightarrow T_{\text{cons}}^{(\text{BVM})}$ и завершить процесс. В противном случае вернуться к шагу а.

Рисунок 1 представляет сеть BVM с 10 индивидами в полной сети.

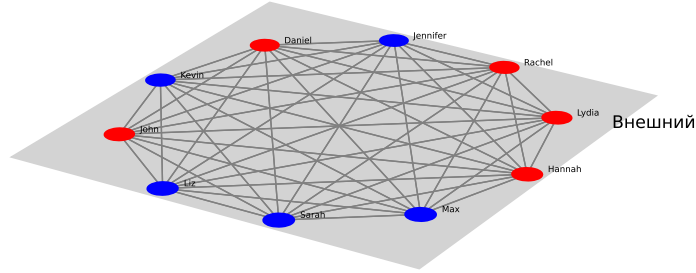


Рис. 1: Полная сеть, соединяющая 10 индивидов в BVM

1.1.2 Скрытая модель избирателя (CVM)

В CVM задана двухслойная сеть, определенная двумя заданными сетями G_1 и G_2 с одинаковым набором из N индивидов. Индивид i представлен узлом E_i во внешнем слое сети G_1 и узлом I_i во внутреннем слое сети G_2 . Таким образом, каждый индивид представлен парой узлов (E_i, I_i) , $i = 1, \dots, N$. Следует отметить, что на структуру сети G_1 не накладывается ограничений, но сеть G_2 всегда является пустой сетью с N индивидами. Двухслойная сетевая структура удовлетворяет следующим условиям:

1. Между узлами I_i и I_j нет связей, $\forall i, j = 1, \dots, N$.
2. Между узлами E_i и I_i должна существовать связь, $\forall i = 1, \dots, N$.

Рисунок 2 представляет такую двухслойную сеть с 10 индивидами, когда внешний слой — полная сеть.

В CVM используется R, B (r, b) для обозначения внешних (внутренних) красных и синих мнений индивидов соответственно. Обозначим множество со-

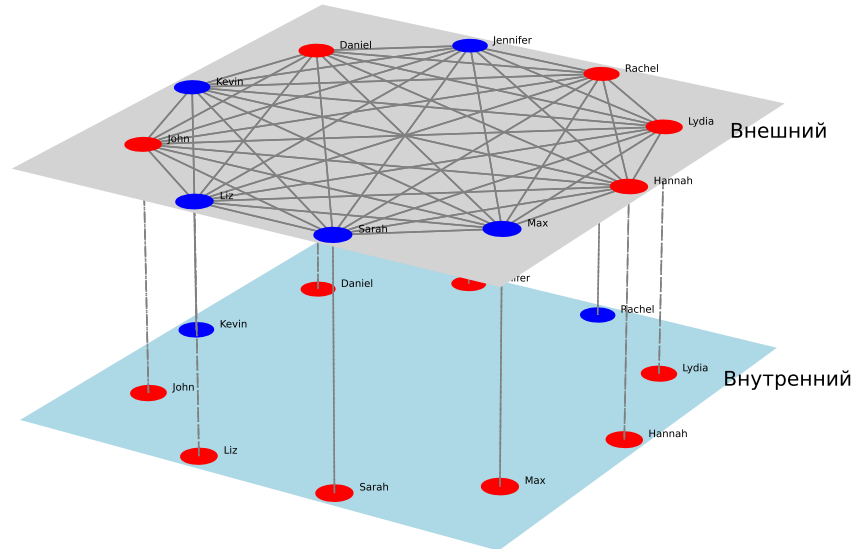


Рис. 2: Двухслойная сетевая структура в CVM с 10 индивидами

стояний мнений индивида через $S = \{Rr, Rb, Bb, Br\}$. Используем следующие обозначения для систем CVM:

- $\omega(\alpha, t) \in S$ — состояние мнения индивида α в момент времени t ;
- c — коэффициент копирования (аналогичный тому, что используется в BVM);
- e — коэффициент экстернализации, то есть коэффициент того, что лицемер решит публично выразить своё внутреннее мнение;
- i — коэффициент интернализации, то есть коэффициент того, что лицемер решит принять своё внешнее мнение;
- r_e — количество индивидов с внешним красным мнением;
- r_i — количество индивидов с внутренним красным мнением;
- r — количество индивидов с одновременно внешним и внутренним красным мнением;
- $\rho_{r_e}, \rho_{r_i}, \rho_r$ — доля индивидов с внешним красным мнением, внутренним красным мнением и красным мнением одновременно во внешнем и внутреннем слоях соответственно;

- T_{cons} — время достижения консенсуса, в (G)CVM оно означает время, необходимое для того, чтобы все индивиды пришли к одному мнению как во внешнем, так и во внутреннем слоях (то есть, $\rho_{r_e} = \rho_{r_i} = \rho_r = 0$ или $\rho_{r_e} = \rho_{r_i} = \rho_r = 1$).

В сети, представленной на рисунке 2, есть пять лицемеров.

Во внешнем слое CVM существует полная сеть, и двухслойная сеть удовлетворяет Условиям 1-2, указанным выше. Эффект лицемерия или когнитивного диссонанса может быть уменьшен с помощью экстернализации или интернализации. При экстернализации лицемер публично выражает своё мнение, а при интернализации принимает ранее выраженное мнение.

Если нас интересует количество лицемеров, можно сосредоточиться на ρ_r , описывающем стохастическую матрицу Q_{CVM} в системе CVM, или можно описать количество индивидов, находящихся в любом состоянии из множества S (последнее рассматривается в [21]). Предположим, что внешний слой в системе CVM — это полный граф, тогда количество индивидов в каждом состоянии определяется следующим образом:

$$\begin{aligned}
 N &= \#Rr + \#Rb + \#Br + \#Bb, \\
 r_e &= \#Rr + \#Rb, \\
 r_i &= \#Rr + \#Br, \\
 r &= \#Rr, \\
 \#Bb &= N - r_e - r_i + r, \\
 \#Rb &= r_e - r, \\
 \#Br &= r_i - r,
 \end{aligned} \tag{3}$$

где $\#s$ — количество индивидов в состоянии $s \in S$.

Можно представить состояние системы с N индивидами с помощью тройки $(\rho_{r_e}, \rho_{r_i}, \rho_r)$. Переходы с положительными вероятностями из одного состояния в другое для описанной системы CVM приведены в таблице 1. Используем Q для обозначения матрицы переходов.

Процедура обновления в CVM похожа на BVM, обобщим различия следующим образом:

- Используем тройку $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ вместо ρ для представления состояния;

Таблица 1: Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе SVM

| Новое состояние (x, y, z) | Как можно достичь нового состояния? | Вероятность | Элемент матрицы переходов Q_{SVM} за единицу времени |
|--|--|--|--|
| $(\rho_{r_e} + \frac{1}{N}, \rho_{r_i}, \rho_r)$ | $Bb \rightarrow Rb$, один индивид в состоянии Bb копирует мнение соседа с внешним мнением R | $\frac{c(N-r_e-r_i+r)r_e}{N(N-1)+r_e+r_i-2r}$ | $\frac{c(N-r_e-r_i+r)r_e}{N}$ |
| $(\rho_{r_e} + \frac{1}{N}, \rho_{r_i}, \rho_r + \frac{1}{N})$ | $Bb \rightarrow Rr$, один индивид в состоянии Bb копирует мнение соседа с внешним мнением R или выражает своё внутреннее мнение | $\frac{c(r_i-r)r_e+e(r_i-r)}{N(N-1)+r_e+r_i-2r}$ | $(r_i - r)(c\frac{r_e}{N} + e)$ |
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r + \frac{1}{N})$ | $Rb \rightarrow Rr$, один индивид в состоянии Rb принимает своё внешнее мнение | $\frac{i(r_e-r)}{N(N-1)+r_e+r_i-2r}$ | $(r_e - r)i$ |
| $(\rho_{r_e} - \frac{1}{N}, \rho_{r_i}, \rho_r)$ | $Rb \rightarrow Bb$, один индивид в состоянии Rb копирует мнение соседа с внешним мнением B или выражает своё внутреннее мнение | $\frac{c(r_e-r)(N-r_e)+e(r_e-r)}{N(N-1)+r_e+r_i-2r}$ | $(r_e - r)(c\frac{N-r_e}{N} + e)$ |
| $(\rho_{r_e} - \frac{1}{N}, \rho_{r_i}, \rho_r - \frac{1}{N})$ | $Rr \rightarrow Br$, один индивид в состоянии Rr копирует мнение соседа с внешним мнением B | $\frac{cr(N-r_e)}{N(N-1)+r_e+r_i-2r}$ | $\frac{cr(N-r_e)}{N}$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r)$ | $Bb \rightarrow Bb$, один индивид в состоянии Bb принимает своё внешнее мнение | $\frac{i(r_i-r)}{N(N-1)+r_e+r_i-2r}$ | $(r_i - r)i$ |
| $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ | обратные ко всем выше | | отрицательная сумма вышеуказанных элементов |

- Как показано в таблице 1, существует 6 возможных изменений состояния, используем ту же схему, что и в BVM, но с другой кусочной функцией для определения изменений состояния.

Основываясь на таблице 1, можем описать систему SVM, выразив ρ_r — долю лицемеров в сети, как это дано в уравнении (4), где $Q_{SVM}[x, y] = \{P(x, y)\}$ представляет вероятность перехода из состояния x (предыдущее значение ρ_r) в состояние y (следующее значение ρ_r). Для $\rho_r \notin \{0, 1\}$, мы имеем уравнение (4) и $P(0, 0) = P(1, 1) = 1$:

$$\begin{aligned}
 P\left(\rho_r, \rho_r + \frac{1}{N}\right) &= Q_{SVM}\left[\left(\rho_{r_e}, \rho_{r_i}, \rho_r\right), \left(\rho_{r_e} + \frac{1}{N}, \rho_{r_i}, \rho_r + \frac{1}{N}\right)\right] \\
 &\quad + Q_{SVM}\left[\left(\rho_{r_e}, \rho_{r_i}, \rho_r\right), \left(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r + \frac{1}{N}\right)\right], \\
 P\left(\rho_r, \rho_r - \frac{1}{N}\right) &= Q_{SVM}\left[\left(\rho_{r_e}, \rho_{r_i}, \rho_r\right), \left(\rho_{r_e} - \frac{1}{N}, \rho_{r_i}, \rho_r - \frac{1}{N}\right)\right] \\
 P\left(\rho_r, \rho_r\right) &= 1 - P\left(\rho_r, \rho_r + \frac{1}{N}\right) - P\left(\rho_r, \rho_r - \frac{1}{N}\right).
 \end{aligned} \tag{4}$$

1.2 Общая скрытая модель избирателя (GSVM)

1.2.1 Обоснование модели

Идея общей скрытой модели избирателя (GSVM) проста. Предположим, что во внутреннем слое существуют связи. Логично, что у человека (индивида) есть свои близкие друзья, с которыми он делится своим истинным (внутренним) мнением. Это означает, что человек и его друзья могут образовать свою собственную группу или клику в сети. Если клики формируются во внутреннем слое, и люди никогда не делятся истинными мнениями с другими людьми за пределами клики, за исключением случаев, когда они связаны во внутреннем слое. Сначала мы определяем вероятности переходов для специального случая внутренней сети, когда она представлена полным графом, затем описываем динамику сети в случае неполного внутреннего слоя.

1.2.2 Симметричный случай: полная внутренняя сеть

Сначала рассмотрим простой случай GCVM, когда как внешний, так и внутренний слои являются полными сетями. Рисунок 3 представляет симметричный случай GCVM с десятью индивидами.

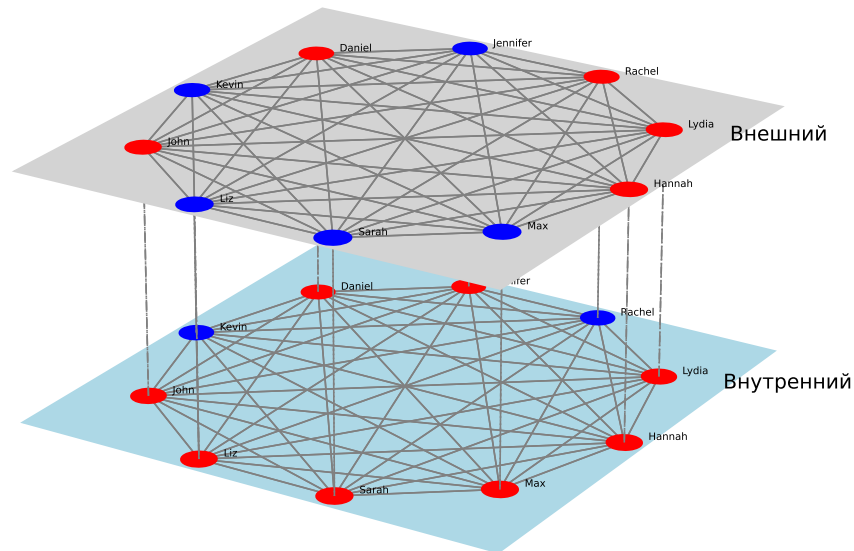


Рис. 3: Представление GCVM с 10 индивидами: симметричный случай

Матрица переходов представлена в таблице 2. Существует восемь случаев переходов состояний с ненулевой вероятностью, в то время как в системе SVM существует шесть таких случаев, поскольку в системе GCVM состояния Bb и Rr могут изменяться на Br и Rb соответственно, что вызвано внутренними взаимодействиями.

По сравнению с SVM, новая модель GCVM предполагает наличие внутренних взаимодействий, которые влияют на процесс передачи внутреннего мнения. В SVM изменения внутреннего мнения могут происходить только через интернализацию (индивид принимает своё внешнее мнение), тогда как в GCVM внутреннее мнение может изменяться как через интернализацию, так и через взаимодействие между внутренними узлами (или взаимодействие индивидов во внутреннем слое). Например, индивиды без лицемерия (индивиды в состояниях Bb или Rr) не могут изменять свои состояния на Br или Rb соответственно в системе SVM, но это возможно в системе GCVM. Кроме того, внутреннее взаимодействие увеличивает время достижения консенсуса и изменяет вероятность победы каждого из мнений.

Таблица 2: Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GСVM в симметричном случае

| Новое состояние (x, y, z) | Как можно достичь нового состояния? | Вероятность | Элемент матрицы переходов $Q_{GСVM}$ за единицу времени |
|--|---|---|---|
| $(\rho_{r_e} + \frac{1}{N}, \rho_{r_i}, \rho_r)$ | $Bb \rightarrow Rb$, один индивид в состоянии Bb копирует мнение соседа с внешним мнением R | $\frac{c(N-r_e-r_i+r)r_e}{2N(N-1)+r_e+r_i-2r}$ | $\frac{c(N-r_e-r_i+r)r_e}{N}$ |
| $(\rho_{r_e} + \frac{1}{N}, \rho_{r_i}, \rho_r + \frac{1}{N})$ | $Br \rightarrow Rr$, один индивид в состоянии Br копирует мнение соседа с внешним мнением R или выражает своё внутреннее мнение | $\frac{c(r_i-r)r_e+c(r_i-r)}{2N(N-1)+r_e+r_i-2r}$ | $(r_i - r)(c\frac{r_e}{N} + e)$ |
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r + \frac{1}{N})$ | $Rb \rightarrow Rr$, один индивид в состоянии Rb принимает своё внешнее мнение или копирует мнение соседа с внутренним мнением r | $\frac{i(r_e-r)+c(r_e-r)r_i}{2N(N-1)+r_e+r_i-2r}$ | $(r_e - r)(i + c\frac{r_i}{N})$ |
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r)$ | $Bb \rightarrow Br$, один индивид в состоянии Bb копирует мнение соседа с внутренним мнением r | $\frac{c(N-r_e-r_i+r)r_i}{2N(N-1)+r_e+r_i-2r}$ | $\frac{c(N-r_e-r_i+r)r_i}{N}$ |
| $(\rho_{r_e} - \frac{1}{N}, \rho_{r_i}, \rho_r)$ | $Rb \rightarrow Bb$, один индивид в состоянии Rb копирует мнение соседа с внешним мнением B или выражает своё внутреннее мнение | $\frac{c(r_e-r)(N-r_e)+c(r_e-r)}{2N(N-1)+r_e+r_i-2r}$ | $(r_e - r)(c\frac{N-r_e}{N} + e)$ |
| $(\rho_{r_e} - \frac{1}{N}, \rho_{r_i}, \rho_r - \frac{1}{N})$ | $Rr \rightarrow Br$, один индивид в состоянии Rr копирует мнение соседа с внешним мнением B | $\frac{cr(N-r_e)}{2N(N-1)+r_e+r_i-2r}$ | $\frac{cr(N-r_e)}{N}$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r)$ | $Br \rightarrow Bb$, один индивид в состоянии Br принимает своё внешнее мнение или копирует мнение соседа с внутренним мнением b | $\frac{i(r_i-r)+c(r_i-r)(N-r_i)}{2N(N-1)+r_e+r_i-2r}$ | $(r_i - r)(i + c\frac{N-r_i}{N})$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r - \frac{1}{N})$ | $Rr \rightarrow Rb$, один индивид в состоянии Rr копирует мнение соседа с внутренним мнением b | $\frac{cr(N-r_i)}{2N(N-1)+r_e+r_i-2r}$ | $\frac{cr(N-r_i)}{N}$ |
| $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ | обратные ко всем выше | единица минус сумма вышеприведенных значений | отрицательная сумма вышепуказанных элементов |

Аналогично, можем представить стохастическую матрицу относительно ρ_r . Для $\rho_r \notin \{0, 1\}$ вычисляем вероятности переходов по уравнению (5), где $P(0, 0) = P(1, 1) = 1$:

$$\begin{aligned}
P\left(\rho_r, \rho_r + \frac{1}{N}\right) &= Q_{GCV M} \left[(\rho_{r_e}, \rho_{r_i}, \rho_r), \left(\rho_{r_e} + \frac{1}{N}, \rho_{r_i}, \rho_r + \frac{1}{N}\right) \right] \\
&\quad + Q_{GCV M} \left[(\rho_{r_e}, \rho_{r_i}, \rho_r), \left(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r + \frac{1}{N}\right) \right], \\
P\left(\rho_r, \rho_r - \frac{1}{N}\right) &= Q_{GCV M} \left[(\rho_{r_e}, \rho_{r_i}, \rho_r), \left(\rho_{r_e} - \frac{1}{N}, \rho_{r_i}, \rho_r - \frac{1}{N}\right) \right] \\
&\quad + Q_{GCV M} \left[(\rho_{r_e}, \rho_{r_i}, \rho_r), \left(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r - \frac{1}{N}\right) \right], \\
P(\rho_r, \rho_r) &= 1 - P\left(\rho_r, \rho_r + \frac{1}{N}\right) - P\left(\rho_r, \rho_r - \frac{1}{N}\right).
\end{aligned} \tag{5}$$

Используем ту же меру, что и в [21, 22], для представления силы красного мнения, а именно

$$m(\rho_{r_e}, \rho_{r_i}) = \frac{i\rho_{r_e} + e\rho_{r_i}}{e + i}, \tag{6}$$

где $m \in [0, 1]$ — взвешенное значение, которое объединяет доли красных мнений во внутренних и внешних слоях, и m можно рассматривать как эквивалент ρ в BVM.

1.2.3 Асимметричные случаи: неполная внутренняя сеть

В этом разделе рассматриваются несколько асимметричных случаев, включая внутреннюю сеть с центральной звездой, две внутренние сети с центральными звёздами и внутренние двухкликковые сети, при этом внешний слой всегда является полной сетью. Предполагается, что обе группы, представленные во внутренних двузвёздных и двухкликковых сетях, равны по размеру.

Внутренний слой: сеть с центральной звездой

В этом случае есть центральный узел во внутреннем слое (граф звезда), как показано на рисунке 4. Взаимодействие во внутреннем слое существу-

ет только между центральным узлом и любым нецентральным узлом. Между нецентральными узлами взаимодействия нет.

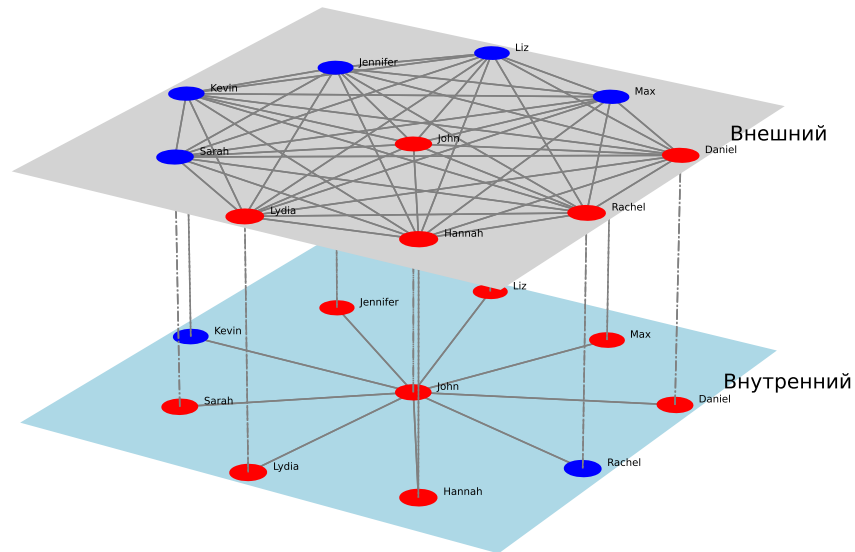


Рис. 4: Представление внутренней сети с центральной звездой

Для внутренней сети с центральной звездой существует 8 типов переходов с ненулевыми вероятностями, описанных в таблице 3.

Если сравнить переходы, указанные в таблицах 2 и 3, можно сделать вывод, что изменения структуры внутренней сети влияют только на скорость переходов следующих состояний: $Rb \rightarrow Rr$, $Bb \rightarrow Br$, $Br \rightarrow Bb$ и $Rr \rightarrow Rb$. Таким образом, в следующих разделах, где рассматриваются другие структуры внутренней сети, будем указывать только интенсивности переходов для этих четырёх случаев. Все остальные интенсивности переходов ($Bb \rightarrow Rb$, $Br \rightarrow Rr$, $Rb \rightarrow Bb$ и $Rr \rightarrow Br$) остаются такими же, как в таблице 2.

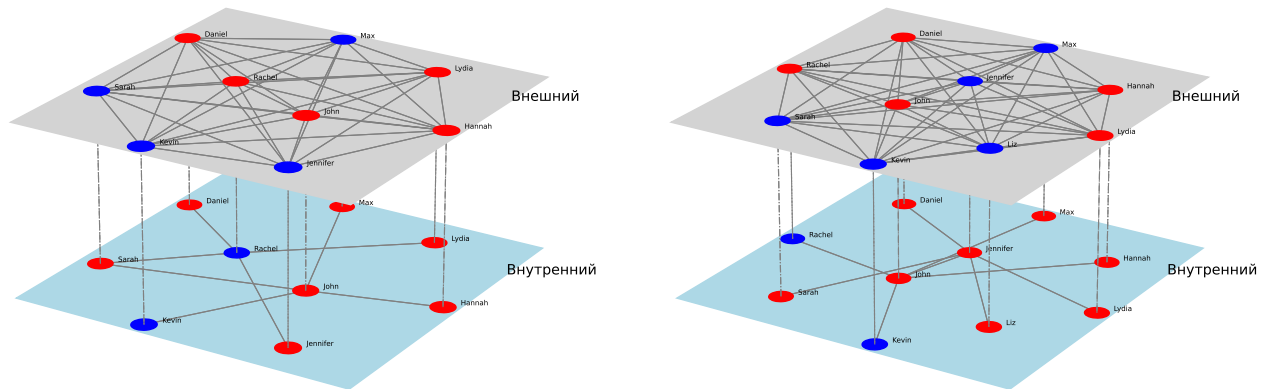
Внутренний слой: двузвёздная сеть

Предположим, что во внутреннем слое имеется две сети, соединённые по принципу звезды, и рассмотрим два возможных случая:

1. Две сети-звезды, имеют один общий узел (общее количество узлов — нечётное число).
2. Две сети-звезды, имеют два общих узла (общее количество узлов — чётное число).

Таблица 3: Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GCVM с внутренней сетью, соединённой по принципу звезды

| Новое состояние (x, y, z) | Как можно достичь нового состояния? | Элемент матрицы переходов Q_{GCVM} за единицу времени |
|--|--|---|
| $(\rho_{r_e} + \frac{1}{N}, \rho_{r_i}, \rho_r)$ | $Bb \rightarrow Rb$, один индивид в состоянии Bb копирует мнение соседа с внешним мнением R | $\frac{c(N-r_e-r_i+r)r_e}{N}$ |
| $(\rho_{r_e} + \frac{1}{N}, \rho_{r_i}, \rho_r + \frac{1}{N})$ | $Br \rightarrow Rr$, один индивид в состоянии Br копирует мнение соседа с внешним мнением R или выражает своё внутреннее мнение | $(r_i - r)(c\frac{r_e}{N} + e)$ |
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r + \frac{1}{N})$ | $Rb \rightarrow Rr$, индивид Rb во внутреннем слое имеет мнение b , он принимает своё внешнее мнение или копирует мнение соседа с внутренним мнением r ; другие индивиды имеют мнение Bb , один из них принимает своё внешнее мнение или копирует авторитета, если у него внутреннее мнение r | $\frac{1}{N}(r_e - r)(i + c\frac{r_i}{N}) + \frac{N-1}{N}(r_e - r)(i + c\frac{r_i}{N^2})$ |
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r)$ | $Bb \rightarrow Br$, индивид Bb с мнением b во внутреннем слое копирует мнение соседа с внутренним мнением r ; другие индивиды имеют мнение Bb , один из них копирует индивида, если у него внутреннее мнение r | $\frac{1}{N}(N - r_e - r_i + r)c\frac{r_i}{N} + \frac{N-1}{N}(r_e - r)(i + c\frac{r_i}{N^2})$ |
| $(\rho_{r_e} - \frac{1}{N}, \rho_{r_i}, \rho_r)$ | $Rb \rightarrow Bb$, один индивид в состоянии Rb копирует мнение соседа с внешним мнением B или выражает своё внутреннее мнение | $(r_e - r)(c\frac{N-r_e}{N} + e)$ |
| $(\rho_{r_e} - \frac{1}{N}, \rho_{r_i}, \rho_r - \frac{1}{N})$ | $Rr \rightarrow Br$, один индивид в состоянии Rr копирует мнение соседа с внешним мнением B | $\frac{cr(N-r_e)}{N}$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r)$ | $Br \rightarrow Bb$, индивид Br с мнением r во внутреннем слое принимает своё внешнее мнение или копирует мнение соседа с внутренним мнением b ; другие индивиды имеют мнение Br , один из них принимает своё внешнее мнение или копирует индивида, если у него внутреннее мнение b | $\frac{1}{N}(r_i - r)(i + c\frac{N-r_i}{N}) + \frac{N-1}{N}(r_i - r)(i + c\frac{N-r_i}{N^2})$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r - \frac{1}{N})$ | $Rr \rightarrow Rb$, авторитет во внутреннем слое имеет мнение Rr , он копирует мнение соседа с внутренним мнением b ; другие индивиды имеют мнение Rr , один из них копирует индивида, если у него внутреннее мнение b | $\frac{r}{N}c\frac{N-r_i}{N} + \frac{r(N-1)}{N}c\frac{N-r_i}{N^2}$ |
| $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ | обратные ко всем выше | отрицательная сумма выше-указанных элементов |



(a) Нечётное количество узлов

(b) Чётное количество узлов

Рис. 5: Представление внутренней двухзвёздной сети

Рисунок 5а показывает простой пример первого случая, где каждая подсеть-звезда делит общий узел «Sarah». Если мы определим количество хабов (центральных узлов) как a , то для этого случая $a = 2$. Также определим ρ_g как долю узлов в одной подсети-звезде относительно общего числа узлов в сети, и $\rho_g = \frac{N+1}{2N}$. Переходы для первого случая внутренней сети, соединённой по принципу двух звёзд, описаны в таблице 4.

Рисунок 5b показывает пример, когда во внутреннем слое — двухзвёздная сеть, где две звезды соединяются одним ребром через свои хабы (центральные узлы) [55]. У подсетей-звёзд два общих узла — «John» и «Jennifer», и каждая подсеть состоит из пяти узлов. В этом случае $\rho_g = \frac{N+2}{2N}$. Переходы описаны в таблице 5.

Внутренний слой: двухкликковая сеть

Предположим, что во внутреннем слое есть две клики, и в этой сети существует одно межкликковое ребро. В каждой клике одинаковое количество индивидов. Возможны два случая:

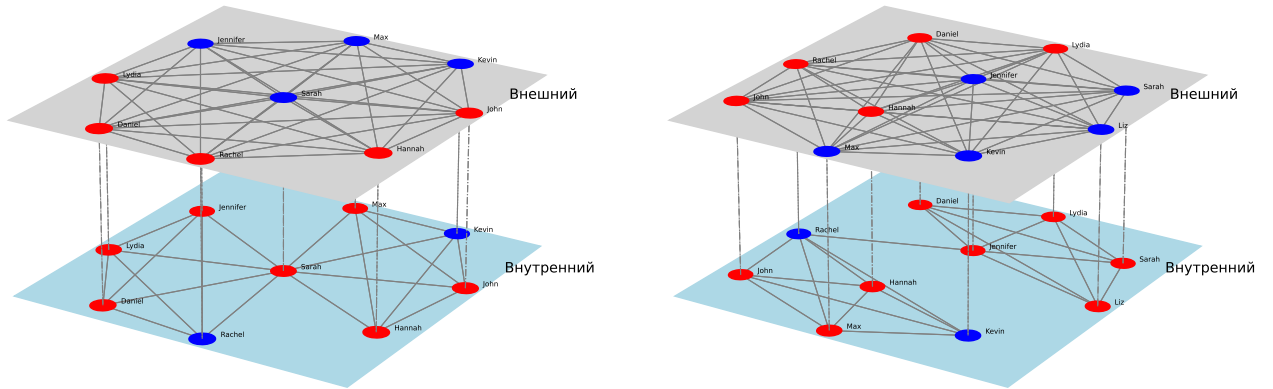
1. Две клики имеют один общий узел (этот общий узел имеет степень $N - 1$);
2. Две клики имеют два общих узла (эти два общих узла имеют степень $\frac{2}{N}$).

Таблица 4: Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GCVM с двухзвёздной внутренней сетью (нечётный случай)

| Новое состояние (x, y, z) | Как можно достичь нового состояния? | Элемент матрицы переходов Q_{GCVM} за единицу времени |
|--|---|---|
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r + \frac{1}{N})$ | $Rb \rightarrow Rr$, индивид с мнением Rb принимает своё внешнее мнение R или копирует мнение соседа с внутренним мнением r ; другие индивиды имеют мнение Rb , один из них принимает своё внешнее мнение или копирует индивида, если у него внутреннее мнение r | $\frac{a}{N}(r_e - r)(i + c\frac{r_i}{N}\rho_g) + \frac{N-a}{N}(r_e - r)i + \frac{N-a+1}{N}(r_e - r)c\frac{r_i}{N^2}$ |
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r)$ | $Bb \rightarrow Br$, индивид с мнением Bb копирует мнение соседа с внутренним мнением r ; другие индивиды имеют мнение Bb , один из них копирует индивида, если у него внутреннее мнение r | $\frac{a}{N}(N - r_e - r_i + r)c\frac{r_i}{N}\rho_g + \frac{N-a+1}{N}(N - r_e - r_i + r)c\frac{r_i}{N^2}$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r)$ | $Br \rightarrow Bb$, индивид с мнением Br принимает своё внешнее мнение или копирует мнение соседа с внутренним мнением b ; другие индивиды имеют мнение Br , один из них принимает своё внешнее мнение или копирует индивида, если у него внутреннее мнение b | $\frac{a}{N}(r_i - r)(i + c\frac{N-r_i}{N}\rho_g) + \frac{N-a}{N}(r_i - r)i + \frac{N-a+1}{N}(r_i - r)c\frac{N-r_i}{N^2}$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r - \frac{1}{N})$ | $Rr \rightarrow Rb$, индивид с мнением Rr копирует мнение соседа с внутренним мнением b ; другие индивиды имеют мнение Rr , один из них копирует индивида, если у него внутреннее мнение b | $r\frac{a}{N}c\frac{N-r_i}{N}\rho_g + r\frac{(N-a+1)}{N}c\frac{N-r_i}{N^2}$ |

Таблица 5: Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GCVМ с внутренней двухзвёздной сетью (чётный случай)

| Новое состояние (x, y, z) | Как можно достичь нового состояния? | Элемент матрицы переходов $Q_{GCVМ}$ за единицу времени |
|--|--|---|
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r + \frac{1}{N})$ | $Rb \rightarrow Rr$, авторитет во внутреннем слое имеет мнение Rb , он принимает своё внешнее мнение или копирует мнение соседа с внутренним мнением r ; другие индивиды имеют мнение Rb , один из них принимает своё внешнее мнение или копирует авторитета, если у него внутреннее мнение r . | $\frac{a}{N}(r_e - r)(i + c\frac{r_i}{N}\rho_g) + \frac{N-a}{N}(r_e - r)(i + c\frac{r_i}{N^2})$ |
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r)$ | $Bb \rightarrow Br$, авторитет во внутреннем слое имеет мнение Bb , он копирует мнение соседа с внутренним мнением r ; другие индивиды имеют мнение Bb , один из них копирует авторитета, если у него внутреннее мнение r . | $\frac{a}{N}(N - r_e - r_i + r)c\frac{r_i}{N}\rho_g + \frac{N-a}{N}(N - r_e - r_i + r)c\frac{r_i}{N^2}$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r)$ | $Br \rightarrow Bb$, авторитет во внутреннем слое имеет мнение Br , он принимает своё внешнее мнение или копирует мнение соседа с внутренним мнением b ; другие индивиды имеют мнение Br , один из них принимает своё внешнее мнение или копирует авторитета, если у него внутреннее мнение b . | $\frac{a}{N}(r_i - r)(i + c\frac{N-r_i}{N}\rho_g) + \frac{N-a}{N}(r_i - r)(i + c\frac{N-r_i}{N^2})$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r - \frac{1}{N})$ | $Rr \rightarrow Rb$, авторитет во внутреннем слое имеет мнение Rr , он копирует мнение соседа с внутренним мнением b ; другие индивиды имеют мнение Rr , один из них копирует авторитета, если у него внутреннее мнение b . | $r\frac{a}{N}c\frac{N-r_i}{N}\rho_g + r\frac{(N-a)}{N}c\frac{N-r_i}{N^2}$ |



(a) Нечётное количество узлов

(b) Чётное количество узлов

Рис. 6: Представление внутренней двухкликковой сети

Рисунок 6а показывает пример двухкликковой сети, когда две клики делят общий узел «Sarah». Степень узла «Sarah» составляет $N-1$. Другие узлы имеют степень $\frac{N-1}{2}$, и $\rho_g = \frac{N+1}{2N}$ в этом случае. Переходы для этого типа системы GCVM приведены в таблице 6.

Рисунок 6б показывает простой пример второго случая двухкликковой внутренней сети, когда в каждой клике есть специальный индивид, через которого клики соединены, то есть существует ребро, соединяющее двух специальных индивидов, которое называется межкликковым ребром. Во втором случае $\rho_g = \frac{N+2}{2N}$. Переходы для этой системы GCVM приведены в таблице 7.

1.2.4 Асимметричный случай: Неполная внешняя сеть

В этом разделе предполагается, что не только внутренний, но и внешний слой могут быть представлены неполным графом. Эти изменения во внешнем слое будут влиять на четыре перехода: $Bb \rightarrow Rb$, $Br \rightarrow Rr$, $Rb \rightarrow Bb$, и $Rr \rightarrow Br$.

Таблица 6: Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GCVM с внутренними двумя кликовыми сетями (нечётный случай)

| Новое состояние (x, y, z) | Как можно достичь нового состояния? | Элемент матрицы переходов Q_{GCVM} за единицу времени |
|--|--|---|
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r + \frac{1}{N})$ | $Rb \rightarrow Rr$, общий узел внутреннего слоя с мнением Rb принимает своё внешнее мнение или копирует мнение соседа с внутренним мнением r ; другие индивиды имеют мнение Rb , один из них принимает своё внешнее мнение или копирует мнение соседа. | $\frac{1}{N}(r_e - r)(i + c\frac{r_i}{N}) + \frac{N-1}{N}(r_e - r)(c\frac{r_i}{N}\rho_g + i)$ |
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r)$ | $Bb \rightarrow Br$, общий узел внутреннего слоя имеет мнение Bb и копирует мнение соседа с внутренним мнением r ; другие индивиды имеют мнение Bb , один из них копирует мнение соседа с внутренним мнением r . | $\frac{1}{N}(N - r_e - r_i + r)c\frac{r_i}{N} + \frac{N-1}{N}(N - r_e - r_i + r)c\frac{r_i}{N}\rho_g$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r)$ | $Br \rightarrow Bb$, общий узел внутреннего слоя имеет мнение Br , принимает своё внешнее мнение или копирует мнение соседа с внутренним мнением b ; другие индивиды имеют мнение Br , один из них принимает своё внешнее мнение или копирует мнение соседа с мнением b . | $\frac{1}{N}(r_i - r)(i + c\frac{N-r_i}{N}) + \frac{N-1}{N}(r_i - r)(i + c\frac{N-r_i}{N}\rho_g)$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r - \frac{1}{N})$ | $Rr \rightarrow Rb$, общий узел внутреннего слоя имеет мнение Rr , копирует мнение соседа с внутренним мнением b ; другие индивиды имеют мнение Rr , один из них копирует мнение соседа с внутренним мнением b . | $r\frac{1}{N}c\frac{N-r_i}{N} + r\frac{(N-1)}{N}c\frac{N-r_i}{N}\rho_g$ |

Таблица 7: Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GCVМ с двухкликвым внутренним слоем (чётный случай)

| Новое состояние (x, y, z) | Как достигается новое состояние? | Элемент матрицы переходов $Q_{GCVМ}$ за единицу времени |
|--|--|--|
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r + \frac{1}{N})$ | $Rb \rightarrow Rr$, общие узлы внутреннего слоя имеют мнение Rb , один из них принимает своё внешнее мнение или копирует мнение соседа с внутренним мнением r ; другие индивиды имеют мнение Rb , один из них принимает своё внешнее мнение или копирует мнение соседа. | $\frac{a}{N}(r_e - r)(i + c\frac{r_i}{a}) + \frac{N-a}{N}(r_e - r)(c\frac{r_i}{aN} + i)$ |
| $(\rho_{r_e}, \rho_{r_i} + \frac{1}{N}, \rho_r)$ | $Bb \rightarrow Br$, общие узлы внутреннего слоя имеют мнение Bb , один из них копирует мнение соседа с внутренним мнением r ; другие индивиды имеют мнение Bb , один из них копирует мнение соседа с внутренним мнением r . | $\frac{a}{N}(N - r_e - r_i + r)c\frac{r_i}{N}\rho_g + \frac{N-a}{N}(N - r_e - r_i + r)c\frac{r_i}{aN}$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r)$ | $Br \rightarrow Bb$, общие узлы внутреннего слоя имеют мнение Br , один из них принимает своё внешнее мнение или копирует мнение соседа с внутренним мнением b ; другие индивиды имеют мнение Br , один из них принимает своё внешнее мнение или копирует мнение соседа с мнением b . | $\frac{a}{N}(r_i - r)(i + c\frac{N-r_i}{a}) + \frac{N-a}{N}(r_i - r)(i + c\frac{N-r_i}{aN})$ |
| $(\rho_{r_e}, \rho_{r_i} - \frac{1}{N}, \rho_r - \frac{1}{N})$ | $Rr \rightarrow Rb$, общие узлы внутреннего слоя имеют мнение Rr , один из них копирует мнение соседа с внутренним мнением b ; другие индивиды имеют мнение Rr , один из них копирует мнение соседа с внутренним мнением b . | $r\frac{a}{N}c\frac{N-r_i}{N}\rho_g + r\frac{(N-a)}{N}c\frac{N-r_i}{aN}$ |

Внешний слой: сеть–цикл

В этом разделе предполагается, что внешний слой представлен циклом. Далее показываем, как изменения структуры внутреннего слоя влияют на переходы. Также рассматриваются внутренние слои, представленные сетями с двумя звёздами и двумя кликами. Представления таких двухслойных сетей для GCVM показаны на рисунках 7, 8, 9.

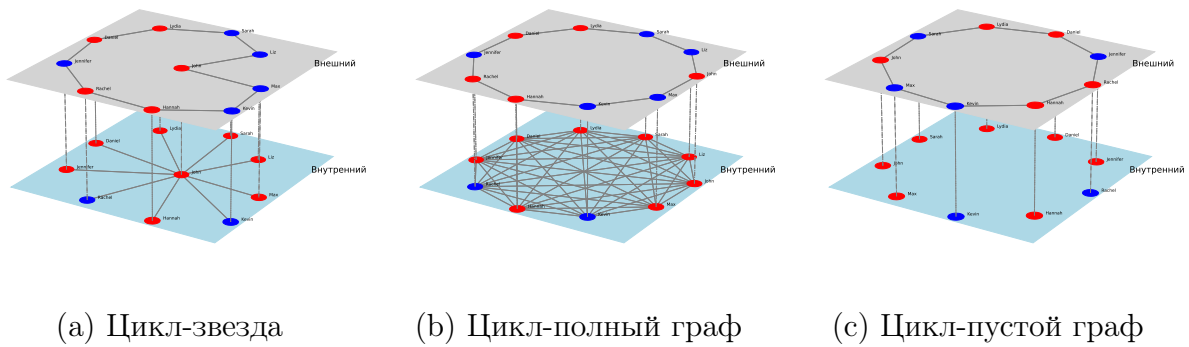


Рис. 7: Представление внешней сети–цикла

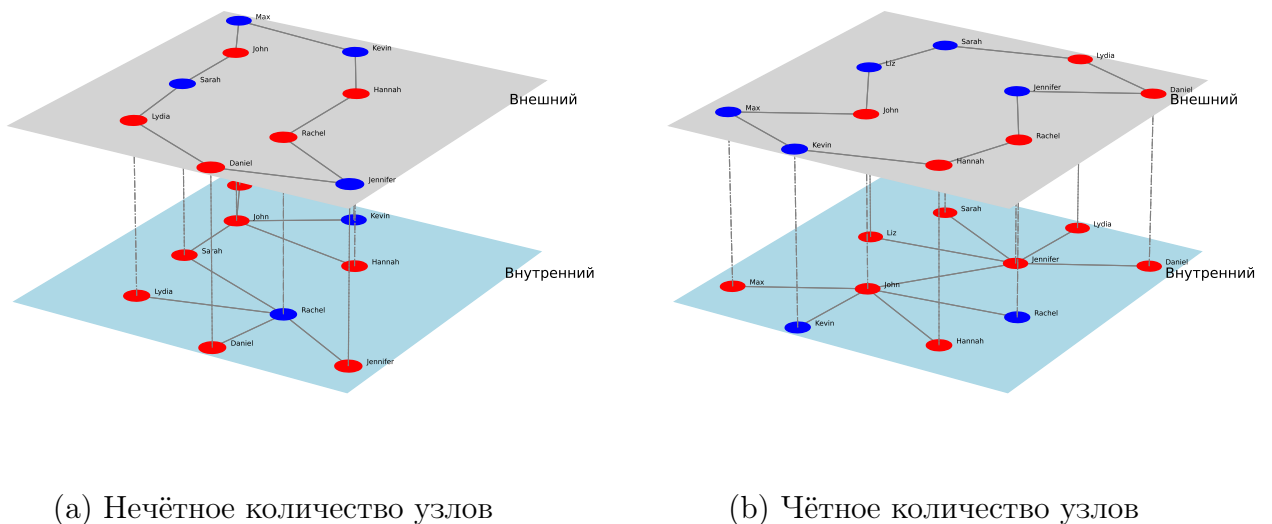


Рис. 8: Представление внешней сети–цикл и внутренней двухзвёздной сети

1. $Bb \rightarrow Rb$: Один индивид в состоянии Bb копирует мнение соседа с внешним мнением R . Скорость изменения: $\frac{c(N-r_e-r_i+r)r_e}{N} \frac{3}{N}$;

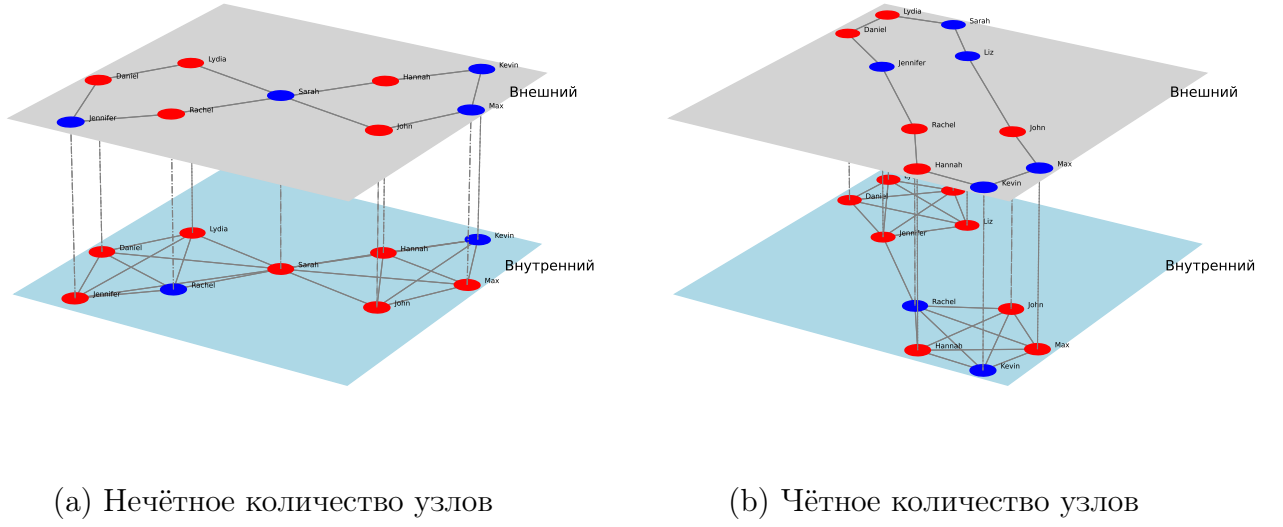


Рис. 9: Представление внешней циклической и внутренней двухкликковой сети

2. $Br \rightarrow Rr$: Один индивид в состоянии Br копирует мнение соседа с внешним мнением R или выражает своё внутреннее мнение. Скорость изменения: $(r_i - r)(c \frac{r_e}{N} \frac{3}{N} + e)$;
3. $Rb \rightarrow Bb$: Один индивид в состоянии Rb копирует мнение соседа с внешним мнением B или выражает своё внутреннее мнение. Скорость изменения: $(r_e - r)(c \frac{N-r_e}{N} \frac{3}{N} + e)$;
4. $Rr \rightarrow Br$: Один индивид в состоянии Rr копирует мнение соседа с внешним мнением B . Скорость изменения: $\frac{cr(N-r_e)}{N} \frac{3}{N}$.

1.3 Эксперименты и результаты

Для моделирования процесса изменения мнений в двухслойной сети воспользуемся методом Монте-Карло. Симуляции организованы следующим образом:

1. Установить случайное зерно, количество прогонов (обозначено как n) и инициализировать модель, задав определённые параметры;

2. Проверить, достигнут ли консенсус для данной модели (т. е. существует ли уникальное мнение в обоих слоях сети), и если да, остановить симуляцию, иначе перейти к шагу 3;
3. Рассчитать скорость перехода для каждого возможного изменения состояния при текущем состоянии;
4. Определить следующее состояние, предполагая, что время между двумя переходами распределено экспоненциально с текущей скоростью изменения;
5. Перейти к шагу 2.

Будем рассматривать время достижения консенсуса и вероятность победы определенного (красного) мнения как ключевые показатели эффективности (КПИ) и исследуем их. Численный эксперимент организован следующим образом:

1. Запускаем симуляции, изменяя параметры, указанные ниже. В частности, фокусируемся на силе красного мнения для каждой симуляции. В этой части можно наблюдать, как сила красного мнения влияет на КПИ. Параметры, использованные в симуляциях, следующие:
 - **Набор 1:** $\rho = 0.75, N = 400, r = 300, c = 1$ для BVM, и $m \approx 0.74, N = 400, r_e = 300, r_i = 100, r = 80, e = 0.01, i = 0.50, c = 1$ для GCVM;
 - **Набор 2:** $\rho = 0.80, N = 500, r = 400, c = 1$ для BVM, и $m \approx 0.79, N = 500, r_e = 400, r_i = 100, r = 80, e = 0.01, i = 0.80, c = 1$ для GCVM;
2. Изменяем структуру внутреннего слоя и показываем, как структура внутренней сети влияет на КПИ;
3. Модифицируем структуру внешнего слоя и наблюдаем влияние различных внешних сетевых структур на КПИ.

Остальная часть этого раздела демонстрирует самые интересные результаты экспериментов.

Рисунок 10 показывает время достижения консенсуса для первых 20 симуляций. Можно увидеть, что время достижения консенсуса для BVM, SVM и

GCVM с полной внешней сетевой структурой находится в одном порядке величины.

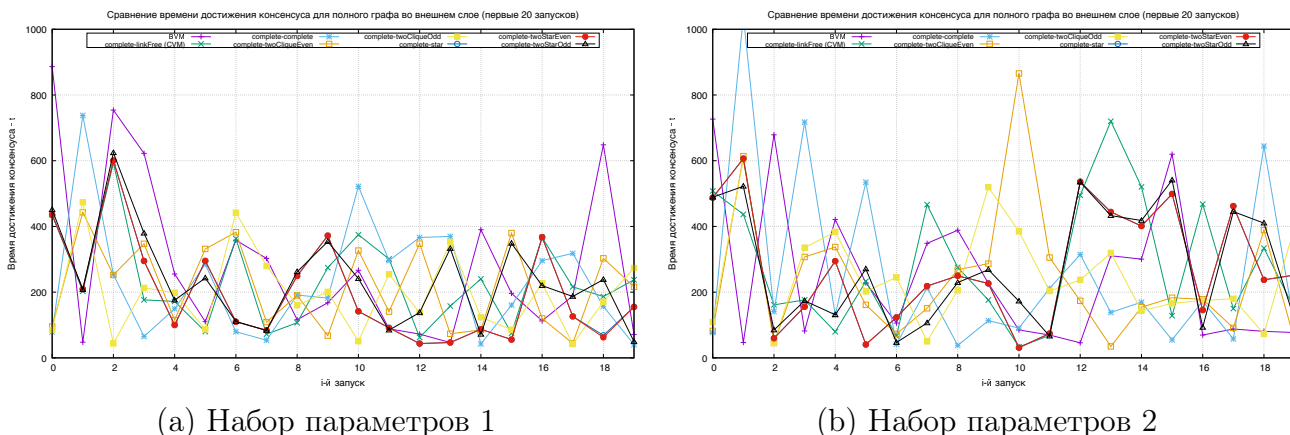


Рис. 10: Время достижения консенсуса для 20 экспериментов, сравнение BVM, SVM и GCVM с полной внешней сетевой структурой

Можно рассмотреть GCVM с неполной внешней структурой и сравнить результаты с ранее описанными моделями. Результаты обобщены на рисунке 11. Подчеркнем неожиданный результат: во внешнем слое структура цикл, которая «намного проще», чем полная структура, но она значительно увеличивает время достижения консенсуса. Заметим разницу между левым и правым рисунками: на левом рисунке отсутствует модель «cycle-complete». Мы не включили эту модель, потому что в наших симуляциях консенсус не достигался при наборе параметров 1. Даже при наборе параметров 2 время достижения консенсуса для структуры «cycle-complete» значительно больше, чем для других моделей.

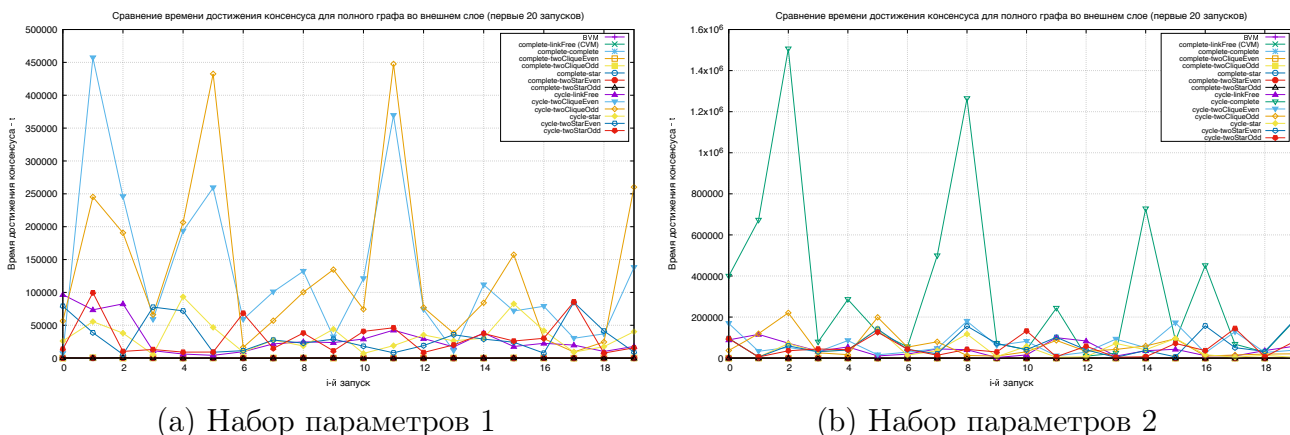


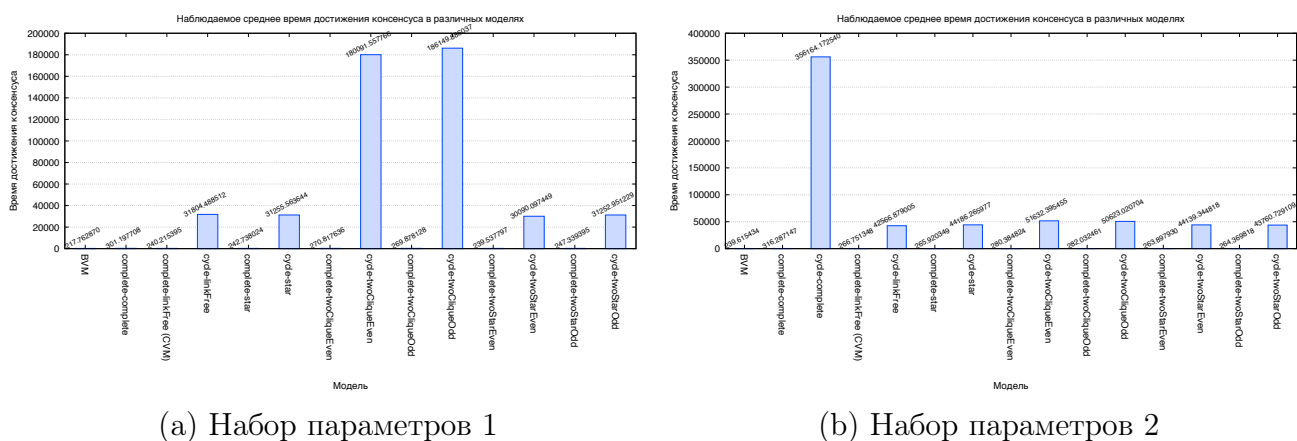
Рис. 11: Время достижения консенсуса для 20 экспериментов, сравнение всех моделей

Рисунок 12 показывает наблюдаемое среднее время достижения консенсуса для всех моделей. Несложно сделать тот же вывод, что и на основе рисунка

ка 11: цикл, структура сети во внешнем слое, значительно увеличивает время достижения консенсуса.

Если сравнить наблюдаемое среднее время достижения консенсуса между BVM, SVM и GCVM с полной внешней структурой, можно сделать следующие выводы:

1. Многослойная сетевая структура общества увеличивает время достижения консенсуса по сравнению с BVM, в которой существует единственный слой коммуникации;
2. Не все внутренние структуры увеличивают время достижения консенсуса в SVM.



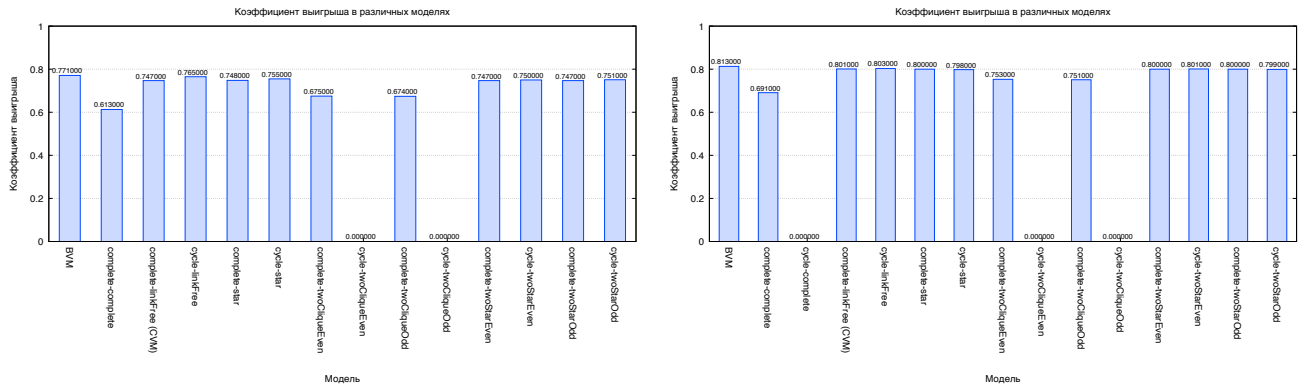
(a) Набор параметров 1

(b) Набор параметров 2

Рис. 12: Наблюдаемое среднее время достижения консенсуса для всех моделей

Рассмотрим рисунок 13, основанный на 1000 симуляциях для каждой модели. Как показано на графике, BVM имеет самый высокий уровень победы для обоих наборов параметров. Случай «cycle-complete» (внешний слой — цикл, внутренний — полный граф) интересен, потому что для набора параметров 1 консенсус не достигнут в наших численных симуляциях, поэтому эта структура пропущена на рисунке 12а. В случае «cycle-complete» с набором параметров 2 консенсус достигается, но время достижения консенсуса чрезвычайно велико по сравнению с другими структурами слоёв. Отметим интересное наблюдение: для внутреннего слоя сложной структуры (полный граф, две клики) внешняя циклическая структура уменьшает вероятность победы красного мнения, которая приближается к нулю, хотя сила начального красного мнения, заданная нашими наборами параметров, очень высока. Это можно объяснить следующим образом: когда структура сети внешнего слоя является циклической, чем ниже

средняя степень узлов, тем выше время достижения консенсуса и ниже вероятность победы. Если сравнить рисунок 13а с рисунком 13б, можно заметить более высокую силу красного мнения и более высокую вероятность его победы. Следовательно, наш вывод представляется разумным.



(а) Набор параметров 1

(б) Набор параметров 2

Рис. 13: Вероятность победы красного мнения для всех моделей

1.4 Заключение к Главе 1

В данной главе формализуется макроскопическая общая скрытая модель избирателя (macroscopic-GCVM), являющаяся расширением скрытой модели избирателя. Используется двухслойная сетевая структура для моделирования социальной сети: внешний слой представляет публичную коммуникационную сеть, а внутренний слой — частную коммуникационную сеть. Если индивиды выражают разные мнения во внутреннем и внешнем слоях, они могут изменить свои мнения через экстернализацию или интернализацию. Моделируется процесс изменения мнений для различных сетевых структур (например, полная, циклическая, звезда, двухзвёздная (чётное и нечётное количество узлов), двухкликерная (чётное и нечётное количество узлов) для внутреннего слоя и полная, циклическая — для внешнего слоя) с различными параметрами, которые представляют разные модели поведения индивидов, и сравниваем результаты с базовой моделью избирателя и скрытой моделью избирателя. Основные выводы, сформулированные на основе численных симуляций, включая некоторые контринтуитивные результаты, следующие:

1. Для некоторых симуляций простая структура внешнего слоя, такая как цикл, создаёт проблему для достижения консенсуса;
2. Если индивиды в социальной сети плохо выражают свои мнения публично (имеют низкое значение параметра e), внутреннее взаимодействие не оказывает значительного влияния на консенсус (включая вероятность победы и время достижения консенсуса).

Глава 2. Анализ времени достижения консенсуса и вероятности победы мнения в двухслойных сетях различных структур при наличии лицемерия

В этой главе представлено определение двухслойных сетей с узлами-репликами. Здесь мы опишем микроскопическую модель скрытого избирателя (microscopic-GCVM), которая является более общей по сравнению с макроскопической моделью скрытого избирателя (macroscopic-GCVM), представленной в главе 1, при допущении, что взаимодействия между узлами во внутреннем слое разрешены. Микроскопическая модель GCVM подходит для любой заданной структуры сети, удовлетворяющей данному определению. Рассматривается влияние свойств сети и поведения индивидов на время достижения консенсуса и вероятность победы. Результаты этой главы опубликованы в [25, 27].

2.1 Многослойная сеть с узлами-репликами

Многослойная сеть — сеть, образованная несколькими слоями-сетями, узлы в которой взаимодействуют друг с другом [56].

В многослойной сети существует взаимно-однозначное соответствие узлов в разных слоях, и соответствующие узлы называются узлами-репликами. Поскольку существует взаимно-однозначное соответствие между узлами в разных слоях, каждый слой состоит из одинакового количества узлов [56].

2.2 Двухслойная сеть с узлами-репликами

Воспользуемся следующими обозначениями для определения двухслойной (внешней и внутренней) сети с узлами-репликами:

1. N : количество узлов в каждом слое;
2. $a_i = (a_i^E, a_i^I)$: взаимно-однозначное соответствие узла i во внешнем и внутреннем слоях, где a_i^E (a_i^I) является представлением узла i во внешнем (внутреннем) слое;
3. $G_E(\mathcal{V}_E, \mathcal{E}_E)$: заданная внешняя сеть, где $\mathcal{V}_E = \{a_i^E\}$ и \mathcal{E}_E — множество индивидов и множество рёбер во внешнем слое соответственно;
4. $G_I(\mathcal{V}_I, \mathcal{E}_I)$: заданная внутренняя сеть, где $\mathcal{V}_I = \{a_i^I\}$ и \mathcal{E}_I — множество индивидов и множество рёбер во внутреннем слое соответственно;
5. $\mathcal{E}_C = \{(a_i^E, a_i^I) | i = 1, \dots, N\}$: рёбра между узлами-репликами.

Двухслойная сеть с N индивидами/агентами может быть определена как¹:

$$G(\mathcal{V}, \mathcal{E}), \quad (7)$$

где $\mathcal{V} = \mathcal{V}_E \cup \mathcal{V}_I$, $|\mathcal{V}_E| = |\mathcal{V}_I| = N$, а $\mathcal{E} = \mathcal{E}_E \cup \mathcal{E}_I \cup \mathcal{E}_C$. Это определение не зависит от конкретной структуры сети, то есть внешние/внутренние сети могут быть разными по структуре.

¹Это определение и соответствующие модели динамики мнений были впервые введены в [27] и далее обсуждены в [25] и [26].

2.3 Модель

2.3.1 Общая скрытая модель избирателя (макроверсия)

Чжао и Парилина [24] предложили GCVM на основе SVM, введённой в [21]. Эти работы используют симуляции для представления процессов передачи мнений в двухслойных сетях. В следующем разделе представлена микроверсия GCVM.

2.3.2 Общая скрытая модель избирателя (микроверсия)

В GCVM мы используем R, B (r, b) для обозначения внешних (внутренних) красных и синих мнений индивидов соответственно. Список обозначений:

- $S = \{Rr, Rb, Br, Bb\}$: множество всех возможных состояний индивида;
- $\omega(a_i, t) \in S$: мнение индивида a_i в момент времени t , где $i = 1, \dots, N$, и $t = 0, 1, \dots$;
- ρ_{r_e} : доля индивидов с красным мнением во внешнем слое;
- ρ_{r_i} : доля индивидов с красным мнением во внутреннем слое;
- ρ_r : доля индивидов с красным мнением в обоих слоях — внутреннем и внешнем;
- r_e : количество индивидов с красным мнением во внешнем слое;
- r_i : количество индивидов с красным мнением во внутреннем слое;
- r : количество индивидов с красным мнением в обоих слоях;
- π_{c_e} : скорость копирования во внешнем слое, то есть того, что индивид скопирует мнение своего внешнего соседа;

- π_{c_i} : скорость копирования во внутреннем слое, то есть того, что индивид скопирует мнение своего внутреннего соседа;
- π_e : скорость экстернализации, то есть того, что лицемер² выберет публичное выражение своего внутреннего мнения;
- π_i : скорость интернализации, то есть того, что лицемер примет своё внешнее мнение.

В разделах 2.3.2–2.3.2 описана GCVM динамики мнений в двухслойной сети. Описание организовано так, чтобы понять, как были проведены численные симуляции, представленные в разделе 2.4.

Инициализация структуры двухслойной сети

Начинаем с задания двух сетей G_E и G_I (эти сети загружаются из файла, пример такого файла, представляющего внешнюю циклическую и внутреннюю звёздную структуры, показан в листинге 1. Затем мы добавляем рёбра между внешними и внутренними вершинами, представляющими индивидов).

Листинг 1: Пример файла графа с внешней циклической и внутренней звездной структурами

```

>>>external
E0 E1
E1 E2
E2 E3
E3 E0
<<<external end
>>>internal
I00 I1
I00 I2
I00 I3
<<<internal end

```

Это приводит к формированию двухслойной сети G , которую мы храним в виде графа смежности.

²Под лицемерием мы понимаем узел, который имеет разные мнения во внешнем и внутреннем слоях, т. е. узлы в состояниях Rb и Bv .

Инициализация начальных состояний индивидов

Связь между количеством индивидов в каждом состоянии представлена в уравнении (3) главы 1. Предполагая равномерное распределение вероятности для каждого агента оказаться в любом состоянии $s \in S$ в начальный момент времени, используем следующее правило для задания начального состояния $\omega(a_i, 0)$ для любого агента a_i в момент времени $t = 0$:

$$\omega(a_i, 0) := f(x) = \begin{cases} Rr, & 0 \leq x < \rho_r, \\ Rb, & \rho_r \leq x < \rho_{r_e}, \\ Br, & \rho_{r_e} \leq x < \rho_{r_e} + \rho_{r_i} - \rho_r, \\ Bb, & \rho_{r_e} + \rho_{r_i} - \rho_r \leq x \leq 1, \end{cases} \quad (8)$$

где $x \sim U(0, 1)$.

Процесс передачи мнений

Можно разделить индивидов на лицемеров и нелицемеров в зависимости от согласованности их внешних и внутренних мнений. Лицемеры — это индивиды, которые имеют разные мнения во внутренних и внешних слоях, а нелицемеры — те, у кого одинаковые мнения в обоих слоях.

Мы фокусируемся на двух показателях при анализе GCVM:

- *время достижения консенсуса*: T_{cons} — время достижения консенсуса в (G)CVM, то есть время, необходимое для того, чтобы все индивиды сформировали одно и то же мнение во внутренних и внешних слоях (т. е. $\rho_{r_e} = \rho_{r_i} = \rho_r = 0$ или 1 для T_{cons});
- *вероятность победы*: ρ — вероятность победы красного мнения в серии симуляций. Для мнения «победить» означает, что в сети больше нет другого мнения (т. е. количество симуляций, при которых красное мнение побеждает синее, делённое на общее количество симуляций).

Прежде чем представить алгоритм GCVM, кратко определим действия, доступные для случайно выбранного индивида a_i :

- *выбор соседа a_j* : Случайным образом выбрать соседа среди всех соседей a_i . Пусть это будет индивид a_j (это предварительное действие для внешнего/внутреннего копирования);

- *внешнее копирование*: a_i копирует внешнее мнение a_j с вероятностью π_{c_e} ;
- *внутреннее копирование*: a_i копирует внутреннее мнение a_j с вероятностью π_{c_i} ;
- *экстернализация*: a_i выражает своё внутреннее мнение с вероятностью π_e (это действие доступно только для лицемеров);
- *интернализация*: a_i принимает своё внешнее мнение с вероятностью π_i (это действие доступно только для лицемеров).

Экстернализация и интернализация не имеют смысла для нелицемеров, поэтому у них есть только два возможных действия (внешнее и внутреннее копирование).

Алгоритм GCVM:

Шаг 1. Инициализировать $t = 0$.

Шаг 2. Выбрать индивида a_i случайным образом из N индивидов в двухслойной сети G ;

Шаг 3. Проверить все допустимые действия индивида a_i (в зависимости от его состояния) и случайным образом выбрать одно из допустимых действий с равными вероятностями:

I) если a_i является лицемером, тогда у него/неё есть четыре возможных действия: (i) внешнее копирование, (ii) внутреннее копирование, (iii) экстернализация и (iv) интернализация. Любое действие выбирается с вероятностью 0.25;

II) если a_i не является лицемером, тогда у него/неё есть только два возможных действия: внешнее или внутреннее копирование. Любое действие выбирается с вероятностью 0.5.

Шаг 4. Сгенерировать случайное число $x \sim U(0, 1)$. Выполнить действие, выбранное на шаге 3:

- а) если на шаге 3 выбрано внешнее копирование и $x < \pi_{c_e}$, тогда a_i копирует внешнее мнение a_j ;
- б) если на шаге 3 выбрано внутреннее копирование и $x < \pi_{c_i}$, тогда a_i копирует внутреннее мнение a_j ;

- с) если на шаге 3 выбрана экстернализация и $x < \pi_e$, тогда a_i выражает своё внутреннее мнение;
- д) если на шаге 3 выбрана интернализация и $x < \pi_i$, тогда a_i принимает своё внешнее мнение.

Шаг 5. Увеличить t на 1. Если достигнут консенсус³, прекратить итерацию. В противном случае вернуться к шагу 2.

2.4 Эксперименты и результаты

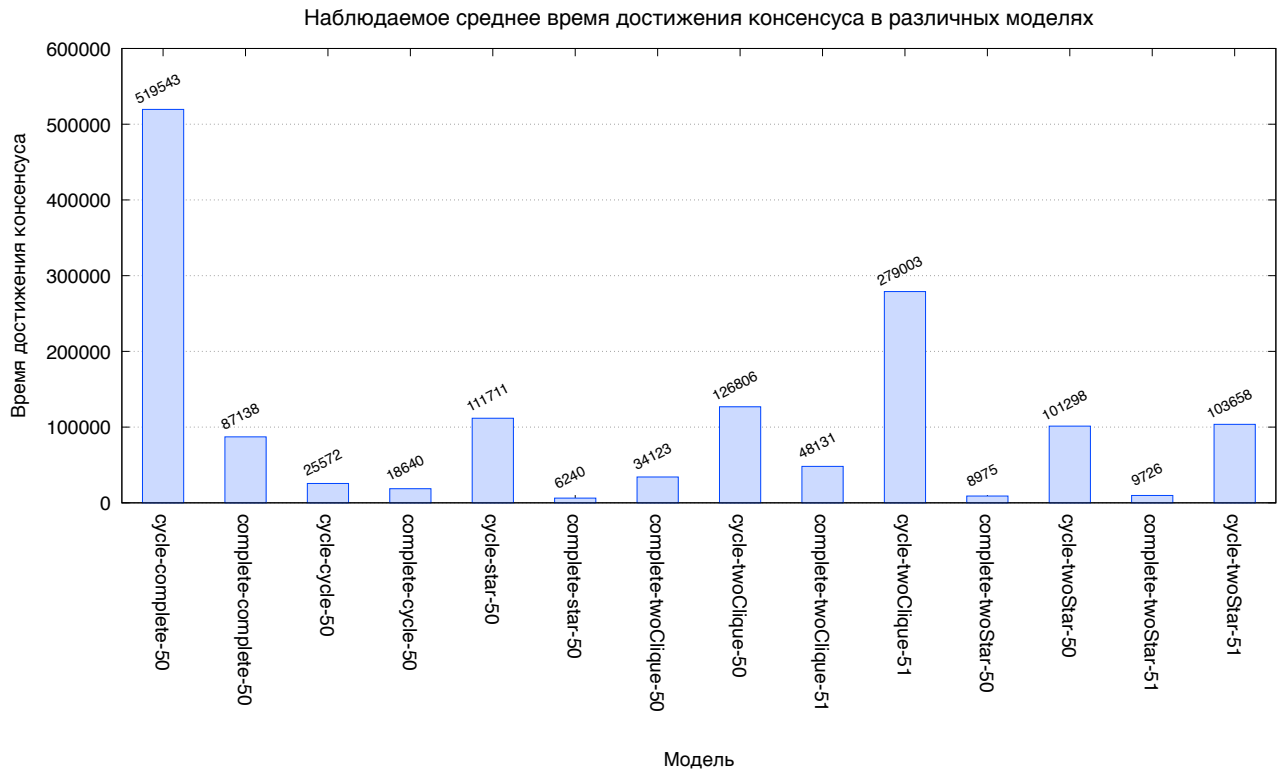
2.4.1 Общее описание эксперимента

Эксперимент сосредоточен на наблюдении влияния структуры внешней сети на вероятность победы мнения и время достижения консенсуса.

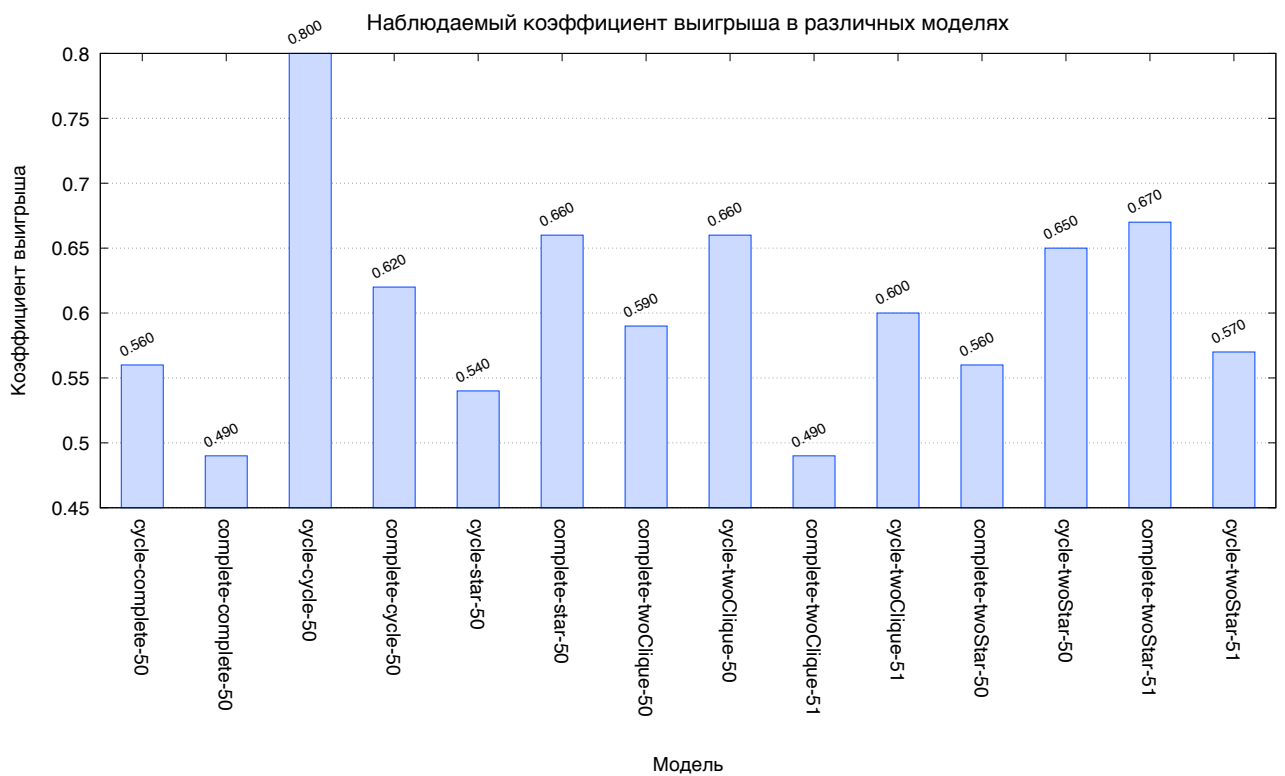
Были рассмотрены два типа внешних структур: цикл и полная сеть, а также семь внутренних структур: цикл, полная сеть, звезда, двухзвёздная сеть (чётный и нечётный случаи) и двухкликровая сеть (чётный и нечётный случаи). Это даёт нам 14 различных комбинаций внешних и внутренних структур, т. е. 14 двухслойных сетей. Большинство структур показано на рисунках с 3 по 9 главы 1, мы не приводим их здесь для экономии места.

Для наших экспериментов зафиксируем следующие параметры: $\rho_{r_e} = 0.75$, $\rho_{r_i} = 0.25$, $\rho_r = 0.2$, $\pi_{c_i} = \pi_{c_e} = 1$, $\pi_e = 0.01$, $\pi_i = 0.5$, $N = 100$. Затем наблюдаем за влиянием структуры внешнего слоя на вероятность победы мнения и время достижения консенсуса (количество итераций) для заданной внутренней структуры. Для каждой модели проводим 100 симуляций и получаем следующие статистические результаты. Модели, представленные на рисунках 14а и 14б, названы как: «внешний слой — внутренний слой — количество индивидов».

³Алгоритм останавливается, когда все индивиды в обоих слоях придерживаются одного и того же мнения, т.е. достигнут консенсус.



(a) Время достижения консенсуса



(b) Вероятность победы

Рис. 14: Наблюдаемое время достижения консенсуса (a) и вероятность победы (b) для моделей с разными структурами двухслойных сетей

Как показано⁴ на рисунке 14а, мы приходим к выводу, что циклические внешние структуры увеличивают время достижения консенсуса по сравнению с полной сетью, определяющей внешнюю структуру (аналогичный результат получен в [24]). Этот вывод справедлив для всех внутренних структур, которые мы исследовали в эксперименте. Кроме того, мы наблюдаем, что внутренняя структура также влияет на время достижения консенсуса. Например, время достижения консенсуса для внутренней структуры «two-star» меньше, чем для внутренней структуры «two-clique».

Однако результаты наблюдений за вероятностью победы существенно отличаются от наших предыдущих работ [24], как показано на рисунке 14b. Можно заметить, что за исключением внутренних структур «two-star-51» и «star-50», циклический внешний слой снижает вероятность победы. Для всех других моделей циклическая внешняя структура положительно влияет на вероятность победы. Возможная причина заключается в спецификации микроскопической модели, то есть действия, которые может выполнять индивид/агент, связаны только с его текущим состоянием.⁵ В макроскопической версии GCVM вероятность каждого возможного действия зависит от общего состояния системы, представленного тройкой $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ (см. [24]).

2.4.2 Основные результаты и наблюдения

На основании выводов, сделанных в разделе 2.4.1, которые указывают на то, что циклическая структура положительно влияет на время достижения консенсуса, мы сформулировали следующие исследовательские вопросы. Серия экспериментов была разработана и проведена для того, чтобы ответить на эти вопросы:

1. Как внешняя структура влияет на ключевые показатели эффективности (KPI), т.е. есть ли особенности сети, которые значительно влияют на KPI?
Для ответов на этих вопросов мы делаем следующее:

⁴Число в названии модели обозначает количество индивидов. Например, 50 — чётный случай, а 51 — нечётный.

⁵Для нелицемеров у индивида есть два возможных действия, а для лицемеров — четыре. Вероятность каждого возможного действия индивида/агента в данный момент времени фиксирована.

- a) расширяем циклическую структуру до полной разными способами,⁶
 - b) сокращаем циклическую структуру до линии и наблюдаем за КРІ,
 - c) проводим статистические тесты, чтобы определить, есть ли значимые различия в распределении КРІ для различных способов построения полного графа (см. пункт а),
 - d) проверяем корреляцию между характеристиками сети и КРІ.
2. Как скорости экстернализации и интернализации влияют на КРІ?
 - a) изменяем π_e от 0.1 до 1.0 с шагом 0.1,
 - b) изменяем π_i от 0.3 до 1.0 с шагом 0.1 (замечание: при $\pi_i = 0.1$ или $\pi_i = 0.2$ в некоторых моделях консенсус не достигается).
 3. Как скорость копирования влияет на КРІ?
 - a) изменяем π_c от 0.1 до 1.0 для обоих слоёв с шагом 0.1.
 4. Как комбинация параметров влияет на КРІ (то есть какие комбинации максимизируют или минимизируют время достижения консенсуса и вероятность победы)?
 - a) используем комбинацию параметров $(\pi_i, \pi_e, \pi_{c_i}, \pi_{c_e})$, где π_i изменяется в диапазоне $[0.3, 1]$, а $\pi_e, \pi_{c_i}, \pi_{c_e}$ изменяются в диапазоне $[0.1, 1]$ с шагом 0.1.

Начинаем с расширения циклической структуры тремя различными способами:

- *нормальный*: найти множество рёбер, присутствующих в полном графе, но отсутствующих в циклическом графе, и добавлять их последовательно в циклический граф до тех пор, пока не получится полный граф;
- *случайный*: найти множество рёбер, присутствующих в полном графе, но отсутствующих в циклическом графе, и добавлять их случайным образом в циклический граф;
- *кратчайший*: найти множество рёбер, присутствующих в полном графе, но отсутствующих в циклическом графе, и добавлять рёбра из списка, которые минимизируют d , где d — средняя длина кратчайших путей между

⁶Циклический граф можно преобразовать в полный, добавив несколько несуществующих рёбер, или превратить в линию, удалив существующее ребро.

всеми парами узлов во внешнем слое. Здесь d рассчитывается следующим образом:

$$d = \sum_{s,t \in \mathcal{V}_E} \frac{d(s,t)}{n_E(n_E - 1)}, \quad (9)$$

где $d(s,t)$ — длина кратчайшего пути между узлами s и t , \mathcal{V}_E — множество узлов во внешнем слое, $n_E = |\mathcal{V}_E|$ — количество узлов во внешнем слое.

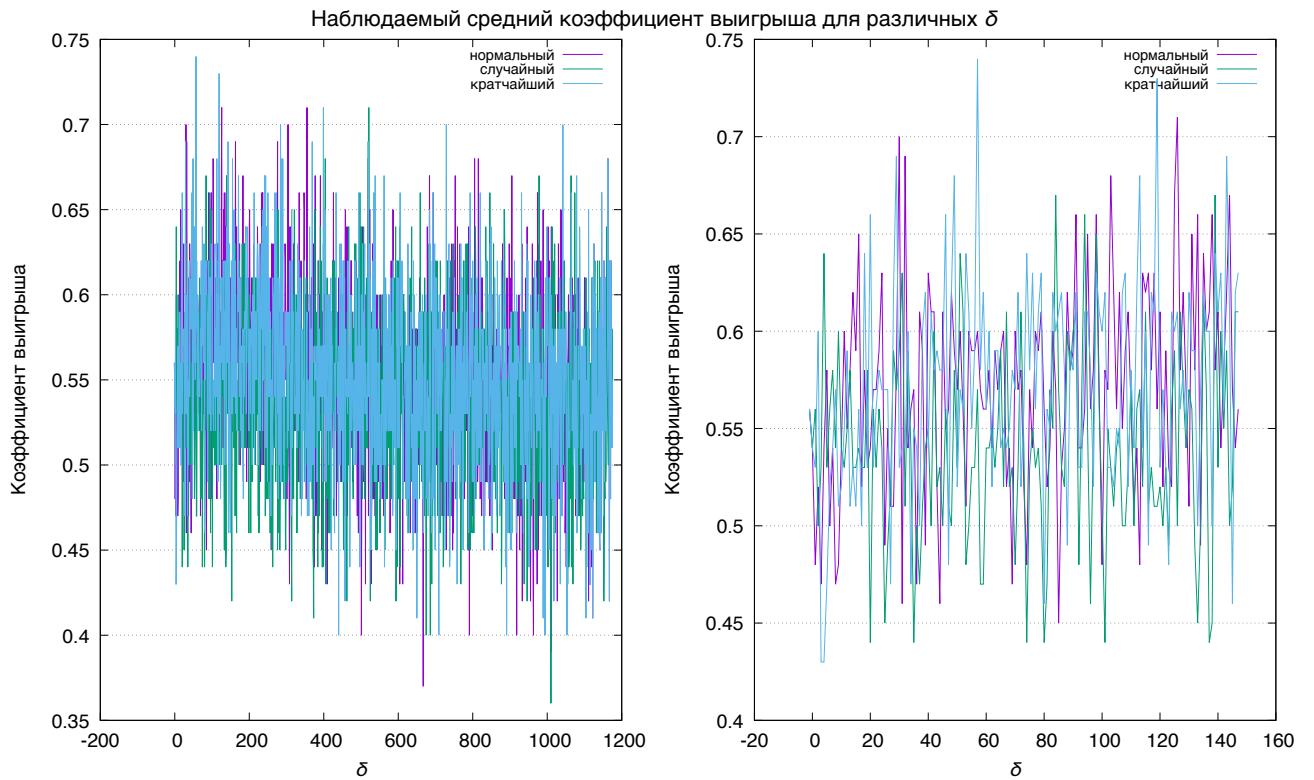
Очевидно, что для неориентированного графа с N узлами циклическая структура имеет только N рёбер, а полная структура — $N \times (N - 1)/2$ рёбер. Таким образом, циклический граф может стать полным графом, добавив $N \times (N - 3)/2$ рёбер. Если удалить одно ребро из циклической структуры, она превратится в линию.

В наших экспериментах мы расширили сетевую структуру от «cycle-complete-50» до «complete-complete-50», определив «cycle+ δ -complete-50» как промежуточную сетевую структуру, где $\delta \in \mathbb{Z}$ означает количество рёбер, которые были добавлены в циклическую структуру на каждой итерации. Здесь $\delta \in [-1, 1175]$, и значение $\delta = -1$ соответствует случаю, когда мы удалили одно ребро из цикла, превращая его в линию. Когда $\delta = 1175$, цикл становится полным графом, то есть максимальное количество рёбер, которые можно добавить в цикл с 50 узлами, равно 1175. В наших экспериментах мы исследовали динамику времени достижения консенсуса и d при увеличении δ .

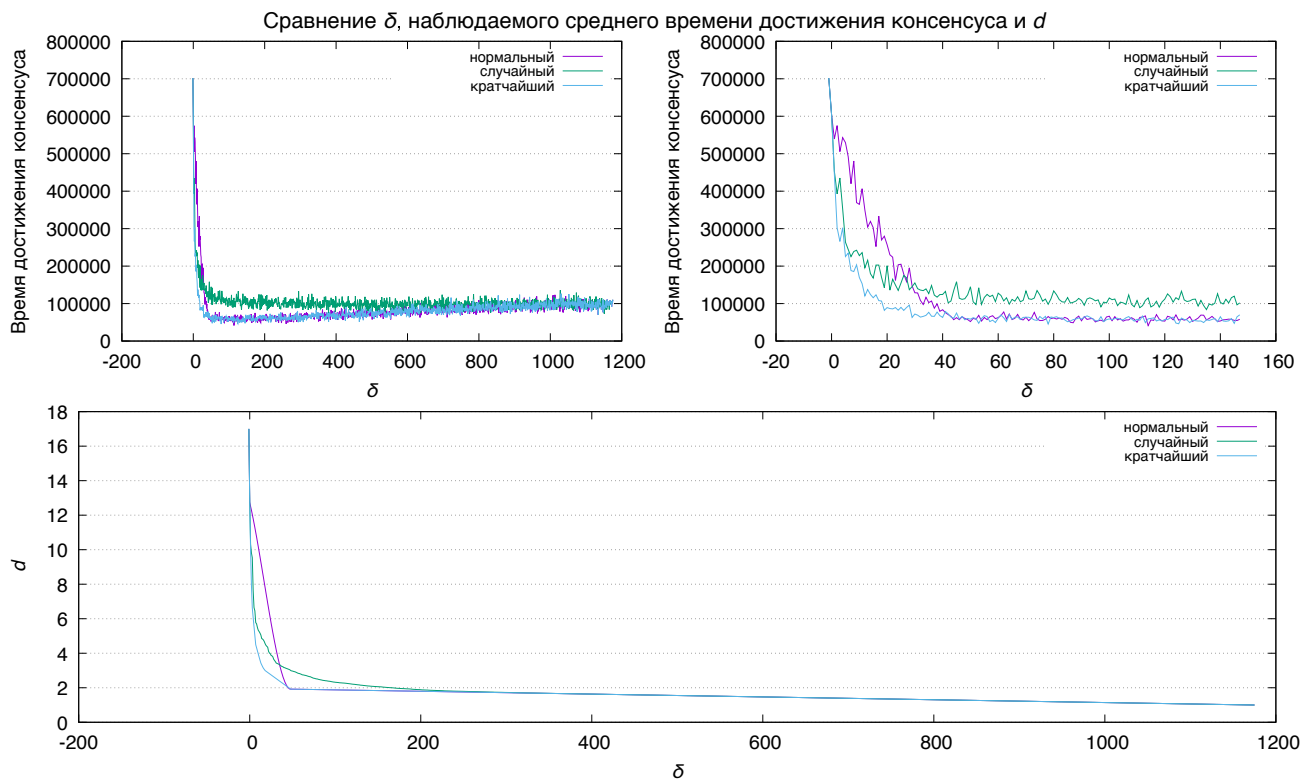
Результаты симуляций показаны на рисунке 15. В частности, на рисунке 15а показано, как d (δ) влияет на вероятность победы для различных способов расширения циклической структуры до полной. Правая часть графика — увеличение первых 150 точек с левого графика, аналогичное действие мы выполнили на рисунках 15b и 16. Если посмотреть на рисунок 15а, можно заметить, что вероятность победы практически представляет собой белый шум при изменении δ .⁷

На рисунке 15b показан тренд времени достижения консенсуса при изменении δ , и легко заметить, что фиолетовая и зелёная линии имеют почти одинаковый тренд. Вначале с увеличением числа добавленных рёбер время достижения консенсуса значительно уменьшается. После достижения определённого уровня увеличение числа рёбер не оказывает значительного влияния на время консенсуса. Первые точки на рисунке 15b соответствуют структуре сети

⁷Вероятность победы является белым шумом для «случайного» расширения с задержкой 1–10, для «кратчайшего» — с задержкой 1–3, но вероятность победы для «нормального» расширения не является белым шумом, что подтверждено тестом Льюнга–Бокса.



(a) Вероятность победы

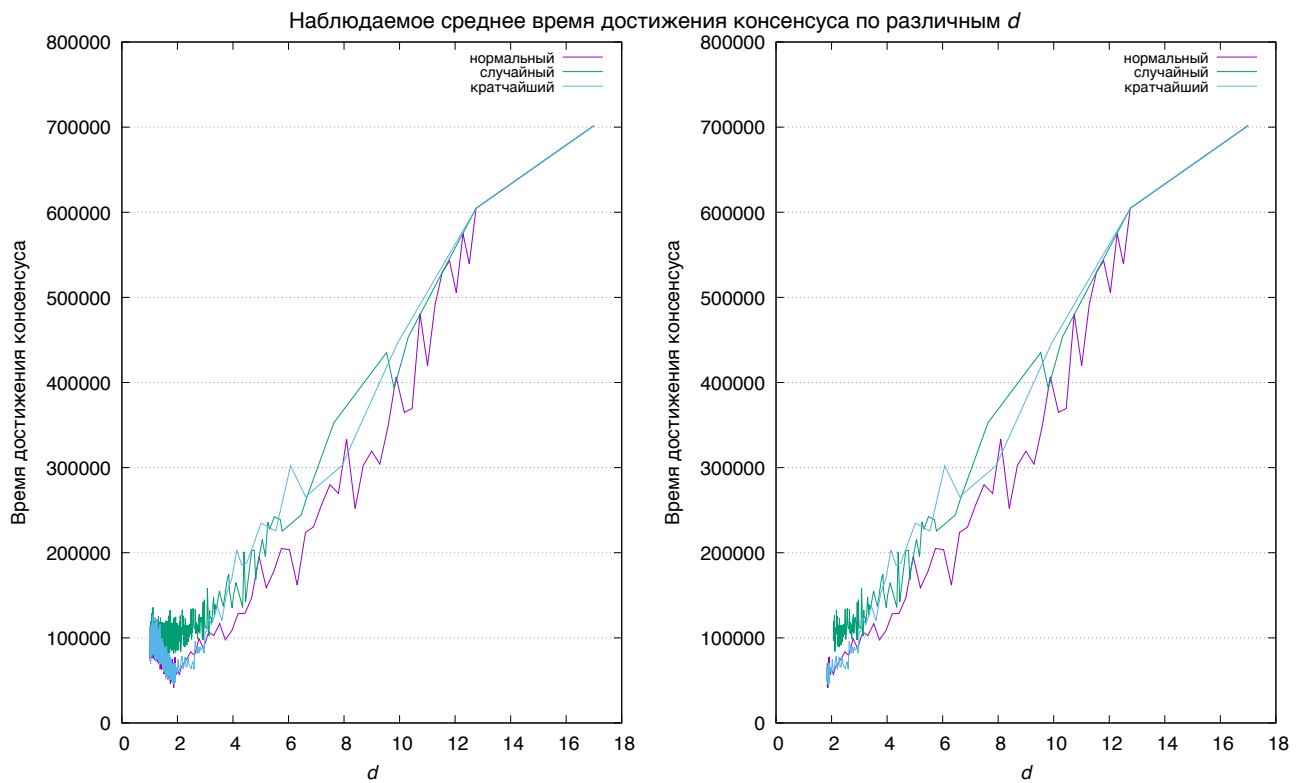


(b) Время достижения консенсуса

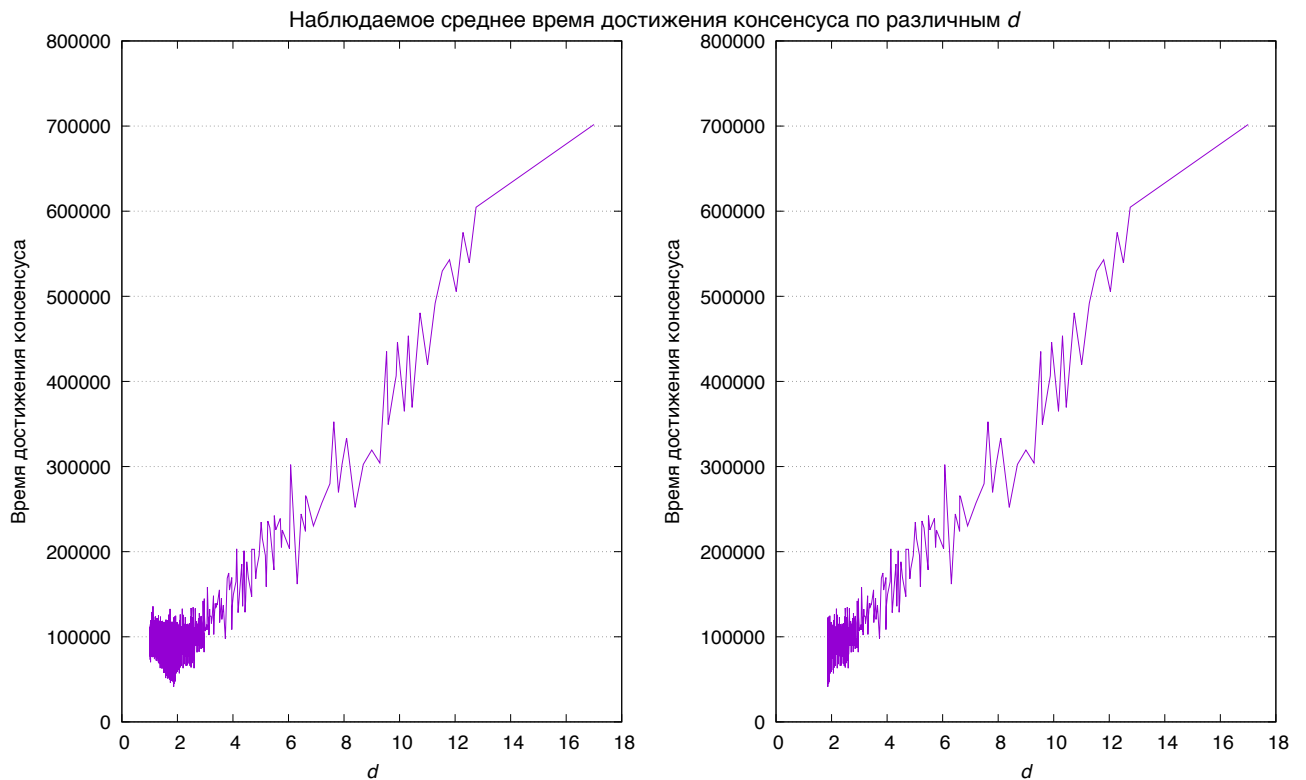
Рис. 15: Вероятность победы (a) и время достижения консенсуса (b) для моделей с различными способами расширения до полной сети

«cycle+1-complete-50», имеющей линейную структуру внешнего слоя. Очевидно, что $d_{\text{line}} > d_{\text{cycle}} > d_{\text{complete}}$. Поэтому мы формулируем гипотезу 1:

Гипотеза 1. Существует значимая корреляция между временем достижения консенсуса и d .



(а) Время достижения консенсуса в зависимости от d (для различных способов расширения)



(б) Время достижения консенсуса в зависимости от d (агрегированные данные)

Рис. 16: Время достижения консенсуса в зависимости от d

Для проверки гипотезы 1 необходимо вычислить d для каждого графа, представленного на рисунке 15b. После расчёта d для каждого графа мы строим третью подфигуру на рисунке 15b. Очевидно, что первая и третья подфигуры имеют схожий тренд.

На рисунке 16 представлена зависимость между d и временем достижения консенсуса. Разница между рисунками 16a и 16b заключается в том, что на рисунке 16a данные сначала сгруппированы по способу расширения графа, а затем отображён тренд, в то время как на рисунке 16b способ расширения графа не указывается, анализ проводится только на основе различий в d и времени достижения консенсуса.

Можно наблюдать приблизительно линейную зависимость на рисунке 16. Далее можем использовать статистические методы для анализа значимости корреляции [57, 58]. Результаты проверки значимости коэффициента корреляции Пирсона (РСС) [59] показаны на рисунке 17. Были сделаны следующие выводы:

- Корреляция между d и временем достижения консенсуса является значимой, коэффициент корреляции Пирсона равен 0.78;
- Для других пар КРІ корреляция не является значимой, абсолютные значения РСС менее 0.15.

Гипотеза 2. *Существуют значительные различия в распределениях КРІ для различных способов построения полного графа.*

Нас интересует, как разные способы расширения до полного графа влияют на распределение КРІ. Эмпирические распределения КРІ показаны на рисунке 18. Следует отметить, что распределения значительно различаются для разных способов расширения графа от цикла до полного. Однако для «нормального» и «кратчайшего» способов расширения распределения КРІ очень схожи. Для дальнейшего анализа мы используем тест однородности Колмогорова–Смирнова [60, 61]. Результаты приведены в таблице 8. Из таблицы 8 видно, что p -значения для всех КРІ при сравнении «нормального» и «кратчайшего» способов расширения превышают 0.05. Делаем следующий вывод: необходимо принять нулевую гипотезу о том, что распределения КРІ для «нормального» и «кратчайшего» способов расширения идентичны.

На данный момент установлено, что различные способы расширения цикла до полного графа влияют на распределение КРІ. Но насколько значим этот

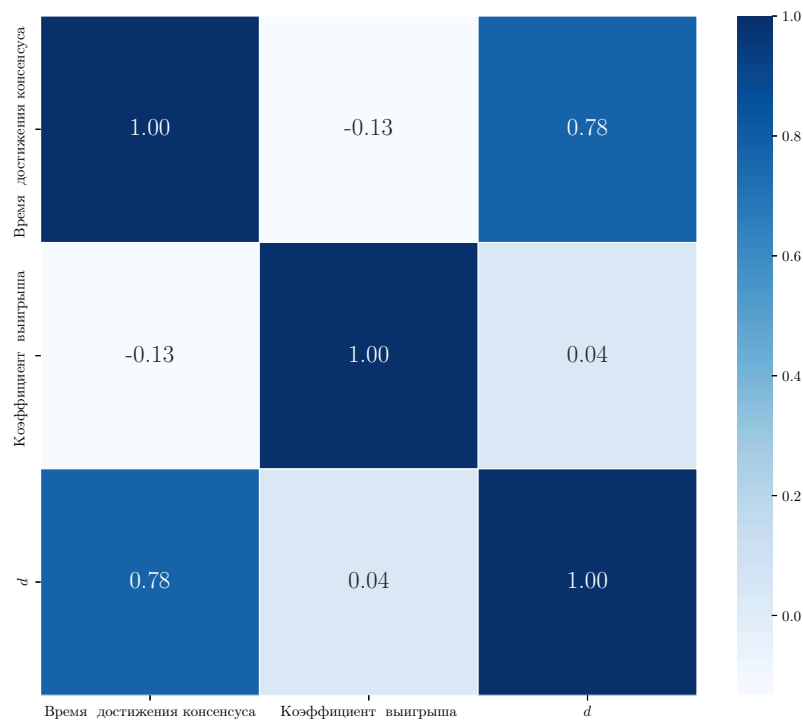


Рис. 17: Коэффициенты корреляции Пирсона

эффект? Как изменяются средние значения и дисперсии? На основе этих вопросов формулируется следующая гипотеза:

Гипотеза 3. *Средние значения и дисперсии КРІ одинаковы для различных способов расширения цикла до полного графа.*

Используются статистические тесты для проверки равенства дисперсий и средних значений. Перед этим проведены тесты на нормальность [62, 63], поскольку некоторые статистические тесты являются параметрическими и предполагают нормальное распределение данных. Результаты тестов на нормальность показаны в таблице 9, где видно, что p -значения для всех КРІ меньше 0.05. Следовательно, необходимо отвергнуть нулевую гипотезу о том, что какой-либо КРІ распределён нормально.

В таблицах 10 и 11 представлены две группы результатов: «тест на равенство дисперсий/средних для всех» показывает, равны ли дисперсии/средние значения для всех трёх способов расширения. «Парные тесты на равенство дисперсий/средних» соответствуют двухвыборочным тестам на равенство дисперсий/средних значений. Поскольку ни один из КРІ не имеет нормального распределения, для проверки равенства дисперсий используется тест Левена [64–67].

Таблица 8: Результаты тестов Колмогорова–Смирнова

| Пара | Статистика p -значения | |
|------------------------------------|--------------------------|-------|
| Время достижения консенсуса | | |
| нормальный vs случайный | 0.557 | 0.0 |
| нормальный vs кратчайший | 0.031 | 0.641 |
| случайный vs кратчайший | 0.574 | 0.0 |
| Вероятность победы | | |
| нормальный vs случайный | 0.114 | 0.0 |
| нормальный vs кратчайший | 0.02 | 0.967 |
| случайный vs кратчайший | 0.121 | 0.0 |
| d | | |
| нормальный vs случайный | 0.121 | 0.0 |
| нормальный vs кратчайший | 0.02 | 0.978 |
| случайный vs кратчайший | 0.12 | 0.0 |

Таблица 9: Результаты тестов на нормальность

Общие результаты тестов на нормальность

| | | КPI | Статистика | p -значения |
|---|-----------------------------|-----------------------------|------------|---------------|
| | | Время достижения консенсуса | 0.5 | 0.0 |
| | | Вероятность победы | 0.996 | 0.0 |
| | | d | 0.396 | 0.0 |
| Результаты тестирования на нормальность по способам расширения | | | | |
| Способ расширения | | КPI | Статистика | p -значения |
| нормальный | Время достижения консенсуса | | 0.379 | 0.0 |
| нормальный | Вероятность победы | | 0.996 | 0.003 |
| нормальный | d | | 0.32 | 0.0 |
| случайный | Время достижения консенсуса | | 0.347 | 0.0 |
| случайный | Вероятность победы | | 0.996 | 0.003 |
| случайный | d | | 0.469 | 0.0 |
| кратчайший | Время достижения консенсуса | | 0.442 | 0.0 |
| кратчайший | Вероятность победы | | 0.995 | 0.001 |
| кратчайший | d | | 0.369 | 0.0 |

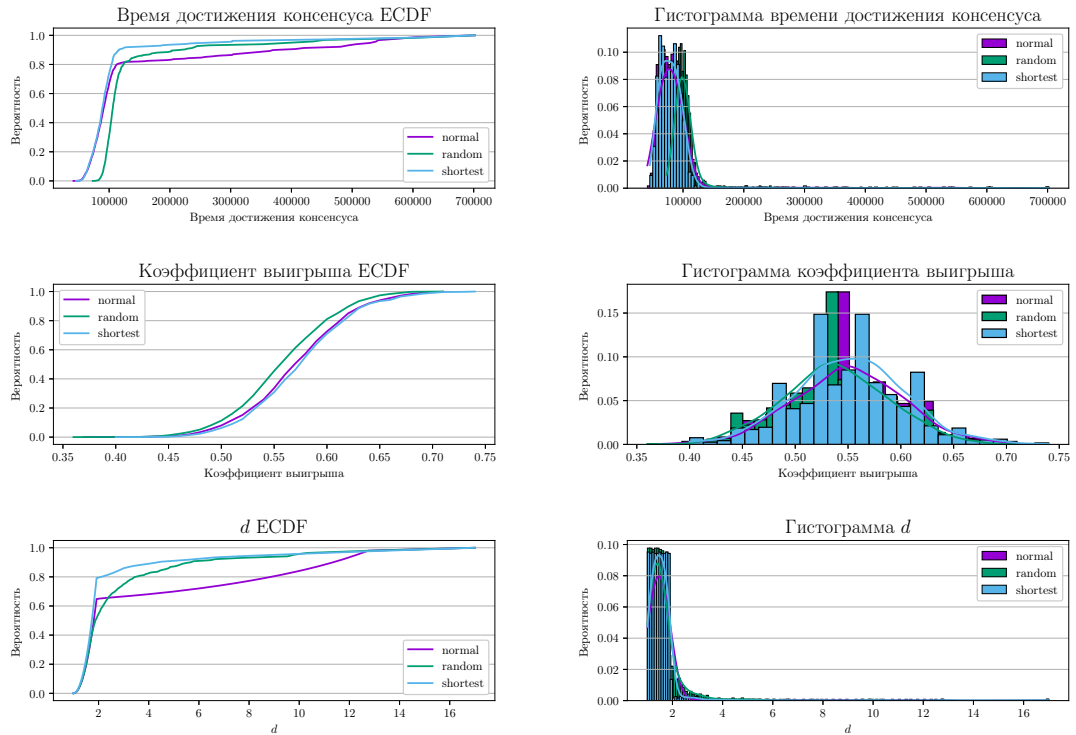


Рис. 18: Функция плотности распределения и эмпирическая функция распределения (ECDF) для различных KPI

На основании данных таблицы 10 делаются следующие выводы:

- дисперсии вероятности победы одинаковы для всех способов расширения (все p -значения тестов Левена больше 0.05);
- нулевая гипотеза о равенстве дисперсий времени достижения консенсуса для всех способов расширения отвергается (все p -значения тестов Левена меньше 0.05);
- дисперсии d одинаковы для нормального и случайного способов расширения.

Поскольку ни один из KPI не имеет нормального распределения, и не все KPI гомоскедастичны, для проверки равенства средних значений используется тест Краскала [68–71]. Результаты тестов приведены в таблице 11, и на основании них делаются следующие выводы:

- средние значения d одинаковы для всех способов расширения, то есть способ расширения цикла до полного графа не влияет на среднее значение

Таблица 10: Результаты тестов на равенство дисперсий

| Тест | Описание | KPI | Статистика | p -значения |
|---|---|-----------------------------|------------|---------------|
| Тест на равенство дисперсий для всех | | | | |
| Тест Флигнера | Независимость от распределения при одинаковых популяциях | Время достижения консенсуса | 195.249 | 0.0 |
| | | Вероятность победы | 1.206 | 0.547 |
| | | d | 23.341 | 0.0 |
| Тест Левена | Более робастный для значительно ненормальной популяции | Время достижения консенсуса | 12.533 | 0.0 |
| | | Вероятность победы | 1.074 | 0.342 |
| | | d | 5.441 | 0.004 |
| Тест Бартлетта | Зависит от нормального распределения | Время достижения консенсуса | 320.877 | 0.0 |
| | | Вероятность победы | 3.584 | 0.167 |
| | | d | 351.795 | 0.0 |
| Парные тесты на равенство дисперсий | | | | |
| Тест Левена для пар | нормальный vs случайного нормальный vs кратчайший случайный vs кратчайший нормальный vs случайный нормальный vs кратчайший случайный vs кратчайший нормальный vs случайный нормальный vs кратчайший случайный vs кратчайший | Время достижения консенсуса | 20.967 | 0.0 |
| | | Время достижения консенсуса | 6.579 | 0.01 |
| | | Время достижения консенсуса | 7.651 | 0.006 |
| | | Вероятность победы | 1.572 | 0.21 |
| | | Вероятность победы | 0.003 | 0.959 |
| | | Вероятность победы | 1.71 | 0.191 |
| | | d | 0.187 | 0.665 |
| | | d | 8.533 | 0.004 |
| | | d | 10.099 | 0.002 |

d (все p -значения тестов Краскала больше 0.05). Для времени достижения консенсуса и вероятности победы средние значения различаются в зависимости от способа расширения;

- нулевая гипотеза о равенстве средних значений времени достижения консенсуса (и вероятности победы) принимается для нормального и кратчайшего способов расширения (все p -значения в тестах Краскала больше 0.05).

Рисунки 19 и 20 показывают, как параметры (π_{c_e} и π_{c_i} на рисунке 19, и π_e и π_i на рисунке 20) влияют на вероятность победы. Можно заметить, что вероятность победы колеблется в определённых пределах, но не слишком сильно. Таким образом, можно предположить, что влияние параметров на вероятность победы ограничено.

Рисунки 21 и 22 показывают, как время достижения консенсуса изменяется с изменением параметров (π_{c_e} и π_{c_i} на рисунке 21, и π_e и π_i на рисунке 22). Можно сделать следующие интересные наблюдения:

- увеличение внешней вероятности копирования π_{c_e} оказывает отрицательное влияние на время достижения консенсуса. Это можно объяснить следующим образом: когда индивид в обществе более склонен прислушиваться-

Таблица 11: Результаты тестов на равенство средних значений

| Тест | Описание | KPI | Статистика | <i>p</i> -значение |
|---|---|---|-----------------------------|---------------------|
| Тест на равенство средних для всех | | | | |
| f_oneway тест | Независимая выборка; каждая выборка из нормально распределенной популяции; гомоскедастичность | Время достижения консенсуса | 99.143 | 0.0 |
| | | Вероятность победы | 23.205 | 0.0 |
| | | <i>d</i> | 4.795 | 0.008 |
| Тест Краскала | Размер выборки должен быть > 5 | Время достижения консенсуса Вероятность победы <i>d</i> | 1100.711 46.919 1.596 | 0.0 0.0 0.45 |
| Тест Александера- Говерна | Независимая выборка; каждая выборка из нормально распределенной популяции; гетероскедастичность | Время достижения консенсуса Вероятность победы <i>d</i> | 261.282 47.419 12.243 | 0.0 0.0 0.002 |
| Парные тесты на равенство средних | | | | |
| Тест Краскала для пар | нормальный vs случайный | Время достижения консенсуса | 795.081 | 0.0 |
| | нормальный vs кратчайший | Время достижения консенсуса | 0.775 | 0.379 |
| | случайный vs кратчайший | Время достижения консенсуса | 854.985 | 0.0 |
| | нормальный vs случайный | Вероятность победы | 32.215 | 0.0 |
| | нормальный vs кратчайший | Вероятность победы | 0.229 | 0.632 |
| | случайный vs кратчайший | Вероятность победы | 37.928 | 0.0 |
| | нормальный vs случайный | <i>d</i> | 1.092 | 0.296 |
| | нормальный vs кратчайший | <i>d</i> | 0.001 | 0.972 |
| | случайный vs кратчайший | <i>d</i> | 1.301 | 0.254 |

ся к мнению своих внешних соседей, это способствует достижению консенсуса;

- при увеличении внутренней вероятности копирования π_{c_i} время достижения консенсуса увеличивается;
- при увеличении вероятности экстернализации π_e время достижения консенсуса сначала увеличивается до максимального значения, а затем уменьшается. Это можно интерпретировать следующим образом: выражение своего мнения до определённого уровня не способствует достижению консенсуса в системе, но после достижения этого порога дальнейшее увеличение желания выразить своё мнение упрощает достижение консенсуса в системе;
- при увеличении вероятности интернализации π_i время достижения консенсуса уменьшается. Это можно объяснить тем, что когда люди охотнее принимают внешнее мнение, это ускоряет достижение консенсуса в системе.

В таблицах 12 и 13 представлены минимальное/максимальное время достижения консенсуса и вероятность победы соответственно для различных комбинаций параметров. Если сравнить левую и правую части в таблицах 12 и 13, можно заметить значительное различие между максимальными и минимальными значениями. Для анализа была выполнена кластеризация методом *k*-means

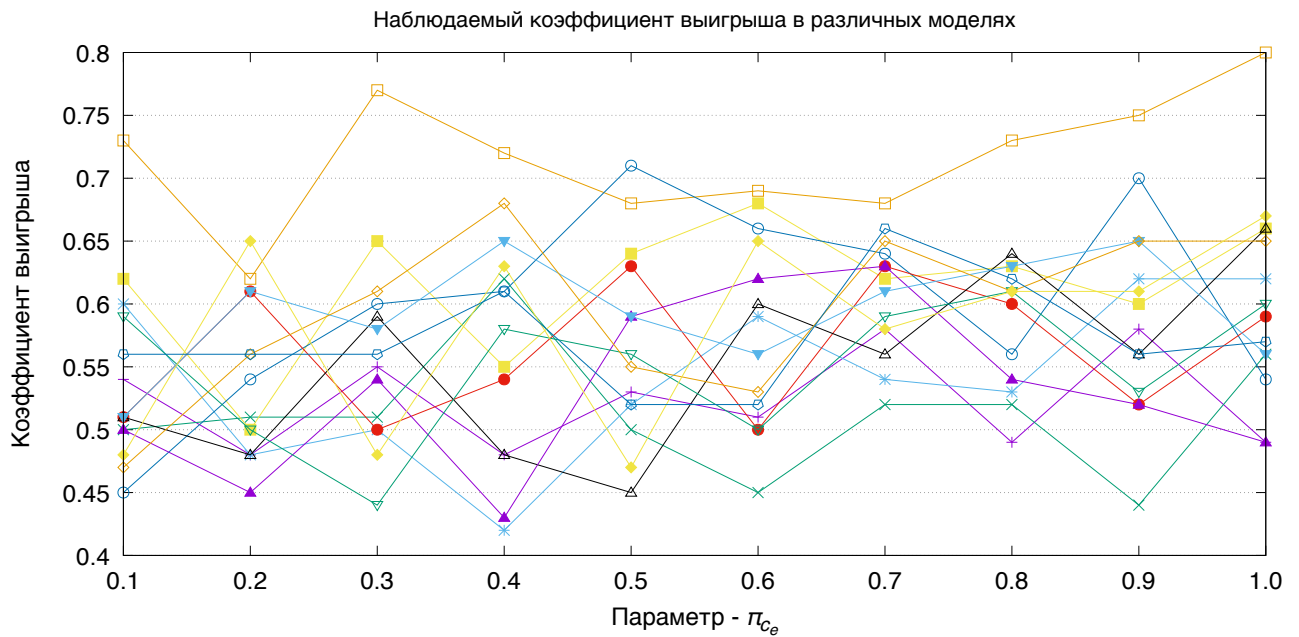
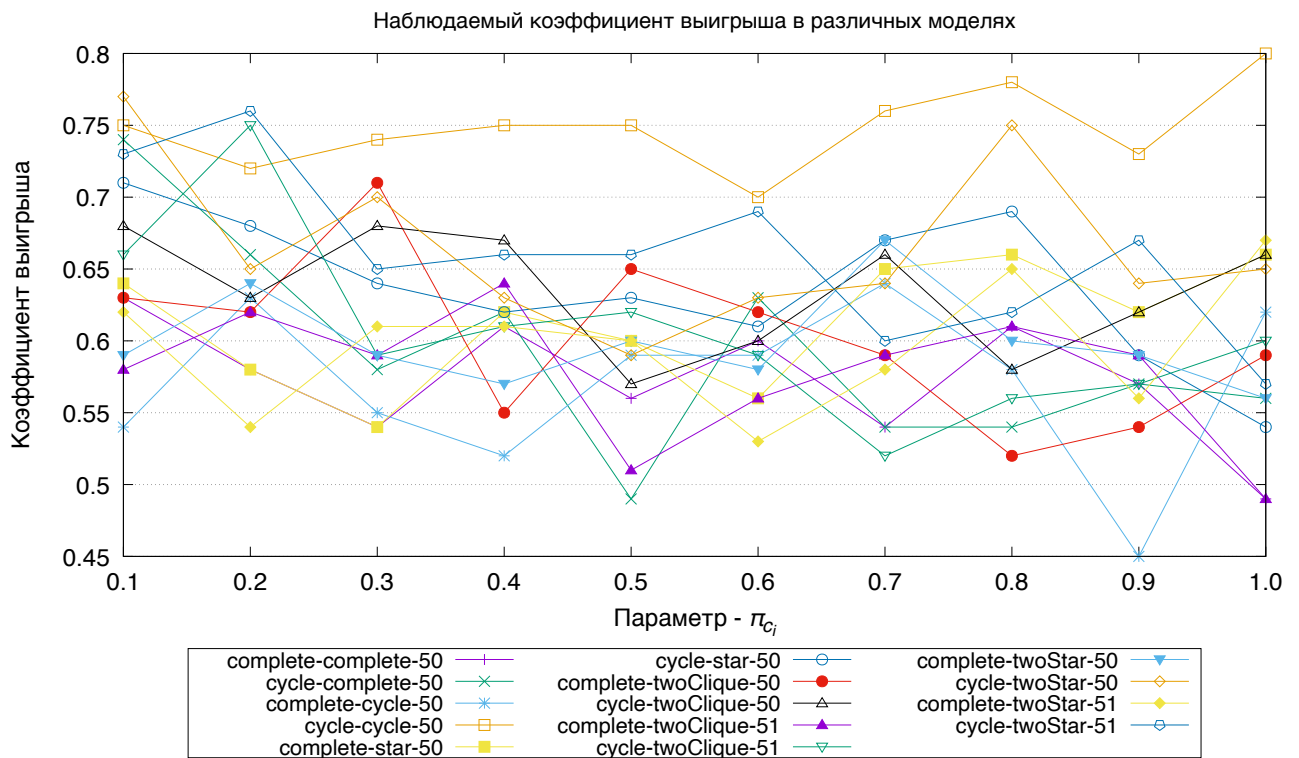
(a) Изменение внешней вероятности копирования π_{c_e} (b) Изменение внутренней вероятности копирования π_{c_i}

Рис. 19: Вероятность победы для различных вероятностей копирования

с входными переменными: время достижения консенсуса и вероятность победы отдельно [72, 73]. Полученные метки кластеров были добавлены к исходным данным. На основе этих меток кластеров был проанализирован распределение соответствующих четырёх параметров (π_{c_e} , π_{c_i} , π_e и π_i).

Таблица 12: Наблюдаемые минимальное и максимальное время достижения консенсуса

| Внешний | Внутренний | N | Max | π_{c_e} | π_{c_i} | π_e | π_i | Min | π_{c_e} | π_{c_i} | π_e | π_i |
|----------|------------|-----|-----|-------------|-------------|---------|---------|-----|-------------|-------------|---------|---------|
| complete | complete | 50 | | 0.1 | 1.0 | 0.1 | 0.3 | | 0.8 | 0.1 | 0.1 | 1.0 |
| cycle | complete | 50 | | 0.6 | 1.0 | 0.1 | 0.3 | | 0.9 | 0.1 | 0.3 | 1.0 |
| complete | cycle | 50 | | 1.0 | 1.0 | 0.4 | 0.3 | | 1.0 | 0.1 | 0.1 | 1.0 |
| cycle | cycle | 50 | | 0.1 | 0.1 | 0.6 | 0.9 | | 1.0 | 0.9 | 0.1 | 1.0 |
| complete | star | 50 | | 1.0 | 0.1 | 0.5 | 0.3 | | 0.1 | 1.0 | 1.0 | 0.3 |
| cycle | star | 50 | | 1.0 | 1.0 | 0.1 | 0.3 | | 0.1 | 1.0 | 1.0 | 0.3 |
| complete | twoClique | 50 | | 0.2 | 1.0 | 0.1 | 0.3 | | 0.1 | 0.7 | 1.0 | 0.3 |
| cycle | twoClique | 50 | | 0.7 | 1.0 | 0.1 | 0.3 | | 0.1 | 1.0 | 1.0 | 0.3 |
| complete | twoClique | 51 | | 0.1 | 1.0 | 0.1 | 0.3 | | 0.9 | 0.1 | 0.1 | 1.0 |
| cycle | twoClique | 51 | | 0.6 | 1.0 | 0.1 | 0.3 | | 0.1 | 1.0 | 1.0 | 0.3 |
| complete | twoStar | 50 | | 1.0 | 0.1 | 0.5 | 0.3 | | 0.1 | 0.9 | 1.0 | 0.3 |
| cycle | twoStar | 50 | | 1.0 | 0.9 | 0.1 | 0.3 | | 0.2 | 1.0 | 1.0 | 0.4 |
| complete | twoStar | 51 | | 1.0 | 0.1 | 0.5 | 0.3 | | 0.1 | 0.9 | 1.0 | 0.3 |
| cycle | twoStar | 51 | | 0.7 | 1.0 | 0.1 | 0.3 | | 0.2 | 1.0 | 1.0 | 0.3 |

Таблица 13: Наблюдаемые минимальная и максимальная вероятность победы

| Внешний | Внутренний | N | Max | π_{c_e} | π_{c_i} | π_e | π_i | Min | π_{c_e} | π_{c_i} | π_e | π_i |
|----------|------------|-----|-----|-------------|-------------|---------|---------|-----|-------------|-------------|---------|---------|
| complete | complete | 50 | | 0.1 | 0.1 | 0.2 | 1.0 | | 0.2 | 0.6 | 1.0 | 0.5 |
| cycle | complete | 50 | | 0.9 | 0.1 | 0.2 | 0.8 | | 0.4 | 0.1 | 0.8 | 0.3 |
| complete | cycle | 50 | | 0.2 | 0.1 | 0.4 | 1.0 | | 0.7 | 1.0 | 0.6 | 0.3 |
| cycle | cycle | 50 | | 0.1 | 0.6 | 0.1 | 0.6 | | 0.2 | 0.9 | 1.0 | 0.6 |
| complete | star | 50 | | 0.1 | 0.4 | 0.2 | 1.0 | | 0.3 | 0.4 | 1.0 | 0.4 |
| cycle | star | 50 | | 0.2 | 0.1 | 0.1 | 0.7 | | 0.1 | 0.2 | 1.0 | 0.4 |
| complete | twoClique | 50 | | 0.5 | 0.7 | 0.4 | 0.8 | | 0.7 | 0.3 | 0.9 | 0.9 |
| cycle | twoClique | 50 | | 0.9 | 0.1 | 0.1 | 1.0 | | 0.2 | 0.8 | 0.6 | 0.5 |
| complete | twoClique | 51 | | 0.1 | 0.1 | 0.1 | 0.9 | | 0.4 | 0.8 | 0.7 | 0.4 |
| cycle | twoClique | 51 | | 0.7 | 0.2 | 0.1 | 0.9 | | 0.1 | 0.3 | 1.0 | 0.5 |
| complete | twoStar | 50 | | 0.1 | 0.4 | 0.2 | 1.0 | | 0.1 | 0.8 | 0.8 | 0.4 |
| cycle | twoStar | 50 | | 0.8 | 0.2 | 0.2 | 0.9 | | 0.3 | 0.3 | 1.0 | 0.3 |
| complete | twoStar | 51 | | 0.1 | 0.1 | 0.1 | 0.7 | | 0.1 | 0.8 | 0.9 | 0.3 |
| cycle | twoStar | 51 | | 0.8 | 0.2 | 0.1 | 1.0 | | 0.1 | 1.0 | 0.9 | 0.4 |

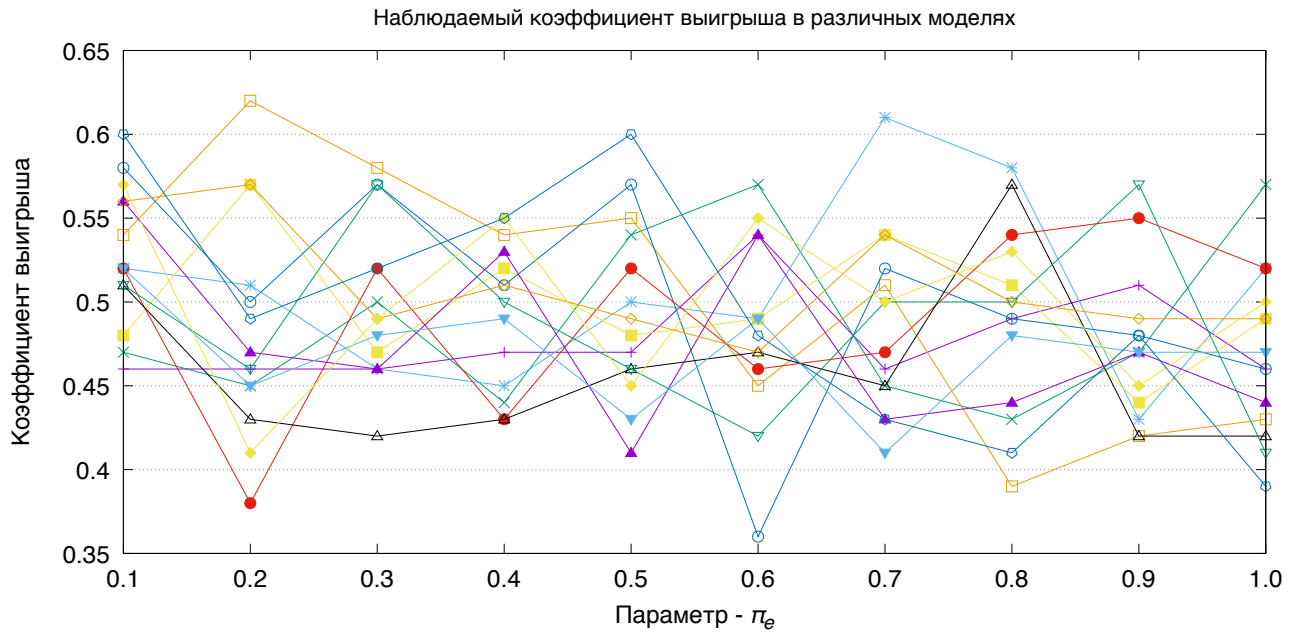
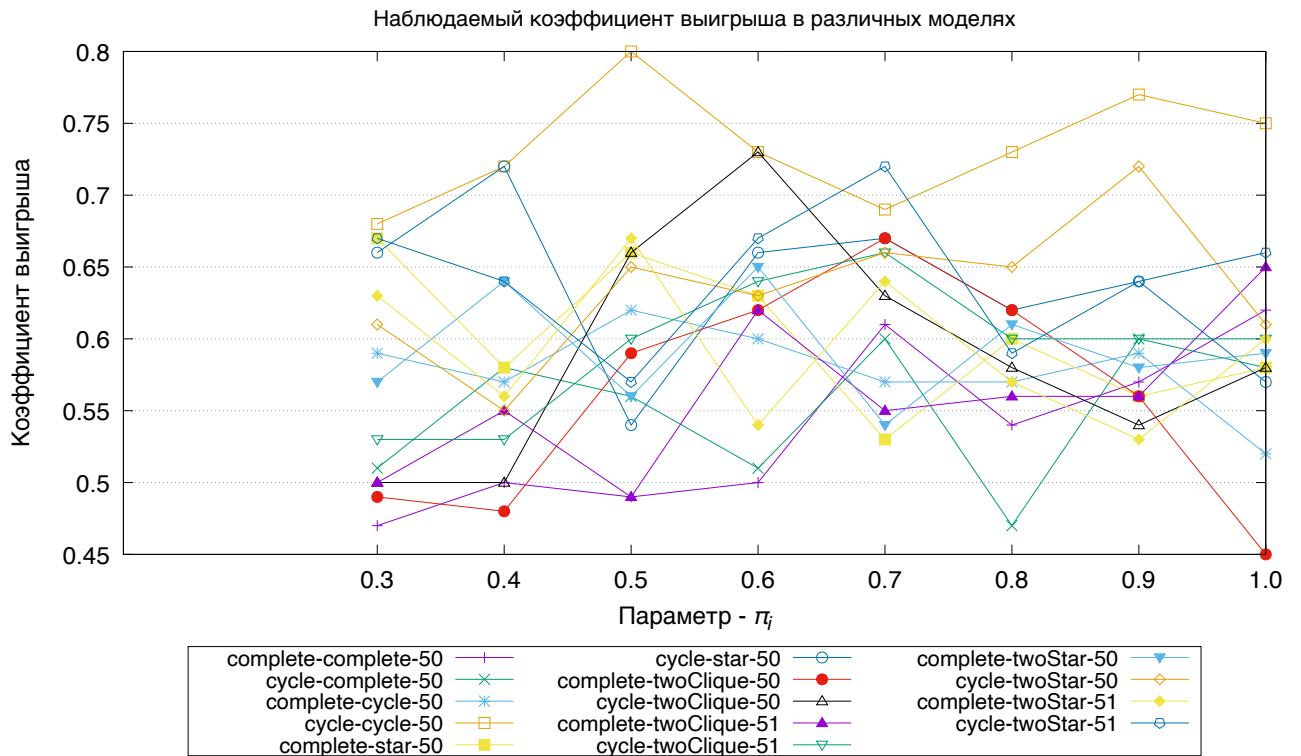
(a) Изменение вероятности экстернализации π_e (b) Изменение вероятности интернализации π_i

Рис. 20: Вероятность победы при различных значениях вероятности экстернализации и интернализации

Идеально, если удастся получить два кластера, поскольку это позволяет определить, какие комбинации параметров приводят к соответственно большим или малым значениям КРІ. На практике задается диапазон значений количества кластеров k от 1 до 20 и вычисляется коэффициент силуэта [74]. В неко-

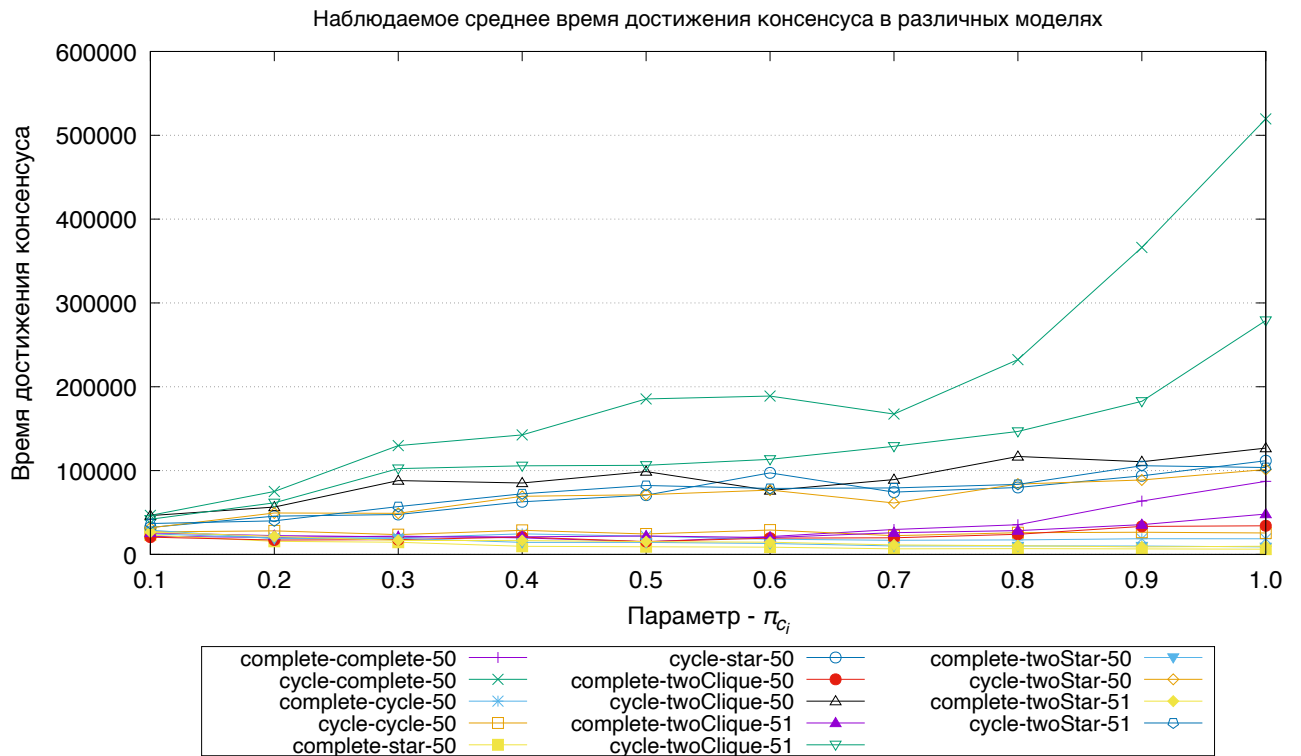
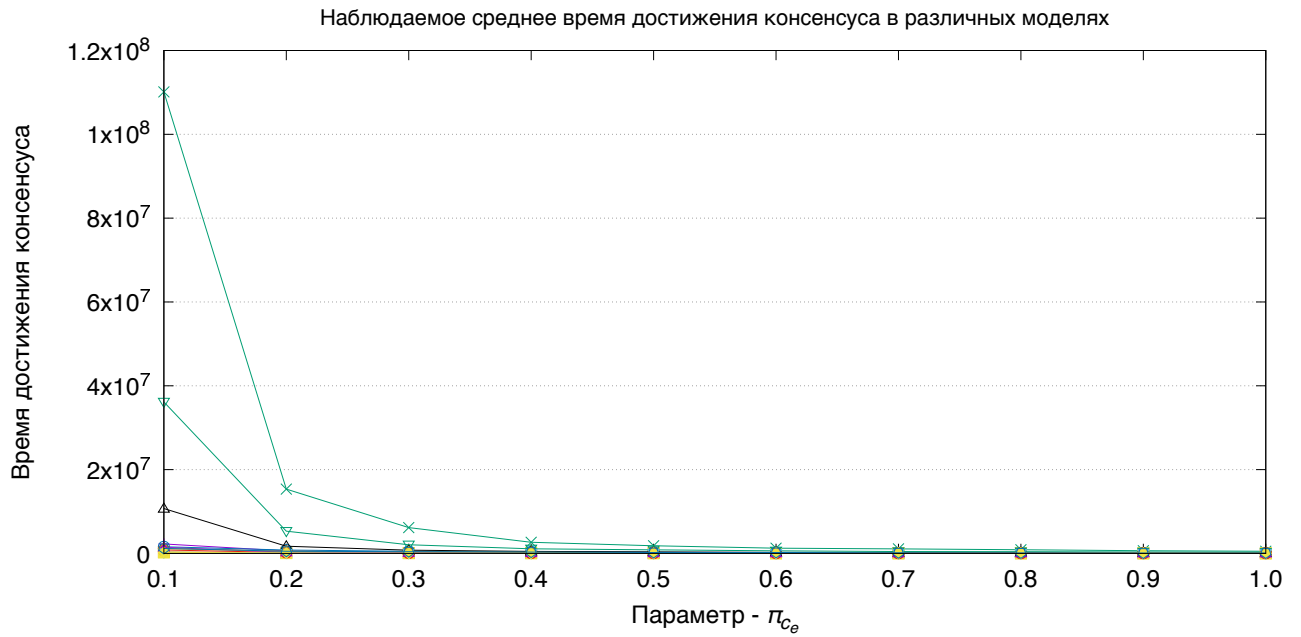


Рис. 21: Время достижения консенсуса для различных вероятностей копирования

тором смысле, оптимальным значением k является то, которое максимизирует коэффициент силуэта.

Поскольку распределения показателей выигрыша очень близки к нормальному распределению, на основе результатов кластеризации показателей

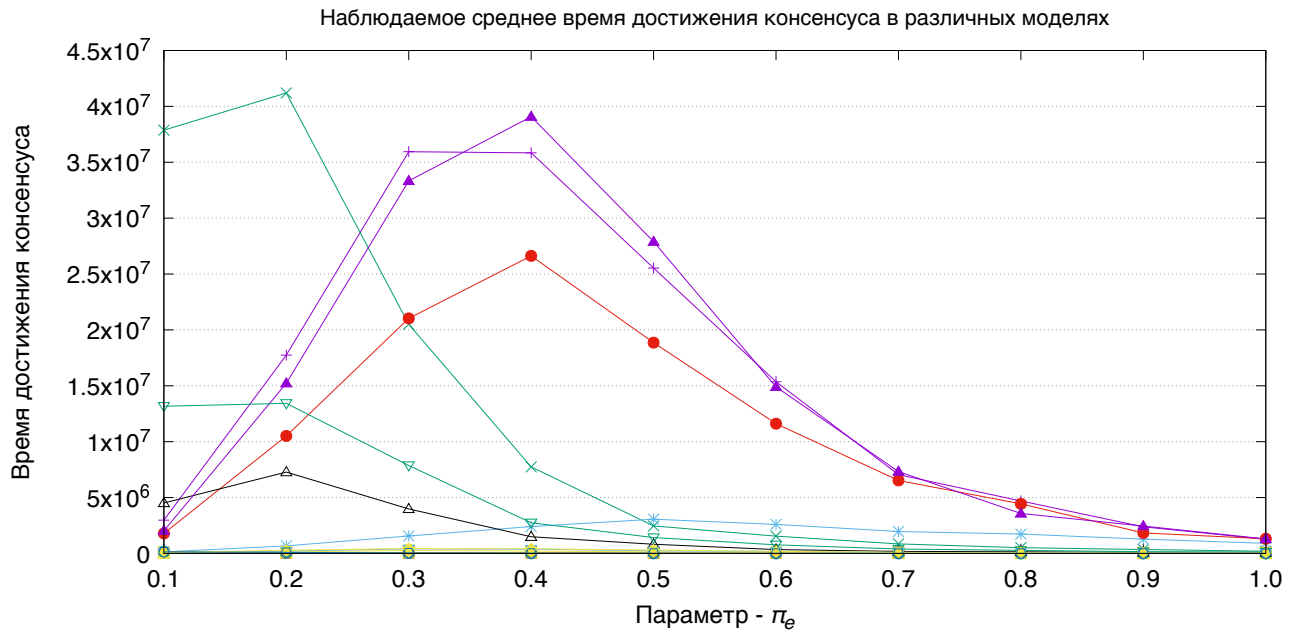
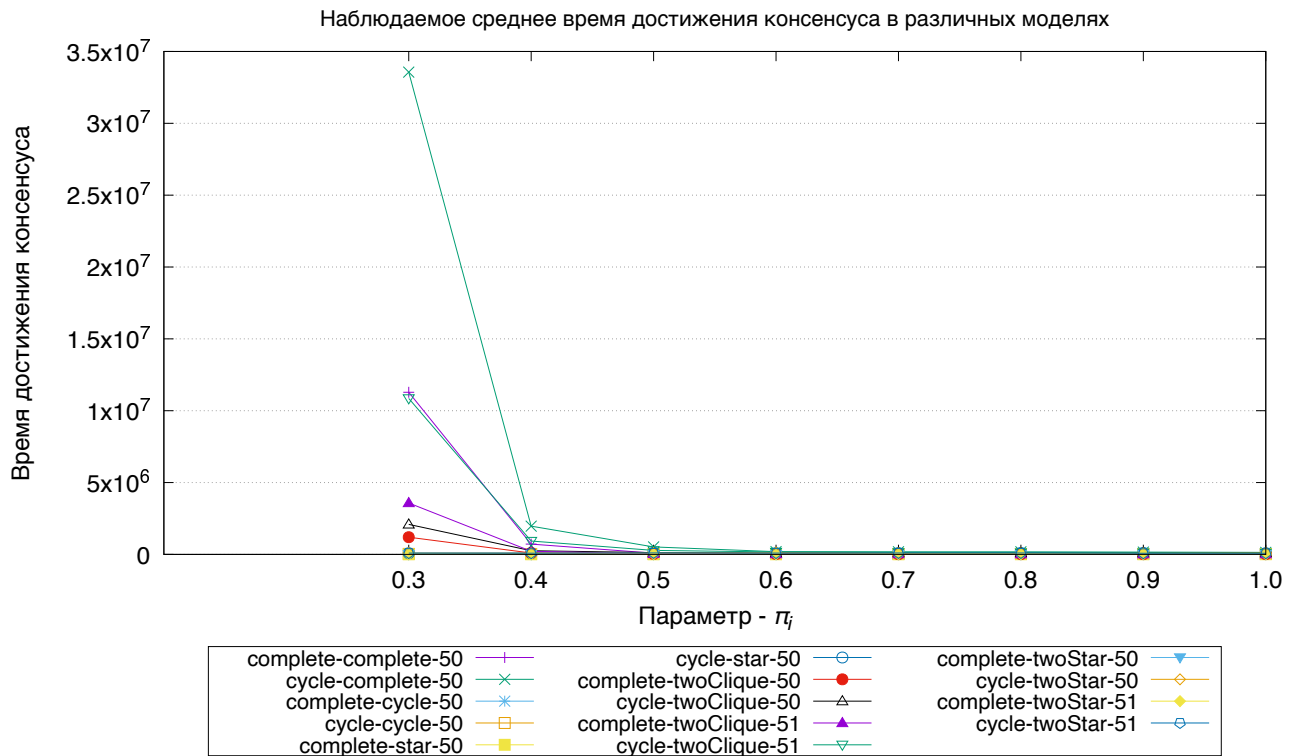
(a) Изменение внешней вероятности экстернализации π_e (b) Изменение внутренней вероятности интернализации π_i

Рис. 22: Время достижения консенсуса при различных вероятностях экстернализации и интернализации

выигрыша не удастся наблюдать каких-либо значительных различий в описательной статистике кластеров (см. таблицу 14). Поэтому в дальнейшем анализе мы сосредоточимся только на кластерном анализе времени достижения консенсуса.

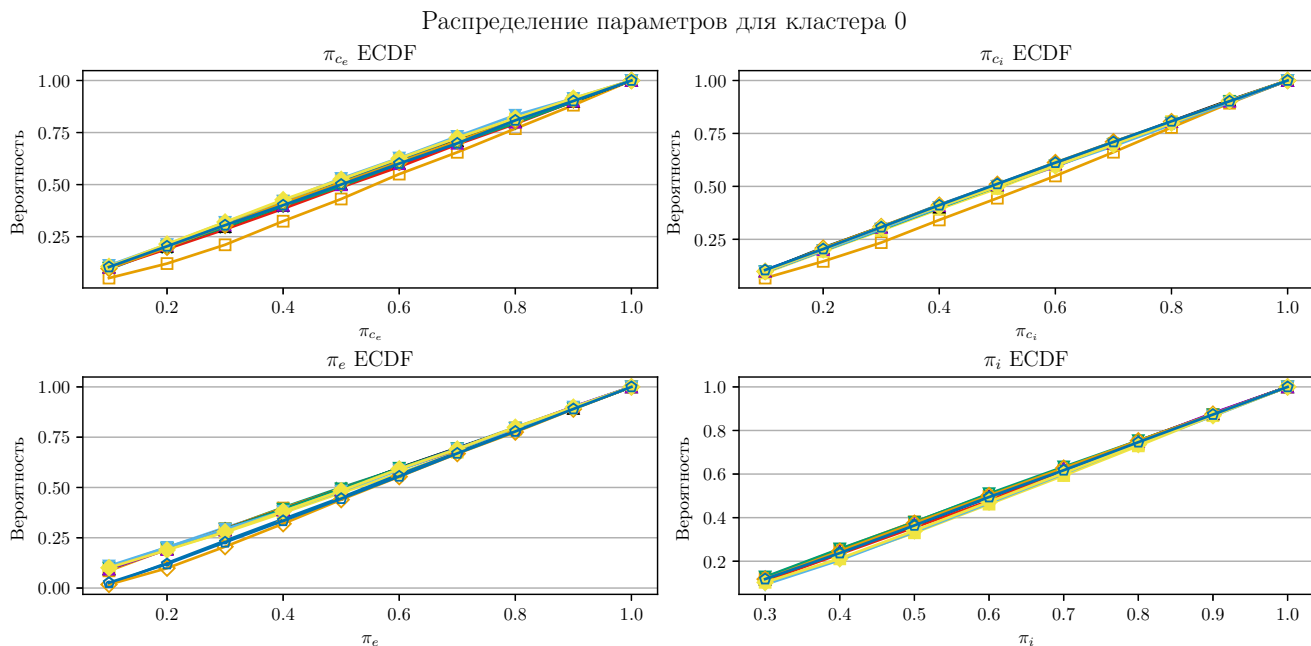
Таблица 14: Описательная статистика кластеров по показателям выигрыша (complete-complete-50)

| Параметры | Кластер | Количество | Среднее | Стандартное отклонение | Min | 25% | 50% | 75% | Max |
|-------------|---------|------------|---------|------------------------|-----|-----|-----|-----|-----|
| π_{c_e} | 0 | 169 | 0.484 | 0.282 | 0.1 | 0.3 | 0.4 | 0.7 | 1.0 |
| | 1 | 238 | 0.540 | 0.287 | 0.1 | 0.3 | 0.5 | 0.8 | 1.0 |
| | 2 | 186 | 0.573 | 0.274 | 0.1 | 0.3 | 0.6 | 0.8 | 1.0 |
| | 3 | 291 | 0.543 | 0.289 | 0.1 | 0.3 | 0.5 | 0.8 | 1.0 |
| | ⋮ | | | | | | | | |
| π_{c_i} | 0 | 169 | 0.487 | 0.306 | 0.1 | 0.2 | 0.4 | 0.7 | 1.0 |
| | 1 | 238 | 0.591 | 0.293 | 0.1 | 0.3 | 0.6 | 0.9 | 1.0 |
| | 2 | 186 | 0.533 | 0.272 | 0.1 | 0.3 | 0.5 | 0.8 | 1.0 |
| | 3 | 291 | 0.547 | 0.290 | 0.1 | 0.3 | 0.5 | 0.8 | 1.0 |
| | ⋮ | | | | | | | | |
| π_e | 0 | 169 | 0.509 | 0.290 | 0.1 | 0.3 | 0.5 | 0.7 | 1.0 |
| | 1 | 238 | 0.581 | 0.304 | 0.1 | 0.3 | 0.6 | 0.9 | 1.0 |
| | 2 | 186 | 0.525 | 0.292 | 0.1 | 0.3 | 0.5 | 0.8 | 1.0 |
| | 3 | 291 | 0.587 | 0.282 | 0.1 | 0.4 | 0.6 | 0.8 | 1.0 |
| | ⋮ | | | | | | | | |
| π_i | 0 | 169 | 0.698 | 0.233 | 0.3 | 0.5 | 0.7 | 0.9 | 1.0 |
| | 1 | 238 | 0.632 | 0.230 | 0.3 | 0.4 | 0.6 | 0.8 | 1.0 |
| | 2 | 186 | 0.631 | 0.237 | 0.3 | 0.4 | 0.6 | 0.8 | 1.0 |
| | 3 | 291 | 0.622 | 0.230 | 0.3 | 0.4 | 0.6 | 0.8 | 1.0 |
| | ⋮ | | | | | | | | |

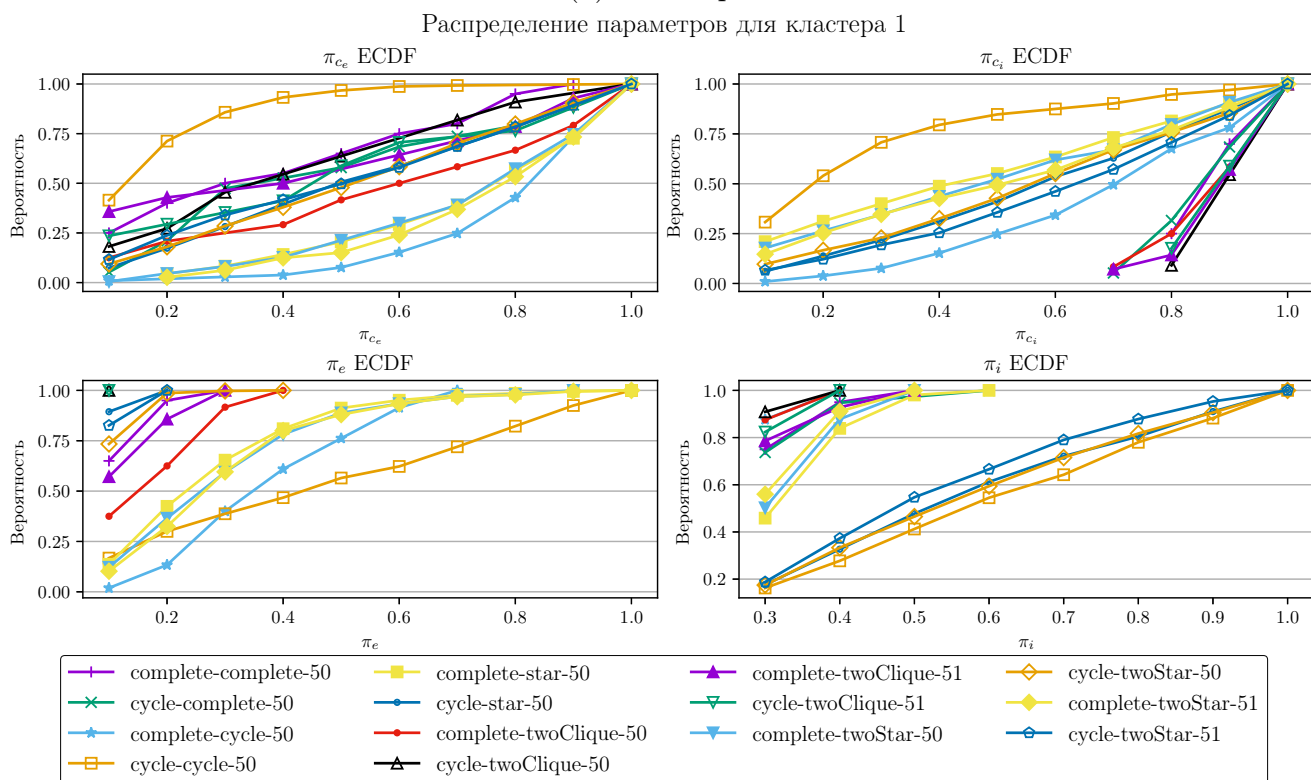
Таблица 15 показывает количество элементов в кластерах времени консенсуса, где было выбрано два кластера, поскольку это максимизирует коэффициент силуэта.

На рисунке 23а распределения времени достижения консенсуса практически совпадают для различных моделей многослойных сетей. В то же время на рисунке 23б можно чётко наблюдать заметные различия. Этот интересный результат предоставляет ценность для понимания того, что разнообразные распределения параметров могут значительно увеличить время достижения консенсуса.

В правом верхнем углу рисунка 23б видно, что когда внутренняя структура является полной или двухкликковой, параметр π_{c_i} значительно отличается по времени достижения консенсуса для этих моделей по сравнению с другими моделями. В левом нижнем углу видно, что для полной внутренней структуры или двухкликковой структуры значение параметра π_e в кластере 1 всегда равно 0.1. Для графиков в левом верхнем и правом нижнем углах можно сделать аналогичные выводы при сравнении.



(a) Кластер 0



(b) Кластер 1

Рис. 23: Эмпирическая функция распределения времени достижения консенсуса в зависимости от параметров π_{c_e} , π_{c_i} , π_e и π_i

Таблица 15: Время достижения консенсуса, размеры кластеров для каждой модели

| | Модель | Кластер 0 | Кластер 1 |
|----|-----------------------|-----------|-----------|
| 0 | complete-complete-50 | 3267 | 20 |
| 1 | cycle-complete-50 | 3241 | 19 |
| 2 | complete-cycle-50 | 3319 | 105 |
| 3 | cycle-cycle-50 | 3027 | 400 |
| 4 | complete-star-50 | 3219 | 205 |
| 5 | cycle-star-50 | 3124 | 312 |
| 6 | complete-twoClique-50 | 3239 | 24 |
| 7 | cycle-twoClique-50 | 3348 | 11 |
| 8 | complete-twoClique-51 | 3904 | 14 |
| 9 | cycle-twoClique-51 | 3831 | 17 |
| 10 | complete-twoStar-50 | 3171 | 197 |
| 11 | cycle-twoStar-50 | 3100 | 372 |
| 12 | complete-twoStar-51 | 3598 | 225 |
| 13 | cycle-twoStar-51 | 3634 | 362 |

2.5 Заключение ко второй главе

В данной главе представлен новый подход (микроскопическая версия GCVM), который моделирует процесс динамики мнений путём создания реальной сети, вместо использования статистико-физических методов, описанных в главе 1. Этот подход применим для любых двухслойных сетей, представленных в (7). Кроме того, были использованы различные способы расширения внешней циклической структуры до полной. Мы выдвигаем гипотезу о том, как способ расширения циклической сети до полного графа влияет на время достижения консенсуса и вероятность победы, основываясь на результатах моделирования, и используем статистические тесты для их проверки. Основные выводы таковы:

- циклическая внешняя структура всегда увеличивает время достижения консенсуса;
- циклическая внешняя структура оказывает положительное влияние на вероятность победы;
- циклическая внешняя структура влияет на время достижения консенсуса за счёт увеличения d , то есть существует сильная линейная зависимость

между d и временем достижения консенсуса. Чем меньше d , тем выше время достижения консенсуса;

- способ расширения цикла до полного графа существенно влияет на время достижения консенсуса и вероятность победы;
- каждый параметр оказывает разное влияние на время достижения консенсуса, но практически не влияет на вероятность победы;
- комбинация параметров (различные модели поведения индивидов) оказывает значительное влияние на время достижения консенсуса.

Глава 3. Меры центральности и динамика мнений в двухслойных сетях с узлами-репликами

В главах 1 и 2 были представлены макроскопическая и микроскопическая версии GCVM. Сетевые структуры, обсуждаемые в главах 1 и 2, являются весьма простыми — например, циклический граф, полный граф, граф-звезда и т. п. В данной главе мы рассмотрим социальную сеть клуба каратэ Закари и расширим её, добавив второй (внутренний) слой коммуникации, чтобы смоделировать динамику мнений в реальных двухслойных сетях на основе микро-GCVM. Основные обозначения и концепции микро-GCVM были представлены во второй главе, поэтому мы не будем их повторять.

В этой главе обсуждается новый подход к сведению двухслойной сети до однослойной взвешенной сети. Также предложены два быстрых и точных алгоритма для однослойной взвешенной сети, которые позволяют приближенно вычислять теоретико-игровые меры центральности, и исследовать связь между мерами центральности и характеристиками процессов динамики мнений на таких сетях. Результаты данной главы опубликованы в статьях [26, 51].

3.1 Метод построения однослойной взвешенной сети по двухслойной сети

Согласно определению двухслойной сети в главе 2 (см. уравнение (7)), двухслойная сеть $G(\mathcal{V}, \mathcal{E})$ состоит из внешней сети $G_E(\mathcal{V}_E, \mathcal{E}_E)$, внутренней сети $G_I(\mathcal{V}_I, \mathcal{E}_I)$ и набора рёбер \mathcal{E}_C , соединяющих узлы между слоями. Двухслойные сети могут быть также представлены матрицей смежности. Матрица смежности двухслойной сети является блочной матрицей, где диагональные блоки — это матрицы смежности внешнего и внутреннего слоёв, а недиагональные блоки — это матрицы смежности рёбер между внешним и внутренним слоями.

Матрица смежности $G(\mathcal{V}, \mathcal{E})$ выглядит следующим образом:

$$A = \begin{bmatrix} A_{EE} & A_{EI} \\ A_{IE} & A_{II} \end{bmatrix} \quad (10)$$

где A_{EE} — матрица смежности внешнего слоя, A_{EI} — матрица смежности рёбер между внешним и внутренним слоями, A_{IE} — матрица смежности рёбер между внутренним и внешним слоями, а A_{II} — матрица смежности внутреннего слоя. Для неориентированных графов в обоих слоях матрица смежности A является симметричной.

Принимая во внимание двухслойную сеть в моделях динамики мнений, представленных в главах 1 и 2, определим скорости копирования мнений одного узла другим:

- π_{c_e} : скорость копирования во внешнем слое, с которой индивид копирует мнение своего внешнего соседа, если они оба выбраны случайным образом;
- π_{c_i} : скорость копирования во внутреннем слое, с которой индивид копирует мнение своего внутреннего соседа;
- π_e : скорость внешней экстернализации, с которой индивид, находящийся в состоянии лицемерия⁸, выбирает публичное выражение своего внутреннего мнения;
- π_i : скорость интернализации, с которой индивид, находящийся в состоянии лицемерия, принимает своё внешнее мнение.

Мы предлагаем способ преобразования двухслойной сети $G(\mathcal{V}, \mathcal{E})$ с заданными параметрами динамики мнений $\pi_{c_e}, \pi_{c_i}, \pi_e, \pi_i$ в однослойную взвешенную сеть. Матрицу весов можно определить следующим образом:

$$W' = \pi_{c_e} \cdot A_{EE} + \pi_{c_i} \cdot A_{II} + \pi_i \cdot \Lambda_E + \pi_e \cdot \Lambda_I, \quad (11)$$

где Λ_E и Λ_I — диагональные матрицы размерности $N \times N$, и элементы на диагонали представляют степени узлов во внешнем и внутреннем слоях соответственно. Кроме того, используем w'_{ij} для обозначения элементов матрицы W' и определяем новую взвешенную сеть $G'(\mathcal{V}', \mathcal{E}', W')$, где $\mathcal{V}' = \{1, 2, \dots, N\}$

⁸Под лицемерием понимается узел, имеющий разные мнения во внешнем и внутреннем слоях.

— множество узлов, $\mathcal{E}' = \{(i, j) \mid w'_{ij} \neq 0, i, j \in \mathcal{V}'\}$ — множество рёбер, а W' — матрица весов.

На основе новой взвешенной сети $G'(\mathcal{V}', \mathcal{E}', W')$ предложим две теоретико-игровые меры центральности, которые будут описаны в разделе 3.3.3.

3.2 Сеть клуба каратэ Закари в двухслойной сетевой постановке

В качестве примера социальной сети рассмотрим сеть клуба каратэ Закари, представляющую дружеские связи между 34 участниками клуба каратэ в университете США в 1970-х годах [75]. Это исследование стало известным в литературе по анализу данных и сетей, так как оно осветило конфликт между управляющим (узел 0) и директором (узел 33), который в конечном итоге привел к расколу клуба на две группы. Однослойная сеть клуба каратэ Закари представлена на рисунке 24. Синие и красные цвета узлов обозначают два мнения в социальной сети.

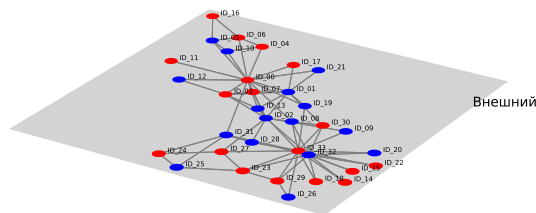


Рис. 24: Однослойная сеть клуба каратэ Закари

На рисунке 25 показано, как однослойная сеть клуба каратэ Закари может быть расширена до двухслойной сети, если добавить внутренний слой коммуникации между агентами. Если рассматривать бинарные модели динамики мнений, то в модели скрытого избирателя (CVM) [21,22] узлы во внутреннем слое не связаны, т.е. внутренний слой представлен пустым графом (см. рисунок 25а), тогда как в обобщённой модели скрытого избирателя (GCVM) [24, 25, 27] может существовать непустая сеть, представляющая внутренние коммуникации агентов. На рисунке 25b показана звездообразная структура внутреннего слоя. Цвета на рисунке 25 отображают мнения индивидов.

Цвет, синий или красный, инициализируется случайным образом для данных параметров, а именно: (i) вероятность того, что у индивида будет начальное красное мнение для базовой модели избирателя в однослойной сети, (ii) вероятности наличия начального красного мнения во внешнем, внутреннем или обоих слоях для моделей SVM и GCVM.

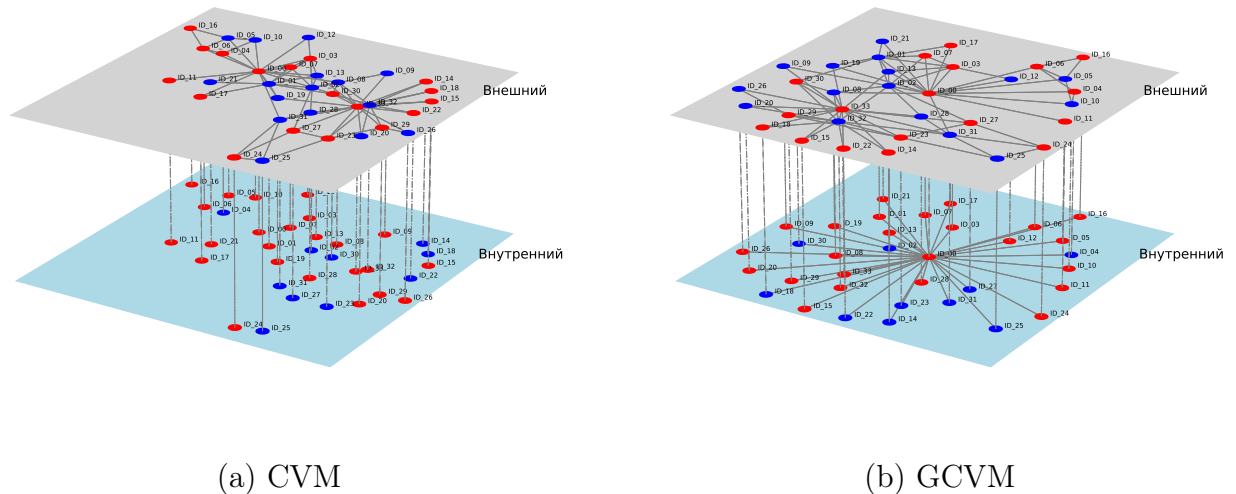


Рис. 25: Двухслойные сети, используемые в SVM и GCVM: (a) SVM: двухслойная сеть с внешним слоем — клуб каратэ Закари и пустым внутренним слоем, (b) GCVM: двухслойная сеть с внешним слоем — клуб каратэ Закари и внутренним слоем в виде звезды.

В базовой модели голосования (BVM, см. [4]), существует только один слой, как показано на рисунке 24, и каждый участник придерживается одного из двух мнений (красного или синего). На каждом шаге случайный участник выбирает случайного соседа и принимает его мнение с вероятностью копирования π_c . Этот процесс повторяется до тех пор, пока все в сети не придут к одному мнению, то есть не будет достигнут консенсус. Процесс динамики мнений в SVM [21, 22] и GCVM [24–27] реализуется на двухслойной сети (см. рисунок 25a и 25b). В SVM участники внутреннего слоя не взаимодействуют, в то время как в GCVM взаимодействие между участниками внутреннего слоя возможно. Процесс динамики мнений в SVM и GCVM аналогичен BVM, но предоставляет больше возможностей: в SVM участники могут публично выражать свое внутреннее мнение с вероятностью экстернализации π_e или принимать внешнее мнение с вероятностью интернализации π_i . Эти опции отсутствуют в BVM, то-

гда как в GCVM участники могут копировать мнение своих внутренних соседей с внутренней вероятностью копирования π_{c_i} , что не разрешено в SVM.

3.3 Меры центральности в однослойных и двухслойных сетях

В этом разделе рассматриваются несколько мер центральности. Некоторые из них определены для однослойных сетей и могут быть применены к двум слоям отдельно, другие учитывают многослойную структуру сети. Также будут введены теоретико-игровые меры центральности и представлен алгоритм для их приближенного вычисления, когда сеть содержит большое количество узлов.

Определение 1. *Средний попарный кратчайший путь для внешнего слоя d_E в двухслойной сети определяется как*

$$d_E = \sum_{s,t \in \mathcal{V}_E} \frac{d_E(s,t)}{n_E(n_E - 1)}, \quad (12)$$

где $d_E(s,t)$ — длина кратчайшего пути между узлами s и t во внешнем слое, \mathcal{V}_E — множество узлов во внешнем слое, $n_E = |\mathcal{V}_E|$ — количество узлов во внешнем слое. Аналогично, можно определить средний попарный кратчайший путь для внутреннего слоя, обозначаемого d_I .

Определение 2. *Плотность графа — это отношение количества рёбер $|\mathcal{E}|$ к максимальному количеству рёбер. Поскольку внутренний слой представлен неориентированным графом, плотность внутреннего графа определяется, как в [76]:*

$$D_I = \frac{2|\mathcal{E}_I|}{|\mathcal{V}_I|(|\mathcal{V}_I| - 1)}. \quad (13)$$

Эти определения будут использоваться в экспериментах для анализа влияния структуры сети на процесс динамики мнений (см. раздел 3.4.2).

3.3.1 Классические меры центральности

В этом разделе кратко представлены некоторые (наиболее известные) меры центральности, определенные для однослойных сетей. В написанном ниже используется обозначение V для множества узлов в однослойной сети.

Степень посредничества

Степень посредничества узла, предложенная в [29], измеряет количество геодезических путей между всеми узлами, проходящих через данный узел. Эта мера отражает уровень участия узла в распространении информации между другими узлами графа. Она рассчитывается по следующей формуле:

$$C_b(v) = \frac{1}{n_b} \sum_{s,t \in V} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}}, \quad (14)$$

где $\sigma_{s,t}$ обозначает количество кратчайших путей между узлами s и t , а $\sigma_{s,t}(v)$ — количество кратчайших путей между узлами s и t , проходящих через узел v . Нормировочный коэффициент $n_b = (|V| - 1)(|V| - 2)$ для $v \notin \{s, t\}$, иначе $n_b = |V|(|V| - 1)$, где $|V|$ — количество узлов в однослойной сети [28]. Если $s = t$, то $\sigma_{s,t} = 1$, а если $v \in \{s, t\}$, то $\sigma_{s,t}(v) = 0$.

Степень посредничества группы

Степень посредничества группы указывает долю кратчайших путей, соединяющих пары узлов, не входящих в группу, которые проходят через группу (см. [77]), и определяется следующей формулой:

$$C_{gb}(X) = \frac{1}{n_{gb}} \sum_{s,t \in V \setminus X} \frac{\sigma_{s,t}(X)}{\sigma_{s,t}}, \quad (15)$$

где $\sigma_{s,t}(X)$ — количество кратчайших путей между узлами s и t , проходящих через некоторые узлы группы X . Нормировочный коэффициент $n_{gb} = (|V| - |X|)(|V| - |X| - 1)$, где $|X|$ — количество узлов в группе X .

Степень близости

В связанном графе степень близости узла u является обратной величиной суммы длины кратчайших путей между u и всеми другими узлами графа [30–32]. При расчете степени близости обычно используется её нормированная форма, представляющая собой среднюю длину кратчайшего пути вместо их суммы. Она вычисляется следующим образом:

$$C_c(u) = \frac{n_c}{\sum_{v \in V \setminus \{u\}} d(v, u)}, \quad (16)$$

где нормировочный коэффициент $n_c = |V| - 1$.

Степень близости группы

Степень близости группы – это обратная величина суммы кратчайших расстояний от группы до всех узлов, находящихся за пределами группы [77–79], она рассчитывается по следующей формуле:

$$C_{gc}(X) = \frac{n_{gc}}{\sum_{v \in V \setminus X} d(v, X)}, \quad (17)$$

где $d(v, X)$ – кратчайшее расстояние между группой X и узлом v . Нормировочный коэффициент $n_{gc} = |V - X|$.

Степень связности

Степень связности узла v [33] определяется следующим образом:

$$C_d(v) = \frac{v_d}{n_d}, \quad (18)$$

где v_d – степень узла v , а нормировочный коэффициент $n_d = |V| - 1$.

Степень связности группы

Степень связности группы – это количество узлов вне группы, соединенных с узлами из этой группы [77, 78]. Нормированная степень связности группы

X рассчитывается по формуле:

$$C_{gd}(X) = \frac{|\{v_i \in V \setminus X | v_i \text{ соединен с } v_j \in X\}|}{n_{gd}}, \quad (19)$$

где нормировочный коэффициент $n_{gd} = |V| - |X|$.

3.3.2 Меры центральности при случайном блуждании

Вторая группа мер центральности основана на случайных блужданиях, простом динамическом процессе, который происходит на сети. Случайные блуждания также могут использоваться для приближенного моделирования других видов диффузионных процессов [34, 80, 81].

Занятость при случайном блуждании

Занятость при случайном блуждании [34] для узла v — это вероятность того, что узел v будет посещён случайным блуждающим во время бесконечно долгого блуждания, и она определяется следующим образом:

$$C_{rwoc}(v) = \lim_{t \rightarrow \infty} \frac{n_v(t)}{t}, \quad (20)$$

где $n_v(t)$ — это количество раз, когда узел v был посещён случайным блуждающим в течение времени t . Для вычисления занятости можно использовать различные стратегии исследования сети, в данной работе используется стратегия равномерного исследования (т.е. каждый узел переходит к своему соседу с равной вероятностью). Во взвешенных сетях вероятности переходов пропорциональны весам рёбер.

Аналитические выражения занятости при случайном блуждании с равномерной стратегией исследования в многослойных сетях представлены в [34].

Нагрузка узла при случайном блуждании

Наиболее распространённая мера посредничества — это центральность посредничества по кратчайшему пути [29, 34], где центральность узла v рассчитывается на основе числа кратчайших путей между всеми парами узлов, проходящих через v . Однако в реальных сетях информация (слухи, сообщения или интернет-пакеты), перемещающаяся по сети, не всегда следуют кратчайшему пути [34, 82, 83]. Нагрузка узла при случайном блуждании v определяется как количество случайных блужданий между любой парой узлов (s, d) , проходящих через узел v [35]:

$$C_{rwb}(v) = \frac{1}{n_{rwb}} \sum_{\substack{s, t \in V \\ s \neq t \\ v \neq s, v \neq t}} \mathbf{1}_{v \in \text{Path}_{s \rightarrow t}}, \quad (21)$$

где $n_{rwb} = 2N(N - 1)$ — нормировочный коэффициент. Индикаторная функция $\mathbf{1}_{v \in \text{Path}_{s \rightarrow t}}$ равна 1, если узел v находится на пути между узлами s и t , и 0 в противном случае. $\text{Path}_{s \rightarrow t}$ — случайный путь между узлами s и t в сети. Повторяя процесс случайного блуждания несколько раз, чтобы получить разные случайные пути, можно вычислить среднюю нагрузку узла при случайном блуждании.

Полезно получить аналитическое выражение нагрузки узла при случайном блуждании через поглощающее случайное блуждание, где поглощающее состояние выбирается как узел назначения d [35, 81]. Расширенное аналитическое выражение нагрузки узла при случайном блуждании для многослойных сетей можно найти в [34].

Близость при случайном блуждании

Вариант степени близости — это близость при случайном блуждании, вычисление которой основано на среднем времени первого достижения (MFPT). MFPT определяется как среднее количество шагов до достижения узла d , начиная с узла s . Чем меньше среднее MFPT, тем быстрее узел доступен для других узлов. Таким образом, узел с меньшим средним MFPT до всех других узлов считается более «центральным» в сети. Близость при случайном блуждании определяется как обратная величина среднего MFPT и вычисляется по формуле:

$$C_{rwcc}(v) = \frac{n - 1}{\sum_{u \in V \setminus \{v\}} \tau_{uv}}, \quad (22)$$

где τ_{uv} — это MFPT от узла u до узла v . Матрицу MFPT можно вычислить аналитически с помощью фундаментальной матрицы Кемени-Снелла Z [84, 85] или с помощью поглощающих случайных блужданий [81, 86].

Аналитические выражения близостей при случайном блуждании для многослойных сетей можно найти в [34].

3.3.3 Теоретико-игровые меры центральности

Центральность на основе вектора Шепли

Вектор Шепли — одно из дележей или решений в кооперативной теории игр, предложен он был Ллойдом Шепли в 1953 году [37]. Он представляет собой меру среднего вклада игрока во все возможные коалиции. Вектор Шепли — это концепция, при которой каждому игроку назначается единственное число в качестве компоненты распределения, вычисляемого по следующей формуле:

$$\phi(i) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} (v(S \cup \{i\}) - v(S)), \quad (23)$$

где $S \subseteq V$ представляет собой коалицию, а значение коалиции S или значение характеристической функции для S обозначается как $v(S)$. Определим характеристическую функцию $v(S)$ как половину суммы взвешенных степеней всех узлов в подграфе, индуцированном коалицией S , а именно:

$$v(S) = \frac{1}{2} \sum_{\{i,j\} \subseteq S} W(i, j), \quad (24)$$

где $W(i, j)$ — это вес ребра между узлами i и j в подграфе, индуцированном коалицией S . Коэффициент $\frac{1}{2}$ используется для коррекции в случае неориентированного графа, так как каждое ребро учитывается дважды при суммировании по всем парам узлов.

Алгоритм 1 описывает, как вычислить вектор Шепли на основе взвешенного графа. Однако вычисление вектора Шепли является сложной вычисли-

Алгоритм 1 Вычисление вектора Шепли на основе взвешенного графа

Требования Граф $G(V, E, W)$ с $n = |V|$ узлами

Гарантии Компонент вектора Шепли $\phi(i)$ для каждого узла $i \in V$

```

1: for all узлов  $i \in V$  do
2:   Инициализация  $\phi(i) \leftarrow 0$ 
3: end for
4: for all узлов  $i \in V$  do
5:   for all подмножеств  $S \subseteq V \setminus \{i\}$  do
6:     Вычислить  $v(S) \leftarrow \sum_{\{j,k\} \subseteq S} W(j, k)$  в подграфе, индуцированном  $S$ 
7:     Вычислить  $v(S \cup \{i\})$  в подграфе, индуцированном  $S \cup \{i\}$ 
8:      $\Delta v(S, i) \leftarrow v(S \cup \{i\}) - v(S)$ 
9:      $\text{coeff} \leftarrow \frac{|S|! \cdot (n - |S| - 1)!}{n!}$ 
10:     $\phi(i) \leftarrow \phi(i) + \text{coeff} \cdot \Delta v(S, i)$ 
11:   end for
12: end for
   return  $\phi(i)$  для всех  $i \in V$ 

```

тельной задачей, особенно для больших сетей с большим количеством коалиций.⁹ Поэтому мы предлагаем новый подход для вычисления приближенного вектора Шепли, описанный в следующем разделе.

Аппроксимированный вектор Шепли

С учетом того, что влияние от других узлов уменьшается с увеличением длины пути, предложены несколько идей для ускорения вычисления вектора Шепли:

1. *Ограничение глубины:* Ограничивая глубину достижимых узлов (reachable nodes), которые рассматриваются, уменьшается количество подмножеств, которые необходимо учитывать.
2. *Итерация по локальным подмножествам:* Итерация по подмножествам только среди достижимых узлов, а не по всему графу, уменьшает количество итераций.
3. *Сэмплирование по количеству достижимых узлов:* Для большого числа достижимых узлов, вычислительная сложность может быть уменьшена за счет случайной выборки, тем самым уменьшая количество подмножеств, по которым происходит итерация.

⁹Для сети с n узлами общее количество коалиций равно 2^n .

Сначала определим $\psi(i, d_{max})$ как множество достижимых узлов узла i до заданной глубины d_{max} за исключением узла i . Аппроксимированный вектор Шепли для узла i на основе множества $\psi(i, d_{max})$ можно вычислить по формуле:

$$\phi_a(i) = \begin{cases} \sum_{S \subseteq \psi(i, d_{max})} \frac{v(S \cup \{i\}) - v(S)}{2^{|\psi(i, d_{max})|}} & \text{если } |\psi(i, d_{max})| < m, \\ \beta \sum_{S \subseteq \psi(i, d_{max})} \frac{v(S \cup \{i\}) - v(S)}{2^{|\psi(i, d_{max})|}} & \text{если } |\psi(i, d_{max})| \geq m, \end{cases} \quad (25)$$

где $\beta = \frac{|\psi(i, d_{max})| + 1}{m + 1}$ — это коэффициент масштабирования, а m — максимальное количество рассматриваемых достижимых узлов.

Для $|\psi(i, d_{max})| \geq m$ производится случайная выборка m узлов из множества $\psi(i, d_{max})$ несколько раз, и компоненты вектора Шепли вычисляются на основе этих выборок. Время выборки $H_{|\psi(i, d_{max})|, m}$ определяется формулой (26) (см. [87]):

$$H_{|\psi(i, d_{max})|, m} = \left(\frac{|\psi(i, d_{max})| + \frac{1}{2}}{m} - \frac{1}{2} \right) (\ln |\psi(i, d_{max})| + \gamma) + \frac{1}{2}, \quad (26)$$

где $\gamma \approx 0.5772156649$ — постоянная Эйлера-Маскерони. Эта формула представляет собой математическое ожидание количества выборок для сбора m узлов из множества $\psi(i, d_{max})$ до тех пор, пока не будут собраны все достижимые узлы¹⁰.

Формула (25) дает хорошую аппроксимацию вектора Шепли, особенно если плотность графа не слишком высока (менее 0.7). Необходимо отметить, что предложенный алгоритм обеспечивает точную оценку отношения аппроксимированной компоненты вектора Шепли к сумме всех его компонент. Зная это отношение и точное значение $v(N)$, аппроксимированный вектор Шепли можно вычислить, используя коэффициент масштабирования ξ :

$$\xi = \frac{v(N)}{\sum_{i \in V} \phi_a(i)}. \quad (27)$$

Шаги вычисления аппроксимированного вектора Шепли во взвешенных графах описаны в алгоритме 2. Используется та же характеристическая функция, что и при вычислении оригинального вектора Шепли, но количество вычислений уменьшается, следуя предложенным идеям.

¹⁰Эту задачу можно рассматривать как обобщенную задачу коллекционера купонов [87].

Более точный аппроксимированный вектор Шепли можно получить, умножив его компоненты $\phi_a(i)$, определенные формулой (25), на коэффициент ξ , определенный формулой (27). Этот шаг включен в алгоритм 2. Результаты его работы приведены в разделе 3.4.1.

Алгоритм 2 Вычисление аппроксимированного вектора Шепли во взвешенном графе

Требования Взвешенный граф $G = (V, E, W)$, ограничение глубины d_{max} , максимальный размер m множества достижимых узлов

Гарантии Аппроксимированный вектор Шепли $(\phi_a(i), i \in V)$

- 1: Инициализировать $\phi_a(i) \leftarrow 0$ для каждого $i \in V$
 - 2: **for** $i \in V$ **do**
 - 3: $\psi(i, d_{max}) \leftarrow$ вычислить или получить всех *reachable nodes* для i до глубины d_{max}
 - 4: **if** $|\psi(i, d_{max})| < m$ **then**
 - 5: **for** каждое подмножество $S \subseteq \psi(i, d_{max}) \setminus \{i\}$ **do**
 - 6: Вычислить $v(S) \leftarrow \sum_{\{j,k\} \subseteq S} W(j, k)$ в подграфе, индуцированном S
 - 7: Вычислить $v(S \cup \{i\})$ в подграфе, индуцированном $S \cup \{i\}$
 - 8: $\Delta v(S, i) \leftarrow v(S \cup \{i\}) - v(S)$
 - 9: $\phi_a(i) \leftarrow \phi_a(i) + \Delta v(S, i)$
 - 10: **end for**
 - 11: $\text{coeff} \leftarrow \frac{1}{2^{|\psi(i, d_{max})|}}$
 - 12: $\phi_a(i) \leftarrow \phi_a(i) \cdot \text{coeff}$, нормализовать $\phi_a(i)$ на основе количества подмножеств
 - 13: **else**
 - 14: Выбрать m узлов случайно из $\psi(i, d_{max})$ и повторить $H_{|\psi(i, d_{max})|, m}$ раз
 - 15: **for** $i = 1$ to $H_{|\psi(i, d_{max})|, m}$ **do**
 - 16: $s_{reachable} \leftarrow$ произвести случайную выборку из m узлов из $\psi(i, d_{max})$,
 - 17: **for** каждое подмножество $S \subseteq s_{reachable} \setminus \{i\}$ **do**
 - 18: Рассчитать $v(S)$ и $v(S \cup \{i\})$, как ранее
 - 19: $\Delta v(S, i) \leftarrow v(S \cup \{i\}) - v(S)$
 - 20: $\phi_a(i) \leftarrow \phi_a(i) + \Delta v(S, i)$
 - 21: **end for**
 - 22: **end for**
 - 23: $\text{coeff} \leftarrow 1/2^{|\psi(i, d_{max})|} / H_{|\psi(i, d_{max})|, m} \cdot \frac{|\psi(i, d_{max})|+1}{m+1}$
 - 24: $\phi_a(i) \leftarrow \phi_a(i) \cdot \text{coeff}$
 - 25: **end if**
 - 26: **end for**
 - 27: Определить коэффициент масштабирования $\xi \leftarrow \frac{v(N)}{\sum_{i \in V} \phi_a(i)}$ \triangleright Для точных результатов
 - 28: **return** $\phi_a(i) \leftarrow \xi \cdot \phi_a(i)$ для всех $i \in V$
-

Центральность на основе вектора Майерсона

Значение или вектор Майерсона было введено Роджером Майерсоном в 1977 году [38] и представляет собой правило распределения выигрышей, когда игроки связаны сетевой структурой. Изменяя метод расчета вектора Шепли, Майерсон учитывает связи в сети, что позволяет отразить влияние сетевой

структуры на кооперативную игру. Рассматривается игра, в которой граф G является деревом, состоящим из N узлов, а характеристическая функция определяется по схеме, предложенной в [88]. Каждое прямое соединение (наличие ребра в сети) приносит коалиции S значение r , где $0 \leq r \leq 1$. Игроки также получают выигрыши от не прямых соединений, которые уменьшаются с увеличением длины пути. Характеристическая функция определяется следующим образом: [40]:

$$v(S) = a_1 r + a_2 r^2 + \dots + a_k r^k + \dots + a_L r^L = \sum_{k=1}^L a_k r^k, \quad (28)$$

где L — максимальное расстояние между двумя узлами в коалиции, a_k — количество путей длины k в данной коалиции, $v(i) = 0, \forall i \in N$.

Мазалов и его соавторы в [89] доказали, что правило распределения, а именно значение Майерсона для невзвешенных графов, имеет вид:

$$Y_i(v, g) = \frac{\sigma_1(i)}{2} r + \frac{\sigma_2(i)}{3} r^2 + \dots + \frac{\sigma_L(i)}{L+1} r^L = \sum_{k=1}^L \frac{\sigma_k(i)}{k+1} r^k, \quad (29)$$

где $\sigma_k(i)$ — количество путей длины k , которые проходят через узел i . Тот же подход к определению распределения может быть применён к взвешенным графам путём преобразования веса ребра в количество путей между двумя узлами, то есть преобразования взвешенного графа в мультиграф [40].

Аппроксимированное значение Майерсона

Так как вычисление значения Майерсона также является «затратным», особенно для больших сетей, то можно рассмотреть правило «шести рукопожатий» [90], которое реализует идею о том, что любым двум людям в мире, которые не знакомы друг с другом, требуется лишь несколько посредников, чтобы установить контакт. Основываясь на этой идее, можно снизить вычислительные затраты, ограничив максимальную глубину L узлов, с которыми узел связан. Для социальной сети чем выше плотность, тем меньше промежуточных узлов требуется для соединения двух узлов. Мы переопределим L следующим

образом:

$$L = \begin{cases} 6, & \text{если } D \leq 0.2, \\ 2, & \text{если } 0.2 < D \leq 0.3, \\ 1, & \text{если } D > 0.3, \end{cases} \quad (30)$$

где D — это плотность сети.

Алгоритм для вычисления значений Майерсона во взвешенных графах представлен в алгоритме 3. Используется та же характеристическая функция, что и в уравнении (29), но максимальная глубина ограничивается для аппроксимации значения Майерсона согласно формуле (30).

Аналогично уравнению (27), можно определить масштабирующий коэффициент ξ для аппроксимированного значения Майерсона следующим образом:

$$\xi = \frac{v(N)}{\sum_{i \in V} Y_i(v, g)}. \quad (31)$$

В обеих формулах (27) и (31) используется $v(N)$, и для его вычисления в случае значения Майерсона необходимо посчитать количество путей для всех длин, т.е. a_1, a_2, \dots, a_L . Это значительно более вычислительно затратно, чем в случае вектора Шепли. Однако после масштабирования с помощью ξ получается более точная аппроксимация, что будет показано в разделе 3.4.1.

3.4 Эксперименты

Структура сети оказывает огромное влияние на ключевые показатели эффективности (КРІ) динамики мнений, реализованной на этой сети. Поэтому определяются несколько характеристик сети, которые, по нашему мнению, имеют наиболее значительную корреляцию с КРІ динамики мнений. Эксперименты организованы следующим образом: в разделе 3.4.1 представлена серия экспериментов, в которых были вычислены меры центральности, основанные на аппроксимированных векторах Шепли и Майерсона (реализации Алгоритмов 2 и 3) для графов с различной плотностью. В разделе 3.4.2 проводится анализ корреляции между КРІ динамики мнений и характеристиками сети.

Алгоритм 3 Вычисление аппроксимированного значения Майерсона для взвешенного графа

Требования Взвешенный граф $G(V, E, W)$, коэффициент дисконтирования r (по умолчанию 0.5), булева переменная $weight$ для учета весов рёбер (по умолчанию True), булева переменная $approximate$ для аппроксимации (по умолчанию True), булева переменная $scale$ для масштабирования (по умолчанию False)

Гарантии Компоненты значения Майерсона $Y_i(v, g)$, $i \in V$

```

1:  $Y_i(v, g) \leftarrow 0$  для каждого  $i \in V$  ▷ Инициализация значений Майерсона
2: if  $approximate$  then
3:    $L \leftarrow \begin{cases} 1, & \text{если } density(G) > 0.3, \\ 2, & \text{если } density(G) > 0.2, \\ 6, & \text{в остальных случаях} \end{cases}$  ▷ Корректировка  $L$  на основе  $density(G)$ .
4: else
5:    $L \leftarrow |V| - 1$  ▷ Без аппроксимации
6: end if
7: for all  $i \in V$  do
8:    $l2c \leftarrow$  Инициализация карты длина->счет для путей через  $i$ 
9:   for all пары  $(start, end)$  в  $V \times V$  do
10:    for all  $path$  среди всех простых путей от  $start$  до  $end$  с  $length \leq L$  do
11:      if  $node \in path$  then
12:         $length \leftarrow len(path) - 1$ 
13:        if  $weight$  then
14:           $l2c[length] \leftarrow l2c[length] + \min_{(u,v) \in path} w(u, v)$ 
15:        else
16:           $l2c[length] \leftarrow l2c[length] + 1$ 
17:        end if
18:      end if
19:    end for
20:  end for
21:  for all  $(length, count) \in l2c$  do
22:     $count \leftarrow count/2$  ▷ Корректировка для устранения двойного подсчета
23:     $Y_i(v, g) \leftarrow Y_i(v, g) + \left( count \cdot \frac{r^{length}}{length+1} \right)$ 
24:  end for
25: end for
26: if  $scale$  then
27:   Определить  $\xi \leftarrow \frac{v(N)}{\sum_{i \in V} Y_i(v, g)}$  ▷ Для более точных результатов
28:   return  $Y_i(v, g) \leftarrow \xi \cdot Y_i(v, g)$  для всех  $i \in V$ 
29: else
30:   return  $Y_i(v, g)$  для всех  $i \in V$ 
31: end if

```

3.4.1 Центральности, основанные на векторе Шепли и значении Майерсона

Из-за вычислительной сложности вектора Шепли и значения Майерсона эксперименты проводятся на сетях, состоящих из 20 узлов, для заданной плотности. Сеть создается следующим образом: случайным образом и многократно выбираются два различных узла из множества узлов и добавляется связь между ними, пока плотность не достигнет желаемого значения. Были разработаны следующие эксперименты для оценки производительности предложенных мер центральности:

1. **Центральность на основе вектора Шепли:** Плотность сети принимает значения: $0.1, 0.2, \dots, 1.0$. Для каждого взвешенного или невзвешенного графа вычисляются точный вектор Шепли и аппроксимированный вектор Шепли. Сравниваются: (i) эти два значения и (ii) время вычисления для этих двух методов.
2. **Центральность на основе значения Майерсона:** Плотность сети принимает значения: $0.1, 0.11, \dots, 0.2$ ¹¹. Для каждого взвешенного или невзвешенного графа вычисляются точное значение Майерсона и аппроксимированное значение Майерсона. Сравниваются: (i) эти два значения и (ii) время вычисления для этих двух методов. Также проводится серия экспериментов на сетях, состоящих из 10 узлов, с плотностью от 0.05 до 1.0, чтобы показать эффективность предложенного алгоритма для аппроксимации значения Майерсона.
3. **Сравнение с классическими мерами центральности:** На основе реального набора данных социальной сети «Клуб карате Закари» была создана двухслойная сеть путем добавления различных внутренних структур сети. Двухслойная сеть сводится к однослойной взвешенной сети с использованием параметров динамики мнений по методу, описанному в разделе 3.1. Наиболее важными узлами в сети являются узлы 0 (инструктор — Mr Hi) и 33 (менеджер — John A). Коэффициент точности или

¹¹Плотность ограничивается множеством $\{0.1, 0.11, \dots, 0.2\}$, так как даже небольшое увеличение плотности значительно увеличивает время вычисления. Это делает выполнение серии экспериментов довольно сложным.

просто точность меры центральности определяется следующим образом:

$$A_c = \frac{|\text{Top 2 nodes according to a centrality measure} \cap \{0, 33\}|}{2} \cdot 100\%. \quad (32)$$

Значение A_c представляет собой процент важных узлов (0 и 33) в двух лучших узлах, определённых рассматриваемыми мерами центральности. Также сравнивается точность A_c предложенных мер центральности с точностью A_c классических мер центральности (степень посредничества и степень близости).

Результаты этой части экспериментов представлены в таблицах 16–23. В таблице 16 приведены результаты для графов с 20 узлами и разной плотностью (от 0.1 до 1.0), для взвешенных и невзвешенных графов. Сравнивается время вычисления точного вектора Шепли (столбец “SV time”) и аппроксимированного вектора Шепли (столбец “ASV time”). В таблице 16 также представлена среднеквадратичная ошибка для аппроксимированного вектора Шепли (столбец “RMSE ASV”) и для отношения аппроксимированной компоненты вектора Шепли к сумме всех его компонент (столбец “RMSE ratio ASV”). Чем меньше значение RMSE, тем точнее получается аппроксимация. На основе результатов таблицы 16 можно сделать следующие выводы:

1. Время вычисления аппроксимированного вектора Шепли значительно меньше, чем время вычисления точного вектора Шепли (см. столбцы “SV time” и “ASV time”).
2. Среднеквадратичная ошибка аппроксимированного вектора Шепли очень мала (см. столбец “RMSE ASV”), а среднеквадратичная ошибка отношения ещё меньше, чем первая (см. столбец “RMSE ratio ASV”). Здесь под отношением подразумеваются нормализованные значения. RMSE для отношения — это RMSE между точными нормализованными векторами Шепли и аппроксимированными нормализованными векторами Шепли. Значение RMSE увеличивается с ростом плотности графа из-за процедуры выборки.
3. Взвешенный граф вносит большую неопределенность в процесс выборки при высокой плотности графа (≥ 0.7), что приводит к несколько большему значению RMSE. Однако, даже в худшем случае (см. граф “20-1.0” в таблице 16), наш алгоритм всё равно демонстрирует высокую точность.

Таблица 16: Результаты для вектора Шепли при применении масштабирующего коэффициента ξ

| Graph | Weighted | SV time | ASV time | RMSE ASV | RMSE ratio ASV |
|--------|----------|------------------|---------------------|----------------------|----------------------|
| 20-0.1 | True | $7.2 \cdot 10^2$ | $2.2 \cdot 10^{-3}$ | $2.6 \cdot 10^{-22}$ | $5.5 \cdot 10^{-27}$ |
| 20-0.2 | True | $8.7 \cdot 10^2$ | $1.9 \cdot 10^{-2}$ | $3.1 \cdot 10^{-22}$ | $5.6 \cdot 10^{-27}$ |
| 20-0.3 | True | $1 \cdot 10^3$ | $1.5 \cdot 10^{-1}$ | $1.6 \cdot 10^{-22}$ | $1.2 \cdot 10^{-27}$ |
| 20-0.4 | True | $1.2 \cdot 10^3$ | $9.3 \cdot 10^{-1}$ | $1.7 \cdot 10^{-22}$ | $6.1 \cdot 10^{-28}$ |
| 20-0.5 | True | $1.3 \cdot 10^3$ | $2.2 \cdot 10^0$ | $7.1 \cdot 10^{-23}$ | $2 \cdot 10^{-28}$ |
| 20-0.6 | True | $1.5 \cdot 10^3$ | $1.9 \cdot 10^1$ | $7.7 \cdot 10^{-23}$ | $1.8 \cdot 10^{-28}$ |
| 20-0.7 | True | $1.6 \cdot 10^3$ | $4.7 \cdot 10^1$ | $3.3 \cdot 10^{-2}$ | $6 \cdot 10^{-8}$ |
| 20-0.8 | True | $1.7 \cdot 10^3$ | $1.7 \cdot 10^2$ | $2 \cdot 10^0$ | $2.9 \cdot 10^{-6}$ |
| 20-0.9 | True | $1.8 \cdot 10^3$ | $2.5 \cdot 10^2$ | $6.9 \cdot 10^0$ | $7.9 \cdot 10^{-6}$ |
| 20-1.0 | True | $1.9 \cdot 10^3$ | $3.2 \cdot 10^2$ | $9.9 \cdot 10^0$ | $9 \cdot 10^{-6}$ |
| 20-0.1 | False | $7.2 \cdot 10^2$ | $3.4 \cdot 10^{-3}$ | $4.5 \cdot 10^{-23}$ | $1.1 \cdot 10^{-26}$ |
| 20-0.2 | False | $8.8 \cdot 10^2$ | $1.6 \cdot 10^{-2}$ | $1.1 \cdot 10^{-22}$ | $7.8 \cdot 10^{-27}$ |
| 20-0.3 | False | $1 \cdot 10^3$ | $9.6 \cdot 10^{-2}$ | $1.1 \cdot 10^{-22}$ | $6.6 \cdot 10^{-27}$ |
| 20-0.4 | False | $1.2 \cdot 10^3$ | $5.5 \cdot 10^{-1}$ | $1.5 \cdot 10^{-22}$ | $1.6 \cdot 10^{-26}$ |
| 20-0.5 | False | $1.3 \cdot 10^3$ | $2.9 \cdot 10^0$ | $1.1 \cdot 10^{-22}$ | $8.6 \cdot 10^{-27}$ |
| 20-0.6 | False | $1.5 \cdot 10^3$ | $2.5 \cdot 10^1$ | $1.4 \cdot 10^{-22}$ | $9.4 \cdot 10^{-27}$ |
| 20-0.7 | False | $1.6 \cdot 10^3$ | $6.2 \cdot 10^1$ | $3.8 \cdot 10^{-4}$ | $2.1 \cdot 10^{-8}$ |
| 20-0.8 | False | $1.7 \cdot 10^3$ | $1.7 \cdot 10^2$ | $1.8 \cdot 10^{-3}$ | $7.7 \cdot 10^{-8}$ |
| 20-0.9 | False | $1.8 \cdot 10^3$ | $2.5 \cdot 10^2$ | $2 \cdot 10^{-3}$ | $6.9 \cdot 10^{-8}$ |
| 20-1.0 | False | $1.9 \cdot 10^3$ | $3.2 \cdot 10^2$ | $4.1 \cdot 10^{-21}$ | $8.6 \cdot 10^{-34}$ |

В таблицах 17 и 18 представлены результаты вычисления значения Майерсона (точного и аппроксимированного) без и с применением масштабирующего коэффициента ξ , определённого в (31), соответственно. В обеих таблицах приведены результаты для графов с 20 узлами и различной плотностью (от 0.1 до 0.2), для взвешенных и невзвешенных графов. Сравнивается время вычисления точного значения Майерсона (столбец “MV time”) и аппроксимированного значения Майерсона (столбец “AMV time”). В этих двух таблицах также представлена среднеквадратичная ошибка для аппроксимированного значения Майерсона (столбец “RMSE AMV”) и для отношения аппроксимированной компоненты значения Майерсона к сумме всех её компонент (столбец “RMSE ratio AMV”). Выводы из таблиц 17 и 18 следующие:

1. Вычислительная сложность центральности на основе точного значения Майерсона быстро растёт с увеличением плотности сети (см. столбец “MV time”).
2. Среднеквадратичная ошибка аппроксимированного значения Майерсона увеличивается с ростом плотности сети (см. столбец “RMSE AMV”), но среднеквадратичная ошибка отношения остаётся очень маленькой. RMSE для отношения — это среднеквадратичная ошибка между нормализованными аппроксимированными и нормализованными точными значениями Майерсона. Поэтому можно рекомендовать использовать аппроксимацию значения Майерсона в качестве меры центральности.

3. Если оценить влияние масштабирующего коэффициента ξ на результаты, можно сравнить “AMV time” в таблицах 17 и 18 и сделать вывод, что без ξ время вычисления значительно меньше, но RMSE AMV выше. В то время как значения RMSE ratio AMV в таблицах 17 и 18 практически одинаковы.
4. Сравнивая таблицы 17 и 18, нельзя заметить значительных различий в результатах для взвешенных и невзвешенных графов.

Таблица 17: Результаты для значения Майерсона без масштабирования с коэффициентом ξ

| Graph | Weighted | MV time | AMV time | RMSE AMV | RMSE ratio AMV |
|---------|----------|----------------------|----------------------|----------------------|----------------------|
| 20-0.1 | True | $1.03 \cdot 10^{-1}$ | $1.62 \cdot 10^{-1}$ | $4.37 \cdot 10^{-4}$ | $1.48 \cdot 10^{-8}$ |
| 20-0.11 | True | $1.91 \cdot 10^{-1}$ | $1.67 \cdot 10^{-1}$ | $8.75 \cdot 10^{-4}$ | $1.7 \cdot 10^{-8}$ |
| 20-0.12 | True | $4.17 \cdot 10^{-1}$ | $2.23 \cdot 10^{-1}$ | $1.54 \cdot 10^{-1}$ | $5.28 \cdot 10^{-7}$ |
| 20-0.13 | True | $1.02 \cdot 10^0$ | $3.56 \cdot 10^{-1}$ | $5.28 \cdot 10^{-1}$ | $8.15 \cdot 10^{-7}$ |
| 20-0.14 | True | $2.21 \cdot 10^0$ | $6.15 \cdot 10^{-1}$ | $2.14 \cdot 10^0$ | $1.93 \cdot 10^{-6}$ |
| 20-0.15 | True | $8.78 \cdot 10^0$ | $9.24 \cdot 10^{-1}$ | $1.89 \cdot 10^1$ | $4.3 \cdot 10^{-6}$ |
| 20-0.16 | True | $1.8 \cdot 10^1$ | $1.48 \cdot 10^0$ | $1.27 \cdot 10^2$ | $1.25 \cdot 10^{-5}$ |
| 20-0.17 | True | $2.87 \cdot 10^1$ | $2.14 \cdot 10^0$ | $4.06 \cdot 10^2$ | $3.24 \cdot 10^{-5}$ |
| 20-0.18 | True | $6.34 \cdot 10^1$ | $2.77 \cdot 10^0$ | $1.55 \cdot 10^3$ | $4.7 \cdot 10^{-5}$ |
| 20-0.19 | True | $1.45 \cdot 10^2$ | $3.58 \cdot 10^0$ | $4.82 \cdot 10^3$ | $8.66 \cdot 10^{-5}$ |
| 20-0.2 | True | $2.94 \cdot 10^2$ | $4.45 \cdot 10^0$ | $1.33 \cdot 10^4$ | $1.35 \cdot 10^{-4}$ |
| 20-0.1 | False | $3.3 \cdot 10^{-1}$ | $1.96 \cdot 10^{-1}$ | $1.12 \cdot 10^{-2}$ | $1.43 \cdot 10^{-6}$ |
| 20-0.11 | False | $3.45 \cdot 10^{-1}$ | $2.16 \cdot 10^{-1}$ | $1.73 \cdot 10^{-2}$ | $1.41 \cdot 10^{-6}$ |
| 20-0.12 | False | $4.87 \cdot 10^{-1}$ | $2.17 \cdot 10^{-1}$ | $3.43 \cdot 10^{-2}$ | $1.52 \cdot 10^{-6}$ |
| 20-0.13 | False | $1.37 \cdot 10^0$ | $3.79 \cdot 10^{-1}$ | $3.95 \cdot 10^{-1}$ | $5.73 \cdot 10^{-6}$ |
| 20-0.14 | False | $3.69 \cdot 10^0$ | $5.81 \cdot 10^{-1}$ | $2.71 \cdot 10^0$ | $1.53 \cdot 10^{-5}$ |
| 20-0.15 | False | $7.65 \cdot 10^0$ | $1.05 \cdot 10^0$ | $1.38 \cdot 10^1$ | $2.75 \cdot 10^{-5}$ |
| 20-0.16 | False | $1.71 \cdot 10^1$ | $1.38 \cdot 10^0$ | $5.94 \cdot 10^1$ | $4.19 \cdot 10^{-5}$ |
| 20-0.17 | False | $4.28 \cdot 10^1$ | $1.62 \cdot 10^0$ | $2.97 \cdot 10^2$ | $6.29 \cdot 10^{-5}$ |
| 20-0.18 | False | $8.25 \cdot 10^1$ | $1.88 \cdot 10^0$ | $8.04 \cdot 10^2$ | $6.73 \cdot 10^{-5}$ |
| 20-0.19 | False | $1.69 \cdot 10^2$ | $2.56 \cdot 10^0$ | $2.72 \cdot 10^3$ | $7.6 \cdot 10^{-5}$ |
| 20-0.2 | False | $3.66 \cdot 10^2$ | $3.57 \cdot 10^0$ | $8.75 \cdot 10^3$ | $1.31 \cdot 10^{-4}$ |

Далее демонстрируется эффективность нашего алгоритма для аппроксимации значения Майерсона в графах с высокой плотностью. Для этого рассматриваются графы с 10 узлами и плотностью от 0.05 до 1.0 с шагом 0.05, для взвешенных и невзвешенных графов. Результаты экспериментов представлены в таблицах 19 и 20 для случаев без и с применением масштабирующего коэффициента, соответственно.

Сравнивая пары таблиц 17 и 18, а также таблиц 19 и 20, можно заключить, что масштабирующий коэффициент ξ оказывает очень ограниченное влияние на уменьшение ошибки аппроксимированного значения Майерсона, и, более того, использование ξ значительно увеличивает время вычисления. Однако RMSE для отношения остаётся очень малым. Поэтому можно рекомендовать использовать аппроксимацию центральности на основе значения Майерсона без

Таблица 18: Результаты для значения Майерсона при применении масштабирования с коэффициентом ξ

| Graph | Weighted | MV time | AMV time | RMSE AMV | RMSE ratio AMV |
|---------|----------|----------------------|----------------------|----------------------|----------------------|
| 20-0.1 | True | $1.23 \cdot 10^{-1}$ | $1.1 \cdot 10^{-1}$ | $1.98 \cdot 10^{-4}$ | $1.48 \cdot 10^{-8}$ |
| 20-0.11 | True | $1.57 \cdot 10^{-1}$ | $2.11 \cdot 10^{-1}$ | $3.21 \cdot 10^{-4}$ | $1.7 \cdot 10^{-8}$ |
| 20-0.12 | True | $4.53 \cdot 10^{-1}$ | $2.21 \cdot 10^{-1}$ | $1.55 \cdot 10^{-2}$ | $5.28 \cdot 10^{-7}$ |
| 20-0.13 | True | $9.1 \cdot 10^{-1}$ | $4.33 \cdot 10^{-1}$ | $3.15 \cdot 10^{-2}$ | $8.15 \cdot 10^{-7}$ |
| 20-0.14 | True | $2.09 \cdot 10^0$ | $6.7 \cdot 10^{-1}$ | $1.08 \cdot 10^{-1}$ | $1.93 \cdot 10^{-6}$ |
| 20-0.15 | True | $7.83 \cdot 10^0$ | $1.24 \cdot 10^0$ | $4.81 \cdot 10^{-1}$ | $4.3 \cdot 10^{-6}$ |
| 20-0.16 | True | $1.65 \cdot 10^1$ | $2.24 \cdot 10^0$ | $3.85 \cdot 10^0$ | $1.25 \cdot 10^{-5}$ |
| 20-0.17 | True | $2.63 \cdot 10^1$ | $3.28 \cdot 10^0$ | $2.36 \cdot 10^1$ | $3.24 \cdot 10^{-5}$ |
| 20-0.18 | True | $5.91 \cdot 10^1$ | $5.91 \cdot 10^0$ | $6.15 \cdot 10^1$ | $4.7 \cdot 10^{-5}$ |
| 20-0.19 | True | $1.41 \cdot 10^2$ | $1.15 \cdot 10^1$ | $2.07 \cdot 10^2$ | $8.66 \cdot 10^{-5}$ |
| 20-0.2 | True | $2.92 \cdot 10^2$ | $2.06 \cdot 10^1$ | $6.48 \cdot 10^2$ | $1.35 \cdot 10^{-4}$ |
| 20-0.1 | False | $2.75 \cdot 10^{-1}$ | $2.43 \cdot 10^{-1}$ | $1.59 \cdot 10^{-3}$ | $1.43 \cdot 10^{-6}$ |
| 20-0.11 | False | $3.64 \cdot 10^{-1}$ | $1.94 \cdot 10^{-1}$ | $1.81 \cdot 10^{-3}$ | $1.41 \cdot 10^{-6}$ |
| 20-0.12 | False | $4.56 \cdot 10^{-1}$ | $2.47 \cdot 10^{-1}$ | $2.65 \cdot 10^{-3}$ | $1.52 \cdot 10^{-6}$ |
| 20-0.13 | False | $1.34 \cdot 10^0$ | $4.13 \cdot 10^{-1}$ | $2.01 \cdot 10^{-2}$ | $5.73 \cdot 10^{-6}$ |
| 20-0.14 | False | $3.33 \cdot 10^0$ | $6.97 \cdot 10^{-1}$ | $1.18 \cdot 10^{-1}$ | $1.53 \cdot 10^{-5}$ |
| 20-0.15 | False | $7.17 \cdot 10^0$ | $1.29 \cdot 10^0$ | $5.64 \cdot 10^{-1}$ | $2.75 \cdot 10^{-5}$ |
| 20-0.16 | False | $1.55 \cdot 10^1$ | $2.01 \cdot 10^0$ | $1.92 \cdot 10^0$ | $4.19 \cdot 10^{-5}$ |
| 20-0.17 | False | $3.93 \cdot 10^1$ | $3.3 \cdot 10^0$ | $7.44 \cdot 10^0$ | $6.29 \cdot 10^{-5}$ |
| 20-0.18 | False | $7.78 \cdot 10^1$ | $5.58 \cdot 10^0$ | $1.6 \cdot 10^1$ | $6.73 \cdot 10^{-5}$ |
| 20-0.19 | False | $1.64 \cdot 10^2$ | $1.06 \cdot 10^1$ | $4.58 \cdot 10^1$ | $7.6 \cdot 10^{-5}$ |
| 20-0.2 | False | $3.43 \cdot 10^2$ | $1.99 \cdot 10^1$ | $2.06 \cdot 10^2$ | $1.31 \cdot 10^{-4}$ |

масштабирующего коэффициента ξ , а лучше использовать его отношение в качестве аппроксимации значения Майерсона.

В таблице 16 результаты для взвешенного графа с плотностью 1.0 показывают наибольшее значение RMSE, а в таблице 18 результаты для взвешенного графа с плотностью 0.2 также показывают наибольшее значение RMSE. Поэтому принято решение тщательно рассмотреть случай с высокой плотностью и исследовать вектор Шепли для взвешенного графа с плотностью 1.0 и значение Майерсона для взвешенного графа с плотностью 0.2. Для анализа представлены компоненты точного и аппроксимированного вектора Шепли в таблице 21 для первого графа и точные и аппроксимированные значения Майерсона в таблице 22 для второго графа. В этих таблицах приведена относительная ошибка, которая рассчитывается как разница между аппроксимированным значением и точным значением, делённая на точное значение, то есть $\frac{\text{approximated value} - \text{exact value}}{\text{exact value}}$.

Анализируя таблицу 21, можно сказать, что наш алгоритм обеспечивает очень точную аппроксимацию вектора Шепли, так что лишь для двух узлов (узлы 5 и 19) абсолютная относительная ошибка превышает 5%.

Анализируя результаты в таблице 22, можно заметить, что некоторые узлы (узлы 1, 2, 8 и 9) имеют высокую относительную ошибку в аппроксимации. Причина заключается в том, что в нашем алгоритме подсчитывается количе-

Таблица 19: Результаты для значения Майерсона без масштабирования с коэффициентом ξ для сетей с 10 узлами

| Graph | Weighted | MV time | AMV time | RMSE AMV | RMSE ratio AMV |
|---------|----------|----------------------|----------------------|----------------------|-----------------------|
| 10-0.05 | True | $1.24 \cdot 10^{-3}$ | $1.59 \cdot 10^{-2}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.1 | True | $1.79 \cdot 10^{-3}$ | $1.58 \cdot 10^{-3}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.15 | True | $1.27 \cdot 10^{-2}$ | $3.17 \cdot 10^{-3}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.2 | True | $7.33 \cdot 10^{-3}$ | $1.05 \cdot 10^{-2}$ | $9.65 \cdot 10^{-6}$ | $7.63 \cdot 10^{-10}$ |
| 10-0.25 | True | $1.49 \cdot 10^{-2}$ | $2.89 \cdot 10^{-3}$ | $3.27 \cdot 10^1$ | $1.64 \cdot 10^{-4}$ |
| 10-0.3 | True | $6.69 \cdot 10^{-2}$ | $3.35 \cdot 10^{-3}$ | $2.79 \cdot 10^2$ | $1.91 \cdot 10^{-4}$ |
| 10-0.35 | True | $1.17 \cdot 10^{-1}$ | $1.78 \cdot 10^{-3}$ | $1.85 \cdot 10^3$ | $5.12 \cdot 10^{-4}$ |
| 10-0.4 | True | $2.6 \cdot 10^{-1}$ | $1.2 \cdot 10^{-3}$ | $4.58 \cdot 10^3$ | $4.08 \cdot 10^{-4}$ |
| 10-0.45 | True | $5.53 \cdot 10^{-1}$ | $1.21 \cdot 10^{-3}$ | $1.05 \cdot 10^4$ | $6.26 \cdot 10^{-4}$ |
| 10-0.5 | True | $1.21 \cdot 10^0$ | $1.37 \cdot 10^{-3}$ | $1.56 \cdot 10^4$ | $5.53 \cdot 10^{-4}$ |
| 10-0.55 | True | $2.56 \cdot 10^0$ | $1.26 \cdot 10^{-3}$ | $3.92 \cdot 10^4$ | $4.49 \cdot 10^{-4}$ |
| 10-0.6 | True | $5.7 \cdot 10^0$ | $1.62 \cdot 10^{-3}$ | $1.11 \cdot 10^5$ | $6.17 \cdot 10^{-4}$ |
| 10-0.65 | True | $1.05 \cdot 10^1$ | $1.33 \cdot 10^{-3}$ | $2.31 \cdot 10^5$ | $7.47 \cdot 10^{-4}$ |
| 10-0.7 | True | $1.99 \cdot 10^1$ | $1.33 \cdot 10^{-3}$ | $5.19 \cdot 10^5$ | $6.84 \cdot 10^{-4}$ |
| 10-0.75 | True | $3.55 \cdot 10^1$ | $6.44 \cdot 10^{-3}$ | $1 \cdot 10^6$ | $5.96 \cdot 10^{-4}$ |
| 10-0.8 | True | $7.36 \cdot 10^1$ | $1.38 \cdot 10^{-3}$ | $3.42 \cdot 10^6$ | $5.9 \cdot 10^{-4}$ |
| 10-0.85 | True | $1.15 \cdot 10^2$ | $1.38 \cdot 10^{-3}$ | $8.11 \cdot 10^6$ | $3.76 \cdot 10^{-4}$ |
| 10-0.9 | True | $1.73 \cdot 10^2$ | $1.39 \cdot 10^{-3}$ | $1.63 \cdot 10^7$ | $3.92 \cdot 10^{-4}$ |
| 10-0.95 | True | $2.48 \cdot 10^2$ | $1.37 \cdot 10^{-3}$ | $1.37 \cdot 10^7$ | $2.74 \cdot 10^{-4}$ |
| 10-1.0 | True | $4.33 \cdot 10^2$ | $1.37 \cdot 10^{-3}$ | $7.31 \cdot 10^7$ | $2.79 \cdot 10^{-4}$ |
| 10-0.05 | False | $1.15 \cdot 10^{-3}$ | $1.06 \cdot 10^{-3}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.1 | False | $2.09 \cdot 10^{-3}$ | $1.99 \cdot 10^{-3}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.15 | False | $3.4 \cdot 10^{-3}$ | $2.84 \cdot 10^{-3}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.2 | False | $7.57 \cdot 10^{-3}$ | $1.08 \cdot 10^{-2}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.25 | False | $3.14 \cdot 10^{-2}$ | $8.37 \cdot 10^{-3}$ | $2.56 \cdot 10^0$ | $1.57 \cdot 10^{-4}$ |
| 10-0.3 | False | $3.12 \cdot 10^{-2}$ | $3.11 \cdot 10^{-3}$ | $8.85 \cdot 10^0$ | $9.04 \cdot 10^{-5}$ |
| 10-0.35 | False | $8.49 \cdot 10^{-2}$ | $1.13 \cdot 10^{-3}$ | $5.98 \cdot 10^1$ | $1.53 \cdot 10^{-4}$ |
| 10-0.4 | False | $2.18 \cdot 10^{-1}$ | $1.33 \cdot 10^{-3}$ | $2.28 \cdot 10^2$ | $1.29 \cdot 10^{-4}$ |
| 10-0.45 | False | $4.74 \cdot 10^{-1}$ | $1.16 \cdot 10^{-3}$ | $6.23 \cdot 10^2$ | $9.58 \cdot 10^{-5}$ |
| 10-0.5 | False | $9.61 \cdot 10^{-1}$ | $1.17 \cdot 10^{-3}$ | $1.81 \cdot 10^3$ | $8.71 \cdot 10^{-5}$ |
| 10-0.55 | False | $1.95 \cdot 10^0$ | $1.21 \cdot 10^{-3}$ | $4.91 \cdot 10^3$ | $7.4 \cdot 10^{-5}$ |
| 10-0.6 | False | $4.15 \cdot 10^0$ | $1.3 \cdot 10^{-3}$ | $1.97 \cdot 10^4$ | $4.98 \cdot 10^{-5}$ |
| 10-0.65 | False | $7.11 \cdot 10^0$ | $1.24 \cdot 10^{-3}$ | $4.79 \cdot 10^4$ | $5.33 \cdot 10^{-5}$ |
| 10-0.7 | False | $1.14 \cdot 10^1$ | $1.33 \cdot 10^{-3}$ | $1.07 \cdot 10^5$ | $9.37 \cdot 10^{-5}$ |
| 10-0.75 | False | $1.86 \cdot 10^1$ | $3.51 \cdot 10^{-3}$ | $2.46 \cdot 10^5$ | $1.6 \cdot 10^{-4}$ |
| 10-0.8 | False | $3.77 \cdot 10^1$ | $1.39 \cdot 10^{-3}$ | $8.53 \cdot 10^5$ | $1.07 \cdot 10^{-4}$ |
| 10-0.85 | False | $5.26 \cdot 10^1$ | $1.29 \cdot 10^{-3}$ | $1.73 \cdot 10^6$ | $1.35 \cdot 10^{-4}$ |
| 10-0.9 | False | $7.45 \cdot 10^1$ | $1.31 \cdot 10^{-3}$ | $3.64 \cdot 10^6$ | $9.85 \cdot 10^{-5}$ |
| 10-0.95 | False | $1.05 \cdot 10^2$ | $1.3 \cdot 10^{-3}$ | $7.74 \cdot 10^6$ | $4.85 \cdot 10^{-5}$ |
| 10-1.0 | False | $1.71 \cdot 10^2$ | $1.28 \cdot 10^{-3}$ | $2.16 \cdot 10^7$ | $0 \cdot 10^0$ |

ство путей не для всех длин ($length \leq L$). Увеличив L , можно получить более точный результат, но это увеличит вычислительную сложность.

Последний пункт в анализе этой серии экспериментов заключается в сравнении точности, определённой по формуле (32), для аппроксимированных векторов Шепли и значений Майерсона, а также для классических мер центральности (посредничества и близости). Эти результаты приведены в таблице 23. Обе меры центральности, предложенные в этой работе, имеют более высокую точность по сравнению с классическими мерами центральности. В частности, аппроксимированная центральность на основе вектора Шепли может определить наиболее важные узлы с точностью 100% для всех исследованных графов. Для аппроксимированной центральности на основе значения Майерсона точность составляет 100% для всех случаев, за исключением сети “karate-twoStar-

Таблица 20: Результаты для значения Майерсона при применении масштабирования с коэффициентом ξ для сетей с 10 узлами

| Graph | Weighted | MV time | AMV time | RMSE AMV | RMSE ratio AMV |
|---------|----------|----------------------|----------------------|-----------------------|-----------------------|
| 10-0.05 | True | $1.24 \cdot 10^{-3}$ | $1.22 \cdot 10^{-3}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.1 | True | $3.08 \cdot 10^{-3}$ | $1.99 \cdot 10^{-3}$ | $6.24 \cdot 10^{-32}$ | $0 \cdot 10^0$ |
| 10-0.15 | True | $4.67 \cdot 10^{-3}$ | $3.73 \cdot 10^{-3}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.2 | True | $1.2 \cdot 10^{-2}$ | $1.74 \cdot 10^{-2}$ | $3.25 \cdot 10^{-6}$ | $7.63 \cdot 10^{-10}$ |
| 10-0.25 | True | $1.46 \cdot 10^{-2}$ | $4.48 \cdot 10^{-3}$ | $1.67 \cdot 10^0$ | $1.64 \cdot 10^{-4}$ |
| 10-0.3 | True | $4.97 \cdot 10^{-2}$ | $8.24 \cdot 10^{-3}$ | $6.05 \cdot 10^0$ | $1.91 \cdot 10^{-4}$ |
| 10-0.35 | True | $1.05 \cdot 10^{-1}$ | $1.69 \cdot 10^{-2}$ | $3.5 \cdot 10^1$ | $5.12 \cdot 10^{-4}$ |
| 10-0.4 | True | $2.57 \cdot 10^{-1}$ | $2.63 \cdot 10^{-2}$ | $6.41 \cdot 10^1$ | $4.08 \cdot 10^{-4}$ |
| 10-0.45 | True | $5.52 \cdot 10^{-1}$ | $7.07 \cdot 10^{-2}$ | $2.14 \cdot 10^2$ | $6.26 \cdot 10^{-4}$ |
| 10-0.5 | True | $1.18 \cdot 10^0$ | $1.27 \cdot 10^{-1}$ | $2.91 \cdot 10^2$ | $5.53 \cdot 10^{-4}$ |
| 10-0.55 | True | $2.5 \cdot 10^0$ | $3.8 \cdot 10^{-1}$ | $5.78 \cdot 10^2$ | $4.49 \cdot 10^{-4}$ |
| 10-0.6 | True | $5.81 \cdot 10^0$ | $5.76 \cdot 10^{-1}$ | $2.15 \cdot 10^3$ | $6.17 \cdot 10^{-4}$ |
| 10-0.65 | True | $1.06 \cdot 10^1$ | $1.24 \cdot 10^0$ | $5.36 \cdot 10^3$ | $7.47 \cdot 10^{-4}$ |
| 10-0.7 | True | $1.98 \cdot 10^1$ | $2.14 \cdot 10^0$ | $1.12 \cdot 10^4$ | $6.84 \cdot 10^{-4}$ |
| 10-0.75 | True | $3.57 \cdot 10^1$ | $3.66 \cdot 10^0$ | $1.9 \cdot 10^4$ | $5.96 \cdot 10^{-4}$ |
| 10-0.8 | True | $7.43 \cdot 10^1$ | $6.77 \cdot 10^0$ | $6.41 \cdot 10^4$ | $5.9 \cdot 10^{-4}$ |
| 10-0.85 | True | $1.16 \cdot 10^2$ | $1.11 \cdot 10^1$ | $9.71 \cdot 10^4$ | $3.76 \cdot 10^{-4}$ |
| 10-0.9 | True | $1.75 \cdot 10^2$ | $1.71 \cdot 10^1$ | $2.03 \cdot 10^5$ | $3.92 \cdot 10^{-4}$ |
| 10-0.95 | True | $2.48 \cdot 10^2$ | $2.45 \cdot 10^1$ | $3.22 \cdot 10^5$ | $3.7 \cdot 10^{-4}$ |
| 10-1.0 | True | $4.33 \cdot 10^2$ | $4.33 \cdot 10^1$ | $6.49 \cdot 10^5$ | $2.79 \cdot 10^{-4}$ |
| 10-0.05 | False | $1.23 \cdot 10^{-3}$ | $1.29 \cdot 10^{-3}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.1 | False | $2.28 \cdot 10^{-3}$ | $2.55 \cdot 10^{-3}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.15 | False | $4.91 \cdot 10^{-3}$ | $1.73 \cdot 10^{-2}$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |
| 10-0.2 | False | $2.15 \cdot 10^{-2}$ | $7.96 \cdot 10^{-3}$ | $2.16 \cdot 10^{-31}$ | $2.65 \cdot 10^{-33}$ |
| 10-0.25 | False | $1.64 \cdot 10^{-2}$ | $9.18 \cdot 10^{-3}$ | $5.35 \cdot 10^{-2}$ | $1.57 \cdot 10^{-4}$ |
| 10-0.3 | False | $4.56 \cdot 10^{-2}$ | $6.48 \cdot 10^{-3}$ | $7.23 \cdot 10^{-2}$ | $9.04 \cdot 10^{-5}$ |
| 10-0.35 | False | $6.45 \cdot 10^{-2}$ | $6.84 \cdot 10^{-3}$ | $3.23 \cdot 10^{-1}$ | $1.53 \cdot 10^{-4}$ |
| 10-0.4 | False | $2.43 \cdot 10^{-1}$ | $1.99 \cdot 10^{-2}$ | $1.02 \cdot 10^0$ | $1.29 \cdot 10^{-4}$ |
| 10-0.45 | False | $4.54 \cdot 10^{-1}$ | $3.73 \cdot 10^{-2}$ | $2 \cdot 10^0$ | $9.58 \cdot 10^{-5}$ |
| 10-0.5 | False | $9.12 \cdot 10^{-1}$ | $8.23 \cdot 10^{-2}$ | $5.1 \cdot 10^0$ | $8.71 \cdot 10^{-5}$ |
| 10-0.55 | False | $1.86 \cdot 10^0$ | $2.11 \cdot 10^{-1}$ | $1.15 \cdot 10^1$ | $7.4 \cdot 10^{-5}$ |
| 10-0.6 | False | $4.29 \cdot 10^0$ | $3.61 \cdot 10^{-1}$ | $3.03 \cdot 10^1$ | $4.98 \cdot 10^{-5}$ |
| 10-0.65 | False | $7.36 \cdot 10^0$ | $6.8 \cdot 10^{-1}$ | $7.83 \cdot 10^1$ | $5.33 \cdot 10^{-5}$ |
| 10-0.7 | False | $1.14 \cdot 10^1$ | $1.51 \cdot 10^0$ | $3.05 \cdot 10^2$ | $9.37 \cdot 10^{-5}$ |
| 10-0.75 | False | $1.94 \cdot 10^1$ | $2.03 \cdot 10^0$ | $1.23 \cdot 10^3$ | $1.6 \cdot 10^{-4}$ |
| 10-0.8 | False | $3.74 \cdot 10^1$ | $3.41 \cdot 10^0$ | $2.9 \cdot 10^3$ | $1.07 \cdot 10^{-4}$ |
| 10-0.85 | False | $5.27 \cdot 10^1$ | $5.04 \cdot 10^0$ | $7.35 \cdot 10^3$ | $1.35 \cdot 10^{-4}$ |
| 10-0.9 | False | $7.45 \cdot 10^1$ | $6.5 \cdot 10^0$ | $1.13 \cdot 10^4$ | $9.85 \cdot 10^{-5}$ |
| 10-0.95 | False | $1.05 \cdot 10^2$ | $9.53 \cdot 10^0$ | $1.19 \cdot 10^4$ | $4.85 \cdot 10^{-5}$ |
| 10-1.0 | False | $1.71 \cdot 10^2$ | $1.63 \cdot 10^1$ | $0 \cdot 10^0$ | $0 \cdot 10^0$ |

34", для которой точность составляет 50%. Ошибка в предсказании одного из самых важных узлов заключается в том, что другой узел (узел 17) является центральным в внутреннем слое по своей конструкции. Для степени посредничества точность равна нулю для сетей с высокой плотностью. Степень близости также показывает худшую точность для этих сетей.

Таблица 21: Точный вектор Шепли против аппроксимированного вектора Шепли для “graph-1.0”

| Nodes | SV | SV ratio | ASV | ASV ratio | Rel. error ASV/SV |
|-------|---------|----------|---------|-----------|------------------------|
| 0 | 52.0000 | 4.96% | 53.3579 | 5.09% | 2.61% |
| 1 | 46.5000 | 4.43% | 46.5090 | 4.43% | $1.83 \cdot 10^{-2}\%$ |
| 2 | 50.5000 | 4.81% | 49.0574 | 4.68% | -2.86% |
| 3 | 53.0000 | 5.05% | 53.6764 | 5.12% | 1.28% |
| 4 | 50.5000 | 4.81% | 50.0131 | 4.77% | -0.96% |
| 5 | 50.0000 | 4.77% | 54.3135 | 5.18% | 8.63% |
| 6 | 50.0000 | 4.77% | 51.2873 | 4.89% | 2.57% |
| 7 | 63.0000 | 6.01% | 62.5960 | 5.97% | -0.64% |
| 8 | 55.0000 | 5.24% | 53.3579 | 5.09% | -2.98% |
| 9 | 44.0000 | 4.19% | 44.4384 | 4.24% | 0.99% |
| 10 | 49.5000 | 4.72% | 50.3316 | 4.80% | 1.68% |
| 11 | 51.0000 | 4.86% | 51.6058 | 4.92% | 1.19% |
| 12 | 64.0000 | 6.10% | 62.9145 | 6.00% | -1.70% |
| 13 | 49.0000 | 4.67% | 49.0574 | 4.68% | 0.12% |
| 14 | 48.0000 | 4.58% | 47.9425 | 4.57% | -0.12% |
| 15 | 63.5000 | 6.05% | 63.0738 | 6.01% | -0.67% |
| 16 | 54.0000 | 5.15% | 54.4728 | 5.19% | 0.87% |
| 17 | 57.0000 | 5.43% | 57.0213 | 5.44% | $3.66 \cdot 10^{-2}\%$ |
| 18 | 53.0000 | 5.05% | 52.0837 | 4.97% | -1.73% |
| 19 | 45.5000 | 4.34% | 41.8899 | 3.99% | -7.93% |

Таблица 22: Точное значение Майерсона против аппроксимированного значения Майерсона для “graph-0.2”

| Nodes | MV | MV ratio | AMV | AMV ratio | Rel. error AMV/MV |
|-------|----------|----------|----------|-----------|-------------------|
| 0 | 143.2672 | 6.53% | 144.0222 | 6.56% | 0.53% |
| 1 | 68.7002 | 3.13% | 49.3370 | 2.25% | -28.18% |
| 2 | 68.0767 | 3.10% | 81.6335 | 3.72% | 19.91% |
| 3 | 104.6799 | 4.77% | 109.9763 | 5.01% | 5.06% |
| 4 | 164.0880 | 7.48% | 157.1853 | 7.16% | -4.21% |
| 5 | 114.6388 | 5.22% | 106.1499 | 4.84% | -7.40% |
| 6 | 16.7705 | 0.76% | 17.8578 | 0.81% | 6.48% |
| 7 | 228.8444 | 10.43% | 251.4307 | 11.46% | 9.87% |
| 8 | 33.7682 | 1.54% | 24.3128 | 1.11% | -28.00% |
| 9 | 36.1832 | 1.65% | 28.1827 | 1.28% | -22.11% |
| 10 | 184.1740 | 8.39% | 192.6532 | 8.78% | 4.60% |
| 11 | 79.7901 | 3.64% | 69.8602 | 3.18% | -12.45% |
| 12 | 56.7906 | 2.59% | 53.0066 | 2.42% | -6.66% |
| 13 | 48.8678 | 2.23% | 46.0495 | 2.10% | -5.77% |
| 14 | 107.3196 | 4.89% | 93.9069 | 4.28% | -12.50% |
| 15 | 63.4039 | 2.89% | 66.5861 | 3.03% | 5.02% |
| 16 | 186.6146 | 8.50% | 200.2029 | 9.12% | 7.28% |
| 17 | 229.3778 | 10.45% | 258.1879 | 11.77% | 12.56% |
| 18 | 162.6473 | 7.41% | 152.4821 | 6.95% | -6.25% |
| 19 | 96.4034 | 4.39% | 91.3826 | 4.16% | -5.21% |

Таблица 23: Точность, определённая по формуле (32), для предложенных в этой работе и классических мер центральности

| graph \ accuracy | betweenness | closeness | Shapley value | Myerson value |
|---------------------|-------------|-----------|---------------|---------------|
| karate-34 | 100% | 50% | 100% | 100% |
| karate-empty-34 | 100% | 50% | 100% | 100% |
| karate-karate-34 | 100% | 50% | 100% | 100% |
| karate-star-34 | 100% | 100% | 100% | 100% |
| karate-twoStar-34 | 100% | 50% | 100% | 50% |
| karate-cycle-34 | 100% | 100% | 100% | 100% |
| karate-twoClique-34 | 0% | 0% | 100% | 100% |
| karate-complete-34 | 0% | 0% | 100% | 100% |

3.4.2 Эксперименты по изучению корреляции характеристик сети и динамики мнений

Мы провели симуляции динамики мнений в однослойной сети клуба карате Закари с использованием BVM и в двухслойной сети с использованием GCVM, где сеть клуба карате Закари выступает в качестве внешнего слоя, а структуры внутренней сети различаются. Список внутренних слоёв, которые используются в нашем анализе:

1. **karate**: сеть клуба карате Закари;
2. **star**: звездообразная структура с узлом 0 в центре;
3. **two-star**: два центральных узла — узлы 0 и 17, узлы 1–16 соединены с узлом 0, узлы 18–33 соединены с узлом 17. Кроме того, узлы 0 и 17 соединены;
4. **cycle**: узел 0 соединён с узлом 1, узел 1 соединён с узлом 2 и так далее. В итоге узел 33 соединён с узлом 0;
5. **two-clique**: узлы 0–16 относятся к первой клике, узлы 17–33 — ко второй клике, и эти две клики соединены через связь между узлами 0 и 17;
6. **complete**: все узлы связаны друг с другом.

Начинаем с симуляций динамики мнений модели BVM для однослойной сети, а затем моделей CVM и GCVM для двухслойных сетей с различными

внутренними структурами, и исследуем их влияние на время достижения консенсуса и вероятность выигрыша мнения. Характеристики сети (в данной работе — меры центральности) также будут меняться с изменением структуры сети.

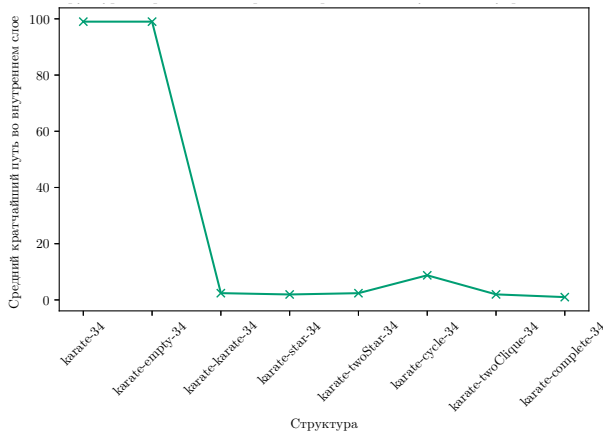
Результаты экспериментов описаны следующим образом:

- На рисунке 26а показано, как изменяется средний внутренний кратчайший путь d_I в зависимости от структуры сети¹².
- На рисунке 26b показано, как изменяется внутренняя плотность в зависимости от структуры сети. Сравнивая рисунки 26а и 26b, можно заметить, что внутренняя плотность D_I и средний внутренний кратчайший путь d_I обратно коррелируют.
- На рисунке 27 показаны различные меры центральности для узлов 0 и 33 для различных структур сети. В левой части рисунка 27 представлены степень посредничества, степень близости, аппроксимированный вектор Шепли и аппроксимированное значение Майерсона для узлов 0 и 33 на упрощённой однослойной взвешенной сети, где веса рассчитаны по формуле (11). В правой части рисунка 27 представлены степень связности группы, степень близости группы, степень посредничества группы и две разные центральности при случайных блужданиях для двухслойной сети с различными внутренними структурами. Сравнивая тренды центральности для узлов 0 и 33 в левой и правой частях рисунка 27, соответственно, можно наблюдать, что оба узла имеют схожие тренды, что подтверждает валидность нашего подхода к упрощению двухслойной сети до однослойной взвешенной сети.
- На рисунке 28 показано, как КРІ изменяются при различных структурах сети. Изучая рисунок 27, можно заметить, что некоторые тренды центральности для узла 33 очень похожи на тренд на рисунке 26а. Рисунок 26b также демонстрирует аналогичный тренд, как и на рисунке 28b.
- На рисунке 28а показано, что структура сети имеет значительное влияние на вероятность выигрыша мнения.

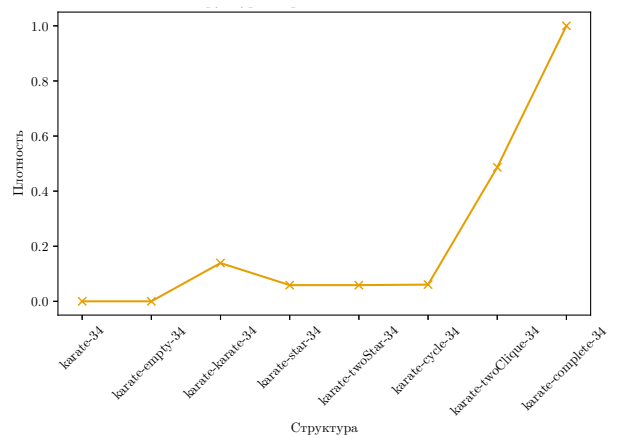
¹²“Karate-34” и “karate-empty-34” относятся к однослойной сети клуба карате Закари и двухслойной сети, где сеть клуба карате Закари — это внешний слой, а внутренний слой отсутствует, соответственно, d_I не существует для этих двух структур, в частности, оно равно бесконечности. Однако на рисунке 26а, вместо бесконечности используется число 99.

- Из рисунков 28b и 26b можно заметить, что существует взаимосвязь между внутренней плотностью D_I и временем достижения консенсуса T_{cons} . Сети с более высокой плотностью, такие как “karate-complete-34”, требуют больше времени для достижения консенсуса, в то время как сети с меньшей плотностью, такие как “karate-empty-34”, достигают консенсуса быстрее. Это можно объяснить тем, что сети с большей плотностью имеют больше связей, что затрудняет быструю доминацию одного мнения.

Основные выводы из приведённых выше результатов: (i) существует обратная корреляция между средним внутренним кратчайшим путём (d_I) и как внутренней плотностью (D_I), так и временем достижения консенсуса (T_{cons}), (ii) подход упрощения двухслойной сети до однослойной взвешенной сети является обоснованным.



(а) Структура в сравнении со средним кратчайшим путем



(б) Структура в сравнении с плотностью

Рис. 26: Средний внутренний кратчайший путь и плотность для различных структур сети

Наш следующий шаг — изучить зависимость структуры сети и ключевых показателей эффективности (KPI) динамики мнений. Для этого были проведены тесты корреляции по результатам наблюдений с использованием SciPy [57, 58]. Мы рассчитали три коэффициента для каждой пары характеристик: коэффициенты корреляции Пирсона [59], Кендалла [91] и Спирмена [92]. Эти коэффициенты представлены в таблице 24, где *, **, *** обозначают уровень значимости 0.05, 0.01 и 0.001 для коэффициента корреляции соответственно¹³.

Из таблицы 24 можно сделать несколько выводов:

¹³Мы выбрали узел 33 в качестве входного для центральности. На самом деле, если выбрать другой узел, выводы остаются актуальными.

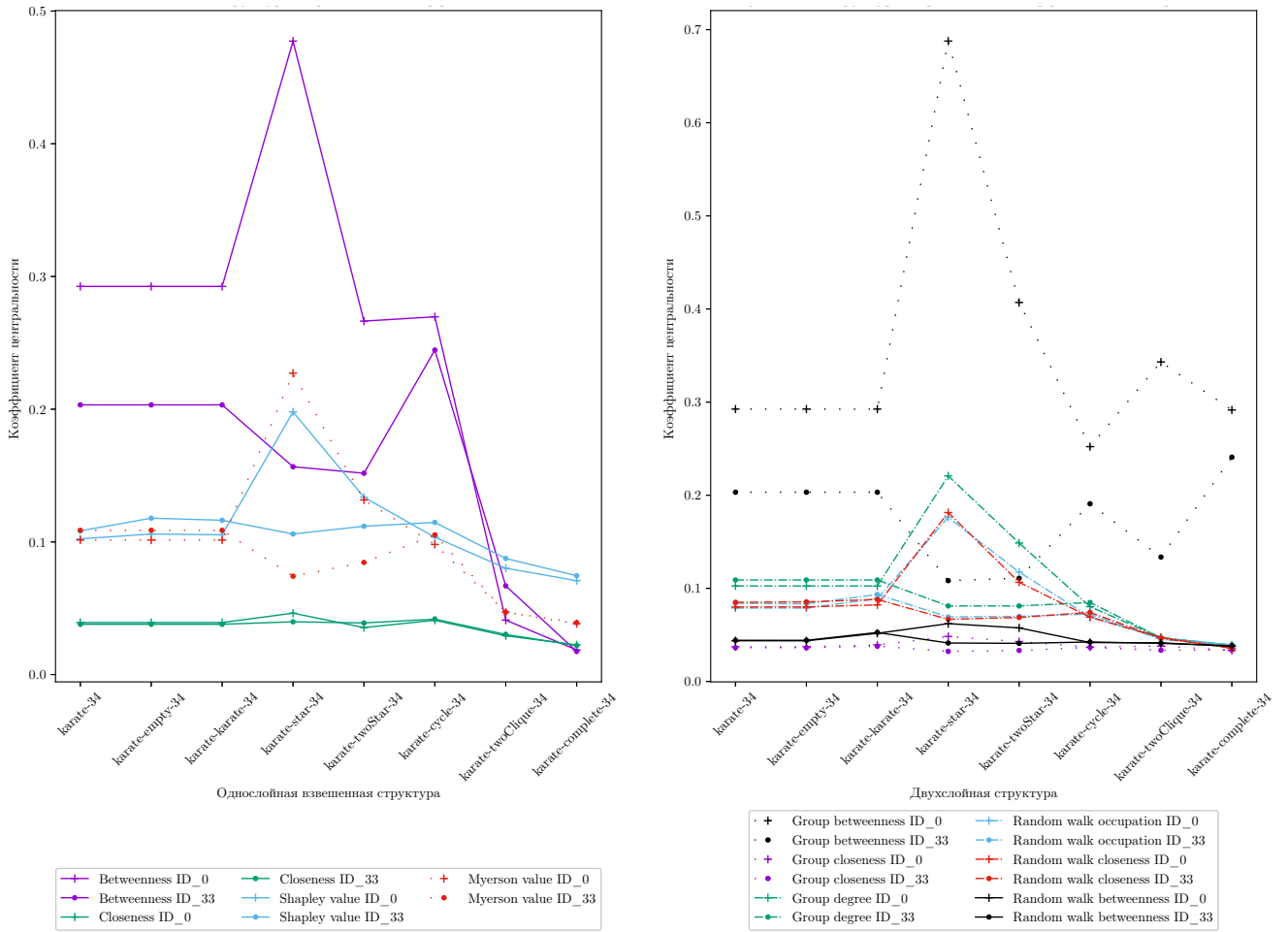
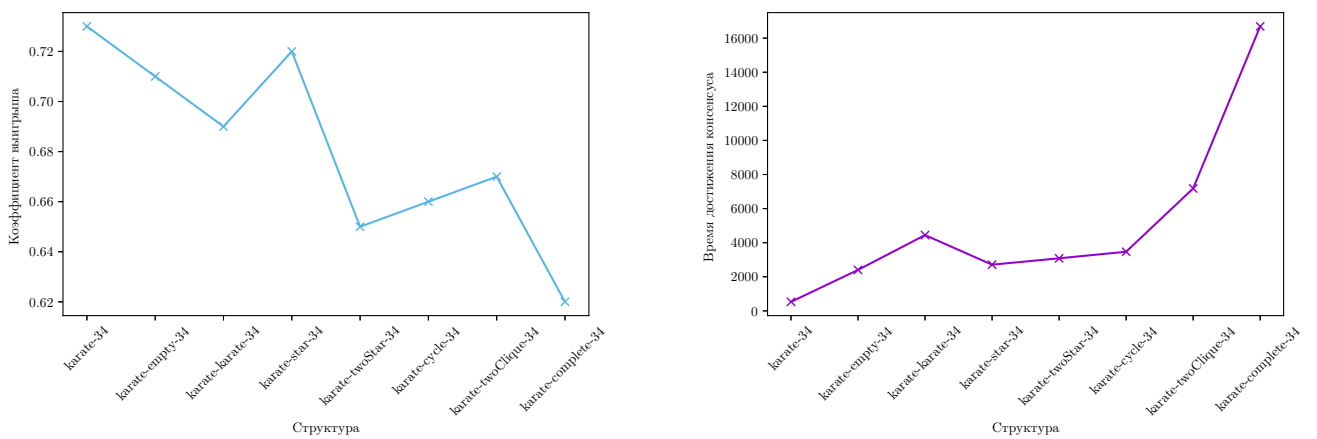


Рис. 27: Меры центральности для различных структур



(а) Структура в сравнении с вероятностью выигрыша мнения

(б) Структура в сравнении с временем достижения консенсуса

Рис. 28: Вероятность выигрыша мнения и время достижения консенсуса для различных структур

1. *Высокая положительная значимая корреляция:* (i) Внутренняя плотность D_I демонстрирует очень сильную и высоко значимую положительную корреляцию с временем достижения консенсуса по всем коэффициентам корреляции (Пирсона, Кендалла и Спирмена), все они больше 0.95. Это указывает на то, что при увеличении D_I время достижения консенсуса T_{cons} значительно увеличивается; (ii) Средний внутренний кратчайший путь d_I имеет значимую положительную корреляцию с большинством мер центральности, что подтверждается коэффициентами корреляции Кендалла и Спирмена, они значимы.
2. *Отрицательная значимая корреляция:* T_{cons} показывает сильную значимую отрицательную корреляцию с мерами центральности сети, такими как центральность посредничества и центральность близости, особенно с центральностью близости (коэффициент корреляции Пирсона равен -0.928). Это позволяет сделать вывод о том, что более высокие значения центральности связаны с меньшим временем достижения консенсуса T_{cons} , что ожидаемо.
3. *Вариативность корреляций:* Различные метрики показывают разные уровни силы корреляции для коэффициентов корреляции Пирсона, Кендалла и Спирмена. Эта вариативность указывает на то, что сила и значимость корреляций могут зависеть от выбранного метода корреляции, вероятно, под влиянием распределений исходных данных.

В заключение, таблица 24 подчёркивает значимые взаимосвязи между определёнными характеристиками сети и показателями динамики мнений, такими как время достижения консенсуса. Центральность «важных» узлов и плотность сети играют ключевую роль для времени достижения консенсуса.

Таблица 24: Коэффициенты корреляции

| | Пирсон | Кендалл | Спирмен |
|---|-----------|---------|----------|
| d_I vs Центральность посредничества | 0.414 | 0.654* | 0.810* |
| d_I vs Центральность близости | 0.231 | 0.192 | 0.270 |
| d_I vs Вектор Шепли | 0.363 | 0.618* | 0.766* |
| d_I vs Значение Майерсона | 0.548 | 0.808** | 0.908** |
| d_I vs Групповая центральность посредничества | 0.366 | 0.185 | 0.157 |
| d_I vs Групповая центральность близости | 0.408 | 0.333 | 0.614 |
| d_I vs Групповая степень | 0.610 | 0.830** | 0.933*** |
| d_I vs Занятость в случайных блужданиях | 0.475 | 0.691* | 0.826* |
| d_I vs Центральность близости в случайных блужданиях | 0.553 | 0.691* | 0.826* |
| d_I vs Центральность посредничества в случайных блужданиях | 0.161 | 0.618* | 0.778* |
| T_{cons} vs Центральность посредничества | -0.841** | -0.491 | -0.537 |
| T_{cons} vs Центральность близости | -0.928*** | -0.340 | -0.464 |
| T_{cons} vs Вектор Шепли | -0.872** | -0.286 | -0.452 |
| T_{cons} vs Значение Майерсона | -0.791* | -0.491 | -0.659 |
| T_{cons} vs Групповая центральность посредничества | 0.407 | 0.255 | 0.299 |
| T_{cons} vs Групповая центральность близости | -0.291 | 0.109 | 0.036 |
| T_{cons} vs Групповая степень | -0.807* | -0.593* | -0.771* |
| T_{cons} vs Занятость в случайных блужданиях | -0.785* | -0.429 | -0.548 |
| T_{cons} vs Центральность близости в случайных блужданиях | -0.834* | -0.357 | -0.524 |
| T_{cons} vs Центральность посредничества в случайных блужданиях | -0.424 | -0.500 | -0.571 |
| D_I vs T_{cons} | 0.983*** | 0.964** | 0.988*** |

3.5 Заключение к главе 3

В этой главе была рассмотрена корреляция нескольких характеристик динамики мнений (включая BVM, SVM и GCVM), реализованной на двухслойных сетях, с характеристиками этих сетей. В качестве сети одного слоя рассматривался клуб каратэ Закари. Было исследовано, как внутренняя структура сети влияет на время достижения консенсуса и вероятность выигрыша мнения, а также, корреляцию этих ключевых показателей эффективности с мерами центральности сети.

В этой главе предложено два быстрых и точных алгоритма для расчёта мер центральности на основе теоретико-игрового подхода. Наши алгоритмы эффективно аппроксимируют теоретические значения этих мер для сетей, для которых точные значения вычислить затруднительно. Оба наших алгоритма

могут выявлять наиболее важные узлы в сети, что было проверено на различных примерах. Идеи нахождения аппроксимированных мер центральности в графах, реализованные в наших алгоритмах, могут быть применены в других области, например, в области искусственного интеллекта. Это будет сделано в следующей главе.

Глава 4. SharpG: новый метод оценки значимости признаков на основе вектора Шепли

В этой главе представлено одно из практических применений разработанных мер центральности, обсужденных в главе 3, в области объяснимого искусственного интеллекта (ХАИ). Мы рассматриваем признаки в моделях машинного обучения как узлы в графе и интерпретируем центральность как значимость признаков. Основываясь на приближённом алгоритме вычисления вектора Шепли, предложенном в главе 3, нами предложен новый метод ХАИ, названный SharpG. Большая часть результатов, представленных в этой главе, опубликована в статье [52].

4.1 Методы объяснимого искусственного интеллекта (ХАИ)

Предполагаем, что имеется выборка наблюдений признаков X_1, \dots, X_M , которые используются для построения модели $f(X_1, \dots, X_M)$ для объяснения, прогнозирования или классификации целевой переменной Y . Из-за непрозрачности «черных ящиков» в анализе данных функциональная форма f может быть неизвестна (в особенности в сложных нейросетевых моделях), и существует актуальная необходимость в оценке значимости признаков в моделях ИИ, применяемых при анализе данных. Существует три наиболее популярных метода ХАИ: Feature Importance (можно перевести как значимость признаков), LIME (локальные интерпретируемые модели, независимые от модели) и SHAP (SHapley Additive exPlanations). Далее будут предоставлены ссылки на соответствующие источники, где эти методы введены или подробно описаны.

Мы используем следующие методы для сравнения с нашим новым методом ХАИ, описанным в разделе 4.2.2:

- **Feature Importance** или **значимость признаков**: Feature Importance — встроенный метод, применяемый к древовидным моделям, таким как деревья решений, случайные леса, градиентный бустинг и т.д. При построении древовидной модели алгоритм автоматически рассчитывает вклад каждого признака в прогноз модели, измеряя влияние каждого признака на предсказанный результат и оценивая его значимость [93].
- **Пермутационная значимость признаков**: Метод оценки значимости признаков, основанный на перестановках, предложен в [94]. Если признак важен, то производительность модели значительно снизится при случайной перестановке, а если признак не важен, его перестановка будет иметь незначительное влияние на производительность модели.
- **LIME**: Это метод объяснения, применяемый к любым моделям, предложенный в [47] и используемый для объяснения значимости переменных в прогнозах для моделей машинного обучения [95]. Метод называется *независимым*, если он не связан с каким-либо конкретным методом машинного обучения (например, используется только для деревьев решений или только к регрессионным моделям, построенным методом наименьших квадратов), но может применяться к большинству моделей машинного обучения.
- **SHAP (SHapley Additive exPlanations)**: Это объясняющий метод, разработанный на основе теории кооперативных игр и предложенный в [48]. SHAP предоставляет механизм объяснения, независимый от модели, который теоретически может быть применен к любой модели машинного обучения.
- **KernelSHAP**: Этот метод основан на LIME и векторе Шепли. Метод использует следующие шаги для упрощения вычислений: (i) генерация случайного числа выборок признаков, (ii) определение данных выборки для каждого подмножества признаков специальным образом, (iii) вычисление весов для каждого подмножества признаков и (iv) решение специально определённой задачи взвешенных наименьших квадратов для нахождения вектора значимости признаков, т.е. приближённого вектора Шепли.
- **Sampling SHAP**: Этот метод вычисляет вектор Шепли при условии независимости признаков и является расширением алгоритма, предложен-

ного в [96]. Вычисления основаны на известной альтернативной формулировке вектора Шепли [97].

4.2 SharG: новый метод ХАІ

4.2.1 Вектор Шепли

Кооперативная игра определяется как (\mathcal{M}, v) , где $\mathcal{M} = \{1, \dots, M\}$ — множество игроков¹⁴, а $v : 2^{\mathcal{M}} \rightarrow \mathbb{R}$ — характеристическая функция, определяющая «силу» любой коалиции игроков, которая является подмножеством множества \mathcal{M} , то есть для любой коалиции или набора игроков $\mathcal{S} \subset \mathcal{M}$, где $S = |\mathcal{S}|$. Значение $v(\mathcal{S})$ представляет выигрыш или силу коалиции \mathcal{S} . Одна из главных задач, которую решает теория кооперативных игр, заключается в нахождении «справедливого» распределения общего выигрыша большой коалиции $v(\mathcal{M})$ между ее членами. Один из возможных дележей предложил Шепли [98] для распределения $v(\mathcal{M})$, и это вектор $\phi = (\phi_1, \dots, \phi_M)$, где ϕ_i — компонента дележа или выигрыш (часть $v(\mathcal{M})$) игрока $i \in \mathcal{M}$, определяемый следующим образом:

$$\phi_i = \sum_{\mathcal{S}: \mathcal{S} \subset \mathcal{M} \setminus \{i\}} \frac{(M - S - 1)! S!}{M!} (v(\mathcal{S} \cup \{i\}) - v(\mathcal{S})), \quad (33)$$

где $(v(\mathcal{S} \cup \{i\}) - v(\mathcal{S}))$ — маргинальный вклад игрока i , если он присоединяется к коалиции \mathcal{S} .

Вектор с компонентами, определёнными формулой (33), называется вектором Шепли, и это единственный вектор, удовлетворяющий четырём аксиомам (эффективности, симметрии, нулевого игрока и аддитивности). Аксиома эффективности означает, что сумма компонент вектора Шепли равна выигрышу большой коалиции \mathcal{M} , то есть $\sum_{i \in \mathcal{M}} \phi_i = v(\mathcal{M})$.

¹⁴В задаче оценки значимости признаков, признак рассматривается как игрок в игре, поэтому мы используем одно и то же обозначение M для множества игроков и признаков.

Приведем вероятностную интерпретацию (см. [99]) вектора Шепли для лучшего понимания того, почему этот вектор может быть применён для измерения значимости признаков в сложных моделях машинного обучения. Рассмотрим i -ю компоненту вектора Шепли, определённую формулой (33). Игрок i получает выигрыш $(v(\mathcal{S} \cup \{i\}) - v(\mathcal{S}))$, когда он присоединяется к случайно сформированной коалиции $\mathcal{S} : \mathcal{S} \subset \mathcal{M} \setminus \{i\}$. Вероятность того, что коалиция \mathcal{S} , содержащая S игроков, будет сформирована, равна $\frac{1}{M} \binom{M-1}{S}$. Предполагается, что все размеры коалиций \mathcal{S} от 0 до $M - 1$ равновероятны, и для данного размера коалиции S подмножества из S игроков также равновероятны. Таким образом, значение ϕ_i , вычисляемое по формуле (33), является ожидаемым выигрышем игрока i в такой вероятностной схеме.

Чтобы связать вектор Шепли с вектором значимости признаков, можно сопоставить множество игроков с множеством признаков, а характеристическую функцию — с некоторой количественной характеристикой предсказания, сделанного моделью машинного обучения с использованием подмножества признаков. Тогда разность $(v(\mathcal{S} \cup \{i\}) - v(\mathcal{S}))$ можно интерпретировать как вклад или выигрыш в качестве предсказания, если мы добавим признак i в подмножество признаков \mathcal{S} . Ожидаемое значение такого выигрыша связано с важностью признака i в тестируемой предсказательной модели машинного обучения.

Основная идея, описанная в этом разделе, заимствована из теории кооперативных игр и реализована в методе SHAP, но из-за сложности вычисления вектора Шепли по формуле (33) в случае большого количества признаков и сложности модели предсказания f , предложены такие алгоритмы, как KernelSHAP и SamplingSHAP, для приближённого вычисления вектора Шепли за счёт сокращения числа вычислений [48, 100].

4.2.2 ShapG (модель на основе вектора Шепли для графов)

В этом разделе представлено описание нового метода ХАИ, называемого *ShapG*, для вычисления значимости признаков в моделях машинного обучения,

основанного на векторе Шепли, определённом на неориентированном взвешенном графе, построенном специальным образом.

Вектор Шепли для неориентированных взвешенных графов

Вычисление вектора Шепли для неориентированного взвешенного графа можно разделить на следующие шаги:

1. Определим неориентированный взвешенный граф $G = (\mathcal{M}, \mathcal{E})$, где \mathcal{M} — множество узлов, ассоциированных с признаками из множества $\mathcal{M} = \{1, \dots, M\}$, и множество рёбер \mathcal{E} без петель. Вес ребра (j, k) , $j \neq k$, равен коэффициенту корреляции Пирсона $W(j, k)$ между признаками j и k , вычисленному по заданной выборке.
2. Для любого подмножества признаков $\mathcal{S} \subset \mathcal{M}$ определяем подграф $G_{\mathcal{S}}$ графа G .
3. Для любого подмножества признаков $\mathcal{S} \subset \mathcal{M}$ определяем значение функции v следующим образом¹⁵:

$$v(\mathcal{S}) = \sum_{\{j,k\} \subseteq G_{\mathcal{S}}} W(j, k). \quad (34)$$

4. Вычисляем вектор Шепли по формуле (33). В результате алгоритм возвращает центральность вектора Шепли для каждого узла (признака).

Поскольку множество признаков \mathcal{M} может быть большим, мы предлагаем подход для приближённого вычисления вектора Шепли с высокой точностью, представленный в [51]. Предложенный метод основан на вычислении весов рёбер и функции (34), где \mathcal{S} является подмножеством признаков из множества \mathcal{M} , а $G_{\mathcal{S}}$ — подграф, порождаемый множеством \mathcal{S} . С минимальной модификацией мы можем применить этот подход для определения нового метода ХАІ, представленного в следующем разделе.

Описание и алгоритм метода SharpG

Метод SharpG можно разделить на следующие шаги:

¹⁵Характеристическая функция (34) была предложена в статье [39], авторы рассматривают кооперативную игру, где характеристическая функция определяется степенью связности группы [77] каждой коалиции [36].

1. Определяем неориентированный взвешенный граф $G = (\mathcal{M}, \mathcal{E})$, где \mathcal{M} — множество узлов, связанных с признаками $\mathcal{M} = \{1, \dots, M\}$ в предсказательной модели, а \mathcal{E} — множество всех возможных рёбер без петель, то есть $\mathcal{E} = \{(i, j) : i \in \mathcal{M}, j \in \mathcal{M}, i \neq j\}$ — полный граф без петель. Вес ребра (j, k) , $j \neq k$, равен коэффициенту корреляции Пирсона $W(j, k)$ между признаками j и k , вычисленному на заданной выборке.
2. Матрица весов $W = \{W(j, k)\}_{(j,k) \in \mathcal{E}}$ обычно является очень плотной, поэтому необходимо уменьшить плотность графа G , чтобы сократить количество дальнейших вычислений. Мы реализуем идею сохранения всех признаков набора данных при минимизации количества рёбер в графе для уменьшения плотности. Соответствующий метод реализован в алгоритме 4. Идея проста: мы строим граф G' , начиная с пустого графа, последовательно выбирая рёбра с наибольшими коэффициентами корреляции Пирсона, данными в матрице W , и добавляя эти рёбра в граф G' , гарантируя, что каждый узел включён в G' хотя бы один раз, а граф G' является связным графом (последнее условие является критерием остановки в алгоритме 4). Результат работы алгоритма 4 — новый граф G' . На следующих шагах мы не используем матрицу весов W ¹⁶.
3. В графе G' определяем подграф G'_S для любого подмножества признаков $\mathcal{S} \subset \mathcal{M}$.
4. Определяем характеристическую функцию $f(\mathcal{S})$, назначая оценку R^2 (для регрессионных моделей) или $F1$ (для моделей классификации) для любого подмножества признаков \mathcal{S} :

$$v(\mathcal{S}) = f(\mathcal{S}). \quad (35)$$

Характеристическая функция, используемая для ХАИ, может быть определена как в (35), где $f(\mathcal{S})$ вычисляется как прогнозное значение, модели, обученной только на признаках из подмножества \mathcal{S} . Для наших целей мы можем использовать R^2 или $F1$ как меру «силы» множества \mathcal{S} .

5. Вычисляем вектор Шепли по формуле (33). Для нахождения вектора Шепли по точной формуле (33) используем алгоритм 5. Если число признаков

¹⁶Мы используем веса только для уменьшения графа G до G' , но не для вычисления вектора Шепли.

велико, используем алгоритм 6 для нахождения приближённого вектора Шепли.

Алгоритм 4 Предварительная обработка данных для метода ShapG

Требования Набор данных с M признаками

Гарантии Матрица смежности $A \in \mathbb{R}^{M \times M}$.

- 1: Вычислить матрицу корреляций Пирсона W для всех признаков. ▷ Это исходная матрица весов для графа G
 - 2: Инициализировать матрицу смежности A нулевой матрицей.
 - 3: $\mathcal{E} \leftarrow$ Список кортежей $(i, j, W(i, j))$, где $i < j$.
 - 4: $\mathcal{E} \leftarrow$ Отсортировать \mathcal{E} по убыванию $|W(i, j)|$.
 - 5: Инициализировать индекс $k = 0$
 - 6: **while** G' не является связным **do** ▷ G' представлен матрицей A
 - 7: $i, j, weight = \mathcal{E}[k]$
 - 8: $A(i, j) \leftarrow 1$
 - 9: $A(j, i) \leftarrow 1$
 - 10: $k \leftarrow k + 1$
 - 11: **end while**
- return** Матрица смежности A для представления графа признаков.
-

Алгоритм 5 описывает вычисление вектора Шепли для признаков $1, \dots, M$ на основе характеристической функции, определённой в (35), следуя шагам, описанным выше¹⁷

Алгоритм 5 Вычисление вектора Шепли на основе графа G'

Требования Граф $G'(\mathcal{M}, \mathcal{E})$ с $M = |\mathcal{M}|$ узлами

Гарантии Компоненты вектора Шепли $\phi(i)$ для каждого узла $i \in \mathcal{M}$

- 1: **for all** узлы $i \in \mathcal{M}$ **do**
 - 2: Инициализировать $\phi(i) \leftarrow 0$
 - 3: **end for**
 - 4: **for all** узлы $i \in \mathcal{M}$ **do**
 - 5: **for all** подмножества $\mathcal{S} \subseteq \mathcal{M} \setminus \{i\}$ **do**
 - 6: Вычислить $v(\mathcal{S}) \leftarrow f(\mathcal{S})$
 - 7: Вычислить $v(\mathcal{S} \cup \{i\}) \leftarrow f(\mathcal{S} \cup \{i\})$
 - 8: $\Delta v(\mathcal{S}, i) \leftarrow v(\mathcal{S} \cup \{i\}) - v(\mathcal{S})$
 - 9: коэффициент $\leftarrow \frac{S!(M-S-1)!}{M!}$
 - 10: $\phi(i) \leftarrow \phi(i) + \text{коэффициент} \cdot \Delta v(\mathcal{S}, i)$
 - 11: **end for**
 - 12: **end for**
- return** $\phi(i)$ для всех $i \in \mathcal{M}$
-

¹⁷Для этого алгоритма не нужно уменьшать плотность графа, так как независимо от структуры графа, алгоритм 5 всегда должен перебирать все возможные коалиции. Алгоритм 5 выполняет итерации для всех возможных коалиций, поэтому для графов G и G' , количество итераций одинаково. Для алгоритма 6 не нужно учитывать все коалиции, а только каждый узел и его достижимые узлы. Таким образом, алгоритм 5 не зависит от графа, но алгоритм 6 зависит от него.

Тем не менее, вычисление вектора Шепли довольно ресурсоёмко, особенно для большого числа признаков и, следовательно, большого числа подмножеств признаков. Поэтому мы предоставляем модифицированный алгоритм, основанный на том факте, что влияние на конкретный узел от других узлов уменьшается с увеличением длины пути, соединяющего их. Алгоритм 6 может использоваться вместо алгоритма 5 с графом G' в качестве входных данных. Алгоритм 6 реализует следующие идеи для ускорения вычисления вектора Шепли:

1. *Ограничение глубины:* Ограничивая глубину достижимых узлов, которые учитываются, уменьшается количество подмножеств, которые необходимо рассматривать.

Задаём параметр d_{max} , который является глубиной графа для любого узла $i \in \mathcal{M}$ для формирования множества достижимых узлов. Определим $\psi(i, d_{max})$ как множество достижимых узлов для узла i на глубине до d_{max} , исключая узел i . Затем вычисляем вектор Шепли для узла/признака i на основе $\psi(i, d_{max})$ по уравнению (25), где $\beta = \frac{|\psi(i, d_{max})|+1}{m+1}$ является коэффициентом масштабирования, а m — максимальное количество достижимых узлов.

2. *Итерация по локальным подмножествам:* Итерация только по подмножествам внутри достижимых узлов, а не по всему графу, уменьшает количество итераций. Мы вычисляем только маргинальный вклад текущего узла в его достижимые узлы из множества $\psi(i, d_{max})$ при заданном d_{max} (см. пункт 1).

В алгоритме 6 используем формулу (35) для определения значений характеристической функции, но ограничиваем глубину достижимых узлов (пункт 2), чтобы уменьшить количество вычислений.

3. *Сэмплирование по количеству достижимых узлов:* При большом количестве достижимых узлов сложность вычислений может быть снижена за счёт случайного сэмплирования, что уменьшает количество итераций по подмножествам.

Когда $|\psi(i, d_{max})| \geq m$, мы выбираем случайную выборку из m узлов из множества $\psi(i, d_{max})$ несколько раз и вычисляем вектор Шепли на основе этих выборок. Размер сэмплирования $H_{|\psi(i, d_{max})|, m}$ задаётся по формуле (26). Значение $H_{|\psi(i, d_{max})|, m}$ — математическое ожидание количества

выборки при каждом сэмплеировании m узлов из множества $\psi(i, d_{max})$ до тех пор, пока не будут собраны все достижимые узлы¹⁸. Как показано в алгоритме 6, процесс сэмплеирования повторяется $H_{|\psi(i, d_{max})|, m}$ раз, после чего берётся среднее значение вычисленных векторов Шепли.

Пункты 1–3 из этого списка уменьшают количество вычислений и вычислительную сложность алгоритма. Основываясь на этих трёх пунктах, строится модифицированный алгоритм 6 метода SharpG¹⁹.

Алгоритм 6 Приближённое значение вектора Шепли на основе графа признаков

Требования Граф $G' = (\mathcal{M}, \mathcal{E})$, ограничение глубины d_{max} , максимальный размер m множества достижимых узлов

Гарантии Вектор Шепли $\phi_a(i)$ для каждого узла $i \in \mathcal{M}$

- 1: Инициализировать $\phi(i)_a \leftarrow 0$ для каждого $i \in \mathcal{M}$
 - 2: **for** $i \in \mathcal{M}$ **do**
 - 3: $\psi(i, d_{max}) \leftarrow$ Вычислить или получить все *reachable nodes* для i до глубины d_{max}
 - 4: **if** $|\psi(i, d_{max})| < m$ **then**
 - 5: **for** каждое подмножество $\mathcal{S} \subseteq \psi(i, d_{max}) \setminus \{i\}$ **do**
 - 6: Вычислить $v(\mathcal{S}) \leftarrow f(\mathcal{S})$
 - 7: Вычислить $v(\mathcal{S} \cup \{i\}) \leftarrow f(\mathcal{S} \cup \{i\})$
 - 8: $\Delta v(\mathcal{S}, i) \leftarrow v(\mathcal{S} \cup \{i\}) - v(\mathcal{S})$
 - 9: $\phi_a(i) \leftarrow \phi_a(i) + \Delta v(\mathcal{S}, i)$
 - 10: **end for**
 - 11: $\text{coeff} \leftarrow \frac{1}{2^{|\psi(i, d_{max})|}}$
 - 12: $\phi_a(i) \leftarrow \phi_a(i) \cdot \text{coeff}$, нормализовать $\phi_a(i)$ на основе числа подмножеств
 - 13: **else**
 - 14: Выбрать m узлов случайным образом из $\psi(i, d_{max})$ и повторить $H_{|\psi(i, d_{max})|, m}$ раз
 - 15: **for** $i = 1$ до $H_{|\psi(i, d_{max})|, m}$ **do**
 - 16: $s_{reachable} \leftarrow$ Случайно выбрать выборку из m узлов из $\psi(i, d_{max})$
 - 17: **for** каждое подмножество $\mathcal{S} \subseteq s_{reachable} \setminus \{i\}$ **do**
 - 18: Вычислить $v(\mathcal{S})$ и $v(\mathcal{S} \cup \{i\})$ как прежде
 - 19: $\Delta v(\mathcal{S}, i) \leftarrow v(\mathcal{S} \cup \{i\}) - v(\mathcal{S})$
 - 20: $\phi_a(i) \leftarrow \phi_a(i) + \Delta v(\mathcal{S}, i)$
 - 21: **end for**
 - 22: **end for**
 - 23: $\text{coeff} \leftarrow 1/2^{|\psi(i, d_{max})|} / H_{|\psi(i, d_{max})|, m} \cdot \frac{|\psi(i, d_{max})|+1}{m+1}$
 - 24: $\phi_a(i) \leftarrow \phi_a(i) \cdot \text{coeff}$
 - 25: **end if**
 - 26: **end for**
 - 27: **return** $\phi_a(i)$ для всех $i \in \mathcal{M}$
-

¹⁸Этот вероятностный сценарий известен как обобщённая задача о сборщике купонов, которая была рассмотрена и изучена в [87].

¹⁹Соответствующий код для метода SharpG можно найти в репозитории GitHub по адресу <https://github.com/vectorsss/sharpG>.

4.3 Эксперименты

4.3.1 Описание наборов данных

Для демонстрации работы нашего метода XAI SharpG мы рассмотрим два набора данных для построения предсказательных моделей: (i) набор данных “housing price” для задачи построения регрессии и (ii) набор данных “H1N1 flu vaccine” для задачи классификации.

Кратко опишем наборы данных:

1. Набор данных “housing price” был собран Бюро переписи населения США для сбора информации о жилье в районе Бостона, штат Массачусетс. Набор данных содержит 13 признаков, включая “уровень преступности на душу населения по городам”, “среднее количество комнат на одно жилище” и “доля населения с низким социальным статусом”, а также целевую переменную — “медианную стоимость домов, занятых владельцами”.
2. Набор данных “H1N1 flu vaccine” предоставлен Национальным центром статистики здравоохранения и взят с сайта DrivenData. Набор данных содержит 35 признаков, включая “возраст человека”, “пол”, “уровень образования” и “знание о вакцине от гриппа H1N1”, а также целевую переменную — бинарное значение того, получил ли человек вакцину от H1N1 или нет.

Эти два набора данных используются для разных задач прогнозирования, но в обоих случаях используются алгоритмы машинного обучения для моделирования и прогнозирования. Для задачи регрессии мы используем R^2 для определения характеристической функции (35) в нашем методе XAI SharpG, а для задачи классификации — $F1$ для определения значений характеристической функции (35).

Следует отметить, что набор данных “housing price” содержит значительно меньше признаков, чем набор данных “H1N1 flu vaccine”. Мы будем сравнивать наш метод SharpG с другими методами XAI не только по количественному значению, но и по времени работы алгоритмов. Такое сравнение помогает оценить эффективность различных методов XAI при обработке данных большого объёма.

4.3.2 Предварительная обработка данных для SharG

Следуем алгоритму 4 для создания графа G' на основе исходного полного графа, соединяющего узлы, представляющие признаки. Алгоритм 4 начинается с пустого графа и последовательно добавляет рёбра, представляющие пары признаков с наибольшими коэффициентами корреляции. Он останавливается, когда все признаки соединены (граф G' должен быть связным), обеспечивая наличие важной структурной информации в графе признаков. На рисунках 29 и 30 представлены тепловые карты коэффициентов корреляции признаков для наборов данных “housing price” и “H1N1”, соответственно. На обоих рисунках, фигура (a) показывает исходную тепловую карту коэффициентов корреляции, а фигура (b) — тепловую карту коэффициентов корреляции между признаками в редуцированном графе G' , вычисленную по алгоритму 4.

Подчеркнем, что используем алгоритм 4 для уменьшения количества рёбер, сохраняя при этом все признаки, тем самым игнорируя “неважные” взаимосвязи между признаками при построении графа G' . Это уменьшит количество итераций для вычисления компонентов вектора Шепли, измеряющих значимость признаков.

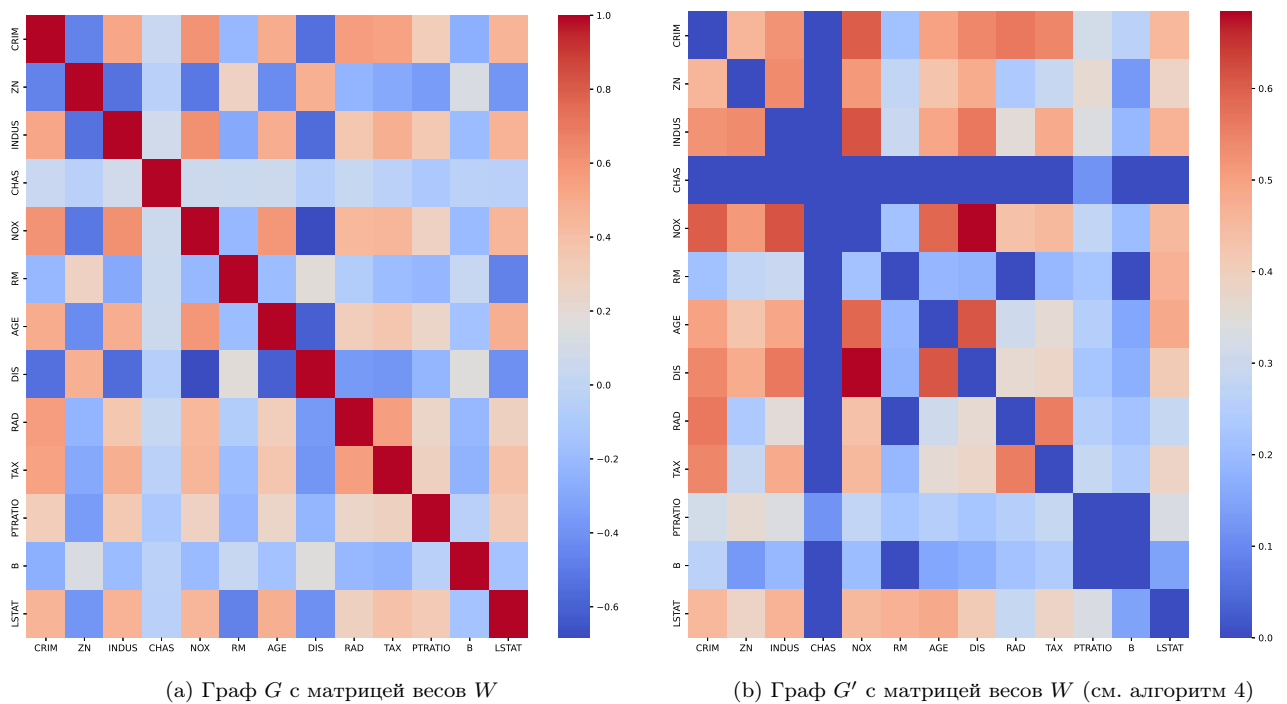


Рис. 29: Тепловая карта коэффициентов корреляции Пирсона для набора данных “housing price”

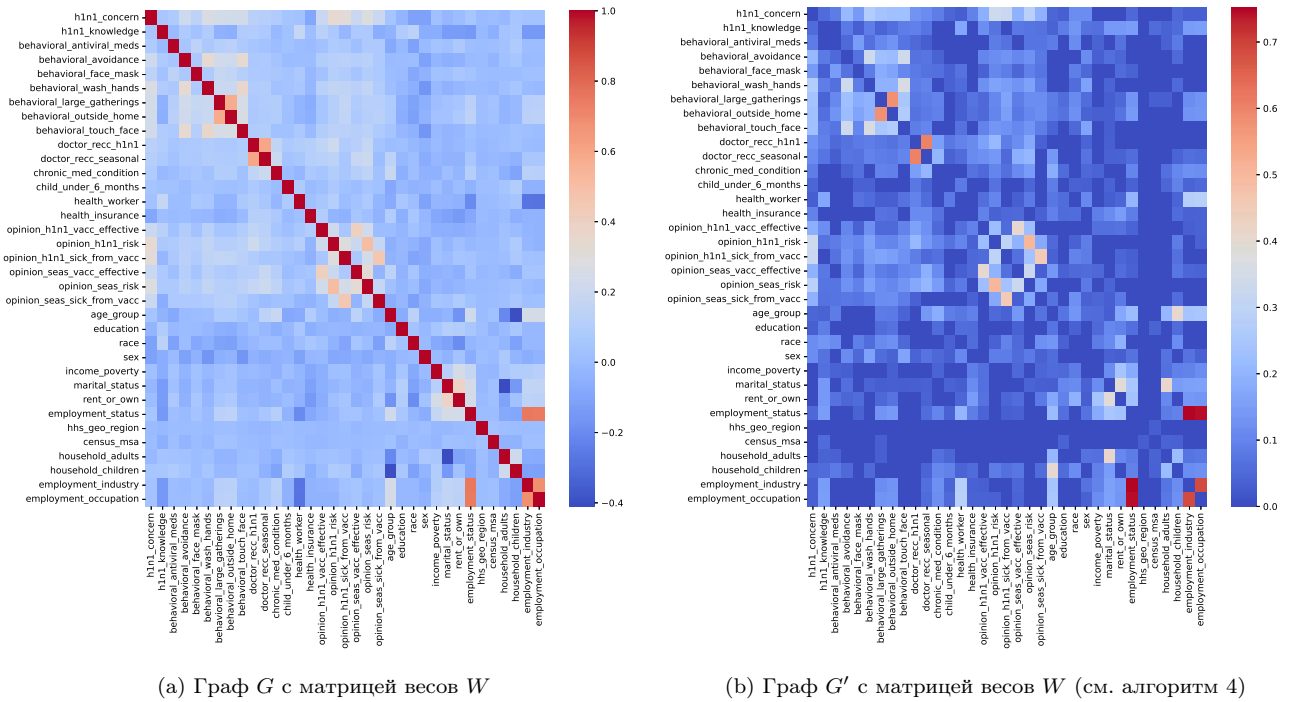


Рис. 30: Тепловая карта коэффициентов корреляции Пирсона для набора данных “H1N1”

Исходный неориентированный граф G с узлами, представляющими признаки набора данных “housing price” (“H1N1”), показан на рисунке 31a (рисунок 32a), в то время как на рисунке 31b (рисунок 32b) изображён неориентированный граф G' , вычисленный по алгоритму 4. Таким образом, мы упрощаем структуры графов и сохраняем важные пары признаков, улучшая эффективность и объяснимость метода SharpG.

4.3.3 Модели предсказания ИИ

Для оценки эффективности нашего метода SharpG применяются несколько методов объяснимого искусственного интеллекта (ХАИ) для объяснения значимости признаков в моделях LightGBM и MLP (многослойный перцептрон). Более того, наш алгоритм SharpG может предоставлять объяснения для сложных моделей ИИ, которые существующие методы ХАИ не могут объяснить за разумное время выполнения. Мы используем ансамблевое обучение и двоичное комбинирование древовидных моделей, моделей нейронных сетей, линейных и

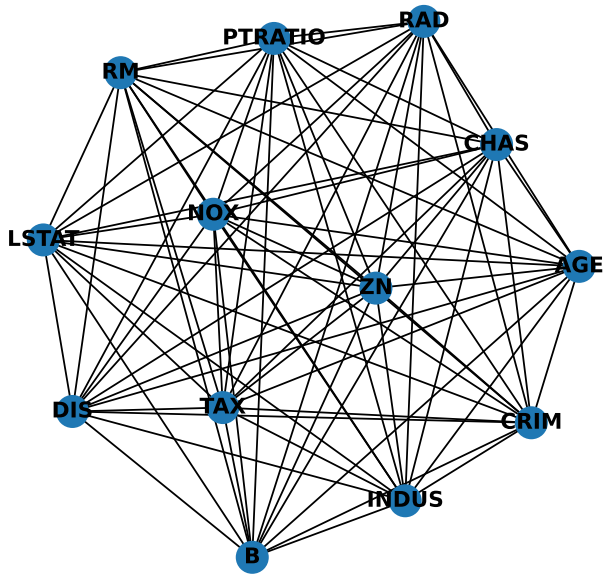
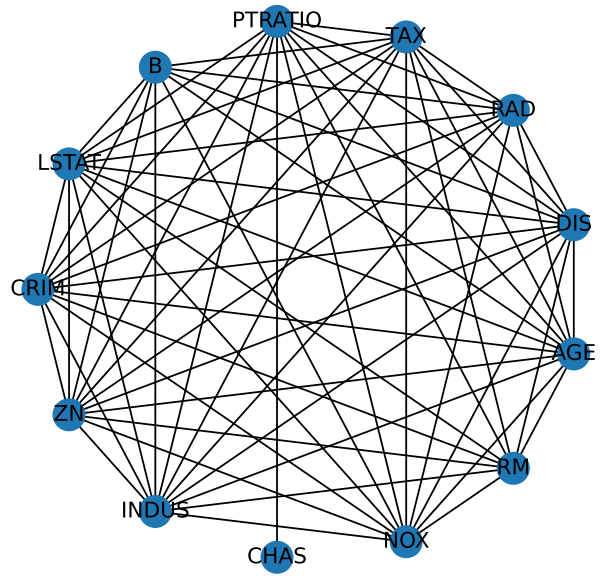
(a) Исходный граф G (b) Редуцированный граф G' , построенный по алгоритму 4

Рис. 31: Граф, соединяющий признаки в наборе данных “housing price”

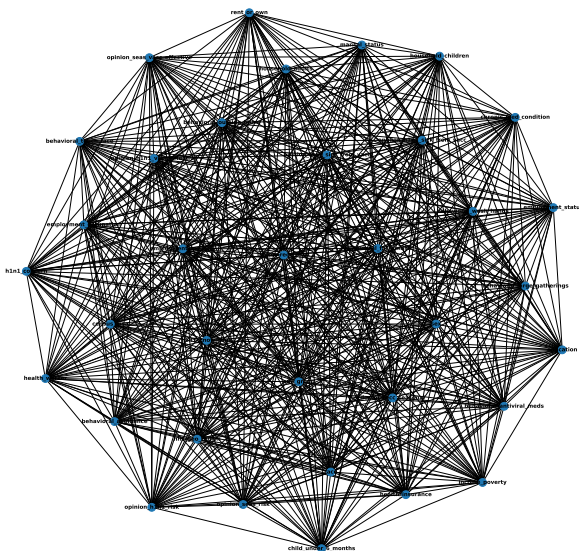
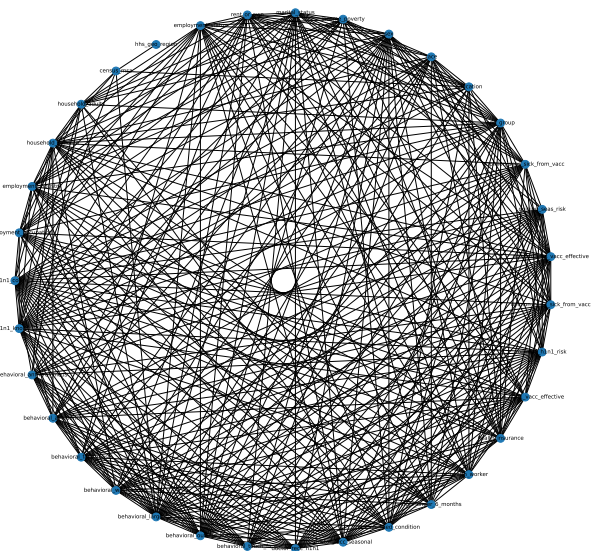
(a) Исходный граф G (b) Редуцированный граф G' , построенный по алгоритму 4

Рис. 32: Граф, соединяющий признаки в наборе данных “H1N1”

машинных моделей для построения гибридных предсказательных моделей. Эти гибридные модели создаются как комбинации различных типов более простых моделей для достижения лучшей производительности и объяснимости. Список построенных моделей (включая их комбинации) и их описания:

- **LightGBM (LGB):** Это эффективный фреймворк Gradient Boosting Decision Tree (GBDT), обладающий высокой скоростью обучения и производительностью. Он широко используется на практике [101].
- **Многослойный перцептрон (MLP):** Это нейронная сеть с прямым распространением, состоящая из полностью связанных нейронов с нелинейными функциями активации [102]. После нескольких тестов мы настроили параметры для получения наилучшей модели прогнозирования. В регрессионной модели (для набора данных “housing price”) используется скрытый слой, содержащий 300 нейронов, скорость обучения установлена на уровне 0.007, функция активации выбрана как ReLU (Rectified Linear Unit), а алгоритм оптимизации — Adam (Adaptive Moment Estimation). Эти параметры обеспечивают наибольшую точность предсказания для набора данных “housing price”. В классификационной модели (для набора данных “H1N1”), для достижения высокой производительности, модель имеет два скрытых слоя: первый с 100 нейронами, а второй — с 50 нейронами. Модель выполняет до 3 итераций для завершения обучения на наборе данных, а функция активации также выбрана как ReLU.
- **Ансамблевое обучение (Stacking):** Это подход для повышения точности прогнозирования за счёт объединения нескольких отдельных моделей [103]. Stacking — это один из методов ансамблевого обучения, где сначала обучаются несколько различных типов базовых моделей с использованием обучающей выборки, а затем метамодель обучается с использованием прогнозов базовых моделей в качестве входных признаков в комбинации с реальными метками [104]. В этой работе мы выбираем Random Forest и XGBoost как базовые модели, а LightGBM — как метамодель.
- **Линейная регрессия — LightGBM (Linear - LGB):** Это классическая модель линейной регрессии и наиболее представительская модель LightGBM для регрессионного прогнозирования на наборе данных “housing price”. Линейные модели эффективно «учитывают» линейные

взаимосвязи между признаками [105], в то время как древовидные модели способны обрабатывать нелинейные взаимосвязи.

- **Логистическая регрессия — LightGBM (Logistic - LGB):** Логистическая регрессия — распространённая линейная модель, используемая для решения задач классификации [106]. Для набора данных “H1N1” мы комбинируем логистическую регрессию и LightGBM для классификации.
- **Линейная регрессия — Многослойный перцептрон (Linear - MLP):** Многослойный перцептрон (MLP) — искусственная нейронная сеть (ANN), состоящая из нескольких слоев связанных нейронов, способная обрабатывать различные типы данных [107]. Мы объединили модель линейной регрессии и модель нейронной сети — MLP для создания регрессионной модели прогнозирования.
- **Логистическая регрессия — Многослойный перцептрон (Logistic - MLP):** Мы комбинируем логистическую регрессию и MLP для создания классификационной модели. Используя логистическую регрессию для первоначальной категоризации входных данных, мы затем применяем MLP для учета более сложных и нелинейных взаимосвязей. Это подход хорош для решения различных сложных задач классификации.
- **Линейная регрессия — К-ближайших соседей (Linear - KNN):** KNN — распространённый алгоритм машинного обучения, широко применяемый для решения задач классификации и регрессии [108]. Модель Linear-KNN может быть использована для прогнозирования как линейных, так и нелинейных взаимосвязей при решении задач регрессии, улучшая гибкость и точность модели.
- **Логистическая регрессия — К-ближайших соседей (Logistic - KNN):** Мы комбинируем логистическую регрессию с алгоритмом KNN, и эта гибридная модель может использовать как линейные взаимосвязи логистической регрессии, так и непараметрические свойства KNN при решении задач классификации.
- **Многослойный перцептрон — LightGBM (MLP - LGB):** Гибридная модель, состоящая из нейронных сетей многослойного перцептрона (MLP) и LightGBM, одновременно использует как способность нейронных сетей

к нелинейному приближению, так и эффективную производительность градиентных бустинговых деревьев решений.

- **К-ближайших соседей — LightGBM (KNN - LGB):** Гибридная модель KNN и LightGBM даёт преимущество высокой производительности и эффективности LightGBM, используя при этом непараметрические свойства KNN.
- **Многослойный перцептрон — К-ближайших соседей (MLP - KNN):** Мы комбинируем нейронную сеть многослойного перцептрона (MLP) с алгоритмом KNN, что позволяет использовать как нелинейные возможности нейронной сети, так и непараметрические свойства KNN для повышения производительности классификации или регрессии.

Цель гибридных моделей заключается в объединении различных типов моделей для повышения производительности и объяснимости. Наш предложенный алгоритм ХАІ объясняет эти сложные модели ИИ и демонстрирует свою широкую применимость. Это означает, что его можно применять к широкому спектру сложных гибридных моделей, и он не ограничен определёнными типами моделей. Следовательно, он независим от модели.

Сначала мы вычисляем значимость признаков для моделей LightGBM и MLP для двух наборов данных в разделе 4.4.1, а затем для более сложных моделей, представленных в списке выше, в разделе 4.4.3.

4.3.4 Оценка методов ХАІ

Один из распространённых способов оценки методов ХАІ — проведение анализа возмущений признаков, удаление признаков в порядке их важности от самого важного к наименее важному и наблюдение за снижением точности или R^2 модели [109].

Когда мы последовательно удаляем признаки на основе их важности, если производительность модели значительно ухудшается, это указывает на то, что признак очень важен. Оценка метода ХАІ таким образом позволяет понять вклад каждого признака в предсказательную производительность модели.

Чтобы доказать эффективность нашего метода ShapG, мы можем использовать этот метод оценки для сравнения ShapG с существующими популярными методами ХАИ. Сначала мы применим предложенный нами метод ХАИ и другие популярные методы ХАИ для генерации результатов объяснения. Затем будем постепенно удалять признаки в соответствии с их важностью и наблюдать изменения в производительности модели. Мы сравниваем этот процесс, используя следующие методы ХАИ: Feature Importance, Permutation Feature Importance, LIME, SHAP (KernelSHAP и SamplingSHAP) и ShapG.

4.4 Результаты и их анализ

4.4.1 Значимость признаков, вычисленная ShapG

Набор данных “housing price”

На рисунке 33 показана значимость признаков, вычисленная алгоритмом ShapG для набора данных “housing price” на основе моделей LightGBM (рисунок 33a) и MLP (рисунок 33b). Для набора данных “housing price” с регрессионным прогнозом на основе LightGBM четыре самых важных признака, полученные ShapG, это “LSTAT (низкий статус населения)”, “RM (среднее количество комнат на одно жилище)”, “NOX (концентрация оксидов азота)”, “PTRATIO (соотношение числа учеников к учителям по городам)”, тогда как с регрессионным прогнозом на основе модели MLP, эти признаки: “LSTAT (низкий статус населения)”, “B (доля чернокожих по городам)”, “RM (среднее количество комнат на одно жилище)”, “PTRATIO (соотношение числа учеников к учителям по городам)”. Три признака совпадают для обеих моделей, а самый важный признак уникален для каждой модели.

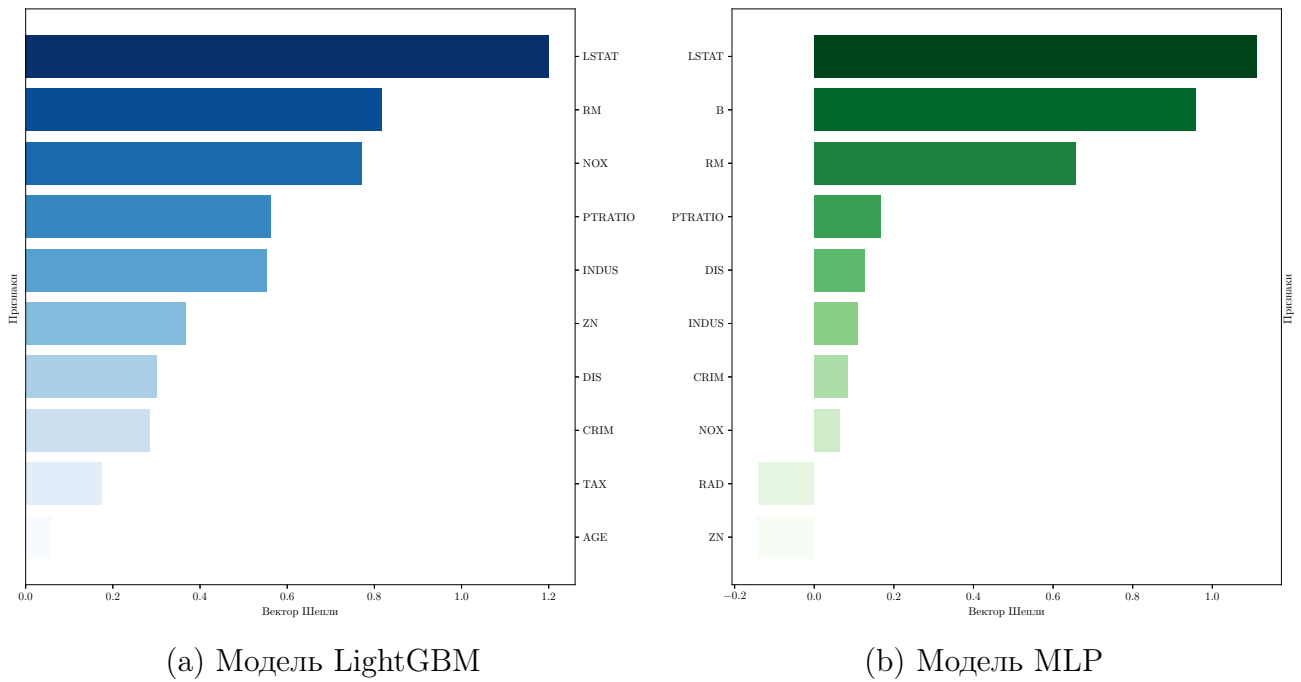


Рис. 33: Значимость признаков в наборе данных “housing price”, вычисленная с помощью SharG

Набор данных “H1N1”

На рисунке 34 показана значимость признаков, вычисленная алгоритмом SharG для набора данных “H1N1” на основе моделей LightGBM (рисунок 34a) и MLP (рисунок 34b). Для набора данных “H1N1” с классификационным прогнозом приближённые значения векторов Шепли для признаков, полученные алгоритмом SharG, отражают важность каждого признака для готовности людей вакцинироваться против H1N1. Для классификации на основе LightGBM пять самых важных признаков — это “doctor recs h1n1 (рекомендация врача вакцинироваться от H1N1)”, “opinion h1n1 risk (мнение респондента о риске заболеть H1N1 без вакцинации)”, “health insurance (наличие медицинской страховки)”, “opinion h1n1 vacc effective (мнение респондента об эффективности вакцины от H1N1)”, “employment occupation (профессия респондента)”.

Для классификации на основе модели MLP пять самых важных признаков — это “doctor recs h1n1 (рекомендация врача вакцинироваться от H1N1)”, “opinion h1n1 risk (мнение респондента о риске заболеть H1N1 без вакцинации)”, “doctor recs seasonal (рекомендация врача вакцинироваться от сезонного гриппа)”, “health insurance (наличие медицинской страховки)”, “health worker (работник здравоохранения)”. Три самых важных признака для обеих моделей LightGBM и MLP совпадают.

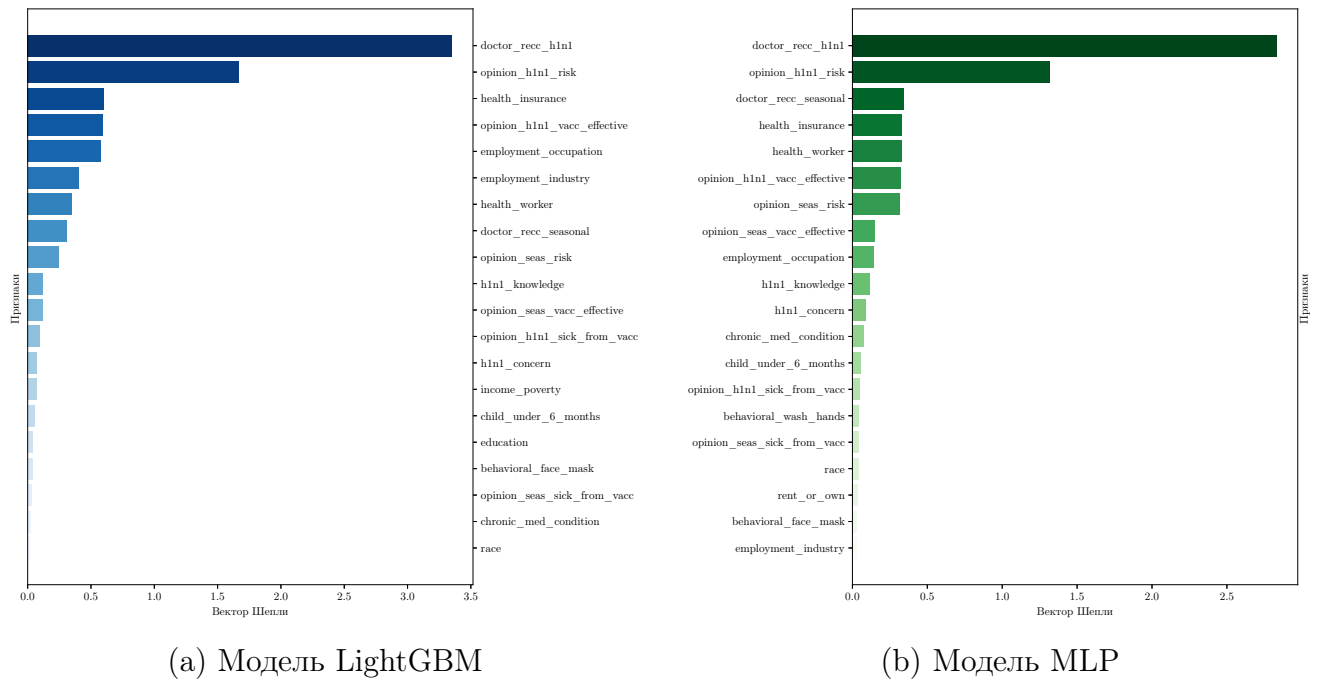


Рис. 34: Значимость признаков в наборе данных “H1N1”, вычисленная с помощью SharpG

Основываясь на подходе ХАИ для объяснения различных моделей ИИ, получены разные результаты значимости объясняющих переменных, так как разные архитектуры моделей могут по-разному обрабатывать данные, извлекать признаки и делать предсказания, что приводит к различиям в результатах объяснения. Когда обе модели ИИ (LightGBM и MLP) выделяют одни и те же самые важные признаки, это позволяет принимать более обоснованные решения и лучше понимать поведение моделей “черного ящика”. Как упоминалось выше, для набора данных “housing price” можно наблюдать, что признаки “LSTAT”, “RM” и “PTRATIO” считаются самыми важными для предсказания стоимости домов как для модели LightGBM, так и для MLP. Для набора данных “H1N1” такими признаками являются “doctor_recc_h1n1”, “opinion_h1n1_risk (мнение респондента о риске заболеть H1N1 без вакцины)” и “health_insurance (наличие медицинской страховки)”, что позволяет разрабатывать соответствующие стратегии для повышения уровня вакцинации.

4.4.2 Оценка методов ХАІ

Чтобы доказать эффективность предложенного нами метода ХАІ SharpG, мы сравниваем результаты его работы с другими существующими методами ХАІ, (i) оценивая все методы на основе анализа возмущений признаков и (ii) измеряя время выполнения для получения результатов.

Мы вводим нумерацию признаков для наборов данных “housing price” и “H1N1” в таблицах 25 и 26 соответственно.

Таблица 25: Нумерация признаков в наборе данных “housing price”

| № | Признак | № | Признак | № | Признак |
|---|---------|----|---------|----|---------|
| 1 | CRIM | 6 | RM | 11 | PTRATIO |
| 2 | ZN | 7 | AGE | 12 | B |
| 3 | INDUS | 8 | DIS | 13 | LSTAT |
| 4 | CHAS | 9 | RAD | | |
| 5 | NOX | 10 | TAX | | |

Таблица 26: Нумерация признаков в наборе данных “H1N1”

| № | Признак | № | Признак |
|----|-----------------------------|----|-----------------------------|
| 1 | h1n1 concern | 19 | opinion seas vacc effective |
| 2 | h1n1 knowledge | 20 | opinion seas risk |
| 3 | behavioral antiviral meds | 21 | opinion seas sick from vacc |
| 4 | behavioral avoidance | 22 | age group |
| 5 | behavioral face mask | 23 | education |
| 6 | behavioral wash hands | 24 | race |
| 7 | behavioral large gatherings | 25 | sex |
| 8 | behavioral outside home | 26 | income poverty |
| 9 | behavioral touch face | 27 | marital status |
| 10 | doctor recc h1n1 | 28 | rent or own |
| 11 | doctor recc seasonal | 29 | employment status |
| 12 | chronic med condition | 30 | hhs geo region |
| 13 | child under 6 months | 31 | census msa |
| 14 | health worker | 32 | household adults |
| 15 | health insurance | 33 | household children |
| 16 | opinion h1n1 vacc effective | 34 | employment industry |
| 17 | opinion h1n1 risk | 35 | employment occupation |
| 18 | opinion h1n1 sick from vacc | | |

В таблице 27 представлены результаты (ранжирование важности признаков), полученные различными методами XAI для набора данных “housing price” на основе моделей LightGBM и MLP. Мы наблюдаем, что SamplingSHAP и KernelSHAP дают одинаковые результаты для обеих моделей, поэтому они объединены в один столбец. Как видно из таблицы 27, для модели LightGBM на наборе данных “housing price” все методы XAI имеют одинаковое ранжирование для первых двух самых важных признаков: “LSTAT (низкий статус населения)” и “RM (среднее количество комнат на одно жилище)”. Методы XAI, применённые к модели MLP, включая SHAP, SamplingSHAP, KernelSHAP и ShapG, дают одинаковое ранжирование для первых двух самых важных признаков: “LSTAT” и “B (доля чернокожих по городам)”, тогда как ранжирование, предоставленное LIME, отличное. Кроме того, метод PFI тоже дает результаты, отличные от результатов других методов XAI в ранжировании второго признака. Мы используем метод оценки XAI для сравнения производительности этих методов XAI, и результаты сравнения показаны на рисунках 35 и 36.

Таблица 27: Ранжирование важности признаков для набора данных “housing price”

Модель LightGBM

| Ранг \ XAI | SHAP | SamplingSHAP (KernelSHAP) | LIME | Feature Importance | ShapG |
|------------|------|---------------------------|------|--------------------|-------|
| Топ 1 | 13 | 13 | 13 | 13 | 13 |
| Топ 2 | 6 | 6 | 6 | 6 | 6 |
| Топ 3 | 8 | 7 | 7 | 1 | 5 |
| Топ 4 | 7 | 8 | 11 | 8 | 11 |
| Топ 5 | 5 | 5 | 10 | 7 | 3 |
| Топ 6 | 1 | 1 | 12 | 11 | 2 |
| Топ 7 | 11 | 11 | 4 | 12 | 8 |

Модель MLP

| Ранг \ XAI | SHAP | SamplingSHAP (KernelSHAP) | LIME | Permutation Feature Importance | ShapG |
|------------|------|---------------------------|------|--------------------------------|-------|
| Топ 1 | 13 | 13 | 12 | 13 | 13 |
| Топ 2 | 12 | 12 | 13 | 7 | 12 |
| Топ 3 | 7 | 7 | 11 | 12 | 6 |
| Топ 4 | 2 | 9 | 3 | 2 | 11 |
| Топ 5 | 9 | 2 | 1 | 9 | 8 |
| Топ 6 | 8 | 8 | 7 | 8 | 3 |
| Топ 7 | 6 | 6 | 6 | 6 | 1 |

В таблице 28 представлены результаты (ранжирование важности признаков), полученные различными методами XAI для моделей LightGBM и MLP, построенных для набора данных “H1N1”. Следует отметить, что мы не используем KernelSHAP для получения результатов объяснения из-за очень длительного

времени выполнения: для модели LightGBM требуется более 72 часов, а для модели MLP — более 654 часов. Как видно из таблицы 28, SamplingSHAP, LIME и ShapG дают одинаковый ранг для самого важного признака — “doctor recs h1n1 (10)” для модели LightGBM на наборе данных “H1N1”. SHAP, SamplingSHAP и ShapG ранжируют первые четыре признака в разном порядке, но набор этих признаков одинаков: “doctor recs h1n1 (10)”, “opinion h1n1 risk (17)”, “health insurance (15)” и “opinion h1n1 vacc effective (16)”. Однако, начиная с пятого признака, ранжирование признаков существенно отличается для различных методов XAI. Тем временем, объяснения, предоставленные методом FI, значительно отличаются от результатов других методов XAI. Для модели MLP видно, что методы SHAP, SamplingSHAP, LIME и Permutation Feature Importance (PFI) ранжируют признак “opinion h1n1 risk (17)” как самый важный, тогда как ShapG ставит этот признак на второе место. Из-за различий в объяснениях становится крайне важно оценивать и сравнивать эти методы XAI. Результаты сравнения показаны на рисунках 37 и 38.

Замечание 1. Отметим, что для модели LightGBM мы используем метод объяснения Feature Importance (FI), в отличие от метода Perturbation Feature Importance (PFI), который используется для модели MLP. Причина заключается в следующем. FI — метод объяснения, встроенный в древовидные модели, и он широко используется для объяснения моделей на основе деревьев. Однако для моделей MLP, которые не имеют древовидных структур для прямого вычисления значимости признаков, влияние признаков может быть оценено с помощью метода PFI более эффективно. Так как метод PFI не зависит от конкретной модели, его можно широко использовать для различных типов моделей в качестве замены метода FI для вычисления важности признаков. Таким образом, в наших экспериментах мы используем FI для модели LightGBM и PFI для модели MLP для обоих наборов данных.

Рисунки 35 и 36 показывают изменения R^2 после последовательного удаления признаков в различных методах XAI для моделей LightGBM и MLP соответственно, построенных для набора данных “housing price”. На основании этих данных можно интуитивно сравнить снижение точности и сделать вывод, что точность объяснения результатов, предоставленных предложенным методом XAI ShapG, выше, чем у других существующих методов.

Рисунки 37 и 38 показывают изменения точности после последовательного удаления признаков на основе различных методов XAI для моделей LightGBM

Таблица 28: Ранжирование важности признаков для набора данных “H1N1”

Модель LightGBM

| Ранг \ XAI | SHAP | SamplingSHAP | LIME | Feature Importance | ShapG |
|------------|------|--------------|------|--------------------|-------|
| Топ 1 | 15 | 10 | 10 | 30 | 10 |
| Топ 2 | 10 | 15 | 15 | 35 | 17 |
| Топ 3 | 16 | 17 | 16 | 34 | 15 |
| Топ 4 | 17 | 16 | 14 | 22 | 16 |
| Топ 5 | 20 | 20 | 27 | 16 | 35 |
| Топ 6 | 19 | 22 | 20 | 17 | 34 |
| Топ 7 | 11 | 30 | 17 | 31 | 14 |
| Топ 8 | 30 | 18 | 35 | 23 | 11 |
| Топ 9 | 35 | 35 | 18 | 20 | 20 |
| Топ 10 | 21 | 34 | 4 | 26 | 2 |

Модель MLP

| Ранг \ XAI | SHAP | SamplingSHAP | LIME | Permutation Feature Importance | ShapG |
|------------|------|--------------|------|--------------------------------|-------|
| Топ 1 | 17 | 17 | 17 | 17 | 10 |
| Топ 2 | 10 | 35 | 10 | 35 | 17 |
| Топ 3 | 20 | 21 | 28 | 10 | 11 |
| Топ 4 | 35 | 15 | 29 | 15 | 15 |
| Топ 5 | 34 | 10 | 32 | 11 | 14 |
| Топ 6 | 15 | 30 | 22 | 31 | 16 |
| Топ 7 | 16 | 31 | 11 | 18 | 20 |
| Топ 8 | 30 | 18 | 8 | 30 | 19 |
| Топ 9 | 27 | 22 | 30 | 29 | 35 |
| Топ 10 | 1 | 29 | 7 | 24 | 2 |

и MLP соответственно, построенных для набора данных “H1N1”. На рисунке 37 видно, что в результатах объяснения для модели LightGBM точность трёх методов: SHAP, SamplingSHAP и ShapG, демонстрирует схожую тенденцию к стабильному снижению. Хотя точность ShapG остаётся неизменной при удалении пяти и шести признаков, и в других ситуациях не наблюдается увеличения, стоит отметить, что точность SHAP увеличивается при удалении первых шести и восьми признаков, тогда как точность SamplingSHAP продолжает увеличиваться при удалении первых шести, семи и восьми признаков. Более того, хотя производительность объяснения ShapG сопоставима с большинством методов XAI, с точки зрения изменения точности после удаления первых десяти признаков, ShapG показывает лучшие результаты. Как видно на рисунке 38, ShapG значительно превосходит другие методы XAI в результатах объяснения на основе модели MLP.

Для сравнения времени выполнения методов XAI, основанных на теории игр, мы сравниваем KernelSHAP, SamplingSHAP и ShapG. Полученные результаты приведены ниже. Таблицы 29 и 30 показывают время выполнения методов

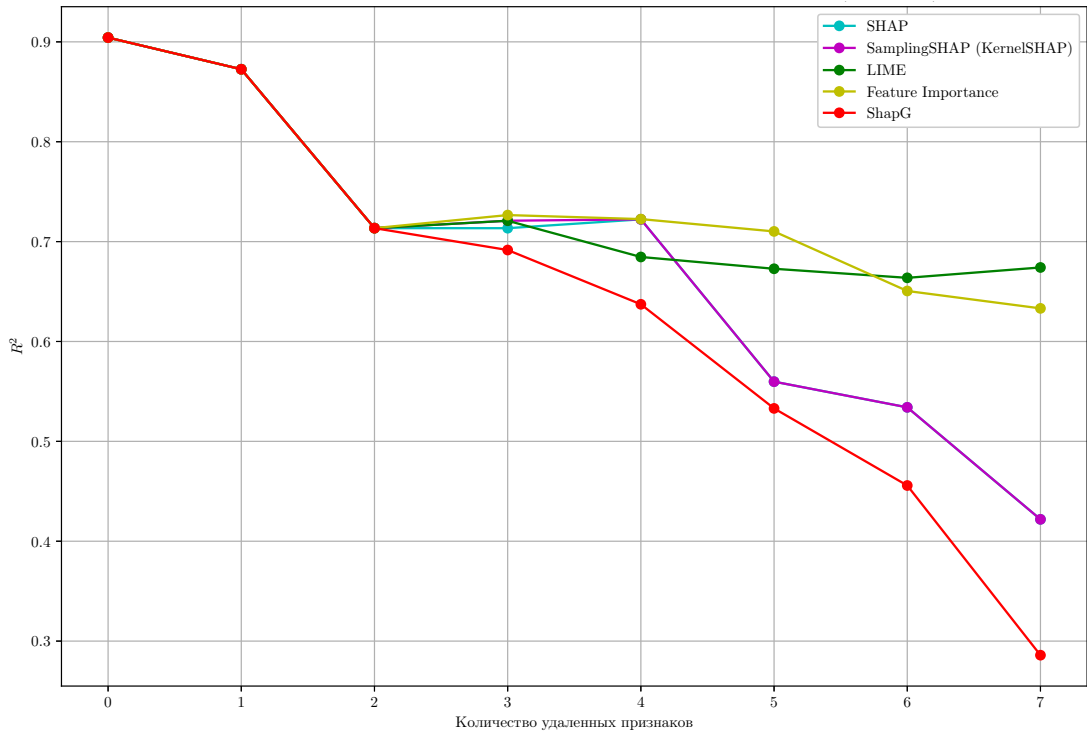


Рис. 35: Сравнение R^2 после удаления признаков на основе различных методов XAI в наборе данных “housing price” (LightGBM)

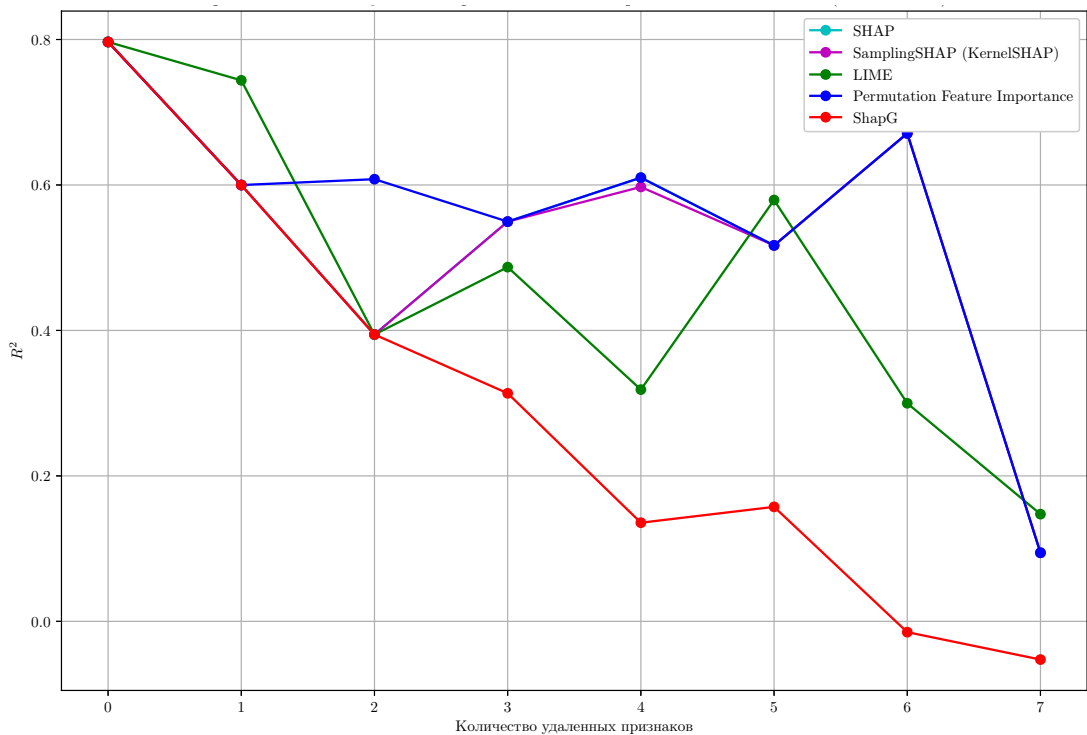


Рис. 36: Сравнение R^2 после удаления признаков на основе различных методов XAI в наборе данных “housing price” (модель MLP)

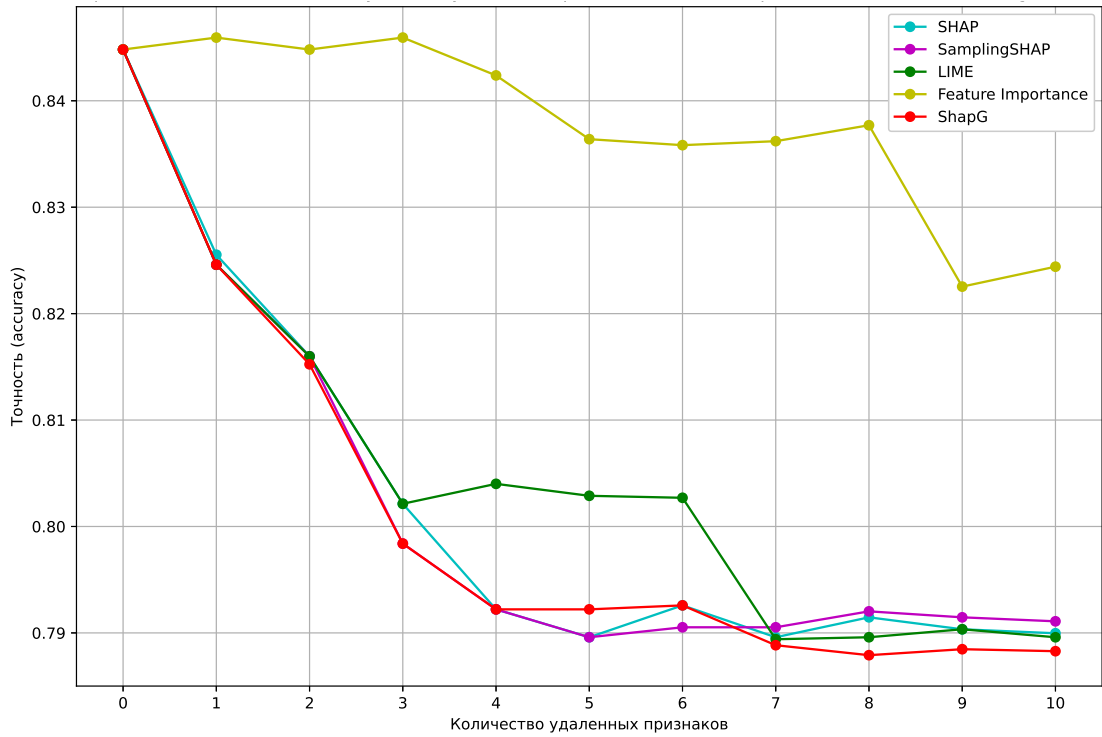


Рис. 37: Сравнение точности после удаления признаков на основе различных методов XAI в наборе данных “H1N1” (LightGBM)

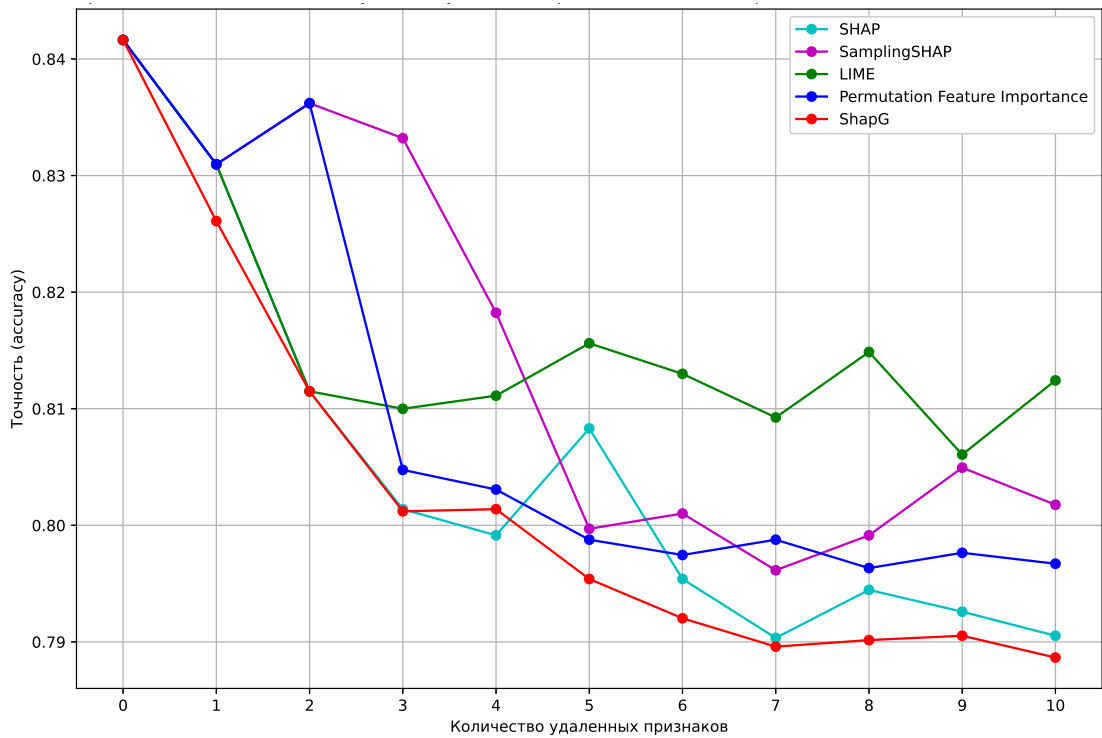


Рис. 38: Сравнение точности после удаления признаков на основе различных методов XAI в наборе данных “H1N1” (модель MLP)

XAI для моделей LightGBM и MLP на наборах данных “housing price” и “H1N1” соответственно. Сравнение ясно показывает, что метод ShapG значительно быстрее, чем такие методы теории игр, как KernelSHAP и SamplingSHAP. На наборе данных “housing price” с 13 признаками ShapG работает более чем в 5 (4.8) раз быстрее, чем KernelSHAP с моделью LightGBM (MLP). Мы не запускали KernelSHAP на наборе данных “H1N1” с 35 признаками, поскольку оценочное время его работы составляет примерно 4219 (39284) минут для модели LightGBM (MLP), что более чем в 161 (685) раз превышает время работы нашего метода ShapG.

Таблица 29: Время выполнения (в секундах) методов XAI для набора данных “housing price”

Модель LightGBM

| Метод XAI | KernelSHAP | SamplingSHAP | ShapG |
|------------------|------------|--------------|--------------|
| Время выполнения | 214.55 с | 42.14 с | 37.73 с |

Модель MLP

| Метод XAI | KernelSHAP | SamplingSHAP | ShapG |
|------------------|------------|--------------|--------------|
| Время выполнения | 901.51 с | 269.85 с | 184.22 с |

Таблица 30: Время выполнения методов XAI для набора данных “H1N1”

Модель LightGBM

| Метод XAI | KernelSHAP | SamplingSHAP | ShapG |
|------------------|------------|--------------|--------------|
| Время выполнения | ~ 4219 мин | 374 мин 87 с | 26 мин 12 с |

Модель MLP

| Метод XAI | KernelSHAP | SamplingSHAP | ShapG |
|------------------|-------------|--------------|--------------|
| Время выполнения | ~ 39284 мин | 364 мин 50 с | 57 мин 22 с |

Мы не рассматриваем SHAP в сравнительном анализе времени выполнения по нескольким причинам:

- При объяснении деревоподобных моделей (например, LightGBM) SHAP использует объяснитель TreeSHAP. TreeSHAP использует свойства самой деревовидной модели для быстрой оценки приближённых значений SHAP, что значительно улучшает вычислительную эффективность. Однако ограничение TreeSHAP заключается в том, что его можно использовать только

для объяснения древовидных моделей. Это означает, что мы не можем применять его к сложным моделям, таким как нейронные сети.

- SHAP может автоматически выбирать наиболее подходящую объясняющую переменную для различных моделей ИИ. Когда мы используем SHAP для объяснения модели MLP, SHAP вызывает `PermutationExplainer` для генерации результатов объяснения. Он работает, перебирая прямые и обратные перестановки признаков, изменяя признаки по одному, таким образом, эффективно оценивая независимый вклад каждого признака в конечный результат. `PermutationExplainer` не строит напрямую подмножества всех возможных признаков и не использует формулу вектора Шепли. Вместо этого он использует метод перестановки признаков для приближённого расчёта значения SHAP.

Сравнивая результаты оценки и время выполнения, наш метод XAI ShapG предоставляет значительное преимущество в эффективности и скорости. Это дополнительно подтверждает реализуемость и полезность алгоритма ShapG. Более того, в следующем разделе будет показано, что его можно использовать для предсказательных моделей с сложной архитектурой, когда другие методы XAI, основанные на теории игр, такие как `KernelSHAP` и `SamplingSHAP`, не работают.²⁰ Это делает метод ShapG высокопроизводительным инструментом для XAI.

4.4.3 Объяснение сложных моделей

Предложенный метод XAI ShapG может использоваться не только для одиночных моделей, таких как `LightGBM` или MLP, как показано в предыдущем разделе, но он также предоставляет глобальные объяснения для более сложных моделей. Эти сложные модели включают одиночные модели со сложными архитектурами, модели ансамблевого обучения, гибридные модели и т.д. Модели, которые мы используем в этом разделе, описаны в Разделе 4.3.3. Метод ShapG

²⁰Они могут дать результат, но время выполнения настолько велико, что делает их невозможными для практического применения, особенно для наборов данных с большим количеством признаков.

может применяться к широкому спектру моделей и может предоставлять объяснения их процессов принятия решений.

Таблицы 31 и 32 представляют ранжирование важности признаков, предоставленное ShapG для гибридных моделей, построенных путём комбинации различных типов моделей ИИ в парах, для наборов данных “housing price” и “H1N1” соответственно. Мы можем наблюдать незначительные различия в ранжировании важности признаков для различных моделей.

Таблица 31: Ранжирование важности признаков, вычисленное ShapG, для сложных моделей ИИ на наборе данных “housing price”

| Ранг \ Модель ИИ | LGB | Stacking | Linear-LGB | Linear-MLP | Linear-KNN | MLP-LGB | KNN-LGB | MLP-KNN |
|------------------|-----|----------|------------|------------|------------|---------|---------|---------|
| Топ 1 | 13 | 13 | 13 | 13 | 13 | 13 | 13 | 13 |
| Топ 2 | 6 | 5 | 6 | 6 | 6 | 6 | 6 | 6 |
| Топ 3 | 5 | 6 | 5 | 2 | 11 | 3 | 5 | 3 |
| Топ 4 | 11 | 3 | 11 | 3 | 3 | 5 | 3 | 11 |
| Топ 5 | 3 | 11 | 3 | 11 | 5 | 2 | 11 | 2 |
| Топ 6 | 2 | 2 | 2 | 8 | 2 | 11 | 2 | 5 |
| Топ 7 | 8 | 10 | 8 | 1 | 10 | 8 | 8 | 1 |

Для набора данных “housing price”, из таблицы 31 можно чётко наблюдать, что признак “LSTAT – низкий статус населения (13)” является самым важным признаком в прогнозировании цен на жильё, и это соответствует результатам, полученным на простых моделях LightGBM и MLP. Кроме того, можно заметить, что признаки “RM – среднее количество комнат на одно жилище (6)”, “INDUS – доля земель, занятых нежилым бизнесом, на город (3)”, и “NOX – концентрация оксидов азота (5)” чаще встречаются среди наиболее важных признаков для всех сложных моделей.

Для набора данных “H1N1” из таблицы 32 можно легко заметить, что “doctor recs h1n1 (10)” и “opinion h1n1 risk (17)” являются двумя наиболее важными признаками, что указывает на их значительное влияние на решение респондента о получении вакцины от гриппа H1N1. Аналогично, можно заметить, что “health insurance (15)” и “opinion h1n1 vacc effective (16)” чаще встречаются среди наиболее важных признаков. Это означает, что данные признаки также являются важными факторами, влияющими на намерение людей вакцинироваться.

Таблицы 33 и 34 показывают время выполнения алгоритма ShapG для объяснения сложных моделей ИИ на наборах данных “housing price” и “H1N1” соответственно. В зависимости от размера набора данных и сложности модели ShapG требует разного времени выполнения. Как и ожидалось, время выполне-

Таблица 32: Ранжирование важности признаков, вычисленное методом SharG, для сложных моделей ИИ на наборе данных “H1N1”

| Ранг \ Модель ИИ | LGB | Stacking | Logistic-LGB | Logistic-MLP | Logistic-KNN | MLP-LGB | KNN-LGB | MLP-KNN |
|------------------|-----|----------|--------------|--------------|--------------|---------|---------|---------|
| Топ 1 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| Топ 2 | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 |
| Топ 3 | 15 | 15 | 16 | 15 | 20 | 15 | 16 | 15 |
| Топ 4 | 16 | 16 | 11 | 14 | 16 | 11 | 15 | 14 |
| Топ 5 | 35 | 35 | 15 | 11 | 15 | 14 | 35 | 11 |
| Топ 6 | 34 | 14 | 14 | 20 | 11 | 35 | 11 | 24 |
| Топ 7 | 14 | 34 | 20 | 16 | 14 | 16 | 34 | 13 |
| Топ 8 | 11 | 11 | 19 | 5 | 24 | 20 | 14 | 20 |
| Топ 9 | 20 | 20 | 13 | 13 | 13 | 34 | 20 | 16 |
| Топ 10 | 2 | 19 | 1 | 1 | 29 | 24 | 2 | 35 |

Таблица 33: Время выполнения SharG для сложных моделей ИИ на наборе данных “housing price”

| Модель ИИ | LGB | Stacking | Linear-LGB | Linear-MLP | Linear-KNN | MLP-LGB | KNN-LGB | MLP-KNN |
|------------------|---------|----------|------------|------------|------------|----------|---------|----------|
| Время выполнения | 37.73 с | 457.37 с | 51.51 с | 364.32 с | 11.03 с | 536.28 с | 51.47 с | 377.75 с |

Таблица 34: Время выполнения SharG для сложных моделей ИИ на наборе данных “H1N1”

| Модель ИИ | LGB | Stacking | Logistic-LGB | Logistic-MLP | Logistic-KNN | MLP-LGB | KNN-LGB | MLP-KNN |
|------------------|-------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Время выполнения | 26 мин 12 с | 249 мин 35 с | 42 мин 31 с | 61 мин 17 с | 115 мин 02 с | 105 мин 03 с | 157 мин 16 с | 156 мин 44 с |

ния для набора данных “H1N1”, содержащего 35 признаков, значительно больше, чем для той же модели, построенной для набора данных “housing price” с 13 признаками. Хотя время выполнения для сложных моделей ИИ с использованием SharG значительно больше, чем для простых моделей (LightGBM или MLP), метод SharG имеет множество значительных преимуществ по сравнению с другими методами ХАИ. Метод SharG не только теоретически применим к любой модели, но и его эффективность проверена в многочисленных экспериментах со сложными моделями ИИ. Несмотря на то, что SHAP, SamplingSHAP и KernelSHAP предполагаются как применимые для объяснения любой модели, в ходе наших экспериментов мы обнаружили, что при применении этих методов к более сложным моделям нейронных сетей или гибридным моделям код часто не выполнялся успешно и не давал корректных результатов объяснения. Эти методы страдают от плохой совместимости со сложными моделями ИИ и требуют более глубоких настроек наборов данных или моделей ИИ. Поэтому для исследователей, не специализирующихся в области ИИ, эти методы не могут легко предоставить объяснимые результаты. При объяснении сложных моделей ИИ SharG по-прежнему способен предоставлять надёжные результаты объяснений.

4.5 Заключение к главе 4

В этой главе мы предложили новый метод объяснимого искусственного интеллекта под названием ShapG, основанный на векторе Шепли для игр на графах. Основная идея этого алгоритма вытекает из предложенных в главе 3 мер центральности. ShapG является независимым глобальным методом объяснения, который вычисляет значимость признаков, строя неориентированный граф признаков, где узлы графа представляют собой признаки, а выборки строятся на основе графа. Алгоритм начинается с пустого графа и последовательно добавляет рёбра, представляющие собой пары признаков с наибольшей корреляцией. Алгоритм останавливается, когда все признаки соединены, чтобы гарантировать, что граф признаков содержит важную структурную информацию. В процессе вычисления вектора Шепли необходимо учитывать только коалиции между каждым узлом и его достижимыми узлами, а не все возможные коалиции. Эта оптимизация повышает эффективность алгоритма.

Мы сравнили ShapG с несколькими популярными методами ХАИ, такими как Feature Importance (FI), Permutation Feature Importance (PFI), LIME, SHAP, SamplingSHAP и KernelSHAP. Метод ShapG демонстрирует отличные результаты объяснения значимости переменных, которые значительно лучше, чем у других методов ХАИ, что подтверждено на двух наборах данных. Кроме того, по сравнению с методами SamplingSHAP и KernelSHAP, также основанными на теории кооперативных игр, ShapG значительно сокращает затраты вычислительных ресурсов по времени выполнения. Эти результаты подтверждают надёжность и широкую применимость нашего метода.

ShapG можно рассматривать как полезный метод ХАИ, который может применяться не только к простым моделям ИИ, но и предоставлять глобальные объяснения для сложных моделей. Он способен надёжно объяснять процессы принятия решений сложными моделями, помогая пользователям лучше понимать эти модели.

Заключение

Диссертация посвящена анализу влияния структуры сети и поведения агентов на динамику мнений в социальных сетях. В ходе исследования динамика мнений при заданных различных сетевых структурах описана статистико-физическими (макроскопическими) моделями или микроскопическими моделями в предположении, что агенты могут взаимодействовать во внутреннем слое. Рассматриваются классические сетевые структуры с точки зрения динамики мнений в макроскопической GCVM и микроскопической GCVM с различными конфигурациями (главы 1 и 2). В главе 2 исследуется, как свойства сети и поведение индивидов влияют на динамику мнений. Для моделирования динамики мнений в двухслойной сети, основанной на реальной социальной сети (сеть карате-клуба Закари), использована микроскопическая GCVM. Влияние центральных важных узлов на динамику мнений обсуждается в главе 3. В этой главе также предложен подход к упрощению двухслойной сети до однослойной взвешенной сети и рассмотрены алгоритмы аппроксимации теоретико-игровых мер центральности. В главе 4 предложен новый метод глобального объяснения переменных для моделей машинного обучения, основанный на теоретико-игровых мерах центральности, представленных в главе 3.

Основные результаты работы следующие:

1. Сформулирован ряд моделей динамики мнений для различных сетевых структур (например, полный граф, циклический граф, звезда, двойная звезда с четным и нечетным количеством вершин, двойная клика с четным и нечетным количеством вершин для внутреннего слоя, а также полный граф и циклический граф для внешнего слоя). Модели основаны на методах статистической физики с привлечением различных параметров, связанных с поведением индивидов. Основные наблюдения: (i) если индивиды в социальной сети не склонны выражать мнения публично, внутреннее взаимодействие не оказывает большого влияния на вероятность выигрыша и время достижения консенсуса; (ii) простая структура внешнего слоя, например циклическая, значительно увеличивает время достижения консенсуса и может затруднять его достижение (см. [24]).
2. Применена предложенная микроскопическая GCVM для моделирования динамики мнений и анализа влияния типов поведения и их комбинаций на

динамику мнений. Результаты моделирования подтвердили корректность предложенной в главе 1 макроскопической модели. Рассмотрены три различных способа расширения циклической структуры внешнего слоя до полной, и проверена гипотеза о влиянии этого расширения на время достижения консенсуса и вероятность выигрыша определенного мнения. Основные результаты: (i) существует сильная линейная зависимость между средним кратчайшим путём и временем достижения консенсуса; (ii) способ расширения циклической структуры до полной существенно влияет на время достижения консенсуса и вероятность выигрыша; (iii) циклическая структура всегда увеличивает время достижения консенсуса; (iv) циклическая структура положительно влияет на вероятность выигрыша мнения; (v) параметры по-разному влияют на время достижения консенсуса, но практически не влияют на вероятность выигрыша. Например, выражение своего мнения до определённого порога неэффективно для достижения консенсуса, но при превышении этого порога желание выражать мнение облегчает системе достижение консенсуса. Когда люди охотнее принимают внешние мнения, это ускоряет консенсус; (vi) комбинация параметров оказывает значительное влияние на время достижения консенсуса (см. [25]).

3. Предложен подход к упрощению двухслойной сети с предопределённой динамикой мнений до однослойной взвешенной сети, применимый для стохастических процессов в двухслойных сетях с аналогичной динамикой. Разработаны алгоритмы аппроксимации теоретико-игровых мер центральности. Результаты тестирования показывают, что метод упрощения структуры до однослойной взвешенной и предложенные алгоритмы эффективны. Обнаружена значительная отрицательная корреляция между центральностями важных узлов и временем достижения консенсуса, что подчёркивает важность лидеров в социальной сети для достижения консенсуса (см. [51]).
4. Предложен новый алгоритм объяснимого искусственного интеллекта, рассматривающий признаки моделей машинного обучения как узлы неориентированного графа, построенного особым образом. Проведён ряд экспериментов, подтвердивших интерпретируемость, применимость и эффективность по времени выполнения. Алгоритм превосходит другие (включая PFI, FI, LIME и SHAP) по всем трём параметрам (см. [52]).

Задачи, сформулированные в диссертации, выполнены, и поставленные цели достигнуты.

Список сокращений и условных обозначений

BVM basic voter model.

CVM concealed voter model.

FI feature importance.

GCVM general concealed voter model.

KNN K-nearest neighbors.

KPI key performance indicator.

LIME local interpretable model-agnostic explanations.

MLP multilayer perceptron.

PCC Pearson correlation coefficient.

PFI permutation feature importance.

SHAP Shapley additive explanations.

ShapG explanations based on Shapley value for graphs.

XAI explainable artificial intelligence.

Список иллюстраций

| | | |
|----|---|----|
| 1 | Полная сеть, соединяющая 10 индивидов в BVM | 27 |
| 2 | Двухслойная сетевая структура в SVM с 10 индивидами | 28 |
| 3 | Представление GСVM с 10 индивидами: симметричный случай | 32 |
| 4 | Представление внутренней сети с центральной звездой | 35 |
| 5 | Представление внутренней двухзвёздной сети | 37 |
| 6 | Представление внутренней двухкликковой сети | 40 |
| 7 | Представление внешней сети–цикла | 43 |
| 8 | Представление внешней сети–цикл и внутренней двухзвёздной сети | 43 |
| 9 | Представление внешней циклической и внутренней двухкликковой сети | 44 |
| 10 | Время достижения консенсуса для 20 экспериментов, сравнение BVM, SVM и GСVM с полной внешней сетевой структурой | 46 |
| 11 | Время достижения консенсуса для 20 экспериментов, сравнение всех моделей | 46 |
| 12 | Наблюдаемое среднее время достижения консенсуса для всех мо- делей | 47 |
| 13 | Вероятность победы красного мнения для всех моделей | 48 |
| 14 | Наблюдаемое время достижения консенсуса (а) и вероятность по- беды (b) для моделей с разными структурами двухслойных сетей | 57 |
| 15 | Вероятность победы (а) и время достижения консенсуса (b) для моделей с различными способами расширения до полной сети | 61 |
| 16 | Время достижения консенсуса в зависимости от d | 62 |
| 17 | Коэффициенты корреляции Пирсона | 64 |
| 18 | Функция плотности распределения и эмпирическая функция рас- пределения (ECDF) для различных КРІ | 66 |
| 19 | Вероятность победы для различных вероятностей копирования | 69 |
| 20 | Вероятность победы при различных значениях вероятности экс- тернализации и интернализации | 71 |
| 21 | Время достижения консенсуса для различных вероятностей копи- рования | 72 |
| 22 | Время достижения консенсуса при различных вероятностях экс- тернализации и интернализации | 73 |

| | | |
|----|--|-----|
| 23 | Эмпирическая функция распределения времени достижения консенсуса в зависимости от параметров π_{c_e} , π_{c_i} , π_e и π_i | 75 |
| 24 | Однослойная сеть клуба каратэ Закари | 80 |
| 25 | Двухслойные сети, используемые в SVM и GCVM: (a) SVM: двухслойная сеть с внешним слоем — клуб каратэ Закари и пустым внутренним слоем, (b) GCVM: двухслойная сеть с внешним слоем — клуб каратэ Закари и внутренним слоем в виде звезды. | 81 |
| 26 | Средний внутренний кратчайший путь и плотность для различных структур сети | 104 |
| 27 | Меры центральности для различных структур | 105 |
| 28 | Вероятность выигрыша мнения и время достижения консенсуса для различных структур | 105 |
| 29 | Тепловая карта коэффициентов корреляции Пирсона для набора данных “housing price” | 119 |
| 30 | Тепловая карта коэффициентов корреляции Пирсона для набора данных “H1N1” | 120 |
| 31 | Граф, соединяющий признаки в наборе данных “housing price” | 121 |
| 32 | Граф, соединяющий признаки в наборе данных “H1N1” | 121 |
| 33 | Значимость признаков в наборе данных “housing price”, вычисленная с помощью SharG | 126 |
| 34 | Значимость признаков в наборе данных “H1N1”, вычисленная с помощью SharG | 127 |
| 35 | Сравнение R^2 после удаления признаков на основе различных методов ХАИ в наборе данных “housing price” (LightGBM) | 132 |
| 36 | Сравнение R^2 после удаления признаков на основе различных методов ХАИ в наборе данных “housing price” (модель MLP) | 132 |
| 37 | Сравнение точности после удаления признаков на основе различных методов ХАИ в наборе данных “H1N1” (LightGBM) | 133 |
| 38 | Сравнение точности после удаления признаков на основе различных методов ХАИ в наборе данных “H1N1” (модель MLP) | 133 |

Список таблиц

| | | |
|----|--|----|
| 1 | Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе SVM | 30 |
| 2 | Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GCVM в симметричном случае | 33 |
| 3 | Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GCVM с внутренней сетью, соединённой по принципу звезды | 36 |
| 4 | Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GCVM с двухзвёздной внутренней сетью (нечётный случай) | 38 |
| 5 | Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GCVM с внутренней двухзвёздной сетью (чётный случай) | 39 |
| 6 | Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GCVM с внутренними двумя кликовыми сетями (нечётный случай) | 41 |
| 7 | Переходы из состояния $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ в системе GCVM с двукликовым внутренним слоем (чётный случай) | 42 |
| 8 | Результаты тестов Колмогорова–Смирнова | 65 |
| 9 | Результаты тестов на нормальность | 65 |
| 10 | Результаты тестов на равенство дисперсий | 67 |
| 11 | Результаты тестов на равенство средних значений | 68 |
| 12 | Наблюдаемые минимальное и максимальное время достижения консенсуса | 70 |
| 13 | Наблюдаемые минимальная и максимальная вероятность победы | 70 |
| 14 | Описательная статистика кластеров по показателям выигрыша (complete-complete-50) | 74 |
| 15 | Время достижения консенсуса, размеры кластеров для каждой модели | 76 |
| 16 | Результаты для вектора Шепли при применении масштабирующего коэффициента ξ | 96 |
| 17 | Результаты для значения Майерсона без масштабирования с коэффициентом ξ | 97 |
| 18 | Результаты для значения Майерсона при применении масштабирования с коэффициентом ξ | 98 |
| 19 | Результаты для значения Майерсона без масштабирования с коэффициентом ξ для сетей с 10 узлами | 99 |

| | | |
|----|--|-----|
| 20 | Результаты для значения Майерсона при применении масштабирования с коэффициентом ξ для сетей с 10 узлами | 100 |
| 21 | Точный вектор Шепли против аппроксимированного вектора Шепли для “graph-1.0” | 101 |
| 22 | Точное значение Майерсона против аппроксимированного значения Майерсона для “graph-0.2” | 101 |
| 23 | Точность, определённая по формуле (32), для предложенных в этой работе и классических мер центральности | 102 |
| 24 | Коэффициенты корреляции | 107 |
| 25 | Нумерация признаков в наборе данных “housing price” | 128 |
| 26 | Нумерация признаков в наборе данных “H1N1” | 128 |
| 27 | Ранжирование важности признаков для набора данных “housing price” | 129 |
| 28 | Ранжирование важности признаков для набора данных “H1N1” | 131 |
| 29 | Время выполнения (в секундах) методов ХАИ для набора данных “housing price” | 134 |
| 30 | Время выполнения методов ХАИ для набора данных “H1N1” | 134 |
| 31 | Ранжирование важности признаков, вычисленное SharpG, для сложных моделей ИИ на наборе данных “housing price” | 136 |
| 32 | Ранжирование важности признаков, вычисленное методом SharpG, для сложных моделей ИИ на наборе данных “H1N1” | 137 |
| 33 | Время выполнения SharpG для сложных моделей ИИ на наборе данных “housing price” | 137 |
| 34 | Время выполнения SharpG для сложных моделей ИИ на наборе данных “H1N1” | 137 |

Список литературы

- [1] Wiener N. Cybernetics // Bulletin of the American Academy of Arts and Sciences. — 1950. — Vol. 3, no. 7. — P. 2–4.
- [2] Wiener N. The human use of human beings: cybernetics and society. — Houghton Mifflin, 1950.
- [3] McKeehan L. W. A contribution to the theory of ferromagnetism // Physical Review. — 1925. — Vol. 26, no. 2. — P. 274.
- [4] Holley R. A., Liggett T. M. Ergodic theorems for weakly interacting infinite systems and the voter model // The annals of probability. — 1975. — P. 643–663.
- [5] Noorazar H. Recent advances in opinion propagation dynamics: A 2020 survey // The European Physical Journal Plus. — 2020. — Vol. 135. — P. 1–20.
- [6] Sznajd-Weron K., Sznajd J. Opinion evolution in closed community // International Journal of Modern Physics C. — 2000. — Vol. 11, no. 06. — P. 1157–1165.
- [7] DeGroot M. H. Reaching a consensus // Journal of the American Statistical Association. — 1974. — Vol. 69, no. 345. — P. 118–121.
- [8] Friedkin N. E., Johnsen E. C. Social influence and opinions // Journal of Mathematical Sociology. — 1990. — Vol. 15, no. 3–4. — P. 193–206.
- [9] Hegselmann R., Ulrich K. Opinion dynamics and bounded confidence: models, analysis and simulation // Journal of Artificial Societies and Social Simulation (JASSS). — 2002. — Vol. 5, no. 3.
- [10] Mixing beliefs among interacting agents / Deffuant G., Neau D., Amblard F., and Weisbuch G. // Advances in Complex Systems. — 2000. — Vol. 3, no. 01n04. — P. 87–98.
- [11] Novel multidimensional models of opinion dynamics in social networks / Parsegov S. E., Proskurnikov A. V., Tempo R., and Friedkin N. E. // IEEE Transactions on Automatic Control. — 2016. — Vol. 62, no. 5. — P. 2270–2285.

- [12] Rogov M. A., Sedakov A. A. Coordinated Influence on the Opinions of Social Network Members // *Autom Remote Control*. — 2020. — Vol. 81. — P. 528–547.
- [13] Mazalov V. V., Parilina E. M. The Euler-Equation Approach in Average-Oriented Opinion Dynamics // *Mathematics*. — 2020. — Vol. 8, no. 3. — Access mode: <https://www.mdpi.com/2227-7390/8/3/355>.
- [14] Kareeva Y. S., Sedakov A. A., Zhen M. Influence in social networks with stubborn agents: From competition to bargaining // *Applied Mathematics and Computation*. — 2023. — Vol. 444. — P. 127790. — Access mode: <https://www.sciencedirect.com/science/article/pii/S009630032200858X>.
- [15] Opinion dynamics in finance and business: a literature review and research opportunities / Zha Q., Kou G., Zhang H., Liang H., Chen X., Li C.-C., and Dong Y. // *Financial Innovation*. — 2020. — Vol. 6. — P. 1–22.
- [16] Krasnoshchekov P. S. The simplest mathematical model of behaviour. Psychology of conformism // *Matematicheskoe Modelirovanie* [In Russian]. — 1998. — Vol. 10, no. 7. — P. 76–92.
- [17] Kozitsin I. V., Belolipetskii A. A. Opinion convergence in the Krasnoshchekov model // *The Journal of Mathematical Sociology*. — 2019. — Vol. 43, no. 2. — P. 104–121.
- [18] Bounded confidence opinion dynamics: A survey / Bernardo C., Altafini C., Proskurnikov A. V., and Vasca F. // *Automatica*. — 2024. — Vol. 159. — P. 111302. — Access mode: <https://www.sciencedirect.com/science/article/pii/S0005109823004661>.
- [19] Éminence Grise Coalitions: On the Shaping of Public Opinion / Bolouki S., Malhamé R. P., Siami M., and Motee N. // *IEEE Transactions on Control of Network Systems*. — 2017. — Vol. 4, no. 2. — P. 133–145.
- [20] A survey on the fusion process in opinion dynamics / Dong Y., Zhan M., Kou G., Ding Z., and Liang H. // *Information Fusion*. — 2018. — Vol. 43. — P. 57–65.

- [21] Gastner M. T., Oborny B., Gulyás M. Consensus time in a voter model with concealed and publicly expressed opinions // *Journal of Statistical Mechanics: Theory and Experiment*. — 2018. — Vol. 2018, no. 6. — P. 063401.
- [22] The impact of hypocrisy on opinion formation: A dynamic model / Gastner M. T., Takács K., Gulyás M., Szvetelszky Z., and Oborny B. // *PloS one*. — 2019. — Vol. 14, no. 6. — P. e0218729.
- [23] Donati S., Zappalà S., González-Romá V. The influence of friendship and communication network density on individual innovative behaviours: a multi-level study // *European Journal of Work and Organizational Psychology*. — 2016. — Vol. 25, no. 4. — P. 583–596.
- [24] Zhao C., Parilina E. M. Opinion Dynamics in Two-Layer Networks with Hypocrisy // *Journal of the Operations Research Society of China*. — 2024. — Mar. — Vol. 12, no. 1. — P. 109–132. — Access mode: <https://doi.org/10.1007/s40305-023-00503-2>.
- [25] Zhao C., Parilina E. M. Analysis of consensus time and winning rate in two-layer networks with hypocrisy of different structures // *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*. — 2024. — Vol. 20, no. 2. — P. 170–192.
- [26] Zhao C., Parilina E. M. Network Structure Properties and Opinion Dynamics in Two-Layer Networks with Hypocrisy // *Mathematical Optimization Theory and Operations Research*. — Cham : Springer Nature Switzerland. — 2024. — P. 300–314.
- [27] Zhao C., Parilina E. M. Consensus time and winning rate based on simulations in two-layer networks with hypocrisy // *2023 7th Scientific School Dynamics of Complex Networks and their Applications (DCNA)*. — 2023. — P. 68–71.
- [28] Mazalov V. V., Chirkova J. V. *Networking games: network forming games and games on networks*. — Academic Press, 2019. — P. 117–118.
- [29] Freeman L. C. A set of measures of centrality based on betweenness // *Sociometry*. — 1977. — P. 35–41.

- [30] Bavelas A. Communication patterns in task-oriented groups // The journal of the acoustical society of America. — 1950. — Vol. 22, no. 6. — P. 725–730.
- [31] Freeman L. C. Centrality in social networks conceptual clarification // Social Networks. — 1978. — Vol. 1, no. 3. — P. 215–239. — Access mode: <https://www.sciencedirect.com/science/article/pii/0378873378900217>.
- [32] Sabidussi G. The centrality index of a graph // Psychometrika. — 1966. — Vol. 31, no. 4. — P. 581–603.
- [33] Powell J., Hopkins M. 9 - Library networks—coauthorship, citation, and usage graphs // A Librarian’s Guide to Graphs, Data and the Semantic Web. — Chandos Publishing, 2015. — Chandos Information Professional Series. — P. 75–81. — Access mode: <https://www.sciencedirect.com/science/article/pii/B9781843347538000099>.
- [34] Random walk centrality in interconnected multilayer networks / Solé-Ribalta A., De Domenico M., Gómez S., and Arenas A. // Physica D: Non-linear Phenomena. — 2016. — Vol. 323. — P. 73–79.
- [35] Newman M. A measure of betweenness centrality based on random walks // Social networks. — 2005. — Vol. 27, no. 1. — P. 39–54.
- [36] Game-theoretic network centrality: A review / Tarkowski M. K., Michalak T. P., Rahwan T., and Wooldridge M. // arXiv preprint arXiv:1801.00218. — 2017.
- [37] Shapley L. S. A value for n-person games // Contribution to the Theory of Games. — 1953. — Vol. 2.
- [38] Myerson R. B. Graphs and cooperation in games // Mathematics of operations research. — 1977. — Vol. 2, no. 3. — P. 225–229.
- [39] Suri N. R., Narahari Y. Determining the top-k nodes in social networks using the shapley value // Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 3. — 2008. — P. 1509–1512.
- [40] Game-theoretic centrality measures for weighted graphs / Mazalov V. V., Avrachenkov K. E., Trukhina L. I., and Tsynguev B. T. // Fundamenta Informaticae. — 2016. — Vol. 145, no. 3. — P. 341–358.

- [41] Mazalov V. V., Khitraya V. A. A modified Myerson value for determining the centrality of graph vertices // Automation and Remote Control. — 2021. — Vol. 82. — P. 145–159.
- [42] Khitraya V. A., Mazalov V. V. Game-Theoretic Centrality of Directed Graph Vertices // Automation And Remote Control. — 2024. — Vol. 85, no. 2.
- [43] Boldi P., Vigna S. Axioms for centrality // Internet Mathematics. — 2014. — Vol. 10, no. 3-4. — P. 222–262.
- [44] Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI / Arrieta A. B., Díaz-Rodríguez N., Del Ser J., Bennetot A., Tabik S., Barbado A., García S., Gil-López S., Molina D., Benjamins R., et al. // Information fusion. — 2020. — Vol. 58. — P. 82–115.
- [45] Plumb G., Molitor D., Talwalkar A. S. Model agnostic supervised local explanations // Advances in neural information processing systems. — 2018. — Vol. 31.
- [46] Global explanations of neural networks: Mapping the landscape of predictions / Ibrahim M., Louie M., Modarres C., and Paisley J. // Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. — 2019. — P. 279–287.
- [47] Ribeiro M. T., Singh S., Guestrin C. ” Why should i trust you?” Explaining the predictions of any classifier // Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. — 2016. — P. 1135–1144.
- [48] Lundberg S. M., Lee S.-I. A unified approach to interpreting model predictions // Advances in neural information processing systems. — 2017. — Vol. 30.
- [49] From local explanations to global understanding with explainable AI for trees / Lundberg S. M., Erion G., Chen H., DeGrave A., Prutkin J. M., Nair B., Katz R., Himmelfarb J., Bansal N., and Lee S.-I. // Nature machine intelligence. — 2020. — Vol. 2, no. 1. — P. 56–67.

- [50] Zhao C. shapG. — PyPI: <https://pypi.org/project/shapG/>; GitHub: <https://github.com/vectorsss/shapG>. — 2024. — July. — (Accessed on 11/16/2024).
- [51] Zhao C., Parilina E. M. Centrality measures and opinion dynamics in two-layer networks with replica nodes // arXiv preprint arXiv:2406.18780v2. — 2024. — 2406.18780v2.
- [52] Zhao C., Liu J., Parilina E. M. ShapG: new feature importance method based on the Shapley value // arXiv preprint arXiv:2407.00506. — 2024. — 2407.00506.
- [53] Свидетельство о государственной регистрации программы для ЭВМ № 2023661532 Российская Федерация : Программа для моделирования динамики распространения бинарных мнений в двухслойных сетях Чжао Ч. (CN) ; федеральное государственное бюджетное образовательное учреждение высшего образования “Санкт-Петербургский государственный университет” (СПбГУ) ; Чжао Чи. — № 2023660681 ; заявл. 24.05.2023 ; опубли. 01.06.2023.
- [54] Festinger L. A theory of cognitive dissonance. — Stanford : Stanford University Press, 1957.
- [55] Synchronization and graph topology / Belykh I., Hasler M., Lauret M., and Nijmeijer H. // International Journal of Bifurcation and Chaos. — 2005. — Vol. 15, no. 11. — P. 3423–3433.
- [56] Bianconi G. Multilayer networks: structure and function. — Oxford university press, 2018.
- [57] SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python / Virtanen P., Gommers R., Oliphant T. E., Haberland M., Reddy T., Cournapeau D., Burovski E., Peterson P., Weckesser W., Bright J., van der Walt S. J., Brett M., Wilson J., Millman K. J., Mayorov N., Nelson A. R. J., Jones E., Kern R., Larson E., Carey C. J., Polat İ., Feng Y., Moore E. W., VanderPlas J., Laxalde D., Perktold J., Cimrman R., Henriksen I., Quintero E. A., Harris C. R., Archibald A. M., Ribeiro A. H., Pedregosa F., van Mulbregt P., and SciPy 1.0 Contributors // Nature Methods. — 2020. — Vol. 17. — P. 261–272.

- [58] `scipy.stats.pearsonr` – SciPy v1.12.0 Manual // <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.pearsonr.html>. — (Accessed on 02/11/2024).
- [59] Student. Probable Error Of A Correlation Coefficient // *Biometrika*. — 1908. — Sep. — Vol. 6, no. 2-3. — P. 302–310. — https://academic.oup.com/biomet/article-pdf/6/2-3/302/49729619/biomet_6_2-3_302.pdf.
- [60] Simard R., L’Ecuyer P. Computing the two-sided Kolmogorov-Smirnov distribution // *Journal of Statistical Software*. — 2011. — Vol. 39. — P. 1–18.
- [61] `scipy.stats.kstest` – SciPy v1.12.0 Manual // <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kstest.html>. — (Accessed on 02/11/2024).
- [62] Shapiro S. S., Wilk M. B. An analysis of variance test for normality (complete samples) // *Biometrika*. — 1965. — Vol. 52, no. 3/4. — P. 591–611.
- [63] `scipy.stats.normaltest` – SciPy v1.12.0 Manual // <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.normaltest.html>. — (Accessed on 02/11/2024).
- [64] Levene H. Contributions to probability and statistics: essays in honor of Harold Hotelling. — Stanford University Press, 1960. — P. 278–292.
- [65] Bartlett M. S. Properties of sufficiency and statistical tests // *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences*. — 1937. — Vol. 160, no. 901. — P. 268–282.
- [66] Fligner M. A., Killeen T. J. Distribution-free two-sample tests for scale // *Journal of the American Statistical Association*. — 1976. — Vol. 71, no. 353. — P. 210–213.
- [67] `scipy.stats.levene` – SciPy v1.12.0 Manual // <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.levene.html>. — (Accessed on 02/11/2024).
- [68] Kruskal W. H., Wallis W. A. Use of ranks in one-criterion variance analysis // *Journal of the American statistical Association*. — 1952. — Vol. 47, no. 260. — P. 583–621.

- [69] Heiman G. W. Understanding research methods and statistics: An integrated introduction for psychology. — Houghton, Mifflin and Company, 2001.
- [70] Alexander R. A., Govern D. M. A new and simpler approximation for ANOVA under variance heterogeneity // Journal of Educational Statistics. — 1994. — Vol. 19, no. 2. — P. 91–101.
- [71] scipy.stats.kruskal - SciPy v1.12.0 Manual // <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kruskal.html>. — (Accessed on 02/11/2024).
- [72] Fast Exact k-Means, k-Medians and Bregman Divergence Clustering in 1D. — 2018. — Apr. — arXiv:1701.07204 [cs]. online; accessed: <http://arxiv.org/abs/1701.07204> (online; accessed: 2023-10-24).
- [73] Scikit-learn: Machine Learning in Python / Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., and Duchesnay E. // Journal of Machine Learning Research. — 2011. — Vol. 12. — P. 2825–2830.
- [74] Rousseeuw P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis // Journal of computational and applied mathematics. — 1987. — Vol. 20. — P. 53–65.
- [75] Zachary W. W. An information flow model for conflict and fission in small groups // Journal of anthropological research. — 1977. — Vol. 33, no. 4. — P. 452–473.
- [76] Dense graph - Wikipedia // https://en.wikipedia.org/wiki/Dense_graph. — (Accessed on 02/20/2024).
- [77] Everett M. G., Borgatti S. P. The centrality of groups and classes // The Journal of mathematical sociology. — 1999. — Vol. 23, no. 3. — P. 181–201.
- [78] Everett M. G., Borgatti S. P. Extending centrality // Models and methods in social network analysis. — 2005. — Vol. 35, no. 1. — P. 57–76.
- [79] Measuring and maximizing group closeness centrality over disk-resident graphs / Zhao J., Lui J. C., Towsley D., and Guan X. // Proceedings of the 23rd International Conference on World Wide Web. — 2014. — P. 689–694.

- [80] Chung F. R. Spectral graph theory. — American Mathematical Soc., 1997. — Vol. 92.
- [81] Newman M. Networks: An Introduction. — Oxford University Press, 2010. — Mar. — ISBN: 9780199206650. — Access mode: <https://doi.org/10.1093/acprof:oso/9780199206650.001.0001>.
- [82] Freeman L. C., Borgatti S. P., White D. R. Centrality in valued graphs: A measure of betweenness based on network flow // Social networks. — 1991. — Vol. 13, no. 2. — P. 141–154.
- [83] Stephenson K., Zelen M. Rethinking centrality: Methods and examples // Social networks. — 1989. — Vol. 11, no. 1. — P. 1–37.
- [84] Lovász L. Random walks on graphs: A survey // Combinatorics, Paul Erdos is Eighty. — 1993. — Vol. 2, no. 1. — P. 1–46. — Access mode: http://scholar.google.de/scholar.bib?q=info:1lcRghStI1oJ:scholar.google.com/&output=citation&hl=de&as_sdt=0,5&ct=citation&cd=3.
- [85] Mean first-passage time for random walks on undirected networks / Zhang Z., Julaiti A., Hou B., Zhang H., and Chen G. // The European Physical Journal B. — 2011. — Vol. 84. — P. 691–697.
- [86] Kemeny J. G., Snell J. L. et al. Finite markov chains. — van Nostrand Princeton, NJ, 1969. — Vol. 26.
- [87] Pólya G. Eine Wahrscheinlichkeitsaufgabe in der Kundenwerbung // Zamm-zeitschrift Fur Angewandte Mathematik Und Mechanik. — 1930. — Vol. 10. — P. 96–97. — Access mode: <https://api.semanticscholar.org/CorpusID:123056679>.
- [88] Jackson M. O. et al. Social and economic networks. — Princeton university press Princeton, 2008. — Vol. 3.
- [89] Mazalov V. V., Trukhina L. I. Generating functions and the Myerson vector in communication networks // Discrete Mathematics and Applications. — 2014. — Vol. 24, no. 5. — P. 295–303.
- [90] Milgram S. The small world problem // Psychology today. — 1967. — Vol. 2, no. 1. — P. 60–67.

- [91] Kendall M. G. The treatment of ties in ranking problems // *Biometrika*. — 1945. — Vol. 33, no. 3. — P. 239–251.
- [92] Spearman C. The proof and measurement of association between two things // *The American journal of psychology*. — 1987. — Vol. 100, no. 3/4. — P. 441–471.
- [93] Louppe G. Understanding random forests: From theory to practice // *arXiv preprint arXiv:1407.7502*. — 2014.
- [94] Permutation importance: a corrected feature importance measure / Altmann A., Tološi L., Sander O., and Lengauer T. // *Bioinformatics*. — 2010. — Vol. 26, no. 10. — P. 1340–1347.
- [95] Mishra S., Sturm B. L., Dixon S. Local interpretable model-agnostic explanations for music content analysis. // *Ismir*. — 2017. — Vol. 53. — P. 537–543.
- [96] Štrumbelj E., Kononenko I. An efficient explanation of individual classifications using game theory // *The Journal of Machine Learning Research*. — 2010. — Vol. 11. — P. 1–18.
- [97] Castro J., Gómez D., Tejada J. Polynomial calculation of the Shapley value based on sampling // *Computers & Operations Research*. — 2009. — Vol. 36, no. 5. — P. 1726–1730.
- [98] Shapley L. S. Notes on the n-person game—ii: The value of an n-person game. — Rand Corporation, 1951.
- [99] Naumova N. I. Shapley Value and its Extensions. — VVM Publishing House, 2017. — P. 60.
- [100] Štrumbelj E., Kononenko I. Explaining prediction models and individual predictions with feature contributions // *Knowledge and information systems*. — 2014. — Vol. 41. — P. 647–665.
- [101] Lightgbm: A highly efficient gradient boosting decision tree / Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q., and Liu T.-Y. // *Advances in neural information processing systems*. — 2017. — Vol. 30.

- [102] Samatin Njikam A. N., Zhao H. A novel activation function for multilayer feed-forward neural networks // *Applied Intelligence*. — 2016. — Vol. 45. — P. 75–82.
- [103] Polikar R. Ensemble learning // *Ensemble machine learning: Methods and applications*. — 2012. — P. 1–34.
- [104] Stacking ensemble learning for short-term electricity consumption forecasting / Divina F., Gilson A., Gómez-Vela F., García Torres M., and Torres J. F. // *Energies*. — 2018. — Vol. 11, no. 4. — P. 949.
- [105] Weisberg S. *Applied linear regression*. — John Wiley & Sons, 2005. — Vol. 528.
- [106] LaValley M. P. Logistic regression // *Circulation*. — 2008. — Vol. 117, no. 18. — P. 2395–2399.
- [107] Taud H., Mas J.-F. Multilayer perceptron (MLP) // *Geomatic approaches for modeling land change scenarios*. — 2018. — P. 451–455.
- [108] KNN model-based approach in classification / Guo G., Wang H., Bell D., Bi Y., and Greer K. // *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2003, Catania, Sicily, Italy, November 3-7, 2003. Proceedings* / Springer. — 2003. — P. 986–996.
- [109] Towards a rigorous evaluation of XAI methods on time series / Schlegel U., Arnout H., El-Assady M., Oelke D., and Keim D. A. // *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)* / Ieee. — 2019. — P. 4197–4201.

Приложение

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2023661532

**«Программа для моделирования динамики
распространения бинарных мнений в двухслойных
сетях» (TL-Net OD)**

Правообладатель: *федеральное государственное бюджетное
образовательное учреждение высшего образования
"Санкт-Петербургский государственный университет"
(СПбГУ) (RU)*

Автор(ы): *Чжао Чи (CN)*



Заявка № 2023660681

Дата поступления 24 мая 2023 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 01 июня 2023 г.

Руководитель Федеральной службы
по интеллектуальной собственности

Ю.С. Зубов