Федеральное государственное бюджетное учреждение науки
Институт проблем управления им. В. А. Трапезникова
Российской академии наук

УДК 519.8

На правах рукописи

Губанов Дмитрий Алексеевич

МОДЕЛИ И МЕТОДЫ ИНФОРМАЦИОННОГО ВЛИЯНИЯ И УПРАВЛЕНИЯ В АКТИВНЫХ СЕТЕВЫХ СТРУКТУРАХ

Специальность 05.13.10 – Управление в социальных и экономических системах

Автореферат диссертации на соискание ученой степени доктора технических наук

Москва – 2021

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институте проблем управления им. В. А. Трапезникова Российской академии наук (ИПУ РАН)

Научный консультант: доктор физико-математических наук
Чхартишвили Александр Гедеванович

Официальные оппоненты: **Леонидов Андрей Владимирович**,
доктор физико-математических наук,
ФГБУН Физический институт
им. П.Н. Лебедева Российской академии наук,
заведующий Лабораторией математического моделирования сложных систем.

Мальцева Светлана Валентиновна,
доктор технических наук, Национальный исследовательский университет
"Высшая школа экономики"
Московский институт электроники
и математики, профессор.

Печников Андрей Анатольевич,
доктор технических наук,
Институт прикладных математических исследований – обособленное подразделение ФГБУН ФИЦ "Карельский научный центр Российской академии наук", руководитель Лаборатории телекоммуникационных систем.

Ведущая организация: ФГБУН Институт системного программирования им. В.П. Иванникова Российской академии наук.

Защита состоится «16» декабря 2021 г. в 16:00 на заседании Диссертационного Совета Д 002.226.02 ФГБУН Института проблем управления им. В. А. Трапезникова РАН по адресу: 117997, Москва, ул. Профсоюзная, 65.

С диссертацией можно ознакомиться в библиотеке и на сайте ИПУ РАН (https://www.ipu.ru).

Автореферат разослан ____________________________________ 2021 г.

Ученый секретарь
Диссертационного Совета Д 002.226.02
кандидат физико-математических наук Мусатова Е.Г.

Общая характеристика работы

Актуальность темы. *Активные сетевые структуры* (АСС) – социальные структуры, состоящие из множества *агентов* (индивидуальных и коллективных субъектов) и определенного на нем множества *отношений* (совокупности связей между агентами, например, коммуникации), являются предметом активных исследований, начиная со второй половины двадцатого века. Частными случаями АСС являются онлайновая социальная сеть, толпа, социальная группа или сеть авторов научных публикаций. При моделировании АСС возникает необходимость их анализа, в том числе, как *сетей информационного влияния* – учета взаимного влияния членов сети и динамики их мнений под воздействием этого влияния.

С середины XX века разрабатываются модели информационного влияния в социальных сетях, имеющие как микроэкономические, так и социально-психологические основания. Развивается направление микроэкономических исследований, в котором моделируется поведение рациональных агентов, наблюдающих действия других агентов и стремящихся устранить неопределенность своих представлений относительно изучаемого вопроса (Bikhchandani S., Vives X., Burguet R., Acemoglu D., Ozdaglar A.). Психологи, однако, отмечают, что индивиды обладают ограниченной рациональностью и допускают систематические ошибки, влияющие на обработку информации (Tversky A., Kahneman D., Simon H.A.). В этом плане более релевантным является «эвристическое» направление исследований, которое основывается на эмпирических закономерностях изменения мнений индивидов и демонстрирует наблюдаемые на практике социально-психологические эффекты (Hunter J.E., French J. R., DeGroot M.H., Friedkin N.E., Deffuant G., Hegselmann R., Krause U., Jackson M.O., Flache A., Новиков Д.А., Чхартишвили А.Г., Проскурников А.В., Михайлов А.П., Петров А.П., и многие другие). Отметим также направление исследований многоагентных систем (Cao Y., Yu W., Ren W., Chen G., Чеботарев П.Ю., Агаев Р.П. и др.), в котором изучаются вопросы анализа и управления многоагентными системами.

В то же время к середине 2000-х годов в теории управления организационными системами были получены важные результаты изучения механизмов информационного управления – воздействия на информированность участников системы (Кульба В.В., Новиков Д.А., Чхартишвили А.Г. и др.). Эти результаты не в полной мере учитывают структуру АСС и ряд других ее свойств, что привело к началу разработки и применения механизмов информационного управления и противоборства в социальных сетях к концу 2000-х (Новиков Д.А., Чхартишвили А.Г.). В настоящее время теоретико-игровые исследования информационного противоборства в социальных сетях проводятся рядом (в основном, судя по открытым публикациям, отечественных) исследователей (Мазалов В.В., Буре В.М., Парилина Е.М., Седаков А.А. и др.), однако эти исследования в целом ограничиваются случаем оказания воздействий на мнения простых (в смысле внутренней структуры) агентов в игре с двумя центрами.

Таким образом, актуальной является разработка моделей информационного влияния в АСС с учетом компонент внутренней структуры агента, а также сопряженных с ними моделей информационного управления и противоборства.

Объектом исследования в диссертационной работе являются активные сетевые структуры, **предметом исследования** – информационное влияние и управление в таких структурах.

Цель работы состоит в разработке и исследовании математических моделей, методов и технологий анализа информационного влияния и синтеза эффективного информационного управления в АСС.

Реализация поставленной цели предполагает решение следующих **основных задач**:

1. Выявление специфики АСС как объектов управления; формулировка задач информационного управления в АСС.

2. Разработка и исследование моделей и методов информационного влияния в АСС с учетом значимых для информационного взаимодействия компонент внутренней структуры агента.

3. Разработка и исследование моделей и методов информационного управления в АСС.

4. Разработка и исследование моделей и методов информационного противоборства в АСС.

5. Разработка и исследование моделей и методов анализа информационного влияния и влиятельности структурных элементов АСС.

6. Разработка технологий анализа информационного влияния в АСС.

7. Внедрение полученных результатов в практику информационного управления активными сетевыми структурами.

Методы исследования. В основе исследования лежит аппарат современной теории управления, в частности подходы и результаты теории игр, теории активных систем, теории принятия решений и исследования операций. Для формализации объекта исследования используются методы математического моделирования.

Научная новизна работы состоит в разработке единого методологического подхода к созданию и исследованию моделей информационного влияния и управления в активных сетевых структурах, заключающегося в общности описания компонент внутренней структуры агентов и иерархическом способе построения (от моделей информационного взаимодействия агентов до моделей управления и противоборства). На основе разработанного подхода

1. Предложена классификация моделей информационного влияния в АСС.

2. Предложено семейство взаимосвязанных математических моделей динамики мнений агентов АСС в условиях информационного влияния со стороны окружения в сети. В базовой модели описывается изменение мнений доверчивых агентов, в более сложных моделях интеллектуальность агентов возрастает: агенты начинают различать доверие к другим агентам и доверие к содержанию их конкретных действий (осторожные агенты), а также вести себя рационально, моделируя информированность и поведение других агентов и сообщая информацию, приводящую к принятию выгодных им решений в АСС (манипулирующие и рефлексивные агенты).

3. Сформулированы и решены задачи информационного управления (воздействия на АСС с целью формирования требуемых мнений агентов) для всех разработанных моделей динамики мнений. В этих задачах предметом управления являются различные компоненты внутренней структуры агента АСС.

4. Предложен подход к моделированию и анализу информационного влияния в АСС на основе совершаемых агентами действий и учета интересов управляющего субъекта (акциональный подход). Формализованы различные случаи влияния и влиятельности агентов и структур АСС, исследованы свойства предложенных функций влияния и влиятельности. Поставлены и решены примеры задач прогноза и управления действиями агентов АСС.

5. Сформулирована задача информационного противоборства, для которой построена общая теоретико-игровая модель

противоборствующих субъектов, оказывающих управляющие информационные воздействия на АСС. Исследован ряд ее частных случаев, в том числе задача распределенного контроля в АСС, для которой охарактеризованы режимы информационной кооперации и информационной войны управляющих субъектов, а также задача «защита-нападение» в АСС, для которой приведены решения для ситуации различной информированности и рефлексии управляющих субъектов.

6. Рассмотрены случаи информационного противоборства в АСС, различающиеся информированностью субъектов, структурой их целевых функций и порядком функционирования. Для них построены и проанализированы теоретико-игровые модели, в которых решением игры является: равновесие в доминантных стратегиях, равновесие Нэша, «контрактное равновесие», равновесие Штакельберга, информационное равновесие и равновесие в безопасных стратегиях.

7. Разработаны прикладные методы и алгоритмы анализа АСС на основе моделей информационного влияния для расчета влияния и влиятельности элементов АСС, выявления структур и устойчивых каналов распространения действий в АСС, анализа защищенности агентов АСС от информационных воздействий, а также выявления сообществ в АСС.

Теоретическая и практическая значимость. Результаты диссертационной работы позволяют разрабатывать и обосновывать механизмы эффективного информационного влияния и управления в АСС. Общность подхода к разработке моделей информационного влияния, управления и противоборства в АСС позволяет распространить полученные теоретические и практические результаты на широкий круг активных сетевых структур и может служить основой переноса решения практических задач из одних областей в другие.

Эффективность использования разработанных в диссертационной работе моделей и методов информационного влияния и управления подтверждена актами и справками о внедрении. Результаты работы также использовались в ряде проектов по спецтематике. Кроме того, получен ряд свидетельств о регистрации программ.

Основные положения, выносимые на защиту.

1. Предложена концепция информационного влияния, управления и противоборства в АСС, в ней рассматривается новый объект

управления – активная сетевая структура. В АСС отсутствует иерархия, а ее элементы (агенты) обладают внутренней активностью, которую проявляют в самостоятельно выбираемых действиях. Концепция позволяет раскрыть объект исследования в его основных сущностных аспектах и перейти от фрагментарной разработки моделей, описывающих информационное влияние и управление, к целостному их рассмотрению.

2. На основе концепции разработан подход к созданию и исследованию моделей информационного влияния и управления в активных сетевых структурах. В зависимости от решаемой задачи в рамках подхода можно сначала комплексировать модель исследуемого объекта (выбрав компоненты внутренней структуры агента), а затем построить требуемую иерархию моделей (от моделей информационного взаимодействия агентов до моделей управления и противоборства).

3. В рамках реализации подхода предложено семейство взаимосвязанных математических моделей динамики мнений агентов АСС. В совокупности модели покрывают все описываемые в концепции компоненты внутренней структуры агента, механизмы формирования мнений агентов и принятия ими решений. Модели позволяют учитывать различные эффекты на микроуровне и – как следствие – моделировать возникновение ключевых социально-психологических макро-эффектов в АСС. Проведен анализ разработанных моделей.

4. Поставлены, исследованы и решены задачи информационного управления АСС, в которых предметом управления являются различные компоненты внутренней структуры агента АСС. Показано, что управляющий субъект в каждом рассматриваемом случае может добиться требуемого состояния АСС.

5. Предложен подход к моделированию и анализу информационного влияния в АСС, который позволяет учесть совершаемые действия в сети и интересы управляющего субъекта. Поставлена и решена задача управления АСС для этого случая. Показано, что разработанные методы расчета влияния и влиятельности агентов АСС обладают рядом хороших свойств, которые позволяют эффективно рассчитывать и сравнивать влияние агентов больших АСС.

6. Поставлены, исследованы и решены задачи информационного противоборства в АСС для различных ситуаций противоборства, различающихся информированностью управляющих

субъектов, структурой их целевых функций и порядком их функционирования. Получены условия согласованности интересов управляющих субъектов (информационной кооперации). Показано, что для ситуаций информационного противоборства, сводимых к биматричным играм, стратегическая рефлексия управляющих субъектов приводит к уменьшению игровой неопределенности и в ряде случаев к увеличению выигрышей игроков.

7. Предложены прикладные методы и алгоритмы анализа АСС на основе моделей информационного влияния. Предложенные методы позволяют анализировать информационный ландшафт АСС на основе предпочтений лица, принимающего решения: выявлять наиболее влиятельные элементы АСС, оценивать подверженность элементов АСС информационным воздействиям, находить каналы влияния в АСС и скрытые информационные сообщества.

Степень достоверности полученных результатов подтверждена анализом отечественных и зарубежных работ по предмету исследования; анализом, обоснованием и апробацией предлагаемых моделей, методов и программных решений; результатами формализации и решения прикладных задач анализа АСС на основе информационного влияния; строгостью применяемого математического аппарата, а также результатами математического и компьютерного моделирования.

Апробация работы. Результаты диссертационной работы докладывались на семинарах Института проблем управления им. В.А. Трапезникова РАН, МГУ, СПбГУ, МФТИ, ФИАН и др.; международной конференции Management of Large-Scale System Development (Москва 2018, 2019, 2020), международной конференции International Conference on Computer Simulation in Physics and beyond (Москва 2020), международной конференции IEEE International Conference on Application of Information and Communication Technologies (Москва 2017), всемирном Конгрессе IFAC (Сеул 2008, Милан 2011), международной конференции UKACC International Conference on Control (Ковентри 2010), международной конференции X International Meeting of the Society for Social Choice and Welfare (Москва 2010), международной конференции European Conference on Operational Research (Рим 2013), международной научно-практической конференции «Теория активных систем» (Москва 2009, 2011, 2014, 2016, 2019), всероссийском совещании по проблемам управления (Москва 2014, 2019), всероссийской мультиконференции по проблемам управления (Москва 2009; Дивноморское 2013, 2015, 2017; Санкт-Петербург 2012, 2020), всероссийской школе-конференции молодых ученых «Управление большими системами» (Ижевск 2009; Пермь 2010; Самара 2016), всероссийской междисциплинарной конференции «Социофизика и инженерия» (Москва 2018), всероссийской научной конференции МФТИ (Долгопрудный 2008, 2010, 2016, 2017) и ряде других конференций.

Публикации. По теме диссертационной работы опубликовано 96 научных работ, в том числе две монографии и 34 статьи в изданиях из списка ВАК.

Личный вклад. Все основные результаты получены автором самостоятельно.

Структура и объем работы. Диссертационная работа состоит из введения, пяти глав, заключения и списка литературы. Работа содержит 307 страниц текста, список литературы включает 420 наименований.

Содержание работы

Во **введении** обоснована актуальность темы диссертационной работы, определены цель и задачи исследования, охарактеризованы используемые методы, описана структура работы и краткое содержание ее разделов.

Первая глава посвящена обсуждению общей проблематики информационного влияния и управления в АСС. В разделе 1.1 дано понятие активной сетевой структуры, определены факторы, влияющие на поведение агента в АСС и кратко обозначены задачи информационного влияния, управления и противоборства.

В разделе 1.2 рассмотрен процесс принятия решений агентом в активной сетевой структуре, испытывающим информационное влияние со стороны своих соседей. Информационное влияние приводит к изменению внутренних представлений агента и, как следствие, может подтолкнуть его к изменению поведения.

Согласно общей модели принятия решений агентом АСС, i-й агент способен выбирать *действие* y_i из множества допустимых действий A_i. В отсутствие внешних воздействий выбор действия определяется его типом $x_i \in X_i$ – *мнением* по некоторому вопросу. В АСС его выбор также определяется информированностью о действиях других агентов в сети. В результате выбора действия y_i агент получает выигрыш $f_i(x_i, y_i, y_{-i})$, где $y_{-i} \in A_{-i} = \prod_{j \neq i} A_j$ – вектор действий других агентов АСС, а f_i: $X_i \times A_i \times A_{-i} \rightarrow \Re^1$ – *целевая функция*, отражающая предпочтения агента. Принимается *гипотеза рационального поведения*, заключающаяся в том, что агент с учетом всей имеющейся у него информации выбирает действия, которые наиболее предпочтительны с точки зрения значений своей целевой функции. В соответствии с гипотезой агент выбирает альтернативу из множества альтернатив, на которых достигается максимум целевой функции: $\mathrm{Arg}\max_{y_i} f_i(x_i, y_i, y_{-i})$ при фиксированных x_i и y_{-i}. Далее модель усложняется: тип агента меняется под влиянием наблюдаемых им действий других агентов, в особенности тех, которым он доверяет (и которые, тем самым, оказывают на него информационное влияние).

Существенными здесь являются следующие компоненты (параметры) агента, которые влияют на тип/состояние агента.

- *Доверие агента к другому агенту* $a_{ij} \in [0, 1]$ определяет его подверженность влиянию действий другого агента в целом.
- *Доверие агента к информации в действии другого агента* $G_i(x_i, y_j)$, G_i: $X_i \times A_j \rightarrow [0,1]$. Из социальной психологии следует, что доверие к информации зависит не только от ее источника, но и от содержания (Майерс Д., 2002).
- *Репутация агента* $r_i \in [0,1]$ определяется как «общественная оценка, общее мнение о качествах, достоинствах и недостатках кого-чего-нибудь» (Ожегов С. И., 2008). Естественной представляется связь репутации агента и доверия к нему.

Взаимосвязь между этими компонентами показана на рис. 1, где компонент «принятие решений» включает в себя предпочтения агента и его цель. В такую схему можно уложить большое число имеющихся моделей поведения агентов, основывающихся на информационном влиянии (например, модели распространения активности и формирования мнений). Различные компоненты такой концептуальной модели взаимовлияния мнений и действий формализованы и исследованы в диссертации.

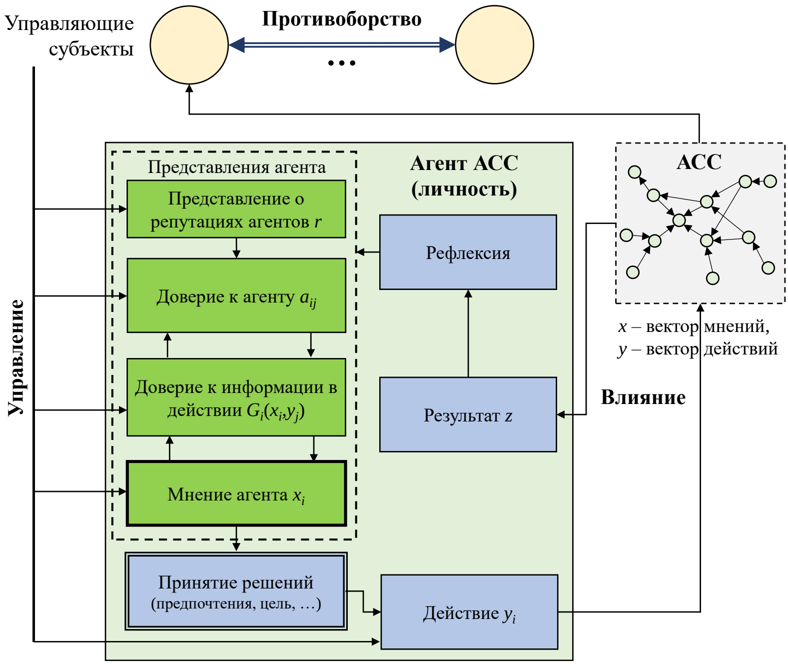

Рисунок 1 – Структура процесса информационного влияния, управления и противоборства

Поведение каждого агента по отношению к своему действию предполагается активным, т. е. агент самостоятельно и целенаправленно (стремясь максимизировать целевую функцию) выбирает свои действия. В зависимости от того, что агент знает на момент принятия решений, какие факторы и взаимосвязи принимает во внимание, возможны различные варианты формализации механизмов *принятия решений агентами АСС*, которые исследуются в диссертации.

- Первый случай – это «*простой*» агент (локально-оптимальный, см. Новиков Д.А., 2020), который меняет свой тип на основании действий своих соседей в сети и действие которого полностью отражает его тип. Рассматриваются доверчивые и осторожные виды простых агентов.
- Второй случай – агент-«*манипулятор*», который стремится достичь своей цели, выбирая для ее достижения оптимальные действия и не меняя при этом свой тип. Действия манипулятора в общем случае не отражают его тип и направлены на изменение типов других

агентов. В случае отсутствия общего знания относительно параметров АСС такой агент может рефлексировать и моделировать информированность и поведение других агентов в рамках своих представлений об их информированности и глубине рефлексии.

В разделе 1.3 в рамках введенной иерархии моделей АСС для описанной модели информационного взаимодействия агентов АСС рассматриваются постановки задач информационного *управления и противоборства*. В общем случае предметом информационного управления могут быть различные компоненты агента (см. рис. 1). В частности, зная механизм принятия решений агентом того или иного вида в АСС, управляющий субъект оказывает воздействие на начальные мнения агентов $u \in U = \prod_{i \in N} U_i$ с целью формирования требуемого мнения X. Тогда *критерием эффективности управления* будет $\Phi(X, u) \to \max_{u \in U}$. Далее приводится постановка задачи информационного противоборства субъектов, воздействующих на АСС каждый в своих интересах, приводится описание этапов исследования игр на сетях и устанавливается, что в зависимости от рассматриваемой ситуации (информированности игроков, порядка их функционирования и т. д.) задача информационного противоборства может быть сведена к той или иной теоретико-игровой модели.

Приведенная в разделах 1.4 и 1.5 классификация моделей поведения агентов в АСС, основывающихся на информационном влиянии, показывает нехватку работ, которые бы систематически рассматривали различные компоненты поведения агентов в АСС и были взаимно сопряжены в рамках иерархии «*информационное влияние → информационное управление → информационное противоборство*». Таким образом, проведенный в первой главе анализ обосновывает актуальность комплекса задач диссертационного исследования.

На основе предложенного подхода во **второй главе** разработано семейство математических моделей динамики мнений агентов АСС. Раздел 2.1 посвящен базовой модели информационного влияния, в которой изучается динамика и формирование мнений доверчивых агентов под влиянием других членов АСС. Входящие в АСС агенты описываются множеством $N = \{1, 2, \ldots, n\}$. Влияние агентов друг на друга задается матрицей прямого влияния A размерности $n \times n$, где $a_{ij} \geq 0$ обозначает степень доверия i-го агента j-му агенту (влияния j-го на i-го). Считается, что выполнено условие нормировки:

$$\forall\ i \in N \sum_{j=1}^{n} a_{ij} = 1 \qquad (2.1)$$

Если i-й агент доверяет j-му, j-й доверяет k-му, то это означает, что k-й агент косвенно влияет на i-го. Для определения агентов, формирующих мнение в сети, вводятся понятия *сообщества* (множество агентов, которое не подвергается внешнему влиянию), *группы* (сообщество агентов, в котором каждый агент влияет или подвергается влиянию каждого другого агента группы прямо или косвенно) и *спутника* (агент, не оказывающий влияния ни на одну из групп).

Мнения агентов в АСС формируются следующим образом. У каждого агента в начальный момент времени имеется мнение по некоторому вопросу, мнение i-го агента отражает вещественное число x_i^0, $i \in N$. Мнение всех агентов сети отражает вектор мнений x^0 размерности n. Агенты в сети взаимодействуют, обмениваясь мнениями. Этот обмен приводит к тому, что мнение каждого агента меняется в соответствии с мнениями агентов, которым он доверяет:

$$x_i^{\tau} = \sum_j a_{ij} x_j^{\tau-1},\ i \in N, \qquad (2.2)$$

где индекс τ обозначает момент времени. Если взаимодействие агентов продолжается достаточно долго, то их мнения могут сойтись к итоговому мнению $X = \lim\limits_{\tau\to\infty} x^{\tau}$ или

$$X = A^{\infty}\, x^0, \qquad (2.3)$$

где $A^{\infty} = \lim\limits_{\tau\to\infty}(A)^{\tau}$ – матрица результирующего влияния. Предполагается, что в каждой группе существует хотя бы один агент $i \in N$, для которого $x_{ii}>0$. Тогда из теории матриц следует справедливость утверждений 2.1–2.4.

<u>Утверждение 2.1</u>. Существует матрица результирующих влияний – предел $A^{\infty} = \lim\limits_{\tau\to\infty}(A)^{\tau}$.

<u>Утверждение 2.2</u>. Мнения агентов стабилизируются, т. е. существует предел $X = \lim\limits_{\tau\to\infty} x^{\tau}$.

<u>Утверждение 2.3</u>. Результирующее влияние любого спутника на любого агента равно нулю. Это, в частности, означает, что начальные мнения спутников не оказывают влияния на итоговые мнения в сети.

<u>Утверждение 2.4</u>. В матрице результирующих влияний строки, соответствующие членам одной группы, совпадают. Это означает, что совпадают итоговые мнения агентов группы.

Следовательно, мнения спутников определяются мнением групп, а в группах мнения стабилизируются и равны.

На основе данной модели информационного влияния в следующих разделах систематически формулируются и решаются различные задачи информационного управления мнениями агентов в АСС. Задача информационного управления понимается как задача формирования требуемого мнения в АСС путем информационного воздействия на отдельных агентов. В разделе 2.2 управляющее воздействие состоит в изменении центром начальных мнений агентов x^0 путем «добавления» вектора управлений $u \in \mathfrak{R}^n$, т. е. мнение i-го агента изменяется с x_i на $x_i + u_i$, $i \in N$. Пусть $u_i \in U_i$, $i \in N$, и $U = \prod_{i\in N} U_i$, тогда итоговые мнения определяются следующим уравнением:

$$X = A^{\infty}(x^0 + u), \qquad (2.4)$$

или в покоординатном виде:

$$X_{ui} = \sum_{j\in N} A_{ij}^{\infty}(x_j^0 + u_j) = \sum_{j\in N} A_{ij}^{\infty}x_j^0 + \sum_{j\in N} A_{ij}^{\infty}u_j,\ i \in N. \qquad (2.5)$$

В силу (2.4) «стабильное» состояние сети линейно по управлению. Считается, что целевая функция центра $\Phi(X, u)$ – критерий эффективности управления – зависит от итоговых мнений агентов и вектора управлений. Тогда задача управления заключается в выборе допустимого вектора управлений, максимизирующего критерий $\Phi(A^{\infty}(x^0 + u), u) \to \max_{u\in U}$. В целевой функции центра согласно традициям теории управления организационными системами выделяются две аддитивные компоненты:

$$\Phi(X, u) = H(X) - c(u), \qquad (2.6)$$

где $H(\cdot)$ – выигрыш центра, зависящий от итоговых мнений агентов, $c(\cdot)$ – затраты на осуществление управляющих воздействий. Показано, что при линейных доходах и затратах задача информационного управления (2.6) сводится к задаче линейного программирования, в оптимальном решении которой информационные воздействия оказываются, в первую очередь, на наиболее влиятельных агентов.

В последующих расширениях базовой модели информационного влияния рассматриваются содержательно более богатые механизмы формирования мнений и принятий решений агентами АСС с учетом их внутренней структуры, и решаются задачи управления.

В разделе 2.3 рассматривается ситуация, когда помимо «простых» агентов в АСС входят управляющие субъекты (например, СМИ), действия которых также влияют на мнения членов АСС. Начальные мнения всех агентов одинаковы и равны $x^0 \in \mathfrak{R}^1$, а граф связей между ними связный и l-регулярный. Каждый агент с некоторой степенью $\alpha \in (0; 1]$ доверяет себе; с некоторой степенью $\beta \in [0; 1]$ ($\alpha + \beta \leq 1$) доверяет внешним источникам информации, а «остаток доверия» $(1 - \alpha - \beta)$ делит поровну между смежными агентами. СМИ сообщает всем агентам одинаковое мнение $u \in \Re^1$. Динамика мнений агентов описывается выражением:

$$x_i^k = \alpha\, x_i^{k-1} + \beta\, u + \frac{(1-\alpha-\beta)}{l} \sum_{j \in N_i} x_j^{k-1}, k = 1,2, \ldots \qquad (2.7)$$

Оказывается, что для любого момента времени мнения агентов не выходят за диапазон, ограниченный их начальным мнением x^0 и управлением u. Предел последовательности мнений агентов при $k \to +\infty$ равен значению управления u. В рассматриваемой модели управление является постоянным и унифицированным. Задача управления заключается в нахождении управления $u(x^*, x^0, T)$, которое при известных начальных мнениях агентов в заданный момент времени T приводит агентов к требуемому мнению x^*. Если ограничения на управление отсутствуют, то решением будет:

$$u(x^*,\ x^0,\ T) = \frac{x^* - x^0(1-\beta)^T}{1-(1-\beta)^T}. \qquad (2.8)$$

При $T \to +\infty$ управление (2.8) стремится к итоговому мнению x^*.

Возможны отличные от выражения (2.7) законы изменения мнений агентов под влиянием друг друга и СМИ. В частности, в разделе 2.3 приведен вариант закона изменения мнения агента, в котором СМИ якобы отражает мнение той части АСС, которая не взаимодействует с данным агентом. Представлено решение задачи управления и приведены иллюстрирующие примеры.

В рассмотренной модели степень доверия агента сообщениям СМИ является постоянной и не зависит от того, насколько сообщения СМИ совпадают с мнением агента или ему противоречат. Условно такой случай соответствует *«доверчивым» агентам*. В разделе 2.3 рассмотрен вариант *«осторожных» агентов*, доверие которых к сообщениям СМИ зависит от содержания этих сообщений (см. рис. 1). Для того чтобы отразить зависимость степени доверия агента действиям СМИ от их содержания, вводится описывающая эту зависимость *функция доверия* $G(x, u)$, где x – мнение агента, u – управление (сообщение). Относительно свойств функции доверия вводятся различные предположения (о монотонности, о точках максимума и т. д.), комбинации которых приводят к различным случаям формализации функции доверия. В частности, можно условно выделить пять случаев, имеющих следующие содержательные интерпретации.

Случай 1 соответствует агенту, который независимо от содержания реагирует на сообщение СМИ (функция доверия – константа).

Случай 2 соответствует агенту-консерватору, у которого вероятность выделить сообщение будет уменьшаться с возрастанием отклонения его мнения от сообщения:

$$G(x, u) = \beta \exp(-\gamma |x-u|),\ \gamma > 0 \tag{2.9}$$

Случай 3 соответствует агенту-новатору, у которого вероятность выделить сообщение будет возрастать с ростом отклонения его мнения от мнения СМИ:

$$G(x, u) = 1-(1-\beta) \exp(-\gamma |x-u|),\ \gamma > 0. \tag{2.10}$$

Случай 4 соответствует агенту-умеренному консерватору, который выделяет и воспринимает информацию СМИ, совпадающую с его мнением, до тех пор, пока различие во мнениях не станет достаточно велико. Но при очень больших отклонениях вероятность того, что агент заметит такую информацию, растет.

$$G(x, u) = \beta\,[1-(1-\exp(-\gamma |x-u|)) \exp(-\gamma |x-u|)]. \tag{2.11}$$

Случай 5 соответствует агенту-умеренному новатору, у которого, пока отличие его мнения от мнения СМИ не слишком велико, вероятность выделить сообщение СМИ только возрастает, но при достаточно больших отклонениях начинает уменьшаться.

$$G(x, u) = (1-\beta) \exp(-\gamma |x-u|)) \exp(-\gamma |x-u|) + \beta \tag{2.12}$$

Для таких агентов исследуется случай переменного управления в АСС. Пусть $u^{0,T-1} = (u^0, u^1, \ldots, u^{T-1}) \in \Re^T$ – последовательность управлений, $x^{0,T} = (x^0, x^1, \ldots, x^T) \in \Re^{T+1}$ – траектория состояний АСС, $T \geq 0$, $F(x^{0,T}, u^{0,T-1})$ – критерий эффективности управления, где $F(\cdot, \cdot)$: $\Re^{(T+1)\ T} \to \Re^1$ – заданная функция. По аналогии с (2.7) управляемая динамика состояний АСС описывается выражением:

$$x^k = G(x^{k-1}, u^{k-1})\, u^{k-1} + (1 - G(x^{k-1}, u^{k-1}))\, x^{k-1},\ k = 1, 2, \ldots . \tag{2.13}$$

В общем виде *задача синтеза оптимального информационного управления* в однородной АСС формулируется как задача поиска такой последовательности управлений динамической системой (2.13), которая максимизирует критерий эффективности:

$$F(x^{0,T}, u^{0,T-1}) \to \max_{u^{0,T-1} \in \Re^T} \tag{2.14}$$

Задача оптимального управления (2.14) может быть решена известными методами, например, при аддитивном по периодам времени критерии эффективности – применением принципа оптимальности Беллмана. Если управление постоянно, то выражение (2.13) примет вид

$$x^k = G(x^{k-1}, u)\, u + (1 - G(x^{k-1}, u))\, x^{k-1},\ k = 1, 2, \dots , \tag{2.15}$$

а задача (2.15) может быть записана как

$$F_0(x^{0,T}, u) \to \max_{u \in \Re^1}, \tag{2.16}$$

т. е. является задачей безусловной скалярной оптимизации, где $F_0(\cdot,\ \cdot)$: $\Re^{T+1} \to \Re^1$ – заданный критерий эффективности в задаче с постоянным управлением.

Важным частным случаем задачи (2.14) является следующая постановка: пусть фиксирован вектор x^* («цель» управления) и заданы затраты $C(u^{0,T-1})$: $\Re^T \to \Re^1$ на управление, а также ограничение $R \geq 0$ на них. Тогда задача (2.14) может быть записана в виде:

$$\begin{cases} ||x^T - x^*|| \ \to \ \min\limits_{u^{1,T}}, \\ \quad C(u^{0,T-1}) \leq R. \end{cases} \tag{2.17}$$

В работе приведены примеры решения задач синтеза оптимального информационного управления, иллюстрирующие зависимость оптимального решения от свойств функции доверия.

В разделе 2.4 вводится понятие *репутации* и исследуется ее роль в осуществлении информационных воздействий, ставятся и решаются задачи управления репутаций. Пусть $r_i \geq 0$ – параметр, описывающий репутацию i-го агента. Вектор репутаций $r = (r_1, r_2, \dots, r_n)$ считается общим знанием среди агентов. Считается также, что сеть представляет собой полный граф, в котором существует агент с ненулевой репутацией, следовательно, в силу результатов раздела 2.1, итоговое мнение является единым для всех агентов АСС. Степень доверия i-го агента j-му агенту определяется как

$$\alpha_{ij} = \frac{r_j}{\sum_{k \in N} r_k}, i, j \in N, \tag{2.18}$$

Динамика мнений агентов записывается в виде

$$x_i^\tau = \frac{1}{R} \sum_{j \in N} r_j x_j^{\tau-1},\ i \in N, \tag{2.19}$$

а итоговое мнение агентов

$$X = \frac{1}{R}\,(r \cdot x^0), \tag{2.20}$$

где $R = \sum_{k \in N} r_k$ – «*коллективная*» *репутация* членов сети.

Выражения (2.18) – (2.20) представляют собой достаточно простую модель влияния в АСС с учетом репутации агентов.

Дальше рассматриваются две постановки задачи управления, в которых центр может достичь нужных мнений, воздействуя *на репутацию* агентов. В первой предполагается, что существует множество $M \subseteq N$ агентов влияния, на репутацию которых может оказывать влияние центр. Начальные мнения всех агентов, а также репутации всех агентов, кроме агентов влияния, считаются фиксированными. Известными считаются затраты центра $c_j(r_j)$ на создание репутации r_j j-го агента влияния, $j \in M$ ($|M| = m$), и зависимость выигрыша центра $H(X)$ от итогового мнения агентов X. Вводятся обозначения $r_M = (r_j)_{j \in M}$ – вектор репутаций агентов влияния, $C_0(r_M) = \sum_{j \in M} c_j(r_j)$ – суммарные затраты центра. Итоговое мнение членов АСС зависит от их начальных мнений и репутаций следующим образом:

$$X(r_M) = \frac{1}{\sum_{i \in N \setminus M} r_i + \sum_{j \in M} r_j} \left[\sum_{i \in N \setminus M} r_i x_i^0 + \sum_{j \in M} r_j x_j^0\right].$$

Целевая функция центра $\Phi(r_M) = H(X(r_M)) - C_0(r_M)$ представляет собой разность между выигрышем и затратами. *Задача управления репутацией* сводится к стандартной оптимизационной задаче:

$$H\left(\frac{\sum_{i \in N \setminus M} r_i x_i^0 + \sum_{j \in M} r_j x_j^0}{\sum_{i \in N \setminus M} r_i + \sum_{j \in M} r_j}\right) - \sum_{j \in M} c_j(r_j) \to \max_{r_M \geq 0}.$$

Более сложной является ситуация, в которой центр является интеллектуальным участником сети (*манипулятором*), его сообщения влияют на его репутацию и тем самым могут помочь в достижении его долгосрочных целей. Сначала рассматривается модель *информационного управления*, в ней некоторый агент (без потери общности первый, с $r_1 > 0$) заинтересован в итоговом мнении агентов X_*. При заданном векторе репутаций и фиксированных мнениях остальных агентов для этого, в силу (2.20), ему достаточно сообщить

$$s_1 = \frac{1}{r_1}\left[R X_* - \sum_{k>1} r_k x_k^0\right]. \qquad (2.21)$$

В разделе 2.4 показано, что (а) чем выше репутация агента, осуществляющего манипулирование, тем больше его возможности по влиянию на итоговое мнение агентов в АСС; (б) чем выше репутации других агентов, тем жестче требования к репутации «манипулятора».

В реальных АСС репутация агентов зависит от предыстории их взаимодействия. Возникает задача описания динамики репутации и процессов целенаправленного ее формирования. Предполагается, что члены АСС последовательно рассматривают и обсуждают T вопросов, по каждому из которых у каждого из агентов имеется свое начальное мнение x_i^τ, $i \in N$, $\tau = \overline{1, T}$. Общим знанием среди агентов являются репутации, начальные и результирующие мнения всех агентов для текущего и всех прошлых периодов. Пусть R^τ – суммарная репутация агентов в начале периода τ, X^τ – результирующее мнение агентов к концу периода τ. Обсуждаемые вопросы являются независимыми, и результирующие мнения определяются так

$$X^\tau = \frac{1}{R^\tau}(r^\tau \cdot x^\tau), \qquad (2.22)$$

где $r^\tau = (r_1^\tau, \ldots, r_n^\tau)$, $x^\tau = (x_1^\tau, \ldots, x_n^\tau)$ – соответственно, вектора репутаций и начальных мнений агентов в начале периода времени τ, $\tau = \overline{1, T}$.

Для описания всей траектории изменения мнений и репутаций агентов необходимо доопределить, как изменяется репутация каждого из агентов в каждом периоде времени. Считается, что репутация любого агента в начале любого периода равна его репутации в конце предыдущего периода, поскольку обсуждаемые вопросы принадлежат примерно одной тематике. В общем случае репутация i-го агента в момент времени τ определяется начальными и результирующими мнениями всех агентов и их репутациями во всех предшествующих периодах:

$$r_i^\tau = F_i(r^1, \ldots, r^{\tau-1}, x^1, \ldots, x^{\tau-1}, X^1, \ldots, X^{\tau-1}),\ i \in N,\ \tau = \overline{2, T}, \qquad (2.23)$$

причем предполагается, что функция $F_i(\cdot)$ монотонно убывает по разности $|x_i^{\tau-1} - X^{\tau-1}|$ и возрастает по предыдущим значениям репутации данного агента. В качестве частного предлагается следующий закон:

$$r_i^\tau = \frac{r_i^{\tau-1}}{\gamma + \beta|x_i^{\tau-1} - X^{\tau-1}|},\ i \in N,\ \tau = \overline{2, T}, \qquad (2.24)$$

где $\gamma \in (0; 1]$, $\beta > 0$ – заданные константы. В соответствии с (2.24) репутация агента в начале некоторого периода времени зависит только от его репутации в предыдущем периоде, а также того, насколько его начальное мнение в предыдущем периоде оказалось отличным от результирующего мнения всех агентов к концу этого периода.

Далее выполняется переход от описания информационного влияния и динамики репутации (уравнения (2.22) и (2.23)) к постановке и решению для рассматриваемой модели задачи управления. Рассматривается случай манипулирования со стороны первого агента, который *манипулирует* своими начальными мнениями по каждому из вопросов для достижения необходимого результирующего мнения в сети по последнему вопросу. Требуется найти последовательность сообщаемых другим агентам начальных мнений первого агента $s_1^1, s_1^2, \ldots, s_1^T$, удовлетворяющую ограничениям $s_1^\tau \geq x_1^{\tau\min}$, $\tau = \overline{1, T}$, и минимизирующую заданную монотонную целевую функцию $F(|X^T - X_*^T|)$, где формирование итогового мнения X_*^T по последнему вопросу интерпретируется как цель управления. В общем случае сформулированная задача является задачей динамического программирования и в каждом конкретном случае решается численно. Вводится эвристика поведения первого агента, когда первый агент должен решить задачу, состоящую из $T - 1$ независимой задачи максимизации репутации и одной задачи выбора своего начального мнения на последнем шаге, представлено решение этих задач. Оказывается, для повышения своей репутации агенту всегда следует высказывать «средневзвешенное» (с учетом репутаций) мнение остального коллектива. Далее приводятся примеры взаимодействия таких манипулирующих агентов.

В разделе 2.4 рассматриваются ситуации, в которых представления агентов о мнениях и/или репутациях друг друга могут быть неполными и/или различающимися. В случае стратегической рефлексии агенты размышляют о том, какие действия выберут оппоненты. В случае информационной рефлексии агенты имеют асимметричную информированность о репутации друг друга, например, представления i-го агента о репутации j-го, представления i-го агента о представлениях j-го агента о репутации k-го агента и т. д. К таким конструкциям применим аппарат теории рефлексивных игр, с помощью которого можно искать *рефлексивные* и *информационные равновесия*, исследовать их стабильность и т. д. В разделе 2.4 приведены соответствующие примеры информационной и стратегической рефлексии агентов в АСС.

В <u>разделах 2.5 и 2.6</u> формулируется и решается задача информационного управления *доверием* членов АСС. В первой постановке управляющие воздействия оказываются на доверие между агентами (раздел 2.5). Управляя взаимным доверием агентов, центр может добиться нужных ему мнений в сети. В отсутствии управления состояние АСС в момент $t \geq 0$ задается соотношением

$$x^t = (A)^t x, \tag{2.25}$$

где $x = x^0$ – начальное состояние сети, A – матрица прямого влияния размерности $n \times n$. Центр осуществляет *управление доверием* путем аддитивного изменения матрицы A – увеличения ее на матрицу управлений $V = \|v_{ij}\|$. Эта матрица принадлежит множеству возможных управлений $\bar{V}$, которое отражает возможности центра по оказанию воздействия на те или иные связи между агентами, а также общие ресурсные ограничения.

Сначала исследуется случай оказания воздействия центром в начальный момент времени. С учетом этого воздействия формула (2.25) приобретает следующий вид:

$$x^t = (A + V)^t x. \tag{2.26}$$

Поскольку добавление V должно сохранять свойство стохастичности матрицы влияния, то на ее выбор накладываются дополнительные ограничения:

$$\begin{cases} \forall i \in N \quad \sum_{j \in N} v_{ij} = 0, \\ \forall i, j \in N \quad -a_{ij} \le v_{ij} \le 1 - a_{ij}. \end{cases} \tag{2.27}$$

Множество матриц, удовлетворяющих (2.27), обозначается $\hat{V}$.

Целевая функция центра $\Phi(x^t, V)$ – критерий эффективности управления – зависит от мнений агентов в момент t и матрицы управлений. Тогда задача управления заключается в выборе допустимой матрицы управлений, максимизирующей критерий:

$$\Phi(x^t, V) \to \max_{V \in \bar{V} \cap \hat{V}}.$$

В общем случае воздействия на взаимное влияние агентов могут оказываться в разные моменты времени, для каждого момента могут быть свои ограничения. Множество возможных управлений в момент τ обозначается через $\bar{V}^\tau$, а матрица управлений – через V^τ. Тогда матрица влияния в момент t рассчитывается по следующей формуле:

$$A^t = A + \sum_{\tau=0}^{t-1} V^\tau, \tag{2.28}$$

а рекуррентная формула вычисления состояния сети:

$$x^{t+1} = (A^t + V^t) x^t. \tag{2.29}$$

При этом допустимыми на горизонте планирования T являются лишь такие управления V^τ, $\tau = 0, \ldots, T-1$, для которых все матрицы A^t, $t = 1, \ldots, T-1$, являются стохастическими по строкам:

$$\begin{cases} \forall i \in N \ \forall t \in \{0, \ldots, T-1\} \quad \sum_{j \in N} v_{ij}^t = 0, \\ \forall i, j \in N \ \forall t \in \{0, \ldots, T-1\} \quad -a_{ij} \le \sum_{\tau=0}^{t} v_{ij}^t \le 1 - a_{ij}. \end{cases} \tag{2.30}$$

Пусть $\hat{V}(T)$ – множество конечных последовательностей матриц $(V^0, \ldots, V^{T-1})$ размерности $n \times n$, удовлетворяющих условиям (2.30). Соотношение (2.26) приобретает вид $x^T = \left(\prod_{t=0}^{T-1} A^{T-t}\right) x$.

Целевая функция центра зависит от итоговых мнений агентов в момент T и матриц управления в моменты времени 0, …, T – 1. Задача управления заключается в выборе допустимой последовательности матриц управления, которая максимизирует критерий эффективности:

$$\Phi(x^T,\, V^0, \dots, V^{T-1}) \to \max_{\substack{V^0 \in \bar{V}^0,\, \dots,\, V^{T-1} \in \bar{V}^{T-1} \\ (V^0,\, \dots,\, V^{T-1}) \in \hat{V}(T)}} . \qquad (2.31)$$

В общем случае задача управления (2.31) является довольно сложной. При этом отмечается следующая особенность управления доверием: если в начальный момент мнения агентов находятся в некотором промежутке, то они и дальше будут находиться в нем при любом управлении.

Утверждение 2.5. При осуществлении управления доверием для любого t = 0, 1, … и любого $i \in N$ справедливо соотношение

$$x_{\min} \le x_i^t \le x_{\max}, \qquad (2.32)$$

где $x_{\min} = \min\{x_1^0, \dots, x_n^0\}$, $x_{\max} = \max\{x_1^0, \dots, x_n^0\}$.

Следствие 2.5.1. Если мнения агентов в начальный момент совпадают, то они не меняются со временем при осуществлении центром управления доверием.

Утверждение 2.5 накладывает серьезные ограничения на возможности центра достигать своих целей посредством управления доверием: мнения агентов ни при каком управлении не могут выйти за пределы $[x_{\min}, x_{\max}]$. Но, наряду с этим, справедливо утверждение 2.6.

Утверждение 2.6. Если не накладывать ограничения на возможные управляющие воздействия центра, то управление доверием позволяет за один шаг сделать мнением каждого агента любое наперед заданное значение $x^* \in [x_{\min}, x_{\max}]$.

В разделе 2.6 формулируется и решается задача *децентрализованного формирования доверия* (весов связей) каждого агента АСС для достижения требуемого согласия в сети, задаваемого как взвешенная комбинация начальных мнений агентов сети. Показывается, что для расчета искомых весов структуры взаимодействий агентам достаточно лишь локальной информированности. В иной интерпретации центр может, скрытно изменяя веса связей агентов, добиться нужного консенсуса в сети.

Под сетевой структурой понимается неориентированный связный граф $G = (N, E)$, в котором $N = \{1, \dots, n\}$ – конечное множество вершин (агентов), $n \geq 2$, а E – совокупность пар вершин (связей между агентами). Множество E задается бинарной матрицей (e_{ij}) размерности $n \times n$, отражающей наличие связи между агентами и имеющей единичные значения на диагонали. Число соседей i-го узла определяется как $v_i = \left|\{j \mid j \neq i, e_{ij} = 1\}\right|$.

Мнение i-го агента о некоем вопросе (параметре ситуации) в начальный момент времени отражает вещественное число θ_i. Агенты в сетевой структуре обмениваются с соседями мнениями о параметре θ и меняют их в соответствии с уравнением (2.2), где $x_i^0 = \theta_i$ – мнение i-го агента в начальный момент времени. Поскольку агенты могут обмениваться информацией лишь при наличии между ними связи, считается, что $a_{ij} > 0$ может выполняться лишь при $e_{ij} = 1$.

Согласно модели центр заинтересован в определенном значении итогового представления агентов. Каждый i-й узел сети обладает значимостью $w_i > 0$. Задача центра формулируется как задача поиска такой матрицы информационного взаимодействия (доверия), для которой при любых начальных мнениях узлов процесс информационного обмена сходится к взвешенной сумме начальных мнений (*агрегированной характеристике сети*):

$$\overline{\theta} = \frac{\sum_{i \in N} w_i \theta_i}{\sum_{i \in N} w_i}. \tag{2.33}$$

Говоря более строго, требуется для заданных сетевой структуры G и набора весов $w_i > 0$ найти такую матрицу доверия $A = (a_{ij})$, что

$$a_{ij} > 0 \Rightarrow e_{ij} = 1 \text{ для любых } i, j \in N, \tag{2.34}$$

$$x_i^{(t)} \xrightarrow[t \to \infty]{} \overline{\theta} \text{ для любого } \theta \text{ и любого } i \in N, \tag{2.35}$$

где $\overline{\theta}$ задается соотношением (2.33). Показано, что можно получить конструктивное доказательство существования матрицы доверия, позволяющей центру достичь требуемого консенсуса агентов.

<u>Утверждение 2.7</u>. Для любой связной сетевой структуры G и любого набора весов $w_i > 0$ существует такая стохастическая матрица $A = (a_{ij})$, удовлетворяющая условию (2.34), что процесс информационного обмена (2.2) удовлетворяет условию (2.35).

Доказывается, что такая матрица $A = (a_{ij})$ может быть построена следующим образом:

$$b_{ij} = e_{ij} \min\left\{\frac{1}{1 + w_i v_j}, \frac{1}{1 + w_j v_i}\right\}, i \neq j, \tag{2.36}$$

$$b_{ii} = 1 - \sum_{j \in N \setminus \{i\}} b_{ij} w_{j,} \tag{2.37}$$

$$a_{ij} = \begin{cases} b_{ij}, & i = j, \\ b_{ij} w_j, & i \neq j. \end{cases} \quad (2.38)$$

Процесс информационного обмена, определяемый матрицей (2.36) – (2.38), приводит агентов к консенсусу относительно значения неопределенного параметра. Этот процесс подразумевает, что состав агентов и связи между ними остаются неизменными. Однако может оказаться, что состав связей между агентами и даже состав агентов меняется. Показано, что в этом случае перезапускать процесс информационного обмена нет необходимости: несмотря на изменения, предельное значение представлений агентов будет тем же самым, как если бы процесс обмена изначально был запущен на финальном графе (утверждение 2.9). Этот факт опирается на следующее свойство представлений агентов.

Утверждение 2.8. Пусть процесс информационного обмена $x^{(t)} = A^t x^{(0)}$, $t \geq 0$, начинается с вектора мнений $x^{(0)} = \theta = (\theta_1, \dots, \theta_n)$, и агенты достигают консенсуса $\overline{\theta} = \frac{\sum_{i \in N} w_i \theta_i}{\sum_{i \in N} w_i}$. Тогда, во-первых, для любого t справедливо следующее равенство:

$$\textstyle\sum_{i \in N} w_i x_i^t = \sum_{i \in N} w_i \theta_i, \quad (2.39)$$

т. е. взвешенная сумма мнений агентов не меняется, в частности

$$\textstyle\sum_{i \in N} w_i x_i^{(t)} = \sum_{i \in N} w_i x_i^{(t+1)}; \quad (2.40)$$

во-вторых, для любого начального $y = (y_1, \dots, y_n)$ такого, что

$$\textstyle\sum_{i \in N} w_i y_i = \sum_{i \in N} w_i \theta_i, \quad (2.41)$$

достигается такой же консенсус $\overline{\theta}$.

Множество агентов N с вектором мнений $x^0 = (\theta_1, \dots, \theta_n)$ в каждый момент $t = 0, 1, \dots, T-1$ разбивается на непересекающиеся подмножества $N = N_1(t) \cup \dots \cup N_{m(t)}(t)$. Агенты из $N_q(t), q = 1, \dots, m(t)$, являются вершинами неориентированного связного графа $G_q(t) = (N_q(t), E_q(t))$. Процесс информационного обмена в момент t осуществляется при помощи $m(t)$ матриц $A_q(t)$. Если в подмножестве $N_q(t)$ ровно один агент, то он ни с кем не взаимодействует, и матрица вырождается в число 1. Если в подмножестве $N_q(t)$ не менее двух агентов, то матрица $A_q(t)$ строится на основе соответствующего графа $G_q(t)$ при помощи (2.36)–(2.38). По сути, этот процесс декомпозируется на несколько процессов, в каждый момент протекающих в непересекающихся связных графах. Однако начиная с некоторого момента T все агенты объединены в связный граф $G_1(t)$, который далее не меняется (и, соответственно, не меняется $A_1(t)$). Таким образом, $m(t) = 1$ при $t \geq T$. Описанный процесс называется *процессом информационного обмена в переменных графах*.

Утверждение 2.9. Описанный выше процесс информационного обмена в переменных графах

$$G_q(t),\ q = 1, \dots, m(t),\ t = 0, 1, \dots, T–1,$$
$$G_1(t) = G_1(T),\ m(t) = 1 \text{ при } t \geq T,$$

приводит к тому же консенсусу, что и процесс информационного обмена в постоянном связном графе $G_1(T)$.

В **третьей главе** рассматривается подход к моделированию и анализу информационного влияния в АСС на основе *действий* и интересов управляющего органа (*акциональный подход*, см. рис. 1), в рамках которого разрабатывается модель распространения действий и формализуются интересы управляющего органа при помощи функции значимости действий агентов АСС. Формализуются различные случаи влияния и влиятельности агентов и структур АСС, доказывается ряд утверждений, связанных с функциями влияния и влиятельности, а также приводятся задачи прогноза и управления. В разделе 3.1 рассматривается формальная модель распространения действий в АСС, в которой участниками сети являются агенты из множества $N = \{1, 2, \dots, n\}$, совершающие действия того или иного вида из фиксированного множества $K = \{1, 2, \dots, k\}$ в те или иные моменты времени из интервала T. Например, в социальной сети видом действия может быть создание поста. Конечное множество действий (например, создание конкретного поста) обозначается через Δ. Каждое действие $a \in \Delta$ характеризуется тремя параметрами – совершившим его агентом, видом действия и моментом времени совершения: $a(i, j, t)$, $i \in N$, $j \in K$, $t \in T$. Определяется функция $\alpha(a)$, которая каждому действию $a \in \Delta$ ставит в соответствие совершившего его агента $\alpha \in N$. На множестве действий задается бинарное отношение частичного порядка «a является причиной b», обозначаемое следующим образом: $a \to b$. Пример такого отношения в социальной сети: a – создание поста, b – создание комментария к посту. Считается, что это отношение удовлетворяет следующим свойствам.

1. Рефлексивность: для любого $a \in \Delta$ справедливо $a \to a$.
2. Транзитивность: если $a \to b$ и $b \to c$, то $a \to c$.
3. Антисимметричность: если $a \to b$ и $b \to a$, то a и b совпадают.

Если $a \to b$ и $a \neq b$, но при этом не существует такого $c \in \Delta$, что $a \to c$ и $c \to b$, то a называется *непосредственной причиной* b (или $a \downarrow b$). Это позволяет выделить класс бинарных отношений, называемых *однозначными*, в которых у каждого действия не более одной непосредственной причины. Если задано множество $A \subseteq \Delta$, то можно определить множество всех действий, являющихся последствиями действий из A: $\pi(A) = \{b \in \Delta \mid \exists\, a \in A\ a \to b\}$. Среди всех действий Δ можно выделить множество Δ^0 *начальных действий*: $\Delta^0 = \{a \in \Delta \mid \forall b \in \Delta\ (b \to a) \Rightarrow (a = b)\}$.

В разделе 3.2 вводится точка зрения центра для расчета значимости действий в АСС и оценки их влияния. Центр исходя из своих предпочтений определяет, какие именно действия агентов являются значимыми. Для формализации его точки зрения вводится в рассмотрение *значимость множества действий* – функция $\Phi(S)$, которая каждому множеству действий $S \subseteq \Delta$ ставит в соответствие вещественное число, $\Phi: 2^{\Delta} \to [0, +\infty)$. Предполагается, что хотя бы какие-то действия обладают положительной значимостью ($\Phi(\Delta)>0$), а функция значимости является монотонно возрастающей:

$$\text{если } A \subset B, \text{ то } \Phi(A) \leq \Phi(B). \qquad (3.1)$$

Для решения конкретных прикладных задач функция значимости должна быть корректно определена (включая выполнение свойства (3.1)). Немаловажным с практической точки зрения является также наличие эффективных алгоритмов расчета ее значения.

В разделах 3.3 и 3.4 предлагается конструктивный подход к определению влияния, основанный на модели распространения действий (раздел 3.1) и учете точки зрения центра (раздел 3.2). Вводится определение влияния мета-агента в АСС, представляющего собой непустое подмножество множества агентов N. Для каждого мета-агента $I \subseteq N$ определяется множество $\delta \subseteq \Delta$ всех совершенных им действий $\delta_I = \{a \in \Delta | \alpha(a) \in I\}$, а также множество совершенных им начальных действий $\delta_I^0 = \{a \in \Delta^0 | \alpha(a) \in I\}$. Неформальное понимание влияния формулируется так: влияние мета-агента $I \subseteq N$ на мета-агента $J \subseteq N$ велико, если деятельность агентов из множества J в достаточно большой степени обусловлена деятельностью агентов из множества I. Это понимание можно формализовать в зависимости от решаемой практической задачи. Рассматриваются вопросы, в зависимости от ответа на каждый из которых понятие влияния формализуется несколько различным образом.

Вопрос А. Оказывают ли влияние все действия (А1), либо только начальные (А2)?

Вопрос В. Интересует ли воздействие на действия мета-агента (В1) или воздействие на последствия действий мета-агента (В2)?

Вопрос С. Считается ли, что максимально возможное влияние на любого мета-агента принимает одно и то же значение, т. е. влияние является нормированной величиной (С1), либо нет (С2)?

Комбинация различных ответов на вопросы А – С приводит к разным случаям и, соответственно, разным вариантам формализации понятия влияния. Рассматривается каждый из случаев (при этом для функции влияния мета-агента *I* на мета-агента *J* используется одно и то же обозначение $\chi(I, J)$).

Случай 1. Сочетанию А1, В1, С1 отвечает функция влияния

$$\chi(I,J) = \begin{cases} \frac{\Phi(\pi(\delta_I)\cap\delta_J)}{\Phi(\delta_J)}, & \Phi(\delta_J) > 0; \\ 0, & \Phi(\delta_J) = 0. \end{cases}$$

Случай 2. Сочетанию А1, В1, С2 отвечает функция влияния

$$\chi(I,J) = \Phi(\pi(\delta_I) \cap \delta_J).$$

Случай 3. Сочетанию А1, В2, С1 отвечает функция влияния

$$\chi(I,J) = \begin{cases} \frac{\Phi(\pi(\delta_I)\cap\pi(\delta_J))}{\Phi(\pi(\delta_J))}, & \Phi(\pi(\delta_J)) > 0; \\ 0, & \Phi(\pi(\delta_J)) = 0. \end{cases}$$

Случай 4. Сочетанию А1, В2, С2 отвечает функция влияния

$$\chi(I,J) = \Phi(\pi(\delta_I) \cap \pi(\delta_J)).$$

Случай 5. Сочетанию А2, В1, С1 отвечает функция влияния

$$\chi(I,J) = \begin{cases} \frac{\Phi(\pi(\delta_I^0)\cap\delta_J)}{\Phi(\delta_J)}, & \Phi(\delta_J) > 0; \\ 0, & \Phi(\delta_J) = 0. \end{cases}$$

Случай 6. Сочетанию А2, В1, С2 отвечает функция влияния

$$\chi(I,J) = \Phi(\pi(\delta_I^0) \cap \delta_J).$$

Случай 7. Сочетанию А2, В2, С1 отвечает функция влияния

$$\chi(I,J) = \begin{cases} \frac{\Phi(\pi(\delta_I^0)\cap\pi(\delta_J))}{\Phi(\pi(\delta_J))}, & \Phi(\pi(\delta_J)) > 0; \\ 0, & \Phi(\pi(\delta_J)) = 0. \end{cases}$$

Случай 8. Сочетанию А2, В2, С2 отвечает функция влияния

$$\chi(I,J) = \Phi(\pi(\delta_I^0) \cap \pi(\delta_J)).$$

Отмечается важный частный случай, когда мета-агент J совпадает со всем множеством агентов ($J = N$) и функция влияния характеризует влияние мета-агента I на всю сеть, которое называется *влиятельностью* и обозначается через $\varepsilon(I) = \chi(I, N)$. В зависимости от поставленной практической задачи влияние и влиятельность вычисляются в соответствии с одним из описанных случаев.

В разделе 3.3 формулируются *свойства функции влияния*.

<u>Утверждение 3.1.</u> Функция влияния $\chi(I, J)$ является монотонной по первому аргументу, т.е. если $I_1 \subseteq I_2$, то для любого J выполняется неравенство $\chi(I_1, J) \leq \chi(I_2, J)$.

Приведенное утверждение означает, что чем «больше» мета-агент, тем больше его влияние, независимо от прочих обстоятельств. Следующие два утверждения относятся к аддитивности функции влияния, которая является важной с вычислительной точки зрения.

<u>Утверждение 3.2.</u> Если бинарное отношение является однозначным, а функция значимости – аддитивной, то в случаях 5–8 функция влияния является аддитивной по первому аргументу, т. е. для любых $I_1, I_2, J \subseteq N$, $I_1 \cap I_2 = \emptyset$, выполняется равенство

$$\chi(I_1 \cup I_2, J) = \chi(I_1, J) + \chi(I_2, J).$$

Влиятельность мета-агента также в данных случаях является аддитивной функцией: для любых непересекающихся множеств I_1, $I_2 \subseteq N$ выполняется равенство $\varepsilon(I_1 \cup I_2) = \varepsilon(I_1) + \varepsilon(I_2)$.

<u>Утверждение 3.3.</u> Если функция значимости является аддитивной, то в случаях 2 и 6 функция влиятельности является аддитивной по второму аргументу, т.е. для любых $I, J_1, J_2 \subseteq N$, $J_1 \cap J_2 = \emptyset$, выполняется равенство $\chi(I, J_1 \cup J_2) = \chi(I, J_1) + \chi(I, J_2)$.

<u>Утверждение 3.4.</u> Если бинарное отношение является однозначным, а функция значимости – аддитивной, то в случае 6 функция влиятельности является аддитивной по обоим аргументам, т. е. для любых $I_1, I_2, J_1, J_2 \subseteq N$, $I_1 \cap I_2 = J_1 \cap J_2 = \emptyset$, выполняется равенство

$$\chi(I_1 \cup I_2, J_1 \cup J_2) = \chi(I_1, J_1) + \chi(I_2, J_1) + \chi(I_1, J_2) + \chi(I_2, J_2).$$

<u>Утверждение 3.5.</u> Если множество последствий действий мета-агента I содержится в множестве последствий действий мета-агента J, то влиятельность мета-агента I не превосходит влиятельности мета-агента J: $\pi(\delta_I) \subseteq \pi(\delta_J) \Rightarrow \varepsilon(I) \leq \varepsilon(J)$ (случаи 1–4),

$$\pi(\delta_I^0) \subseteq \pi(\delta_J^0) \Rightarrow \varepsilon(I) \leq \varepsilon(J) \text{ (случаи 5–8).}$$

Утверждение 3.6. Если влияние i-го агента на любого из агентов не меньше влияния на него j-го агента, то в случае 3 влиятельность i-го агента не меньше влиятельности j-го агента:

$$\forall k \in N \ \chi(i,k) \geq \chi(j,k) \Rightarrow \varepsilon(i) \geq \varepsilon(j).$$

Если функция значимости является аддитивной, то утверждение является верным и для случаев 2 и 6.

Утверждение 3.7. Если влияние мета-агента I на мета-агента J больше влияния мета-агента J на мета-агента I, то в случае 3 влиятельность мета-агента I больше влиятельности мета-агента J:

$$\chi(I,J) > \chi(J,I) \Rightarrow \varepsilon(I) > \varepsilon(J).$$

Утверждение 3.8. Если в результате совершения новых действий в сети ($\Delta' \supseteq \Delta$) влияние мета-агента I на мета-агента J увеличится, то в случаях 2, 4, 6 и 8 увеличится и влиятельность мета-агента I:

$$\chi_{\Delta'}(I, J) > \chi_{\Delta}(I, J) \Longrightarrow \varepsilon_{\Delta'}(I) > \varepsilon_{\Delta}(I).$$

Раздел 3.4 посвящен определению влияния структур АСС, включающих в себя узлы и связи между ними. Связи между индивидами так же важны с точки зрения управления состоянием сети, поскольку являются каналами распространения информации, формирующей мнения людей и побуждающей их к действиям.

Умея рассчитывать влияние и влиятельность элементов активных сетевых структур согласно акциональной модели, можно ставить задачи прогноза и управления АСС (см. раздел 3.5). Акциональная модель позволяет рассчитывать влиятельность в различных разрезах: тематическом, целевых аудиторий и т. д. При прогнозировании интерес представляет в первую очередь временной разрез. Рассматриваются равные последовательные интервалы времени τ_1, τ_2, …, τ_m. Поскольку каждое действие в сети осуществляется в определенный момент времени, при вычислении влиятельности агентов можно учитывать только действия, совершенные в эти интервалы. Пусть $\delta_i(\tau) = \bigcup_{a\in\Delta,\ j\in K,\ t\in\tau} a(i,j,t)$ – действия, совершенные i-ым агентом в течение интервала τ, тогда влиятельность этих действий определяется так $\varepsilon_i(\tau) = \Phi(\pi(\delta_i(\tau)))$.

Прогноз влиятельности i-го агента в следующий интервал времени τ_{m+1} предлагается составлять стандартными методами анализа временных рядов на основе ряда значений влиятельности в предыдущие интервалы: $\varepsilon_i(\tau_1)$, $\varepsilon_i(\tau_2)$,…, $\varepsilon_i(\tau_m)$. Далее предполагается, что центр может выполнить прогноз влиятельности агентов в следующий интервал времени ($\tilde{\varepsilon}$). *Управляющее воздействие* заключается в том, чтобы зарезервировать под свои нужды частоту действий агентов (от 0 до 1): $u = (u_1, u_2, \dots, u_N)$.

Выигрыш центра зависит от влиятельности агентов и зарезервированных частот: $H(\tilde{\varepsilon}, u) = \sum_{i \in N} \tilde{\varepsilon}_i u_i$. Целевая функция центра записывается в традиционном для теории управления организационными системами виде: $\mathcal{F}(\tilde{\varepsilon}, u) = H(\tilde{\varepsilon}, u) - c(u)$, где $c(u)$ – затраты центра на осуществление управляющих воздействий, которые можно определить как $c(u) = \sum_{i \in N} \frac{u_i}{1-u_i}$, т.е. чем больше зарезервированная частота, тем больше затраты центра. Задача управления сводится к следующей оптимизационной задаче:

$$\mathcal{F}(\tilde{\varepsilon}, u) = \sum_{i \in N} \tilde{\varepsilon}_i u_i - \frac{u_i}{1-u_i} \to \max_{\{0 \le u_i < 1\}}.$$

Решение этой задачи следующее: чем влиятельнее агент, тем большая его частота резервируется центром.

В разделе 3.6 показана связь разработанных моделей и методов расчета влиятельности с известными мерами сетевой центральности.

В **четвертой главе** на основе разработанных моделей информационного влияния и управления исследуется информационное противоборство управляющих субъектов. В разделе 4.1 рассматривается общая теоретико-игровая модель информационного противоборства в АСС и исследуется ряд ее частных случаев. Задается множество *игроков*, имеющих возможность влиять на начальные мнения агентов АСС и заинтересованных в формировании определенных итоговых мнений. Игроки активны и имеют собственные интересы. Описывается возникающая между игроками игра. Для начала вводятся обозначения: $M = \{1, 2, \dots, m\}$ – множество игроков, $u_{ij} \in U_{ij} = [-r_{ij}; R_{ij}]$ – действие j-го игрока по изменению мнения i-го агента, $r_{ij}, R_{ij} \ge 0$, $\boldsymbol{u} = ||u_{ij}||$, $\boldsymbol{u}_j = (u_{1j}, u_{2j}, \dots, u_{nj}) \in U_j = \prod_{i \in N} U_{ij}$, $u_i = \sum_{j \in M} u_{ij}$, $u = (u_1, u_2, \dots, u_n)$ – вектор воздействий, $g_j(X)$: $\Re^n \to \Re^1$ – целевая функция j-го игрока, $i \in N, j \in M$. Считается, что воздействия игроков на мнения агентов аддитивны. Тогда итоговое мнение будет

$$X_i(\boldsymbol{u}) = \sum_{j \in N} A^{\infty}_{ij} \left[x^0_j + \sum_{k \in M} u_{jk}\right] = \sum_{j \in N} A^{\infty}_{ij} x^0_j + \sum_{j \in N} A^{\infty}_{ij} \sum_{k \in M} u_{jk},$$
$$i \in N. \quad (4.1)$$

Каждый из игроков в общем случае может влиять на начальные мнения всех агентов. Тогда $G_j(\boldsymbol{u}) = g_j(X_1(\boldsymbol{u}), X_2(\boldsymbol{u}), \dots, X_n(\boldsymbol{u}))$, $j \in M$. Игроки выбирают действия однократно, одновременно и независимо. Задается игра $\Gamma = (M, \{U_j\}_{j \in M}, \{G_j(\cdot)\}_{j \in M})$ в нормальной форме, определяемая множеством игроков, их множествами допустимых действий и целевыми функциями. Для игры в нормальной форме можно исследовать ее равновесия, определять «на ней» различные виды игр, что показано на соответствующих примерах.

Во многих реальных системах один и тот же агент оказывается подчинен одновременно нескольким управляющим органам – центрам. Такая ситуация называется *распределенным контролем* (Новиков Д.А., 1999). Центры, осуществляющие управление агентом, оказываются вовлеченными в «игру», равновесие которой имеет сложную структуру. В частности, выделяются два устойчивых режима взаимодействия центров – режим сотрудничества и режим конкуренции. В *режиме сотрудничества* центры действуют совместно, что позволяет добиваться требуемых результатов в сети с использованием минимального количества ресурсов. В *режиме конкуренции*, который возникает, если цели центров различаются достаточно сильно, ресурсы расходуются неэффективно. Раздел 4.2 посвящен исследованию распределенного контроля в АСС, когда управляющие субъекты могут иметь несовпадающие интересы. В нем исследуются условия согласованности интересов управляющих органов – когда они смогут договориться между собой, каковы должны быть формируемые мнения агентов.

Пусть $\{N_i\}_{i \in K}$ – совокупность подмножеств множества агентов N, где N_i – множество агентов, на которые может оказывать информационные воздействия i-й центр, $i \in K$; $K_j = \{k \in K \mid j \in N_k\}$ – множество центров, которые могут оказывать информационные воздействия на j-го агента, $j \in N$; $c_i(y^0, x)$ – затраты на изменение мнения i-го агента с y_i^0 на x_i, $i \in N$; $H_i(x)$ – предпочтения i-го центра на множестве мнений агентов, $i \in K$; $\sigma_{ij}(y^0, x)$ – затраты i-го центра на осуществление воздействий на j-го агента, $j \in N_i$, $i \in K$.

Целевая функция i-го центра имеет вид

$$\Phi_i(\{\sigma_{ij}(\cdot)\}_{j\in N_i}, y^0, x) = H_i(x) - \sum_{j\in N_i} \sigma_{ij}(y^0, x)\,, i \in K, \quad (4.2)$$

а целевая функция j-го агента:

$$f(\{\sigma_{ij}(\cdot)\}_{i\in K_j}, y) = \sum_{i\in K_j} \sigma_{ij}(y^0, x) - c_i(y^0, x). \quad (4.3)$$

Центры одновременно и независимо выбирают свои управляющие воздействия и сообщают их агентам. Рассматриваются Парето-эффективные равновесия Нэша игры центров, т. е. исследуются стратегии центров вида

$$\sigma_{ij}(y^0, x) = \begin{cases} \lambda_{ij}, & y_j = x_j \\ 0, & y_j \neq x_j \end{cases}, j \in N_i,\ i \in K. \quad (4.4)$$

Содержательно, центры договариваются о сотрудничестве, т. е. о том, что они будут вместе формировать мнения агентов x и нести затраты. Из условий оптимальности по Парето следует, что сумма затрат центров должна быть равна затратам агента, то есть:

$$c_i(y^0, x) = \sum_{j \in K_i} \lambda_{ji},\ i \in N. \quad (4.5)$$

Условие (4.5) означает, что центры должны распределить между собой затраты на изменение мнений каждого из агентов.

Определяется полезность i-го центра от «самостоятельного» взаимодействия с агентом

$$W_i = \max_x [H_i(x_{N_i}, y^0_{-N_i}) - \sum_{j \in N_i} c_i(y^0, x)],\ i \in K \quad (4.6)$$

и полезность в режиме сотрудничества

$$W_0 = \max_x [\sum_{i \in K} H_i(x) - \sum_{j \in N} c_j(y^0, x)]. \quad (4.7)$$

Вводится обозначение $\lambda = \|\lambda_{ij}\|$, через

$$S = \{x \in \Re^n_+ | \exists \lambda \in \Re^{nk}_+ : H_i(x) - \sum_{j \in N_i} \lambda_{ij} \geq W_i, i \in K, c_i(y^0, x) = \sum_{j \in K_i} \lambda_{ij}, i \in N\} \quad (4.8)$$

обозначается множество таких векторов мнений агентов, для реализации которых сотрудничество выгодно для центров. Множество пар векторов $x \in S$ и соответствующих матриц затрат центров λ называется *областью компромисса* в задаче распределенного управления АСС:

$$\Lambda = \{x \in \Re^n_+, \exists \lambda \in \Re^{nk}_+ | H_i(x) - \sum_{j \in N_i} \lambda_{ij} \geq W_i, i \in K, c_i(y^0, x) = \sum_{j \in K_i} \lambda_{ij}, i \in N\}. \quad (4.9)$$

Режим сотрудничества (*информационная кооперация*) по определению имеет место, если область компромисса (4.9) не пуста: $\Lambda \neq \varnothing$. Доказывается справедливость следующего утверждения.

Утверждение 4.1. Согласование интересов управляющих органов, осуществляющих информационные воздействия на членов АСС, возможно тогда и только тогда, когда

$$\max_x \left[\sum_{i \in K} H_i(x) - \sum_{j \in N} c_j(y^0, x)\right] \geq \sum_{i \in K} \max_x \left[H_i(x_{N_i}, y^0_{-N_i}) - \sum_{j \in N_i} c_i(y^0, x)\right]. \quad (4.10)$$

Условие (4.10) гарантирует возможность согласования интересов управляющих органов. Если оно не выполнено, то имеет место режим конкуренции. Если считать, что воздействия центров не «интерферируют», то есть агент соглашается принять мнение того центра, который предложил максимальное поощрение, не обращая внимания на информацию от других центров, то будет иметь место аукционное решение. Содержательно, режим конкуренции соответствует *информационной войне*, победителем в которой будет центр, имеющий максимальный ресурс (4.6). В этом случае для определения итоговых мнений вводится обозначение $x^i = \underset{x}{\operatorname{argmax}}\left[H_i\left(x_{N_i}, y^0_{-N_i}\right) - \sum_{j \in N_i} c_i(y^0, x)\right]$, $i \in K$, центры упорядочиваются в порядке убывания $\{W_i\}$: $W_1 \geq W_2 \geq \ldots \geq W_k$ и приводится утверждение 4.2.

Утверждение 4.2. Если условие (4.10) не выполнено, то мнение членов АСС, сложившееся в результате информационных воздействий, будет $(x_{N_1}, y^0_{-N_1})$

В разделе 4.2 рассмотрена ситуация информационного противоборства игроков в сети, в которой постепенно формируются мнения относительно, как правило, сложного и неоднозначного вопроса. Важной также является ситуация информационного противоборства, в которой происходит быстрое распространение *мнений простого вида* (фактоидов) и стоит задача определения оптимальных стратегий защиты и нападения. В разделе 4.3 формулируется и исследуется задача информационного противоборства в форме распространения информационной эпидемии и защиты от нее с учетом различной информированности и рефлексии управляющих субъектов. Связи между агентами в АСС задаются симметричной матрицей $G = (g_{km})_{k,\, m \,\in\, N}$, где $g_{km} = 1$ (ненулевое доверие), если между агентами k и m имеется связь, либо агенты совпадают; иначе $g_{km} = 0$. Наряду с агентами в ситуации участвуют два игрока – A и B. Игрок B стремится распространить в сети некоторое мнение. Для этого он может выбрать одного из агентов и «инфицировать» его, далее «инфекция» распространяется по сети. Формально: имеется последовательность моментов времени $\tau = 0, 1, \ldots$ Если в момент τ имеется множество инфицированных агентов $S_\tau \subset N$, то в следующий момент инфицированными окажутся все агенты из множества:

$$S_{\tau+1} = \{m \in N \mid \exists\, k \in S_\tau\ g_{km} = 1\}. \tag{4.11}$$

Игрок A стремится противодействовать распространению. Он проводит периодический мгновенный мониторинг сети, в ходе которого выявляет множество инфицированных агентов. Выявив инфекцию, игрок A может мгновенно ее остановить. Стратегией игрока B в данной игре является выбор агента $j \in N$, с которого он начинает инфицирование сети. Стратегией игрока A является выбор периода мониторинга – целого неотрицательного числа i. В общем случае множеством инфицированных агентов $\delta(i, j)$ при выборе игроками A и B стратегий i и j соответственно является множество S_i, определяемое за i шагов из рекуррентного соотношения (4.11) с начальным значением $S_1=\{j\}$. Стратегии игроки выбирают одновременно и независимо, т. е. разыгрывается игра в нормальной форме. Предполагается, что каждый агент $k \in N$ обладает для игроков определенной ценностью (a_k для игрока A и b_k для игрока B), а затраты игрока A на мониторинг с периодичностью i составляют c_i. При этих предположениях выигрыши игроков A и B при выборе пары стратегий (i, j) составляют

$$f_{ij} = -\sum_{k\in\delta(i,j)} a_k - c_i, \text{ и} \tag{4.12}$$

$$h_{ij} = \sum_{k\in\delta(i,j)} b_k. \tag{4.13}$$

При этом структура сети (матрица G) и параметры a_k, b_k, $k \in N$; c_i, $i = 0, 1, \ldots$, являются общим знанием игроков A и B.

Для данной модели приводится алгоритм построения биматрицы игры, которая полностью описывает ситуацию информационного противоборства. Показано, что в этой игре может отсутствовать равновесие Нэша в чистых стратегиях. В то же время для АСС, задаваемой полным графом, верно следующее утверждение.

Утверждение 4.3. В произвольной игре информационного противоборства на полном графе существует равновесие Нэша.

Далее исследуется вопрос о том, приведет ли стратегическая рефлексия в биматричной игре к преимуществам для игроков. Пусть $I = \{1, 2, ..., m\}$ – множество действий первого игрока, $J = \{1, 2, ..., n\}$ – множество действий второго игрока. Считается, что выполняются условия, обеспечивающие однозначное соответствие между рангом рефлексии агента и его действием: матрицы выигрышей (f_{ij}, h_{ij}) таковы, что у каждого игрока существует единственный наилучший ответ на любое действие оппонента; и максимальный гарантированный результат каждого игрока достигается ровно на одном действии. Каждый игрок может выбрать конечный ранг своей рефлексии. Это приводит к выбору соответствующего действия: обладая нулевым рангом, первый игрок выбирает гарантирующую стратегию –

действие $i_0 = \arg\max_{i\in I}\min_{j\in J} f_{ij}$, а обладая рангом $k \geq 1$ – действие $i_k = \arg\max_{i\in I} f_{ij_{k-1}}$. Аналогично для действий второго игрока: $j_0 = \arg\max_{j\in J}\min_{i\in I} h_{ij}$ – при нулевом ранге; $j_k = \arg\max_{j\in J} h_{i_{k-1}j}$ – при ранге $k \geq 1$. Справедливо следующее утверждение.

Утверждение 4.4 (Новиков Д.А., Чхартишвили А.Г., 2003). В биматричных играх неограниченное увеличение ранга рефлексии нецелесообразно, т. е. существует ранг рефлексии, превышение которого не приводит к новым действиям агентов. Максимальный целесообразный ранг не превышает max {min {n, m + 1}, min {m, n + 1}}.

Из утверждения 4.4 следует, что множество допустимых действий по выбору ранга конечно. Поэтому можно перейти из исходной игры к *игре рангов* стратегической рефлексии, в которой стратегией игрока является выбор ранга стратегической рефлексии. В этой игре справедливыми являются следующие утверждения.

Утверждение 4.5. Матрица выигрышей в игре рангов является подматрицей матрицы исходной биматричной игры.

Утверждение 4.6. Для произвольной биматричной игры переход к игре рангов не приводит к появлению новых равновесий.

Утверждение 4.7. В игре рангов есть не более двух равновесий.

Таким образом, игра рангов не увеличивает числа равновесий и может привести к исчезновению равновесий исходной игры. Показано, что в некоторых случаях любой исход игры рангов дает обоим игрокам лучший результат, чем равновесие Нэша исходной игры.

В разделе 4.4 демонстрируется сведение различных задач анализа информационного противоборства в АСС к задачам теории игр. Считается, что мнения агентов меняются согласно выражению (2.2) и что взаимодействие агентов продолжается достаточно долго для того, чтобы можно было воспользоваться оценкой вектора их итоговых мнений (2.3). При этом каждый агент хоть сколько-нибудь доверяет всем остальным агентам, т.е. $a_{ij} > 0$, $i, j \in N$. В рамках этого предположения все строки матрицы результирующего влияния A^∞ одинаковы (элементы $r_j = a_{ij}^\infty > 0$, $i, j \in N$ интерпретируются как *влиятельность* агентов) и одинаковы итоговые мнения всех агентов ($X = x_i$, $i \in N$, $X \in \mathfrak{R}^1$), то есть выражение (2.3) примет вид:

$$X = \sum_{j\in N} r_j x_j^0. \quad (4.14)$$

Предметом *информационного управления* являются начальные мнения агентов, целью – требуемые субъектом значения итоговых мнений агентов. Вводятся два игрока, каждый из которых может влиять на начальные мнения некоторых агентов: $F \subseteq N$ – множество *агентов влияния* первого игрока, $S \subseteq N$ – множество агентов влияния второго игрока, $F \cap S = \varnothing$. Предполагается, что информационное управление является унифицированным, то есть у всех агентов из множества F формируется начальное мнение $u \in U$, а у всех агентов из множества S формируется начальное мнение $v \in V$, где U и V – отрезки $\Re^1$. С учетом обозначений $r_F = \sum_{i \in F} r_i$, $r_S = \sum_{j \in S} r_j$ и $X^0 = \sum_{k \in N \setminus (F \cup S)} r_k x_k^0$ выражение (4.14) принимает вид:

$$X(u, v) = r_F\, u + r_S\, v + X^0, \tag{4.15}$$

то есть итоговое мнение членов АСС линейно зависит от управлений u и v, входящих соответственно с весами, которые определяются суммарной влиятельностью агентов влияния. Зависимость (4.15) итогового мнения агентов от управляющих воздействий позволяет формулировать теоретико-игровую модель взаимодействия игроков, осуществляющих эти воздействия. Для этого необходимо определить их целевые функции. Целевая функция первого (второго) игрока $f_F(u, v) = H_F(X(u, v)) - c_F(u)$ $(f_S(u, v) = H_S(X(u, v)) - c_S(v))$ определяется разностью между его «доходом», зависящим от итогового мнения агентов, и затратами на осуществление управления.

Совокупность $\Gamma = \{f_F(u, v), f_S(u, v), u \in U, v \in V\}$ целевых функций и множеств допустимых действий двух игроков задает семейство *игр*, различия между которыми порождаются конкретизацией информированности игроков и порядка функционирования. Если описание игры Γ и выражение (4.15) являются общим знанием среди игроков, которые выбирают свои действия однократно, одновременно и независимо, получим игру в нормальной форме, для которой можно искать и исследовать равновесия Нэша, их эффективность по Парето и т. д. Фиксация последовательности выбора игроками своих действий приводит к иерархической игре. Отказ от гипотезы об общем знании дает рефлексивную игру и т. п. Далее приводятся формализации таких ситуаций.

«Антагонистическая» игра. В качестве «статус-кво» выбирается нулевое значение мнений агентов ($X^0 = 0$) и считается, что первый игрок заинтересован в максимизации итогового мнения

($H_F(X) = X$), а второй – в минимизации ($H_S(X) = – X$), причем «ресурсы управления» у игроков одинаковы ($U = V = [d; D]$, $d < -1 < 1 < D$), как и функции затрат ($c_F(u) = u^2 / 2$, $c_S(v) = v^2 / 2$). Тогда целевые функции игроков сепарабельны по соответствующим действиям, значит, при одновременном независимом выборе игроками действий существует равновесие в доминантных стратегиях – РДС (u^d, v^d), где $u^d = r_F$, $v^d = – r_S$. Оказывается, что РДС не эффективно по Парето, а точка Парето неустойчива относительно индивидуальных отклонений игроков. Если существует третья сторона, обладающая правом контролировать выполнение игроками взятых на них обязательств, то можно найти условия, при которых игрокам выгодно выполнять обязательства, ведущие к устойчивой реализации точки Парето (влиятельности агентов влияния игроков не должны различаться слишком сильно).

Игра с «непротивоположными» интересами. Далее рассматривается игра в нормальной форме, отличающаяся от описанной выше только функциями «дохода» игроков: $H_F(X) = X – 2\,X^2$, $H_S(X) = X – X^2$ (первый игрок хотел бы добиться итогового мнения $X_F = 0.25$, а второй – мнения $X_S = 0.5$). Целевые функции игроков уже не сепарабельны по соответствующим действиям, поэтому находится равновесие Нэша:

$$u^* = \frac{r_F - 2r_F(r_S)^2}{4(r_F)^2+2(r_S)^2+1},\ v^* = \frac{r_S + 2r_S(r_F)^2}{4(r_F)^2+2(r_S)^2+1}\,.$$

Таким образом, чем больше суммарная репутация (влиятельность) агентов влияния второго игрока, тем меньше равновесное управляющее воздействие первого игрока и больше равновесное управляющее воздействие второго игрока, т. е. первый игрок стремится достичь меньшего мнения, а второй игрок – большего.

Далее рассматривается иерархическая игра типа Γ_1, в которой игроки имеют те же самые целевые функции, и первый игрок обладает правом первого хода. Второй игрок на своем ходе уже знает действие u первого игрока и максимизирует свой выигрыш, выбирая действие $v^*(u) = \frac{r_S - 2r_S r_F u}{2(r_S)^2+1}$. Множество выбора второго игрока состоит из единственного действия. Гарантирующая стратегия первого игрока в игре Γ_1 и его стратегия в равновесии Штакельберга:

$$u^* = \frac{r_F - 2r_F(r_S)^2}{4(r_F)^2+4(r_S)^4+4(r_S)^2+1};\ v^* = \frac{2(r_S)^3-(2(r_F)^2+1)r_S}{4(r_F)^2+4(r_S)^4+4(r_S)^2+1}.$$

В рефлексивной игре рассматриваются целевые функции, отличающиеся от описанных выше только функциями затрат игроков: $c_F(u) = u^2 / (2q_F)$, $c_S(v) = v^2 / (2q_S)$, где $q_F = 1$ и $q_S = 1/2$ – «эффективности» игроков. Предполагается, что каждый игрок знает свою эффективность, первый игрок считает, что общим знанием является $q_S = 1$, второй игрок информирован об этом и знает истинную эффективность первого игрока. Граф такой рефлексивной игры имеет вид: $2 \leftarrow 1 \leftrightarrow 12$. Тогда первый игрок в соответствии с равновесием Нэша выберет $u^* = \frac{r_F - 2r_F(r_S)^2}{4(r_F)^2+2(r_S)^2+1}$. Исходя из этого, второй агент выберет свой наилучший ответ $v^* = \frac{0.5r_S(1+2(r_F)^2)(1+2(r_S)^2)}{(1+(r_S)^2)(4(r_F)^2+2(r_S)^2+1)}$. При этом в сети установится следующее итоговое мнение $X = \frac{(r_F)^2+(r_S)^4+0.5(r_S)^2}{(1+(r_S)^2)(4(r_F)^2+2(r_S)^2+1)}$, которое в общем случае будет отличаться от ожидаемого первым игроком $X^1 = \frac{(r_F)^2+(r_S)^2}{4(r_F)^2+2(r_S)^2+1}$, то есть информационное равновесие не является стабильным. Для заданного выше графа рефлексивной игры информационное равновесие является стабильным только в двух случаях: либо представление первого игрока об эффективности второго истинно, либо суммарная репутация агентов влияния второго игрока равна нулю.

Равновесие в безопасных стратегиях. Рассматривается игра, в которой $H_F(X(u,v)) = \begin{cases} h_F > 0, & X \geq \hat{X} \\ 0, & X < \hat{X} \end{cases}$, $H_S(X(u,v)) = \begin{cases} h_S > 0, & X < \hat{X} \\ 0, & X \geq \hat{X} \end{cases}$, $c_F(u) = u$, $c_S(v) = v$, $U = V = [d; D]$, $d < -1 \leq 1 < D$, причем $h_F > D$, $h_v > |d|$. Содержательно, первый игрок заинтересован в принятии некоторого решения (для чего необходимо, чтобы мнение членов сети превысило порог $\hat{X}$), второй игрок заинтересован в блокировании этого решения. Для определенности считается, что $r_F D + r_S d + X^0 > \hat{X}$. В такой игре не существует равновесия Нэша, но существует *равновесие в безопасных стратегиях* (РБС), которое имеет вид: $((\hat{X} - r_S d - X^0)/r_F + \varepsilon; 0)$, где ε – произвольно малая строго положительная константа. Содержательно РБС таково, что первый игрок обеспечивает принятие нужного ему решения, причем второй игрок, даже если выберет максимально возможные действия, все равно не сможет изменить результат.

В **пятой главе** разработаны технологии анализа АСС на основе предложенных ранее моделей информационного влияния. В разделе

5.1 рассматриваются прикладные задачи анализа и управления активными сетевыми структурами.

В разделе 5.2 описана технология анализа АСС, предназначенная для решения задач мониторинга и анализа конкретной АСС, прогнозирования ее состояния и последующего информационного управления. Рассмотрены основные объекты концептуальной модели, позволяющей описывать и моделировать информационное взаимодействие в реальных АСС (этой модели соответствуют в различных аспектах математические модели глав 2–5). Изложен конструктивный подход к анализу АСС, заключающийся в поэтапной формулировке содержательных вопросов и гипотез, разработке формализованных показателей, разработке технологии сбора, передачи, хранения и обработки данных АСС, выработке рекомендаций и осуществлении управляющих воздействий, обеспечивающих требуемое состояние АСС. Рассмотрена технология мониторинга и анализа АСС, в разделе 5.3 приведено описание реализующего технологию программного комплекса анализа активных сетевых структур.

Вошедшие в состав программного комплекса методы представлены в разделах 5.4–5.7. В разделе 5.4 рассмотрены методы расчета влияния и влиятельности агентов АСС. Для расчета вводится следующая схема:

1) конкретизируются предпочтения центра относительно исходных данных и формализации функции влияния;
2) выполняется сбор и хранение исходных данных АСС;
3) собранные данные структурируются;
4) рассчитывается влияние и влиятельность агентов АСС.

Приведены результаты расчета влиятельности пользователей русскоязычного сегмента социальных сетей Facebook и ВКонтакте. Показано, что большую влиятельность имеет небольшое число пользователей. В частности, для сети ВКонтакте совокупная влиятельность всего одного процента наиболее влиятельных пользователей составляет 94–96% общей влиятельности всех пользователей. Таким образом, предложенный способ расчета влиятельности позволяет эффективно выявлять небольшое множество пользователей, оказывающих наибольшее влияние на действия остальных пользователей сети в рамках заданных центром предпочтений. На примере сети Reddit показаны существенные различия между разработанными вариантами формализации влияния.

В разделе 5.5 предлагается подход к выявлению таких связей между агентами, которые играют заметную роль в распространении действий в сети и образуют каналы распространения активности. Приведены методы расчета влияния связей АСС, а также методы выявления структур и устойчивых каналов распространения активности. Показаны результаты расчета влияний связей и связных структур каналов для сети ВКонтакте.

В разделе 5.6 приведен метод анализа защищенности агентов АСС от информационных воздействий. Рассматривается степень защищенности одного агента сети (цели) от информационных воздействий другого агента (источника). Приведены примеры модельных расчетов защищенности агентов.

В разделе 5.7 приведен метод выявления информационных сообществ агентов АСС, которые характеризуются источниками информации, оказывающими на них воздействия. Как показано в разделе 5.4, в онлайновых социальных сетях сравнительно мала доля пользователей с заметной влиятельностью, поэтому при выявлении сообществ в качестве признаков агента можно использовать оказываемые на него влияния со стороны наиболее влиятельных агентов. Схема предлагаемого метода такова.

1) Рассчитывается влиятельность агентов АСС и выбирается l наиболее влиятельных.

2) Формируются стохастические векторы влияний на всех агентов АСС со стороны влиятельных агентов. Полученные значения влияния представляются в виде матрицы $P = (p_{ij})$ размерности $n \times l$.

3) Вводится расстояние (метрика) между агентами $d(p_u, p_v) = 1 - \sum_{j=1}^{l} \min(p_{uj}, p_{vj})$, где p_u (p_v) – вектор влияний на агента u (v).

4) Выявляются информационные сообщества на основе рассчитанных расстояний. При этом для поиска центров сообществ используется оригинальный вычислительно эффективный алгоритм.

Приведены примеры выявления сообществ.

В **заключении** сформулированы основные выводы и результаты диссертационного исследования.

Основные результаты и выводы

На основании выполненных исследований предложен единый подход к разработке и исследованию моделей и методов информационного управления в активных сетевых структурах, заключающийся в общности описания компонент внутренней структуры агентов, иерархическом способе построения и применении концепции взаимного информационного влияния управляемых субъектов в АСС. На основе предложенного подхода:

1) разработан комплекс математических моделей информационного влияния агентов АСС: в базовой модели описывается изменение мнений простых агентов под информационным влиянием других членов АСС, в более сложных рассматривается изменение мнений с учетом репутации агентов и доверия к содержанию сообщений, а также исследуется поведение интеллектуальных агентов с учетом их информированности;

2) разработан метод моделирования и анализа информационного влияния в АСС на основе действий и интересов управляющего органа, в рамках которого разработана модель распространения действий, формализованы интересы управляющего органа, формализованы различные варианты влияния и влиятельности агентов и структур АСС, и доказан ряд утверждений, связанных с функциями влияния и влиятельности;

3) сформулированы и решены задачи информационного управления для всех разработанных моделей информационного влияния, в этих задачах предметом управления являются различные компоненты внутренней структуры агентов;

4) построена общая теоретико-игровая модель информационного противоборства в АСС, исследован ряд ее частных случаев, включая задачу распределенного контроля в АСС, для которой получены условия согласования интересов управляющих органов и охарактеризованы режимы информационной кооперации и войны;

5) сформулирована и исследована задача информационного противоборства в форме распространения информационной эпидемии и защиты от нее с учетом различной информированности и рефлексии управляющих субъектов; получены результаты в области теории рефлексивных игр, связанные с уменьшением числа равновесий;

6) различные задачи анализа информационного противоборства сведены к задачам теории игр, в которых равновесием игры является: равновесие в доминантных стратегиях, равновесие Нэша, «контрактное равновесие», равновесие Штакельберга, информационное равновесие и равновесие в безопасных стратегиях;

7) разработаны прикладные методы анализа АСС на основе моделей информационного влияния: методы расчета влияния и влиятельности агентов АСС, методы расчета влияния связей активных сетевых структур, методы выявления структур и устойчивых каналов распространения активности, метод анализа защищенности агентов активных сетевых структур от информационных воздействий, метод выявления сообществ агентов активных сетевых структур;

8) предложен подход к мониторингу и анализу АСС, разработана соответствующая технология мониторинга и анализа АСС.

Основные публикации по теме диссертации

Статьи в изданиях из списка ВАК

1. Губанов Д.А. Влияние в социальных сетях: варианты формализации // Управление большими системами. 2020. № 85. С. 51–71.
2. Губанов Д.А., Петров И.В. Информационные сообщества в социальных сетевых структурах. Ч.1. От основного понятия к математическим моделям формирования // Проблемы управления. 2021. 1. С. 15–23.
3. Губанов Д.А., Петров И.В. Информационные сообщества в социальных сетевых структурах. Ч.2. Математические сетевые модели формирования сообществ // Проблемы управления. 2021. № 2. С. 18–32.
4. Губанов Д.А., Петров И.В., Чхартишвили А.Г. Многомерная модель динамики мнений в социальных сетях: индексы поляризации // Проблемы управления. 2020. № 3. С. 26–33.
5. Бызов Л.Г., Губанов Д.А., Козицин И.В., Чхартишвили А.Г. Идеальный политик для социальной сети: подход к анализу идеологических предпочтений пользователей //Проблемы управления. 2020.№4. С.15-26.
6. Губанов Д.А., Новиков Д.А. Методы извлечения и анализа терминологических структур смежных предметных областей (на примере методологии) // Онтология проектирования. 2018. Т. 8, № 3 (29). С. 347–365.
7. Губанов Д.А., Чхартишвили А.Г. Связи дружбы и комментирования пользователей социальной сети Facebook // Управление большими системами. 2014. Вып. 52. С. 69–84.
8. Базенков Н.И., Губанов Д.А. Обзор информационных систем анализа социальных сетей // Управление большими системами. 2013.

Вып. 41. С. 357–394.
9. Губанов Д.А., Коргин Н.А., Новиков Д.А. Модели нечеткой сетевой экспертизы // Системы управления и информационные технологии. 2010. № 4. С. 13–18.
10. Губанов Д.А., Новиков Д.А. Модели унифицированного информационного управления в однородных социальных сетях // Управление большими системами. 2010. № 30.1. С. 722–742.
11. Новиков Д.А., Губанов Д.А., Чхартишвили А.Г. Информационные войны и социальные сети // Информационные войны. 2010. №3. С.44-53.
12. Губанов Д.А., Новиков Д.А. Модели распределенного контроля в социальных сетях // Системы управления и информационные технологии. 2009. № 3.1. С. 124–129.
13. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Модели влияния в социальных сетях // Управление большими системами. 2009. № 27. С. 205-281.
14. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Модели репутации и информационного управления в социальных сетях // Управление большими системами. 2009. № 26.1. С. 209–234.
15. Губанов Д.А., Чхартишвили А.Г. Об одной модели информационного противоборства в социальной сети // Системы управления и информационные технологии. 2009. № 3 (37). С. 13–16.
16. Губанов Д.А., Чхартишвили А.Г. О стратегической рефлексии в биматричных играх // Управление большими системами. 2008. Выпуск 21. С. 49–57.
17. Gubanov D. A study of a complex model of opinion dynamics in social networks / Journal of Physics: Conf. Series. Moscow: IOP Publishing Ltd., 2021. Vol. 1740. P. 1-6
18. Gubanov D.A., Chkhartishvili A.G. Influence Levels of Users and Meta-Users of a Social Network // Automation and Remote Control. 2018. Vol. 79, Iss. 3. P. 545–553.
19. Gubanov D.A., Mikulich L.I., Naumkina T.S. Language games in investigation of social networks: Finding communities and influential agents // Automation and Remote Control. 2016. Vol.77, Iss.1. P.144–158.
20. Gubanov D.A., Chkhartishvili A.G. A conceptual approach to online social networks analysis // Automation and Remote Control. 2015. Vol. 76, Iss. 8. P. 1455–1462.
21. Gubanov D.A., Chkhartishvili A.G. An actional model of user

influence levels in a social network // Automation and Remote Control. 2015. Vol. 76, Iss. 7. P. 1282–1290.
22. Gubanov D.A. A Study of Formal and Informal Relations of Russian-Speaking Facebook Users // Communications in Computer and Information Science. 2014. Vol. 436. P. 85-90.
23. Gubanov D.A., Novikov D.A., Makarenko A.V. Analysis methods for the terminological structure of a subject area // Automation and Remote Control. 2014. 75 (12). P. 2231-2247.
24. Gubanov D.A., Novikov D.A., Kalashnikov A.O. Game-theoretic models of informational confrontation in social networks // Automation and Remote Control. 2011. Vol. 72, № 9. P. 2001-2008.
25. Novikov D.A., Gubanov D.A., Chkhartishvili A.G. Informational influence and informational control models in social networks // Automation and Remote Control. 2011. 72 (7). C. 1557-1567.
26. Gubanov D.A. A Study of Formalizations of User Influence in Actional Model / Proc. of the 13th Intern. Conf. "Management of Large-Scale System Development" (MLSD). Moscow: IEEE, 2020.
27. Gubanov D.A., Petrov I.V. Multidimensional Model of Opinion Polarization in Social Networks / Proc. of the 12th Intern. Conf. "Management of Large-Scale System Development" (MLSD). Moscow: IEEE, 2019.
28. Gubanov D.A., Chkhartishvili A.G. On Approaches to Identifying Information Spread Channels in Online Social Networks / Proc. of the 12th Intern. Conf. "Management of Large-Scale System Development" (MLSD). Moscow: IEEE, 2019.
29. Gubanov D.A., Chkhartishvili A.G. Controlled Consensus in a Social Network with Simple Agents / Proc. of the 11th Intern. Conf. "Management of Large-Scale System Development" (MLSD). Moscow: IEEE, 2018.
30. Gubanov D.A., Zhilyakova L.Yu. Double-threshold Model of the Activity Spreading in a Social Network / Proc. of the 11th IEEE Intern. Conf. on Application of Information and Communication Technologies (AICT2017). Moscow: IEEE, 2017. Vol. 2. P. 267-270.
31. Gubanov D.A., Chkhartishvili A.G. Analysis of User Influence Types in Online Social Networks: An Example of VKontakte / Proc. of the 11th IEEE Intern. Conf. on Application of Information and Communication Technologies (AICT2017). Moscow: IEEE, 2017. Vol. 1.
32. Gubanov D.A., Chkhartishvili A.G. Models of information opinion and trust control of social network members / Proc. of the 18th IFAC World

Congress (Milano, 2011). Milan: Intern. Federation of Automatic Control (IFAC), 2011. P. 1991-1996.
33. Gubanov D.A. An Approach to Knowledge Management in Research Organization. Seoul: IFAC Publication, 2008. P. 8119 - 8123.

Монографии

34. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. 3-е изд., перераб. и дополн. М.: МЦНМО, 2018. 224 р. Перевод: Gubanov D.A., Novikov D.A., Chkhartishvili A.G. Social Networks: Models of information influence, control and confrontation. Cham, Switzerland: Springer International Publishing, 2019. 158 p.
35. Новиков Д.А., Губанов Д.А., Коргин Н.А., Райков А.Н. Сетевая экспертиза. 2-е изд. М.: ЭГВЕС, 2011. 166 с. Перевод: Gubanov D.A., Korgin N.A., Novikov D.A., Райков А.Н. E-Expertise: Modern Collective Intelligence. Switzerland: Springer, 2014. 112 p.

Доклады на конференциях

36. Губанов Д.А., Чхартишвили А.Г. О понятии информационного сообщества в социальной сети / Труды 13-й Мультиконференции по проблемам управления (МКПУ-2020). СПб.: АО «Концерн «ЦНИИ «Электроприбор», 2020. С.158-161.
37. Губанов Д.А., Козицин И.В., Чхартишвили А.Г. О выявлении идейно-политических предпочтений пользователей городских сообществ в онлайновой социальной сети / Труды 10-й Междунар. социол. Грушинской конф. «Жить в России. Жить в мире. Социология повседневности». М.: ВЦИОМ, 2020. С. 291–297.
38. Губанов Д.А., Чхартишвили А.Г. О влиятельности структур социальной сети / Труды 13-го Всероссийского совещания по проблемам управления (Москва, 2019). М.: ИПУ РАН, 2019. С. 2291–2294.
39. Губанов Д.А., Петров И.В., Чхартишвили А.Г. О моделировании динамики и оценке поляризации мнений в социальных сетях / Материалы междунар. науч.-практ. конф. «Теория активных систем - 50 лет» (ТАС-50, Москва). М.: ИПУ РАН, 2019. С. 510–519.
40. Губанов Д.А., Бойко Л.М. Об одном подходе к выявлению информационных предпочтений политически вовлеченных пользователей онлайновой социальной сети / Материалы междунар. науч.-практ. конф. «Теория активных систем - 50 лет». М.: ИПУ РАН, 2019. С. 503–509.

41. Губанов Д.А., Чхартишвили А.Г. Акциональная модель: о направлениях исследования социальных сетевых структур / Материалы 10-й Всероссийской мультиконференции по проблемам управления (МКПУ-2017, Таганрог). Таганрог: Издательство Южного федерального университета, 2017. Т. 1. С. 193–195.
42. Губанов Д.А., Чхартишвили А.Г. О методе кластеризации пользователей онлайновых социальных сетей на основе оказываемого на них влияния / Труды 60-й Всероссийской науч. конф. МФТИ (Москва, 2017). М.: МФТИ, 2017. Радиотехника и компьютерные технологии. С. 71–72.
43. Губанов Д.А., Чхартишвили А.Г. О расчете нормированной влиятельности пользователей онлайновой социальной сети в соответствии с акциональной моделью / Тезисы 59-й науч. конф. МФТИ с междунар. участием (Долгопрудный, 2016). М.: МФТИ, 2016.
44. Губанов Д.А. О взаимосвязи связей дружбы и комментирования в социальной сети Facebook / Материалы междунар. науч.-практ. конф. «Теория активных систем» (TAS'2014, Москва). М.: ИПУ РАН, 2014. С. 203–205.
45. Губанов Д.А. Об одной модели информационного взаимодействия в социальных сетях / Труды междунар. науч.-практ. конф. «Теория активных систем» (TAS'2011, Москва). М.: ИПУ, 2011. С. 235–237.
46. Губанов Д.А. Об одной модели информационной эпидемии в социальной сети / Пермь: ППУ, 2010.
47. Gubanov D.A., Korgin N.A., Novikov D.A. Models of Reputation Dynamics in Expertise by Social Networks / Proc. of the UKACC Intern. Conf. on CONTROL (Coventry, 2010). Coventry: IET, 2010. P. 203-210.
48. Gubanov D.A., Korgin N.A., Novikov D.A. Network expertise and dynamics of reputation / Proc. of X Intern. Meeting of the Society for Social Choice and Welfare. Moscow: HSE. 2010. P. 27.
49. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Нечеткие модели влияния в социальных сетях / Труды 6-й Всероссийской школы-конф. молодых ученых «Управление большими системами» (УБС'2009, Ижевск). Ижевск: ООО «Бон Анца», 2009. С. 141–145.
50. Губанов Д.А. Подход к имитационному моделированию информационного влияния в социальных сетях /М.: ИПУ РАН, 2009. С.84-90.
51. Губанов Д.А., Чхартишвили А.Г. Теоретико-игровые задачи

управления в линейных социальных сетях / Proc. of the 3rd Intern. Conf. on Game Theory and Management (GTM-2009, St. Petersburg). Петрозаводск: ИПМИ РАН, 2009. С. 18–21.

Патенты и свидетельства

52. Губанов Д. А. Программа для извлечения и анализа терминологических структур смежных предметных областей: Свидетельство о государственной регистрации программы для ЭВМ № 2019665358 РФ; Зарег. 22.11.2019.
53. Губанов Д. А. Программа для расчета влияния и влиятельности пользователей социальных сетей на основе акциональной модели: Свидетельство о государственной регистрации программы для ЭВМ № 2019665357 РФ; Зарег. 22.11.2019.
54. Губанов Д. А. Программа для расчета и сравнительного анализа мер центральности узлов в связных неориентированных сетях: Свидетельство о государственной регистрации программы для ЭВМ № 2019665415 РФ; Зарег. 22.11.2019.

Личный вклад автора в работах, опубликованных в соавторстве, заключается в следующем: в [4, 10, 14, 25, 27, 29, 32, 36, 39] – формализация и исследование математических моделей формирования представлений в условиях информационного влияния и управляющих воздействий, в [9, 12, 34, 35, 47, 48, 49] – участие в создании и исследовании математической модели, в [18, 21, 38] – подход к моделированию и анализу информационного влияния в АСС на основе действий и интересов управляющего органа, исследование свойств информационного влияния, в [15, 24, 25, 51] – постановка задачи информационного противоборства, исследование частных случаев общей модели противоборства, в [2, 3, 11, 13, 34, 35] – обзор математических моделей и методов информационного влияния, система классификации моделей и методов, в [19, 28, 31, 40, 42, 43] – разработка и исследование методов анализа информационного влияния, в [5, 6, 7, 18, 21, 23, 28, 30, 31, 37, 40, 43] – проведение экспериментальных исследований, в [8, 11, 20, 41] – общая схема исследования задач информационного влияния, управления и противоборства в АСС, технология анализа информационного влияния.