

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное  
учреждение высшего образования  
ЮЖНЫЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ

**МОДЕЛИ УПРАВЛЕНИЯ  
УСТОЙЧИВЫМ РАЗВИТИЕМ  
АКТИВНЫХ СИСТЕМ И ИХ  
ПРИЛОЖЕНИЯ**

Монография

Под редакцией Г.А. Угольницкого

Ростов-на-Дону

Издательство Южного федерального университета

2019

УДК 504.06 + 517.977 + 519.83

ББК 22.18 65.23 65.29

*Издание осуществлено при поддержке Российского научного фонда, проект №17-19-01038*

Рецензенты:

декан факультета управления ЮФУ, д.э.н., профессор Анопченко Т.Ю.,  
директор Института проблем управления РАН, член-корреспондент РАН,  
д.т.н., профессор Новиков Д.А.

Модели управления устойчивым развитием активных систем и их приложения: монография / Под ред. Г.А. Угольницкого. - Ростов-на-Дону: Издательство Южного федерального университета, 2019. - 330 с.

ISBN

В монографии излагаются результаты развития авторской концепции управления устойчивым развитием активных систем, основанной на использовании иерархических дифференциально-игровых моделей и информационных технологий их анализа. Предложены динамические обобщения моделей стимулирования и согласования общественных и частных интересов. Наряду с теоретическими, рассматриваются прикладные модели управления эколого-экономическими, организационными, социальными системами, а также методы их исследования: стохастическое оптимальное управление, имитационное моделирование, эвристические алгоритмы различного типа. В частности, проанализированы модели влияния и управления в социальных сетях, управления университетами как активными системами, информационно-аналитическая система управления водными ресурсами региона. Монография отражает результаты научных исследований и может использоваться в методических целях. В силу междисциплинарного характера книги её целевая аудитория весьма широка и включает научных и практических работников, экспертов и специалистов, аспирантов и магистрантов по направлениям "Прикладная математика и информатика", "Системный анализ и управление", "Экология и природопользование", "Конфликтология", "Менеджмент", "Управление персоналом", "Государственное и муниципальное управление".

Публикуется в авторской редакции.

© Коллектив авторов, 2019

© Южный федеральный университет, 2019

## ВВЕДЕНИЕ

Цель настоящего издания состоит в монографическом изложении результатов, полученных в ходе выполнения научно-исследовательского проекта "Разработка комплексной теории управления устойчивым развитием активных систем", поддержанного грантом Российского научного фонда №17-19-01038 (руководитель проекта Г.А. Угольницкий, основные исполнители Г.И. Белявский, Д.Б. Рохлин, А.Б. Усов). Основу книги составляют статьи авторского коллектива, опубликованные в 2017-2019 годах и частично ранее, имеются некоторые неопубликованные результаты.

Согласно авторской концепции, устойчивое развитие активной системы означает выполнение требований гомеостаза для управляемой динамической подсистемы при условии согласования интересов осуществляющих управление активных агентов.

Проблема согласования интересов занимает центральное место в политике, экономике и организационном управлении. При математической формализации этой проблемы используются теория игр, теория группового выбора, методы оптимизации и исследования операций. В российской науке ведущие позиции в этой области занимают теория активных систем и теория управления организационными системами (В.Н. Бурков, Д.А. Новиков), информационная теория иерархических систем (Ю.Б. Гермейер, В.А. Горелик, Ф.И. Ерешко, А.Ф. Кононенко), теория дифференциальных игр в форме характеристической функции (Л.А. Петросян, В.В. Мазалов, В.В. Захаров, Н.А. Зенкевич). Аналогичные зарубежные исследования относятся к теории контрактов (мотивации) и дизайну механизмов (Л. Гурвиц, Ж.-Ж. Лаффон, Р. Майерсон, Э. Маскин, Ж. Тироль).

Формальное изучение гомеостаза базируется на теории динамических систем и имитационном моделировании. Большой вклад в эту область внесли российские исследования по математическому моделированию эколого-экономических систем (А.Б. Горстко, В.И. Гурман, В.Ф. Крапивин, Н.Н.

Моисеев, Е.В. Рюмина, Ю.М. Свирежев, А.М. Тарко). Среди зарубежных работ следует отметить теорию живучести (viability) Ж.-П. Обена и многочисленные публикации по теории устойчивого развития (sustainability) (У. Адамс, С. Жанрено, Л. Дуайен, У. Кларк, В. Мартине). Здесь также активно используются дифференциально-игровые модели (М. Бретон, Ж. Заккур, Нго Ван Лонг).

Представляется необходимым конвергентный синтез перечисленных направлений научного поиска и их приложений. Настоящая монография комплексно развивает указанные исследования, особенно теорию активных систем, на базе авторского подхода, детально изложенного в предыдущей книге (Угольницкий 2016).

Первая глава монографии посвящена методологии управления устойчивым развитием активных систем. В первом параграфе изучается ключевое с точки зрения согласования интересов понятие активной системы: его возникновение, различные интерпретации, математическая формализация. Прослеживается связь активной системы с авторским понятием иерархически управляемых динамических систем, в этой связи вводится новое понятие расширенной активной системы. Приводится спецификация активных систем в различных предметных областях: организационные, территориальные, эколого-экономические активные системы. Формулируются и обсуждаются принципы функционирования активных систем.

Второй параграф анализирует понятие устойчивого развития и механизмы управления, применяемые для его достижения. Описываются история возникновения концепции устойчивого развития применительно к взаимоотношениям общества и природы и возможности её обобщения для произвольных активных систем. Устойчивое развитие активных систем трактуется как выполнение требований гомеостаза для динамической системы, управляемой заинтересованными в этом (мотивированными) активными агентами. Представлена математическая формализация понятий

гомеостаза и мотивированности (согласования интересов), для иерархически управляемых динамических систем дополнительно рассматривается понятие системной согласованности и показатели её измерения.

Третий параграф дополняет и суммирует полученные результаты в виде методики управления устойчивым развитием активных систем. Приводятся постановки задач управления для активных систем различной структуры, что находит отражение в соответствующих динамических теоретико-игровых моделях и их принципах оптимальности. Дается краткий обзор полученных к настоящему времени результатов моделирования в различных предметных областях, разбирается иллюстративный пример.

Замысел работ, вошедших во вторую главу монографии, состоит в динамическом обобщении моделей стимулирования и согласования интересов активных агентов. Это направление составляет основной вклад авторского коллектива в развитие современной теории активных систем. В первом параграфе рассматривается базовая модель теории управления организационными системами "Центр - агент" в дискретной стохастической постановке на бесконечном интервале времени. Показано, что при соответствующих модельных предположениях сохраняет силу основной результат для статической модели: возможность построения механизма стимулирования, в котором Центр компенсирует агенту его затраты при выполнении планового задания (возможно, с мотивирующей надбавкой) и ничего не платит в противном случае. При этом плановое задание находится в результате решения задачи оптимального управления для целевой функции, равной разности между доходом Центра и затратами агента.

Во втором параграфе эта же идея проводится для древовидных систем управления типа "Центр - несколько агентов". В первых двух моделях изучается марковская игра стимулируемых агентов при частных и более общих предположениях относительно множеств стратегий. Показано, что и здесь удаётся обобщить как аналогичный результат для статической модели с древовидной структурой, так и результат предыдущего параграфа для одного

агента. В третьей модели исследуется стимулирование агента Центром в динамической модели оптимальной эксплуатации биологических ресурсов с неопределёнными параметрами, где Центр считается дальновидным, а агент - близоруким. Получен результат о виде решения, приведены иллюстративные численные примеры.

Третий параграф посвящён динамическому обобщению моделей согласования общественных и частных интересов (СОЧИ-моделей) для ряда предметных областей: управления качеством мелководных экосистем, рыболовства, продвижения инноваций в организациях. Здесь основным методом исследования служит имитационное моделирование.

Третья глава монографии посвящена методам решения задач управления устойчивым развитием активных систем и соответствующим алгоритмам. Первые два параграфа описывают применение методов случайного поиска (генетические алгоритмы, алгоритм имитации отжига, бинарное деление) к решению дифференциальных иерархических игр, в том числе моделирующих механизмы стимулирования при распределении ресурсов. Подход основан на идее о том, что на практике управление Центра обладает "липшицевым" свойством, т.е. не может меняться резко произвольным образом, а либо сохраняется, либо увеличивается или уменьшается на небольшую заданную величину. Тогда использовать стратегию предъявляемой агентам альтернативы поощрения и сурового наказания (как в стандартных моделях стимулирования) не удаётся, и надо искать другие, адаптивные способы решения задачи. Во втором параграфе эта идея реализуется с использованием онлайн-подхода, когда Центр пошагово "обучается" на неизвестных ему заранее стратегиях оптимальных ответов агентов. Дается сравнительный анализ результатов численных расчётов на иллюстративных примерах.

Третий параграф посвящён методу качественно репрезентативных сценариев имитационного моделирования. Основная трудность применения мощной технологии компьютерной имитации состоит в необходимости

обоснования перехода от рассмотрения необозримого или даже бесконечного полного факторного плана вычислительных экспериментов к разумным неполным. Важная идея предлагаемого метода состоит в том, что для получения вполне приемлемой качественной картины прогноза достаточно использовать небольшое число значений каждой управляющей переменной, отражающих характерные принципиально различные варианты воздействия на управляемую динамическую систему. Описывается формализация условий проверки применимости этой идеи, близких к условиям внутренней и внешней устойчивости, приводятся примеры её численной реализации для различных моделей.

Четвёртая глава монографии посвящена приложениям теории и методов управления устойчивым развитием активных систем в различных предметных областях. В первом параграфе рассматриваются модели влияния и управления на сетях. Известные модели анализа и прогноза групповых мнений формулируются в маркетинговой интерпретации, приводятся авторские алгоритмы исследования моделей. Даются постановки задач оптимального и конфликтного управления мнениями целевой аудитории на сетях с учётом требований устойчивого развития, намечаются пути их решения на основе описанного выше метода качественно репрезентативных сценариев имитационного моделирования.

Во втором параграфе представлены модели управления устойчивым развитием учреждений высшего образования. Дан анализ университета как активной системы, предложены постановки ряда задач управления устойчивым развитием учреждений высшего образования на уровне кафедры, факультета и университета в целом. Эти задачи формулируются как статические модели дискретной оптимизации и динамические теоретико-игровые модели с учётом требований устойчивого развития. В частности, подробно исследованы теоретико-игровые модели борьбы с экономической коррупцией, актуальные для данной предметной области.

В третьем параграфе описан проект ИАС ВР РО - информационно-аналитической системы управления водными ресурсами (на примере Ростовской области). Ключевая особенность авторского замысла заключается в решении посредством ИАС задач управления устойчивым развитием с использованием теории управления организационными системами и реализации в составе ИАС ВР РО ряда механизмов управления. Кроме того, ИАС ВР РО обеспечивает ввод, хранение, обработку и использование данных мониторинга водных ресурсов. Разработка и реализация ИАС представляется наиболее перспективным направлением практического решения задач управления устойчивым развитием активных систем, материальным воплощением конвергентного синтеза разнородных моделей и методов на базе информационных технологий.

Усилия авторов по написанию книги распределились следующим образом:

введение, глава 1, заключение: Г.А. Угольницкий;

2.1: Д.Б. Рохлин, Г.А. Угольницкий;

2.2: Д.Б. Рохлин, Г.А. Угольницкий (при участии А.Б. Усова);

2.3: Г.А. Угольницкий, А.Б. Усов (при участии А.В. Никитиной, М.В. Пучкина, И.С. Семёнова, А.И. Сухинова, А.Е. Чистякова);

3.1, 3.2: Г.И. Белявский, Н.В. Данилова, Г.А. Угольницкий;

3.3: Г.А. Угольницкий, А.Б. Усов;

4.1: М.Т. Агиева (при участии Ю.В. Бабичевой, Н.М. Окулист);

4.2: М.Х. Мальсагов (при участии Г.А. Угольницкого, А.Б. Усова);

4.3: Г.А. Угольницкий, А.Б. Усов (при участии М.В. Пучкина).

Авторы благодарят Южный федеральный университет за издание настоящей монографии.



## **ГЛАВА 1. МЕТОДОЛОГИЯ УПРАВЛЕНИЯ УСТОЙЧИВЫМ РАЗВИТИЕМ АКТИВНЫХ СИСТЕМ**

Первая глава монографии посвящена методологии управления устойчивым развитием активных систем. В первом параграфе изучается ключевое понятие активной системы: его возникновение, различные интерпретации, математическая формализация. Прослеживается связь активной системы с авторским понятием иерархически управляемых динамических систем. Приводится спецификация активных систем в различных предметных областях: организационные, территориальные, эколого-экономические активные системы. Формулируются и обсуждаются принципы функционирования активных систем.

Второй параграф анализирует понятие устойчивого развития и механизмы управления, применяемые для его достижения. Описываются история возникновения концепции устойчивого развития применительно к взаимоотношениям общества и природы и возможности её обобщения для произвольных активных систем. Устойчивое развитие активных систем трактуется как выполнение требований гомеостаза для динамической системы, управляемой заинтересованными в этом (мотивированными) активными агентами. Представлена математическая формализация понятий гомеостаза и мотивированности, для иерархически управляемых динамических систем дополнительно рассматривается понятие системной согласованности и показатели её измерения.

Третий параграф дополняет и суммирует полученные результаты в виде методики управления устойчивым развитием активных систем. Приводятся постановки задач управления для активных систем различной структуры, что находит отражение в соответствующих динамических теоретико-игровых моделях и их принципах оптимальности. Дается краткий обзор полученных к настоящему времени результатов моделирования, разбирается иллюстративный пример.

## 1.1. Активные системы

Понятие активного элемента (агента) введено В.Н. Бурковым в 1969 году. Под активностью понимается наличие у агента собственных интересов, способность к формулированию целей и выбору способов их достижения, исходя из существующей обстановки и наличных ресурсов. Таким образом, агент ведёт себя стратегически; в частности, он может сознательно искажать передаваемую остальным агентам информацию, уклоняться от выполнения нежелательных действий, предлагать другим агентам компенсацию за реализацию выгодных ему стратегий. Взаимодействующие активные элементы образуют активные системы, которые в соответствии с известным философским принципом "целое больше суммы частей" приобретают новые свойства и цели. Для реализации этих целей вводится специальный активный элемент (Центр в терминологии В.Н. Буркова), руководящий всей системой. В статье (Бурков и Опойцев 1974) был выдвинут принцип метаигрового синтеза, согласно которому задача управления активной системой состоит в выборе Центром такого механизма, который максимизировал бы выигрыш Центра в равновесии Нэша в игре активных элементов. Эта идея ведёт к так называемому принципу открытого управления (Бурков и Лернер 1971). В силу активности элементов их индивидуально оптимальные стратегии не образуют глобально оптимальный план для системы в целом. Добиться выполнения глобально оптимального плана практически невозможно, поскольку активные элементы искажают необходимую информацию в своих интересах. Поэтому Центру приходится решать задачу глобальной оптимизации на множестве планов, оптимальных для элементов. Это может привести к потерям, но обеспечивает достоверную информацию.

В дальнейшем эти идеи оформились в теорию активных систем (ТАС) (Бурков 1977; Бурков и Новиков 1999). В 1970-80-х годах ТАС активно развивалась как в теоретическом аспекте, так и в области приложений. Был развит принцип согласованного планирования и построены реальные

системы управления на его основе. Была разработана и внедрена теория многоканальных активных систем, активно исследовались задачи распределения ресурсов, серьёзное развитие получили системы оценки и стимулирования производственной деятельности, противозатратные механизмы налогообложения и ценообразования и механизмы обмена ресурсами.

В XXI веке в ИПУ РАН на основе ТАС создана теория управления организационными системами (Новиков 2007; Механизмы 2011). В рамках этой теории обоснована классификация задач управления организационными системами, систематизированы и исследованы механизмы стимулирования, планирования и информационного управления в организациях, предложены решения задач защиты от манипулирования, активной экспертизы, распределения ресурсов. Теория управления организационными системами доказала свою перспективность в приложениях к различным областям: управлению динамическими системами (Новиков и др. 2002), эколого-экономической проблематике (Бурков и др. 2008), развитию инноваций (Новиков и Иващенко 2006), рефлексивному поведению (Новиков и Чхартишвили 2013), управлению в социальных сетях (Губанов и др. 2010), формированию организационных структур (Губко 2006). Продолжается деятельность по внедрению полученных результатов.

Другое направление изучения активных систем - информационная теория иерархических систем (ИТИС) (Гермейер 1975; Горелик и Кононенко 1982; Горелик и др. 1991), в основе которой лежит предложенный Ю.Б. Гермейером подход к исследованию операций (Гермейер 1971) на основе принципа гарантированного результата. Идея этого принципа состоит в том, что анализ сложных ситуаций принятия решений осуществляется с позиции одного из участников (Центра в терминах ТАС), в интересах которого проводится исследование соответствующих операций. Именно, находится наибольший выигрыш, который этот выделенный игрок может обеспечить себе при самых неблагоприятных условиях, порождаемых всеми

действующими факторами: остальными активными агентами, внешней средой, его собственными недостаточными знаниями. Гарантированный результат - это выигрыш выделенного агента в антагонистической игре с остальным миром, в которой этот агент имеет право первого хода - выбора и передачи информации другим агентам.

Ещё одну фундаментальную идею ИТИС выражает известная теорема Гермейера (Горелик и Кононенко 1982): ведущий (Центр) поощряет ведомого (агента), если последний выбирает желаемую для ведущего стратегию, и наказывает в противном случае, что позволяет обеспечить согласование интересов с позиции ведущего при учёте интересов ведомого. На случай динамических игр эта идея обобщена А.Ф. Кононенко и М.А. Гореловым (Горелик и др. 1991; Горелов и Кононенко 2014а,б; 2015). В общей теории дифференциальных игр (Dockner et al. 2000) указанный план реализуется с помощью так называемых триггерных стратегий. Их основная идея состоит в том, что агенты соглашаются следовать определённой траектории и поддерживают это соглашение путем наказания любого «отклониста» (defector).

В зарубежной литературе тематика активных систем исследуется в рамках теории контрактов, дизайна механизмов, теории группового выбора, теории общественного благосостояния. В теории контрактов, или теории мотивации (contract theory, theory of incentives) базовой моделью служит схема принципал-агент, отражающая иерархическое взаимодействие субъектов с асимметричной информацией (Bolton and Dewatripont 2004; Laffont and Martimort 2002; Salanie 1997; Юдкевич и др. 2002). Выделяются три основные модификации этой модели: модель неблагоприятного отбора (adverse selection) и её разновидность – модель информативных сигналов (signaling), модель постконтрактного оппортунистического поведения или морального риска (moral hazard), модель неполных контрактов (nonverifiability).

В модели *adverse selection* в момент заключения контракта принципал не знает некоторых существенных параметров агента (тип агента). Проблема заключается в выявлении информации и предложении агенту оптимального контракта, зависящего от его типа. В модели *signaling* агент может послать принципалу сообщение о своём типе, естественно, не бесплатно, то есть это сообщение является стратегией агента. Наиболее важна модель *moral hazard*, в которой агент может предпринимать неизвестные принципалу действия уже после заключения контракта. Задача принципала заключается в том, чтобы побудить агента действовать приемлемым для принципала образом. В модели *nonverifiability* все переменные известны обоим участникам, но не могут быть записаны в контракт, так как их значения не верифицируются судом (арбитром). Тут существенное значение приобретает возможность двустороннего пересмотра контракта (*renegotiation*) и проектирования специальных механизмов взаимодействия (*mechanism design*). Разумеется, что все модели допускают обобщение на случай нескольких участников и исследованы в различных постановках (дискретных и непрерывных, детерминированных и стохастических и т.п.). При этом получен ряд важных математических результатов, среди которых принцип раскрытия информации (*revelation principle*), аналогичный принципу открытого управления.

Еще одно важное направление в указанной области - дизайн механизмов (*mechanism design*), выросший на основе теории игр и теории группового выбора (*Algorithmic Game Theory 2007*; Николенко 2009). Дизайн механизмов решает всё ту же задачу побуждения активных агентов к выбору действий, выгодных Центру.

Итак, под *активной системой* понимается структурированная совокупность активных элементов (агентов). Наряду с этим, целесообразно рассматривать *расширенную* активную систему, которая включает исходную активную систему и управляемую ею динамическую систему, не содержащую активных элементов (рис. 1.1.1).

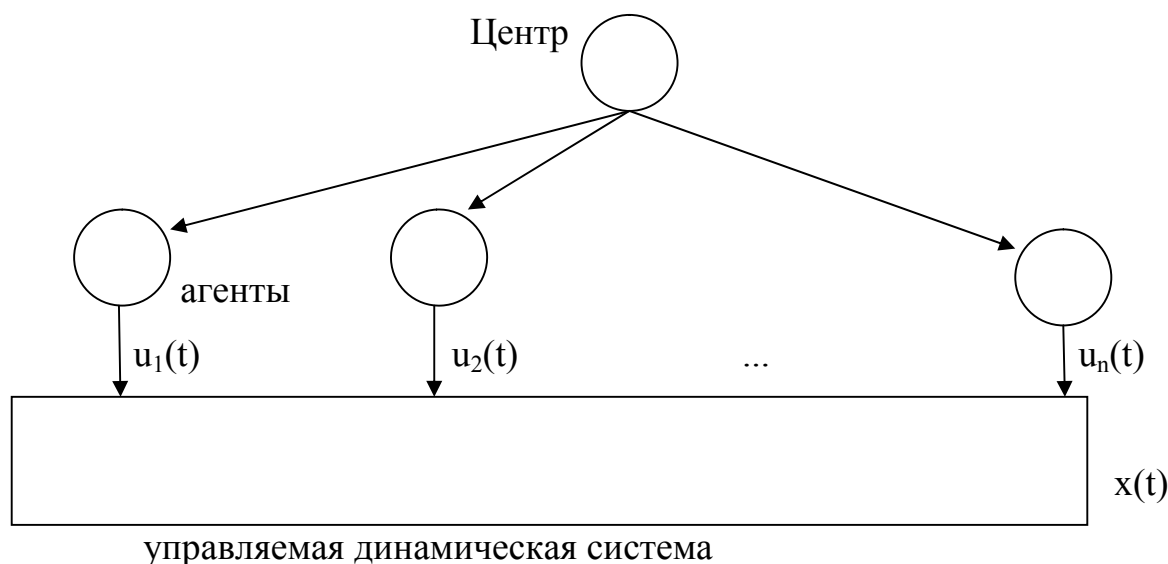


Рис. 1.1.1. Расширенная активная система

Понятие расширенной активной системы аналогично введённому ранее авторскому понятию иерархически управляемой динамической системы (Угольницкий 1999), простейший вариант которой показан на рис. 1.1.2.

Простейшая конфигурация иерархически управляемой динамической системы (ИУДС) включает три элемента:

- субъект воздействия верхнего уровня (Ведущий, Центр);
- субъект воздействия нижнего уровня (Ведомый, агент);
- управляемая динамическая система (УДС).

Предполагается, что взаимоотношения между элементами ИУДС устроены именно так, как показано на рис. 1.1.2: Ведущий воздействует на Ведомого, а Ведомый на УДС (непосредственное воздействие Ведущего на УДС отсутствует). Ведущего и Ведомого вместе можно рассматривать как совокупный субъект воздействия на УДС (активную систему), имеющий иерархическую структуру, этим объясняется название ИУДС. Конечно, в общем случае возможны более сложные конфигурации, включающие в качестве иерархического субъекта воздействия на УДС древовидные, ромбовидные и другие управляющие структуры и их сочетания.

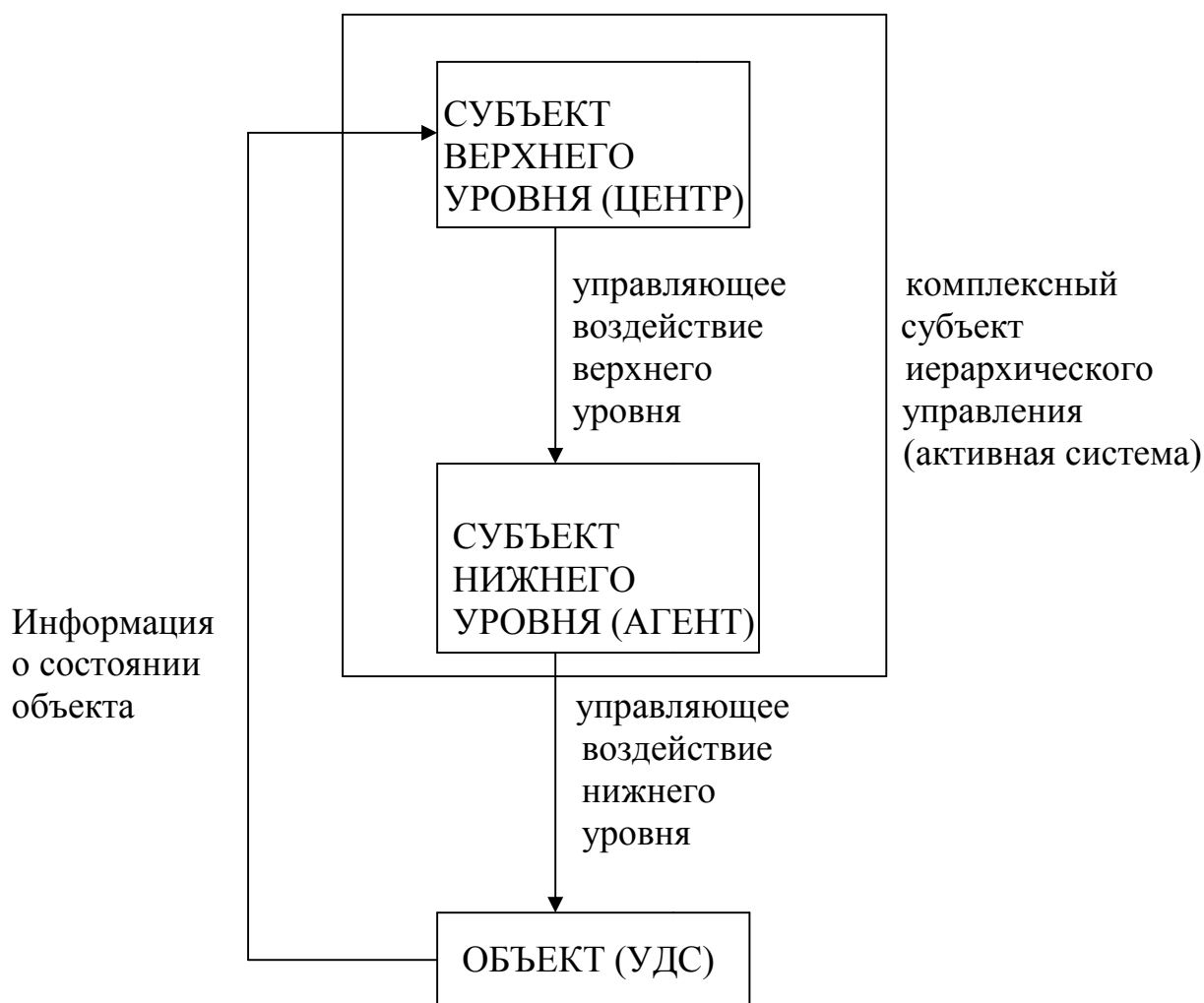


Рис. 1.1.2. Иерархически управляемая динамическая система

Смысл вводимого понятия ИУДС заключается в следующем. Воздействуя на УДС, Ведомый преследует определённые цели, которые в общем случае не совпадают с некоторыми объективно существующими целями поддержания УДС в заданном состоянии (условиями гомеостаза). Сама УДС, будучи пассивным объектом без участия людей, не может обеспечить достижение этих целей. Именно поэтому нужен Ведущий, который способен воздействовать на Ведомого с указанными целями. Тем самым ИУДС служат инструментом разрешения динамических конфликтов в активных системах с учетом требований их устойчивого развития.

Помимо основных структурных элементов, показанных на рис. 1.1.2, в состав понятия ИУДС входят следующие аспекты.

*Динамика управляемой системы.* Поскольку воздействие на УДС оказывает только Ведомый, то какая-либо специфика по сравнению с классической моделью УДС (Новиков 2016) здесь отсутствует. УДС характеризуется вектором состояния, компоненты которого изменяются во времени в силу саморазвития, управляющих воздействий Ведомого и неконтролируемых внешних факторов (случайных или неопределённых). Начальное состояние УДС предполагается известным.

*Объективная цель управления.* Здесь имеется в виду объективность по отношению к Ведомому, оказывающему воздействие на УДС в своих субъективных интересах. Субъектом, стремящимся достичь указанной цели, выступает Ведущий - выразитель объективных интересов. Это могут быть интересы общества или другой системы, более широкой, нежели пара "Ведомый-УДС". Объективная цель управления (гомеостаз УДС) может пониматься и соответственно формулироваться по-разному (более подробно эта тема обсуждается в следующем параграфе).

*Действия, интересы и возможности Ведомого.* Интересы Ведомого описываются стремлением к максимизации некоторой целевой функции, зависящей в общем случае от состояния системы и управляющих воздействий Ведомого и Ведущего. Множество допустимых управляющих воздействий Ведомого характеризует его возможности по управлению динамической системой. В общем случае это множество, как и целевая функция Ведомого, зависит от состояния системы и действий Ведущего; последнее обстоятельство обуславливает иерархическую природу отношений между Ведущим и Ведомым. Поскольку речь идет об управлении динамической системой, то целевая функция Ведомого также динамическая, т.е. представляет собой функционал (интеграл по периоду прогнозирования или сумму целевых функций на каждом шаге). Это обстоятельство принципиально, поскольку требует соизмерения разновременных затрат и эффектов (дисконтирования). Период прогнозирования может быть конечным или бесконечным. В первом случае обычно следует вводить



дополнительное условие, касающееся состояния системы за пределами периода прогнозирования, иначе оптимальное решение может привести к уничтожению системы в последний момент времени. По сути дела, это требование относится к объективной цели управления.

В содержательном плане воздействие Ведомого на УДС может быть весьма разнообразным. В контексте рассматриваемой проблемы наиболее естественно трактовать это воздействие как "эксплуатацию" УДС, использование её ресурсов для достижения цели Ведомого (максимизации его целевой функции). Цель Ведомого - максимизация некоторой функции полезности. Зачастую эта полезность имеет экономическую природу (максимум прибыли, минимум затрат), хотя вполне возможны и другие интерпретации.

*Действия, интересы и возможности Ведущего.* Главная цель Ведущего состоит в обеспечении некоторых объективных требований к состоянию УДС. Если указанная цель для Ведущего единственна, то его естественно назвать бескорыстным. Однако у Ведущего могут быть дополнительные субъективные интересы, к удовлетворению которых он стремится при *непрерывном* условии достижения главной цели. Как и в случае Ведомого, эти интересы выражаются в максимизации некоторой функции полезности. Такого Ведущего будем называть небескорыстным. Поскольку в сделанных предположениях Ведущий не оказывает прямого воздействия на УДС, то он добивается своих целей через управление Ведомым. Параметры воздействия определяются множеством допустимых управлений Ведущего. Если управляющие параметры Ведущего могут принимать любые физически возможные значения, то будем говорить, что его ресурсы не ограничены; в противном случае, конечно, они ограничены. Таким образом, с точки зрения целей можно различать бескорыстных и небескорыстных Ведущих, а с точки зрения возможностей управления - Ведущих с ограниченными и неограниченными ресурсами.

Наиболее богатый набор примеров ИУДС дают эколого-экономические системы. Взаимодействие общества и природы на различных уровнях формализуется с помощью понятия природно-хозяйственной (эколого-экономической) системы, концептуальная схема которой показана на рис. 1.1.3.

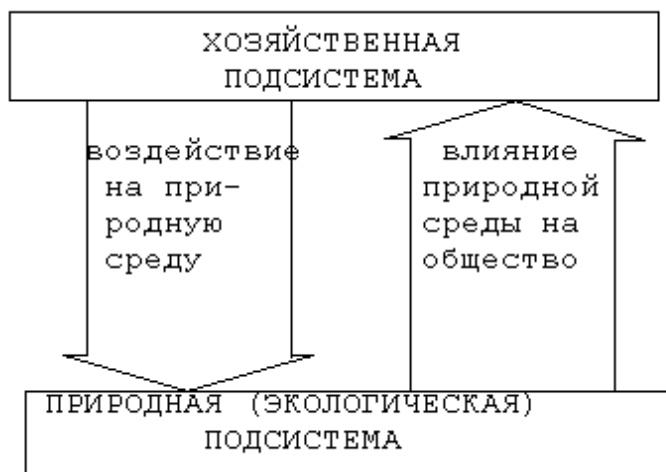


Рис. 1.1.3. Схема природно-хозяйственной системы

Уже на этом, самом общем уровне можно сделать определённые заключения о структуре природно-хозяйственной системы. Это контур, образованный двумя иерархическими подструктурами: с одной стороны, хозяйственная подсистема воздействует на экологическую; с другой стороны, экологическая подсистема оказывает влияние на хозяйственную. При этом воздействие на природную среду более важно с точки зрения последствий как для природы, так и для человеческого общества, и в этом смысле можно говорить об иерархии природно-хозяйственной системы в целом (хозяйственная подсистема как управляющая и экологическая как управляемая).

Таким образом, природно-хозяйственная система включает следующие аспекты: хозяйственную подсистему; экологическую подсистему; влияние природной среды на общество; воздействие общества на природную среду.

В состав *хозяйственной* подсистемы входят следующие элементы и связи: хозяйственная деятельность (промышленность, энергетика, сельское, лесное, водное хозяйство, строительство, их взаимодействие); население (населённые пункты, демографические процессы, рекреация и др.); правовая и административная регуляция (экологическое право, нормативные документы в области охраны окружающей среды и использования природных ресурсов, органы охраны окружающей среды и контроля её качества и т.п.). Таким образом, название "хозяйственная" подсистема не вполне точно, т.к. следует учитывать ряд аспектов социального характера. Поэтому, сохраняя устоявшееся название "природно-хозяйственная система", лучше говорить о ее "социально-экономической" подсистеме.

Ведущую роль при рассмотрении *экологической* подсистемы играет понятие биогеоценоза. Согласно В.Н. Сукачёву, "биогеоценоз - это совокупность на известном протяжении земной поверхности однородных природных явлений (атмосферы, горной породы, растительности, животного мира и мира микроорганизмов, почвы и гидрологических условий), имеющая свою особую специфику взаимодействий этих слагающих ее компонентов и определённый тип обмена веществом и энергией их между собой и с другими явлениями природы..." (Основы лесной биогеоценологии 1964:23).

*Влияние природной среды на общество* осуществляется по следующим направлениям: предоставление физического базиса (места) для различных видов деятельности человека; обеспечение человека ресурсами для его жизнедеятельности (воздух, вода, пища) с учётом их качества; предоставление ресурсов для производственной деятельности (сырьё, материалы, энергия); защита от космических излучений.

Наконец, *воздействие общества на природную среду* происходит в следующих основных формах: загрязнение окружающей среды; потребление природных ресурсов; охрана среды и восстановление ее ресурсов.

В связи с вышеизложенным при анализе эколого-экономических (природно-хозяйственных) систем представляется полезным использовать концепцию ИУДС. В соответствии с ней, природная (экологическая) подсистема трактуется как УДС, а в хозяйственной (социально-экономической) подсистеме выделяются Ведомый - непосредственный источник антропогенного воздействия на природную среду, и Ведущий, контролирующий это воздействие в целях соблюдения определённых экологических требований. Такую систему естественно назвать иерархически управляемой экологической системой (ИУЭС). Таким образом, ИУЭС представляет собой частный случай ИУДС и даёт богатый набор примеров ИУДС (Угольницкий 2016).

Другой важный класс ИУДС образуют организационные системы. Организацию целесообразно рассматривать как ИУДС, схема простейшего варианта которой представлена на рис. 1.1.2. При этом в качестве УДС выступает процесс, осуществление которого есть социальная функция организации. А вот идентификация Ведущего и Ведомого допускает различные варианты решения (Угольницкий 2002б).

Естественный подход заключается в том, чтобы считать Ведущим высшее руководство организации (предполагая любую организацию иерархически структурированной), а Ведомым – её коллектив. Поскольку требования гомеостаза отождествляются со стратегическими целями развития организации, то её руководство наиболее всего подходит на роль субъекта реализации этих целей, в то время как отдельные члены организации и их группы более склонны ориентироваться на достижение текущих экономических целей. Однако иногда и высшее руководство отдаёт предпочтение текущим субъективным интересам по сравнению с долгосрочными целями устойчивого развития организации. В этом случае роль выразителя объективных общеорганизационных интересов должно брать на себя государство (обычно в лице своих отраслевых, региональных или местных органов управления).

Кроме того, в организациях со сложной структурой схема ИУДС, показанная на рис. 1.1.2, также усложняется. Ведущий и Ведомый утрачивают монолитность и разделяются на множество целеустремлённых субъектов, обладающих своими целями и возможностями их достижения, а также критериями оценки эффективности различных способов достижения цели. Поэтому вместо линейной цепочки возникают более сложные древовидные, ромбовидные и комбинированные иерархические структуры.

Таким образом, необходимость иерархического управления устойчивым развитием организационных систем обусловлена следующими причинами: 1) несовпадением объективных стратегических общеорганизационных целей и субъективных краткосрочных целей членов организации; 2) сложной структурой организации, определяющей необходимость дополнительного согласования общеорганизационных, групповых и индивидуальных интересов.

Объективно на роль Ведущего в современных российских условиях более всего подходит государство, политика которого должна быть направлена на обеспечение условий устойчивого развития предприятий и организаций. Для достижения этой цели государственные органы располагают широким спектром воздействий, включающим кредитные, налоговые, информационные, протекционистские инструменты, социально-трудовую политику, меры по экологическому нормированию и т.д. вплоть до банкротства предприятий, не отвечающих заданным требованиям. Однако учёт государства в качестве субъекта управления усложняет структуру модели ИУДС, делая ее многоуровневой (Угольницкий 2016).

Очень интересным примером ИУДС может служить «двухэтажная» модель человека экономической теории самоконтроля, в которой на первом этаже расположен «деятель» (doer), ориентированный на достижение краткосрочных, сугубо эгоистических целей, а на втором – «планировщик» (planner), который задаёт «деятелю» правила поведения, ориентируясь при этом на долгосрочные цели (Thaler and Shefrin 1981).

Наконец, важный класс ИУДС образуют территориальные системы. На рис. 1.1.4 представлена схема региональной расширенной активной системы (Ougolnitsky 2017).

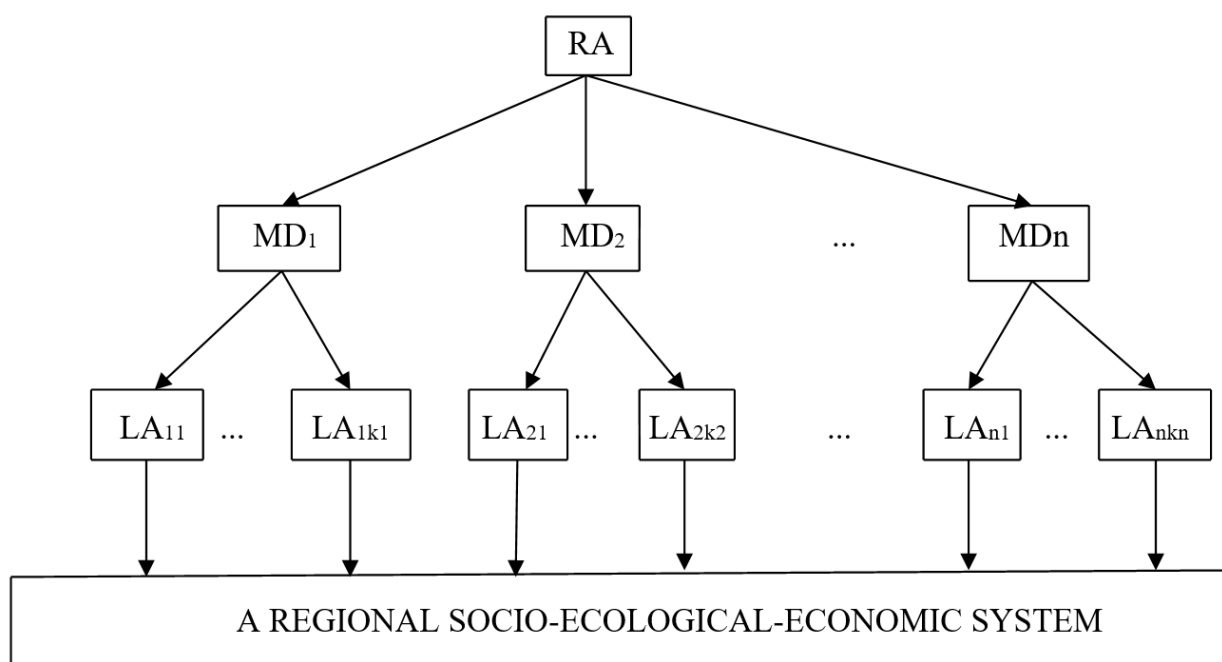


Рис. 1.1.4. Региональная социо-эколого-экономическая (расширенная активная) система

На верхнем уровне управления находится региональная администрация (RA, Центр). Средний уровень представлен муниципальными образованиями (MD). Возможен более сложный вариант, когда городские муниципальные образования и районы в их составе рассматриваются отдельно. Нижний уровень управления образован местными активными агентами (LA): предприятиями, организациями, индивидуальными предпринимателями. Объект управления - региональная социо-эколого-экономическая система, рассматриваемая как ИУДС.

В большинстве случаев LA - это экономические агенты, имеющие финансовые, трудовые и иные ресурсы и стремящиеся к максимизации своего дохода (прибыли). Однако, некоммерческие организации и местные органы управления также могут рассматриваться как LA с

неэкономическими целями. Экономические ЛА, в свою очередь, могут иметь социальные и экологические цели.

МД решают задачи социально-экономического развития соответствующих территорий с учётом экологических требований, используя определённые бюджетные ресурсы (Лексин и Швецов 2000). Цели и возможности РА структурно аналогичны таковым для МД, отличаясь от них объёмами и источниками финансирования. РА могут принимать собственные законы и имеют дополнительные средства воздействия на МД и ЛА.

Объект управления характеризуется тремя группами показателей, описывающих соответственно социальную сферу, экономический потенциал и состояние окружающей среды. Для определения конкретных значений этих показателей используются данные официальной статистики, которые могут дополняться экспертными оценками. Заметим, что только ЛА оказывают непосредственное влияние на региональную социо-эколого-экономическую систему. Органы управления верхних уровней лишь создают некоторые административные, законодательные и экономические рамочные условия этого влияния.

Закономерности изменения показателей состояния в силу естественных факторов и воздействия ЛА описываются динамическими балансовыми соотношениями. Их левая часть содержит новое значение показателя, а правая равна старому значению плюс приход минус расход, порождаемые рассматриваемыми процессами. Здесь используются модели математической экологии и экономики, а также динамические модели прогнозирования регионального уровня (Моделирование 2001).

Технически, административные управления РА и МД включают законодательные ограничения деятельности подчинённых им органов управления (нормы, квоты, стандарты качества). Экономические управления - это налоговые ставки, параметры привилегий, субсидий, дотаций.

Соответствующая трёхуровневая теоретико-игровая модель регионального управления имеет вид

$$J_0 = \sum_{t=1}^T e^{-\rho t} g_0(r^t, s^t, p^t, q^t, u^t, x^t) \rightarrow \max \quad (1.1.1)$$

$$r^t \in R; \quad s^t \in S; \quad (1.1.2)$$

$$J_i = \sum_{t=1}^T e^{-\rho t} g_i(r_i^t, p_i^t, q_i^t, u_i^t, x^t) \rightarrow \max \quad (1.1.3)$$

$$p_i^t \in P_i(r_i^t); \quad q_i^t \in Q_i(s_i^t); \quad (1.1.4)$$

$$J_{ij} = \sum_{t=1}^T e^{-\rho t} g_{ij}(r_i^t, p_{ij}^t, u_i^t, x^t) \rightarrow \max \quad (1.1.5)$$

$$u_{ij}^t \in U_{ij}(s_i^t, q_{ij}^t); \quad j = 1, \dots, k_i; \quad i = 1, \dots, n; \quad (1.1.6)$$

$$x^{t+1} = x^t + f(x^t, u^t), \quad x^0 = x_0; \quad (1.1.7)$$

$$r^t = (r_1^t, \dots, r_n^t); \quad s^t = (s_1^t, \dots, s_n^t);$$

$$p^t = (p_1^t, \dots, p_n^t); \quad q^t = (q_1^t, \dots, q_n^t); \quad p_i^t = (p_{i1}^t, \dots, p_{ik_i}^t); \quad q_i^t = (q_{i1}^t, \dots, q_{ik_i}^t);$$

$$u^t = (u_1^t, \dots, u_n^t) = (u_{11}^t, \dots, u_{1k_1}^t, \dots, u_{n1}^t, \dots, u_{nk_n}^t).$$

Здесь  $T$  - период рассмотрения (в годах);  $\rho$  - коэффициент дисконтирования;  $J_0, J_i, J_{ij}$  и  $g_0, g_i, g_{ij}$  - суммарные и текущие функции выигрыша RA, MD, и LA соответственно;  $r^t, s^t$  - экономические и административные управления RA в году  $t$ ;  $R, S$  - соответствующие множества допустимых управлений;  $p_i^t, q_i^t$  - экономические и административные управления  $i$ -го MD в году  $t$ ;  $P_i, Q_i$  - соответствующие множества допустимых управлений;  $u_{ij}^t$  - управление  $ij$ -го LA в году  $t$ ;  $U_{ij}$  - соответствующее множество допустимых управлений;  $x^t$  - вектор состояния региональной социо-эколого-экономической системы в году  $t$ ;  $x_0$  - вектор начальных значений показателей состояния в базовом году;  $f$  - набор моделей динамики состояния;  $n$  - число MD;  $k_i$  - число LA в  $i$ -м MD.



Модель (1.1.1) - (1.1.7) представляет собой иерархическую разностную игру, дискретная форма которой ориентирована на исследование с помощью имитационного моделирования по методу сценариев (Law and Kelton 2000). Решения игры (1.1.1)-(1.1.7) понимаются в смысле Штакельберга (Basar and Olsder 1999). Как правило, РА выбирает только  $s^t$  при фиксированном  $r^t$  (административное управление, или принуждение) или наоборот (экономическое управление, или побуждение). Аналогично, при заданных  $s^t$  или  $r^t$   $MD_i$  выбирают  $q_i^t$  при фиксированном  $p_i^t$  или наоборот. Наконец, при заданных  $q_i^t$  и  $p_i^t$ , а также  $s^t$  и  $r^t$   $LA_{ij}$  выбирают свои управления  $u_{ij}^t$ . На самом деле, имитационное моделирование есть практически единственная возможность решения игры (1.1.1)-(1.1.7) в общей постановке. Упрощённые версии могут быть исследованы посредством стандартных методов, таких как уравнения Гамильтона-Якоби-Беллмана или принцип максимума Понтрягина совместно с численными методами.

Можно сформулировать следующие принципы функционирования активных систем.

1. *Принцип активности агента*: каждый агент стремится достичь своих целей наилучшим образом, исходя из наличных ресурсов, в т.ч. информации о ситуации.

Математически это можно записать как

$$J_i(u_i, u_{-i}) \rightarrow \max, \quad u_i : B_i \rightarrow U_i, \quad i = 1, \dots, n. \quad (1.1.8)$$

Здесь  $J_i$  - целевая функция (функция выигрыша)  $i$ -го агента,  $U_i$  - множество его допустимых действий,  $(u_1, \dots, u_n)$  - множество стратегий агентов,  $u_{-i} = (u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_n)$ . Важно подчеркнуть, что: 1) выигрыш агента зависит не только от его собственных действий, но и от действий других агентов, т.е. рассматривается игровая постановка; 2) стратегия агента  $u_i$  - это функция, значения которой - допустимые действия агента, а область определения  $B_i$  - множество всей доступной агенту информации о ситуации, целях и возможностях других агентов.

2. *Принцип иерархического согласования интересов*: решение задачи управления активной системой ищется с позиции Центра, который максимизирует свой выигрыш на множестве наихудших для него ситуаций, порождаемых оптимальной реакцией агентов на выбранное Центром управляющее воздействие.

Математическая формулировка данного принципа имеет вид

$$J = \max_{v \in V} \min_{u \in R(v)} J(v, u), \quad (1.1.9)$$

где  $J$  - функционал выигрыша Центра,  $v, u$  - векторы управляющих стратегий Центра и агентов соответственно;  $V$  - множество допустимых стратегий Центра;  $R(v)$  - множество оптимальных ответов агентов на выбранную Центром стратегию,  $\forall v R(v) \subseteq U = U_1 \times \dots \times U_n$ . Если  $\max$  и  $\min$  в формуле (1.1.9) не достигаются, то их можно заменить на  $\sup$  и  $\inf$  соответственно.

Этот принцип хорошо известен в различных формулировках: принцип гарантированного результата Ю.Б. Гермейера, принцип открытого управления В.Н. Буркова, *revelation principle* в теории контрактов. В случае нескольких агентов обычно предполагается, что  $R(v) = NE(v)$ . т.е. оптимальная реакция представляет собой равновесие Нэша в игре агентов (особенно это подчёркивается в принципе открытого управления).

3. *Принцип невозможности системного выбора*: множество агентов не способно прийти к коллективному решению, удовлетворяющему всем желательным свойствам.

Математическую формулировку этого принципа дают широко известные теоремы Эрроу и Гиббарда-Саттертуэйта в теории коллективного выбора. Рассмотрим множество альтернатив  $A$  (кандидатов) и множество активных агентов  $I$  (выборщиков). Обозначим через  $L$  множество линейных порядков на  $A$  (изоморфное множеству перестановок), так что  $\prec \in L$  есть общий порядок на  $A$  (антисимметричный и транзитивный). Тогда  $a \succ_i b$  означает, что  $i$  предпочитает альтернативу  $a$  альтернативе  $b$ .

*Определение.* Функция  $F : L^n \rightarrow L$  называется *функцией общественного благосостояния*, а  $f : L^n \rightarrow A$  - *функцией общественного выбора*.

*Определение.* Функция общественного благосостояния называется *анонимной*, если  $\forall \prec \in L F(\prec, \dots, \prec) = \prec$ , т.е. если все выборщики имеют одинаковые предпочтения, то и групповое предпочтение такое же.

Выборщик  $i$  называется *диктатором* для функции общественного благосостояния  $F$ , если  $\forall \prec_1 \dots \prec_n \in L F(\prec_1 \dots \prec_n) = \prec_i$ . Функция не диктаторская, если ни один из выборщиков не диктатор.

Функция общественного благосостояния *независима от несвязанных альтернатив*, если  $\forall a, b \in A$  групповое предпочтение между  $a$  и  $b$  зависит только от предпочтений отдельных выборщиков между  $a$  и  $b$ .

*Теорема Эрроу.* Пусть  $|A| \geq 3$ . Тогда любая анонимная и независимая от несвязанных альтернатив функция общественного благосостояния диктаторская.

*Определение.* Функция общественного выбора  $f$  *стратегически манипулируема* выборщиком  $i$ , если для некоторых  $\prec_1 \dots \prec_n \in L$  и некоторого  $\prec'_i$  справедливо  $a \prec_i a'$ , где  $a = f(\prec_1, \dots, \prec_i, \dots, \prec_n)$ ,  $a' = f(\prec_1, \dots, \prec'_i, \dots, \prec_n)$ . Функция  $f$  называется *мотивационно совместимой* (incentive compatible), если она не манипулируема.

*Определение.* Выборщик  $i$  называется *диктатором* для функции общественного выбора  $f$ , если  $\forall \prec_1, \dots, \prec_n \in L, \forall a \neq b, a \succ_i b \Rightarrow f(\prec_1, \dots, \prec_n) = a$ . Функция общественного выбора *диктаторская*, если для нее есть диктатор.

*Теорема Гиббарда-Саттертуэйта.* Пусть  $|A| \geq 3$ ,  $f$  - мотивационно совместимая функция общественного выбора, отображающая  $L$  на  $A$ . Тогда  $f$  диктаторская.

Собственно, все усилия дизайна механизмов направлены на такое изменение исходной модели, которое позволяет избежать категоричного запрета этой теоремы (и теоремы Эрроу) и всё-таки построить недиктаторскую функцию, защищённую от манипулирования.

Наиболее распространены попытки введения "денег", т.е. количественной оценки интенсивности предпочтений. В этом случае механизм не только выбирает альтернативу, но и определяет платежи агентам для каждого выбора.

Зададим предпочтения агента  $i$  функцией оценки  $v_i : A \rightarrow R$ , где  $v_i \in V_i$  - некоторому известному множеству.

*Определение. Механизмом (прямого раскрытия информации)* называется функция общественного выбора  $f : V_1 \times \dots \times V_n \rightarrow A$  вместе с вектором платёжных функций  $p_1, \dots, p_n$ , где  $p_i : V_1 \times \dots \times V_n \rightarrow R$  - платёж игроку  $i$ .

*Определение. Механизм  $(f, p_1, \dots, p_n)$  мотивационно совместимый*, если  $\forall i \in I \forall v_1 \in V_1, \dots, v_n \in V_n \forall v'_i \in V_i v_i(a) - p_i(v_i, v_{-i}) \geq v_i(a') - p_i(v'_i, v_{-i})$ , где  $a = f(v_i, v_{-i})$ ,  $a' = f(v'_i, v_{-i})$ .

*Определение. Механизм  $(f, p_1, \dots, p_n)$  называется механизмом Викри - Кларка - Гроувза*, если:

- $f(v_1, \dots, v_n) \in \text{Arg max}_{a \in A} \sum_i v_i(a)$ ;
- существуют функции  $h_1, \dots, h_n$ , где  $h_i : V_{-i} \rightarrow R$  (т.е. не зависящие от  $v_i$ ), такие, что  $\forall v_1 \in V_1, \dots, v_n \in V_n \quad p_i(v_1, \dots, v_n) = h_i(v_{-i}) - \sum_{j \neq i} v_j(f(v_1, \dots, v_n))$ .

*Теорема Викри - Кларка - Гроувза.* Механизмы Викри-Кларка-Гроувза мотивационно совместимы.

Таким образом, в отличие от общей постановки, введение денег позволяет построить защищённый от манипулирования механизм, который к тому же максимизирует общественное благосостояние (Algorithmic Game Theory 2007).

4. *Принцип стратегического информирования:* передача агентом информации другим агентам - это его стратегия, направленная на достижение собственных целей. В частности, агентам может быть выгодно искажать передаваемую информацию.

Рассмотрим в качестве примера стратегического голосования следующий профиль предпочтений активных агентов на множестве альтернатив  $\{x,y,z\}$ :

Группа А (три агента)	Группа В (два агента)	Группа С (два агента)
x	z	y
y	x	z
z	y	x

По критерию большинства будет избрана альтернатива x. Агенты из группы С, не желая избрания наихудшей для них альтернативы, могут исказить свои предпочтения, поставив на первое место альтернативу z. Тогда возникает следующий профиль, в котором победителем становится z.

Группа А (три агента)	Группа В (два агента)	Группа С (два агента)
x	z	z
y	x	y
z	y	x

Заметим, что и для группы В такое искажение предпочтений агентов из С (манипулирование) оказывается выгодным. Интересные исторические примеры манипулирования приведены в (Алескеров и др. 2006).

## 1.2. Устойчивое развитие и методы управления

Словосочетание “sustainable development”, переведённое на русский язык как «устойчивое развитие», впервые было введено в научный и публицистический обиход в докладе «Всемирная стратегия охраны природы», подготовленном Международным союзом охраны природы и природных ресурсов в 1980 году. Уже в этом докладе развитие определялось как «модификация биосферы и использование людских, финансовых, возобновляемых и невозобновляемых природных ресурсов для удовлетворения потребностей людей и улучшения «качества жизни»; чтобы развитие было устойчивым, следует учитывать не только его экономические аспекты, но и социальные и экологические факторы...» и отмечалось, что «сохранение природы – это такое управление использованием человеком ресурсов биосферы, которое может принести устойчивые прибыли современному поколению, не подвергая при этом сомнению потенциальные возможности удовлетворения потребностей будущих поколений» (Дрейер и Лось 1997:8).

Однако серьёзное внимание к проблематике устойчивого развития (УР) было привлечено лишь после публикации доклада «Наше общее будущее» (1987), подготовленного Комиссией ООН по окружающей среде и развитию – так называемой «комиссией Брундтланд». В этом докладе УР определяется как развитие, «которое обеспечивает нужды современного поколения и не подрывает возможностей для будущих поколений удовлетворять их потребности» (Our Common Future 1987:43). Выводы «комиссии Брундтланд» легли в основу решений, принятых на Конференции ООН по окружающей среде и развитию (Рио-де-Жанейро, 1992). В документах Рио-92 УР определяется как позволяющее на долговременной основе обеспечить стабильное экономическое развитие, не приводя к деградации природной среды, что гарантирует удовлетворение потребностей не только настоящего, но и будущих поколений.

Во второй половине 1990-х годов в проблематике УР прочно утвердилась концепция «трех столпов» (three pillars), означающая необходимость одновременного учёта и балансирования экологических, экономических и социальных целей.

Несмотря на активное обсуждение концепции УР и ряд принятых официальных решений, единство в определении и трактовке этого термина до сих пор отсутствует. Уже в работе (Pezzey 1989) собраны более 60 определений УР, предложенных разными авторами. Приведём некоторые из более поздних определений: «устойчивое развитие – это такое экономическое развитие, которое не подрывает природную базу для будущих поколений и возрастает в расчёте на душу населения»; «устойчивое развитие – это экономическое развитие, обеспечивающее устойчивость окружающей среды и устойчивый, постоянный экономический рост»; «коэволюция человека и биосферы, т.е. такое взаимоотношение природы и общества, которое допускает их совместное развитие»; «путь максимизации долговременных выгод для человечества»; «повышение качества жизни людей, проживающих в пределах поддерживающей ёмкости экосистем» и т.д. (Данилов-Данильян и Лосев 2000:104). Развёрнутое синтетическое определение предлагает Г. Сдасюк: "Устойчивое развитие – это многоуровнево-иерархический управляемый процесс коэволюционного развития природы и общества (при массовом и осознанном участии населения), цель которого – обеспечить здоровую, производительную жизнь в гармонии с природой ныне живущим и будущим поколениям на основе охраны и обогащения культурного и природного наследия" (Переход к устойчивому развитию... 2002:18). Ещё в 2000 году Международный институт УР опубликовал список, включающий более сотни наборов принципов УР, каждый из которых используется некоторой организацией в общественном, академическом, производственном и иных секторах. Подробный анализ проблематики УР, а также математических моделей и систем поддержки решений в этой области дан в (Угольницкий 2016).

Изучение существующих концепций УР позволяет сделать следующие выводы (Угольницкий 2010). Во-первых, УР подразумевает необходимость одновременной реализации экологического и экономического императивов. Для УР человечества требуется не только соблюдение допустимой антропогенной нагрузки на окружающую природную среду, но и экономическое развитие, обеспечивающее удовлетворение материальных потребностей людей на уровне, который определяется исторически конкретными социально-культурными условиями общественной жизни. Одностороннее доминирование экономического императива приводит к неадекватным концепциям «природоохранной экономики», а одностороннее доминирование экологического императива – к столь же неадекватным концепциям «нулевого роста». Соответственно, при математическом моделировании УР на любом уровне следует в равной мере учитывать показатели, закономерности и критерии, относящиеся к экологической и экономической подсистемам.

Во-вторых, следует признать пользу математического моделирования для решения проблемы УР. Хотя большинство моделей не позволяет дать точный прогноз динамики природно-хозяйственных объектов и сформулировать исчерпывающие рекомендации для лиц, принимающих решения, модели играют исключительно важную роль в уточнении понятий, становлении единого языка, формировании концептуального багажа учёных, политиков, предпринимателей и широкой общественности. Для отдельных процессов эколого-экономического взаимодействия модели позволяют получить количественные оценки.

В-третьих, УР не может быть достигнуто само по себе и требует специальных целенаправленных управленческих усилий. Как правило, непосредственные субъекты воздействия на окружающую природную среду преследуют цели, отличные от целей УР и даже противоречащие им. Поэтому необходимы действия верхних уровней иерархической системы управления, направленные на обеспечение условий УР. Хотя кардинальный



путь решения проблем УР - это экологическое воспитание и повышение соответствующего самосознания, в обозримом будущем выполнение экологического императива потребует серьёзных мер стимулирования и даже принуждения. В математических моделях отражение механизмов иерархического управления возможно с помощью аппарата теории игр.

В рамках развиваемой авторами концепции, УР активной системы есть совместное выполнение условий гомеостаза для управляемой входящими в её состав активными агентами динамической системы и мотивированности агентов. Начнём с понятия гомеостаза (Угольницкий 2016).

Активные системы должны противостоять внешним воздействиям или приспосабливаться к ним, обеспечивая сохранение условий своего существования и целенаправленного развития. Как считал французский физиолог К. Бернар, «постоянство внутренней среды есть условие свободной жизни организма» (Бернар 1878). В 1932 году У. Кэннон ввел термин «гомеостаз», под которым он понимал поддержание относительного динамического постоянства всего организма (Cannon 1932). Нельзя не заметить сходства понятия гомеостаза как динамического постоянства организма и концепции УР как сочетания экономического развития (динамики) и экологической стабильности (постоянства) на уровне биосферы в целом и ее отдельных подсистем. На наш взгляд, понятие гомеостаза также можно и нужно рассматривать не только на уровне отдельного организма, но и на уровне популяции, экосистемы и эколого-экономических систем более высокого уровня, а также применительно к произвольным активным системам. Здесь речь идёт о входящей в состав расширенной активной системы динамической системе, управляемой активными агентами.

Назовем *гомеостазом* область значений существенных характеристик системы, в которой возможно ее нормальное существование и развитие. Функционирование любой динамической системы характеризуется набором показателей, значения которых изменяются во времени. Требование гомеостаза в первом приближении означает, что все показатели состояния

системы в течение рассматриваемого периода времени (достаточно длительного или даже бесконечного) принимают значения из заданных диапазонов, а в частном случае – принимают точечные заданные значения. Например, хорошо известны точечные требования к физиологическим параметрам человеческого организма – температура 36.6 градусов Цельсия, артериальное давление 120 на 80 миллиметров ртутного столба и т.д. В то же время допустимы небольшие отклонения от указанных эталонных значений, которые образуют приемлемые диапазоны функционирования здорового организма. Аналогичным образом можно сформулировать точечные и интервальные требования, определяющие гомеостаз любой расширенной активной системы.

Представляется естественным обобщение этого понятия для «производственного организма» - промышленного предприятия, а также организаций других типов, также рассматриваемых как «социальные организмы». Так, применительно к макроэкономическому уровню А. Бузгалин отмечает: "Проблема в том, чтобы экономико-политические власти страны ставили в качестве целей не просто «увеличение ВВП» или задачу обогнать по его объёму, скажем, ФРГ (страну, которая почти в 2 раза меньше РФ по численности населения), а прежде всего качественные индикаторы, объединённые в дерево целей. Чтобы эти индикаторы были не разбросаны, а сведены в систему задач, за выполнение которых через соответствующий срок перед гражданами должны отчитаться те, кто реально руководит экономикой"<sup>1</sup>.

Однако создание методики измерения гомеостаза, позволяющей делать обоснованные заключения о факте и тем более степени его наличия в данной организации, оказывается далеко не простым делом, что послужило предметом пристального внимания основоположника концепции "менеджмента качества" У.Э. Деминга (2007).

---

<sup>1</sup> <https://vz.ru/opinions/2019/2/15/964068.html>

Как измерить степень соответствия текущих значений показателей процессов их плановым значениям, оценить тем самым результативность или «гомеостатичность» организации? Естественно подсчитать количество пунктов одного или всех планов, намеченных к выполнению на данный момент (период), и количество выполненных пунктов. Отношение второго числа к первому можно считать степенью реализации запланированной деятельности, или результативностью (Шадрин 2005:189). Однако практическое применение этого подхода вызывает ряд вопросов. Допустим, что из десяти намеченных пунктов плана выполнены восемь: достаточно этого для результативности или нет? А 95 из 100? Более того, ведь каждый пункт, в свою очередь, может быть выполнен в той или иной степени. Например, срок сдачи жилого дома намечен на 23 февраля. Если акт приёмки в эксплуатацию действительно подписан не позже 23 февраля, то результативность сомнений не вызывает. А если 24-го? А 7 марта?

При ответах на эти вопросы можно выделить два подхода. Традиционный подход основан на так называемых допусках. Пусть длина некоторой детали должна равняться 25 миллиметрам. Тогда результативность считается достигнутой, а изделие качественным, если его реальная длина принадлежит диапазону, например, от 24.8 до 25.2 мм. Аналогично, дом считается сданным в срок, если акт приёмки подписан с задержкой, например, не более двух недель или месяца от указанной в договоре даты (и в этом случае штрафные санкции на застройщика не налагаются). Результативность считается достигнутой (а процесс качественным), если из десяти пунктов плана выполнено не менее девяти.

Основанный на допусках подход прост, понятен и действительно используется на практике в подавляющем большинстве случаев. Более того, он имеет под собой теоретическую основу, поскольку любой сложный процесс неизбежно находится под воздействием целого ряда случайных факторов, в силу чего его результат обязательно будет получен с некоторым отклонением от планового значения (не говоря уже о погрешности самих

измерений – этот вопрос здесь не рассматривается). Поэтому говорить можно только о жёсткости требований к допускам – будет ли это 0.2 мм, 0.02 или 0.002 мм, и т.д.

Тем не менее, Деминг подверг основанный на допусках подход суровой критике и предложил альтернативный подход к оценке качества. Конечно, выдающийся статистик Деминг не отрицает объективности возникновения погрешностей при осуществлении любого процесса. Более того, он настаивает на том, что операциональное определение «истинного» значения чего бы то ни было невозможно, так как наблюдаемое численное значение зависит от используемых определений и операций, которые будут различаться у разных экспертов в данной предметной области (Деминг 2007:261). Деминг и его комментатор Г. Нив приводят многостраничные обсуждения, говорящие о субъективности и тем самым, по их мнению, бессмысленности попыток установления численных значений допусков.

Таким образом, истинное значение недостижимо, а попытки установления диапазонов для его оценки некорректны. Выход из этого порочного круга Деминг видит в последовательном применении принципа постоянного улучшения - одного из фундаментальных принципов менеджмента качества. Если требуемое значение длины детали составляет 25 мм, то можно ли получить это значение для любой из изготовленных десяти тысяч или ста тысяч деталей? Видимо, нет, но Деминг призывает к этому стремиться, вместо того чтобы заведомо ограничиваться требованием попадания длины в диапазон 25 плюс-минус 0.2 мм, плюс-минус 0.02 мм или какой-либо ещё произвольно заданный. Тем самым плановое значение показателя трактуется как идеал, достижение которого возможно в пределе, а стремление к этому пределу позволяет мобилизовать физические и интеллектуальные возможности работников и реализовать принцип постоянного улучшения в полной мере. По мнению Деминга, основанный на допусках подход расхолаживает работников, устраняет из их деятельности «сверхзадачу». По сути, альтернативный подход Деминга усиливает

требования традиционного подхода, основанного на допусках. Сам Деминг заявляет: «Использование допусков ... – не ошибка. Просто этого недостаточно» или «Конечно же, мы не хотим нарушать допусков, но мы должны делать лучше, чем предписывается ими». Допуски становятся начальной точкой отсчёта, а не конечной целью деятельности (Нив 2007:157,161).

Подробный обзор подходов к формализации понятия гомеостаза приведён в (Угольницкий 2016). Авторский подход состоит в следующем.

В математическом моделировании набор существенных показателей функционирования динамической системы называется вектором состояния (фазовым вектором) и обозначается  $x(t) = (x_1(t), \dots, x_n(t))$ . Его компоненты (переменные состояния)  $x_i(t), i = 1, \dots, n$  - это значения выбранных показателей, характеризующих состояние системы в момент времени  $t$  с той точки зрения и степенью подробности, которые обусловлены целями и возможностями моделирования. Развитие динамической системы с вектором состояния  $x(t)$  при определённом наборе управляющих воздействий  $u(t)$  описывается уравнением динамики

$$\frac{dx}{dt} = f(x(t), u(t)). \quad (1.2.1)$$

Если известна идеальная траектория УДС  $x^*(t)$ , то область гомеостаза есть криволинейная трубка

$$S_\varepsilon(x^*(t)) = \{x(t) : \|x(t) - x^*(t)\| \leq \varepsilon, t \in [0, T]\}.$$

В частности, при  $x^*(t) \equiv x^*$  (есть идеальная точка) область гомеостаза представляет собой цилиндр

$$C_\varepsilon(x^*) = \{x(t) : \|x(t) - x^*\| \leq \varepsilon, t \in [0, T]\}.$$

Если по каждой координате отдельно задаются допуски относительно идеальных значений, то область гомеостаза - это множество

$$P_\varepsilon(x^*) = \{x(t) : |x_i(t) - x_i^*| \leq \varepsilon_i, i = 1, \dots, n, t \in [0, T]\}.$$

В условной статической постановке при наличии идеальной точки область гомеостаза есть шар, а при наличии допусков - параллелепипед.

Далее, пусть  $X^* \in \{S_\varepsilon, C_\varepsilon, P_\varepsilon\}$ , т.е. область гомеостаза описывается одной из предложенных параметризаций. Тогда условие гомеостаза можно сформулировать различными способами, например:

$$\forall t \in [0, T]: x(t) \in X^*, \quad (1.2.2)$$

$$\lim_{t \rightarrow \infty} x(t) = x^*. \quad (1.2.3)$$

Условие (1.2.2) можно интерпретировать как устойчивость по Лагранжу, а условие (1.2.3) - как предельную устойчивость идеальной точки. В совокупности при  $X^* = C_\varepsilon(x^*)$  эти условия дают асимптотическую устойчивость равновесия  $x^*$  (идеальной точки) по Ляпунову.

Таким образом, если развитие УДС моделируется уравнением (1.2.1), а совокупность плановых значений показателей её состояния – равновесием  $x^*$ , то традиционный подход формализуется как требование нейтральной устойчивости равновесия (или устойчивости по Лагранжу), а подход постоянного улучшения Деминга – как требование его асимптотической устойчивости по Ляпунову. Требование нейтральной устойчивости можно считать слабой формой условия гомеостаза активной системы, а требование асимптотической устойчивости – его сильной формой. Асимптотическая устойчивость (сильное условие гомеостаза) включает в себя нейтральную (слабое условие). Иными словами, если выполняется сильное условие гомеостаза, то выполняется и слабое, но обратное не всегда верно. Поэтому подход Деминга усиливает подход, основанный на допусках (табл. 1.2.1).

Содержательно гомеостаз как одно из условий УР означает, что компоненты вектора состояния принимают значения из плановых диапазонов. Применительно к организационным системам выполнение условий гомеостаза означает обеспечение требований к качеству.

## Условия гомеостаза и их организационная интерпретация

Условия гомеостаза	Организационная интерпретация	Математическая формализация
Слабая форма	Подход к качеству, основанный на допусках	$\forall t \in [0, T]: x(t) \in X^*$ (нейтральная устойчивость по Ляпунову или устойчивость по Лагранжу)
Сильная форма	Принцип постоянного улучшения качества Деминга	Дополнительно $\lim_{t \rightarrow \infty} x(t) = x^*$ (асимптотическая устойчивость по Ляпунову)

Формализация условий гомеостаза в терминах теории устойчивости имеет следующие преимущества. Во-первых, даются операциональные определения обоих подходов к достижению качества на основе формальной модели УДС. Во-вторых, в рамках этих определений однозначно устанавливается, что основанный на допусках подход образует первую часть принципа постоянного улучшения Деминга (поскольку нейтральная устойчивость есть необходимое условие асимптотической). Правда, у самого Деминга можно встретить и более категоричные высказывания: «Следование требованиям допусков – надёжный путь к упадку» (Нив 2007:168), но последовательный подход представляется более взвешенным. В-третьих, при использовании модели УДС (1.2.1) появляется возможность решения вопроса о нейтральной или асимптотической устойчивости на основе формальных критериев, разработанных в теории устойчивости (Угольницкий 2016).

Возможен и несколько иной подход к формализации условий гомеостаза. Потребуем, чтобы

$$|\{t \in [0, T]: x(t) \notin X^*\}| \leq c, \quad (1.2.4)$$

т.е. число выходов за пределы области гомеостаза не превышало заданной константы. В стохастической постановке это же условие можно записать в виде

$$P\{|\{t \in [0, T] : x(t) \notin X^*\}| > c\} \leq \varepsilon. \quad (1.2.5)$$

Главная проблема здесь состоит в том, что обеспечение требований гомеостаза в любой форме должно осуществляться активными агентами, и если они в этом не заинтересованы, то любые стратегии перехода к УР останутся лишь декларациями. Поэтому ключевую роль в поддержании УР играет мотивированность активных агентов.

Активные системы включают большое количество способных к стратегическому поведению агентов с несовпадающими интересами. Поэтому ключевой проблемой управления УР активных систем выступает согласование этих интересов в форме некоторого компромисса, в той или иной степени устраивающего всех агентов при выполнении условий гомеостаза. В идеале согласование интересов должно приводить к кооперации, обеспечивающей совместные действия агентов во имя общесистемных целей УР с получением синергетического эффекта.

История, теория и модели согласования интересов подробно анализируются в (Угольницкий 2016). Изложим текущее состояние подхода.

Смысл *мотивированности* агентов состоит в том, что они должны быть заинтересованы в выполнении условий гомеостаза. В исходной ситуации большое количество активных агентов воздействует на УДС с уравнением динамики (1.2.1) в своих интересах (рис. 1.2.1).

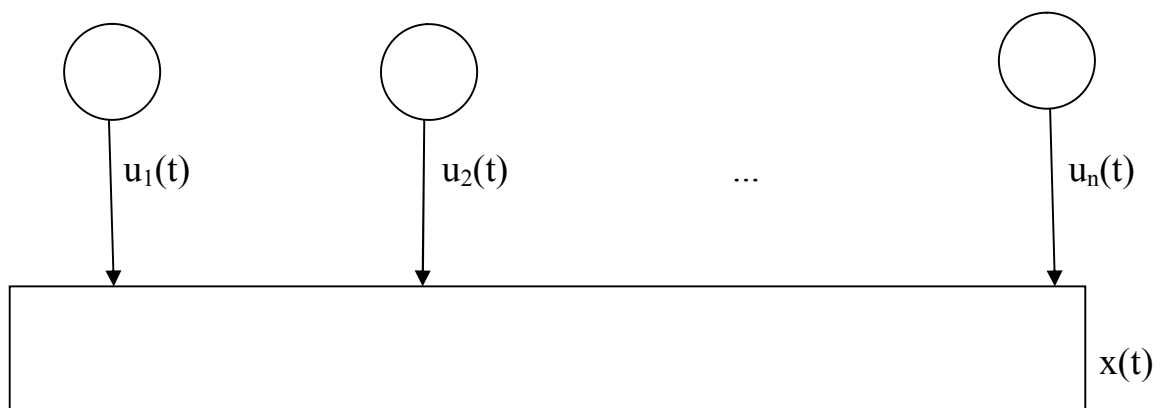


Рис. 1.2.1. Динамическая система, управляемая активными агентами

Важно подчеркнуть, что интересы агентов



$$J_i(u_i(\cdot), u_{-i}(\cdot)) \rightarrow \max, u_i(t) \in U_i, i = 1, \dots, n, t \in [0, T] \quad (1.2.6)$$

в общем случае не имеют никакого отношения к условиям гомеостаза типа (1.2.2), выполнение которых возможно при выборе управлений из множества

$$U^* = \{u(t) : x(t) \in X^*, t \in [0, T]\}. \quad (1.2.7)$$

Поэтому для создания мотивированности требуется управление.

Традиционно выделяются три группы методов управления:

1) принуждение, при котором субъект заставляет объекта способствовать достижению своей цели, не принимая во внимание цели и интересы объекта;

2) побуждение, при котором субъект создает объекту такие условия, что последнему выгодно способствовать достижению цели субъекта и невыгодно обратное;

3) убеждение, при котором субъектно-объектное взаимодействие организуется так, что объект добровольно и заинтересованно стремится к достижению единой цели совместно с субъектом управления, то есть их отношения приобретают субъектно-субъектный характер (Ледяев 2000).

Таким образом, существуют два качественно различных пути к мотивированности. В первом случае субъект управления мотивированностью может вообще отсутствовать (по крайней мере явно), а активные агенты добровольно и осознанно принимают условия гомеостаза (убеждение). При убеждении нет ни угрозы негативных санкций, ни внешних мотивирующих стимулов: агент сам принимает решение согласиться с предлагаемыми ему аргументами (Ледяев 2000:291).

Здесь возможны два варианта: независимые действия агентов или их кооперация. При независимом принятии решений агентами формализацией приемлемого компромисса их интересов естественно считать множество равновесий Нэша NE в дифференциальной игре (1.2.1), (1.2.6). Тогда мотивированность объективно возможна, если  $NE \cap U^* \neq \emptyset$ , т.е. существует набор управлений агентов, одновременно обеспечивающий приемлемый для них компромисс интересов и выполнение условий гомеостаза.

При кооперации агенты объединяются в максимальную коалицию, действующую как единый игрок и совместно максимизирующую суммарный функционал выигрыша:

$$J(u(\cdot)) = \sum_{i=1}^n J_i(u(\cdot)) \rightarrow \max \quad (1.2.8)$$

при ограничениях

$$u(t) \in U, u(t) = (u_1(t), \dots, u_n(t)), U = U_1 \times \dots \times U_n, t \in [0, T] . \quad (1.2.9)$$

Здесь интересы агентов реализуются путём решения задачи оптимального управления для максимальной коалиции, т.е. выбором элемента из множества

$$\bar{U} = \text{Arg max}_{u(\cdot) \in U} J(u(\cdot)) . \quad (1.2.10)$$

Такая ситуация показана на рис. 1.2.2. Теперь мотивированность объективно возможна, если выполняется условие  $\bar{U} \cap U^* \neq \emptyset$ , т.е. максимизация суммарного выигрыша множества агентов совместима с требованиями гомеостаза.

Очевидно, что идеальная предпосылка для мотивированности через убеждение - это выполнение условий  $NE \subseteq U^*$  при независимых агентах или  $\bar{U} \subseteq U^*$  при их кооперации. Тогда альтернативы гомеостазу просто нет.

Если условие  $NE \cap U^* \neq \emptyset$  при независимых агентах или  $\bar{U} \cap U^* \neq \emptyset$  при их кооперации нарушается, то достичь мотивированности к гомеостазу через убеждение невозможно. Не исключено также, что указанные пересечения не пусты, но всё же в силу каких-то причин добровольного принятия условий гомеостаза агентами отдельно или в совокупности не происходит.

Тогда остаётся второй, намного более распространённый в жизни путь обеспечения мотивированности - иерархическое управление, осуществляемое в форме принуждения или побуждения. В этом случае обязательно вводится специальный субъект управления (Центр), отвечающий за выполнение условий гомеостаза. Соответствующая ситуация показана на рис. 1.2.3.

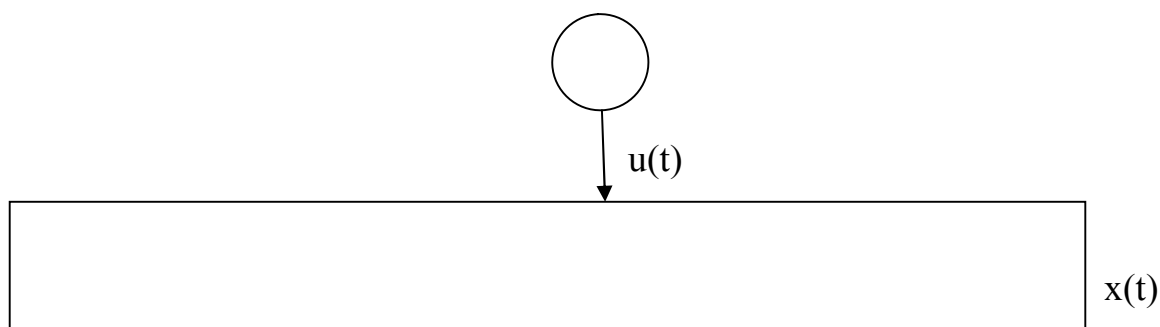


Рис. 1.2.2. Динамическая система, управляемая максимальной коалицией агентов

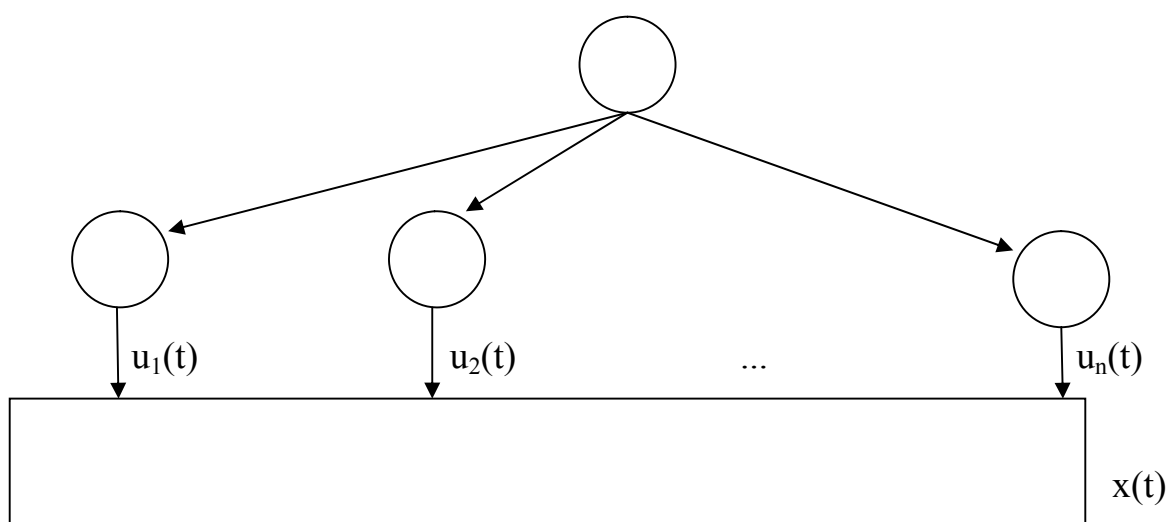


Рис. 1.2.3. Иерархически управляемая динамическая система

Сначала предположим, что агенты кооперируются и создают максимальную коалицию. Тогда возникает базовая модель иерархической игры "Центр-Агент" (Принципал-Агент, Ведущий-Ведомый) между субъектами верхнего и нижнего уровней управления.

В случае принуждения субъект верхнего уровня ставит субъекта нижнего уровня в такие условия, в которых последний вынужден применять желаемую для первого стратегию. Такое воздействие носит насильственный характер и осуществляется на административно-законодательной основе. Интересы субъекта нижнего уровня при этом совершенно не учитываются, т.е. отношения между субъектами имеют субъектно-объектную природу. Если описать отношения между субъектами как иерархическую игру

Ведущего и Ведомого, то принуждение означает воздействие Ведущего на множество допустимых стратегий Ведомого.

Принуждение имеет место при существенном несовпадении интересов субъекта и объекта; в этом случае субъект обеспечивает подчинение объекта путём угрозы применения негативных санкций (Ледяев 2000). Многие авторы дают сходные определения принуждения: "способность А заставить Б действовать в соответствии со своими намерениями и вопреки желаниям Б путем негативного воздействия на Б для достижения его повиновения или путем угрозы негативного воздействия на Б, если тот не повинуется" (De Crespigny 1968:196); "принудительная власть О над П обусловлена ожиданием П наказания со стороны О в случае неповиновения" (French and Raven 1959:157). Таким образом, принуждение характеризуется следующими чертами: 1) А и Б имеют существенно различные (даже антагонистические) предпочтения по данному поводу, поэтому Б не заинтересован в выполнении команды А; 2) однако подчинение является для Б меньшим злом, чем неподчинение, поскольку в случае реализации А негативных санкций Б потеряет больше; 3) А предпочитает не приводить свою угрозу в исполнение, хотя и способен сделать это (Ледяев 2000:286).

Понятно, что принуждение - не лучший путь к мотивированности, поскольку тут она обеспечивается насильственно, под угрозой санкций. Тем не менее, при слабой заинтересованности объекта в выполнении требований гомеостаза этот метод может оказаться необходимым.

Намного естественнее создавать мотивированность через побуждение, где главную роль играет вознаграждение, которое объект получает от субъекта в обмен на сотрудничество. Де Креспиньи определяет побуждение как "способность А заставить Б действовать в соответствии со своими интенциями (так как он иначе не стал бы действовать) путем обеспечения Б тем, что его [Б] привлекает с целью достижения его повиновения, или выполняя обещание, когда Б повинуется" (De Crespigny 1968:198). Как и при принуждении, при побуждении объект нуждается во внешнем стимуле, без

которого он не стал бы выполнять команды субъекта. Однако если при принуждении подчинение достигается угрозой применения негативных санкций (наказания), то побуждение означает обещание вознаграждения (поощрения). Хотя формально эти методы воздействия можно отождествить, позитивные и негативные санкции всё же различаются по ряду аспектов.

Во-первых, принуждение основано на силовых ресурсах и обычно ассоциируется с насилием, применяемым властью (руководством), в то время как побуждение использует утилитарные ресурсы и носит экономический характер. Во-вторых, существенно различается психологическая реакция объекта: реакцией на принуждение обычно являются страх, тревога и сопротивление, а на побуждение – надежда и сотрудничество. Поэтому побуждение лучше обеспечивает легитимацию выдвигаемых требований, но зато требует большей затраты ресурсов (Ледяев 2000:288-290). Точную характеристику побуждению дал А. Смит: «... человек постоянно нуждается в помощи своих ближних, но тщетно было бы ожидать её лишь от их расположения. Он скорее достигнет своей цели, если обратится к их эгоизму и сумеет показать им, что в их собственных интересах сделать для него то, что он требует от них» (Смит 1962:27-28).

В случае побуждения Ведущий создает условия, в которых Ведомому более выгодно применять желаемую для Ведущего стратегию, нежели не делать этого. Этот метод имеет природу экономического стимулирования и обычно реализуется посредством механизма штрафов и поощрений с обратной связью. Здесь частично учитываются интересы Ведомого. При теоретико-игровом моделировании побуждение означает воздействие Ведущего на функцию выигрыша Ведомого. Общая характеристика методов иерархического управления приведена в таблице 1.2.1 (Угольницкий 2016).

Математическая формализация равновесий принуждения и побуждения в программных стратегиях для различных информационных регламентов в базовой модели "Ведущий-Ведомый" дана в (Угольницкий 2016).

## Характеристика методов иерархического управления

	Принуждение	Побуждение	Убеждение
Общее описание метода	Насильственное обеспечение выбора Ведомым стратегии, желаемой для Ведущего	Желаемая для Ведущего стратегия более выгодна Ведомому, чем нежелательные	Ведомый добровольно и осознанно выбирает стратегию, желаемую для Ведущего
Природа воздействия	Административно-законодательное	Экономическое	Социально-психологическое
Тип отношений	Субъектно-объектные	Субъектно-объектные с частичным учетом интересов Ведомого	Субъектно-субъектные
Математическая формализация	Воздействие Ведущего на множество допустимых стратегий Ведомого	Воздействие Ведущего на функцию выигрыша Ведомого	Переход Ведущего и Ведомого к кооперации для совместной максимизации суммарной функции выигрыша

Теперь вернёмся к более общему случаю нескольких агентов, воздействие которых на УДС контролируется Центром для обеспечения гомеостаза в древовидной системе управления (рис. 1.2.3). При математической формализации это даёт иерархическую дифференциальную игру ведущего (leader) с несколькими ведомыми (followers). Для построения классификации информационных структур в таких играх удобно брать три признака, характеризующие стратегию ведущего (в частности, способ обеспечения УР):

1) отсутствие/наличие обратной связи стратегии ведущего по состоянию УДС. Этот признак принимает два основных значения:

*программные стратегии* (open-loop, OL), которые зависят только от текущего момента времени  $t$ , и *позиционные стратегии* (closed-loop, CL), зависящие от игровой позиции  $(t, x(t))$  (Basar and Olsder 1999);

2) отсутствие/наличие обратной связи стратегии ведущего по стратегиям ведомых. В первом случае возникают *прямые игры Штакельберга* (игры Гермейера типа  $\Gamma_1$ ), во втором - *обратные игры Штакельберга* (игры Гермейера типа  $\Gamma_2$ ) (Кононенко 1977; Горелик и др. 1991; Горелов и Кононенко 2014а,б, 2015; Olsder 2009);

3) методы иерархического управления (рассмотренные выше принуждение и побуждение).

В свою очередь, на нижнем уровне возможен один из трёх режимов поведения (которые надо учитывать при реализации УР):

(а) изоляция, когда ведомые действуют независимо и приходят в своей игре в нормальной форме к равновесию Нэша;

(б) кооперация, при которой они объединяют ресурсы в составе максимальной коалиции и путём совместных действий максимизируют суммарный функционал выигрыша;

(в) сотрудничество, означающее добровольную максимизацию ведомыми функционала выигрыша ведущего (убеждение).

Рассмотрим для определённости теоретико-игровую модель вида

$$J = \int_0^T e^{-\rho t} g(y(t)) dt + e^{-\rho T} G(x(T)) \rightarrow \max \quad (1.2.1)$$

$$p(t) \in P; q(t) \in Q; \quad (1.2.2)$$

$$J_i = \int_0^T e^{-\rho t} g_i(y(t)) dt + e^{-\rho T} G_i(x(T)) \rightarrow \max \quad (1.2.3)$$

$$u_i(t) \in U_i(q_i(t)), i \in N; \quad (1.2.4)$$

$$\dot{x} = f(x(t), u(t)), x(0) = x_0; \quad (1.2.5)$$

$$x(t) \in X^*, 0 \leq t \leq T. \quad (1.2.6)$$

Здесь  $J$  и  $J_i$ ,  $g$  и  $g_i$  - интегральные и текущие функционалы выигрыша ведущего и ведомых соответственно;  $T$  - продолжительность игры (которая может быть бесконечной, тогда терминальный член не используется);  $\rho$  - коэффициент дисконтирования;  $y(t) = (q(t), p(t), u(t), x(t))$ ;  $q(t)$  - вектор управляющих переменных принуждения;  $p(t)$  - вектор управляющих переменных побуждения;  $u(t) = (u_1(t), \dots, u_n(t))$  - вектор управлений ведомых (той же размерности, что и  $q(t)$ );  $N = \{1, \dots, n\}$  - множество ведомых;  $x(t)$  - вектор состояния УДС;  $X^*$  - множество состояний, удовлетворяющих требованиям гомеостаза; соотношения (1.2.5) описывают динамику управляемой системы; условия (1.2.6) формулируют требования гомеостаза. Предполагается, что  $Q, P, U$  - компактные подмножества соответствующих метрических пространств;  $J$  и  $J_i$  непрерывны по своим аргументам; вектор-функция  $f$  из (1.2.5) удовлетворяет условиям Каратеодори. Таким образом, (1.2.1) - (1.2.6) - иерархическая дифференциальная игра с фазовыми ограничениями.

Для объяснения предложенной классификации используем следующие две таблицы (пояснения в тексте).

Таблица 1.2.2

Базовые информационные структуры в иерархических дифференциальных играх

	Без обратной связи по управлениям ведомых (прямые игры Штакельберга)	С обратной связью по управлениям ведомых (обратные игры Штакельберга)
Без обратной связи по состоянию УДС (игры в программных стратегиях)	Прямая игра Штакельберга в программных стратегиях $\Gamma_{1t}$	Обратная игра Штакельберга в программных стратегиях $\Gamma_{2t}$
С обратной связью по состоянию УДС (игры в позиционных стратегиях)	Прямая игра Штакельберга в позиционных стратегиях $\Gamma_{1x}$	Обратная игра Штакельберга в позиционных стратегиях $\Gamma_{2x}$



В таблице 1.2.1 показаны типы стратегий ведущего с использованием обозначений из работы (Кононенко, 1977). В игре  $\Gamma_{1t}$  ведущий в процессе игры не получает информации ни о действиях ведомых, ни о состоянии УДС. Поэтому ведущий выбирает и сообщает ведомым программную стратегию  $z(t), 0 \leq t \leq T$ . В игре  $\Gamma_{2t}$  ведущий до начала игры знает о выбранных ведомыми программных управлениях, поэтому его стратегия представляет собой оператор  $z(t, u(t))$ , ставящий в соответствие каждому набору функций  $u(t)$  программную стратегию  $z(t)$ . В игре  $\Gamma_{1x}$  ведущий выбирает и сообщает ведомым набор позиционных стратегий  $z(t, x(t))$ . Наконец, в игре  $\Gamma_{2x}$  ведущий до начала игры узнает выбранные ведомыми позиционные стратегии  $u(t, x(t))$ , поэтому он выбирает и сообщает им позиционную зависимость  $z(t, x(t), u(t, x(t)))$ .

Таким образом, в играх  $\Gamma_{2t}$  и  $\Gamma_{2x}$  присутствует обратная связь по управлениям ведомых, а в играх  $\Gamma_{1x}$  и  $\Gamma_{2x}$  - по состоянию УДС. Регламенты  $\Gamma_{1t}$  и  $\Gamma_{2t}$  определяют прямые игры Штакельберга в программных и позиционных стратегиях соответственно, регламенты  $\Gamma_{1x}$  и  $\Gamma_{2x}$  - аналогичные обратные игры Штакельберга.

Таблица 1.2.3

Максимальные гарантированные выигрыши ведущего в играх с несколькими ведомыми для различных информационных структур

Ведущий	Бездействие (p-const, q-const)	Принуждение (p-const)		Побуждение (q-const)	
		$\Gamma_{1t}, \Gamma_{1x}$ (ST)	$\Gamma_{2t}, \Gamma_{2x}$ (GER)	$\Gamma_{1t}, \Gamma_{1x}$ (ST)	$\Gamma_{2t}, \Gamma_{2x}$ (GER)
Изоляция (NE)	$J_{NE}^0$	$J_{NE}^{COMP-ST}$	$J_{NE}^{COMP-GER}$	$J_{NE}^{IMP-ST}$	$J_{NE}^{IMP-GER}$
Кооперация (C)	$J_C^0$	$J_C^{COMP-ST}$	$J_C^{COMP-GER}$	$J_C^{IMP-ST}$	$J_C^{IMP-GER}$
Сотрудничество (max)	$J_{max}^0$	$J_{max}^{COMP-ST}$	$J_{max}^{COMP-GER}$	$J_{max}^{IMP-ST}$	$J_{max}^{IMP-GER}$

Рассмотрим более подробно таблицу 1.2.3. В иерархических дифференциальных играх в качестве принципа оптимальности принимается выбор ведущим максимальной гарантирующей стратегии с учётом оптимальной реакции ведомых (принцип гарантированного результата Гермейера). В таблице 1.2.3 показаны соответствующие гарантированные выигрыши ведущего для перечисленных информационных структур. В случае изоляции предполагается, что оптимальные ответы ведомых описываются множеством равновесий Нэша  $NE$  в их игре в нормальной форме. При кооперации оптимальная реакция ведомых - это множество  $C$  точек максимума их суммарного функционала выигрыша. Наконец, при сотрудничестве оптимальная реакция ведомых отождествляется с максимизацией функционала выигрыша ведущего.

В случае принуждения ведущий выбирает вектор стратегий  $q$  ( $p - const$ ) в режимах  $\Gamma_{1t}, \Gamma_{1x}$  (прямые игры Штакельберга, ST) или  $\Gamma_{2t}, \Gamma_{2x}$  (обратные игры Штакельберга, GER). Стратегии могут быть программными (OL) (регламенты  $\Gamma_{1t}, \Gamma_{2t}$ ) или позиционными (CL) (регламенты  $\Gamma_{1x}, \Gamma_{2x}$ ). Аналогично, при побуждении ведущий выбирает вектор стратегий  $p$  ( $q - const$ ) для тех же регламентов. В вырожденном случае бездействия  $p$  и  $q$  постоянны (управление ведущего отсутствует).

Теперь можно строго определить выигрыши ведущего для указанных информационных регламентов (т.е. решения соответствующих игр), основываясь на принципе гарантированного результата. При бездействии  $J_{NE}^0 = \inf_{u \in NE} J(u)$ ,  $J_C^0 = \inf_{u \in C} J(u)$ ,  $J_{\max}^0 = \inf_{u \in U} J(u)$ . В случае принуждения для прямых игр Штакельберга получаем

$$J_{NE}^{COMP-ST} = \sup_{q \in Q} \inf_{u \in NE(q)} J(q, u), \quad (1.2.7)$$

$$J_C^{COMP-ST} = \sup_{q \in Q} \inf_{u \in C(q)} J(q, u), \quad (1.2.8)$$

$$J_{\max}^{COMP-ST} = \sup_{q \in Q} \sup_{u \in U(q)} J(q, u), \quad (1.2.9)$$

а для обратных игр Штакельберга

$$J_{NE}^{COMP-GER} = \sup_{\tilde{q} \in \tilde{Q}} \inf_{u \in NE(\tilde{q})} J(\pi(\tilde{q}, u)), \quad (1.2.10)$$

$$J_C^{COMP-GER} = \sup_{\tilde{q} \in \tilde{Q}} \inf_{u \in C(\tilde{q})} J(\pi(\tilde{q}, u)), \quad (1.2.11)$$

$$J_{\max}^{COMP-GER} = \sup_{\tilde{q} \in \tilde{Q}} \sup_{u \in U(\tilde{q})} J(\pi(\tilde{q}, u)), \quad (1.2.12)$$

$$\tilde{Q} = \{\tilde{q} : U \rightarrow Q\}, \pi : \tilde{Q} \times U \rightarrow Q \times U$$

Решения для побуждения строятся аналогично. Удобно также определить глобально-максимальный выигрыш ведущего

$$J_{\max}^* = \sup_{p \in P} \sup_{q \in Q} \sup_{u \in U(q)} J(p, q, u) \quad (1.2.13)$$

Для оценки сравнительной эффективности методов иерархического управления в различных информационных структурах построим набор показателей системной согласованности. Для каждого столбца таблицы 1.2.3 определим показатели:

$$K_{NE}^{(\cdot)} = J_{NE}^{(\cdot)} / J_{\max}^{(\cdot)}; \quad K_C^{(\cdot)} = J_C^{(\cdot)} / J_{\max}^{(\cdot)}; \quad K_{NE}^{(\cdot)*} = J_{NE}^{(\cdot)} / J_{\max}^*; \quad K_C^{(\cdot)*} = J_C^{(\cdot)} / J_{\max}^*, \quad (1.2.14)$$

где верхний индекс  $(\cdot)$  обозначает строку. Заметим, что показатель  $K_{NE}^0 = PA$  определяет *цену анархии* в смысле Papadimitriou (2001) для динамической постановки задачи. Конечно,  $K_{\max}^{(\cdot)} = 1$ .

Итак, можно сформулировать следующие выводы по управлению УР активных систем. Наряду с требованиями к состоянию УДС (условиями гомеостаза) необходимо выдвигать требования к воздействию на неё (условия мотивированности), поскольку условия гомеостаза не выполняются сами собой и нуждаются в активном заинтересованном субъекте реализации. В отсутствие такого субъекта условия гомеостаза обречены оставаться абстрактными и декларативными. Если Центр отсутствует или бескорыстен (имеет единственной целью обеспечение гомеостаза), то возможны следующие три ситуации. Обозначим  $OPT \in \{NE, C, \dots\}$  - некоторый принцип

оптимальности взаимодействия агентов,  $U^*$  - множество управлений агентов, обеспечивающих гомеостаз.

1.  $OPT \subseteq U^*$ . Задача управления УР решается гарантированно, интересы агентов полностью совместимы с требованиями гомеостаза.
2.  $OPT \cap U^* \neq \emptyset$ . Задача управления УР может быть решена, но интересы агентов лишь частично совместимы с требованиями гомеостаза, для его обеспечения нужны дополнительные усилия.
3.  $OPT \cap U^* = \emptyset$ . Задача управления УР неразрешима, поскольку агенты не станут действовать вопреки своим интересам. В этом случае придётся либо трансформировать интересы агентов, либо соглашаться с ослаблением требований гомеостаза.

Если Центр наряду с обеспечением гомеостаза имеет собственные интересы, то его мотивированность следует учитывать дополнительно с помощью рассмотренных выше решений иерархических дифференциальных игр с Центром в роли ведущего. Для измерения мотивированности Центра удобно использовать введённые показатели системной согласованности. Итоговая характеристика вариантов решения задачи управления УР активной системы при наличии Центра приведена в таблице 1.2.4.

Полученное с учётом интересов всех агентов решение игры для своей практической реализуемости должно обладать свойством *динамической устойчивости*, обеспечивающим невыгодность отклонения любого агента от достигнутого соглашения в течение всего времени его осуществления. Слабая динамическая устойчивость (time consistency) означает, что решение исходной игры остается оптимальным в любой подыгре вдоль оптимальной траектории, а сильная динамическая устойчивость (subgame perfectness) требует сохранения оптимальности решения вдоль любой допустимой траектории. Надо отметить, что решения иерархических дифференциальных игр обычно не удовлетворяют свойству динамической устойчивости (даже слабой), поэтому для его обеспечения нужно применять специальные методы

регуляризации (в кооперативных играх - процедуры распределения дележа)  
(Петросян и др. 2012).

Таблица 1.2.4

Учёт интересов Центра и активных агентов при обеспечении гомеостаза

Интересы агентов	Центр может обеспечить	Центр может обеспечить
Интересы Центра	$OPT \cap U^* \neq \emptyset$	$OPT \subseteq U^*$
Индекс системной согласованности $K_{NE}^{(\cdot)*} \geq K_{\min}$	Возможность решения задачи управления УР при достаточном учёте интересов Центра	Гарантированное решение задачи управления УР при достаточном учёте интересов Центра
Индекс системной согласованности $K_{NE}^{(\cdot)*} = 1$	Возможность решения задачи управления УР при полном учёте интересов Центра	Гарантированное решение задачи управления УР при полном учёте интересов Центра

### 1.3. Методика решения задачи управления устойчивым развитием активных систем

Суммируем и дополним приведённые выше результаты, используя статью (Угольницкий 2019). Под *активной системой* понимается структурированная совокупность активных элементов. Примерами активных систем служат организационные, экономические, социальные и иные сложные динамические системы с участием людей.

*Активный элемент (агент)* - субъект, имеющий собственные интересы (цели) и возможности их достижения. Активный агент способен к целеполаганию и стратегическому поведению, т.е. выбору действий, направленных на достижение цели, что может быть связано с сознательным искажением передаваемой другим активным агентам информации (манипуляцией). Примеры активных агентов с учетом принципа системной относительности: предприниматель, фирма, избиратель, партия, муниципальное образование, страна, член семьи и т.д.

Активные агенты действуют в соответствии с собственными интересами. Наряду с этим, существуют цели и требования, важные для активной системы в целом, но совсем не обязательно совпадающие с интересами отдельных агентов. Эти условия назовём *гомеостазом* активной системы. Например, расположенное рядом с водным объектом предприятие заинтересовано в максимизации прибыли, а не в соблюдении экологических требований к состоянию этого объекта. Сотрудник организации в большей степени заинтересован в повышении зарплаты и экономии усилий, чем в достижении долгосрочных целей развития этой организации. Поэтому обычно возникает *Центр (принципал, ведущий)* - выделенный активный элемент, выражающий интересы активной системы в целом и отвечающий за выполнение условий гомеостаза. Примеры: правительство (парламент, президент) страны, руководство организации, администрация региона, партийный лидер, глава семьи и т.д.

Анализ проблемы управления УР активных систем позволяет сформулировать следующие эмпирические принципы.

1. *Принцип экономической рациональности*: интересы каждого активного агента, в т.ч. Центра, целиком и полностью характеризуются стремлением к максимизации выигрыша с учетом имеющихся ограничений.

Надо отметить, что ряд авторов (Д. Канеман и А. Тверски, Г. Саймон и др.) развивают более общую концепцию ограниченной рациональности, учитывающую неполноту информации принимающего решения субъекта и ряд других факторов.

2. *Принцип ответственности Центра*: главная цель Центра заключается в обеспечении условий гомеостаза, при этом он может иметь дополнительные частные интересы.

3. *Принцип необходимости управления*: в общем случае интересы отдельных агентов не совпадают с требованиями гомеостаза.

Основная трудность обеспечения гомеостаза состоит именно в том, что активные агенты далеко не всегда заинтересованы в выполнении его условий, в силу чего формулирование требований гомеостаза в официальных документах по УР остаётся простой декларацией. Поэтому устойчивое развитие наряду с условиями гомеостаза обязательно должно включать учет интересов обеспечивающих его активных агентов (системную согласованность, мотивированность). Совместное выполнение условий гомеостаза и системной согласованности (мотивированности агентов) и означает *устойчивое развитие* активной системы.

Идеальным способом достижения УР выступает *убеждение* - добровольная и осознанная интериоризация требований гомеостаза всеми агентами. Примерами убеждения могут служить экологическое сознание, социально ответственный бизнес, благотворительность, волонтерство, разоружение и т.п.

Однако, на практике убеждение пока остается в основном делом будущего, поэтому приходится использовать иерархическое управление

Центра путем принуждения и/или побуждения. *Принуждение* включает административно-законодательные механизмы управления, явно ограничивающие множества возможных действий агентов так, чтобы они не могли нарушить гомеостаз. *Побуждение* (стимулирование) состоит в использовании экономических механизмов, делающих желательные для Центра гомеостатические действия более выгодными для агентов.

Приведём постановки задач управления УР с использованием методов принуждения, побуждения и убеждения для активных систем различной структуры.

### *Независимое поведение равноправных агентов*

В исходной модели множество равноправных агентов одновременно и независимо друг от друга воздействует на некоторый динамический объект в своих интересах. Вместе с тем, существуют объективные требования (условия гомеостаза), которым должно удовлетворять состояние объекта с точки зрения активной системы в целом. Такая ситуация формализуется как дифференциальная игра нескольких лиц в нормальной форме вида

$$J_i(u(\cdot)) = \int_0^{\infty} e^{-\rho t} g_i(x(t), u(t), t) dt \rightarrow \max \quad (1.3.1)$$

$$u_i(t) \in U_i, \quad i \in N; \quad (1.3.2)$$

$$\dot{x} = f(x(t), u(t), t), \quad x(0) = x_0. \quad (1.3.3)$$

Здесь  $N = \{1, \dots, n\}$  - множество активных агентов;  $g_i, J_i$  - мгновенная функция и интегральный функционал выигрыша агента  $i \in N$ ;  $u_i(t)$  - управляющее действие (управление) агента  $i \in N$  в момент  $t$ ;  $U_i$  - множество допустимых управлений агента  $i \in N$ ;  $u(t) = (u_1(t), \dots, u_n(t))$  - набор управляющих воздействий агентов;  $U = U_1 \times \dots \times U_n$ ;  $x(t)$  - значение переменной состояния в момент  $t$ ;  $x_0$  - начальное состояние;  $f$  - заданная функция, определяющая



изменение состояния во времени;  $\rho \in [0,1]$  - коэффициент дисконтирования. В общем случае фазовая переменная и управление могут быть векторными величинами. Бесконечный период времени более адекватен для концепции устойчивого развития, хотя возможны и постановки при  $T < \infty$ . Предполагаются выполненными следующие условия (Горелик, Кононенко 1982): множества  $U_i$  компактны в соответствующих векторных пространствах; вектор-функция  $f$  непрерывна по  $t$  и непрерывно дифференцируема по остальным параметрам; существует константа  $\kappa > 0$  такая, что при любых  $x, t \geq 0, u \in U$  имеет место неравенство  $\|f(x, u, t)\| \leq \kappa(1 + \|x\|)$ ; функции  $g_i$  непрерывны по всем аргументам; классам допустимых позиционных стратегий принадлежат кусочно-непрерывные по  $t$  и гладкие по  $x$  при всех  $t \geq 0$  функции, удовлетворяющие ограничению (1.3.2). Условие гомеостаза можно записать в виде

$$x(t) \in X^* \subseteq R^n, t \in [0, \infty). \quad (1.3.4)$$

Следует подчеркнуть, что в исходной постановке условие (1.3.4) внешнее для множества агентов, поэтому изначально модель нельзя рассматривать как дифференциальную игру (1.3.1)-(1.3.3) с дополнительными фазовыми ограничениями (1.3.4).

Будем считать решением игры (1.3.1)-(1.3.3) множество равновесий Нэша  $NE$ , считая  $NE \neq \emptyset$  (иначе нечего обсуждать). В частности,  $NE \neq \emptyset$ , если  $U_i$  выпуклы и компактны, а  $g_i(u_i, u_{-i})$  вогнуты по  $u_i$  на  $U_i$  для всех  $u_{-i} \in U_{-i}$ ,  $i \in N$ .

Обозначим  $U^* = \{u \in U : \forall t x(t) \in X^*\}$  - множество управлений, обеспечивающих гомеостаз, и вновь предположим, что  $U^* \neq \emptyset$ , иначе рассмотрение теряет смысл. Если  $U_{DG} = NE \cap U^* \neq \emptyset$ , то задача управления устойчивым развитием (1.3.1)-(1.3.4) имеет решение. Управления  $u \in U_{DG}$  обеспечивают гомеостаз, при этом интересы агентов также учитываются через равновесие Нэша в их игре (мотивированность).

Для более детального анализа проблемы согласования интересов введем функционал общественного благосостояния

$$J(u(\cdot)) = \sum_{i \in N} J_i(u(\cdot)).$$

Тогда, если выполняется условие

$$\forall u^{NE} \in NE \setminus U_{DG} \exists u \in U_{DG} : J(u) \geq J(u^{NE}), \quad (1.3.5)$$

то требования гомеостаза (1.3.4) полностью совместимы с общественным благосостоянием. Если же верно противоположное утверждение

$$\exists u^{NE} \in NE \setminus U_{DG} \forall u \in U_{DG} : J(u) < J(u^{NE}), \quad (1.3.6)$$

то требования гомеостаза не полностью совместимы с общественным благосостоянием. Далее, если справедливо утверждение

$$\forall i \in N \forall u^{NE} \in NE \setminus U_{DG} \exists u \in U_{DG} : J_i(u) \geq J_i(u^{NE}), \quad (1.3.7)$$

то требования гомеостаза полностью совместимы с частными интересами всех агентов, а если верно противоположное утверждение

$$\exists i \in N \exists u^{NE} \in NE \setminus U_{DG} \forall u \in U_{DG} : J_i(u) < J_i(u^{NE}), \quad (1.3.8)$$

то требования гомеостаза несовместимы с частными интересами некоторых агентов. Очевидно, что (1.3.7)  $\rightarrow$  (1.3.5), (1.3.6)  $\rightarrow$  (1.3.8).

Если  $U_{DG} = NE \cap U^* = \emptyset$ , то задача управления УР (1.3.1)-(1.3.4) в исходной постановке неразрешима.

Если  $U_{DG} \neq \emptyset$ , но справедливо (1.3.6) или хотя бы (1.3.8), то задача разрешима, но общественное благосостояние в целом или по крайней мере интересы некоторых агентов при выполнении условий гомеостаза пострадают. В этих случаях возможны два способа изменения постановки задачи, рассматриваемые далее.

### Кооперация агентов и метод убеждения

Независимые агенты могут вступить в кооперацию, образовать коалицию и совместно максимизировать суммарный функционал выигрыша (общественное благосостояние) по всем управляющим переменным. Тогда вместо дифференциальной игры в нормальной форме (1.3.1)-(1.3.3) возникает задача оптимального управления

$$J(u(\cdot)) \rightarrow \max \quad (1.3.9)$$

с ограничениями (1.3.2)-(1.3.3).

Введем множество  $\bar{U} = \text{Arg max}_{u(\cdot) \in U} J(u(\cdot))$ . Если  $U_{oc} = U^* \cap \bar{U} \neq \emptyset$ , то задача управления УР активных систем в кооперативной постановке имеет решение. Заметим, что здесь всегда  $U_{oc} \neq \emptyset \Rightarrow \forall \bar{u} \in \bar{U} \setminus U_{oc} \exists u \in U_{oc} : J(u) = J(\bar{u})$ , т.е. условие  $U_{oc} \neq \emptyset$  обеспечивает согласованность гомеостаза с общественным благосостоянием.

Метод убеждения заключается в том, что все агенты добровольно и осознанно соглашаются обеспечивать условие гомеостаза (1.3.4). В этом и только в этом случае задача управления УР активных систем при независимых агентах описывается дифференциальной игрой в нормальной форме с фазовыми ограничениями (1.3.1)-(1.3.4).

Наконец, возможен комплексный вариант, при котором агенты кооперируются и соглашаются обеспечивать условие гомеостаза (1.3.4). Тогда для максимальной коалиции агентов возникает задача оптимального управления с фазовыми ограничениями (1.3.9), (1.3.2)-(1.3.4).

При кооперации возникает дополнительная задача распределения выигрыша максимальной коалиции между игроками. Она решается путем построения дифференциальной игры в форме характеристической функции на основе исходной игры, выбора некоторого кооперативного принципа оптимальности (С-ядро, вектор Шепли и др.) и обеспечения динамической устойчивости этого решения с помощью предложенной Л.А. Петросяном процедуры распределения дележа (Петросян и др. 2012).

### Иерархическое управление для достижения устойчивого развития

Если задача управления в исходной постановке неразрешима, а метод убеждения на практике неприменим, то вводится Центр, основная цель которого - обеспечение гомеостаза посредством воздействия на множества допустимых управлений агентов (принуждение) или их функционалы выигрыша (побуждение). В этом случае возникает иерархическая дифференциальная игра

$$J_0(p(\cdot), q(\cdot), u(\cdot)) = \int_0^{\infty} e^{-\rho t} g_0(x(t), p(t), q(t), u(t), t) dt \rightarrow \max \quad (1.3.10)$$

$$p(t) \in P, q(t) \in Q; \quad (1.3.11)$$

$$J_i(p(\cdot), q(\cdot), u(\cdot)) = \int_0^{\infty} e^{-\rho t} g_i(x(t), p_i(t), u(t), t) dt \rightarrow \max \quad (1.3.12)$$

$$u_i(t) \in U_i(q_i(t)), \quad i \in N; \quad (1.3.13)$$

в силу ограничений (1.3.3) и (1.3.4). Здесь  $g_0, J_0$  - мгновенная функция и интегральный функционал выигрыша Центра;  $p(t) = (p_1(t), \dots, p_n(t))$  - вектор управлений побуждения;  $q(t) = (q_1(t), \dots, q_n(t))$  - вектор управлений принуждения. Важно отметить, что в данном случае модель представляет собой дифференциальную игру с фазовыми ограничениями, поскольку выполнение условия гомеостаза (1.3.4) есть основная цель Центра. Отметим также, что теперь множество равновесий Нэша в игре агентов есть  $NE(p, q)$ , так как решения зависят от действий Центра. Предполагаются выполненными те же математические требования, что и для модели (1.3.1)-(1.3.3).

Введем множества управлений агентов

$$U^{\max} = \text{Arg} \max_{p \in P, q \in Q} \max_{u \in U(q)} J_0(p, q, u) \quad - \text{ доставляющих глобальный максимум}$$

функционалу выигрыша Центра;

$$U_{NE}^{\max} = \mathit{Arg} \max_{p \in P, q \in Q} \max_{u \in NE(p, q)} J_0(p, q, u) \quad - \quad \text{обеспечивающих согласование}$$

интересов агентов с интересами Центра без учета гомеостаза;

$$U_{\max}^* = \mathit{Arg} \max_{p \in P, q \in Q} \max_{u \in U^*} J_0(p, q, u) \quad - \quad \text{обеспечивающих гомеостаз при учете}$$

интересов Центра, но не агентов;

$$U_{NE}^*(p, q) = NE(p, q) \cap U^* \quad - \quad \text{обеспечивающих гомеостаз при учете}$$

интересов агентов, но не Центра;

$$U_{HDG}(p, q) = U_{NE}^{\max} \cap U_{\max}^* \cap U_{NE}^*(p, q) \quad - \quad \text{обеспечивающих гомеостаз при}$$

полной системной согласованности интересов.

Если  $\exists (p, q) \in P \times Q : U_{HDG} \neq \emptyset$ , то задача управления устойчивым развитием (1.3.10)-(1.3.13), (1.3.3)-(1.3.4) имеет решение в иерархической постановке. Непустота множеств  $U_{\max}^*, U_{NE}^*(p, q)$  дает теоретическую возможность решения задачи, но на практике эта возможность сомнительна, поскольку не учитываются либо интересы агентов, либо интересы Центра.

К *аналитическим* методам решения дифференциальных игр относятся принцип максимума Понтрягина и использование уравнений Гамильтона-Якоби-Беллмана (Differential Games 2000; Петросян и др. 2012; Basar, Olsder 1999; Математическая теория 1983; Optimal Control of Nonlinear Processes 2008]. Эти методы позволяют найти решение игры (равновесие Нэша или Штакельберга) в явном виде, но в силу высокой сложности задачи они реально применимы лишь для ограниченных классов игр (например, линейных по состоянию, линейно-квадратических и т.п.). Для иерархических дифференциальных игр в (Горелик и др. 1991) предложен метод нахождения  $\varepsilon$ -оптимальной гарантирующей стратегии Центра, основанный на применении поощрения агентов, когда они выбирают желательные для него стратегии, и наказания в противном случае, что обобщает известную теорему Гермейера для статической модели (Горелик, Кононенко 1982). Однако, этот метод требует дополнительной модификации для случая нескольких агентов с учетом требований гомеостаза, предложенной в (Угольницкий, Усов 2014).

В зарубежной литературе близкий подход связан с использованием триггерных стратегий (Differential Games 2000).

В общем случае для решения задачи управления УР по сравнению с аналитическими методами более эффективны *численные методы и имитационное моделирование*. В статье (Бежавский и др. 2016) рассматривается применение алгоритмов эволюционного моделирования к решению задач распределения ресурсов в организационных системах. Предлагаемый подход обладает следующими особенностями: 1) явно описывается динамика ресурса в зависимости от управления Центра; 2) Центр распределяет ресурс между агентами пропорционально их действиям, что побуждает агентов к выбору более напряженного плана; 3) управление Центра не может меняться резко, что формализуется как липшицево свойство функции управления. Представляется, что это предположение выполняется для большинства реальных организационно-экономических систем; 4) на основе предыдущей гипотезы разрабатывается генетический алгоритм нахождения оптимальной стратегии Центра в иерархической дифференциальной игре. В работе (Ougolnitsky, Usov 2018) описан метод качественно репрезентативных сценариев (КРС) имитационного моделирования, который подробно рассматривается в п. 3.2 монографии.

*Информационные регламенты дифференциальных игр* определяются следующими парами классификационных признаков:

- программные стратегии  $v(t) = \varphi(t)$  или позиционные стратегии  $v(t) = \varphi(t, x(t))$ ;

- стратегии без обратной связи по управлению  $v(t) = \varphi(t)$  или с таковой  $v(t) = \varphi(t, u(t))$ . В первом случае возникают игры Гермейера типа  $\Gamma_{1t}$  или  $\Gamma_{1x}$  (игры Штакельберга), во втором - игры Гермейера типа  $\Gamma_{2t}$  или  $\Gamma_{2x}$  (обратные игры Штакельберга) (Горелик и др. 1991; Olsder 2009);

- принуждение  $v(t) = q(t)$  или побуждение  $v(t) = p(t)$ ;

- детерминированная или стохастическая модель.

*Индекс системной согласованности.* Условия согласованности интересов (1.3.5), (1.3.7) можно проверять в более слабой форме с помощью индекса системной согласованности

$$SCI = J_0^{\max} - J_0^{NE}, \quad (1.3.14)$$

где  $J_0^{\max} = \max_{p \in P, q \in Q} \max_{u \in U(q)} J_0(p, q, u)$ ,  $J_0^{NE} = \max_{p \in P, q \in Q} \min_{u \in NE(p, q)} J_0(p, q, u)$ .

Тогда системная согласованность имеет место при  $SCI = 0$ . Условие  $SCI \approx 0$  говорит о некотором приближении к системной согласованности.

*Некоторые частные типы Центра.* Иногда при решении задачи (1.3.3)-(1.3.4), (1.3.10)-(1.3.13) удобно выделять:

- бескорыстный Центр  $J_0(\cdot) \equiv J_0$ , заинтересованный только в обеспечении условия гомеостаза (1.3.4). Тогда множество  $U^{\max}$  не нужно и остается проверка условия  $\exists(p, q) \in P \times Q : U(p, q) \neq \emptyset$ ;

- максимизирующий общественное благосостояние Центр  $J_0(\cdot) = J(\cdot) = \sum_{i \in N} J_i(u(\cdot))$ . Тогда интересы Центра и агентов сонаправлены.

Частным случаем задачи управления УР можно считать модель стимулирования

$$J_0 = \int_0^{\infty} e^{-\rho t} \left[ H(x(t), u(t)) - \sum_{i \in N} p_i(x(t), u(t)) \right] dt \rightarrow \max_{p(\cdot)}, \quad (1.3.15)$$

$$J_i = \int_0^{\infty} e^{-\rho t} [p_i(x(t), u(t)) - h_i(x(t), u(t))] dt \rightarrow \max_{u_i(\cdot)}, i \in N, \quad (1.3.16)$$

в силу (1.3.3) и (1.3.4),

где  $H(x, u)$ - функция дохода Центра, возрастающая и вогнутая по обоим аргументам,  $H(0,0) = 0$ ;  $h_i(x, u)$ - функция затрат  $i$ -го агента, возрастающая и выпуклая по обоим аргументам,  $h_i(x, u) = 0$ ;  $p_i(x, u)$ - неотрицательная функция

компенсации затрат агента Центром,  $J_0, J_i$  - функционалы выигрыша Центра и агента соответственно.

Исходя из результата, полученного в (Новиков 2007) для статической модели стимулирования, естественно ожидать, что оптимальный механизм стимулирования должен иметь вид

$$p_i^*(x^*(t), u^*(t)) = \begin{cases} h_i(x^*(t), u_i^*(t), u_{-i}(t)) + \varepsilon_i(t), & u_i(t) = u_i^*(t), \\ 0, & \text{иначе, } i \in N, \end{cases} \quad (1.3.17)$$

где  $u^*(t)$  - решение вспомогательной задачи оптимального управления

$$\int_0^{\infty} e^{-\rho t} \left[ H(x(t), u(t)) - \sum_{i \in N} h_i(x(t), u(t)) \right] dt \rightarrow \max \quad (1.3.18)$$

в силу (1.3.2)-(1.3.4),  $\varepsilon_i(t)$  - неотрицательная мотивирующая надбавка  $i$ -му агенту как функция времени. Подробнее эта идея развивается в пп. 2.1-2.2.

Представляет интерес динамическое обобщение моделей сочетания общественных и частных интересов (СОЧИ-моделей) (Горбанёва 2019; Горбанёва, Угольницкий 2014, 2015, 2016; Gorbaneva, Ougolnitsky 2013, 2015, 2018) с учетом требований гомеостаза. В статической версии модели агенты распределяют имеющиеся у них ресурсы  $r_i > 0$  между инвестициями в производство некоторого общественного блага в размере  $0 \leq u_i \leq r_i$  и частной деятельностью в размере соответственно  $r_i - u_i$ . Частная деятельность приносит агенту доход в размере  $p_i(r_i - u_i)$ , а совокупные инвестиции всех агентов влекут производство общественного блага в объеме  $c(u_1, \dots, u_n)$ , из которого потом агент получает долю  $s_i c(u_1, \dots, u_n)$ ,  $0 \leq s_i \leq 1$ ,  $i \in N$ . Таким образом, функция выигрыша агента есть  $g_i(u) = p_i(r_i - u_i) + s_i c(u)$ ,  $u = (u_1, \dots, u_n)$ .

Динамическая СОЧИ-модель управления УР имеет вид

$$J_0 = \int_0^{\infty} e^{-\rho t} [s_0(t)x(t) - D(q(t))] dt \rightarrow \max \quad (1.3.19)$$

$$\sum_{i=0}^n s_i(t) = 1, s_i(t) \geq 0, 0 \leq q_i(t) \leq r_i, i \in N; \quad (1.3.20)$$



$$J_i = \int_0^{\infty} e^{-\rho t} [p_i(r_i - u_i(t)) + s_i(t)x(t)] dt \rightarrow \max \quad (1.3.21)$$

$$q_i(t) \leq u_i(t) \leq r_i, i \in N; \quad (1.3.22)$$

$$\dot{x} = c(u(t)) - \mu x(t), x(0) = x_0 \quad (1.3.23)$$

с учетом требования гомеостаза (1.3.4).

Здесь  $q(t) = (q_1(t), \dots, q_n(t))$ ;  $u(t) = (u_1(t), \dots, u_n(t))$ ;  $q_i(t)$  - управление принуждения Центра, ограничивающее снизу индивидуализм  $i$ -го агента;  $D(q)$  - возрастающая выпуклая функция затрат Центра на принуждение,  $D(0) = 0$ ;  $s_i(t)$  - управление побуждения Центра (доля соответствующего агента или самого Центра при  $i=0$  в общественном благе);  $x$  - величина общественного блага (переменная состояния);  $\mu$  - коэффициент амортизации общественного блага. Более адекватна, но и более сложна версия модели (1.3.19)-(1.3.23) с динамическими ресурсами, когда  $r_i = r_i(t)$ ,

$$\dot{r}_i = k_i g_i(x(t), u(t)), r_i(0) = r_{i0}, \quad (1.3.24)$$

$k_i \geq 0$  - заданная константа или  $k_i = k_i(t)$ ,  $0 \leq k_i \leq 1$  - еще одно управление агента,  $i \in N$ . Для решения поставленных задач целесообразно использовать численные методы и имитационное моделирование.

Изложенная концепция управления УР активных систем применена к решению прикладных задач в ряде различных предметных областей. Задачи управления на сетях (Губанов и др. 2010; Белов, Новиков 2018; Робертс 1986) развиты применительно к маркетингу с учетом требований устойчивого развития в статье (Агиева 2018).

В работах (Угольницкий и др. 2014, 2015) построена статическая трехуровневая теоретико-игровая модель системы контроля водяного балласта судов. Используются методы иерархического управления (побуждение и принуждение) при одновременном учёте условий поддержания системы в заданном состоянии. Проводится сравнение

результатов исследования модели с точки зрения игр Гермейера  $\Gamma_1$  и  $\Gamma_2$ . Приведены примеры численных расчетов в ряде характерных случаев.

В статье (Назиров и др. 2015) представлена математическая модель, описывающая деятельность субъектов вертикальной маркетинговой системы с учетом экологических требований. В роли субъектов управления выступают: производитель, посредник, с которым у производителя заключен договор комиссии, и торговое предприятие, реализующее в розничной сети продукцию производителя. Предполагается, что основной целью производителя является выполнение экологических требований. Предложен алгоритм построения равновесия Штакельберга в игре трех лиц с учетом требования гомеостаза экологической подсистемы. В качестве метода иерархического управления используется метод побуждения. Приведен ряд характерных примеров с последующей интерпретацией полученных результатов.

Статья (Никитина и др. 2015) посвящена построению и исследованию дифференциально-игровой модели предотвращения заморов в мелководных водоемах. Предложены алгоритмы исследования модели в случае информационных регламентов динамических игр Гермейера  $\Gamma_{1x}$  и  $\Gamma_{2x}$ . Задача решается численно с помощью разработанного параллельного алгоритма, учитывающего архитектуру суперЭВМ с распределенной памятью. Предлагаемый алгоритм численного решения поставленной задачи на суперЭВМ с использованием метода k-средних позволяет существенно сократить время работы программного комплекса, численно реализующего модельную задачу динамики взаимодействующих популяций в Азовском море. Разработанные модели используются для прогнозирования изменения биомассы биологических популяций в мелководных водоемах с учетом требований УР.

Статья (Никитина и др. 2016) демонстрирует применение концепции управления УР к задаче борьбы с эвтрофикацией мелководных водоемов (на примере Азовского моря). При описании динамики состояния водоема

используются уравнения в частных производных, которые решаются численно методом конечных разностей. Решается динамическая задача минимизации затрат на поддержание экосистемы водоема в заданном состоянии, которое интерпретируется как требование гомеостаза. Приведены примеры численных расчетов, проведен анализ полученных результатов.

В статье (Угольницкий и др. 2017) исследуются динамические теоретико-игровые модели двухуровневых систем управления с учетом условий УР. В качестве механизмов иерархического управления рассматриваются методы принуждения и побуждения. Даны определения равновесий и приведены алгоритмы их построения на основе имитационного моделирования для различных информационных регламентов. Проведен сравнительный анализ эффективности указанных механизмов управления для модели экосистемы мелководного водоема (на примере Азовского моря).

Статья (Онопrienko и др. 2016) посвящена исследованию динамических СОЧИ-моделей с использованием метода имитационного моделирования. Для игр Гермейера описаны и реализованы следующие регламенты иерархического управления: 1) программные стратегии, побуждение; 2) программные стратегии, принуждение; 3) программные и позиционные стратегии, побуждение; 4) программные и позиционные стратегии, принуждение.

В статье (Rokhlin et al. 2017) исследована двухуровневая модель оптимального рыболовства при неопределенности параметров. Изучен случай двух иерархически упорядоченных агентов: государства (регулятор) и рыболовного предприятия (рыбак). Регулятор считается дальновидным, а рыбак близоруким. Динамика рыбной популяции описана нелинейным разностным уравнением. Предполагается, что это уравнение и функция выигрыша рыбака содержат параметры, неизвестные регулятору. Доказано, что функция значения этой задачи удовлетворяет уравнению Айзекса-Беллмана, описана оптимальная стратегия регулятора. Приведены два иллюстративных примера.

В статьях (Ougolnitsky 2017; Ougolnitsky et al. 2018) рассматривается применение системного подхода к управлению УР региона на базе теоретико-игровых моделей и информационных технологий. Определяется понятие региональной активной системы. Понятие УР региона дополняется требованием системной согласованности. Устойчивое развитие означает, что хозяйственная деятельность должна обеспечивать достаточные высокие значения социально-экономических индикаторов, не нарушая в то же время требований экологического равновесия. Формально, все показатели состояния региональной социо-эколого-экономической системы должны находиться в заданном диапазоне. В свою очередь, системная согласованность означает, что при достижении своих целей Центр должен максимально учитывать интересы активных агентов. Для количественной оценки системной согласованности используется соответствующий индекс. Описаны административные и экономические механизмы управления, обеспечивающие системную согласованность.

В работах (Анопченко и др. 2017; Горбанёва и др. 2018) исследуется задача управления, которая заключается в распределении двумя соседними субъектами средств между развитием своей и общей (трансграничной) территории. Для координации деятельности вводится специальный орган управления (координатор, центр). Экономический механизм исследуется в двух вариантах (управление долей участия в доходе от развития общей территории и распределение ресурса). Приводится детальный анализ указанных механизмов, а также организационно-экономическая интерпретация для конкретных задач территориального управления.

В статье (Угольницкий, Усов 2016) исследуются динамические СОЧИ-модели инновационного развития корпорации. Динамика системы описывается нелинейным разностным уравнением. При исследовании предложенной модели инновационного развития используется имитация и метод перебора областей допустимых управлений субъектов с некоторым шагом. Основной вклад работы - сравнительный анализ эффективности

методов иерархического управления для информационных регламентов Штакельберга/Гермейера при принуждении/побуждении (четыре регламента) с помощью индексов системной согласованности. Приняты следующие предположения относительно информационного регламента иерархической игры: все игроки используют программные стратегии; ведущий выбирает и сообщает ведомым экономические либо административные управления, которые могут быть только функциями времени (игры Гермейера  $\Gamma_{1t}$ ) либо зависеть также от управлений ведомых (игры Гермейера  $\Gamma_{2t}$ ); при известных стратегиях ведущего ведомые одновременно и независимо выбирают свои стратегии, что приводит к равновесию Нэша в игре ведомых. За конечное число итераций предложенный алгоритм имитационного моделирования позволяет построить приближенное решение модели или сделать вывод, что равновесия не существует. Достоверность и эффективность предложенного алгоритма следуют из свойств методов сценариев и прямого упорядоченного перебора с постоянным шагом. Предлагаемая модель носит универсальный характер и может быть использована для научно обоснованной поддержки программ инвестиционного развития компаний всех отраслей экономики. Специфика конкретной компании учитывается в ходе идентификации модели. Получен ряд содержательных выводов относительно сравнительной эффективности методов иерархического управления инновациями.

В работах (Dyachenko et al. 2015; Нор-Аревян и др. 2018) описывается междисциплинарная методология исследования социального партнерства на примере дополнительного профессионального образования. Использование математического аппарата дифференциальных игр позволяет провести количественный сравнительный анализ эффективности различных способов организации социального партнерства (изоляция, иерархия, кооперация), учесть необходимость согласования частных и общественных интересов. Обсуждаются методика идентификации математических моделей с использованием данных социологических опросов и сценарный подход к решению дифференциальных игр на базе компьютерной имитации.

Рассмотрим упрощённый пример, иллюстрирующий введённые понятия. Пусть имеется множество  $N = \{1, \dots, n\}$  агентов-природопользователей, для определённости эксплуатирующих рыбную популяцию. Динамика популяции с учетом вылова описывается уравнением

$$\dot{x} = \left[ a - \sum_{j \in N} u_j(t) \right] x(t), x(0) = x_0. \quad (1.3.25)$$

Здесь  $x(t)$  - биомасса популяции в году  $t$ ;  $x_0$  - начальное значение биомассы;  $a > 0$  - коэффициент естественного прироста популяции;  $u_i(t)$  - доля вылова  $i$ -м агентом,

$$0 \leq u_i(t) \leq 1, i \in N. \quad (1.3.26)$$

Интересы каждого агента описываются стремлением к максимизации выигрыша

$$J_i = \int_0^T e^{-\rho t} u_i(t) x(t) dt \rightarrow \max, \quad (1.3.27)$$

где  $T$  - период рассмотрения,  $0 \leq \rho \leq 1$  - коэффициент дисконтирования.

Условие гомеостаза для рыбной популяции имеет вид

$$\forall t x(t) \geq x^*, \quad (1.3.28)$$

где  $x^*$  - критическое значение биомассы популяции.

Вообще говоря, агенты не заинтересованы в выполнении условия гомеостаза (1.3.28). Найдем равновесие Нэша в игре агентов (1.3.25)-(1.3.27).

Функция Гамильтона для  $i$ -го агента есть

$$H_i(x, u_i, \lambda_i) = u_i x + \lambda_i \left( a - \sum_{j \in N} u_j \right), i \in N.$$

В силу линейности модели имеем

$$\frac{\partial H_i}{\partial u_i} = (1 - \lambda_i) x,$$

т.е. оптимальные управления могут быть  $u_i = 0$  или  $u_i = 1$ . Поскольку при  $u_i = 0$  выигрыш агента  $J_i = 0$ , то имеет смысл рассматривать случай  $u_i = 1$ .

Тогда уравнение динамики (1.3.25) принимает вид

$$\dot{x} = (a - n)x(t), x(0) = x_0,$$

откуда

$$x(t) = x_0 e^{(a-n)t}, J_i = x_0 \int_0^T e^{(a-\rho-n)t} dt = \frac{x_0 (e^{(a-\rho-n)T} - 1)}{a - \rho - n} > 0.$$

Таким образом, действительно  $NE = \{(1, \dots, 1)\}$ . Для нахождения множества  $U^*$  надо решить неравенство

$$x_0 e^{\left(a - \sum_{j \in N} u_j(t)\right)t} \geq x^*,$$

откуда получаем условие

$$\forall t \left( a - \sum_{j \in N} u_j(t) \right) t \geq \ln \frac{x^*}{x_0}.$$

Логарифм в правой части всегда неположителен, поскольку по биологическому смыслу  $x_0 \geq x^*$  (начальная численность биомассы не меньше критической). Поэтому удобнее переписать полученное условие в виде

$$\forall t \left( \sum_{j \in N} u_j(t) - a \right) t \leq \ln \frac{x_0}{x^*}. \quad (1.3.29)$$

Если скобка в левой части отрицательна, то всегда  $U^* \neq \emptyset$ , иначе непустота этого множества зависит от соотношения параметров.

Наконец, рассмотрим множество  $U_{DG} = NE \cap U^*$ . Оно не пусто, если выполняется условие

$$\forall t (n - a)t \leq \ln \frac{x_0}{x^*}. \quad (1.3.30)$$

Таким образом, вопрос о непустоте множества  $U_{DG}$  и тем самым о разрешимости задачи управления УР (1.3.25)-(1.3.28) зависит от соотношения значений параметров модели. Вероятность выполнения условия (1.3.30) тем выше, чем больше коэффициент естественного прироста рыбной популяции, меньше число рыбаков и больше отношение начальной численности биомассы популяции к ее критическому значению.

Кооперация агентов приводит к задаче оптимального управления

$$J = \sum_{i \in N} J_i \rightarrow \max \quad (1.3.31)$$

с ограничениями (1.3.25)-(1.3.26). Поскольку функция Гамильтона для этой задачи имеет вид

$$H(x, u, \lambda) = x \sum_{j \in N} u_j + \lambda \left( a - \sum_{j \in N} u_j \right) x,$$

то ее решение совпадает с решением игры (1.3.25)-(1.3.27), т.е.  $\bar{U} = NE = \{(1, \dots, 1)\}$ , и непустота множества  $U_{oc} = U^* \cap \bar{U}$  вновь определяется условием (1.3.30).

Теперь введём Центр, который интерпретируется как орган управления некоторого уровня, отвечающий за выполнение экологических требований. Пусть функционал выигрыша Центра определяется формулой (1.3.31). Тогда экономические интересы Центра и агентов совпадают и заключаются в максимизации вылова, однако главная цель Центра - обеспечение требования гомеостаза (1.3.28). Для выполнения этого условия Центр использует метод принуждения, а именно назначает квоты на вылов рыбы

$$0 \leq u_i(t) \leq q_i(t) \leq 1, i \in N. \quad (1.3.32)$$

Очевидно, тогда равновесие Нэша в игре агентов (1.3.25), (1.3.27), (1.3.32) примет вид

$$NE = \{(q_1(t), \dots, q_n(t))\}.$$

Имеем также  $U^{\max} = U_{NE}^{\max} = NE$ . Множества  $U_{\max}^*$  и  $U_{NE}^*(q)$  определяются условиями



$$\forall t \left( \sum_{j \in N} q_j(t) - a \right) t = \ln \frac{x_0}{x^*}. \quad (1.3.33)$$

Наконец,

$$U_{HDG}(q) = U_{NE}^{\max} \cap U_{\max}^* \cap U_{NE}^*(q) = \left\{ (q_1(t), \dots, q_n(t)) : \forall t \left( \sum_{i \in N} q_i(t) - a \right) t = \ln \frac{x_0}{x^*} \right\}.$$

Таким образом, для выполнения требования гомеостаза суммарный вылов со временем придется уменьшать.

Итак, в первой главе монографии рассмотрена концепция управления УР активных систем, под которыми понимаются произвольные динамические системы с участием людей, состоящие из активных элементов. Активные элементы способны к самостоятельному целеполаганию, стратегическому поведению и сознательному искажению информации в собственных интересах. Устойчивое развитие предполагает выполнение требований гомеостаза (некоторых динамических условий на состояние объекта управления) и системной согласованности, или мотивированности (учёта интересов активных агентов, призванных обеспечивать гомеостаз).

Математическая формализация задачи управления УР активных систем проводится с помощью дифференциальных игр в нормальной форме, форме характеристической функции и с иерархической структурой, где условие гомеостаза отражается через фазовые ограничения. Приводятся постановки таких задач и указываются методы их решения. Дан обзор ряда прикладных задач в различных предметных областях, решаемых на основе предлагаемой концепции. Приведен упрощённый иллюстративный пример, демонстрирующий введённые в статье понятия, в частности, показывающий возможность непустоты различных множеств управлений.

Методика исследования задачи управления УР активных систем выглядит следующим образом.

1. Построить дифференциальную игру в нормальной форме (ДИНФ) активных агентов вида (1.3.1)-(1.3.3). Найти множество НЕ равновесий Нэша в этой игре, исследовать его зависимость от параметров модели. Найти множество управлений агентов  $U^*$ , обеспечивающих выполнение условия гомеостаза (1.3.4):  $U^* = \{u(t) \in U : \forall t x(t) \in X^*\}$ , исследовать его зависимость от параметров модели. Проверить выполнение условия разрешимости задачи управления УР активной системы:  $U_{SM} = NE \cap U^* \neq \emptyset$ .

2. Ввести суммарный функционал выигрыша  $J = \sum_{i \in N} J_i$  и найти множество

кооперативных решений  $\bar{U} = \{u(t) \in U : J(u) = \max_{v \in U} J(v)\}$ , исследовать его

зависимость от параметров модели. Проверить выполнение условия

разрешимости задачи управления УР активной системы:  $U_{SM} = \bar{U} \cap U^* \neq \emptyset$ .

Вычислить индекс системной согласованности  $SCI = J(\bar{u}) - J_{\min}^{NE}$ ,

$\bar{u} \in \bar{U}$ ,  $J_{\min}^{NE} = \min_{u \in NE} J(u)$ .

3. Добавить Центр и получить иерархическую дифференциальную игру в нормальной форме (ИДИНФ) с учетом требования гомеостаза (1.3.4) вида (1.3.10)-(1.3.13). Выяснить условия разрешимости задачи управления УР

$\exists (p, q) \in P \times Q : U_{SM}(p, q) = NE(p, q) \cap U^*(q) \neq \emptyset$ . Вычислить

$J_0^{SM} = \max_{p \in P, q \in Q} \min_{u \in U_{SM}(p, q)} J_0(p, q, u)$  и найти  $SCI = J_0^{\max} - J_0^{SM}$ ,  $J_0^{\max} = \max_{p \in P, q \in Q} \max_{u \in U(q)} J_0(p, q, u)$ .

Вообще говоря, исследование проводится для восьми возможных информационных структур (программные или позиционные стратегии, прямые или обратные игры Штакельберга, принуждение или побуждение), но реально в конкретной ситуации надо рассматривать одну-две структуры.

4. Построить и проанализировать дифференциальные игры в форме характеристической функции (ДИХФ) на основе ДИНФ (1.3.1)-(1.3.3) и иерархические дифференциальные игры в форме характеристической функции (ИДИХФ) на основе игры (1.3.10)-(1.3.13). Выяснить условия динамической устойчивости решения (как правило, вектора Шепли) и при необходимости провести процедуру распределения дележа.

## 5. Методы решения:

- ДИНФ: принцип максимума Понтрягина, уравнения Гамильтона-Якоби-Беллмана, численные методы (генетические алгоритмы, имитация отжига и др.), метод качественно репрезентативных сценариев (КРС) имитационного моделирования;
- ИДИНФ: методика А.Ф. Кононенко в авторской редакции, метод КРС;
- ДИХФ, ИДИХФ: нахождение вектора Шепли и проверка его динамической устойчивости по определению или с помощью метода КРС, процедура распределения дележа по Л.А. Петросяну;
- структурная и числовая идентификация модели: стандартные подходы с использованием метода КРС в крайних случаях.

## 6. Основные области приложений указанной концепции:

- управление организационными системами;
- территориальное управление;
- оптимальная эксплуатация биоресурсов;
- продвижение инноваций;
- финансовая математика;
- сетевые модели маркетинга;
- распределение ресурсов, в т.ч. с учетом коррупции.

## ГЛАВА 2. ДИНАМИЧЕСКИЕ МОДЕЛИ СТИМУЛИРОВАНИЯ И СОГЛАСОВАНИЯ ИНТЕРЕСОВ

Общий замысел работ, вошедших во вторую главу монографии, состоит в динамическом обобщении моделей стимулирования и согласования интересов активных агентов. В первом параграфе рассматривается базовая модель теории управления организационными системами "Центр - агент" в дискретной стохастической постановке на бесконечном интервале времени. Показано, что при соответствующих модельных предположениях сохраняет силу основной результат для статической модели: возможность построения механизма стимулирования, в котором Центр компенсирует агенту его затраты при выполнении планового задания (возможно, с мотивирующей надбавкой) и ничего не платит в противном случае. При этом плановое задание определяется в результате решения задачи оптимизации для целевой функции, равной разности между доходом Центра и затратами агента.

Во втором параграфе эта же идея проводится для древовидных систем управления типа "Центр - несколько агентов". В первых двух моделях изучается марковская игра стимулируемых агентов при частных и более общих предположениях относительно множеств стратегий. Показано, что и здесь удаётся обобщить как аналогичный результат для статической модели с древовидной структурой, так и результат предыдущего параграфа для одного агента. В третьей модели исследуется стимулирование агента Центром в динамической модели оптимальной эксплуатации биоресурсов с неопределёнными параметрами, где Центр считается дальновидным, а агент - близоруким. Получен результат о виде решения, приведены примеры.

Третий параграф посвящён динамическому обобщению моделей согласования общественных и частных интересов (СОЧИ-моделей) для ряда предметных областей: управления качеством мелководных экосистем, рыболовства, продвижения инноваций в организациях. Здесь основным методом исследования служит имитационное моделирование.

## 2.1. Базовая модель "Центр - агент"<sup>2</sup>

Мотивация к труду (стимулирование) по очевидным причинам составляет одну из ключевых проблем менеджмента. Теоретико-игровые модели стимулирования служат предметом изучения теории управления организационными системами (Бурков и Опойцев 1974; Механизмы 2011; Новиков 2007) и теории контрактов (Laffont, Martimort 2002). При этом основное внимание уделяется статическим моделям, описывающим оптимальные механизмы стимулирования для базовой модели "Центр (принципал, ведущий, регулятор) - агент (ведомый, исполнитель)" и различных ее модификаций с несколькими агентами. В теории управления организационными системами (Новиков 2007) показано, что в базовой модели оптимальный механизм стимулирования (спецификация решения статической игры Гермейера  $\Gamma_2$ ) с мотивирующей надбавкой компенсирует затраты агента при выборе им желаемого для Центра действия, определяемого максимизацией разности между доходом Центра и затратами агента. Этот результат с соответствующими дополнениями сохраняет силу и для задач стимулирования с несколькими агентами (Новиков 2007).

В теории управления организационными системами динамические модели стимулирования практически не рассматриваются. В статье (Новиков и Шохина 2003) получено решение задачи синтеза оптимальной функции стимулирования для одноэлементной динамической активной системы для связанного и несвязанного стимулирования, а также для многоэлементных динамических активных систем. При этом считается, что центр реализует программное управление в повторяющейся игре. В книге (Новиков и др. 2002) описаны результаты применения этих результатов в прикладных задачах управления активными системами. В теории контрактов при рассмотрении динамических постановок основной акцент делается на

---

<sup>2</sup> Параграф написан на основе статьи: Рохлин Д.Б., Угольницкий Г.А. Равновесие Штакельберга в динамической модели стимулирования с полной информацией // Автоматика и телемеханика, 2018, 4, 152-166.

обеспечение защиты от возможного пересмотра контрактов (их динамической устойчивости) (Laffont, Martimort 2002).

Математический аппарат исследования иерархических дифференциальных игр двух лиц на классе стратегий с обратной связью по состоянию и управлению разработан А.Ф. Кононенко и его соавторами (Кононенко 1977; Горелов, Кононенко 2015). В основе данного подхода лежит построение взаимовыгодной программы действий, реализация которой обеспечивается механизмами поощрения-наказания, предлагаемыми ведущим игроком ведомому. В зарубежной литературе аналогичная идея обозначается терминами incentives или threats. Работ в этой области достаточно много (Ho et al. 1982; Li et al. 2002), однако динамические версии модели стимулирования авторам не известны. Между тем представляется естественным обобщить указанные выше результаты на динамический случай в более общей постановке.

Рассмотрим управляемую систему с дискретным временем, пространство состояний которой есть компактное метрическое пространство  $\mathcal{X}$ , наделенное борелевской  $\sigma$ -алгеброй  $\mathcal{B}(\mathcal{X})$ . Предположим, что множество допустимых действий  $\Gamma(x)$  - непустое замкнутое подмножество интервала  $[0,1]$ . Многозначное отображение  $\Gamma: \mathcal{X} \mapsto 2^{[0,1]}$  предполагается непрерывным в смысле стандартного определения (Sundaram 1996, Ch.9); (Papageorgiou and Kyritsi-Yiallourou 2009, Proposition 6.1.2c). Динамика системы описывается переходным ядром  $q(B|x, a)$ , т.е. если система находится в состоянии  $x \in \mathcal{X}$ , то при использовании действия  $a \in \Gamma(x)$  вероятность того, что следующее состояние окажется в множестве  $B \in \mathcal{B}(\mathcal{X})$ , равна  $q(B|x, a)$ . Для каждой пары  $(x, a) \in \mathcal{X} \times [0,1]$  функция  $q(\cdot | (x, a))$  является вероятностной мерой на  $\mathcal{B}(\mathcal{X})$ , и для каждого  $B \in \mathcal{B}(\mathcal{X})$  функция  $q(B | \cdot, \cdot)$  на  $\mathcal{X} \times \mathcal{A}$  борелевская.

В рассматриваемой задаче считается, что  $a$  - усилие исполнителя, направленное на управление системой и приводящее к затратам  $g(x, a)$ . Одновременно регулятор, который непосредственно не управляет системой, но может оказывать экономическое воздействие на исполнителя, получает

выигрыш  $f(x, a)$ . Предполагается, что функции  $f, g: \mathcal{X} \times [0,1] \mapsto \mathbb{R}_+$  непрерывны,  $f(x, 0) = g(x, 0) = 0$  и

$$\max_{a \in [0,1]} (f(x, a) - g(x, a)) > 0. \quad (2.1.1)$$

Указанное экономическое воздействие состоит в том, что регулятор сообщает исполнителю стимулирующую функцию  $c$ , принадлежащую семейству  $\mathcal{L}$  функций  $c: \mathcal{X} \times [0,1] \mapsto \mathbb{R}_+$ , полунепрерывных сверху.

Напомним, что функция  $c$  называется полунепрерывной сверху, если множества  $\{(x, a): c(x, a) \geq \lambda\}$ ,  $\lambda \in \mathbb{R}$  замкнуты.

Исполнитель использует позиционные управления  $u$  из семейства  $\mathcal{U}$  борелевских функций  $u: \mathcal{X} \mapsto [0,1]$ , удовлетворяющих условию  $u(x) \in \Gamma(x)$ . Фиксируя такую функцию и начальное состояние  $x$ , получаем вероятностную меру  $P_{x,u}$  на пространстве  $(\mathcal{A} \times \mathcal{X})^\infty$  траекторий, которая формально может быть записана следующим образом (см., напр., (Hernandez-Lerma and Lasserre 1996, Appendix C)):

$$P_{x,u}(dx_0, da_0, dx_1, da_1, \dots) = \delta_x(dx_0) \delta_{u(x_0)}(da_0) q(dx_1 | x_0, a_0) \delta_{u(x_1)}(da_1) q(dx_2 | x_1, a_1) \dots,$$

где  $\delta_y$  - мера Дирака, сосредоточенная в точке  $y$ . Математическое ожидание по мере  $P_{x,u}$  обозначим через  $E_{x,u}$ .

В рассматриваемой далее игре между регулятором и исполнителем считается, что если регулятор выбирает стимулирующую функцию  $c \in \mathcal{L}$ , а исполнитель - управление  $u \in \mathcal{U}$ , то дисконтированные доходы регулятора и исполнителя определяются формулами

$$J_1(x, c, u) = E_{x,u} \sum_{t=0}^{\infty} \beta^t (f(x_t, a_t) - c(x_t, a_t)),$$

$$J_2(x, c, u) = E_{x,u} \sum_{t=0}^{\infty} \beta^t (c(x_t, a_t) - g(x_t, a_t)),$$

где  $\beta \in [0,1)$  - коэффициент дисконтирования. Регулятор является лидером, а исполнитель - ведомым. Передавая исполнителю функцию  $c$ , регулятор вычисляет оптимальную стратегию  $u^*$  последнего, а затем максимизирует

свой выигрыш по всем стимулирующим функциям  $c$ . Точная постановка задачи дана ниже.

Как известно, при фиксированной функции  $c \in \mathcal{L}$  оптимальная стратегия исполнителя может быть найдена с помощью метода динамического программирования. Предположим, что переходное ядро слабо непрерывно по  $(x, a)$ :

$$\int_{\mathcal{X}} h(y) q(dy|x_n, a_n) \rightarrow \int_{\mathcal{X}} h(y) q(dy|x, a) \quad \text{при} \quad (x_n, a_n) \rightarrow (x, a)$$

для любой непрерывной функции  $f: \mathcal{X} \mapsto \mathbb{R}$ . Тогда функция Беллмана

$$V_2(x; c) := \sup_{u \in \mathcal{U}} J_2(x, u, c) \tag{2.1.2}$$

исполнителя есть единственное полунепрерывное сверху решение уравнения

$$V_2(x; c) = \sup_{a \in \Gamma(x)} \{c(x, a) - g(x, a) + \beta \int_{\mathcal{X}} V_2(y; c) q(dy|x, a)\},$$

а его оптимальные стратегии  $u^* \in \mathcal{U}$  существуют и в точности описываются соотношением

$$u^*(x) \in \arg \max_{a \in \Gamma(x)} \{c(x, a) - g(x, a) + \beta \int_{\mathcal{X}} V_2(y; c) q(dy|x, a)\}.$$

При сделанных выше предположениях это утверждение вытекает из общих результатов (Maitra 1968; Schal 1993, Proposition 2.1; Bertsekas and Shreve 1996, Ch.9; Feinberg and Lewis 2007, Proposition 3.1). Если функция  $c$  непрерывна, то  $V_2(\cdot; c)$  также будет непрерывной.

Обозначим множество оптимальных стратегий  $u^*$  исполнителя через  $T(c)$ . Рассмотрим игру Штакельберга, в которой регулятор оценивает свой выигрыш исходя из пессимистического сценария, считая, что исполнитель использует одну из своих оптимальных стратегий:

$$G(x, c) = \inf_{u \in T(c)} J_1(x, c, u). \tag{2.1.3}$$

Оптимальный выигрыш регулятора при этом равен



$$V_1(x) = \sup_{c \in \mathcal{L}} G(x, c).$$

Назовем  $V_1$  ценой игры (лидера). Функция  $\bar{c} \in \mathcal{L}$  называется решением Штакельберга, если  $G(x, \bar{c}) = V_1(x)$ ,  $x \in \mathcal{X}$  (Basar and Olsder 1999, Def. 4.6).

Как будет показано ниже, решения Штакельберга в рассматриваемой игре не существует. В связи с этим введём также понятие  $\varepsilon$ -решения. Для  $\varepsilon > 0$  функция  $\bar{c}_\varepsilon \in \mathcal{L}$  называется  $\varepsilon$ -решением Штакельберга (Basar and Olsder 1999, Def. 4.7), если

$$G(x, \bar{c}_\varepsilon) \geq V_1(x) - \varepsilon.$$

При этом пара  $(\bar{c}_\varepsilon, u)$ ,  $u \in T(c)$  называется  $\varepsilon$ -равновесием Штакельберга.

В данной постановке задачи предполагается, что регулятор обладает полной информацией о действиях исполнителя. Если же регулятору известно только состояние системы, то естественно рассматривать класс позиционных стратегий  $c = c(x)$ .

Рассмотрим вспомогательную задачу оптимального управления с целевой функцией

$$J(x, v) = E_{x,v} \sum_{t=0}^{\infty} \beta^t (f(x_t, a_t) - g(x_t, a_t))$$

и тем же множеством  $\mathcal{U}$  допустимых управлений. Положим

$$V(x) = \sup_{v \in \mathcal{U}} J(x, v). \quad (2.1.4)$$

Как и задача исполнителя (2.1.2), данная задача имеет решение  $\bar{v} \in \mathcal{U}$ . Кроме того, поскольку функции  $f$ ,  $g$  непрерывны, то функция Беллмана  $V$  непрерывна. Условие (2.1.1) обеспечивает положительность  $V$ :

$$V(x) > 0, \quad x \in \mathcal{X}.$$

Далее, если решение  $\bar{v} \in \mathcal{U}$  задачи (2.1.4) единственно, то функция  $\bar{v}$  непрерывна. Это вытекает из представления

$$\{\bar{v}(x)\} = \operatorname{argmax}_{a \in \Gamma(x)} \{f(x, a) - g(x, a) + \beta \int_x V(y)q(dy|x, a)\}$$

и теоремы Берга о максимуме (Sundaram 1996, Theorem 9.17).

Заметим, что для любых  $c \in \mathcal{L}$ ,  $u^* \in T(c)$  верно неравенство

$$J_2(x, c, u^*) \geq J_2(x, c, 0) \geq 0.$$

Отсюда следует, что

$$G(x, c) \leq J_1(x, c, u^*) \leq (J_1 + J_2)(x, c, u^*) = E_{x, u^*} \sum_{t=0}^{\infty} \beta^t (f(x_t, a_t) - g(x_t, a_t)) \leq V(x),$$

$$V_1(x) = \sup_{c \in \mathcal{L}} G(x, c) \leq V(x). \quad (2.1.5)$$

Следующий результат показывает, что на самом деле в (2.1.5) имеет место равенство.

**Теорема 2.1.1.1** Пусть существует непрерывное оптимальное решение  $\bar{v} \in \mathcal{U}$  задачи (2.1.4). Тогда для описанной игры Штакельберга на  $(\mathcal{L}, \mathcal{U})$  справедливы следующие утверждения.

- Цена игры совпадает с оптимальным значением задачи (2.1.4):

$$V_1(x) = V(x) \quad (2.1.6)$$

- Полунепрерывная сверху функция

$$\bar{c}_\varepsilon(x, a) = g(x, a) + \varepsilon(1 - \beta)I_{\{a=\bar{v}(x)\}}, \quad \varepsilon > 0 \quad (2.1.7)$$

есть  $\varepsilon$ -решение. При этом  $T(c_\varepsilon) = \{\bar{v}\}$ .

*Доказательство.* Полунепрерывность сверху функции  $c_\varepsilon$  вытекает из непрерывности  $\bar{v}$ . Покажем, что  $\bar{v}$  - единственная оптимальная стратегия исполнителя, соответствующая  $\bar{c}_\varepsilon$ , т.е.  $T(\bar{c}_\varepsilon) = \{\bar{v}\}$ . Заметим, что  $\bar{c}_\varepsilon - g \leq (1 - \beta)\varepsilon$ . Следовательно,

$$V_2(x; \bar{c}_\varepsilon) = \sup_{u \in \mathcal{U}} J_2(x, \bar{c}_\varepsilon, u) \leq \sum_{t=0}^{\infty} \beta^t (1 - \beta) \varepsilon = \varepsilon. \quad (2.1.8)$$

Вычислим выигрыш исполнителя при использовании стратегии  $\bar{v}$ .

Учитывая, что  $\bar{c}_\varepsilon(x, \bar{v}(x)) = g(x, \bar{v}(x)) + \varepsilon(1 - \beta)$ , находим

$$J_2(x, \bar{c}_\varepsilon, \bar{v}) = E_{x, \bar{v}} \sum_{t=0}^{\infty} \beta^t (\bar{c}_\varepsilon(x_t, a_t) - g(x_t, a_t)) = \sum_{t=0}^{\infty} \beta^t (1 - \beta) \varepsilon = \varepsilon. \quad (2.1.9)$$

В силу (2.1.8) это означает, что  $V_2(x; \bar{c}_\varepsilon) = \varepsilon$  и  $\bar{v} \in T(\bar{c}_\varepsilon)$ . В частности, функция Беллмана  $V_2$  равна константе и

$$\begin{aligned} & \arg \max_{a \in \Gamma(x)} \{ \bar{c}_\varepsilon(x, a) - g(x, a) + \beta \int_{\mathcal{X}} V_2(y; c) q(dy|x, a) \} \\ &= \arg \max_{a \in \Gamma(x)} \{ \bar{c}_\varepsilon(x, a) - g(x, a) \} = \{ \bar{v}(x) \}. \end{aligned}$$

Следовательно,  $T(\bar{c}_\varepsilon) = \{ \bar{v} \}$ .

Вычислим соответствующий выигрыш регулятора. С учетом (2.1.9) имеем

$$\begin{aligned} V_1(x) &\geq J_1(x, \bar{c}_\varepsilon, \bar{v}) = E_{x, \bar{v}} \sum_{t=0}^{\infty} \beta^t (f(x_t, a_t) - \bar{c}_\varepsilon(x_t, a_t)) \\ &= E_{x, \bar{v}} \sum_{t=0}^{\infty} \beta^t (f(x_t, a_t) - g(x_t, a_t)) - \varepsilon = V(x) - \varepsilon. \end{aligned} \quad (2.1.10)$$

Из неравенств (2.1.5), (2.1.10) вытекает, что  $\bar{c}_\varepsilon$  есть  $\varepsilon$ -оптимальное решение Штакельберга и имеет место равенство (2.1.6).

Полученный результат можно описать следующим образом. Регулятор должен представить, что он сам занимается управлением системой и его затраты совпадают с затратами исполнителя. Определив оптимальное управление  $\bar{v}$ , регулятор должен мотивировать исполнителя к реализации данного управления, объявив стимулирующую функцию  $\bar{c}_\varepsilon$ . Если исполнитель будет действовать оптимально с точки зрения своих интересов, то выигрыш регулятора  $J_1(x, \bar{c}_\varepsilon, \bar{v}) = V(x) - \varepsilon$  будет отличаться на  $\varepsilon$  от его оптимального выигрыша (см. (2.1.10)), а выигрыш исполнителя будет равен  $\varepsilon$  (см. (2.1.8), (2.1.9)).

Указанное в теореме 2.1.1  $\varepsilon$ -решение, конечно, не единственное.

Например, функции

$$\bar{c}_\varepsilon^1(x, a) = g(x, a) + \varepsilon(1 - \beta)(1 - |a - \bar{v}(x)|),$$

$$\bar{c}_\varepsilon^2(x, a) = (g(x, \bar{v}(x)) + \varepsilon(1 - \beta))I_{\{a=\bar{v}(x)\}}$$

также будут  $\varepsilon$ -решениями. Заметим, что если  $\bar{v}$  непрерывна, то и функция  $\bar{c}_\varepsilon^1$  непрерывна.

В случае конечного  $\mathcal{X}$  условие непрерывности  $\bar{v}$  выполняется автоматически. Как отмечено выше, в общем случае непрерывность оптимального решения вытекает из его единственности. Обсуждение достаточных условий единственности можно найти в работе (Cruz-Suarez et al 2004). Не будем останавливаться на этом более подробно, так как в полученном ниже более общем результате (теорема 2.1.2) выполнения условия непрерывности  $\bar{v}$  не требуется.

**Замечание 2.1.1. 2** *Покажем, что решения Штакельберга не существует. В условиях теоремы 1 предположим, что  $\bar{c}$  - такое решение. Пусть для некоторых  $u \in T(\bar{c})$ ,  $x \in \mathcal{X}$  выигрыш исполнителя положителен:  $J_2(x, \bar{c}, u) > 0$ . Тогда, поскольку*

$$(J_1 + J_2)(x, \bar{c}, u) = E_{x,u} \sum_{t=0}^{\infty} \beta^t (f(x_t, a_t) - g(x_t, a_t)) \leq V(x),$$

то получаем противоречие с определением решения Штакельберга:

$$J_1(x, \bar{c}, u) < V_1(x) = V(x).$$

Таким образом, если  $\bar{c}$  - решение Штакельберга, то

$$J_2(x, \bar{c}, u) = 0, \quad x \in \mathcal{X}, \quad u \in T(\bar{c}).$$

Но так как  $J_2(x, c, 0) \geq 0$ , то  $0 \in T(\bar{c})$ , что также ведет к противоречию:

$$J_1(x, \bar{c}, 0) = 0 < V(x).$$

**Замечание 2.1.2. 3** *В рассмотренной игре регулятор рассматривает пессимистический сценарий: см. (2.1.3). Такая игра Штакельберга известна как "слабая" (Breton et al. 1988). В сильной игре Штакельберга лидер*

рассматривает оптимистический сценарий, при котором его выигрыш равен  $\sup_{u \in T(c)} J_1(x, c, u)$ . В рассматриваемом случае такая игра тривиальна. Действительно, полагая  $c = g$ , заключаем, что выигрыш исполнителя равен нулю при любой стратегии. Таким образом,  $T(g) = \mathcal{U}$ . Далее, для  $u \in \mathcal{U}$  выигрыш регулятора равен

$$J_1(x, g, u) = J(x, u) = E_{x,u} \sum_{t=0}^{\infty} \beta^t (f(x_t, a_t) - g(x_t, a_t)).$$

Выбирая в качестве  $u$  любое решение  $\bar{v}$  задачи (2.1.5), исполнитель обеспечит регулятору выигрыш  $V(x)$ . Но из неравенства (2.1.6) вытекает, что выигрыш регулятора не может быть больше этого значения при любой стратегии  $c \in \mathcal{L}$ . Таким образом,  $g$  - сильное решение Штакельберга:

$$J_1(x, g, \bar{v}) = \sup_{c \in \mathcal{L}} \sup_{u \in T(c)} J_1(x, c, u) = V(x).$$

Очевидный недостаток сильного решения здесь состоит в том, что наиболее естественной оптимальной стратегией исполнителя становится "нулевая" стратегия  $u^* = 0$ , не требующая от него приложения усилий. При этом выигрыш регулятора также будет нулевым.

Наложенное на стимулирующую функцию  $c$  условие полунепрерывности сверху связано с желанием обеспечить существование решения задачи исполнителя. Рассмотрим ситуацию, когда это требование доставляет неудобства. Если задача (2.1.5) не имеет непрерывного решения  $\bar{v}$ , то функция  $\bar{c}_\varepsilon$  (а также любая из функций  $\bar{c}_\varepsilon^1, \bar{c}_\varepsilon^2$ ), может не быть полунепрерывной сверху. Попытка перейти к полунепрерывной сверху оболочке  $\bar{c}_\varepsilon^*$  не приводит к успеху, так как решение задачи

$$\bar{c}_\varepsilon^*(x, a) - g(x, a) \rightarrow \max_{a \in \Gamma(x)} \quad (2.1.11)$$

может быть не единственным. Пусть, например,  $\mathcal{X} = [0, 1]$ , функция  $\bar{v}$  непрерывна справа и  $\bar{v}(x-) := \lim_{y \uparrow x} \bar{v}(y) < \bar{v}(x)$ ,  $x > 0$ . Ясно, что  $\bar{c}_\varepsilon^* \leq g + (1 - \beta)\varepsilon$ . С другой стороны,

$$\bar{c}_\varepsilon^*(x, \bar{v}(x-)) \geq \lim_{y \uparrow x} \bar{c}_\varepsilon(y, \bar{v}(y)) = \lim_{y \uparrow x} g(y, v(y)) + \varepsilon(1 - \beta) = g(x, \bar{v}(x-)) + \varepsilon(1 - \beta).$$

Следовательно,

$$\bar{c}_\varepsilon^*(x, \bar{v}(x-)) = g(x, \bar{v}(x-)) + (1 - \beta)\varepsilon$$

и функция (2.1.11) достигает максимума не только в точке  $\bar{v}(x)$ , но и в точке  $\bar{v}(x-)$ . Это приводит к тому, что  $T(\bar{c}_\varepsilon^*) \neq \{\bar{v}\}$  и рассуждения из доказательства теоремы 2.1.1 не проходят.

В то же время, для функции  $\bar{c}_\varepsilon$  доказательство теоремы 2.1.1 формально работает и в том случае, когда она не полунепрерывна сверху. Чтобы сделать  $c_\varepsilon$  допустимой стратегией в общем случае, естественно попытаться расширить класс  $\mathcal{L}$ , включив в него все борелевские функции. При этом решение задачи исполнителя может не существовать и естественно считать, что он будет использовать приближенно оптимальные решения. Однако, как показано в работе (Blackwell 1965, Example 2), для общей борелевской функции  $c$  в задаче (2.1.2) может не существовать даже  $\varepsilon$ -оптимальных стратегий из класса  $\mathcal{U}$ .

В связи с этим удобным оказывается рассматривать классы универсально измеримых функций. Расширим класс  $\mathcal{L}$  до класса  $\mathcal{M}$  всех универсально измеримых функций  $c: \mathcal{X} \times [0,1] \mapsto \mathbb{R}_+$ . Как известно (Shreve and Bertsekas 1979, Theorem 2; Bertsekas and Shreve 1996, Proposition 9.19), в этом случае  $\varepsilon$ -оптимальные стратегии также существуют в классе универсально измеримых функций  $u: \mathcal{X} \mapsto [0,1]$ ,  $u(x) \in \Gamma(x)$ , который обозначим через  $\mathcal{V}$ . Рассмотрим множество  $\eta$ -оптимальных решений исполнителя в классе  $\mathcal{V}$ :

$$T_\eta(c) = \{u \in \mathcal{V}: J_2(x, u, c) \geq \sup_{v \in \mathcal{V}} J_2(x, v, c) - \eta\}, \quad \eta \geq 0.$$

Следуя известному определению (Morgan 1989, Def. 4.1; Patrone 1995, Def. 11], элемент  $\bar{c}_{\varepsilon, \eta} \in \mathcal{M}$  назовем  $(\varepsilon, \eta)$ -решением Штакельберга, если

$$\inf_{u \in T_\eta(\bar{c}_{\varepsilon, \eta})} J_1(x, \bar{c}_{\varepsilon, \eta}, u) \geq \sup_{c \in \mathcal{M}} \inf_{u \in T_\eta(c)} J_1(x, c, u) - \varepsilon.$$

Обозначим через  $\mathcal{N}$  множество всех последовательностей  $(\varepsilon_n, \eta_n) > 0$ , удовлетворяющих условию  $(\varepsilon_n, \eta_n) \rightarrow 0$ . Величину

$$V_1(x) = \sup_{(\varepsilon_n, \eta_n)_{n=1}^{\infty} \in \mathcal{N}(\varepsilon_n, \eta_n) \rightarrow 0} \liminf \{J_1(x, \bar{c}_{\varepsilon_n, \eta_n}, u) : u \in T_{\eta_n}(\bar{c}_{\varepsilon_n, \eta_n})\} \quad (2.1.12)$$

назовем ценой игры (лидера). Заметим, что рассматриваемая игра некорректна в смысле определений (Morgan 1989), так как равновесия Штакельберга не существует. Тем не менее, имеет место следующий результат.

**Теорема 2.1.2. 4** *Для описанной игры Штакельберга на  $(\mathcal{M}, \mathcal{V})$  имеют место следующие утверждения.*

- Цена игры совпадает с оптимальным значением задачи (2.1.4):

$$V_1(x) = V(x).$$

- При

$$\eta < \frac{1-\beta}{1+\beta} \varepsilon \quad (2.1.13)$$

функция (2.1.7) является  $(\varepsilon, \eta)$ -решением, и  $T_\eta(\bar{c}_\varepsilon) = \{\bar{v}\}$ .

*Доказательство.* Верхняя оценка (2.1.5) цены игры сохраняет силу. Для доказательства теоремы достаточно установить, что  $T_\eta(\bar{c}_\varepsilon) = \{\bar{v}\}$  при выполнении условия (2.1.13), так как тогда неравенство (2.1.10)

$$\inf_{u \in T_\eta(\bar{c}_\varepsilon)} J_1(x, \bar{c}_\varepsilon, u) = J_1(x, \bar{c}_\varepsilon, \bar{v}) \geq V(x) - \varepsilon$$

позволяет в определении (2.1.12) выбрать последовательность  $(\varepsilon_n, \eta_n)$ , например, следующим образом:  $(1/n, (1 - \beta)/(2(1 + \beta)n))$ .

Доказательство теоремы 2.1.1 показывает, что  $\bar{v} \in T(\bar{c}_\varepsilon) \subset T_\eta(\bar{c}_\varepsilon)$ . При доказательстве соотношения  $T_\eta(\bar{c}_\varepsilon) = \{\bar{v}\}$  мы следуем идеям (Montes-De-Oca

and Lemus-Rodriguez 2010), которые в данном случае работают с существенными упрощениями. Пусть

$$H(x, a) = \varepsilon(1 - \beta)I_{\{a=\bar{v}(x)\}} + \beta \int_{\mathcal{X}} V_2(y; \bar{c}_\varepsilon)q(dy|x, a).$$

Рассмотрим произвольную  $\eta$ -оптимальную стратегию  $u$  исполнителя:

$$J_2(x, \bar{c}_\varepsilon, u) \geq V_2(x; \bar{c}_\varepsilon) - \eta.$$

Имеем

$$H(x, \bar{v}(x)) = \varepsilon(1 - \beta) + \beta \int_{\mathcal{X}} V_2(y; \bar{c}_\varepsilon)q(dy|x, \bar{v}(x)),$$

$$H(x, u(x)) = \varepsilon(1 - \beta)I_{\{u(x)=\bar{v}(x)\}} + \beta \int_{\mathcal{X}} V_2(y; \bar{c}_\varepsilon)q(dy|x, u(x)). \quad (2.1.14)$$

Но  $V_2 = \varepsilon$ , поэтому

$$H(x, \bar{v}(x)) - H(x, u(x)) = \varepsilon(1 - \beta)I_{\{u(x) \neq \bar{v}(x)\}}. \quad (2.1.15)$$

С другой стороны, с учетом оптимальности  $\bar{v}$  и принципа динамического программирования имеем

$$H(x, \bar{v}(x)) = V_2(x; \bar{c}_\varepsilon), \quad (2.1.16)$$

$$J_2(x, \bar{c}_\varepsilon, u) = \varepsilon(1 - \beta)I_{\{u(x)=\bar{v}(x)\}} + \beta \int_{\mathcal{X}} J_2(y, \bar{c}_\varepsilon, u)q(dy|x, u(x)). \quad (2.1.17)$$

Используя равенства (2.1.14), (2.1.16), (2.1.17), находим

$$\begin{aligned} |H(x, \bar{v}(x)) - H(x, u(x))| &= |V_2(x; \bar{c}_\varepsilon) - J_2(x, \bar{c}_\varepsilon, u) + \beta \int_{\mathcal{X}} J_2(y, \bar{c}_\varepsilon, u)q(dy|x, u(x)) \\ &\quad - \beta \int_{\mathcal{X}} V_2(y; \bar{c}_\varepsilon)q(dy|x, u(x))| \leq (1 + \beta)\eta. \end{aligned} \quad (2.1.18)$$

Последнее неравенство вытекает из свойства  $\eta$ -оптимальности  $u$ . Если  $u(x) \neq \bar{v}(x)$  при некотором  $x$ , то неравенства (2.1.15), (2.1.18) противоречат друг другу при выполнении условия (2.1.13).

Полученный результат затрагивает также важную проблему, касающуюся последствий (с точки зрения регулятора) использования исполнителем приближенно оптимальных стратегий. Теорема 2.1.2 указывает на то, что при жёстком регулировании, которое заложено в стимулирующую



стратегию (2.1.7), исполнитель может отклониться от управления  $\bar{v}$ , в котором заинтересован регулятор, только за счет больших убытков.

Всюду выше предполагалось, что игроки используют одинаковые коэффициенты дисконтирования. Предположим теперь, что исполнитель использует, вообще говоря, другой коэффициент дисконтирования  $\gamma$ :

$$J_2(x, c, u) = E_{x,u} \sum_{t=0}^{\infty} \gamma^t (c(x_t, a_t) - g(x_t, a_t)) \rightarrow \max_{u \in \mathcal{U}} \gamma \in [0,1).$$

Из доказательств теорем 2.1.1, 2.1.2 ясно, что для любого  $\gamma \in [0,1)$  по-прежнему будут верны соотношения  $\bar{v} \in T(\bar{c}_\varepsilon) \subset T_\eta(\bar{c}_\varepsilon)$  и  $\{\bar{v}\} = T_\eta(\bar{c}_\varepsilon)$  при

$$\eta < \frac{1-\gamma}{1+\gamma} \varepsilon, \quad (2.1.19)$$

а также неравенство (2.1.10):  $V_1 \geq V - \varepsilon$ . Таким образом, цены  $V_1$  обеих описанных выше игр будут не меньше  $V$ . Поскольку рассуждение, которое привело к верхней оценке (2.1.5) цены игры, здесь неприменимо, возникает вопрос, может ли выигрыш регулятора быть больше?

**Гипотеза. 5** *Теоремы 2.1.1, 2.1.2 сохраняют силу (при замене (2.1.13) условием (2.1.19)) при любом коэффициенте дисконтирования  $\gamma \in [0,1)$ .*

Покажем, что данная гипотеза верна в тривиальном случае "близорукого" исполнителя:  $\gamma = 0$ . Пусть  $c \in \mathcal{L}$  или  $c \in \mathcal{M}$ . Для  $\eta$ -оптимального решения  $u^* \in \mathcal{U}$  задачи исполнителя

$$c(x, a) - g(x, a) \rightarrow \max_{a \in \Gamma(x)}$$

верно неравенство  $c(x, u^*(x)) - g(u^*(x)) \geq c(x, 0) - \eta$ . Выигрыш регулятора тогда удовлетворяет оценке

$$\begin{aligned} J_1(x, c, u^*) &= E_{x,u^*} \sum_{t=0}^{\infty} \beta^t (f(x_t, a_t) - c(x_t, a_t)) \\ &\leq E_{x,u^*} \sum_{t=0}^{\infty} \beta^t (f(x_t, a_t) - g(x_t, a_t) - c(x_t, 0) + \eta) \leq V(x) + \eta. \end{aligned} \quad (2.1.20)$$

Рассмотренным в теоремах 2.1.1, 2.1.2 постановкам задач соответствуют случаи  $\eta = 0$  и  $\eta > 0$ . В любом из них из (2.1.20) получаем верхнюю оценку цены игры:  $V_1 \leq V$ . Таким образом, регулятор не может получить выигрыш больше  $V$  в игре против близорукого исполнителя.

Интуиция, подкрепляющая гипотезу, говорит о том, что ситуация не может стать лучше, даже если исполнитель не близорукий.

Итак, предложена динамическая постановка базовой задачи стимулирования "Центр - агент". Показано, что  $\varepsilon$ -оптимальная стратегия Центра основана на решении задачи оптимального управления с целевым функционалом, равным разности функций дохода ведущего и затрат ведомого, т.е. получено непосредственное обобщение соответствующего результата для статической модели стимулирования (Новиков 2007). При этом в доказательстве явно не используется построение механизма поощрения-наказания, как в работах (Кононенко 1977; Горелов и Кононенко 2015) или работах из обзора (Но 1982). Теорема первоначально доказана для полунепрерывных сверху функций и затем распространена на универсально измеримые функции. Интересно, что  $\varepsilon$ -добавка в определении найденной  $\varepsilon$ -оптимальной стратегии Центра может трактоваться как мотивирующая надбавка, которая тем больше, чем сильнее близорукость исполнителя ( $\beta \rightarrow 0$ ).

## 2.2. Модели с несколькими агентами

Теория мотивации (theory of incentives) развивается крайне интенсивно на протяжении последних четырёх десятилетий (Laffont, Martimort 2002; Muthoo 1999; Myerson 1979, 1982, 1983). Действительно, адекватные механизмы мотивации играют ключевую роль в успешном менеджменте. Наиболее распространённой математической моделью мотивационных механизмов служат так называемые обратные игры Штакельберга (Li, Sethi 2017; Olsder 2009a,b). В этой модели ведущий игрок (Центр) сообщает одному или нескольким ведомым (агентам) свою стратегию как функцию его/их управляющих действий. Оптимальная стратегия Центра максимизирует его выигрыш на множестве наилучших ответов агента (-ов). При нескольких агентах их наилучшие ответы обычно определяются как множество равновесий Нэша в игре агентов в нормальной форме.

По сути, наиболее представительная математическая теория обратных игр Штакельберга была построена Ю.Б. Гермейером (1975) в статической постановке и развита А.Ф. Кононенко (1977) в динамическом случае. Согласно предложенному ими подходу, Центр поощряет агента за следование оптимальному плану и наказывает за отклонение от него. Доказано, что такой механизм управления образует  $\varepsilon$ -оптимальную стратегию Центра. Применительно к менеджменту эта идея была специфицирована Д.А. Новиковым (2007) в его теории управления организационными системами. Соответствующая модель мотивации использует специальные функции выигрыша. Центр максимизирует разность между своим доходом и вознаграждением агентам, а агенты - разность между этим вознаграждением и своими трудовыми затратами. Для статического случая показано, что в оптимальном механизме управления Центр компенсирует агенту его затраты (возможно, с мотивирующей надбавкой), если последний выполняет оптимальный для Центра план, и ничего не платит в противном случае. В свою очередь, оптимальный план Центра

максимизирует разность между его доходом и затратами агента. Этот результат был обобщён для различных схем взаимодействия агентов.

В предыдущем параграфе приведены результаты авторов по обобщению задачи стимулирования на динамический случай. В статье (Rokhlin and Ougolnitsky 2019a) динамическая модель стимулирования из предыдущего параграфа обобщается на случай нескольких агентов.

Рассмотрим марковскую игру, определяемую:

- конечным пространством состояний  $X$ ;
- конечным множеством действий игроков  $A = A_1 \times \dots \times A_N$ ;
- множественно-значными отображениями  $x \rightarrow \Gamma_i(x) \subset A_i$ , описывающими допустимые действия игрока  $i$  в состоянии  $x \in X$ ;
- переходным ядром  $q(y|x, a)$  таким, что

$$\sum_{y \in X} q(y|x, a) = 1, \quad q(y|x, a) \geq 0, \quad x \in X, a \in A;$$

- мгновенными выигрышами игроков  $r_i: X \times A \mapsto \mathbb{R}$ ;
- коэффициентом дисконтирования  $\beta \in [0, 1)$ .

Предполагается, что игроки используют стационарные марковские стратегии, задаваемые распределениями  $\pi_i(a|x)$ :

$$\sum_{a_i \in \Gamma_i(x)} \pi_i(a|x) = 1, \quad \pi_i(a|x) \geq 0, \quad x \in X.$$

Любой набор  $\pi = (\pi_1, \dots, \pi_N)$  порождает единственную вероятностную меру  $P_{x, \pi}$  on  $(X \times A)^\infty$  по формуле

$$P_{x, \pi}(x_0, a_0, \dots, x_n, a_n) = \delta_x(x_0) \prod_{t=0}^{n-1} \prod_{i=1}^N \pi_i(a_t^i | x_t) q(x_{t+1} | x_t, a_t) \prod_{i=1}^N \pi_i(a_n^i | x_n),$$

где  $\delta_x(y) = 1, y = x; \delta_x(y) = 0, y \neq x$ . Положим  $\pi = (\pi_1, \dots, \pi_N)$  и обозначим через  $E_{x, \pi}$  математическое ожидание по мере  $P_{x, \pi}$ . Ожидаемый дисконтированный выигрыш игрока  $i$  равен

$$J_i(x, \pi) = E_{x, \pi} \sum_{t=0}^{\infty} \beta^t r_i(x_t, a_t).$$

Тем самым описана динамическая игра, в которой на каждом шаге  $t$  на основе информации о состоянии  $x_t$  игроки независимо выбирают случайные действия  $a_t^i \in \Gamma_i(x_t)$  согласно распределениям  $\pi_i(\cdot | x_t)$ . После этого система переходит в новое состояние  $x_{t+1}$  с вероятностью  $q(x_{t+1} | x_t, a_t)$ , и игроки получают выигрыши  $r_i(x_t, a_t)$ . Ситуация  $\pi^* = (\pi_1^*, \dots, \pi_N^*)$  называется равновесием Нэша, если

$$J_i(x, \pi_i, \pi_{-i}^*) \leq J_i(x, \pi^*), \quad x \in X, \quad i = 1, \dots, N.$$

Мы используем стандартное обозначение  $(\pi_i, \pi_{-i}^*)$  для ситуации, полученной из  $\pi^*$  заменой  $\pi_i^*$  на  $\pi_i$ . Известно, что в данной игре равновесие Нэша существует (Dutta and Sundaram 1998; Theorem 3.1). Каждая компонента  $\pi_i^*$  равновесия  $\pi^*$  есть решение задачи оптимизации

$$J_i(x, \pi_i, \pi_{-i}^*) \rightarrow \max_{\pi_i} \quad (2.2.1)$$

Перепишем эту целевую функцию следующим образом:

$$\begin{aligned} J_i(x, \pi_i, \pi_{-i}^*) &= \sum_{t=0}^{\infty} \beta^t \sum_{y, a} r_i(y, a) P_{(x, \pi_i, \pi_{-i}^*)}(x_t = y, a_t = a) \\ &= \sum_{t=0}^{\infty} \beta^t \sum_{y, a} r_i(y, a) \sum_{x_0, a_0} \dots \sum_{x_{t-1}, a_{t-1}} \delta_x(x_0) \pi_i(a_0^i | x_0) \prod_{j \neq i} \pi_j^*(a_0^j | x_0) q(x_1 | x_0, a_0) \times \dots \\ &\times \pi_i(a_{t-1}^i | x_{t-1}) \prod_{j \neq i} \pi_j^*(a_{t-1}^j | x_{t-1}) q(y | x_{t-1}, a_{t-1}) \prod_{j \neq i} \pi_j^*(a^j | y) \pi_i(a^i | y) \\ &= \sum_{t=0}^{\infty} \beta^t \sum_{y, a^i} \sum_{x_0, a_0^i} \dots \sum_{x_{t-1}, a_{t-1}^i} \delta_x(x_0) \pi_i(a_0^i | x_0) \sum_{a_0^{-i}} \prod_{j \neq i} \pi_j^*(a_0^j | x_0) q(x_1 | x_0, a_0) \times \dots \\ &\times \pi_i(a_{t-1}^i | x_{t-1}) \sum_{a_{t-1}^{-i}} \prod_{j \neq i} \pi_j^*(a_{t-1}^j | x_{t-1}) q(y | x_{t-1}, a_{t-1}) \pi_i(a^i | y) \sum_{a^{-i}} \prod_{j \neq i} \pi_j^*(a^j | y) r_i(y, a). \end{aligned}$$

Суммирование проводится по  $y \in X$ ,  $a_k^i \in \Gamma_i(x_k)$ ,  $a^i \in \Gamma_i(y)$ . Введём обозначение

$$\begin{aligned} r_{-i}^*(y, a^i) &= \sum_{a^{-i}} \prod_{j \neq i} \pi_j^*(a^j | y) r_i(y, a), \quad q_{-i}^*(y | x, a^i) = \\ &= \sum_{a^{-i}} \prod_{j \neq i} \pi_j^*(a^j | x) q(y | x, a). \end{aligned} \quad (2.2.2)$$

Тогда

$$J_i(x, \pi_i, \pi_{-i}^*) = \sum_{t=0}^{\infty} \beta^t \sum_{y, a^i} \sum_{x_0, a_0^i} \dots \sum_{x_{t-1}, a_{t-1}^i} \delta_x(x_0) \pi_i(a_0^i | x_0) q_{-i}^*(x_1 | x_0, a_0) \times \dots \\ \times \pi_i(a_{t-1}^i | x_{t-1}) q_{-i}^*(y | x_{t-1}, a_{t-1}) \pi_i(a^i | y) r_{-i}^*(y, a^i).$$

Определим вероятностную меру

$$P_{x, \pi_i}^{*, -i}(x_0, a_0^i, \dots, x_n, a_n^i) = \delta_x(x_0) \prod_{t=0}^{n-1} \pi_i(a_t^i | x_t) q_{-i}^*(x_{t+1} | x_t, a_t^i) \pi_i(a_n^i | x_n)$$

на декартовом произведении  $(X \times A_i)^\infty$ . Легко видеть, что

$$J_i(x, \pi_i, \pi_{-i}^*) = \sum_{t=0}^{\infty} \beta^t \sum_{y, a^i} r_{-i}^*(y, a^i) P_{x, \pi_i}^{*, -i}(x_t = y, a_t^i = a^i) \\ = E_{x, \pi_i}^{*, -i} \sum_{t=0}^{\infty} \beta^t r_{-i}^*(x_t, a_t^i),$$

где  $E_{x, \pi_i}^{*, -i}$  есть математическое ожидание относительно  $P_{x, \pi_i}^{*, -i}$ . Таким образом, задача (2.2.1) соответствует марковскому процессу с пространством состояний  $X$ , множествами действий  $x \rightarrow \Gamma_i(x) \subset A_i$ , переходным ядром и выигрышами (2.2.2) и коэффициентом дисконтирования  $\beta \in [0, 1)$ . Как хорошо известно, функция значения

$$V_i(x, \pi_{-i}^*) = \sup_{\pi_i} J_i(x, \pi_i, \pi_{-i}^*),$$

где супремум берётся по всем рандомизированным марковским стратегиям  $\pi_i$ , есть единственное решение уравнения

$$V_i(x, \pi_{-i}^*) = \max_{a^i \in \Gamma_i(x)} \left\{ r_{-i}^*(x, a^i) + \beta \sum_{y \in X} q(y | x, a^i) V_i(y, \pi_{-i}^*) \right\} \\ = \max_{p \in \Delta(\Gamma_i(x))} \left\{ \sum_{a^i \in \Gamma_i(x)} r_{-i}^*(x, a^i) p(a^i) + \beta \sum_{y \in X} \sum_{a^i \in \Gamma_i(x)} q(y | x, a^i) p(a^i) V_i(y, \pi_{-i}^*) \right\},$$

где  $\Delta(\Gamma_i(x))$  - множество вероятностных мер на конечном множестве  $\Gamma_i(x)$ .  
 Более того, для оптимальной рандомизированной марковской стратегии  $\bar{\pi}_i$  в силу принципа оптимальности имеет место равенство

$$V_i(x, \pi_{-i}^*) = \sum_{a^i \in \Gamma_i(x)} r_{-i}^*(x, a^i) \bar{\pi}_i(a^i | x) + \beta \sum_{y \in X} \sum_{a^i \in \Gamma_i(x)} q(y | x, a^i) \bar{\pi}_i(a^i | x) V_i(y, \pi_{-i}^*).$$

Итак, оптимальные рандомизированные марковские стратегии в точности описываются соотношением

$$\bar{\pi}_i(a^i | x) \in \arg \max_{p \in \Delta(\Gamma_i(x))} \left\{ \sum_{a^i \in \Gamma_i(x)} r_{-i}^*(x, a^i) p(a^i) + \beta \sum_{y \in X} \sum_{a^i \in \Gamma_i(x)} q(y | x, a^i) p(a^i) V_i(y, \pi_{-i}^*) \right\}.$$

Будем рассматривать игроков как производителей. Их выигрыши имеют вид  $r_i = c_i(x, a) - g_i(x, a)$ , где  $g_i$  - функция затрат  $i$ -го производителя (агента), а  $c_i$  - неотрицательная функция стимулирования, выбираемая управляющим (Центром). Действие  $a^i$  трактуется как объём выпуска  $i$ -го производителя. Считается, что при нулевом объёме выпуска затраты  $i$ -го производителя равны нулю независимо от объёмов выпуска других агентов:

$$g_i(x, (0, a^{-i})) = 0. \quad (2.2.3)$$

Для заданного набора  $c = (c_1, \dots, c_N)$  агенты с целевыми функциями

$$J_i(x, \pi, c) = E_{x, \pi} \sum_{t=0}^{\infty} \beta^t (c_i(x_t, a_t) - g_i(x_t, a_t))$$

участвуют в игре, исход которой есть равновесие Нэша. Обозначим через  $T(c)$  множество таких равновесий. Ожидаемый дисконтированный выигрыш Центра равен

$$J_L(x, \pi, c) = E_{x, \pi} \sum_{t=0}^{\infty} \beta^t \left( f(x_t, a_t) - \sum_{i=1}^N c_i(x_t, a_t) \right),$$

где  $f(x, a)$  - его мгновенный выигрыш. Цель Центра состоит в максимизации дисконтированного выигрыша, соответствующего "наихудшему" из равновесий Нэша, по всем наборам функций стимулирования:

$$G(x, c) = \inf_{\pi \in T(c)} J_L(x, \pi, c) \rightarrow \max_c.$$

Эта задача может быть классифицирована как обратная игра Штакельберга, поскольку стратегии Центра зависят от стратегий агентов. Назовём

$$V_L(x) = \sup_c G(x, c)$$

значением Центра. Набор  $\bar{c}_\varepsilon$  называется  $\varepsilon$ -решением Штакельберга, если

$$V_L(x) - \varepsilon \leq G(x, \bar{c}_\varepsilon), \quad x \in X.$$

Пара  $(\bar{c}_\varepsilon, \pi)$ ,  $\pi \in T(\bar{c}_\varepsilon)$  называется  $\varepsilon$ -равновесием Штакельберга.

Рассмотрим вспомогательный марковский процесс с пространством состояний  $X$ , множествами действий  $x \rightarrow \Gamma(x) = \Gamma_1(x) \times \dots \times \Gamma_N(x) \subset A$ , переходным ядром  $q$ , выигрышами

$$r(x, a) = f(x, a) - \sum_{i=1}^N g_i(x, a)$$

и коэффициентом дисконтирования  $\beta \in [0, 1)$ . Целевая функция этой задачи

$$J(x, \pi) = E_{x, \pi} \sum_{t=0}^{\infty} \beta^t \left( f(x_t, a_t) - \sum_{i=1}^N g_i(x_t, a_t) \right) \quad (2.2.4)$$

соответствует гипотетическому выигрышу Центра, который он мог бы получить от участия в процессе производства, не прибегая к услугам производителей и считая свои затраты равными затратам последних.

Из экономических соображений оптимальный выигрыш

$$V(x) = \sup_{\pi} J(x, \pi) \quad (2.2.5)$$



есть верхняя граница для  $V_L(x)$ , поскольку попытка переложить выпуск продукции на плечи производителей, используя экономический механизм стимулирования, не может вести к большей выгоде. Оказывается, более того, что эти величины совпадают.

Как известно, среди оптимальных решений (2.2.5) существует детерминированная марковская стратегия

$$\bar{\pi}_i(y|x) = \delta_{\bar{u}^i(x)}(y),$$

$$\text{где } \bar{u}^i(x) \in \arg \max_{a \in \Gamma(x)} \{r(x, a) + \beta \sum_{y \in X} q(y|x, a)V(y)\},$$

а  $V$  есть единственное решение уравнения динамического программирования

$$V(x) = \max_{a \in \Gamma(x)} \{r(x, a) + \beta \sum_{y \in X} q(y|x, a)V(y)\}. \quad (2.2.6)$$

*Теорема 2.2.1* (Rokhlin and Ougolnitsky 2019a). Для описанной выше обратной игры Штакельберга:

(i) значение Центра совпадает с (2.2.5):

$$V_L(x) = V(x), \quad x \in X.$$

(ii) функция

$$\bar{c}_\varepsilon^i(x, a) = g_i(x, a) + (1 - \beta) \frac{\varepsilon}{N} I_{\{a^i = \bar{u}^i(x)\}}, \quad \varepsilon > 0$$

есть  $\varepsilon$ -решение, и соответствующее равновесие Нэша единственно:

$$T(\bar{c}_\varepsilon) = \{\delta_{\bar{u}^1(x)}, \dots, \delta_{\bar{u}^1(x)}\}.$$

Согласно теореме 2.2.1, оптимальная стратегия Центра имеет нормативный характер. Центр должен: 1) решить задачу оптимального управления (2.2.4); 2) сообщить заданный оптимальный план  $\bar{u}^i(x)$  каждому агенту; 3) компенсировать затраты агентов со стимулирующей надбавкой  $\varepsilon$ .

К сожалению, реализация этой стратегии требует знания функций затрат  $g_i$  всех агентов и переходного ядра  $q$ . Хотя эти предположения далеки от реальности, значение  $V_L(x) = V(x)$  может служить полезной точкой

отсчёта. Рассмотрим иллюстративный пример, в котором стратегия Центра имеет очень простую структуру.

Предположим, что Центр - это руководитель большой компании, производящей сложные изделия. Заказы  $(a_1, \dots, a_N)$  на компоненты для них размещаются среди  $N$  поставщиков. Каждый поставщик производит компоненты только одного типа, и его производственные издержки задаются неубывающей функцией  $g_i(a_i)$ . Чтобы упростить обозначения, допустим, что в производстве единицы сложного изделия используется один экземпляр каждой компоненты.

Пусть спрос определяется последовательностью независимых одинаково распределённых неотрицательных случайных переменных  $\xi_t$  со значениями из множества  $\mathbb{Z}_+ = \{0, 1, \dots\}$ . Динамическая переменная состояния  $x_t \in \mathbb{Z}_+$  описывает количество изделий:

$$\begin{aligned} x_{t+1} &= \max\{x_t + u_t - \xi_{t+1}, 0\}, \quad u_t = \min\{a_{1,t}, \dots, a_{N,t}\} \\ u_t &\in [0, M - x_t]. \end{aligned} \quad (2.2.7)$$

Здесь  $u_t$  - количество изделий, производимых в момент времени  $t$ . Верхняя граница в (2.2.7) определяется ёмкостью склада. Пусть  $P$  - рыночная цена на единицу изделия. Выигрыш Центра равен разности между прибылью от продажи изделий в момент времени  $t + 1$  и затратами на их хранение за период  $[t, t + 1)$ :

$$f(X_t, a_t, \xi_{t+1}) = \min\{X_t + u_t, \xi_{t+1}\}P - h(x_t + u_t), \quad (2.2.8)$$

Функция затрат на хранение  $h$  неубывающая. Ясно, что Центру невыгодно заказывать разные объёмы компонент у различных поставщиков, поскольку при этом количества  $a^i - u_t$  теряются. Поэтому в соответствии с описанной выше схемой Центр решает задачу стохастического оптимального управления

$$J(x, \pi) = E_{x, \pi} \sum_{t=0}^{\infty} \beta^t \left( \min\{x_t + u_t, \xi_{t+1}\} \cdot P - h(x_t + u_t) - \sum_{i=1}^N g_i(u_t) \right)$$

с пространством состояний  $\{0, \dots, M\}$ . Соответствующая функция значения (2.2.5) удовлетворяет уравнению динамического программирования

$$V(x) = \max_{u \in \{0, \dots, M-x\}} (P \cdot E \min\{x + u, \xi\} - h(x + u) - g(u) + \beta EV(\max\{x + u - \xi, 0\})),$$

где  $g(u) = \sum_{i=1}^N g_i(u_i)$ , и  $\xi$  распределено как  $\xi_t$ . Далее, пусть  $P(\xi = k) = p_k$ ,  $k \in \mathbb{Z}_+$ . Тогда

$$V(x) = \max_{u \in \{0, \dots, M-x\}} \left( \sum_{k=0}^{x+u-1} p_k (kP + \beta V(x + u - k)) + \left( 1 - \sum_{k=0}^{x+u-1} p_k \right) ((x + u)s + \beta V(0)) - h(x + u) - g(u) \right), \quad x \in \{0, \dots, M\}.$$

Это уравнение решалось численно методом итерации по значению (Puterman 1994) для распределения Пуассона  $p_k = \lambda^k e^{-\lambda} / k!$  для спроса и следующих функций затрат:

$$g(u) = \begin{cases} K + Cu, & u > 0 \\ 0, & u = 0, \end{cases} \quad h(z) = Az.$$

Значения  $K > 0$  соответствуют фиксированным издержкам на заказы. В зависимости от наличия таких издержек, получены два хорошо известных типа оптимальных стратегий: базовая стратегия пополнения

$$u^*(x) = \begin{cases} S - x, & x \leq S \\ 0, & x \geq S \end{cases} \quad (2.2.9)$$

для  $K = 0$  и  $(s, S)$ -стратегия:

$$u^*(x) = \begin{cases} S - x, & x \leq s \\ 0, & x \geq s \end{cases} \quad (2.2.10)$$

для  $K > 0$ .

Для  $M = 40, \lambda = 15, P = 14, C = 2, A = 1$  результаты представлены на рис. 2.2.1. Видно, что оптимальная стратегия заказа Центра определяется одним или двумя параметрами:  $S$  или  $(s, S)$ . Однако  $\varepsilon$ -оптимальный механизм стимулирования, определяемый теоремой 2.2.1, требует знания коэффициентов  $K_i, C_i, i = 1, \dots, N$  функций затрат поставщиков, которые по предположению имеют вид

$$g_i(u) = \begin{cases} K_i + C_i u^i, & u^i > 0 \\ 0, & u^i = 0, \end{cases}$$

где  $\sum_{i=1}^N K_i = K, \sum_{i=1}^N C_i = C$ . Действительно, по теореме 2.2.1 имеем

$$\bar{c}_\varepsilon^i(x, a) = g_i(a^i) + (1 - \beta) \frac{\varepsilon}{N} I_{\{a^i = u^*(x)\}}, \quad \varepsilon > 0.$$

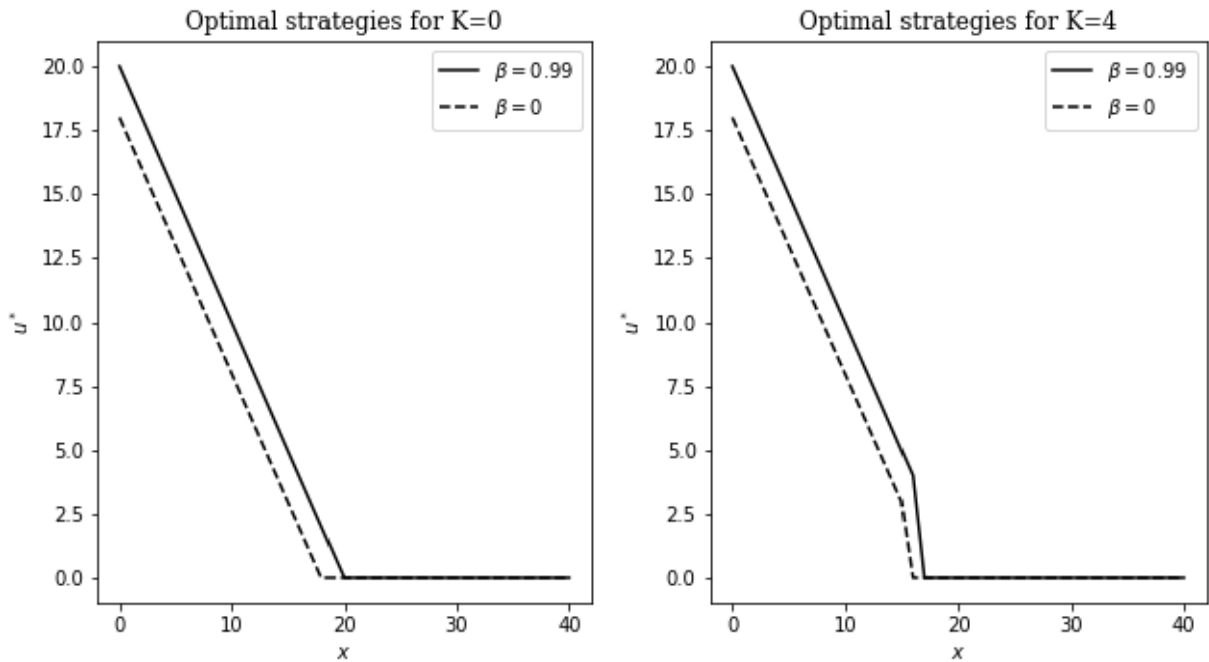


Рис. 2.2.1. Базовые стратегии пополнения (левая часть) и  $(s, S)$ -стратегии (правая часть)

Итак, изучена динамическая стохастическая задача стимулирования для случая нескольких агентов, участвующих в марковской игре. Показано, что для реализации оптимального механизма стимулирования Центр должен решить задачу оптимального управления на максимизацию разности между

своим доходом и суммой компенсационных платежей агентам, сообщить им найденный оптимальный план и покрыть затраты агентов со стимулирующей надбавкой  $\varepsilon$ . С одной стороны, этот результат обобщает (для случая конечных множеств состояний и действий) аналогичный результат из (Рохлин и Угольницкий 2018), с другой стороны - соответствующие результаты для нескольких агентов в статическом случае (Новиков 2007). Основная практическая трудность состоит в том, что реализация данной  $\varepsilon$ -оптимальной стратегии требует знания функций затрат всех агентов и переходного ядра игры (Rokhlin and Ougolnitsky 2019a).

Динамическая модель стимулирования с несколькими агентами в более общей постановке исследована в статье (Rokhlin and Ougolnitsky 2019б). Рассмотрим игру, в которой цель управляющего (Центра) состоит в максимизации дохода большой корпорации посредством распределения средств между  $m$  производителями (агентами). Центр сообщает каждому агенту неотрицательную стимулирующую (мотивирующую) функцию  $c_i(x, a)$ , зависящую от состояния системы  $x$  (например, рыночной цены на производимый товар) и действий производителей  $a = (a^1, \dots, a^m)$  (например, объёмов выпуска). На каждом шаге игры производители независимо выбирают свои действия  $a_t^i$  и получают вознаграждения  $r_i(x_t, a_t) = c_i(x_t, a_t) - g_i(x_t, a_t)$ , где  $g_i$  - производственные издержки. Выигрыш Центра (корпорации) на каждом шаге равен  $f(x_t, a_t) - \sum_{i=1}^m c_i(x_t, a_t)$ , где  $f$  можно трактовать как доход от продаж. Стохастический "закон движения" переменной состояния  $x_t$  определяется переходным ядром  $q$ : неформально,  $P(x_{t+1} \in B | x_t, a_t) = q(B | x_t, a_t)$ .

Выигрыш каждого агента оценивается на бесконечном периоде с общим коэффициентом дисконтирования  $\beta$ . Таким образом,

$$E \sum_{t=0}^{\infty} \beta^t \left( f(x_t, a_t) - \sum_{i=1}^m c_i(x_t, a_t) \right) \rightarrow \max$$

есть целевая функция Центра, а

$$\mathbb{E} \sum_{t=0}^{\infty} \beta^t (c_i(x_t, a_t) - g_i(x_t, a_t)) \rightarrow \max$$

- целевые функционалы агентов. На каждый набор  $(c_1, \dots, c_m)$  множество агентов реагирует равновесием Нэша в соответствующей дисконтированной стохастической игре. Центр проводит оптимизацию по функциям  $c_i$  из некоторого подходящего класса. Из (Новиков и Шохина 2003; Рохлин и Угольницкий 2018) известно, что оптимальный механизм управления Центра состоит в экономической мотивации агентов к выбору стратегий  $a_t^i = \bar{u}_i(x_t)$ , где  $\bar{u} = (\bar{u}_1, \dots, \bar{u}_n)$  - оптимальная стационарная детерминированная марковская стратегия в марковской задаче принятия решения

$$\mathbb{E} \sum_{t=0}^{\infty} \beta^t \left( f(x_t, a_t) - \sum_{i=1}^m g_i(x_t, a_t) \right) \rightarrow \max.$$

При переходе к случаю нескольких агентов возникают некоторые технические трудности, связанные с существованием стационарного равновесия Нэша. Для преодоления этих трудностей модифицируем класс функций стимулирования по сравнению с (Рохлин и Угольницкий 2018), выбирая их непрерывными по действиям. Кроме того, рассмотрим игры с более грубым переходным ядром (He, Sun 2017). Вместо этого можно было использовать другие предположения относительно переходного ядра  $q$ , обеспечивающие существование стационарного марковского равновесия Нэша (Jaśkiewicz, Nowak 2017).

Пусть  $(\Omega, \mathcal{F})$  - измеримое пространство,  $(Y, \tau)$  - топологическое пространство. Функция  $F: \Omega \times Y \rightarrow \mathbb{R}$  называется функцией Каратеодори, если функция  $F(\cdot, y)$   $\mathcal{F}$ -измерима для всех  $y \in Y$  и функция  $F(\omega, \cdot)$   $\tau$ -непрерывна для всех  $\omega \in \Omega$  (Aliprantis, Border 2006; опр. 4.50). Если  $(Y, \tau)$  - сепарабельное метризуемое пространство, то такая функция  $F$  совместно измерима (Aliprantis, Border 2006; лемма 4.51). Обозначим через  $\mathcal{C}_b(\Omega \times Y)$  множество равномерно ограниченных функций Каратеодори. Напомним

также, что стандартное борелевское пространство - это измеримое пространство, изоморфное некоторому борелевскому подмножеству польского пространства (полного сепарабельного метризуемого топологического пространства) (Srivastava 1998).

Пусть  $I = \{1, \dots, m\}$  - множество игроков. Дисконтированная стохастическая игра определяется:

- стандартным борелевским пространством состояний  $(X, \mathcal{B}(X))$  с его борелевской  $\sigma$ -алгеброй  $\mathcal{B}(X)$ ;
- сепарабельными метрическими пространствами  $(\bar{A}_i, \tau_i)$  действий игроков  $i \in I$ ;
- компактно-значными отображениями  $x \mapsto A_i(x) \subset \bar{A}_i$ . Множество  $A_i(x)$  описывает допустимые действия  $i$ -го игрока в состоянии  $x \in X$ ;  
Предполагается, что многозначные отображения  $x \mapsto A_i(x)$  измеримы (Ну, Papageorgiou 1997, гл.2, опр.1.1), т.е.  $\{x \in X: A_i(x) \cap U \neq \emptyset\} \in \mathcal{B}(X)$  для любого открытого множества  $U \subset \bar{A}_i$ ;
- функциями выигрыша  $r_i \in \mathcal{C}_b(X \times \bar{A})$ , где  $\bar{A} = \bar{A}_1 \times \dots \times \bar{A}_m$  снабжено топологией произведения  $\tau$ ;
- вероятностью перехода  $Q(\cdot | \cdot)$  из  $X \times A$  в  $X$  (Bogachev 2007, опр. 10.7.1), что означает:
  - функция  $(x, a) \mapsto Q(B|x, a)$   $\mathcal{B}(X) \times \mathcal{B}(A)$ -измерима для всех  $B \in \mathcal{B}(X)$ ,
  - функция  $B \mapsto Q(B|x, a)$  есть вероятностная мера на  $\mathcal{B}(X)$  для всех  $(x, a) \in X \times A$ .

Предполагается, что функция  $a \mapsto \int w(y)Q(dy|x, a)$  непрерывна для любого  $x \in X$  и любой ограниченной измеримой по Борелю функции  $w$  на  $X$ ;

- коэффициентом дисконтирования  $\beta \in [0, 1)$ .

Считается, что игроки используют стационарные марковские стратегии, которые можно отождествить с переходными вероятностями  $\sigma_i$  из

$X$  в  $\bar{A}_i$  такими, что  $\sigma_i(x)(A_i(x)) = 1$ . Для  $x \in X$  каждый набор  $\sigma = (\sigma_i)_{i \in I}$  порождает вероятностную меру

$$P_{x,\sigma}(dx_0 da_0 \dots dx_t da_t) = \delta_x(dx_0) \prod_{i \in I} \sigma_i(x_0)(da_0^i) \times \dots \times Q(dx_1|x_0, a_0) \dots Q(dx_t|x_{t-1}, a_{t-1}) \prod_{i \in I} \sigma_i(x_t)(da_t^i) \quad (2.2.11)$$

на пространстве последовательностей  $(x_t, a_t)_{t \in \mathbb{Z}_+}$ ,  $(x_t, a_t) \in X \times \bar{A}$ , снабжённом  $\sigma$ -алгеброй произведения.

Ожидаемый дисконтированный выигрыш игрока  $i$  равен

$$J_i(x, \sigma) = E_{x,\sigma} \sum_{t=0}^{\infty} \beta^t r_i(x_t, a_t). \quad (2.2.12)$$

Исход  $\sigma^* = (\sigma_i^*)_{i \in I}$  называется равновесием Нэша, если

$$J_i(x, \sigma^*) \geq J_i(x, \sigma_{-i}^*, \sigma_i), \quad i \in I$$

для любой стратегии  $\sigma_i$  и всех  $x \in X$ . Здесь, как обычно,  $\sigma_{-i} = (\sigma_j)_{j \in (I \setminus i)}$ .

Формально, марковский процесс принятия решений есть стохастическая игра с одним игроком:  $m = 1$ . Опуская индекс “1”, будем в этом случае записывать целевую функцию (2.2.12) в виде

$$J(x, \sigma) = E_{x,\sigma} \sum_{t=0}^{\infty} \beta^t r(x_t, a_t).$$

Обозначим через  $v(x) = \sup_{\sigma} J(x, \sigma)$  соответствующую функцию значения.

*Теорема 2.2.2.* Для описанных марковских процессов принятия решений справедливы следующие утверждения:

(i)  $v$  есть единственное решение уравнения Беллмана

$$v(x) = \sup_{a \in A(x)} \{r(x, a) + \beta \int_X v(y) Q(dy|x, a)\} \quad (2.2.13)$$

в пространстве ограниченных измеримых по Борелю функций на  $X$ .



(ii) существует оптимальная стратегия  $\sigma^*(x)(dy) = \delta_{u^*(x)}(dy)$ , которую можно отождествить с измеримым по Борелю селектором

$$u^*(x) \in \arg \max_{a \in A(x)} \left\{ r(x, a) + \beta \int_X v(y) Q(dy|x, a) \right\}.$$

(iii) если  $\sigma^*$  - оптимальная стратегия:  $v(x) = J(x, \sigma^*)$ , то

$$v(x) = \int_{A(x)} \left( r(x, a) + \beta \int_X v(y) Q(dy|x, a) \right) \sigma^*(x)(da). \quad (2.2.14)$$

Доказательства (i), (ii) можно найти в (Himmelberg et al. 1976; Jaśkiewicz, Nowak 2018). Соотношение (2.2.14) известно как принцип динамического программирования или "фундаментальное уравнение".

Если  $(\sigma_i^*)_{i \in I}$  - равновесие Нэша в игре  $m$  лиц,  $m > 1$ , то каждое  $\sigma_i^*$  есть оптимальное решение задачи оптимизации

$$J_i(x, \sigma_{-i}^*, \sigma_i) \rightarrow \max_{\sigma_i} \quad (2.2.15)$$

Простые вычисления показывают, что

$$J_i(x, \sigma_i, \sigma_{-i}^*) = E_{s, \sigma_i, \sigma_{-i}^*} \sum_{t=0}^{\infty} \beta^t r_i(x_t, a_t^i; \sigma_{-i}^*),$$

где

$$r_i(x, a^i; \sigma_{-i}^*) = \int_{\bar{A}_{-i}} r_i(x, a^i, a^{-i}) \sigma_{-i}^*(da^{-i}) \quad (2.2.16)$$

и математическое ожидание  $E_{s, \sigma_i, \sigma_{-i}^*}$  берётся по мере, порождаемой переходными вероятностями

$$Q_{\sigma_{-i}^*}(B|x, a^i) = \int_{\bar{A}_{-i}} Q(B|x, a^i, a^{-i}) \sigma_{-i}^*(x)(da^{-i}) \quad (2.2.17)$$

и стратегиями  $\sigma_i$  на пространстве последовательностей  $(x_t, a_t^i)_{t \in \mathbb{Z}_+}$ ,  $(x_t, a_t^i) \in X \times \bar{A}_i$  аналогично (2.2.11).

Отсюда следует, что (2.2.15) - это марковский процесс, удовлетворяющий предположениям теоремы 2.2.2. Пусть  $V_{\sigma_{-i}^*}(x) = \sup_{\sigma_i} J(x, \sigma_i, \sigma_{-i}^*)$  - соответствующая функция значения. Поскольку  $\sigma_i^*$  есть

оптимальное решение:  $V_{\sigma_{-i}^*}(x) = J(x, \sigma_i^*, \sigma_{-i}^*)$ , то из принципа оптимальности (2.2.14) и уравнения Беллмана (2.2.13) получаем

$$\begin{aligned} V_{\sigma_{-i}^*}(x) &= \int_{A(x)} (r_i(x, a^i; \sigma_{-i}^*) + \beta \int_X V_{\sigma_{-i}^*}(y) Q_{\sigma_{-i}^*}(dy|x, a^i)) \sigma_i^*(x)(da^i) \\ &\geq r_i(x, a^i; \sigma_{-i}^*) + \beta \int_X V_{\sigma_{-i}^*}(y) Q_{\sigma_{-i}^*}(dy|x, a^i), \quad a^i \in A_i(x). \end{aligned} \quad (2.2.18)$$

Для фиксированного  $x$  и  $(\sigma_i^*)_{i \in I}$  рассмотрим одношаговую игру  $\Gamma(x, \sigma^*)$  на  $A_1(x) \times \dots \times A_m(x)$ , где выигрыш  $i$ -го игрока равен

$$H_i(x, a) = r_i(x, a^i; \sigma_{-i}^*) + \beta \int_X V_{\sigma_{-i}^*}(y) Q_{\sigma_{-i}^*}(dy|x, a^i).$$

Из (2.2.18) следует, что  $(\sigma_i^*(x))_{i \in I}$  есть равновесие Нэша в смешанных стратегиях в игре  $\Gamma(x, \sigma^*)$ . Тем самым доказан следующий хорошо известный результат: аналогичные утверждения приведены в (He, Sun 2017; Jaśkiewicz, Nowak 2018).

*Теорема 2.2.3.* Для дисконтированной стохастической игры с функционалами выигрыша (2.2.12) и стационарными марковскими стратегиями  $(\sigma_i^*)_{i \in I}$  следующие утверждения эквивалентны:

- (i)  $(\sigma_i^*)_{i \in I}$  есть равновесие Нэша;
- (ii) каждое  $\sigma_i^*$ ,  $i \in I$  есть оптимальное решение для марковского процесса с целевым функционалом (2.2.15), переходным ядром (2.2.17) и выигрышами (2.2.16);
- (iii)  $(\sigma_i^*(x))_{i \in I}$  есть равновесие Нэша в смешанных стратегиях в игре  $\Gamma(x, \sigma^*)$  для всех  $x \in X$ .

Некоторые дополнительные предположения обеспечивают существование стационарного марковского равновесия Нэша (Jaśkiewicz, Nowak 2018). Мы опираемся на результаты (He, Sun 2017). Допустим, что:

(А)  $Q(\cdot | x, a)$  абсолютно непрерывна относительно вероятностной меры  $\lambda$  на  $(X, \mathcal{B}(X))$  для всех  $(x, a) \in X \times \bar{A}$ . Соответствующая плотность вероятности  $(x, a, y) \mapsto q(y|x, a)$  предполагается  $\mathcal{B}(X) \times \mathcal{B}(\bar{A}) \times \mathcal{B}(X)$ -измеримой. Здесь  $\mathcal{B}(\bar{A})$  - борелевская  $\sigma$ -алгебра на топологическом пространстве  $(\bar{A}, \tau)$ .

(Б) для всех  $x \in X$  отображение  $a \mapsto Q(\cdot | x, a)$  непрерывно по норме полной вариации:

$$\lim_{a_n \rightarrow a} \sup_{B \in \mathcal{B}(X)} |Q(B|x, a_n) - Q(B|x, a)| = 0.$$

Рассмотрим вероятностное пространство  $(\Omega, \mathcal{F}, P)$ . Под- $\sigma$ -алгебра  $\mathcal{G} \subset \mathcal{F}$  называется *множественно более грубой*, чем  $\mathcal{F}$ , если для всех  $D \in \mathcal{F}$  с  $P(D) > 0$  существует множество  $D_0 \subset D$ ,  $D_0 \in \mathcal{G}$  такое, что  $P(D_0 \Delta D_1) > 0$  для любого  $D_1 \in \{D' \cap D : D' \in \mathcal{G}\}$  (эквивалентно,  $\mathcal{F}$  не имеет  $\mathcal{G}$ -атома под  $P$ : терминология взята из (He, Sun 2017, 2018)). Стохастическая игра имеет более грубое переходное ядро, если существует  $\sigma$ -алгебра  $\mathcal{G} \in \mathcal{B}(X)$  такая, что  $q(\cdot | x, a)$   $\mathcal{G}$ -измерима для всех  $(x, a) \in X \times \bar{A}$  и  $\mathcal{G}$  множественно более грубая, чем  $\mathcal{B}(X)$ .

*Теорема 2.2.4.* Пусть выполнены предположения (А), (Б) и игра имеет более грубое переходное ядро. Тогда существует равновесие Нэша  $(\sigma_i^*)_{i \in I}$ .

В (He, Sun 2018) теорема 2.2.4 сформулирована и доказана при более общей гипотезе существования *разложимо более грубого* переходного ядра.

Пусть выполнены условия теоремы 2.2.4. Формальное описание рассматриваемой игры имеет следующий вид.

(I) Центр выбирает набор  $c = (c_i)_{i \in I}$  неотрицательных функций стимулирования  $c_i \in \mathcal{C}_b(X \times \bar{A})$ .

(II) Множество  $m \geq 1$  агентов с функциями вознаграждения

$$r_i(x, a) = c_i(x, a) - g_i(x, a), \quad 0 \leq g_i \in \mathcal{C}_b(X \times \bar{A})$$

и выигрышами

$$J_i(x, \sigma, c) = E_{x, \sigma} \sum_{t=0}^{\infty} \beta^t (c_i(x, a) - g_i(x, a))$$

разыгрывает описанную стохастическую игру, исход которой есть (стационарное марковское) равновесие Нэша  $(\sigma_i^*)_{i \in I}$ .

(III) Центр получает выигрыш

$$J_L(x, \sigma^*, c) = E_{x, \sigma^*} \sum_{t=0}^{\infty} \beta^t \left( f(x_t, a_t) - \sum_{i=1}^N c_i(x_t, a_t) \right),$$

где  $f \in \mathcal{C}_b(X \times \bar{A})$ .

Обозначим через  $T(c)$  множество равновесий Нэша для данной функции стимулирования  $c$ . Цель Центра состоит в максимизации своего выигрыша для наихудшего из равновесий Нэша:

$$G(x, c) = \inf_{\sigma^* \in T(c)} J_L(x, \sigma^*, c).$$

Задача такого типа известна как слабая игра Штакельберга (Breton et al. 1988). Назовём

$$V_L(x) = \sup\{G(x, c) : c_i \in \mathcal{C}_b(X \times \bar{A}), i \in I\}$$

значением Центра. Набор  $\bar{c}_\varepsilon$  называется  $\varepsilon$ -решением Штакельберга, если

$$V_L(x) - \varepsilon \leq G(x, \bar{c}_\varepsilon), \quad x \in X.$$

Пара  $(\bar{c}_\varepsilon, \sigma^*)$ ,  $\sigma^* \in T(\bar{c}_\varepsilon)$  называется  $\varepsilon$ -равновесием Штакельберга.

Рассмотрим вспомогательный марковский процесс:

$$J(x, \sigma) = E_{x, \sigma} \sum_{t=0}^{\infty} \beta^t \left( f(x_t, a_t) - \sum_{i=1}^N g_i(x_t, a_t) \right) \rightarrow \max_{(\sigma_i)_{i \in I}} \quad (2.2.19)$$

Это задача Центра, который проводит максимизацию по наборам  $\sigma = (\sigma_i)_{i \in I}$ . Согласно теореме 2.2.2 эта задача имеет оптимальное решение вида  $\bar{\sigma} = (\delta_{\bar{u}_i(x)})_{i \in I}$ , которое можно отождествить с измеримым по Борелю селектором

$$(\bar{u}_i(x))_{i \in I} \in \arg \max_{a \in A(x)} \left\{ f(x, a) - \sum_{i=1}^m g_i(x, a) + \beta \int_X V(y) Q(dy|x, a) \right\}.$$

Здесь  $V(x) = \sup_{\sigma} J(x, \sigma)$  - функция значения задачи (2.2.19). Оно совпадает с оптимальным выигрышем Центра при производстве без участия агентов.

Будем считать, что если объём выпуска агента равен нулю, то у него нет производственных издержек:

$$g_i(x, 0, a^-) = 0.$$

*Теорема 2.2.5.* В предположениях теоремы 2.2.4 справедливы следующие утверждения:

(i)  $V_L(x) = V(x)$ ;

(ii) Набор  $\bar{c}_\varepsilon = (c_{i,\varepsilon})_{i \in I}$ ,

$$c_{i,\varepsilon}(x, a) = g_i(x, a) + \frac{\varepsilon}{m}(1 - \beta)(1 - |\zeta^i - \bar{u}_i(x)|)^+, \quad y^+ := \max\{y, 0\}$$

есть  $\varepsilon$ -решение Штакельберга. Соответствующее равновесие Нэша единственно:  $T(\bar{c}_\varepsilon) = \{(\delta_{\bar{u}_i(x)})_{i \in I}\}$ .

Для случая единственного агента ( $m = 1$ ) аналогичные результаты получены в (Рохлин, Угольницкий 2018). Однако, стимулирующая надбавка в (Рохлин, Угольницкий 2018) разрывна в  $a$ :

$$c_\varepsilon(x, a) - g(x, a) = \varepsilon(1 - \beta)I_{\{a = \bar{u}(x)\}} \quad (2.2.20)$$

(индекс “1” опущен). Далее, был доказан аналог теоремы 2.2.3 либо при более сильном предположении о том, что во вспомогательной задаче (2.2.19) существует непрерывная оптимальная стратегия  $\bar{u}$  (кстати, это требует наличия топологии в пространстве состояний), либо при использовании понятия  $(\varepsilon, \eta)$ -решения Штакельберга и класса универсально измеримых функций симулирования  $c$ . Принятая сейчас гипотеза о том, что функции затрат  $g_i$ , функция дохода  $f$  и функции стимулирования  $c_i$  принадлежат классу  $\mathcal{C}_b(X \times \bar{A})$ , ведёт к более естественным и простым результатам. С другой стороны, в (Рохлин, Угольницкий 2018) показано, что в случае стимулирующей надбавки (2.2.20) агент может отклониться от  $\bar{u}$  только за счёт “больших” потерь. Таким образом, надбавка данного вида имеет свои преимущества.

В случае конечных пространств состояний и действий, где не возникает трудностей с точки зрения теории меры, теорема 2.2.5 остаётся верной для разрывной стимулирующей надбавки

$$c_{i,\varepsilon}(x, a) - g_i(x, a) = \frac{\varepsilon}{m} (1 - \beta) I_{\{a^i = \bar{u}_i(x)\}}.$$

В (He, Sun 2017) отмечено, что из теоремы 2.2.4 следует существование стационарного коррелированного марковского равновесия Нэша при условиях (А), (Б). Придерживаясь в основном (He, Sun 2017), кратко изложим этот момент следующим образом. Рассмотрим расширенное пространство состояний  $X' = X \times [0,1]$ , снабжённое  $\sigma$ -алгеброй произведения  $\mathcal{B}(X) \otimes \mathcal{B}([0,1])$  и мерой произведения  $\lambda' = \lambda \otimes \eta$ , где  $\eta$  - мера Лебега на  $\mathcal{B}([0,1])$ . В соответствующей модели на каждом шаге все игроки получают сигнал  $z_t \in [0,1]$ . Эти сигналы представляют собой независимые случайные переменные, равномерно распределённые на  $[0,1]$ . Переходная вероятность принимает вид

$$Q'(B \times C | x, z, a) = Q(B | x, a) \eta(C).$$

Для плотности  $q'(\cdot | x, z, a)$  of  $Q'(\cdot | x, z, a)$  относительно  $\lambda'$  имеем

$$q'(y, u | x, z, a) = q(y | x, a).$$

Функция  $q'(\cdot | x, z, a) = q(\cdot | x, a)$  измерима относительно  $\sigma$ -алгебры  $\mathcal{G}' = \mathcal{B}(X) \otimes \{\emptyset, L\}$  для всех  $x, a$ , и эта  $\sigma$ -алгебра множественно более грубая, чем  $\mathcal{B}(X) \otimes \mathcal{B}([0,1])$ . Таким образом, новая модель удовлетворяет условию более грубого переходного ядра и обладает стационарным марковским равновесием Нэша. По определению, это означает существование коррелированного равновесия в исходной модели, удовлетворяющей предположениям (А), (Б).

Итак, если в схеме (I) – (III), описывающей игру Штакельберга, на этапе (II) заменить равновесие Нэша коррелированным равновесием, то все утверждения теоремы 2.2.5 остаются верными, если отказаться от условия

более грубого переходного ядра. Доказательство фактически не меняется, поскольку оно не зависит от пространства состояний.

Итак, результаты (Рохлин и Угольницкий 2018) распространены на случай нескольких агентов в более общей стохастической постановке, чем в (Rokhlin and Ougolnitsky 2019a). Сначала Центр не рассчитывает на участие агентов и относит их затраты на себя. После определения оптимальных стратегий производства из соответствующего марковского процесса Центр предлагает агентам экономическую мотивацию, обеспечивающую реализацию ими этих стратегий. В дальнейшем предполагается рассмотреть случай нескольких Центров и постановки всех задач в непрерывном времени (Rokhlin and Ougolnitsky 2019b).

В статье (Ougolnitsky et al. 2017) исследована двухуровневая модель оптимальной эксплуатации рыбной популяции при параметрической неопределённости. Изучен случай двух иерархически упорядоченных субъектов управления: государственный контрольный орган (Центр) и рыболовное предприятие (агент). Центр считается дальновидным, а агент - близоруким. Динамика рыбной популяции описывается нелинейным разностным уравнением. Предполагается, что это уравнение и функция выигрыша агента зависят от неизвестных Центру параметров. Доказано, что функция значения этой задачи удовлетворяет уравнению Айзекса-Беллмана и описывает оптимальную налоговую стратегию Центра. Приведены два иллюстративных примера.

Математическое моделирование использования возобновляемых природных ресурсов (например, рыбы) активно развивается начиная со второй половины прошлого века. Поскольку неограниченный доступ к рыбным ресурсам может привести к перелову, то исследования в основном связаны с различными механизмами регулирования: назначением квот, ограничением суммарного вылова, единоличными правами собственности и т.д. (Arnason 2009; Clark 2006). В настоящей работе рассматривается только схема налогообложения, в которой предполагается, что агенты близоруки, а

Центр стремится максимизировать общую дисконтированную величину налоговых платежей. Заметим, что это нетрадиционный критерий. Идея состоит в том, что максимальный дисконтированный налоговый доход на бесконечном периоде может обеспечить лишь экологически рациональная стратегия природопользования, а близоруких агентов можно побудить к её реализации посредством специального налогового механизма.

Учитывать неопределённость в моделях можно различными способами: стохастические эффекты, неопределённость параметров или модели в целом, величины запаса биомассы или вылова (Charles 1998; Walters 1978). Имеется обширная литература, посвящённая анализу воздействия неопределённости на стратегии эксплуатации биоресурсов и соответствующую динамику биомассы (Clark, Kirkwood 1986; Charles 1998; Mitra, Roy 2006; Olson, Roy 2000; Sethi et al. 2005). Ключевой вопрос здесь состоит в оптимальности природоохранных стратегий. Вообще говоря, неопределённость - один из важнейших факторов, приводящих к истощению рыбных популяций.

Будем предполагать, что точные значения параметров уравнения динамики биомассы и функций выигрыша агентов неизвестны. Используем грубый детерминированный (минимаксный) подход, при котором оптимизация проводится для наихудших сценариев. Интерес к таким моделям рыболовства вырос в последнем десятилетии (Anderies J.M. et al. 2007; Doyen, Vene 2003; Woodward, Shaw 2008; Херапачеас, Росета-Палма 2013). Смысл использования грубых методов управления заключается в получении решений, устойчивых к неточностям модели.

С другой стороны, настоящая работа развивает идеи (Угольницкий 2016; П'ичев et al. 2000; Rokhlin, Usov 2017), где рассмотрены двухуровневые модели рыболовства с дальновидным Центром и близоруким агентом. Отметим, что фактически речь идёт об "обобщённом" агенте, включающем нескольких природопользователей.

Модель оптимальной эксплуатации рыбной популяции строится в рамках концепции иерархически управляемых динамических систем (см.



пункт 1.1). Под Центром понимается государственный регулирующий орган, под агентом - рыболовное предприятие. Принуждение реализуется посредством квот на вылов рыбы, а побуждение - через налоги или тарифы. Сравнительный анализ этих методов управления проводится во многих работах по экономике природопользования (Long 2010).

Функционирование системы определяется следующими правилами.

1. Оба субъекта управления стремятся максимизировать свои целевые функционалы.

2. Агент близорук, т.е. максимизирует свой выигрыш на одном шаге, в то время как дальновидный Центр максимизирует свой дисконтированный доход на бесконечном горизонте. Такая цель может быть достигнута только посредством рационального природопользования с учётом экологических соображений.

3. Центр первым выбирает свою стратегию и сообщает её агенту.

4. Предполагается, что некоторые параметры модели неизвестны обоим игрокам.

5. Динамика рыбной популяции описывается нелинейным разностным уравнением.

6. В качестве метода управления Центр использует побуждение (устанавливает налог).

Предположим, что изменение во времени биомассы популяции  $X_t$  задаётся разностным уравнением

$$X_{t+1} = b(X_t - u_t, \alpha_t), \quad t \in \mathbb{Z}_+ = \{0, 1, \dots\},$$

где  $(\alpha_t)_{t \geq 0}$  - произвольная последовательность со значениями на компактном множестве  $A \subset \mathbb{R}$ . Оно отражает неполноту информации о динамике системы. Функция  $b: [0, 1] \times A \rightarrow [0, 1]$  непрерывна, а интервал  $[0, 1]$  инвариантен относительно отображения  $x \mapsto b(x, \alpha)$  для всех  $\alpha \in A$ .

Очевидно, в последнем предположении инвариантный интервал  $[0, 1]$  можно масштабировать произвольно: выбор верхней границы условен.

Последовательность  $(u_t)_{t \geq 0}$  со значениями в  $[0, X_t]$  описывает "сбор урожая" (величину вылова). Близорукий агент в каждый момент времени максимизирует свою целевую функцию

$$p(u_t, \gamma_t) - q_t u_t$$

по переменной  $u_t \in [0, X_t]$ . Здесь  $q_t \in [0, 1]$  (как и в случае  $x$ , выбор верхней границы здесь произволен) - это устанавливаемая Центром налоговая ставка, причём параметр  $\gamma_t \in \Gamma$  целевой функции агента неизвестен Центру.

Пусть множество  $\Gamma \subset \mathbb{R}$  компактно, функция  $p: [0, 1] \times \Gamma \mapsto \mathbb{R}$  непрерывна, а функция  $v \mapsto p(v, \gamma)$  строго вогнута на  $[0, 1]$ . Тогда в силу теоремы Берга о максимуме (Sundaram 1996) единственная точка максимума  $u(x, \gamma, q)$  функции

$$v \mapsto p(v, \gamma) - qv$$

на отрезке  $[0, x]$  непрерывна по всем переменным, т.е. как отображение

$$u: [0, 1] \times \Gamma \times [0, 1] \mapsto [0, 1].$$

Управление Центра - это последовательность  $(q_t)_{t \in \mathbb{Z}_+}$  функций  $q_t: [0, 1] \mapsto [0, 1]$ . Налог  $q_t(X_t)$  зависит только от состояния системы, поскольку последовательности  $\alpha_t, \gamma_t$  неизвестны Центру. Если стратегия агента выбирается так, как определено выше, то уравнение динамики принимает вид

$$X_{t+1} = b(X_t - u(X_t, q_t(X_t), \gamma_t), \alpha_t), \quad t \in \mathbb{Z}_+, \quad X_0 = x. \quad (2.2.21)$$

Последовательность  $X$ , определяемую (2.2.21), обозначим  $X^{x, q, \alpha, \gamma}$ . Цель Центра заключается в максимизации дисконтированной суммы налога на бесконечном горизонте планирования, что решает проблему сохранения экологического равновесия:

$$J(x, q, \alpha, \gamma) = \sum_{t=0}^{\infty} \beta^t u(X_t^{x, q, \alpha, \gamma}, q_t(X_t^{x, q, \alpha, \gamma}), \gamma_t) q_t(X_t^{x, q, \alpha, \gamma}) \rightarrow \max.$$

Введём соответствующую функцию значения:

$$v(x) = \sup_{(q_t)_{t \geq 0}} \inf_{(\alpha_t, \gamma_t)_{t \geq 0}} J(x, q, \alpha, \gamma).$$

Заметим, что сначала задаётся последовательность функций  $q = (q_t)_{t \geq 0}$ ,  $q_t: [0, 1] \mapsto [0, 1]$ , затем вычисляется инфимум по всем последовательностям

$(\alpha, \gamma) = (\alpha_t, \gamma_t)_{t \geq 0}$  со значениями на компактном множестве  $A \times \Gamma$ , и наконец берётся супремум по всем  $q$ .

Значение  $v(x)$  можно трактовать как нижнюю цену игры с нулевой суммой между Центром и Природой, которая выбирает наихудшие последовательности  $(\alpha, \gamma)$ . С точки зрения Центра этот критерий чрезвычайно консервативен, поскольку на самом деле Природа не заинтересована играть против него. Альтернативный подход состоит в вычислении инфимума по фиксированному семейству распределений на пространстве  $(A \times \Gamma)^{\mathbb{Z}_+}$ .

*Теорема 2.2.6* (Ougolnitsky et al. 2017). Функция  $v$  есть единственное решение уравнения Айзекса-Беллмана

$$v(x) = \sup_{q \in [0,1]} \inf_{(\alpha, \gamma) \in A \times \Gamma} \{u(x, q, \gamma)q + \beta v(b(x - u(x, q, \gamma), \alpha))\}, \quad x \in [0,1]. \quad (2.2.22)$$

Для любой функции  $q^*: [0,1] \mapsto [0,1]$ , удовлетворяющей условию

$$v(x) = \inf_{(\alpha, \gamma) \in A \times \Gamma} \{u(x, q^*(x), \gamma)q^*(x) + \beta v(b(x - u(x, q^*(x), \gamma), \alpha))\},$$

стационарная последовательность управлений  $q_n = q^*$  оптимальна:

$$v(x) = \inf_{(\alpha_t, \gamma_t)_{t \geq 0}} \sum_{t=0}^{\infty} \beta^t u(X_t^{x, q^*, \alpha, \gamma}, q_t(X_t^{x, q^*, \alpha, \gamma}), \gamma_t) q_t^*(X_t^{x, q^*, \alpha, \gamma}).$$

*Пример 2.2.1.* Предположим, что

$$p(u) = \ln u, \quad b = (x - u)^\alpha, \quad \alpha \in [\underline{\alpha}, \bar{\alpha}], \quad 0 < \underline{\alpha} \leq \bar{\alpha} < 1.$$

Таким образом, неопределённость присутствует только в уравнении динамики, в то время как целевая функция агента известна. Перепишем уравнение (2.2.22):

$$v(x) = \sup_{q \in [0,1]} \inf_{\alpha \in [\underline{\alpha}, \bar{\alpha}]} \{q \min\{1/q, x\} + \beta v((x - \min\{1/q, x\})^\alpha)\}.$$

Пусть  $x \in [0,1)$ , тогда  $1/q > x$ ,

$$v(x) = \sup_{q \in [0,1]} \inf_{\alpha \in [\underline{\alpha}, \bar{\alpha}]} \{qx + \beta v(0)\} = x, \quad (2.2.23)$$

и максимум достигается при  $q^*(x) = 1$ . Значения  $\alpha^*(x)$  можно выбирать произвольно.

Для  $x \in [1,2)$  имеем  $x - \min\{1/q, x\} \leq 1$ . Следовательно,  $\alpha^*(x) = \bar{\alpha}$ ,

$$\begin{aligned} v(x) &= \sup_{q \in [0,1]} \{\min\{1, qx\} + \beta(x - \min\{1/q, x\})\bar{\alpha}\} \\ &= \min\{1, x\} + \beta(x - \min\{1, x\})\bar{\alpha} = 1 + \beta(x - 1)\bar{\alpha}, \end{aligned} \quad (2.2.24)$$

и  $q^*(x) = 1$ . Пусть  $x \in [n, n + 1)$ ,  $n \geq 2$ . Обозначим

$$I(x, q) = \inf_{\alpha \in [\underline{\alpha}, \bar{\alpha}]} \{q \min\{1/q, x\} + \beta v((x - \min\{\frac{1}{q}, x\})^\alpha)\},$$

и представим  $v(x)$  как

$$v(x) = \max \left\{ \sup_{q \in [0, \frac{1}{x}]} I(x, q), \sup_{q \in [\frac{1}{x}, \frac{1}{x-1}]} I(x, q), \sup_{q \in [\frac{1}{x-1}, 1]} I(x, q) \right\}.$$

Простые вычисления дают

$$\begin{aligned} \sup_{q \in [0, \frac{1}{x}]} I(x, q) &= \sup_{q \in [0, \frac{1}{x}]} \inf_{\alpha \in [\underline{\alpha}, \bar{\alpha}]} \{qx + \beta v(0)\} = 1, \\ \sup_{q \in [\frac{1}{x}, \frac{1}{x-1}]} I(x, q) &= \sup_{q \in [\frac{1}{x}, \frac{1}{x-1}]} \inf_{\alpha \in [\underline{\alpha}, \bar{\alpha}]} \{1 + \beta v((x - 1/q)^\alpha)\} = \sup_{q \in [\frac{1}{x}, \frac{1}{x-1}]} \{1 + \beta v((x - 1/q)^\alpha)\} \\ &= 1 + \beta v(1) = 1 + \beta. \end{aligned}$$

Здесь дважды использован тот факт, что функция  $v$  неубывающая и  $v(1) = 1$ . Далее,

$$\begin{aligned} \sup_{q \in [\frac{1}{x-1}, 1]} I(x, q) &= \sup_{q \in [\frac{1}{x-1}, 1]} \inf_{\alpha \in [\underline{\alpha}, \bar{\alpha}]} \{1 + \beta v((x - 1/q)^\alpha)\} = \sup_{q \in [\frac{1}{x-1}, 1]} \{1 + \beta v((x - 1/q)^\alpha)\} \\ &= 1 + \beta v((x - 1)^\alpha). \end{aligned}$$

Итак,

$$v(x) = \max\{1, 1 + \beta, 1 + \beta v((x - 1)^\alpha)\} = 1 + \beta v((x - 1)^\alpha), \quad x \in [n, n + 1), \quad (2.2.25)$$

и  $q^*(x) = 1$ .

Таким образом, функция значения вычисляется рекуррентно по формулам (2.2.23)-(2.2.25). Оптимальный налог не зависит от  $x$ :  $q^*(x) = 1$ , и наихудшая реакция Природы

$$\alpha^*(x) = \begin{cases} \bar{\alpha}, & x \in [1,2], \\ \underline{\alpha}, & x \geq 2. \end{cases}$$

Заметим, что стратегия агента имеет вид  $u(x, \alpha) = \min\{1, x\}$ , поэтому для  $x \leq 1$  вся рыба вылавливается за один шаг, и  $\alpha^*(x)$  не играет никакой роли.

Соответствующее уравнение динамики имеет вид

$$x_{t+1} = \begin{cases} 0, & x_t \in [0,1], \\ (x_t - 1)\bar{\alpha}, & x_t \in [1,2], \\ (x_t - 1)\underline{\alpha}, & x_t \geq 2. \end{cases}$$

Непосредственно проверяется, что  $x_t$  становится равным нулю за конечное число шагов. В частности, это имеет место при  $\underline{\alpha} = \bar{\alpha}$ .

Заметим, что если агент не близорук (например, владеет водным объектом) и решает задачу оптимального управления

$$x_{t+1} = (x_t - u_t)^\alpha, \quad t \geq 0, \\ \sum_{t=0}^{\infty} \beta^t \ln u_t \rightarrow \max_{u_t \in [0, x_t]},$$

то порождаемая оптимальным управлением траектория асимптотически стремится к  $b(\hat{x})$ , где  $\hat{x}$  - единственный корень уравнения

$$\beta b'(x) = 1 \tag{2.2.26}$$

(Иlichev et al. 2000). В этом случае

$$x_t \rightarrow (\beta\alpha)^{\frac{\alpha}{1-\alpha}} > 0, \quad t \rightarrow \infty.$$

приватизация водного объекта выгоднее налогообложения.

Описанный в примере 2.2.1 экологически нежелательный исход лучше всего объяснить нарушением условия

$$p'(0) \leq 1, \tag{2.2.27}$$

которое теперь будем предполагать выполненным.

Если параметр  $\gamma$  фиксирован и известен Центру, т.е. множество  $\Gamma$  состоит из единственной точки, то существует взаимно-однозначное соответствие между налогом  $q \in [p'(x), p'(0)]$  и выловом  $u(x, q) \in [0, x]$ :  $p'(u(x, q)) = q$ . Нет смысла выбирать  $q < p'(x)$ , поскольку вылов в этом случае тот же, что и при  $q = p'(x)$ , а доход Центра меньше. Бессмысленно также рассматривать  $q > p'(0)$ , поскольку уже при  $q = p'(0)$  вылов равен нулю.

Таким образом, условие (2.2.27) означает, что Центр выбирает величину вылова напрямую, и задача оптимального управления принимает вид

$$\inf_{(\alpha_t)_{t \geq 0}} \sum_{t=0}^{\infty} p'(u_t(X_t))u_t(X_t) \rightarrow \max,$$

$$X_{t+1} = b(X_t - u_t, \alpha_t), \quad u_t \in [0, X_t], \quad t \in \mathbb{Z}_+, \quad X_0 = x.$$

Максимизация проводится по всем последовательностям функций  $u_t: [0, 1] \mapsto [0, 1]$ , удовлетворяющих условию  $u_t(x) \leq x$ .

Обычно функция  $p'(u)u$  наследует вогнутость функции  $p$ . Например,

$$p = \ln(1 + u), \quad up'(u) = \frac{u}{1+u},$$

$$p = \frac{1}{\gamma}(1 + u)^\gamma, \quad \gamma \in (0, 1), \quad up'(u) = \frac{u}{(1+u)^{1-\gamma}}.$$

Следовательно, если  $b$  вогнута, по крайней мере для фиксированного  $\alpha$  оптимальная траектория будет сходиться к значению  $b(\hat{x})$ , где  $\hat{x}$  есть решение (2.2.26). Таким образом, в большинстве случаев описанный метод оптимального налогообложения не нарушает условия гомеостаза.

**Пример 2.2.2.** Пусть функция дохода агента  $p: [0, 1] \times \Gamma \mapsto \mathbb{R}$  имеет вид

$$p(u, \gamma) = \frac{1}{\gamma}u^\gamma, \quad \gamma \in (0, 1).$$

Тогда его целевая функция есть

$$\frac{1}{\gamma_t}u_t^{\gamma_t} - q_t u_t \rightarrow \max,$$

где максимум ищется по  $u_t \in [0, X_t]$ . Решение этой простой задачи оптимизации:

$$u_t^* = \begin{cases} q_t^{1/(\gamma-1)}, & \text{если } x_t^{\gamma-1} < q_t, \\ u_t^* = x_t, & \text{иначе.} \end{cases}$$

Для моделей Ферхюльста-Пирла  $b = \alpha x(1 - x/K)$  ( $K$  - ёмкость среды) и Риккера  $b = \alpha x e^{-x}$  уравнение Айзекса-Беллмана (2.2.22) решалось численно для следующего набора входных данных:  $K = 10$ ;  $\beta = 0.6$ ;  $\gamma \in [0.2; 0.8]$ ;  $\alpha \in [0.2; 0.8]$ . Наихудшие реакции Природы суть  $\alpha = \underline{\alpha} = 0.2$  и  $\gamma = \underline{\gamma} = 0.2$ . Оптимальные налоги представлены на рис. 2.2.3 и 2.2.4.

Верхняя, средняя и нижняя кривые отвечают значениям  $\beta = 0.9$ ,  $\beta = 0.6$  и  $\beta = 0.1$  соответственно. Заметим, что оптимальный налог в модели Ферхюльста-Пирла всегда выше, чем в модели Риккера, и последний намного менее чувствителен к значению дисконта, чем первый. Оптимальные налоги - убывающие по  $x$  функции в соответствии с экономической интуицией, поскольку ценность рыбы убывает с ростом её биомассы.

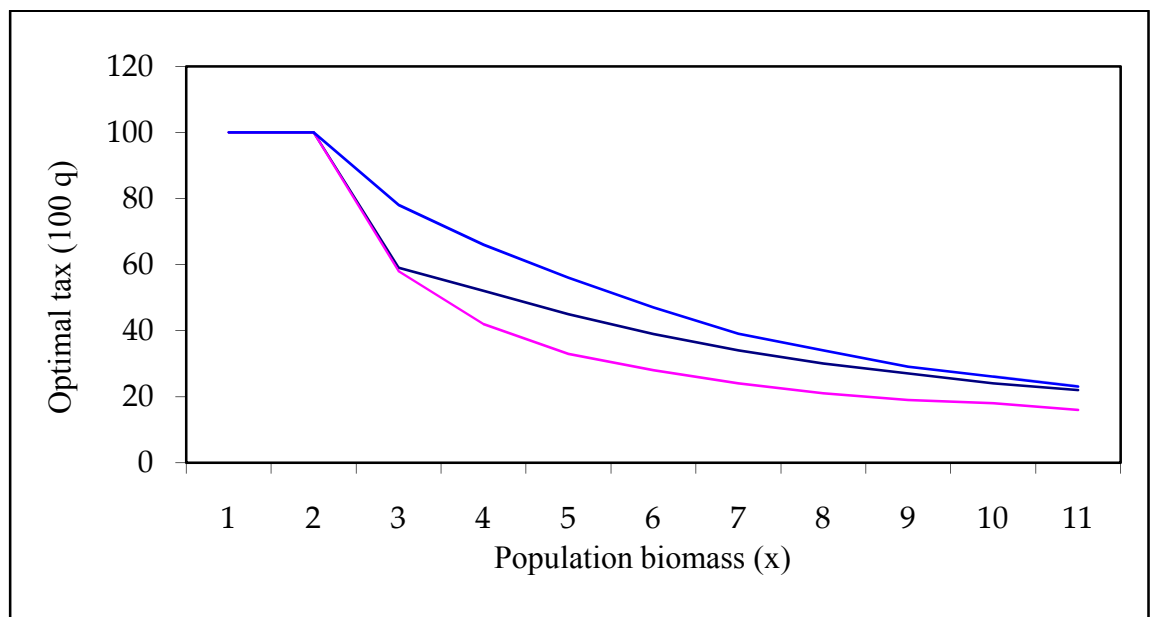


Рис. 2.2.3. Зависимость оптимального налога от величины биомассы в модели Ферхюльста-Пирла

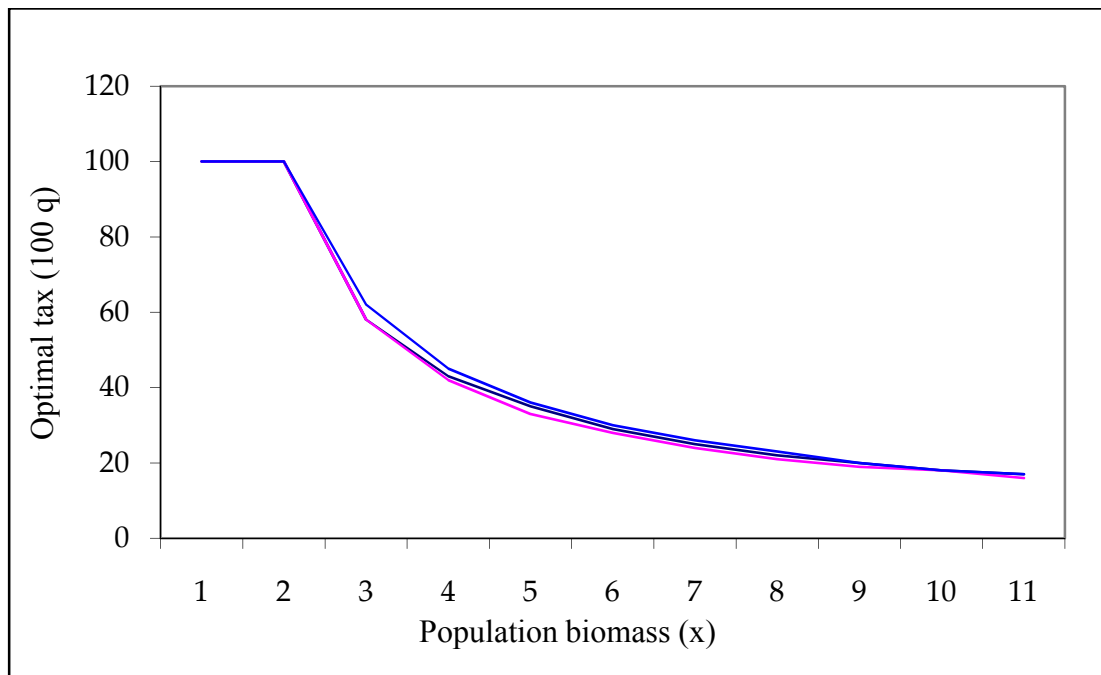


Рис. 2.2.4. Зависимость оптимального налога от величины биомассы в модели Риккера

Итак, исследована двухуровневая модель типа "Центр - агент" оптимальной эксплуатации рыбных ресурсов для случая близорукого агента. Уравнение динамики и целевые функции агентов содержат неопределённые по Найту параметры. Функция значения охарактеризована как единственное решение уравнения Айзекса-Беллмана. Доказана теорема о виде оптимальной налоговой стратегии Центра, которую можно вычислить с помощью стандартных алгоритмов динамического программирования.

Приведены два примера. В первом из них неопределённость присутствует только в уравнении динамики. Выведена рекуррентная формула для функции Беллмана, явно найдены оптимальные стратегии агента, наихудшая реакция Природы и траектория развития популяции. Показано, что при выбранных входных данных популяция вымирает за конечное время. Однако, этот неприятный эффект исчезает при приватизации водного объекта агентом, поскольку тогда он перестаёт быть близоруким. Таким образом, в данном примере приватизация лучше налогообложения с экологической точки зрения.



Во втором примере уравнение Айзекса-Беллмана решалось численно. Проведён краткий сравнительный анализ результатов счёта для моделей популяционной динамики Ферхюльста-Пирла и Риккера.

Итак, предложен механизм оптимальной эксплуатации рыбных ресурсов с учётом экологических требований и методы его реализации (Ougolnitsky et al. 2017).

### 2.3. Прикладные СОЧИ-модели

Статические модели согласования общественных и частных интересов (СОЧИ-модели) описаны в работах (Горбанёва 2019; Горбанёва и Угольницкий 2014, 2015, 2016; Горбанёва и др. 2018; Gorbaneva and Ougolnitsky 2013, 2015, 2017). В СОЧИ-моделях каждый активный агент распределяет свой ресурс между инвестициями в частную деятельность и создание некоторого общественного (совместного) продукта. Соответственно, выигрыш агента представляет собой сумму доходов от частной деятельности и участия в доходах от использования общественного продукта, и агенты максимизируют этот выигрыш путём выбора оптимального распределения ресурса. Вводится функция общественного благосостояния, равная сумме функций выигрыша агентов. Тогда СОЧИ-модель системно согласована, если значение функции общественного благосостояния в наихудшем из равновесий Нэша в игре агентов совпадает с глобально максимальным значением. Показано, что системная согласованность достигается только в том случае, когда агенты относятся к одному из двух типов: индивидуалисты (тратят весь ресурс на поддержку частной деятельности) и коллективисты (тратят весь ресурс на производство общественного продукта).

Для управления системной согласованностью вводится Центр, максимизирующий общественное благосостояние. Получен ряд результатов относительно эффективности применяемых Центром механизмов управления согласованием интересов, в том числе с учётом административной и экономической коррупции. Рассмотрено приложение статических СОЧИ-моделей к решению задач управления развитием территориальных социально-экономических систем.

По аналогии с идеей обобщения статических моделей стимулирования (пункты 2.1, 2.2), в настоящем параграфе предлагается динамическое

обобщение СОЧИ-моделей, при этом основным инструментом исследования пока выступает имитационное моделирование.

В статье (Угольницкий и Усов 2016) исследуются динамические теоретико-игровые модели инновационного развития корпорации. Эти модели основаны на согласовании частных и общественных интересов агентов. Предполагается, что структура интересов каждого агента включает как частную (личные интересы), так и общественную (интересы компании в целом, в первую очередь отражающие необходимость ее инновационного развития) составляющие. Агенты могут делить персональные ресурсы между этими направлениями. Динамика системы описывается разностным уравнением. При исследовании предложенной модели инновационного развития используется компьютерная имитация и метод перебора областей допустимых управлений субъектов с некоторым шагом. Основной вклад работы - сравнительный анализ эффективности методов иерархического управления для информационных регламентов прямых и обратных игр Штакельберга (игр Гермейера  $\Gamma_1$  и  $\Gamma_2$ ) при принуждении/побуждении (четыре регламента) с помощью индексов системной согласованности. Предлагаемая модель носит универсальный характер и может быть использована для научно обоснованной поддержки программ инновационного развития компаний всех отраслей экономики. Специфика конкретной компании учитывается в ходе идентификации модели, которая представляет собой отдельную сложную задачу и предполагает анализ системы официальной отчетности компании и применение экспертных оценок ее специалистов. Приняты следующие предположения относительно информационного регламента иерархической игры: все игроки используют программные стратегии; ведущий выбирает и сообщает ведомым экономические либо административные управления, которые могут быть только функциями времени (прямые игры Штакельберга) либо зависеть также от управлений ведомых (обратные игры Штакельберга); при известных стратегиях ведущего ведомые одновременно и независимо выбирают свои стратегии, что

приводит к равновесию Нэша в игре ведомых. За конечное число итераций предложенный алгоритм имитационного моделирования позволяет построить приближённое решение модели или сделать вывод, что равновесия не существует. Достоверность и эффективность предложенного алгоритма следуют из свойств методов сценариев и прямого упорядоченного перебора с постоянным шагом. Получен ряд выводов относительно сравнительной эффективности методов иерархического управления инновациями.

Разработка и использование инноваций определяют магистральный путь устойчивого развития организаций любого типа. В РФ утвержден перечень крупных компаний с государственным участием, которые должны разработать и реализовать Программы инновационного развития (ПИР).

Важнейшим аспектом реализации ПИР служит создание системы управления инновационной деятельностью компании. Дело в том, что инновации – это внедрение новшеств, а внедрение – процесс насильственного проникновения инородных элементов в сопротивляющуюся среду. Теория и практика инноватики убедительно показывают, что попытки внедрения новшеств обязательно вызывают противодействие у значительной части сотрудников компании, заинтересованных прежде всего в достижении собственных целей и сохранении существующего порядка. Система управления инновационной деятельностью призвана реализовать механизмы согласования частных интересов сотрудников организации с общей целью разработки и использования инноваций (Новиков и Иващенко 2006). Без таких механизмов «мотивации инноваций» инновационная система окажется лишённой реального субъекта разработки и сопровождения и будет обречена на неудачу.

Предлагаемая динамическая СОЧИ-модель инновационного развития корпорации ориентирована в первую очередь на согласование интересов участников инновационного процесса, описывая в то же время его динамику.

Предполагается, что структура интересов каждого агента включает как частную (личные интересы), так и общественную (интересы компании в

целом, в первую очередь отражающие необходимость ее инновационного развития) составляющие. Агенты могут делить персональные ресурсы между этими направлениями. Отношения между агентами строятся на основе иерархии – имеется ведущий и несколько ведомых. Близкие модели изучаются в экономике общественных благ (Long 2010). При этом используется аппарат дифференциальных игр (Горелик и др. 1991; Зенкевич и др. 2009; Basar and Olsder 1999; Basar and Zhu 2011; Dockner et al. 2000; Jorgensen and Zaccour 2004; Long 2010). В иерархических дифференциальных играх существенно различие между прямыми играми Штакельберга, когда стратегия ведущего представляет собой постоянное управление, и обратными играми Штакельберга, где эта стратегия (механизм управления) зависит от управлений ведомых.

Концепция иерархического управления на базе математического моделирования описана в (Угольницкий 2016). Принуждение предполагает воздействие ведущего игрока (субъекта управления устойчивым развитием) на множество допустимых управлений ведомого (субъекта воздействия на управляемую динамическую систему), а побуждение – на его функцию выигрыша (Угольницкий и др. 2014, 2015).

Модели согласования интересов исследовались нами в (Ougolnitsky 2014). Здесь эти модели использованы в новой предметной области (инновационное развитие). Кроме того, в отличие от (Ougolnitsky 2014) теперь динамика системы описывается не дифференциальным, а разностным уравнением. При исследовании предложенной модели инновационного развития используется имитация и метод перебора областей допустимых управлений субъектов с некоторым шагом (Ougolnitsky and Usov 2018).

Итак, рассматривается двухуровневая иерархическая модель инновационного развития корпорации, включающая: субъекта управления верхнего уровня (ведущего);  $n$  субъектов управления нижнего уровня (ведомых, агентов инновации); управляемую динамическую систему.

Ведущий воздействует на каждого из ведомых. Его цель состоит в поддержании системы в некотором заданном состоянии, обеспечивающем её устойчивое функционирование. Добиться этого он может не единственным способом, поэтому, кроме того, он стремится к максимизации своего выигрыша. Взаимоотношения между субъектами управления данной системы носят следующий характер: Ведущий воздействует на Ведомых, Ведомые – на управляемую систему. Непосредственное воздействие Ведущего на последнюю отсутствует.

Целевой функционал Ведущего имеет вид

$$J_0(\cdot) = \sum_{t=1}^T e^{-\rho t} [x(t) - D(q(t))] \rightarrow \max . \quad (2.3.1)$$

Для воздействия на ведомых ведущий распоряжается функциями  $\{w_i(t)\}_{i=1}^n$ . При принуждении  $w_i(t) = q_i(t)$   $i = 1, 2, \dots, n$ , а при побуждении –  $w_i(t) = s_i(t)$ . Здесь  $t$  - время;  $T$  - конечный горизонт времени, до которого ведётся моделирование;  $s_i$  - доля дохода (вознаграждение)  $i$ -го агента от использования инноваций;  $q_i$  - возможные административные ограничения управлений ведомых (величин  $u_i$ ) снизу;  $u_i$  - управления ведомых (доля рабочего времени  $i$ -го агента, используемая на развитие и продвижение инноваций);  $x = x(t)$  - доход от реализации производимой продукции в момент времени  $t$ ;  $\rho$  - коэффициент дисконтирования;  $J_0$  - общий дисконтированный доход корпорации в целом на периоде  $[0, T]$  с учётом затрат  $D(q)$  на административный контроль (при принуждении);  $s = (s_1, s_2, \dots, s_n)$ ;  $q = (q_1, q_2, \dots, q_n)$ ;  $D$  - непрерывная выпуклая функция,  $D(0) = 0$ . Считаем, что контроль за ведомыми проводится только при принуждении, а при побуждении  $D(q) \equiv 0$ .

Ограничения на управления ведущего имеют вид

- при принуждении

$$0 \leq q_i(t) \leq r_i; \quad (2.3.2)$$

- при побуждении

$$s_i(t) \geq 0; \sum_{i=1}^n s_i(t) = 1; \quad (2.3.3)$$

где  $r_i$  - временные ресурсы  $i$ -го агента.

Целевые функционалы Ведомых берутся в виде

$$J_i(\cdot) = \sum_{t=1}^T e^{-\rho t} [p_i(r_i - u_i(t)) + s_i(t)x(t)] \rightarrow \max; \quad i = 1, 2, \dots, n. \quad (2.3.4)$$

Максимизация (2.3.4) проводится по функции  $u_i$ ,  $i \in N$ . Здесь  $N = \{1, 2, \dots, n\}$  - множество агентов инноваций. Это множество может рассматриваться на разных уровнях и включать сотрудников, подразделения, фирмы в составе корпорации;  $r_i - u_i$  - доля рабочего времени  $i$ -го агента, идущая на реализацию его частных интересов, не связанных с инновациями;  $p_i$  - функция дохода  $i$ -го агента от реализации его частных интересов;  $J_i$  - общий дисконтированный доход  $i$ -го агента на периоде  $[0, T]$ ;

Ограничения на управления ведомых имеют вид

$$q_i(t) \leq u_i(t) \leq r_i; \quad i = 1, 2, \dots, n. \quad (2.3.5)$$

Здесь при побуждении  $q_i(t) = 0$ ,  $i = 1, 2, \dots, n$ .

Динамика системы описывается разностным уравнением вида

$$x(t+1) = x(t) + h(x(t)) + \sum_{i=1}^n g_i(u_i(t)), \quad x(0) = x_0; \quad i = 1, \dots, n; \quad t = 0, 1, \dots, T-1. \quad (2.3.6)$$

Здесь  $g_i$  - функция вклада  $i$ -го агента в увеличение дохода корпорации за счет инноваций;  $h$  - функция динамики дохода в отсутствие инноваций;  $x_0$  - доход от реализации производимой продукции в начальный момент времени.

Итак, исследуется модель (2.3.1) – (2.3.6), которая представляет собой динамическую неантагонистическую игру  $n+1$  агента при наличии иерархии в отношениях между ними. Близкие к (2.3.1) - (2.3.6) постановки задач рассматриваются в динамических моделях экономики общественных благ. В

этих работах функции выигрыша игроков также включают две составляющие: личное потребление агента и его полезность от производства общественного блага. Так, в работе (Fershtman and Nitzan, 1991) показано, что программные стратегии по сравнению с позиционными приводят к более высокому стационарному запасу общественного блага. В статье (Dockner and Long, 1993) установлено, что кроме линейных позиционных стратегий существуют также пары нелинейных, которые образуют равновесие Нэша в позиционных стратегиях. Wirl (1996) показал, что использование нелинейных стратегий позволяет добиться лучшего исхода, чем в (Fershtman and Nitzan, 1991). В работах (Yanase, 2005, 2007) описывается долевое распределение общественного дохода посредством субсидий. Исследуется вопрос о том, какие постоянные субсидии обеспечивают долгосрочное социально оптимальное значение общественного блага. В статье (Benchekroun and Long, 2008) вводится уровень кооперации как переменная состояния, косвенно влияющая на целевые функционалы эгоистичных агентов и динамику. Чем больше этот уровень, тем выше доверие индивидов друг к другу и соответствующая готовность к инвестициям в общественное благо, преодолевающая проблему безбилетника без использования угроз. В работе (Kemp and Long, 2009) дополнительно учитывается удовлетворенность донора от личного вклада в общественное благо.

В целом, авторы указанных работ уделяют основное внимание получению аналитических результатов при существенных упрощениях исходных моделей, численные методы используются в основном для иллюстрации. В настоящей статье вопросы существования, единственности и устойчивости равновесий не изучаются. Функционал выигрыша ведущего игрока не представляет собой утилитаристскую функцию общественного благосостояния, а включает фазовую переменную и затраты на административное управление. Акцент сделан на сравнительном анализе методов иерархического управления в достаточно общей модели



посредством численных расчетов, в связи с чем получены содержательные выводы.

С математической точки зрения модель (2.3.1) – (2.3.6) представляет собой динамическую неантагонистическую игру, в которой динамика описывается разностным уравнением. Ее аналитическое решение может быть получено, как правило, лишь в некоторых частных случаях. В общем случае для получения решения применяются численные методы и имитационное моделирование. Примем следующие предположения относительно информационного регламента игры:

- все игроки используют программные стратегии;
- ведущий выбирает и сообщает ведомым экономические управления (2.3.3) либо административные управления (2.3.2), которые могут быть только функциями времени (прямые игры Штакельберга) либо зависеть также от управлений ведомых (обратные игры Штакельберга);
- при известных стратегиях ведущего ведомые одновременно и независимо выбирают свои стратегии, что приводит к равновесию Нэша в игре ведомых. Будем считать, что множество равновесий Нэша не пусто.

Приведем алгоритмы построения равновесий для модели (2.3.1) – (2.3.6) при разных информационных регламентах. Для моделей общего вида эти алгоритмы были сформулированы в (Угольницкий и Усов 2013, 2014).

Алгоритм построения равновесия в прямой игре Штакельберга при побуждении (принуждении) состоит в следующем.

1. Ведущий игрок выбирает программную стратегию вида  $s(t) = \{s_i(t)\}_{i=1}^N \in S = \{s_i(t) \geq 0; i = 1, \dots, N; s_1(t) + \dots + s_N(t) = 1, t \geq 0\}$  (при побуждении) или  $q(t) = \{q_i(t)\}_{i=1}^N \in Q = \{0 \leq q_i(t) \leq r_i; i = 1, \dots, N; t \geq 0\}$  (при принуждении) и сообщает ее остальным игрокам.
2. Зная выбранную Ведущим игроком стратегию, остальные игроки разыгрывают между собой динамическую игру (2.3.4) – (2.3.6), решением

которой служит равновесие Нэша  $NE(w(\cdot))$ , где при побуждении  $w(\cdot) = s(\cdot)$ , а при принуждении  $w(\cdot) = q(\cdot)$ .

3. Ведущий игрок выбирает свою программную стратегию таким образом, чтобы максимизировать выигрыш (2.3.1) на множестве равновесий Нэша  $NE(w(\cdot))$ .

Алгоритм построения равновесия в обратной игре Штакельберга при побуждении (принуждении) состоит в следующем.

1) Ведущий игрок определяет стратегии наказания Вемых, если они отказываются с ним сотрудничать. Для этого в динамической игре (2.3.4) - (2.3.6) находятся равновесия Нэша в зависимости от управлений Ведущего -  $NE(w(\cdot))$ , где при побуждении  $w(\cdot) = s(\cdot)$ , а при принуждении  $w(\cdot) = q(\cdot)$ .

Допустим, что число равновесий Нэша при фиксированном управлении Ведущего  $w(t)$  равно  $L_w$ . Обозначим их через

$$(u_i^{NE}(t))_k = (u_i^{NE}(w(t), t))_k; i = 1, \dots, n; k = 1, 2, \dots, L_w \quad \text{и}$$

$$u_k^{NE}(t) = \left\{ (u_i^{NE}(t))_k \right\}_{i=1}^n.$$

Вводится стратегия наказания Вемого Ведущим

$$w^P(t) = \left\{ w_i^P(t) \right\}_{i=1}^n; \quad w_i^P(t) = \arg \min_{w(t) \in W} \max_{1 \leq k \leq L_w} J_i((u_i^{NE})_k(t), x(t)); i = 1, \dots, n.$$

(2.3.7)

Находятся величины максимальных выигрышей Вемых в случае, когда они отказываются сотрудничать с Ведущим и он применяет стратегию наказания:

$$(L_F)_i = \min_{w(t) \in W} \max_{1 \leq k \leq L_w} J_i((u_i^{NE})_k(t), x(t)); i = 1, \dots, n. \quad (2.3.8)$$

2) Решается задача оптимального управления (2.3.1) - (2.3.3), (2.3.5), (2.3.6) с дополнительными условиями  $(L_F)_i < J_i(u_i(t), x(t)); i = 1, \dots, n$ .

Максимум (2.3.1) ищется по  $2n$  функциям  $w_i(t), u_i(t); i = 1, \dots, n$ . Решение указанной задачи оптимального управления обозначим  $\{w_i^R(t), u_i^R(t)\}_{i=1}^n$ , где  $w_i^R(t)$  - стратегия поощрения  $i$ -го Вемого Ведущим.

3) Ведущий предъявляет каждому Ведомому стратегию с обратной

$$\text{связью } w_i(t) = \begin{cases} w_i^R(t), & \text{если } u_i(t) = u_i^R(t) \text{ для } \forall t \in [0, \infty), \\ w_i^P(t), & \text{иначе.} \end{cases}$$

4) При экономически разумных Ведомых равновесие имеет вид

$$\{w_i^R(t), u_i^R(t)\}_{i=1}^n. \quad (2.3.9)$$

В случае входных данных общего вида решение иерархических игр для указанных информационных регламентов ищется с помощью имитационного моделирования по методу сценариев.

Этот алгоритм для задачи (2.3.1) - (2.3.6) в случае прямой игры Штакельберга состоит в следующем.

1. Область допустимых управлений Ведущего ((2.3.3) при побуждении и (2.3.2) при принуждении) разбивается при каждом  $t=1, \dots, T$  для каждого ведомого ( $i=1, \dots, N$ ) на  $G$  частей точками  $q_{it}^{(z_{it})} = z_{it}r_{it}/G$  при принуждении и  $s_{it}^{(z_{it})} = z_{it}/G$  при побуждении ( $z_{it}=0, 1, \dots, G$ ). В случае побуждения при

выборе значений  $z_{it}$  учитывается условие  $\sum_{i=1}^N s_{it}^{(z_{it})} = 1$ , которое должно

выполняться при  $t=1, \dots, T$ . В результате у ведущего при  $t=1, \dots, T$  для каждого ведомого имеется не более, чем  $(G+1)$  сценарий. Задается текущая стратегия ведущего из множества его допустимых управлений  $\left\{s_{it}^{(z_{it})}\right\}_{i,t=1}^{N(T)}$ .

2. Области допустимых управлений Ведомых (2.3.5) разбиваются при каждом  $t=1, \dots, T$  на  $K$  частей точками  $u_{it}^{(k_{ij})} = q_{it}^{(z_{it})} + k_{it}(r_{it} - q_{it}^{(z_{it})})/K$ ;

$k_{it}=0, 1, \dots, K$  ( $i=1, \dots, N$ ). В случае побуждения -  $q_{it}^{(z_{it})} = 0$ . Путём прямого упорядоченного перебора управлений Ведомых ( $k_{it}=0, 1, \dots, K$ ) при каждом  $t=1, 2, \dots, T$  строится равновесие Нэша, то есть находятся сеточные функции

$$\left\{u_{it}^{NE}(w_{it}^{(z_{it})})\right\}_{t=1}^T; \quad i=1, 2, \dots, N, \text{ при которых (2.3.4) с учетом (2.3.5), (2.3.6)}$$

принимают наибольшие значения среди перебираемых.

4. Найденные на предыдущем шаге стратегии ведомых  $\left\{ \mu_{it}^{NE}(w_{it}^{(z_{it})}) \right\}_{t=1}^T; (i=1,2,\dots,N)$  подставляются в (2.3.1), (2.3.6). Ведущий игрок выбирает свою программную стратегию таким образом, чтобы максимизировать выигрыш (2.3.1) на множестве равновесий Нэша  $\left\{ \mu_{it}^{NE}(w_{it}^{(z_{it})}) \right\}_{t=1}^T; (i=1,2,\dots,N)$  с учетом (2.3.2) или (2.3.3) и (2.3.6). Если число перебираемых управлений (сценариев) Ведущего не исчерпано, то проводится выбор его новой стратегии. После этого необходимо вернуться на пункт 3 алгоритма.

5. Таким образом, определяется приближение к равновесию Нэша, т.е. набор

$$\text{величин } \left\{ \left\{ w_{it}^* \right\}_{t=1}^T, \left\{ \mu_{it}^{NE}(w_{it}^*) \right\}_{t=1}^T \right\}_{i=1}^N. \quad (2.3.10)$$

При выборе новой текущей стратегии одного из Ведомых используется метод прямого упорядоченного перебора с постоянным шагом (Когай и Фадеев 2001). Интервал неопределенности при нахождении, например, стратегии  $i$ -го Ведомого при побуждении есть  $\varepsilon = \frac{r_i}{K+1}$ .

За конечное число итераций предложенный алгоритм имитационного моделирования позволяет построить приближенное решение модели (2.3.1) — (2.3.6) или сделать вывод, что равновесия не существует. Достоверность и эффективность предложенного алгоритма следуют из свойств методов сценариев и прямого упорядоченного перебора с постоянным шагом.

Алгоритм имитационного моделирования на основе метода сценариев для задачи (2.3.1) - (2.3.6) в случае обратной игры Штакельберга состоит в следующем.

1. Область допустимых управлений Ведущего ((2.3.3) при побуждении и (2.3.2) при принуждении) разбивается при каждом  $t=1,\dots,T$  для каждого ведомого ( $i=1,\dots,N$ ) на  $G$  частей точками  $q_{it}^{(z_{it})} = z_{it}r_{it} / G$  при принуждении

и  $s_{it}^{(z_{it})} = z_{it} / G$  при побуждении ( $z_{it} = 0, 1, \dots, G$ ). В случае побуждения при выборе значений  $z_{it}$  учитывается условие  $\sum_{i=1}^N s_{it}^{(z_{it})} = 1$ , которое должно

выполняться при  $t = 1, \dots, T$ . В результате у ведущего в каждый момент времени для каждого ведомого имеется не более, чем  $(G+1)$  сценарий.

2. Области допустимых управлений Ведомых (2.3.5) разбиваются при каждом  $t = 1, \dots, T$  на  $K$  частей точками  $u_{it}^{(k_{it})} = q_{it}^{(z_{it})} + k_{it}(r_{it} - q_{it}^{(z_{it})}) / K$ ;

$k_{ij} = 0, 1, \dots, K$  ( $i = 1, \dots, N$ ). При побуждении -  $q_{it}^{(z_{it})} = 0$ . Путём перебора

стратегий Ведущего, т.е. сеточных функций  $(w_{it}^{(z_{it})})_{i,t=1}^{N(T)}$ , ( $z_{it} = 0, 1, \dots, G$ ), и

реакций ( $k_{it} = 0, 1, \dots, K$ ) каждого Ведомого ( $i = 1, 2, \dots, N$ ) при каждом

$t = 1, 2, \dots, T$  определяются стратегии наказания Ведомых Ведущим

$w^P = \{w_{it}^P\}_{i,t=1}^{N(T)}$  и гарантированный выигрыш Ведомых, если они

отказываются сотрудничать с Ведущим  $(L_F)_i; i = 1, \dots, N$ .

3. Для каждого управления (сценария) Ведущего  $(w_{it}^{(z_{it})})_{i,t=1}^{N(T)} \in W$ ,  $z_{it} = 0, 1, \dots, G$  путём прямого упорядоченного перебора возможных стратегий ( $k_{it} = 1, 2, \dots, K$ ) каждого Ведомого ( $i = 1, 2, \dots, N$ ) при каждом  $t = 1, 2, \dots, T$  находится максимум (2.3.1) при условиях (2.3.5), (2.3.2)

или (2.3.3) и дополнительных условиях  $(L_F)_i < J_i(\{u_{it}\}_{t=1}^T, x(t))$ . Решения

указанных задач оптимального управления обозначим  $\left\{u_i^R(\{w_{it}^{(z_{it})}\}_{t=1}^T)\right\}_{i=1}^N$ . Из

них выбирается пара  $(\{w_{it}^{(*)}\}_{i,t=1}^{N(T)}, \left\{u_i^R(\{w_{it}^{(*)}\}_{t=1}^T)\right\}_{i=1}^N)$ , доставляющая максимум

(2.3.1).

4. При экономически разумных Ведомых равновесие имеет вид  $(\{w_{it}^{(*)}\}_{i,t=1}^{N(T)}, \{u_i^R(\{w_{it}^{(*)}\}_{t=1}^T)\}_{i=1}^N)$ .

В сделанных предположениях возможны четыре информационных регламента для модели (2.3.1) - (2.3.6) – прямые и обратные игры Штакельберга, каждая в случае побуждения и принуждения. Численное исследование на тестовых примерах проведено в каждом случае.

Примем следующие обозначения для выигрышей ведущего ( $J_{NE}^{ind}$ ) и индексов системной согласованности ( $K_{NE}^{ind} = J_{NE}^{ind} / J_{\max}^{ind}$ ) при побуждении или принуждении в прямой игре Штакельберга ( $ind = comp-st$  или  $ind=imp-st$ ) или обратной игре Штакельберга ( $ind = comp-ger$  или  $ind=imp-ger$ ) соответственно;  $J_{\max}^{ind}$  - выигрыш ведущего в соответствующей игре в случае, когда все ведомые максимизируют (2.3.1).

Численное исследование проводилось согласно алгоритмам, сформулированным выше, для следующего набора входных функций:

$$D(q(t)) = \sum_{i=1}^n \bar{C}_i(t) \frac{q_i(t)}{r_i - q_i(t) + \varepsilon}; \bar{C}_i(t), \bar{\bar{C}}_i(t) - \text{заданные функции; } \varepsilon \ll 1;$$

$$g_i(u) = A_i u^2; h(x) = B \sqrt{x} \quad \text{и} \quad \text{следующих} \quad \text{входных} \quad \text{данных:}$$

$$\bar{C}_i(t) = 40; \bar{\bar{C}}_i(t) = 40; (i = 1, 2); p_1 = 10; p_2 = 15; r_1 = 10; r_2 = 15; x_0 = 100; A_1 = 0,7; A_2 = 0,5; T = 3; \rho = 0.1; B = 0.3.$$

В таблицах 2.3.1 - 2.3.4 расчёты 1 приведены для указанных выше входных данных; расчёты 2 - для указанных входных данных и  $\bar{C}_1(t) = \bar{C}_2(t) = 10$ ; расчёты 3 и 4 – для  $p_1 = 50; p_2 = 75$ . и  $r_1 = 50; r_2 = 65$  соответственно; расчёты 5 и 6 – для  $r_1 = 1; r_2 = 1,5$  и  $x_0 = 10$ .

В таблицах 2.3.1, 2.3.2 приведены результаты исследования прямых игр Штакельберга в случае побуждения ( $q_i = 0; i = 1, 2$ ) и принуждения

( $s_i = 1/2; i = 1, 2$ ) соответственно, а в таблицах 2.3.3, 2.3.4 - обратных игр Штакельберга. Во всех примерах оптимальные стратегии агентов зависят от выбранного метода управления и реализуемого информационного регламента.

Таблица 2.3.1

Результаты счёта в случае побуждения для прямой игры Штакельберга

Расчёты	1,2	3	4	5	6
$u_1^*(t_k)(k = 1, 2, 3)$	10; 10; 0	0; 0; 0	50; 50; 50	0; 0; 0	10; 10; 0
$u_2^*(t_k)(k = 1, 2, 3)$	0; 0; 0	0; 0; 0	65; 65; 65	0; 0; 0	0; 0; 0
$s_1^*(t_k)(k = 1, 2, 3)$	0; 2/3; 1	0; 0; 0	0; 1/3; 1/3	0; 0; 0	0; 2/3; 1
$s_2^*(t_k)(k = 1, 2, 3)$	1; 1/3; 0	1; 1; 1	1; 2/3; 2/3	1; 1; 1	1; 1/3; 0
$x(T)$	251	109	11735	109	157
$J_1^{IMP-ST}$ (т.р.)	395	1232	5039	24	2569
$J_2^{IMP-ST}$ (т.р.)	778	3033	13667	316	669
$J_{NE}^{IMP-ST}$ (т.р.)	545	260	18706	260	315
$J_{max}^{IMP-ST}$ (т.р.)	1136	1136	18706	269	907
$K_{NE}^{IMP-ST}$	0.48	0.23	1	0.97	0.35

Таблица 2.3.2

Результаты счёта в случае принуждения для прямой игры Штакельберга

Расчёты	1	2	3	4	5	6
$u_1^*(t_k)(k = 1, 2, 3)$	10; 10; 0	10; 10; 6,6	6,6; 0; 0	50; 50; 50	0; 0; 0	10; 10; 0
$u_2^*(t_k)(k = 1, 2, 3)$	15; 10; 0	15; 10; 10	10; 10; 0	65; 65; 65	0; 0; 0	15; 10; 0

$q_1^*(t_k)(k=1,2,3)$	3,3; 6,6; 0	3,3; 6,6; 6,6	6,6; 0; 0	0; 0; 0	0; 0; 0	3,3; 6,6; 0
$q_2^*(t_k)(k=1,2,3)$	10; 10; 0	10; 10; 10	10; 10; 0	0; 0; 0	0; 0; 0	10; 10; 0
$x(T)$	416	497	242	11735	109	323
$J_1^{IMP-ST}$ (т.п.)	525	506	1201	9353	155	411
$J_2^{IMP-ST}$ (т.п.)	679	598	1750	9353	185	565
$J_{NE}^{IMP-ST}$ (т.п.)	681	878	331	18706	260	452
$J_{max}^{IMP-ST}$ (т.п.)	1136	1136	1136	18706	269	907
$K_{NE}^{IMP-ST}$	0,6	0,77	0,29	1	0,97	0,5

Таблица 2.3.3

Результаты счёта в случае побуждения для обратной игры Штакельберга

Расчёты	1,2	3	4	5	6
$u_1^*(t_k)(k=1,2,3)$	10;10;10	0;0;0	50; 50; 50	0;0;0	10; 10; 10
$u_2^*(t_k)(k=1,2,3)$	15;15;5	0;0;0	65; 65; 65	0;0;0	15; 15; 10
$s_1^*(t_k)(k=1,2,3)$	0;0;2/3	1/3;1/3;1/3	0;0;1/3	1/3;1/3;1/3	0;0;2/3
$s_2^*(t_k)(k=1,2,3)$	1;1;1/3	1/3;1/3;1/3	1;1;2/3	1/3;1/3;1/3	1;1;1/3
$x(T)$	562	109	11735	109	505
$J_1^{IMP-ST}$ (т.п.)	277	1232	2898	24	249
$J_2^{IMP-ST}$ (т.п.)	895	2772	15808	55	666
$J_{NE}^{IMP-ST}$ (т.п.)	1062	260	18706	260	861
$J_{max}^{IMP-ST}$ (т.п.)	1136	1136	18706	269	907
$K_{NE}^{IMP-ST}$	0,93	0.23	1	0.97	0.95



Таблица 2.3.4

Результаты счёта в случае принуждения для обратной игры Штакельберга

Расчёты	1,2,3	4	5	6
$u_1^*(t_k)(k=1,2,3)$	10;10;10	50; 50; 50	1;1;1	10;10;10
$u_2^*(t_k)(k=1,2,3)$	15;15;15	65; 65; 65	1.5;1.5;1.5	15;15;15
$q_1^*(t_k)(k=1,2,3)$	0;0;0	0; 0; 0	0; 0; 0	0;0;0
$q_2^*(t_k)(k=1,2,3)$	0;0;0	0; 0; 0	0; 0; 0	0;0;0
$x(T)$	662	11735	114	568
$J_1^{IMP-ST}$ (т.р.)	568	9353	132	453
$J_2^{IMP-ST}$ (т.р.)	568	9353	133	453
$J_{NE}^{IMP-ST}$ (т.р.)	1136	18706	269	907
$J_{max}^{IMP-ST}$ (т.р.)	1136	18706	269	907
$K_{NE}^{IMP-ST}$	1	1	1	1

На основе проведенных расчётов можно сделать следующие выводы.

1. В данной постановке для достаточно широкого класса входных данных интересы агентов нижнего и верхнего уровней в прямой игре Штакельберга при побуждении плохо согласованы, индекс системной согласованности невелик. Например, при малом доходе от производимой продукции в начальный момент времени или малом доходе агента от реализации своих частных интересов он близок к 0.2 – 0.3. В случае принуждения в прямой игре Штакельберга ситуация для большинства входных данных меняется, индекс системной согласованности возрастает. В обратной игре Штакельберга для большинства входных данных индекс системной

согласованности близок к единице при побуждении, а при принуждении просто равен единице. Именно в этом состоит преимущество информационного регламента обратной игры Штакельберга по сравнению с прямой. Но для реализации регламента обратных игр Штакельберга ведущий должен обладать значительными возможностями воздействия на ведомых.

2. С ростом затрат ведущего на административный контроль при принуждении индекс системной согласованности, как и доход ведущего, растут, доходы ведомых – падают. С ростом доходов от реализации частных интересов индекс системной согласованности и доход ведущего падают, доходы ведомых растут. С ростом временных ресурсов ведомых индекс системной согласованности, как и доходы всех субъектов, растут. С ухудшением начального состояния системы доходы субъектов падают, а индекс системной согласованности растет.

3. При реализации информационного регламента обратной игры Штакельберга модель полностью согласована при принуждении для широкого класса входных функций (индекс системной согласованности равен единице).

4. Выбор более эффективного метода иерархического управления (принуждение или побуждение) зависит от входных параметров модели. В большинстве случаев больший доход ведущему и лучшую системную согласованность обеспечивает принуждение, и только в отдельных случаях – побуждение.

5. В прямой игре Штакельберга, в отличие от обратной, с ростом дохода от производимой продукции в начальный момент, временных ресурсов агентов или доходов агентов от реализации своих частных интересов, индекс системной согласованности растет, а при большом (больше 1000) доходе от производимой продукции в начальный момент времени он близок к единице.

6. С ужесточением административного контроля за агентами при принуждении в прямой игре Штакельберга, в отличие от обратной, индекс системной согласованности падает.

7. Увеличение доходов ведомых от реализации своих частных интересов приводит к резкому росту их доходов в прямой игре Штакельберга и практически не меняет их доходы для достаточно широкого класса входных данных в обратной игре. Доход ведущего при этом падает. Ведущему невыгодно увеличение доходов агентов от частной деятельности.

8. Увеличение временных ресурсов ведомых выгодно и им, и ведущему в играх обоих типов. Ужесточение административного контроля за агентами при принуждении в играх обоих типов увеличивает доход ведущего, доходы ведомых не возрастают.

Таким образом, математическое моделирование инновационного развития корпораций помогает выбрать информационный регламент и метод иерархического управления, которые обеспечивают максимальный экономический эффект от внедрения инноваций для всех субъектов и лучшую системную согласованность (Угольницкий и Усов 2016). Рассмотрение данной тематики продолжено в статье (Угольницкий и Усов 2019).

В статье (Sukhinov et al. 2017) исследуются динамические СОЧИ-модели двухуровневых систем управления с учётом условий устойчивого развития. В качестве механизмов иерархического управления рассматриваются методы принуждения и побуждения. Даны определения равновесий и приведены алгоритмы их построения на основе имитационного моделирования для различных информационных регламентов. Проведен сравнительный анализ эффективности указанных механизмов управления для модели экосистемы мелководного водоема (на примере Азовского моря).

При анализе динамических иерархических игр в качестве основы используется подход А.Ф. Кононенко и его соавторов (Кононенко 1977; Горелик и др. 1991). В (Кононенко 1977; Горелик и др. 1991) описана структура оптимальных управлений и предложен метод решения динамических игр, основанный на разработке взаимовыгодной программы и применении санкций при отклонениях от нее. Важное значение здесь имеет

анализ информационных регламентов динамических игр, т.е. предположений о порядке ходов игроков и их информированности о действиях других игроков и состоянии управляемой системы. Развитие указанного подхода осуществлено авторами (Угольницкий и Усов 2013, 2014). В настоящей статье предложена реализация алгоритмов поиска решений иерархических дифференциальных игр в программных стратегиях с помощью имитационного моделирования.

Для содержательной иллюстрации и численной реализации предлагаемого подхода используется модель мелководного водоема (на примере Азовского моря), включающая гидродинамическую и гидробиологическую подсистемы. Описание исходной модели и ее идентификации приведено в (Никитина и Семёнов 2014; Сухинов и Чистяков 2012); Сухинов и др. 2011, 2012).

Для классификации информационных регламентов иерархических дифференциальных игр двух лиц используются три признака (основания). Во-первых, это отсутствие или наличие обратной связи по состоянию управляемой системы, которое приводит к программным или позиционным стратегиям соответственно. Во-вторых, это отсутствие или наличие обратной связи по управлению, т.е. информированности ведущего игрока о действиях ведомого в момент принятия решения. В совокупности получаем матрицу информационных регламентов иерархических дифференциальных игр двух лиц, подробно исследованных А.Ф. Кононенко с соавторами в (Кононенко 1977; Горелик и др. 1991) (табл. 2.3.5).

В (Угольницкий и Усов 2014) предложено учитывать третий признак классификации: воздействие ведущего игрока на множество допустимых управлений ведомого игрока (принуждение) или на функцию выигрыша ведомого (побуждение), что содержательно с некоторыми оговорками соответствует административным либо экономическим методам управления. Тем самым, получаем восемь классов моделей иерархического управления. Применительно к задаче управления устойчивым развитием (УР)

дополнительно рассматриваются некоторые требования к состоянию управляемой динамической системы (УДС), интерпретируемые как условия ее гомеостаза. Математически эти условия представляют собой фазовые ограничения дифференциально-игровой модели.

Таблица 2.3.5

Базовые информационные регламенты иерархических дифференциальных игр двух лиц

	Без обратной связи по управлению (прямые игры Штакельберга)	С обратной связью по управлению (обратные игры Штакельберга)
Без обратной связи по состоянию (программные стратегии)	Прямые игры Штакельберга в программных стратегиях ( $\Gamma_{1t}$ )	Обратные игры Штакельберга в программных стратегиях ( $\Gamma_{2t}$ )
С обратной связью по состоянию (позиционные стратегии)	Прямые игры Штакельберга в позиционных стратегиях ( $\Gamma_{1x}$ )	Обратные игры Штакельберга в позиционных стратегиях ( $\Gamma_{2x}$ )

Рассмотрим базовую модель иерархической системы управления, включающую ведущего и ведомого игроков и УДС. Ведомый воздействует на УДС, исходя исключительно из своих частных интересов, что может привести к нарушению требований УР. Выполнение этих требований является главной целью ведущего, который, кроме того, также может иметь частные интересы. Требования УР формализуются как принадлежность траектории УДС определенной области пространства состояний. Для обеспечения условий УР ведущий может использовать механизмы принуждения и побуждения. Исходная теоретико-игровая модель управления УР имеет вид

$$J_L(y(t)) = \int_0^T e^{-\alpha t} g_L(y(t)) dt \rightarrow \max, \quad (2.3.11)$$

$$q(t) \in Q; p(t) \in P, \quad (2.3.12)$$

$$J_F(y(t)) = \int_0^T e^{-\alpha t} g_F(y(t)) dt \rightarrow \max, \quad (2.3.13)$$

$$u(t) \in U, \quad (2.3.14)$$

$$\dot{x} = f(u(t), x(t)), x(0) = x_0, \quad (2.3.15)$$

$$x(t) \in X^*, t \geq 0. \quad (2.3.16)$$

Здесь индексами  $L$  и  $F$  обозначены ведущий (Leader) и ведомый (Follower) игроки;  $J_L$  и  $J_F$ ,  $g_L$  и  $g_F$  – соответственно их интегральные и мгновенные функции выигрыша;  $T$  – период рассмотрения, который может быть как конечным, так и бесконечным (второй вариант более адекватен задаче управления УР);  $\alpha$  – коэффициент дисконтирования;  $y(t) = (q(t), p(t), u(t), x(t))$ ;  $q(t)$  – вектор управлений принуждения, применяемых ведущим;  $p(t)$  – вектор управлений побуждения ведущего;  $u(t)$  – вектор управлений ведомого (той же размерности, что и  $q(t)$ );  $x(t)$  – вектор состояния УДС;  $X^*$  – множество состояний УР системы; соотношения (2.3.15) описывают динамику последней при заданных начальных условиях и управлениях ведомого; условия (2.3.16) выражают требования УР применительно к УДС. Предполагается, что  $Q, P, U$  – компактные подмножества соответствующих метрических пространств;  $J_L$  и  $J_F$  непрерывны по своим аргументам; для вектор-функции  $f$  из (2.3.15) выполнены условия Каратеодори.

Математически задача управления УР представляет собой дифференциальную игру (2.3.11)–(2.3.15) с фазовыми ограничениями (2.3.16). Решение игры (2.3.11)–(2.3.15) выражает приемлемый компромисс между субъектами управления, удовлетворяющий требованиям (2.3.16) к состоянию УДС, и тем самым означает решение задачи управления УР. Вид стратегий ведущего в зависимости от информационного регламента игры показан в табл. 2.3.6.

## Стратегии ведущего игрока

	$\Gamma_{1t}$	$\Gamma_{2t}$	$\Gamma_{1x}$	$\Gamma_{2x}$
Принуждение	$q(t)$	$q(t, u(t))$	$q(t, x(t))$	$q(t, x(t), u(t, x(t)))$
Побуждение	$p(t)$	$p(t, u(t))$	$p(t, x(t))$	$p(t, x(t), u(t, x(t)))$

Рассмотрим информационное расширение игры (2.3.11)–(2.3.16), соответствующее методу принуждения для игры вида  $\Gamma_{1t}$ . В этом случае управления побуждения не используются. Ведущий воздействует на множество допустимых стратегий ведомого (но не на его функцию выигрыша), выбирая и сообщая ему программную стратегию  $q(t)$ . Одним из естественных примеров такого воздействия служат неравенства типа  $u(t) \leq q(t)$  или  $q(t) \leq u(t)$ , ограничивающие допустимые управления ведомого сверху или снизу в зависимости от специфики задачи. Эти ограничения “физически” не могут быть нарушены ведомым. Кроме того, ведущий не получает информации о действиях ведомого.

Зная  $q(t)$ , ведомый выбирает свой оптимальный ответ из множества

$$R_F^{comp}(q(t)) = \left\{ u(t) \in U : J_F^{comp}(q(t), u(t), x(t)) = \max_{v(t) \in U} J_F^{comp}(q(t), v(t), x(t)) \right\}.$$

Гарантированный результат ведущего есть

$$\gamma_L^{comp} = \sup_{q(t) \in Q} \inf_{u(t) \in R_F^{comp}(q(t))} J_L^{comp}(q(t), u(t), x(t)),$$

где траектория  $x(t)$  порождается соответствующим управлением  $u(t)$ .

Равновесием принуждения (compulsion) с учетом требований УР для игры вида  $\Gamma_{1t}$  назовем ситуацию

$$(q^*(t), \hat{u}^*(t)) \in \hat{B}R_F : J_L^{comp}(\pi(q^*(t), \hat{u}^*(t), x^*(t))) = \gamma_L^{comp},$$

$$\hat{B}R_F = \left\{ (q(t), \hat{u}(t)) \in Q \times \hat{U} : \hat{u}(t) = u(q(t)) \in R_F^{comp}(q(t)) \right\}$$

при условии (2.3.16), где траектория  $x^*(t)$  порождается стратегией  $\tilde{u}^*(t) = u(q^*(t))$ ,  $\pi$  – функция проекции на исходную область определения (Угольницкий и Усов 2014).

Построенное информационное расширение аналогично игре  $\Gamma_{1t}$  (Кононенко 1977; Горелик и др. 1991) с двумя особенностями: ведущий воздействует на множество допустимых стратегий ведомого и требуется выполнение условия (2.3.16). Отметим, что аналогично статической игре  $\Gamma_1$  здесь оптимальная реакция ведомого на выбор ведущего представляет собой оператор  $\tilde{u}(t) = u(q(t))$ , множеством значений которого является  $R_F^{comp}(q(t))$ .

Теперь рассмотрим информационное расширение игры (2.3.11)–(2.3.16), соответствующее методу побуждения для игры  $\Gamma_{2t}$ . Управления принуждения не используются. Ведущий воздействует на функцию выигрыша ведомого, выбирая и сообщая ему механизм управления в стратегиях с обратной связью по управлению, но без обратной связи по состоянию  $\tilde{p}(t) = p(u(t))$ . Наиболее естественным примером такого оператора служит механизм поощрения – наказания

$$\tilde{p}(t) = \begin{cases} \text{поощрение ведомого, если } u(t) \text{ устраивает ведущего,} \\ \text{наказание ведомого, если иначе.} \end{cases}$$

Зная стратегию  $\tilde{p}(t)$ , ведомый выбирает оптимальный ответ из множества

$$R_F^{imp}(\tilde{p}(t)) = \{u(t) \in U : J_F^{imp}(p(u(t)), u(t), x(t)) = \max_{v(t) \in U} J_F^{imp}(p(v(t)), v(t), x(t))\}.$$

Здесь вновь считаем, что  $u(t) = \tilde{u}(\tilde{p}(t))$ . Гарантированный выигрыш ведущего равен

$$\gamma_L^{imp} = \sup_{\tilde{p}(t) \in \tilde{P}} \inf_{u \in R_F^{imp}(\tilde{p})} J_L^{imp}(\pi(\tilde{p}(t), u(t), x(t))), \tilde{P} = \{\tilde{p} : U \rightarrow P\}, \tilde{U} = \{\tilde{u} : \tilde{P} \rightarrow U\},$$

$\pi : \tilde{P} \times \tilde{U} \times X \rightarrow P \times U \times X$  – проекция, траектория  $x(t)$  порождается стратегией  $\tilde{u}(t) = u(\tilde{p}(t))$  (Угольницкий и Усов 2014). Равновесие побуждения (impulsion) с учетом требований УР для игры  $\Gamma_{2t}$  есть ситуация



$$(\tilde{p}^*(t), \tilde{u}^*(t)) \in B\tilde{R}_F : J_L^{imp}(\pi(\tilde{p}^*(t), \tilde{u}^*(t), x^*(t))) = \gamma_L^{imp},$$

$$B\tilde{R}_F = \{(\tilde{p}(t), \tilde{u}(t)) \in \tilde{P} \times \tilde{U} : u(p(u(t))) \in R_F^{imp}(\tilde{p}(t))\}$$

при условии (2.3.16), где траектория  $x^*(t)$  порождается стратегией  $\tilde{u}^*(t)$ . Построенное информационное расширение аналогично игре  $\Gamma_{2t}$  (Кононенко 1977; Горелик и др. 1991 с дополнительным требованием (2.3.16). Равновесия в остальных шести моделях определяются аналогично, алгоритмы их нахождения приведены в (Угольницкий и Усов 2014).

Рассмотрим нелинейную пространственно-неоднородную трехмерную модель экосистемы мелководного водоёма: “рыба – фитопланктон – зоопланктон – питательные вещества – детрит”, которая описывается системой дифференциальных уравнений в частных производных в области  $G$ , представляющей собой замкнутый бассейн, ограниченный невозмущённой поверхностью водоёма  $\Sigma_0$ , дном  $\Sigma_H = \Sigma_H(x, y)$  и цилиндрической поверхностью  $\sigma$ , для интервала  $0 \leq t \leq T$ ;  $\Sigma = \Sigma_0 \cup \Sigma_H \cup \sigma$  – кусочно-гладкая граница области  $G$  (Никитина и Семёнов 2014).

Уравнения биологической кинетики имеют вид:

$$\begin{aligned} \frac{\partial X}{\partial t} + \text{div}(UX) &= \mu_X \Delta X + \frac{\partial}{\partial z} (v_X \frac{\partial X}{\partial z}) + g_X(X, S) - (\delta_X Z + \varepsilon_X + \sigma_X P)X, \\ \frac{\partial Z}{\partial t} + \text{div}(UZ) &= \mu_Z \Delta Z + \frac{\partial}{\partial z} (v_Z \frac{\partial Z}{\partial z}) + g_Z(X, Z) - (\varepsilon_Z + \delta_Z)Z, \\ \frac{\partial S}{\partial t} + \text{div}(US) &= \mu_S \Delta S + \frac{\partial}{\partial z} (v_S \frac{\partial S}{\partial z}) + \gamma_S \varepsilon_D D - g_S(X, S) + B(S_p - S) + f, \\ \frac{\partial D}{\partial t} + \text{div}(UD) &= \mu_D \Delta D + \frac{\partial}{\partial z} (v_D \frac{\partial D}{\partial z}) - g_D(D, P) + \varepsilon_X X + \varepsilon_Z Z - \varepsilon_D D, \\ \frac{\partial P}{\partial t} + \text{div}(U_P P) &= \mu_P \Delta P + \frac{\partial}{\partial z} (v_P \frac{\partial P}{\partial z}) + g_P(X, D, P) - (\varepsilon_P + \delta_P + w)P, \\ \frac{\partial u_P}{\partial t} + \text{div}(U_P u_P) &= \mu_u \Delta u_P + \frac{\partial}{\partial z} (v_u \frac{\partial u_P}{\partial z}) - \alpha_u u_P + k_D \text{grad } D + k_X \text{grad } X, \end{aligned} \quad (2.3.17)$$

где  $X, Z, S, D, P$  – концентрации фитопланктона (*Coscinodiscus*), зоопланктона (*Copepoda*), биогенного вещества, детрита, пеленгаса;  $g_i$  – трофические функции для субстанций  $i \in \{D, X, Z, S, P\}$ ;  $\gamma_S$  – доля  $S$ , находящегося в биомассе  $X$ ;  $\varepsilon_Z, \varepsilon_P$  – коэффициенты смертности  $Z$  и  $P$  соответственно;  $\varepsilon_X$  – коэффициент, учитывающий смертность и метаболизм  $X$ ;  $\delta_X$  – убыль  $X$  за счет выедания рыбами  $Z$ ;  $\delta_Z$  – убыль зоопланктона за счет выедания;  $\delta_P$  – убыль пеленгаса за счет выедания;  $\delta_Z$  – убыль  $Z$  за счет выедания;  $\delta_P$  – убыль  $P$  за счет выедания;  $S_P$  – предельно возможная концентрация биогенного вещества;  $f = f(x, y, z, t)$  – функция источника загрязнения;  $B$  – удельная скорость поступления загрязняющего вещества;  $\varepsilon_D$  – коэффициент разложения детрита;  $\beta_D$  – скорость потребления органических остатков пеленгасом;  $\mu_i, \nu_i$  – диффузионные коэффициенты в горизонтальном и вертикальном направлениях субстанции  $i \in \{D, X, Z, S, P\}$  соответственно;  $\Delta$  – двумерный оператор Лапласа;  $u$  – поле скоростей водного потока;  $U = u + u_{0j}$  – скорость конвективного переноса вещества;  $U_P = u + u_P$  – скорость конвективного переноса  $P$ ;  $u_P$  – скорость движения рыбы относительно воды;  $k_D, k_X$  – коэффициенты таксиса;  $\mu_u, \nu_u$  – коэффициенты горизонтальной и вертикальной составляющей диффузии скорости таксиса;  $\alpha_u$  – коэффициент инерционного движения рыбы;  $u_{0j}$  – скорость осаждения  $j$ -й субстанции под действием силы тяжести,  $j \in \{X, Z, S, D\}$ ;  $w(x, y, z, t)$  – доля вылова рыбы.

Введём вектор-функцию  $\varphi = (X, Z, S, D, P)$  и вектор  $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5)$  с пятью компонентами. Тогда начальные и граничные условия для (7) примут вид:

начальные

$$\varphi(x, y, z, 0) = \varphi_0(x, y, z), \quad (x, y, z) \in \bar{G}, \quad (2.3.18)$$

граничные :

$$\varphi = 0 \text{ на } \sigma, \text{ если } u_n < 0; \quad \frac{\partial \varphi}{\partial n} = 0 \text{ на } \sigma, \text{ если } u_n \geq 0; \quad (2.3.19)$$

$$\frac{\partial \varphi}{\partial z} = 0 \text{ на } \Sigma_0; \quad \frac{\partial \varphi}{\partial z} = -\phi \text{ на } \Sigma_H; \quad \phi_i = \varepsilon_i \varphi_i, \quad i = 1, 2, 3, 4, 5.$$

Здесь  $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5$  – неотрицательные постоянные, среди которых  $\varepsilon_1, \varepsilon_3, \varepsilon_5$  учитывают опускание планктона и пеленгаса на дно и их затопление;  $\varepsilon_2, \varepsilon_4$  описывают поглощение биогенного вещества и детрита донными отложениями.

Построим на основе этой модели иерархическую дифференциальную игру двух лиц – ведущего (контрольный орган службы рыбного хозяйства) и ведомого (рыболовное предприятие). Целевые функционалы игроков и ограничения на управления имеют вид:

функционалы ведущего и ведомого соответственно:

$$J_L = \int_0^{\infty} \int_G e^{-rt} [M(P(x, y, z, t) - P^*(x, y, z, t))^2 + C(q(x, y, z, t), w(x, y, z, t))] dv dt \rightarrow \min, \quad (2.3.20)$$

$$J_F = \int_0^{\infty} \int_G e^{-rt} [a(1 - s(x, y, z, t))w(x, y, z, t) - 0,5bw^2(x, y, z, t)]P(x, y, z, t) dv dt \rightarrow \max; \quad (2.3.21)$$

ограничения на управления ведущего при принуждении и побуждении соответственно:

$$0 \leq q(x, y, z, t) \leq 1; \quad 0 \leq t \leq T; \quad (x, y, z) \in \bar{G}, \quad (2.3.22)$$

$$0 \leq s(x, y, z, t) \leq 1; \quad 0 \leq t \leq T; \quad (x, y, z) \in \bar{G}, \quad (2.3.23)$$

и ведомого

$$0 \leq w(x, y, z, t) \leq q(x, y, z, t); \quad 0 \leq t \leq T; \quad (x, y, z) \in \bar{G}. \quad (2.3.24)$$

Здесь, повторимся,  $w(x, y, z, t)$  – доля вылова рыбы (управление ведомого);  $q(x, y, z, t)$  – квота вылова (в долях, управление принуждения ведущего);  $C(q, w)$  – выпуклая функция затрат на контроль выполнения квоты;  $s(x, y, z, t)$  – коэффициент штрафа за перелов (в долях, управление побуждения ведущего);  $a$  – цена единицы биомассы рыбы;  $b$  – коэффициент затрат на вылов;  $P^*(x, y, z, t)$  – оптимальное с точки зрения УР значение биомассы

рыбы;  $M$  – коэффициент штрафа, налагаемого на ведущего при отклонении текущего значения биомассы от идеального.

Таким образом, принуждение в этой модели состоит в назначении квот на вылов рыбы и связано с затратами, а побуждение – в штрафе за перелов. Условие УР имеет вид  $\forall t (x, y, z) \in \bar{G}, P(x, y, z, t) = P^*(x, y, z, t)$  или в более слабой форме  $\forall t (x, y, z) \in \bar{G}, |P(x, y, z, t) - P^*(x, y, z, t)| \leq \delta; \delta \ll 1$ ; при его нарушении на ведущего налагается штраф с коэффициентом  $M$ . В зависимости от реальной ситуации, определяющей информационный регламент игры, стратегии ведущего, направленные на обеспечение УР, могут выбираться из табл. 2.3.6.

Для сравнительного анализа эффективности механизмов управления УР проведены численные расчеты по модели (2.3.17) – (2.3.24). Поскольку аналитическое решение этой модели затруднено, расчеты выполнялись с помощью имитационного моделирования по сценариям, отражающим правдоподобные гипотезы относительно возможных стратегий управления. Рассматривались информационные регламенты  $\Gamma_{1t}, \Gamma_{2t}$  для принуждения ( $C$ ) и побуждения ( $I$ ); регламенты  $\Gamma_{1t}, \Gamma_{2t}$  как более сложные и менее соответствующие практическим процедурам принятия решений пока опущены. Используется небольшое число сценариев, отражающих основные качественно различные управления игроков с учетом специфики моделируемой системы (Ougolnitsky and Usov 2018).

Было взято пять вариантов управлений ведущего в каждой точке пространственной области  $\bar{G}$  в каждый момент времени  $t \in [0, T]$ :

при побуждении ( $q(x, y, z, t) = 1, \forall t \in [0, T]; (x, y, z) \in \bar{G}$ ) –  $s_1(x, y, z, t) \equiv 0,05$  (малый штраф за перелов);  $s_2(x, y, z, t) \equiv 0,25$ ;  $s_3(x, y, z, t) \equiv 0,5$ ;  $s_4(x, y, z, t) \equiv 0,75$  (промежуточные по величине штрафы);  $s_5(x, y, z, t) \equiv 0,95$  (значительный штраф);

при принуждении (функция  $s(x, y, z, t)$  задана)  $q_1(x, y, z, t) \equiv 0,05$  (малая квота вылова);  $q_2(x, y, z, t) \equiv 0,25$ ;  $q_3(x, y, z, t) \equiv 0,5$ ;  $q_4(x, y, z, t) \equiv 0,75$  (промежуточные квоты);  $q_5(x, y, z, t) = 0,95$  (большая).

Для ведомого:

при побуждении рассматривалось пять сценариев:

$w_1(x, y, z, t) \equiv 0,05$  (малый вылов);  $w_2(x, y, z, t) \equiv 0,25$ ;  $w_3(x, y, z, t) \equiv 0,5$ ;  
 $w_4(x, y, z, t) \equiv 0,75$  (промежуточные величины вылова);  $w_5(x, y, z, t) \equiv 0,95$   
(значительный вылов);

при принуждении число сценариев зависит от сценария, выбранного ведущим, а именно при:

1.  $q(x, y, z, t) = q_1(x, y, z, t) \equiv 0,05$  у ведомого есть только один сценарий  $w_1(x, y, z, t) \equiv 0,05$ ;

2.  $q(x, y, z, t) = q_2(x, y, z, t) \equiv 0,25$  — два сценария  $w_1(x, y, z, t) \equiv 0,05$ ;  
 $w_2(x, y, z, t) \equiv 0,25$ ;

3.  $q(x, y, z, t) = q_3(x, y, z, t) \equiv 0,5$  — три сценария  $w_1(x, y, z, t) \equiv 0,05$ ;  
 $w_2(x, y, z, t) \equiv 0,25$ ;  $w_3(x, y, z, t) \equiv 0,5$ ;

4.  $q(x, y, z, t) = q_4(x, y, z, t) \equiv 0,75$  — четыре сценария  $w_1(x, y, z, t) \equiv 0,05$ ;  
 $w_2(x, y, z, t) \equiv 0,25$ ;  $w_3(x, y, z, t) \equiv 0,5$ ;  $w_4(x, y, z, t) \equiv 0,75$ ;

5.  $q(x, y, z, t) = q_5(x, y, z, t) \equiv 0,95$  — пять сценариев  $w_1(x, y, z, t) \equiv 0,05$ ;  
 $w_2(x, y, z, t) \equiv 0,25$ ;  $w_3(x, y, z, t) \equiv 0,5$ ;  $w_4(x, y, z, t) \equiv 0,75$ ;  $w_5(x, y, z, t) \equiv 0,95$ .

Заметим, что рассмотрение небольшого числа вариантов управлений каждого из субъектов (малое, промежуточные, значительное) позволяет качественно учесть весь диапазон возможных стратегий ведомого и ведущего. В результате при побуждении в каждой точке пространственной области  $\bar{G}$  в каждый момент времени  $t \in [0, T]$  получено двадцать пять различных сценариев  $CS_{kl}(x, y, z, t)$ ;  $k, l = 1, 2, 3, 4, 5$ : при принуждении имеем пятнадцать сценариев:  $CS_{kl}(x, y, z, t)$ , а именно при:

принуждении  $CS_{kl}(x, y, z, t) = (q_k(x, y, z, t), w_l(x, y, z, t))$ ,

побуждении  $CS_{kl}(x, y, z, t) = (s_k(x, y, z, t), w_l(x, y, z, t))$ .

Выделены точки пространственной области  $(x_i^0, y_i^0, z_i^0) \in \bar{G}$  ( $i = 1, 2, \dots, L_v$ ) и моменты времени  $0 \leq t_j^0 \leq T$  ( $j = 1, 2, \dots, L_T$ ), в которые может проводиться вылов ( $L_v$  – количество таких пространственных точек;  $L_T$  – число моментов времени).

Предлагаются следующие алгоритмы имитационного моделирования иерархической двухуровневой системы.

Алгоритм имитационного моделирования для модели (2.3.17)–(2.3.24) в случае прямой игры Штакельберга  $\Gamma_{1t}$  и при принуждении, и при побуждении основан на подходе (Угольницкий и Усов 2014) и состоит в следующем.

1. Задаются вид и значения всех входных функций и параметров исследуемой модели.

2. Задаётся текущая стратегия ведущего, в качестве которой выбирается один из сформулированных выше сценариев ( $s_1 - s_5$  или  $q_1 - q_5$ ).

3. Путем перебора возможных сценариев ведомого ( $w_1 - w_5$ ) находится его лучший ответ на выбранную стратегию ведущего – стратегия, при которой целевой функционал (2.3.21) принимает наибольшее значение среди перебираемых. При этом система уравнений (2.3.17) – (2.3.19) решается численно.

4. Сравнивается пара стратегий (текущая стратегия ведущего и оптимальный ответ на нее ведомого) с лучшей (с точки зрения функционала (2.3.20)) для ведущего парой стратегий на данный момент. Если необходимо, то текущая стратегия становится оптимальной.

5. Если число просмотренных сценариев для ведущего не исчерпано, то проводится выбор его новой стратегии и возврат на п. 3 алгоритма.

Теперь формулируем алгоритм имитационного моделирования для модели (2.3.17)–(2.3.24) в случае обратной игры Штакельберга  $\Gamma_{2t}$  для

побуждения и принуждения. Он, как и предыдущий алгоритм, основан на подходе (Угольницкий и Усов 2014) и состоит в следующем.

1. Задаются вид и значения всех входных функций и параметров исследуемой модели.

2. Вводится стратегия наказания ведомого ведущим

при принуждении

$$q^P = \left\{ q^P(x_i^0, y_i^0, z_i^0, t_j^0) \right\}_{i=1; j=1}^{L_V; L_T} : q^P = \arg \inf_{0 \leq w \leq q} J_F(w, x);$$

при побуждении

$$s^P = \left\{ s^P(x_i^0, y_i^0, z_i^0, t_j^0) \right\}_{i=1; j=1}^{L_V; L_T} : s^P = \arg \inf_{0 \leq s \leq 1} J_F(s, w, x) .$$

Для модели (2.3.17)–(2.3.24) эта стратегия на множестве рассматриваемых сценариев находится явно путем полного перебора возможных стратегий субъектов управления и определяется формулами при:

$$\text{побуждении } s^P(x_i^0, y_i^0, z_i^0, t_j^0) = 0,95, \text{ принуждении } q^P(x_i^0, y_i^0, z_i^0, t_j^0) = 0,05 .$$

Путём перебора возможных сценариев ведомого вычисляется величина его гарантированного выигрыша, если он отказывается сотрудничать с ведущим при:

$$\text{принуждении } L_F^{comp} = \sup_{\substack{0 \leq w_k(x, y, z, t) \leq q^P(x, y, z, t) \\ k=1, 2, 3, 4, 5}} J_F(w_k, x);$$

$$w_k^P = \arg \sup_{\substack{0 \leq w_k(x, y, z, t) \leq q^P(x, y, z, t) \\ k=1, 2, 3, 4, 5}} J_F(w_k, x);$$

побуждении

$$L_F^{imp} = \sup_{\substack{w_k(x, y, z, t) \\ k=1, 2, 3, 4, 5}} J_F(w_k, s^P, x);$$

$$w_k^P = \arg \sup_{\substack{w_k(x, y, z, t) \\ k=1, 2, 3, 4, 5}} J_F(w_k, s^P, x).$$

3. Путём перебора возможных сценариев  $CS_{kl}(x, y, z, t)$ ,  $k, l = 1, \dots, 5$ , ведущий находит лучшую с точки зрения (2.3.20) стратегию при условии выполнения неравенства  $J_F(w, s, q) > L_F^{comp}$  (при принуждении) и  $J_F(w, s, q) >$

$L_F^{imp}$  (при побуждении). Эта стратегия называется стратегией поощрения, а последнее неравенство обеспечивает преимущество поощрения перед наказанием. Обозначим решение этой задачи через  $(q^R, w^R)$  (в случае принуждения) и  $(s^R, w^R)$  (при побуждении).

4. Если в п. 3 стратегию поощрения найти не удаётся, т.е. не удаётся на множестве сценариев выполнить неравенство  $J_F(w, s, q) > L_F^{comp}$  (при принуждении) и  $J_F(w, s, q) > L_F^{imp}$  (при побуждении), то равновесие в модели (2.3.17)–(2.3.19) имеет вид:  $(q^P(x_i^0, y_i^0, z_i^0, t_j^0), w^P(x_i^0, y_i^0, z_i^0, t_j^0))$  (при принуждении) и  $(s^P(x_i^0, y_i^0, z_i^0, t_j^0), w^P(x_i^0, y_i^0, z_i^0, t_j^0))$  (при побуждении)  $(x_i^0, y_i^0, z_i^0) \in \bar{G}$  ( $i = 1, 2, \dots, L_V$ );  $0 \leq t_j^0 \leq T$  ( $j = 1, 2, \dots, L_T$ ).

В противном случае равновесие следующее:

$(q^R(x_i^0, y_i^0, z_i^0, t_j^0), w^R(x_i^0, y_i^0, z_i^0, t_j^0))$  (при принуждении) и  $(s^R(x_i^0, y_i^0, z_i^0, t_j^0), w^R(x_i^0, y_i^0, z_i^0, t_j^0))$  (при побуждении).

Для каждого сценария и при принуждении, и при побуждении в случае игр  $\Gamma_{1t}$  и  $\Gamma_{2t}$  система уравнений (2.3.17)–(2.3.19) решается численно на основе неявных консервативных разностных схем, имеющих второй порядок аппроксимации относительно шага пространственной сетки и первый – относительно временного шага (Сухинов и др. 2012). Трёхмерный вектор движения водной среды с учетом реальной геометрии – рельефа дна, формы береговой линии и типичных метеоусловий, характерных для вегетационного периода, стоков рек Дон и Кубань, водообмена через Керченский пролив и других факторов, и оценка коэффициентов микротурбулентного обмена определялись на основе численного моделирования с использованием метода поправки к давлению в блоке гидродинамических моделей (Сухинов и др. 2011). Результаты работы данного блока были входными для блока гидробиологических моделей. Возникающие системы сеточных уравнений, содержащие порядка  $10^6$ – $10^9$  неизвестных и являющиеся плохо обусловленными системами линейных алгебраических уравнений, численно



решались на каждом временном слое вариантом итерационного попеременно-треугольного метода вариационного типа, базирующимся на методе минимальных поправок (Сухинов и Чистяков 2012). Поскольку численное решение данной задачи, в особенности на подробных сетках (порядка  $10^7$  узлов и более), не могло быть получено за приемлемое время на обычных (последовательных) компьютерах, то привлекалась супервычислительная система ЮФУ в г. Таганроге, имеющая 2048 ядер и 4 Тбайт оперативной памяти (Никитина и Семёнов 2014). Задача решается с помощью разработанного параллельного алгоритма, учитывающего архитектуру суперЭВМ с распределенной памятью (Сухинов и др. 2011; Никитина и Семёнов 2014).

Все расчёты проводились для начального распределения полей течений в Азовском море при северном ветре. Временной интервал  $T$  равен 155 суткам (после начала вегетационного периода фитопланктона). Количество точек возможного вылова:  $L_V = 50$  и  $L_T = 20$ . Использовался модельный вид функции штрафа

$$C(q, w) = C_1(q - w)/((w + v)(q + v)); C_1, v = \text{const}; v = 0,001,$$

для которого выполнены все свойства функции затрат на контроль за соблюдением квоты, например, выпуклость по обоим аргументам, убывание при  $w \rightarrow q$ . Здесь  $C_1$  – масштабирующий коэффициент, а число  $v$  обеспечивает выполнение условия  $C(0,0) = 0$ .

Численные расчеты проводились для сценариев, когда стратегии субъектов постоянны во всех точках области  $\bar{G} \quad \forall t$ , т.е.  $CS_{kl}(x, y, z, t) = CS_{kl}^0$  ( $CS_{kl}^0 = \text{const}$ ) в случае  $\mu_D = 5 \cdot 10^{-11} \text{ см}^2 / \text{мин}$ ;  $\nu_D = 10^{-11} \text{ см}^2 / \text{мин}$ ;  $\varepsilon_D = 1,9 \cdot 10^{-5} \text{ мин}^{-1}$ ;  $\beta_D = 0,1 \text{ мин}^{-1}$ ;  $\mu_D = 5 \cdot 10^{-11} \text{ см}^2 / \text{мин}$ ;  $\varepsilon_4 = 0,4 \text{ мин}^{-1}$ ;  $\varepsilon_5 = 0,47 \text{ мин}^{-1}$ ;  $\varepsilon_P = 1,16 \cdot 10^{-3} \text{ мин}^{-1}$ ;  $P^0(x, y, z, t) \equiv 0,09 \text{ мг} / \text{см}^3$ ;  $P(x, y, z, 0) \equiv 0,05 \text{ мг} / \text{см}^3$ ;  $a = b = 2 \text{ тыс. руб}$ ;  $C_1 = 50 \text{ тыс. руб}$ ;  $\xi_P = 0,8 \text{ мин}^{-1}$ ;  $\mu_P = 1,5 \cdot 10^{-3} \text{ см}^2 / \text{мин}$ ;  $\delta_P = 0,05 \text{ мин}^{-1}$ ;  $M = 0,2 \text{ (тыс. руб} \cdot \text{см}^6) / \text{мг}^2$ ;  $\gamma_P = 0,125 \text{ мин}^{-1}$ .

Результаты для информационных регламентов игр  $\Gamma_{1t}$  и  $\Gamma_{2t}$  в случае принуждения и побуждения при выбранном сценарии (стратегии ведущего и ведомого) приведены в табл. 2.3.7. Здесь расходы ведущего и доходы ведомого выражены в тысячах рублей и записаны в виде  $J_L/J_F$ . В случае принуждения предполагалось, что  $s(x, y, z, t) \equiv 0,5$ , при побуждении  $q(x, y, z, t) \equiv 1$ . Прочерки в табл. 2.3.7 означают, что такие сценарии при принуждении невозможны (при побуждении 25 сценариев, а при принуждении – только 15). Так как сценарий фиксирован, то результаты не зависят от выбранного информационного регламента и для игр  $\Gamma_{1t}$  и  $\Gamma_{2t}$  совпадают.

Расходы ведущего и доходы ведомого в “лучшем” из рассмотренных сценариев (двадцати пяти при побуждении и пятнадцати при принуждении) для различных регламентов в случае некоторых дополнительных входных данных приведены в табл. 2.3.8.

Результаты счёта показывают, что равновесия в играх  $\Gamma_{1t}$  и  $\Gamma_{2t}$  могут как совпадать, так и отличаться. Заметим, что функционал (2.3.20) строился при естественных предположениях о свойствах и виде функций, отражающих затраты ведущего в модели управления УР мелководных водоемов. Проведённые расчеты показали, что в данной задаче для широкого класса входных функций, соответствующих реальным практическим данным, при принуждении разницы в использовании информационных регламентов игр  $\Gamma_{1t}$  и  $\Gamma_{2t}$  не наблюдается (если равновесие в игре  $\Gamma_{2t}$  существует). В этом случае предпочтительнее использовать регламент игры  $\Gamma_{1t}$ , так как он предполагает меньшую зависимость ведомого от ведущего и реализуется для любого вида входных функций.

При побуждении независимо от входных параметров модели информационный регламент игры  $\Gamma_{1t}$  предпочтительнее для ведомого (в смысле получаемого дохода), а игры  $\Gamma_{2t}$  – для ведущего.

## Результаты расчётов по сценариям

Регламенты	Сценарии				
	$CS_{11}^0$	$CS_{12}^0$	$CS_{13}^0$	$CS_{14}^0$	$CS_{15}^0$
$C_{1t}, C_{2t}$	6/20	--	--	--	--
$I_{1t}, I_{2t}$	335635/40	53851/228	18021/580	6207/1430	7820/7474
	$CS_{21}^0$	$CS_{22}^0$	$CS_{23}^0$	$CS_{24}^0$	$CS_{25}^0$
$C_{1t}, C_{2t}$	17723/20	12/103	--	--	--
$I_{1t}, I_{2t}$	335635/31	53851/172	18021/414	6207/933	7820/4330
	$CS_{31}^0$	$CS_{32}^0$	$CS_{33}^0$	$CS_{34}^0$	$CS_{35}^0$
$C_{1t}, C_{2t}$	79576/20	8995/103	39/207	--	--
$I_{1t}, I_{2t}$	335635/20	53851/103	18021/207	6207/311	7820/313
	$CS_{41}^0$	$CS_{42}^0$	$CS_{43}^0$	$CS_{44}^0$	$CS_{45}^0$
$C_{1t}, C_{2t}$	185547/20	26941/103	6785/207	209/311	--
$I_{1t}, I_{2t}$	335635/9	53851/34	18021/0	6207/-312	7820/-3542
	$CS_{51}^0$	$CS_{52}^0$	$CS_{53}^0$	$CS_{54}^0$	$CS_{55}^0$
$C_{1t}, C_{2t}$	302088/20	47752/103	15415/207	4767/311	6333/393
$I_{1t}, I_{2t}$	335635/1	53851/-20	18021/-165	6207/-808	7820/-3542

С ростом коэффициента штрафа, налагаемого на ведущего при отклонении текущего значения биомассы от идеального, цены единицы

биомассы рыбы или при уменьшении коэффициента затрат на вылов разница при использовании информационных регламентов игр  $\Gamma_{1t}$  и  $\Gamma_{2t}$  становится более существенной.

С ростом коэффициента затрат на вылов или коэффициента штрафа, налагаемого на ведущего при отклонении текущего значения биомассы от идеального, доходы обоих субъектов при разных информационных регламентах падают, причем в случае регламента игры  $\Gamma_{1t}$  доходы падают более резко. С ростом цены единицы биомассы рыбы доход ведомого растет, ведущего может не меняться, а может и расти.

Таблица 2.3.8

Результаты расчётов в лучшем из рассмотренных сценариев для различных регламентов

Регламенты	Примеры					
	Пример	$M = 20000$ ·руб см <sup>6</sup> мг <sup>2</sup>	$M = 20$ руб см <sup>6</sup> мг <sup>2</sup>	$b = 20$ .руб	$b =$ 20000.руб	$b =$ 10000.руб
$C_{1t}$	6/20	602/20	0.6/20	6/21	6/10	6/108
$C_{2t}$	6/20	602/20	0.6/20	6/21	6/10	6/108
$I_{1t}$	7280/7474	55132/34	1580/7479	7280/7479	335635/30	7280/67314
$I_{2t}$	6207/1430	21967/580	1580/7479	6207/2354	335635/30	6207/10885

Наряду с двадцатью пятью сценариями при побуждении и пятнадцатью при принуждении, расчёты проводились и при меньшем числе сценариев ведущего и ведомого при побуждении бралось три сценария ведущего и ведомого (малое, среднее, значительное). При принуждении у ведомого в зависимости от сценария, выбранного ведущим, бралось не более трех сценариев. В этом случае разницы при использовании информационных

регламентов игр  $\Gamma_{1t}$  и  $\Gamma_{2t}$  не наблюдалось ни при принуждении, ни при побуждении для широкого класса входных функций.

На рис. 2.3.1 показано отклонение численности рыбной популяции от начального состояния в фиксированной точке пространства для входных данных примера с течением времени. На рис. 2.3.2 – то же отклонение, но в зависимости от цены единицы биомассы  $a$ . Здесь сплошная линия соответствует информационному регламенту  $I_{1t}$ , штриховая –  $I_{2t}$ , а пунктирная – регламентам  $C_{1t}$  и  $C_{2t}$ .

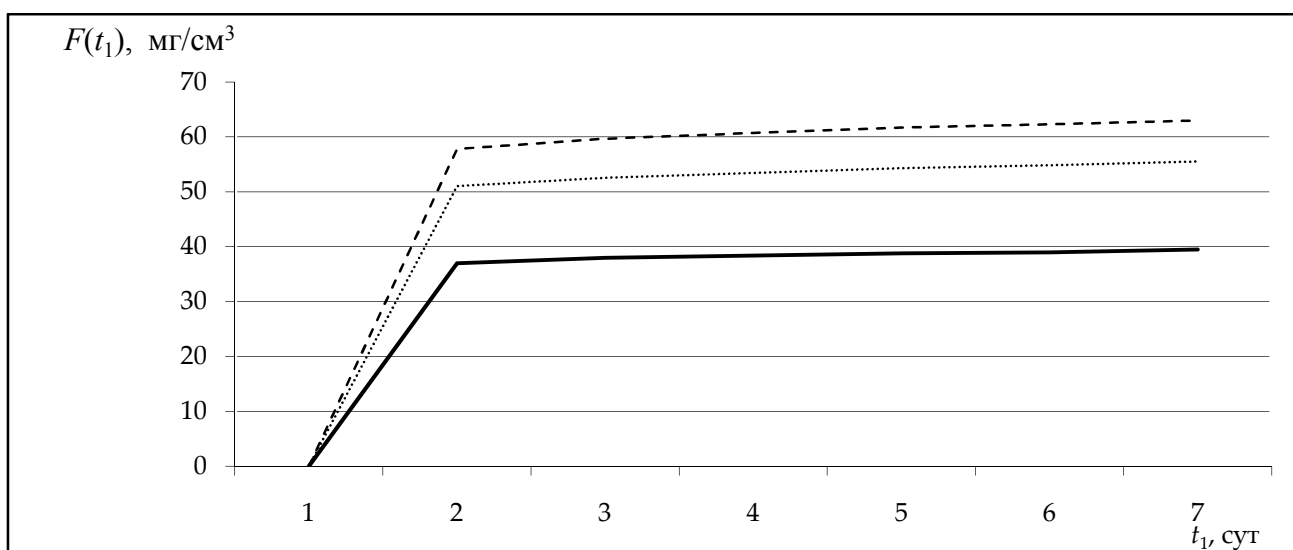


Рис. 2.3.1. Изменение отклонения численности рыбной популяции от начального состояния в фиксированной точке пространства для входных данных примера с течением времени. Здесь

$$t_1 = 200 \cdot t; F(t_1) = 10 \cdot (P(x, y, z, t_1) - P(x, y, z, 0)).$$

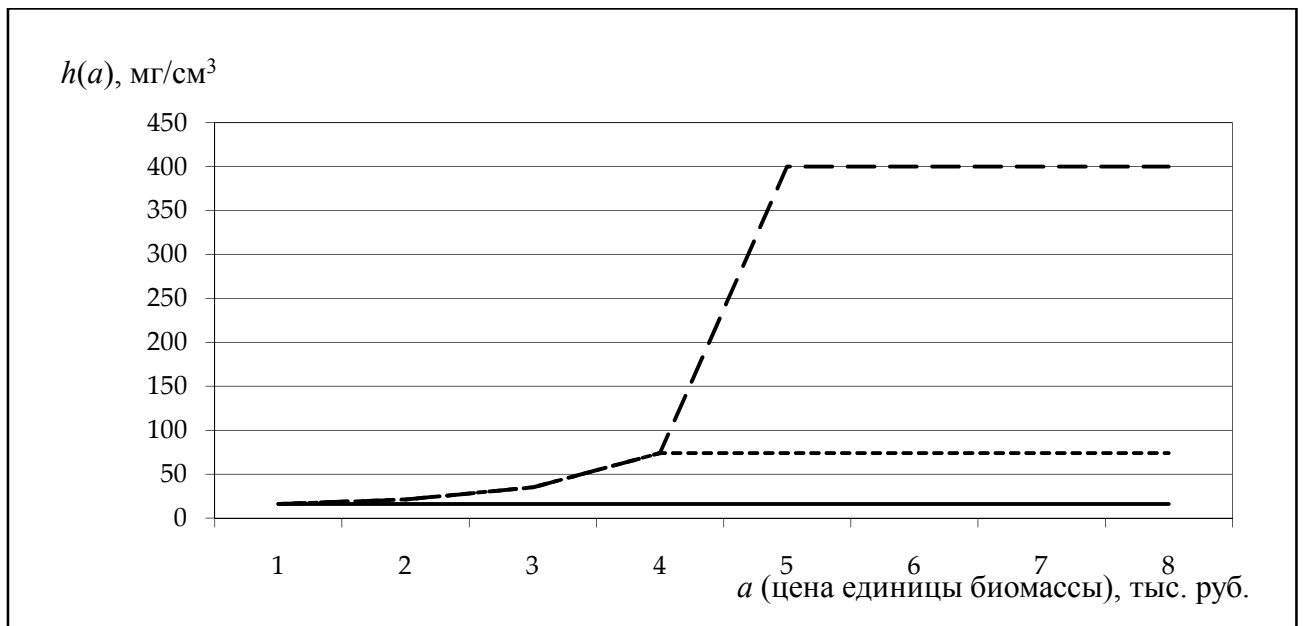


Рис. 2.3.2. Изменение отклонения численности рыбной популяции от начального состояния в фиксированной точке пространства в фиксированный момент времени для входных данных примера с ростом цены единицы биомассы. Здесь  $h(a) = 100 \cdot (P(x, y, z, t) - P(x, y, z, 0))$ .

Результаты счёта показывают, что минимальная численность рыбной популяции наблюдается при использовании информационных регламентов  $C_{1t}$  и  $C_{2t}$ , а максимальная –  $I_{1t}$ . С течением времени численность популяции в рассмотренном примере вначале резко возрастает, а потом держится на одном уровне. С ростом цены единицы биомассы численность популяции не уменьшается при любом информационном регламенте.

Итак, даны уточнённые по сравнению с (Угольницкий и Усов 2014) определения и проведено исследование равновесий принуждения и побуждения с учетом требований УР в двухуровневых системах управления для различных информационных регламентов соответствующих иерархических дифференциально-игровых моделей. Методы принуждения и побуждения при иерархическом управлении (Угольницкий 2016) формализуются как информационные расширения игры двух лиц, основанные на регламентах  $\Gamma_{1t}$ ,  $\Gamma_{2t}$ , описанных в (Кононенко 1977; Горелик и

др. 1991). Требования УР учитываются дополнительно как фазовые ограничения в дифференциальной игре. По сравнению с предыдущими публикациями решения иерархических дифференциальных игр находились посредством имитационного моделирования по специальным регламентам.

Основные результаты статьи относятся к сравнительному анализу указанных решений игр, полученных на примере численного исследования модели экосистемы мелководного водоема (2.3.17)–(2.3.24). В общем случае реализация различных информационных регламентов приводит к разным результатам как при принуждении, так и при побуждении. При небольшом числе сценариев у ведущего разницы в использовании информационных регламентов игр  $\Gamma_{1t}$  и  $\Gamma_{2t}$  может не быть. В этом случае информационный регламент игры  $\Gamma_{1t}$  предпочтительнее, поскольку в игре  $\Gamma_{2t}$  не всегда удаётся построить решение, для этого у ведущего должны быть значительные рычаги воздействия на ведомого. Тогда ведущему удаётся сделать для ведомого стратегию поощрения предпочтительнее стратегии наказания. Равновесия при разных информационных регламентах зависят от входных параметров модели, причём в общем случае информационный регламент игры  $\Gamma_{2t}$  для ведущего предпочтительнее, чем  $\Gamma_{1t}$ . Его расходы в этом случае не выше, чем при регламенте игры  $\Gamma_{1t}$ . Для ведомого – наоборот: предпочтительнее (в смысле получаемого дохода) информационный регламент игры  $\Gamma_{1t}$ . Реализация информационного регламента игры  $\Gamma_{2t}$  возможна только, если у ведущего больше “власти” над ведомым, чем в игре  $\Gamma_{1t}$ .

Полученные результаты могут найти практическое применение при решении задач управления УР реальных динамических активных систем (эколого-экономических, организационных, социальных) (Sukhinov et al., *Automation and Remote Control*, 2017, 78(6), 1059-1071).

### ГЛАВА 3. МЕТОДЫ И АЛГОРИТМЫ УПРАВЛЕНИЯ УСТОЙЧИВЫМ РАЗВИТИЕМ АКТИВНЫХ СИСТЕМ

Третья глава монографии посвящена методам решения задач управления устойчивым развитием активных систем и соответствующим алгоритмам. Первые два параграфа описывают применение методов случайного поиска (генетические алгоритмы, алгоритм имитации отжига, бинарное деление) к решению дифференциальных иерархических игр, в том числе моделирующих механизмы стимулирования при распределении ресурсов. Генетический алгоритм и алгоритм имитации отжига основаны на идее о том, что управление Центра обладает "липшицевым" свойством, т.е. не может меняться резко, а либо сохраняется, либо увеличивается или уменьшается на небольшую заданную величину. Тогда использовать стратегию альтернативы поощрения и сурового наказания (как в стандартных моделях стимулирования) не удаётся, и надо искать другие, адаптивные способы решения задачи. Бинарное деление лишено этого ограничения, а именно, управление центра может быть кусочно-непрерывной функцией. Однако алгоритм бинарного деления накладывает дополнительные ограничения на функционал ведущего. Во втором параграфе условие Липшица сочетается с использованием онлайн-подхода, когда Центр пошагово "обучается" на неизвестных ему заранее стратегиях оптимальных ответов агентов. Дается сравнительный анализ результатов численных расчётов на иллюстративных примерах.

Третий параграф посвящён методу качественно репрезентативных сценариев имитационного моделирования. Основная трудность применения мощного метода компьютерной имитации состоит в необходимости обоснования перехода от рассмотрения необозримого или даже бесконечного полного факторного плана вычислительных экспериментов к разумным неполным. Идея предлагаемого метода состоит в том, что для получения вполне приемлемой качественной картины прогноза достаточно



использовать небольшое число значений каждой управляющей переменной, отражающих характерные принципиально различные варианты воздействия на управляемую динамическую систему. Описывается формализация этой идеи, приводятся примеры её реализации для различных моделей.

### 3.1. Алгоритмы случайного поиска

В последние десятилетия серьёзное внимание привлекают методы эволюционного моделирования (Гладков и др., 2006; Емельянов и др., 2003). Основной замысел работы (Белявский и др. 2016) состоит в использовании генетических алгоритмов при решении иерархических дифференциальных игр с различными информационными регламентами. Этот подход перекликается с методами эволюционных игр (Game Theory, 2012).

Для описания вычислительного метода остановимся на прямой игре Штакельберга при побуждении и программной стратегии ведущего (см. пункт 1.2):

$$J_L = \int_0^T e^{-\rho t} g_L(p(t), u(t), x(t)) dt + e^{-\rho T} S_L(x(T)) \rightarrow \max, p(t) \in P$$

$$J_F = \int_0^T e^{-\rho t} g_F(p(t), u(t), x(t)) dt + e^{-\rho T} S_F(x(T)) \rightarrow \max, u(t) \in U$$

$$\dot{x} = f(x(t), u(t)), \quad x(0) = x_0.$$

Здесь  $L$  – ведущий,  $F$  – ведомый. Далее описываются два варианта применения генетического алгоритма.

*Смешанный вариант.* Рассмотрим регламент  $\Gamma_{it}$  при следующих условиях:

а) стратегия лидера  $p$  удовлетворяет условию Липшица на интервале  $[0, T]$ :

$$|p(t) - p(s)| \leq \alpha |t - s|; \quad (3.1.1)$$

б) существует эффективный алгоритм решения редуцированной задачи. Второе условие, которое будет конкретизировано позже, подчёркивает, что основная вычислительная сложность задачи состоит в нахождении

оптимальной стратегии ведущего. Остановимся на первом условии, которое означает, что существует такая положительная константа  $\alpha$ , для которой при  $\forall t \in [0, T]$  выполняется неравенство (3.1.1), утверждающее, что возможности ведущего ограничены. Определим кусочно-постоянную траекторию. Для этого рассмотрим равномерное разбиение интервала  $[0, T]$  с шагом разбиения  $h = \frac{T}{N}$ . Кусочно-постоянная траектория определяется рекуррентно

следующими равенствами:

$$\begin{aligned} p_N(x, \Delta, t) &= x, t \in [0, t_1], p_N(x, \Delta, t) = p_N(x, \Delta, t_k) + \alpha h \delta_{k-1}, \\ t &\in (t_k, t_{k+1}], k = 1, \dots, N-1. \end{aligned} \quad (3.1.2)$$

В (3.1.2)  $\Delta = (\delta_i)_{i=1}^{N-2}, \delta_i \in \{-1, 0, 1\}$ . Для дальнейшего понадобится следующий результат.

*Теорема 3.1.1* (Белявский и др. 2016). Для любого положительного числа  $\varepsilon$  существуют такие  $N^*, x^*$  и  $\Delta^*$ , что  $\sup_{t \in [0, T]} |p(t) - p_N(x^*, \Delta^*, t)| \leq \varepsilon$ .

Пример аппроксимации приведён на рис. 3.1.1.

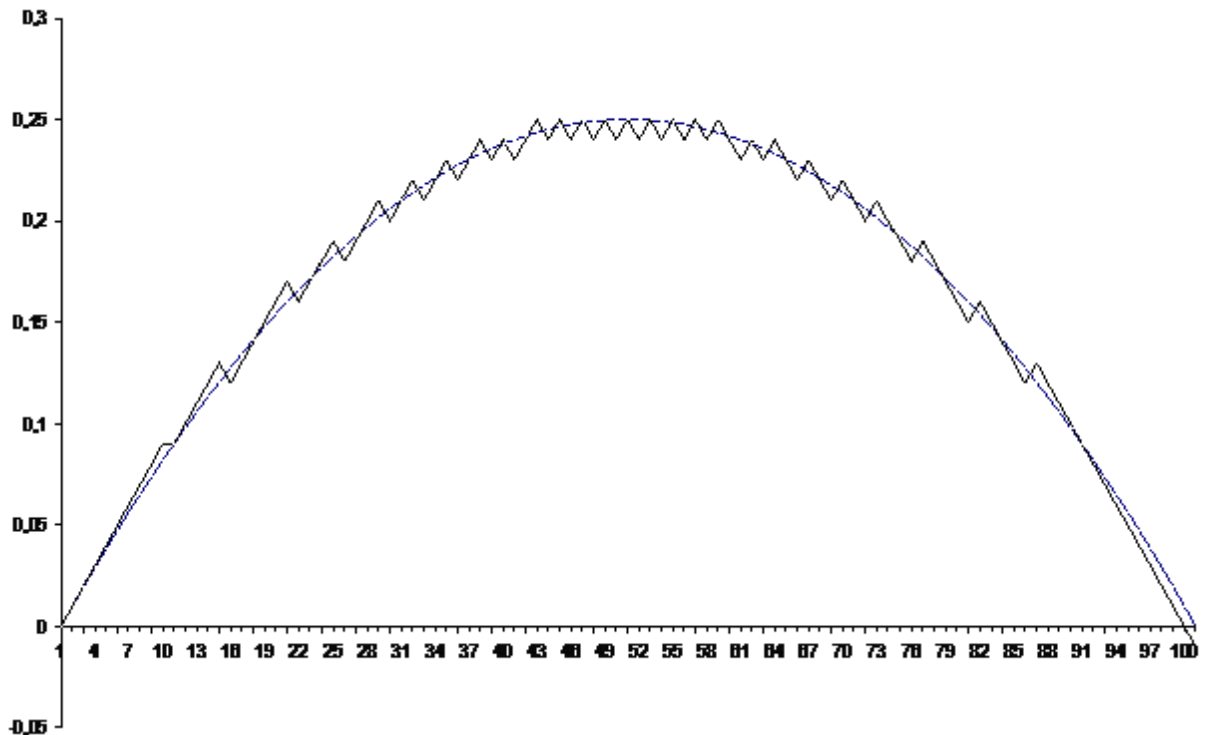


Рис. 3.1.1. Аппроксимация функции  $x(1-x)$  на интервале  $[0,1]$

Теорема 3.1.1 в сочетании с условием Липшица для функции  $g_F(x, y, z)$  по переменной  $y$  позволяет рассмотреть приближённую задачу:

$$J_L = \int_0^T e^{-\rho t} g_F(p_N(x, \Delta, t), u(t), x(t)) dt + e^{-\rho T} S_L(x(T)) \rightarrow \max, x, \Delta; \quad (3.1.3)$$

$$J_F = \int_0^T e^{-\rho t} g_L(p_N(x, \Delta, t), u(t), x(t)) dt + e^{-\rho T} S_L(x(T)) \rightarrow \max, u(t) \in U.$$

$$\dot{x} = f(x(t), u(t)), \quad x(0) = x_0.$$

Характер вектора  $\Delta$  позволяет использовать для решения задачи (3.1.3) стандартный генетический алгоритм следующего вида.

1. Формирование случайным образом начальной популяции  $\Pi = (\Delta_i)_{i=1}^M$ .  
Вычисление здоровья каждой особи этой популяции.
2. Начало цикла.
3. Размножение.
4. Мутация.
5. Вычисление здоровья каждой новой особи.
6. Формирование новой популяции (селекция).
7. Если выполняются условия остановки, то конец цикла, иначе на начало цикла.

Остановимся на пункте 5. Чтобы вычислить здоровье особи  $\Delta$ , необходимо решить следующую редуцированную задачу:

$$J_L = \int_0^T e^{-\rho t} g_F(p_N(x, \Delta, t), u(t), x(t)) dt + e^{-\rho T} S_L(x(T)) \rightarrow \max, x; \quad (3.1.4)$$

$$J_F = \int_0^T e^{-\rho t} g_L(p_N(x, \Delta, t), u(t), x(t)) dt + e^{-\rho T} S_L(x(T)) \rightarrow \max, u(t) \in U;$$

$$\dot{x} = f(x(t), u(t)), \quad x(0) = x_0$$

Редукция заключается в том, что задача ведущего становится одномерной. Ранее было приведено условие: для редуцированной задачи существует

эффективный способ её решения. Если это условие не выполняется, то при дополнительных предположениях о том, что функция  $u(t)$  удовлетворяет условию Липшица и функция  $g_F(x, y, z)$  удовлетворяет условию Липшица по переменной  $y$ , задача ведомого заменяется на приближённую:

$$J_F = \int_0^T e^{-\rho t} g_F(p_N(x, \Delta, t), u_{\bar{N}}(\bar{x}, t, \bar{\Delta}), x(t)) dt + e^{-\rho T} S_F(x(T)) \rightarrow \max, \bar{x}, \bar{\Delta};$$

и генетический алгоритм можно применить для задачи ведомого. Такой вычислительный метод, при котором генетический алгоритм применяется последовательно как для задачи ведущего, так и для задачи ведомого, естественно назвать *общим генетическим алгоритмом*.

К основным достоинствам генетического алгоритма относится прежде всего то, что для его применения не требуется каких-либо особых свойств целевого функционала. Недостаток алгоритма – медленная скорость сходимости, как и для всякого случайного поиска. Поэтому смешанный алгоритм, в котором случайный поиск сочетается с традиционными вычислительными методами, выглядит более перспективным по сравнению с общим генетическим алгоритмом (Белявский и др. 2012). В генетическом алгоритме на каждой итерации получается более здоровая популяция, поскольку отсеиваются нездоровые особи, что приводит в зависимости от задачи к увеличению или уменьшению функционала на каждом шаге. Поэтому эволюционное моделирование выглядит предпочтительней по сравнению с имитационным при большом числе сценариев.

Рассмотрим следующий иллюстративный пример, основанный на работе (Kornienko and Ougolnitsky 2014). Пусть динамика некоторой эксплуатируемой популяции описывается моделью Мальтуса:  $\dot{x} = (a(x) - u(t))x(t)$ ,  $x(0) = x_0$ , где  $x(t)$  – биомасса популяции;  $u(t) \in [0,1]$  – доля биомассы, собираемой в урожай,  $a(x)$  – убывающая функция. Устойчивость развития характеризуется величиной  $\Phi(|x(t) - x_0|)$ , где  $\Phi(x)$  – возрастающая

неотрицательная функция. Функционал ведущего –

$$\int_0^T [ku(t)x(t) - cq^2(t) - \Phi(|x(t) - x_0|)]dt, \text{ функционал ведомого} - J_F = \int_0^T u(t)x(t)dt.$$

Игра Штакельберга на основе принуждения записывается как

$$\begin{aligned} J_L &= \int_0^T [ku(t)x(t) - cq^2(t) - \Phi(|x(t) - x_0|)]dt \rightarrow \max_{0 \leq q \leq 1}, \\ J_F &= \int_0^T u(t)x(t)dt \rightarrow \max_u, \\ \dot{x} &= (a(x) - u(t))x(t), x(0) = x_0, \\ 0 &\leq u(t) \leq 1 - q(t). \end{aligned} \tag{3.1.5}$$

Оптимальное решение задачи ведомого при фиксированном  $q : u^* = 1 - q$ .

В результате задача ведущего имеет вид:

$$\begin{aligned} J_L &= \int_0^T [k(1 - q(t))x(t) - c(1 - q(t))^2 - \Phi(|x(t) - x_0|)]dt \rightarrow \max_{0 \leq q \leq 1}, \\ \dot{x} &= (a(x) - 1 + q(t))x(t), x(0) = x_0. \end{aligned} \tag{3.1.6}$$

Применение генетического алгоритма предполагает решение одномерной редуцированной задачи:

$$\begin{aligned} J_L^\Delta(y) &= \int_0^T [k(1 - q(y, \Delta, t))x_N(y, \Delta, t) - c(1 - q(y, \Delta, t))^2 - \Phi(|x_N(y, \Delta, t) - x_0|)]dt \rightarrow \max_{y \in D_\Delta} \tag{3.1.7} \\ \dot{x}_N &= (a(x_N(y, \Delta, t)) - 1 + q_N(y, \Delta, t))x_N(y, \Delta, t), x_N(y, \Delta, 0) = x_0. \end{aligned}$$

В (3.1.7)  $D_\Delta = \{y : 0 \leq q_N(y, \Delta, t) \leq 1\}$ . Если  $D_\Delta = \emptyset$ , то соответствующей особи присваивается значение здоровья, равное  $-\infty$ .

Рассмотрим дифференциальное уравнение в задаче (3.1.7) в предположении, что  $a$  удовлетворяет условию Липшица. Это предположение позволяет получить приближённое решение уравнения. Определим последовательность  $x_N^i(y, \Delta, t)$  при помощи рекуррентных уравнений:

$$x_N^0(y, \Delta, t_0) = x_0, x_N^i(y, \Delta, t) = x_N^{i-1}(y, \Delta, t_i) \exp((A_{i-1} - 1 + B_{i-1})(t - t_i)), i = 1, \dots, N. \tag{3.1.8}$$

В (3.1.8) использованы следующие обозначения:  $A_j = a(x_N^j(y, \Delta, t_j)), B_j = \alpha \sum_{i=1}^j \delta_i$ .

Приближённое решение дифференциального уравнения выражается через данную последовательность следующим образом:

$$x_N(t) = x_N^i(y, \Delta, t), t \in [t_i, t_{i+1}], i = 0, \dots, N-1. \quad (3.1.9)$$

Функционал ведущего для приближённого решения дифференциального уравнения выглядит следующим образом:

$$J_L^\Delta(y) = \sum_{i=1}^N \left[ \frac{k(1-y-B_{i-1})x_{N-1}^i(y, \Delta, t_i)[\exp(A_{i-1}-1+y+B_{i-1})h-1]}{A_{i-1}-1+y+B_{i-1}} \right] + \quad (3.1.10)$$

$$+ h \sum_{i=1}^N \left[ \Phi(|x_N^i(y, \Delta, t_i) - x_0|) - c(1-y-B_{i-1})^2 \right]$$

Параметры модели приведены в таблице.

$T$	1
$a(x)$	$a_0 - a_1x$
$x_0$	1
$\Phi( x )$	$M( x )$

На рис. 3.1.2 приведены результаты расчётов.

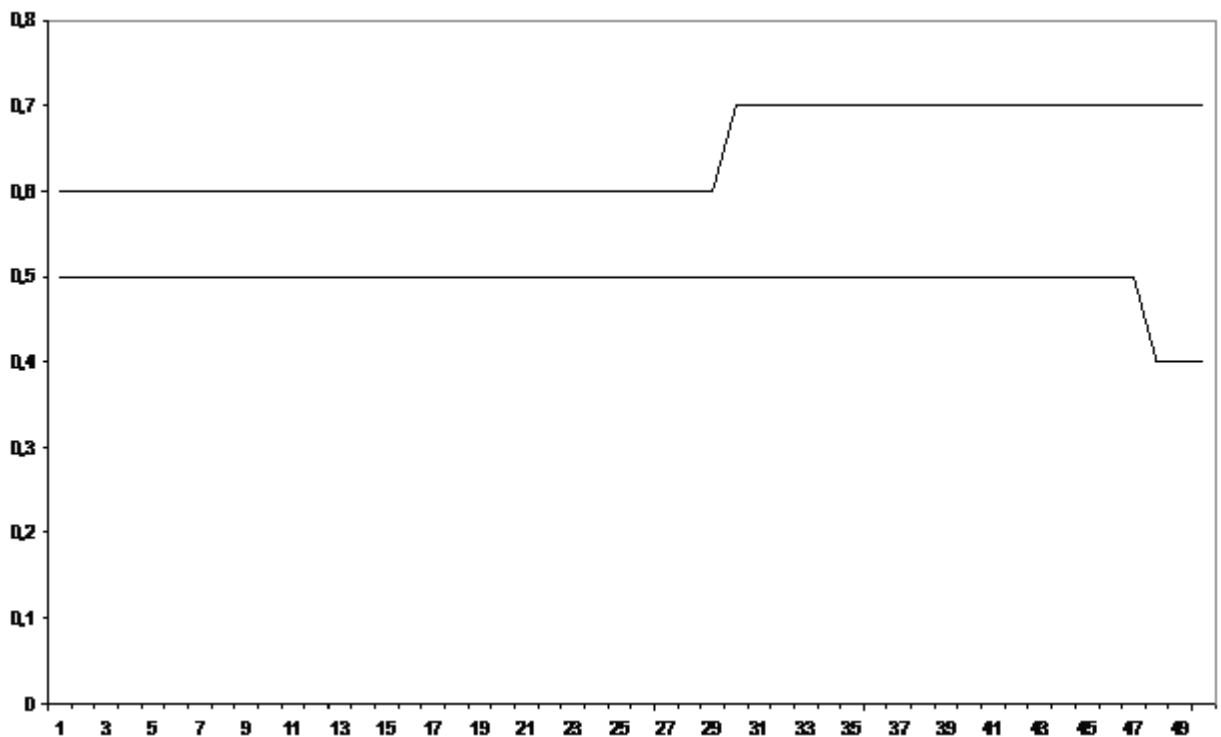


Рис. 3.1.2. Результаты расчётов (верхняя линия соответствует значениям:  $a_0 = 0.5, a_1 = 0.1, M = 5$ ; нижняя линия –  $a_0 = 0.5, a_1 = 0.01, M = 10$ ; в первом случае значение целевого функционала ведущего равно 0.46, во втором 0.48). Вычисления выполнены при  $k = 1, c = 0.1$ .

При  $M \rightarrow \infty, a_1 \rightarrow 0$  равновесное решение будет стремиться к  $(1 - a_0, a_0)$  (см. Kornienko and Ougolnitsky 2014), что соответствует графику оптимального решения, полученного генетическим алгоритмом (Белявский и др. 2016).

В статье (Белявский и др. 2018) предлагается динамическая теоретико-игровая постановка задачи распределения ресурсов в организационной системе. Рассматривается применение алгоритмов эволюционного моделирования к решению таких задач. Распределению ресурсов посвящена обширная литература по теоретико-игровому моделированию. Собственно, вся теория кооперативных игр занимается решением именно этой проблемы - распределению выигрыша максимальной коалиции между всеми игроками (Печерский, Яновская 2004; Petrosjan, Zenkevich 2016).

В экономике общественных благ исследуется проблема распределения индивидуальных ресурсов агентов между их деятельностью в частных интересах и производством общественного блага (Bergstrom et al. 1986; Voadway et al. 1989a,b). Это направление активно развивается и сейчас. Так, статья (Christodoulou et al. 2015) посвящена механизму пропорционального распределения делимых ресурсов, а в работе (Kahana, Klunover 2016) получены условия оптимального распределения ресурсов между трудовой деятельностью и отдыхом в случае, когда агенты имеют одинаковые функции выигрыша, но различные возможности и нетрудовые доходы. Обзор динамических постановок таких задач приведён в монографии (Long 2010).

В российской литературе основополагающая роль в этой области принадлежит статье (Гермейер, Ватель 1974), где изучены модели, в которых функции выигрыша всех агентов состоят из двух частей – общественной

(одинаковой для всех агентов) и частной составляющей. Гермейер и Ватель (1974) показали, что если эта функция имеет вид свёртки по минимуму, то при естественных предположениях в игре существует Парето-оптимальное равновесие по Нэшу. Исследование игр с учетом согласования частных и общественных интересов продолжено, например, в работах (Кукушкин 1991; Kukushkin 1994; Gorbaneva, Ougolnitsky 2018).

Механизмы распределения ресурсов в статической постановке служат предметом изучения теории управления организационными системами (Новиков 2007). Динамические модели распределения ресурсов в этой теории практически не рассматриваются, и наша постановка предлагается впервые. В силу сложности возникающей дифференциальной игры для её решения используются генетические алгоритмы.

Эволюционное моделирование и генетические алгоритмы описаны в монографиях (Емельянов и др. 2003; Гладков и др. 2006). В статье (Белявский и др. 2012) представлен алгоритм гибридного метода обучения искусственных нейронных сетей. В работе авторов (Белявский и др. 2016) предложен генетический алгоритм решения игры Гермейера с одним ведомым игроком и управлением, удовлетворяющим условию Липшица. Альтернативный подход к решению иерархических дифференциальных игр изложен в (Угольницкий и Усов 2013).

Предлагаемый подход к теоретико-игровому моделированию распределения ресурсов в иерархической системе "Центр - агенты" обладает следующими особенностями:

- 1) явно описывается динамика ресурса в зависимости от управления Центра;
- 2) Центр распределяет ресурс между агентами пропорционально их действиям, что побуждает агентов к выбору более напряжённого плана;
- 3) управление Центра не может меняться резко, что формализуется как липшицево свойство функции управления. Представляется, что это



предположение выполняется для большинства реальных организационно-экономических систем;

4) на основе предыдущей гипотезы разрабатывается генетический алгоритм нахождения оптимальной стратегии Центра в иерархической дифференциальной игре.

Согласно теореме 3.1.1, если  $|p(t) - p(s)| \leq \alpha|t - s|, \forall t, s \in [0, T]$ , то для любого положительного числа  $\varepsilon$  существуют такие  $N^*, x^*$  и  $\Delta^*$ , что

$\sup_{t \in [0, T]} |p(t) - p_N(x^*, \Delta^*, t)| \leq \varepsilon$ , где  $p_N$  - кусочно-постоянная функция вида

$$p_N(x, \Delta, t) = x, t \in [0, t_1] \quad p_N(x, \Delta, t) = p_N(x, \Delta, t_k) + \alpha h \delta_{k-1},$$

$$t \in (t_k, t_{k+1}], k = 1, \dots, N-1; \quad \Delta = (\delta_i)_{i=1}^{N-2}, \delta_i \in \{-1, 0, 1\}.$$

При этом

$$x^* = p(0), N^* = \left\lceil \frac{2\alpha T}{\varepsilon} \right\rceil + 1, \delta_1^* = \text{sign}(p(t_1) - p(0)), \delta_k^* = \text{sign}(p(t_k) - p_k(p(0), \Delta^{k,*}, t_k)),$$

$$k = 2, \dots, N-1.$$

При равномерном разбиении погрешность связана с длиной интервала разбиения  $h$  равенством  $\alpha h = 2\varepsilon$ . Если исходить из того, что константа Липшица  $\alpha$  заранее неизвестна, то и число интервалов не может быть выбрано. В подобных случаях во многих вычислительных методах, например, при численном интегрировании, применяется метод половинного деления и сравнения результатов. Если результаты совпадают, то процесс прекращается. Этот метод естественно внедряется в предложенный в (Белявский и др. 2016) алгоритм.

В генетических алгоритмах половинное деление интервалов разбиения приводит к удвоению битов хромосомы. Естественно при этом предусмотреть наследование оптимальной популяции, которое может быть реализовано различными способами. Один из них заключается в следующем.

Пусть оптимальная популяция на  $j$ -й итерации  $M_j = (\Delta_k^j)_{k=1}^L$ . На следующей итерации начальная популяция формируется так, что каждый бит

хромосомы  $\Delta_k^j$  транслируется в два бита:  $\delta_{k,i}^j \rightarrow (\xi, \delta_{k,i}^j)$  и  $\delta_{k,i}^j \rightarrow (\delta_{k,i}^j, \xi)$ . Случайная величина  $\xi$  принимает одно из трёх значений  $\{-1, 0, 1\}$  с вероятностями  $\{q_1, q_2, q_3\}$ , которые, как и число  $L$ , относятся к параметрам алгоритма. В результате происходит удвоение особей в начальной популяции, что выглядит естественно в связи с расширением зоны поиска решения.

Возможно неравномерное разбиение интервала. Пусть длина  $i$ -го интервала разбиения  $[t_i, t_{i+1})$  равна  $h_i$  и локальная константа Липшица

$$\alpha_i = \sup_{s, t \in [t_i, t_{i+1}], s \neq t} \frac{|p(t) - p(s)|}{|t - s|}. \quad \text{В этом случае погрешность } 2 \max_i \alpha_i h_i = \varepsilon.$$

Естественно строить неравномерное разбиение таким образом, чтобы  $\max_i \alpha_i h_i$  был минимальным на множестве возможных разбиений. Вычисление такого разбиения может оказаться более сложной задачей, чем начальная, для которой применяется генетический алгоритм. Однако, в некоторых случаях хорошее разбиение всё же удастся построить.

Пусть  $p(t) = A \ln(1+t)$ ,  $A \neq 0$  и разбиение получается в результате использования рекуррентных равенств:

$$h_i = h(1+t_{i-1}), t_i = t_{i-1} + h_i, i = 1, 2, \dots, k; t_0 = 0; k = \max\{j: t_j \leq T\}.$$

Последний интервал разбиения:  $[t_k, T]$ . Для этого разбиения  $\varepsilon = 2 \max_i \alpha_i h_i = 2|A|h$ , где  $h$  - параметр разбиения. Аналогичного результата можно достичь, если подобрать такую константу  $A > 0$ , что на интервале

$$[0, T] \sup_{t \neq s} \frac{|p(t) - p(s)|}{|\ln(t) - \ln(s)|} \leq A, \quad \text{так как для всех } t \text{ и } s > t \text{ из интервала } [0, T]$$

справедливо:  $p(t) - A(\ln s - \ln t) \leq p(s) \leq p(t) + A(\ln s - \ln t)$ .

К сожалению, константа  $A$ , так же как и константа Липшица  $\alpha$ , может быть неизвестной, поэтому следует увеличивать число интервалов разбиения и сравнивать полученные результаты. Это увеличение достигается за счёт

постепенного уменьшения параметра разбиения  $h$ . Наследование оптимальной популяции при данном неравномерном распределении невозможно. Положительно то, что число интервалов разбиения растёт линейно.

Рассмотрим динамическую задачу распределения ресурсов в организационной системе и генетический алгоритм ее решения. Имеется ведущий игрок (Центр), который распределяет ресурс между агентами, чем компенсирует их затраты. Ресурс (затраты Центра) рассматривается как фазовая переменная. Количество ресурса, выделяемого в момент времени  $t$ , равно  $m(t)$ . Динамика ресурса определяется дифференциальным уравнением:

$$dm(t) = \sigma m(t) dp(t), \quad (3.1.11)$$

с начальным значением  $m_0$ . Параметр  $\sigma$  - локальная изменчивость ресурса,  $p(t)$ - управление Центра, о котором шла речь в предыдущем разделе. Дифференциал управления  $dp(t)$  может не существовать, поскольку не всякая функция, удовлетворяющая условию Липшица, дифференцируема. Поэтому (3.1.11) следует понимать в интегральном смысле:

$$m(t) = m_0 + \sigma \int_0^t m(s) dp(s), \quad \text{где используется интеграл Лебега-Стилтьеса.}$$

Аналогичным (3.1.11), но не совпадающим с ним точно, будет определение ресурса следующим образом:  $m(t) = m_0 \exp(\sigma p(t))$ .

В каждый момент времени ресурс полностью распределяется среди агентов. Получив свою долю ресурса, каждый агент  $i = 1, 2, \dots, M$  выбирает некоторое действие  $x^i(t)$ . Порождаемые действиями затраты агентов описываются выпуклыми по второму аргументу функциями:  $f^i(t, x^i(t))$ ,  $f^i(t, 0) = 0$ . Компенсации агентам определяются механизмом пропорционального распределения:

$$v^i(x(t), m(t)) = \frac{\gamma^i(x(t))m(t)}{\sum_j \gamma^j(x(t))}, x(t) = (x^k(t))_{k=1}^M. \text{ Функция } \gamma - \text{возрастающая, } \gamma(0) = 0.$$

Выигрыш  $i$ -го агента при плане  $x(t) = (x^i(t))_{i=1}^M$  и ресурсе  $m(t)$ :

$$\Psi^i(x^i, m) = \int_0^T (v^i(x(t), m(t)) - f^i(t, x^i(t))) dt. \quad (3.1.12)$$

Выигрыш Центра равен:

$$F(s, m) = \int_0^T f(s(t), m(t)) dt, s(t) = \sum_{i=1}^M x^i(t), \quad (3.1.13)$$

где  $f$  - возрастающая функция по первому аргументу и убывающая по второму, поскольку Центр заинтересован в максимальных суммарных действиях агентов и минимизации выдаваемого им ресурса.

Период времени  $[0, T]$ , на котором реализуется игра, разбивается на  $N$  интервалов. При равномерном разбиении  $N = 2^k$ . В процессе работы алгоритма  $k$  увеличивается на единицу. Начинаем, например, с  $k=1$ . При неравномерном разбиении начинаем с  $h = \sqrt{T+1} - 1$ , что соответствует  $N=2$ , далее  $h$  заменяется на  $zh$ ,  $0 < z < 1$ , где  $z$  относится к параметрам алгоритма. Дискретный аналог уравнения (3.1.13) имеет вид

$$m_j = m_{j-1}(1 + \sigma_N \delta_j), \sigma_N < 1. \quad (3.1.14)$$

Здесь и далее рассматривается равномерное разбиение, неравномерное анализируется аналогично с небольшими изменениями. В (3.1.14)  $\sigma_N = K/N$ , где  $K$  следует отнести к параметрам алгоритма. При достаточно большом  $N$  влияние параметра  $K$  незначительно. Величины  $m_j = m(t_j)$ ,  $\delta_j \in \{-1, 0, 1\}$ . Отметим, что  $m_j > 0$ . Дискретный выигрыш  $i$ -го агента на  $j$ -м временном интервале имеет вид

$$\psi^i(x, m) = \sum_{j=0}^{N-1} \left( v^i(x_j, m_j) - f_j^i(x_j^i) \right). \quad (3.1.15)$$

$$\text{В (3.1.15) } x_j = x(t_j) = \left( x^i(t_j) \right)_{i=1}^M = \left( x_j^i \right)_{i=1}^M, \quad f_j^i(x_j^i) = f^i(t_j, x^i(t_j)), \quad m_j$$

определяются Центром. Отметим, что выигрыш агента есть сепарабельная функция по переменным  $x_j, x_j^i$ , не связанным с переменными  $x_k, x_k^i$  при  $k \neq j$ , поэтому задача каждого агента распадается на  $N$  независимых задач, где  $j$ -я задача решается относительно переменных с индексом  $j$ . Далее, чтобы не загромождать выкладки, будет опущен «временной» индекс  $j$ , и вместо  $m_j$  использовано обозначение  $l$ . Таким образом, задача агентов есть некооперативная игра с индивидуальными выигрышами:

$$\psi^i(x) = \frac{\gamma(x^i)l}{\sum_j \gamma(x^j)} - f^i(x^i), x = \left( x^i \right)_{i=1}^M. \quad (3.1.16)$$

Для случая  $\gamma(u) = u$  игра с выигрышами (3.1.16) подробно изучена в работах (Cornes, Hartley 2005; Cornes, Sato 2016).

Если добавить естественные условия:  $\gamma(u)$  - непрерывная возрастающая функция,  $\gamma(0) = 0$ , то замена переменных  $\gamma(x^i) = y^i$  приводит (3.1.16) к виду:

$$\psi^i(y, l) = \frac{y^i l}{\sum_j y^j} - f^i \circ \gamma^{-1}(y^i), y = \left( y^i \right)_{i=1}^M,$$

поэтому дальше будет рассматриваться игра с индивидуальными выигрышами:

$$\psi^i(x) = \frac{x^i l}{\sum_j x^j} - f^i(x^i), x = \left( x^i \right)_{i=1}^M \quad (3.1.17)$$

Установлено (Cornes, Sato 2016), что если функции затрат  $f^i(x)$  строго выпуклые и дифференцируемые, то существует единственная равновесная по

Нэш стратегия  $\bar{x} : \forall i = 1, \dots, M \forall x_i \in X_i \psi^i(\bar{x}_1, \dots, x_i, \dots, \bar{x}_M) \leq \psi^i(\bar{x}_1, \dots, \bar{x}_i, \dots, \bar{x}_M)$ . Более того, функции  $\psi^i(\bar{x}_1, \dots, x_i, \dots, \bar{x}_M)$  строго вогнуты. Следуя (Cornes, Sato 2016), определим функции  $z_i(s)$  из соотношений:

$$\begin{aligned} s^2 \left( f^i(z^i) \right)' &= l(s - z^i), s \left( f^i(0) \right)' < l \\ z^i = 0, s \left( f^i(0) \right)' &\geq l \end{aligned} \quad (3.1.18)$$

Соотношения (3.1.18) однозначно определяют функции  $z_i(s)$ .

Рассмотрим уравнение

$$\sum_{i=1}^M z^i(s) = s, s > 0. \quad (3.1.19)$$

Уравнение (3.1.19) имеет единственное положительное решение  $\bar{s} = \sum_{i=1}^M \bar{x}^i$ , где  $\bar{x}^i = z^i(\bar{s})$  - равновесные по Нэшу стратегии. Отметим, что использованная при решении игры переменная  $s$  по смыслу совпадает с  $s(t)$  из (3.1.13).

Дискретный выигрыш Центра выражается следующим образом:

$$\Phi(m) = \sum_{j=0}^{N-1} f(\bar{s}_j(m_j), m_j),$$

где  $\bar{s}_j(m_j)$  - сумма равновесных решений игры для  $j$ -го интервала.

Уравнение (3.1.14) позволяет выразить выигрыш Центра непосредственно через вектор  $\Delta_{N-1} = (\delta_j)_{j=1}^{N-1}$ .

Таким образом, стратегией Центра будем считать вектор  $\Delta_{N-1}$ . Для вычисления оптимальной стратегии Центра используется генетический алгоритм. Приведём результаты численных расчётов.

### 3.1.1. Генетический алгоритм

Рассмотрим дифференциальную игру Центра и  $M$  агентов

$$F(s, m) = \int_0^T (\sqrt{s(t)} - m(t)) dt \rightarrow \max, p;$$

$$\psi^i(x(\cdot), m(\cdot)) = \int_0^T \{v^i(x(t), m(t)) - \mu_i(t)[x^i(t)]^2\} dt \rightarrow \max, x^i(t) \geq 0, i = 1, \dots, M;$$

$$v^i(x(t), m(t)) = \frac{x^i(t)m(t)}{\sum_{j=1}^M x^j(t)}, i = 1, \dots, M; m(t) = m_0 + \int_0^t m(s)dp(s).$$

В функции выигрыша агентов входит коэффициент  $\mu_i(t)$ , который позволяет отразить зависимость затрат  $i$ -го агента от времени.

Для хромосомы  $\Delta_{N-1}$  определим последовательность  $m_j$ , используя уравнение (3.1.14), и рассмотрим  $j$ -ю игру с индивидуальными

выигрышами:  $\frac{x_j^i m_j}{\sum_{k=1}^M x_j^k} - \mu_j^i (x_j^i)^2$ . Для вычисления равновесного решения

определим функции  $z_j^i(s_j)$ :  $z_j^i(s_j) = \frac{m_j s_j}{2\mu_j^i s_j^2 + m_j}$ . Сумма равновесных решений:

$\bar{s}_j = \sum_{i=1}^M \bar{x}_j^i$  находится из уравнения  $\sum_{i=1}^M \frac{1}{2\mu_j^i s_j^2 + m_j} = 1/m_j$ . Определим здоровье

хромосомы  $\Delta_{N-1}$ :  $F(\Delta_{N-1}) = \sum_{j=0}^{N-1} (\sqrt{\bar{s}_j} - m_j^2)$ . На рис. 3.1.3 приведены результаты вычислений.

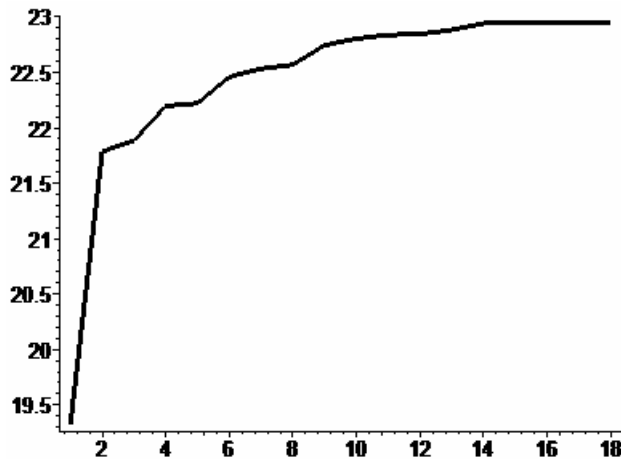


Рис. 3.1.3. Зависимость суммарного здоровья популяции от числа итераций

Ось абсцисс – число итераций генетического алгоритма. Ось ординат – суммарное здоровье популяции. Вычисления производились при следующих

значениях параметров модели:  $m_0 = 0.1, N = 8, \sigma_N = 1/N, M = 2,$

$$\mu_j^1 = \frac{1}{2 + 0.1 \cdot j}, \mu_2^j = 1 - \mu_2^j.$$

Параметры алгоритма: размер популяции  $L = 20$ , число пар при скрещивании  $R = 6$ , вероятность обмена генами для родительской пары  $q = 0.2$ , мутирующая особь выбирается равновероятно из популяции и вероятность мутации гена мутирующей особи  $Q = 0.2$ . В таблице представлена зависимость максимума целевой функции центра от числа точек разбиения. Стабилизация наступает при шестнадцати точках разбиения.

$N$	2	4	8	16	32
$F$	0.7764	0.9546	0.9661	0.9671	0.9671

### 3.1.2. Динамическое программирование и генетический алгоритм

Рассмотрим дифференциальную игру Центра и  $M$  агентов

$$F(s, m) = \int_0^T (\sqrt{s(t)} - m(t)) dt \rightarrow \max, p;$$

$$\psi^i(x(\cdot), m(\cdot)) = \int_0^T \{v^i(x(t), m(t)) - \mu(t)[x^i(t)]^2\} dt \rightarrow \max, x^i(t) \geq 0, i = 1, \dots, M;$$

$$v^i(x(t), m(t)) = \frac{x^i(t)m(t)}{\sum_{j=1}^M x^j(t)}, i = 1, \dots, M;$$

$$m(t) = m_0 + \int_0^t m(s) dp(s).$$

В этом примере коэффициент  $\mu$  в функциях выигрыша агентов не зависит от  $i$ . Это приводит к тому, что при вычислении равновесного решения  $j$ -й игры функции  $z_j(s_j) = \frac{m_j s_j}{2\mu_j s_j^2 + m_j}$  также не зависят от  $i$ , что позволяет найти точное выражение для суммы равновесных решений  $j$ -й игры агентов:



$\bar{s}_j = \sqrt{\frac{(M-1)m_j}{2\mu_j}}$ . Задача Центра заключается в следующем:

$$\sum_{j=0}^{N-1} (\sqrt{\bar{s}_j} - m_j) \rightarrow \max, \Delta_{N-1}, m_j = m_{j-1} (1 + \sigma_N \delta_j), m_0 = a. \quad (3.1.20)$$

Для решения задачи (3.1.20) используем динамическое программирование. Введём обозначения:  $f_j(m_j) = \sqrt{\bar{s}_j} - m_j$ ,  $L_j$  - множество возможных значений  $m_j$ . Непосредственно проверяется следующее

*Утверждение.* Множество возможных значений

$$L_0 = \{a\}, L_j = \left\{ u : u = a(1 - \sigma_N)^{r_j} (1 + \sigma_N)^{q_j}, r_j + q_j \leq j \right\}, j = 1, 2, \dots, N-1, r_j, q_j - \text{целые}$$

неотрицательные числа, включая ноль. Определим функцию Беллмана:

$$B_k(u) = \max_{\Delta_k^{N-1}} \sum_{j=k}^{N-1} f_j(m_j), \Delta_{k+1}^{N-1} = (\delta_i)_{k+1}^{N-1}, m_i = m_{i-1} (1 + \sigma_N \delta_i), i = k, \dots, N-1, m_{k-1} = u \in L_{k-1}.$$

Уравнения Беллмана имеют вид:

$$B_{k-1}(u) = \max_{\delta \in \{-1, 0, 1\}} (f_{k-1}(u(1 + \sigma_N \delta)) + B_k(u(1 + \sigma_N \delta))), u \in L_{k-2}, k = N-1, \dots, 2,$$

$$B_0(a) = f_0(a) + \max_{\delta \in \{-1, 0, 1\}} (B_1(a(1 + \sigma_N \delta))),$$

$$B_{N-1}(u) = \max_{\delta \in \{-1, 0, 1\}} f_{N-1}(u(1 + \sigma_N \delta)), u \in L_{N-2}, \delta_k^*(u) = \arg \max_{\delta \in \{-1, 0, 1\}} B_k(u(1 + \sigma_N \delta)),$$

$$\delta_1^*(a) = \arg \max_{\delta \in \{-1, 0, 1\}} B_1(a(1 + \sigma_N \delta)).$$

Результаты вычислений представлены на рис. 3.1.4 и 3.1.5. Решалась задача с одними и теми же параметрами:  $\mu_j^1 = \frac{1}{2 + 0.1 \cdot j}$ ,  $\sigma_N = 1/N$ , число интервалов разбиения  $N = 16$ , число агентов  $M = 2$ ,  $a = 0.3$ . На рис. 3.1.4 приведены результаты расчёта, выполненные генетическим алгоритмом, на рис. 3.1.5 - методом динамического программирования.

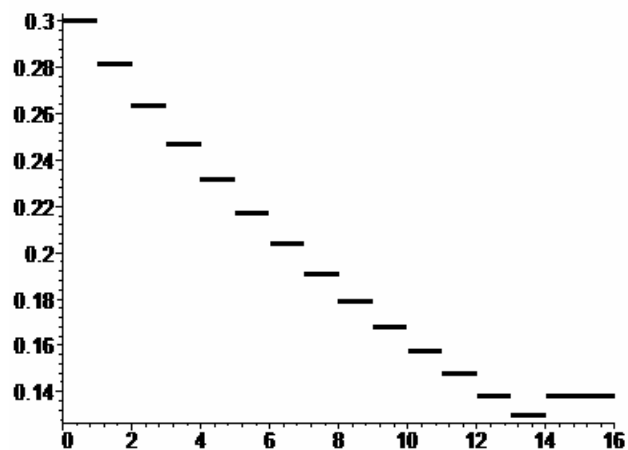


Рис. 3.1.4. Расчёты, выполненные генетическим алгоритмом

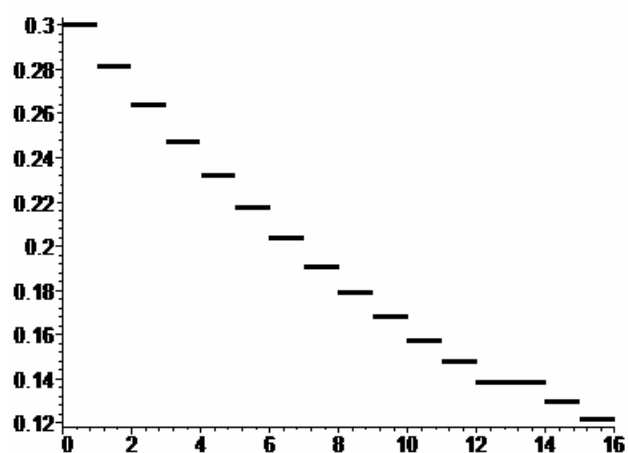


Рис. 3.1.5. Расчёты, выполненные методом динамического программирования

Оси абсцисс – номера интервалов разбиения. Оси ординат – значения распределяемого ресурса. Решение, приведённое на рис. 3.1.5, является точным. Из рисунков следует, что результаты расчётов отличаются незначительно. Время вычисления в обоих методах примерно одинаково. Значение функции выигрыша при использовании генетического алгоритма совпадает с оптимальным значением до четырёх знаков после запятой.

Метод динамического программирования, рассмотренный в этом разделе, без существенных изменений можно применить для решения примера из раздела 3.1.1.

### 3.1.3. Аналитическое решение.

Рассматривается задача:

$$F(s, m) = \int_0^T (\sqrt{s(t)} - m(t)) dt \rightarrow \max, p(t) \geq 0;$$

$$\psi^i(x(\cdot), m(\cdot)) = \int_0^T \{v^i(x(t), m(t)) - [x^i(t)]^2\} dt \rightarrow \max, x^i(t) \geq 0, i = 1, \dots, M;$$

$$v^i(x(t), m(t)) = \frac{x^i(t)m(t)}{\sum_{j=1}^M x^j(t)}, i = 1, \dots, M;$$

$$m(t) = m_0 + \int_0^t m(s) dp(s).$$

При вычислении равновесного решения  $j$ -й игры функции  $z_j(s_j) = \frac{m_j s_j}{2s_j^2 + m_j}$ , сумма равновесных решений  $\bar{s}_j = \sqrt{\frac{(M-1)m_j}{2}}$ . В результате задача Центра:

$$\sum_{j=0}^{N-1} \left( \left( \frac{(M-1)m_j}{2} \right)^{1/4} - m_j \right) \rightarrow \max, \Delta_{N-1}, m_j = m_{j-1} (1 + \sigma_N \delta_j), m_0 = a.$$

Рассмотрим функцию  $f(x) = \left( \frac{(M-1)x}{2} \right)^{1/4} - x$ , которая является вогнутой. Элементарный анализ позволяет найти значение  $x^* = \frac{2}{(M-1)\sqrt[3]{256}}$ , при котором достигается максимум  $f(x)$ . Отсюда следует, что решение задачи Центра выглядит следующим образом:  $a^* = x^*, \delta_j^* = 0, j = 1, \dots, N-1$ .

### 3.1.4. Терминальный критерий. Генетический алгоритм

Рассматривается следующая задача:

$$F(S_T, L_T) = \sqrt{S_T} - L_T \rightarrow \max, p$$

$$\psi^i(x(\cdot), m(\cdot)) = \int_0^T \{v^i(x(t), m(t)) - [x^i(t)]^2\} dt \rightarrow \max, x^i(t) \geq 0, i = 1, \dots, M;$$

$$v^i(x(t), m(t)) = \frac{x^i(t)m(t)}{\sum_{j=1}^M x^j(t)}, i = 1, \dots, M;$$

$$m(t) = m_0 + \int_0^t m(s)dp(s), S_T = \int_0^T s(t)dt, L_T = \int_0^T m(t)dt .$$

Дискретная задача центра в этом примере отличается от задачи центра в предыдущем примере видом функции выигрыша:

$$\sqrt{\sum_{j=0}^{N-1} \sqrt{\frac{(M-1)m_j}{2}}} - \sum_{j=0}^{N-1} m_j \rightarrow \max, \Delta_{N-1}, m_j = m_{j-1}(1 + \sigma_N \delta_j), m_0 = a .$$

Применение динамического программирования приводит к вычислительно трудоёмкой задаче. Вычислительные затраты растут как показательная функция от числа интервалов разбиения. Результаты расчётов с применением генетического алгоритма при равномерном разбиении приведены на рис. 3.1.6.

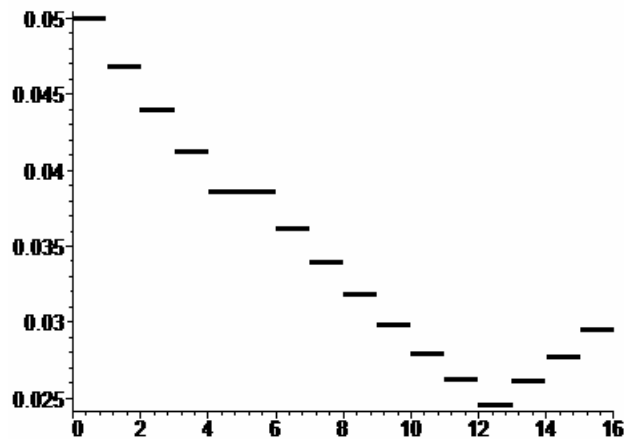


Рис. 3.1.6. Генетический алгоритм. Терминальный критерий. Равномерное разбиение

Оптимальное значение целевой функции равно 0.8903. Значения параметров:  $a = 0.05, M = 2, N = 16, \sigma_N = 1/N$ .

Итак, пропорциональный механизм представляется наиболее естественным способом распределения ограниченных ресурсов между активными агентами, что обуславливает внимание к этой постановке задачи. В статическом случае механизмы пропорционального распределения

подробно исследованы в теории управления организационными системами (Новиков 2007).

Однако, с теоретической и практической точек зрения более интересна динамическая постановка задачи распределения ресурсов, один из вариантов которой в различных модификациях рассмотрен в настоящем параграфе (альтернативный подход предложен в главе 2). В этом варианте ресурс трактуется как фазовая переменная. Важное предположение состоит в том, что ведущий игрок (центр) не может резко менять значения управляющей переменной, что отвечает реальной практике принятия решений в организационно-экономических системах, а с математической точки зрения позволяет использовать условие Липшица.

При этом для решения возникающей дифференциальной игры используются генетические алгоритмы. Приведённые результаты модельных расчетов позволяют надеяться, что вычислительные методы, использующие генетические алгоритмы, применимы для решения динамической задачи распределения ресурсов, особенно в тех случаях, когда метод динамического программирования неэффективен из-за вычислительной сложности (Белявский и др. 2018).

Теперь исследуем алгоритмы имитации отжига и бинарного деления, следуя (Belyavsky and Danilova 2019). Рассмотрим динамическую игру Штакельберга с одним Центром и несколькими агентами. Модель содержит следующие элементы: состояние игры  $(x_0, x(t)) \in R^2$ , стратегии  $(u(t), v(t)) \in R^{r+1}$ ,

выигрыш Центра  $\int_0^T g_0(x_0, u, v) dt$ , выигрыши агентов  $\int_0^T g_i(x, u, v) dt$ ;  $u$  –

управление Центра;  $v_i$  – реакция агента  $i$ . Определим задачу Центра как вычисление

$$\max_u \int_0^T g_0(x_0, u, v) dt, \text{ with constraint } dx_0(t) = f_0(x_0, u) dt, x_0(0) = x_0^0. \quad (3.1.21)$$

Можно также добавить условие гомеостаза  $x(t) \in X$ , например, в виде  $(x_0^* - x(t))^2 \leq a$ . Это условие удобно выразить с помощью штрафа:

$k \int_0^T (x_0^* - x(t))^2 dt$ , включив его в функционал выигрыша Центра, который тогда

примет вид:  $\int_0^T [g_0(x_0, u, v) - k(x_0^* - x(t))^2] dt$ . Задачи агентов ставятся в виде

$$\max_{v_i} \int_0^T g_i(x, u, v) dt, \text{ with constraints } dx_i(t) = f_i(x_i, v_i) dt, x_i(0) = x_i^0. \quad (3.1.22)$$

Предполагается, что игру (3.1.21), (3.1.22) можно преобразовать в статическую игру Штакельберга в бесконечномерных линейных пространствах:

$$J_0(u, v) \rightarrow \max_u; J_i(u, v) \rightarrow \max_{v_i}, i = 1, 2, \dots, r. \quad (3.1.23)$$

Иначе говоря, для любых допустимых стратегий  $(u, v)$  имеется алгоритм вычисления состояния системы  $(x_0, x_i)$  и всех выигрышей.

Допустим, что функции  $u$  и  $v_i$  принадлежат пространству Банаха  $B[0,1]$  ограниченных функций с равномерной нормой:  $\|f\| = \sup_{t \in [0,1]} f(t)$ . В игре (3.1.23)

Центр выбирает стратегию  $u$  и сообщает её агентам. В свою очередь, агенты выбирают свои стратегии как оптимальный ответ на стратегию Центра из множества равновесий Нэша в игре агентов в нормальной форме:  $v(u) \in N(u)$ .

Тогда задача Центра имеет вид

$$\max_u \min_{v \in N(u)} J_0(u, v), \quad (3.1.24)$$

при отказе агентов от сотрудничества, и

$$\max_u \max_{v \in N(u)} J_0(u, v), \quad (3.1.25)$$

при сотрудничестве. Обозначим через  $\Phi(u)$  решение внутренней задачи в (3.1.24) от (3.1.25). Тогда задача Центра принимает вид

$$\max_u \bar{J}_0(u), \quad (3.1.26)$$

где  $\bar{J}_0(u) = J_0(u, \Phi(u))$ .

Рассмотрим конечномерную аппроксимацию задачи (3.1.26). Достаточное условие возможности такой аппроксимации состоит в непрерывности функционала  $J_0(u)$  на множестве допустимых решений. Эта непрерывность обеспечивается условием Липшица  $|\bar{J}_0(u) - \bar{J}_0(w)| \leq L \|u - w\|$ , которое следует из двух неравенств:

$$|J_0(u_2, v) - J_0(u_1, w)| \leq L_u \|u_2 - u_1\| + L_v \|v - w\|_r, \|\Phi(u_2) - \Phi(u_1)\|_r \leq L_\Phi \|u_2 - u_1\|, \quad (3.1.27)$$

Второе неравенство в (3.1.27) труднее всего проверить.

Рассмотрим в качестве первого класса допустимых управлений подмножество множества ограниченных функций вида

$$L^1([0,1]) = \left\{ u \in B[0,1] : \exists \alpha, \sup_{|t-s|} \frac{|u(t) - u(s)|}{|t-s|} \leq \alpha, 0 \leq t \leq 1, 0 \leq s \leq 1, t \neq s \right\}. \quad (3.1.28)$$

Предполагается, что Центр в некотором смысле ограничен в своих действиях и поэтому выбирает управления из этого класса функций. Иными словами, Центр не может совершать "резких движений". Следующий результат даёь основу для предлагаемого метода конечномерной аппроксимации.

*Теорема 3.1.2.* Если  $u \in L^1([0,1])$ , то существует последовательность

$$u^n = u_0 + \alpha \sum_{i=1}^n \delta_i^n(u) I_{\{t > \tau_i\}}, \delta_i^n(u) \in \{-1, 0, 1\}, \tau_i = i/n, i = 0, 1, \dots, n, \quad (3.1.29)$$

которая сходится к  $u$  по норме пространства  $B[0,1]$ . Это означает, что подмножество

$$\bar{L}[0,1] = \left\{ u \in L[0,1] : \exists (u_0, \alpha, n, \delta), u = u_0 + \alpha \sum_{i=1}^n \delta_i I_{\{t > \tau_i\}}, \delta_i \in \{-1, 0, 1\}, \tau_i = i/n, i = 0, 1, \dots, n \right\}$$

плотно в  $L_1[0,1]$ .

Допустим, что исходная задача (3.1.26) с дополнительным ограничением  $u \in L_1[0,1]$  имеет решение. Тогда конечномерная аппроксимация задачи (3.1.26) с указанным ограничением есть задача оптимизации

$$\max_{u \in \bar{L}[0,1]} \bar{J}_0(u) = \max_{u_0 \in R, \alpha, \Delta} \bar{J}_0(u_0, \alpha, \Delta), \Delta = (\delta_i)_{i=1}^n. \quad (3.1.30)$$

Заметим, что задача (3.1.30) имеет решение, если для любого заданного  $\bar{\Delta}$  задача  $\max_{u_0 \in R, \alpha} \bar{J}_0(u_0, \alpha, \bar{\Delta})$  имеет решение. Пусть это условие выполнено. Смысл конечномерной аппроксимации определяется следующим результатом.

*Теорема 3.1.3.* Пусть  $p^* = \max_{u \in L[0,1]} \bar{J}_0(u) = \bar{J}_0(u^*)$ ,  $q^* = \max_{u \in \bar{L}[0,1]} \bar{J}_0(u) = \bar{J}_0(\bar{u}^*)$ . Тогда для любого  $\varepsilon > 0$  справедливо  $p^* - q^* \leq \varepsilon$ .

Действительно, из непрерывности функционала  $J_0(u)$  и плотности множества  $\bar{L}([0,1])$  в множестве  $L^1([0,1])$  вытекают следующие неравенства:  $p^* - q^* = J_0(u^*) - J_0(\bar{u}^*) \leq J_0(u^*) - J_0(\bar{u}_1) \leq \varepsilon$ . Чтобы удовлетворить последнее неравенство, надо выбрать  $\bar{u}_1 \in \bar{L}([0,1])$  достаточно близким к  $u^*$ .

Второй класс допустимых управлений - это множество ограниченных функций, имеющих на отрезке  $[0,1]$  конечное число точек разрыва.

Обозначим это множество через  $L^2([0,1])$ . Определим множество

$$\bar{L}^2([0,1]) = \left\{ u \in B[0,1] : \exists (n, (c_i)_{i=1}^n, (\tau_i)_{i=1}^n), 0 = \tau_0 < \tau_1 < \dots < \tau_n = 1, f(t) = c_1 I_0(t) + \sum_{i=1}^n c_i I_{(\tau_{i-1}, \tau_i]}(t) \right\}.$$

Очевидно, что множество  $\bar{L}^2$  плотно в множестве  $L^2$ . При непрерывности функционала  $\bar{J}_0(u)$  на множестве  $L^2([0,1])$  задача

$$\max_{u \in \bar{L}^2[0,1]} \bar{J}_0(u) \quad (3.1.31)$$



есть конечномерная аппроксимация задачи (3.1.26) с дополнительным ограничением  $u \in L^2[0,1]$ , если обе задачи имеют решения.

Проанализируем переменные задачи (3.1.30). Первые две переменные  $\alpha$  и  $u_0$  - действительные числа, а третья переменная  $\Delta$  принимает значения на конечном множестве последовательностей  $S = \{(\delta_i)_i^n : \delta_i \in \{-1,0,1\}\}$ . Если функционал  $\bar{J}_0(u)$  удовлетворяет глобальному условию Липшица с константой  $L$ , и  $K = \sup \alpha$  в определении (3.1.28), то для заданного  $\varepsilon$  число элементов в последовательности равно  $n = \left\lfloor \frac{KL}{\varepsilon} \right\rfloor + 1$  (Белявский и др. 2016). Таким образом, рассматриваемая задача связана с вычислением максимума функции  $F(\Delta) = \max_{u_0, \alpha} \bar{J}_0(u_0, \alpha, \Delta)$  на конечном множестве  $S$ . Задачи этого типа могут быть эффективно решены с помощью алгоритмов эволюционного моделирования. Генетический алгоритм был рассмотрен выше (Белявский и др. 2016, 2018), поэтому теперь обратимся к алгоритму имитации отжига (Jones 2008).

Этот алгоритм стартует с начального  $\Delta$  и начальной температуры  $T = T_s$ . Итерации имеют следующий вид.

1. Новое значение  $\bar{\Delta}$  вычисляется в окрестности текущего  $\Delta$ .
2. Если  $F(\bar{\Delta}) \geq F(\Delta)$ , то  $\Delta := \bar{\Delta}$ , иначе  $\Delta := \bar{\Delta}$  с вероятностью

$$p = \exp\left(\frac{F(\bar{\Delta}) - F(\Delta)}{T}\right).$$

3. Установить температуру  $T := qT$ , где  $0 < q < 1$ .
4. Повторять итерации, пока  $(T \geq T_f) \wedge (|\Delta F| > \bar{\varepsilon})$ .

Теперь предположим, что функционал  $\bar{J}_0(u)$  аддитивен, т.е. для произвольного разбиения отрезка  $[0,1]$ :  $0 = \tau_0 < \tau_1 < \dots < \tau_n = 1$  имеет место неравенство:

$$\bar{J}_0(u) = \bar{J}_0 \left( \sum_{i=1}^n u(t) I_{[\tau_{i-1}, \tau_i]}(t) \right) = \sum_{i=1}^n \bar{J}_0 \left( u(t) I_{[\tau_{i-1}, \tau_i]}(t) \right). \quad (3.1.32)$$

Алгоритм бинарного деления имеет следующий вид.

1. инициализация. Пусть задано начальное разбиение отрезка:  $[0, 1/2) \cup [1/2, 1]$ .

Рассмотрим аппроксимацию  $u^1(t) = c_{1,0} I_{[0, 1/2]}(t) + c_{1,1} I_{(1/2, 1]}(t)$ . Из всех

аппроксимаций этого вида выберем наилучшую, используя свойство

аддитивности функционала  $\bar{J}_0(u^1) = \bar{J}_0(c_{1,0} I_{[0, 1/2]}) + \bar{J}_0(c_{1,1} I_{(1/2, 1]})$ . Для этого решим

две независимые задачи:  $\max_c \bar{J}_0(c I_{[0, 1/2]})$ ,  $\max_c \bar{J}_0(c I_{(1/2, 1]})$ .

2. Итерации. Итерация с номером  $n$  состоит в следующем. Пусть текущее разбиение  $0 = \tau_0^{(n)} < \tau_1^{(n)} < \dots < \tau_n^{(n)} = 1$ , а соответствующая текущая аппроксимация

$u^{(n)}(t) = c_{n,0} I_0(t) + \sum_{i=1}^n c_{n,i} I_{[\tau_{i-1}^{(n)}, \tau_i^{(n)}]}(t)$ . Определим последовательность

$b_{n,j} = \frac{1}{\tau_j^{(n)} - \tau_{j-1}^{(n)}} \bar{J}_0 \left( c_{n,j} I_{[\tau_{j-1}^{(n)}, \tau_j^{(n)}]} \right)$  и соответствующие вероятности выбора интервала:

$p_{n,j} = \frac{b_{n,j}}{\sum_{k=1}^n b_{n,k}}$ ,  $j = 1, 2, \dots, n$ . Интервал выбирается случайно согласно

распределению вероятностей  $p_{n,j}$ . Выбранный интервал с номером  $j$ , а

именно  $[\tau_{j-1}^{(n)}, \tau_j^{(n)}]$ , разбивается на два интервала равной длины

$\left[ \tau_{j-1}^{(n)}, \frac{\tau_{j-1}^{(n)} + \tau_j^{(n)}}{2} \right]$ ,  $\left[ \frac{\tau_{j-1}^{(n)} + \tau_j^{(n)}}{2}, \tau_j^{(n)} \right]$ , и вычисляется новая аппроксимация

$u^{n+1}(t) = c_{n+1,0} I_0 + \sum_{i=1}^{n+1} c_{n+1,i} I_{[\tau_{i-1}^{(n+1)}, \tau_i^{(n+1)}]}(t)$ , где  $c_{n+1,i} = c_{n,i}$ , если  $i = 1, \dots, j-1$ ;

$c_{n+1,j} = \arg \max_c \bar{J}_0 \left( c I_{\left[ \tau_{j-1}^{(n)}, \frac{\tau_{j-1}^{(n)} + \tau_j^{(n)}}{2} \right]} \right)$ ,  $c_{n+1,j+1} = \arg \max_c \bar{J}_0 \left( c I_{\left[ \frac{\tau_{j-1}^{(n)} + \tau_j^{(n)}}{2}, \tau_j^{(n)} \right]} \right)$ ;

$c_{n+1,i} = c_{n,i-1}$ , если  $i = j+2, \dots, n+1$ . Новое разбиение вычисляется как  $\tau_i^{(n+1)} = \tau_{i-1}^{(n)}$ , если  $i = 0, \dots, j-1$ ,  $\tau_j^{(n+1)} = \frac{\tau_{j-1}^{(n)} + \tau_j^{(n)}}{2}$ , и  $\tau_i^{(n+1)} = \tau_{i-1}^{(n)}$ , если  $i = j+1, \dots, n+1$ . Итерации повторяются, пока изменения не станут достаточно малыми.

Этот алгоритм монотонный. Значит, последовательность  $\bar{J}_0(u_n)$  сходится, если функционал выигрыша Центра ограничен сверху. Дополнительно достигается минимизация числа точек разбиения интервала.

В качестве первого примера рассмотрим динамическую игру Штакельберга, уже сведённую к виду (3.1.26). Эта игра описывает распределение ресурса между производителями (см., например, Новиков 2007). Количество ресурса, распределяемое Центром в момент времени  $t$ , обозначается через  $u(t)$ . Каждый агент получает свою часть ресурса  $u_i(t)$ ; очевидно, что  $u(t) = \sum_{i=1}^r u_i(t)$ . Получив ресурс, каждый агент производит продукцию  $v_i(t)$  так, чтобы максимизировать текущий выигрыш. Центр стремится максимизировать суммарный выпуск продукции посредством механизма пропорционального распределения ресурса  $u_i(t) = \gamma_i(t)v_i(t)$ . Тогда  $\sum_{i=1}^r \gamma_i(t)v_i(t) = u(t)$ . Текущий выигрыш  $i$ -го агента вычисляется как разность между его доходом и производственными затратами, или  $P_i(t) = \gamma_i(t)v_i(t) - \varphi_i(v_i(t))$ . Функция затрат  $\varphi_i(x)$  - это выпуклая неубывающая функция на  $R^+$ , причём  $\varphi_i(0) = 0$ . Итак, возникает вспомогательная игра Штакельберга в виде:

$\max_{\gamma \geq 0} \sum_{i=1}^r v_i \text{ при ограничении } \sum_{i=1}^r \gamma_i v_i = u$ $\max_{v_i \geq 0} [\gamma_i v_i - \varphi_i(v_i)].$	(3.1.33)
---	----------

Для простоты будем опускать индекс  $t$ . Важно отметить, что в данной формулировке игра агентов распадается на независимые задачи

оптимизации. Если предположить, что  $\gamma_i = \gamma$ , то из равенства  $\sum_{i=1}^r \gamma_i v_i = u$

следует равенство  $\gamma = \frac{u}{\sum_{i=1}^r v_i}$ , и решение игры агентов состоит в вычислении

равновесия Нэша в игре с индивидуальными задачами оптимизации агентов:

$$\max \left[ \frac{uv_i}{\sum_{j=1}^r v_j} - \varphi_i(v_i) \right] \text{ (см. Christodoulou et al. 2015).}$$

Рассмотрим игру с функциями затрат  $\varphi_i(x) = \mu_i x^2$ . В этом случае решение (3.1.31) принимает вид

$$v_i = \frac{\gamma_i}{2\mu_i}, \gamma_i = \sqrt{2u / \sum_{j=1}^r \frac{1}{\mu_j}}. \quad (3.1.34)$$

Заметим, что все  $\gamma_i$  не зависят от  $i$ .

Текущий выигрыш Центра определяется как  $R\left(\sum_i v_i(t), u(t)\right)$ . Функция  $R(x, y)$  возрастает по первому аргументу и убывает по второму. Общий выигрыш получается интегрированием текущих:  $\int_0^T R\left(\sum_{i=1}^r v_i, u\right) dt$ . Из (3.1.34)

получаем общий выигрыш в виде  $\bar{J}_0(u) = \int_0^T R\left(\frac{1}{2} \sqrt{2u / \sum_{j=1}^r \frac{1}{\mu_j}} \sum \frac{1}{\mu_i}, u\right) dt$ . Обозначим

$A(t) = \sqrt{\frac{1}{2} \sum_{i=1}^r \frac{1}{\mu_i(t)}}$ . Тогда  $\bar{J}_0(u) = \int_0^T R(A\sqrt{u}, u) dt$ . Положим  $F(x, y) = x - y$ , тогда

$\bar{J}_0(u) = \int_0^T [A\sqrt{u} - u] dt$ . Очевидно, что функция  $u^*(t) = \frac{A^2(t)}{4}$  максимизирует  $\bar{J}_0(u)$ .

Эта функция удовлетворяет обоим дополнительным ограничениям, если  $\mu_i(t) > 0$ .

Применим алгоритм имитации отжига к игре с двумя агентами при  $\mu_1(t) = t^2 + 1$ ,  $\mu_2(t) = 2t^2 + 1$ . Для этого алгоритма разбиение интервала задано.

Вычислим  $F(\Delta) = \max_{u_0, \alpha} \left[ \sum_{i=1}^n \left( a_i \sqrt{u_0 + \alpha g_i(\Delta)} - (u_0 + \alpha g_i(\Delta)) \Delta t \right) \right]$ . Из вогнутости по  $u_0$

и  $\alpha$  следует, что оптимальные значения - это решения алгебраической

системы двух уравнений:  $\sum_{i=1}^n \frac{a_i}{\sqrt{u_0 + \alpha g_i(\Delta)}} = 2, \sum_{i=1}^n \left( \frac{a_i}{\sqrt{u_0 + \alpha g_i(\Delta)}} - 2g_i(\Delta) \Delta t \right) = 0$ . В

этих формулах  $a_i = \int_{\tau_{i-1}}^{\tau_i} A(t) dt, g_i(\Delta) = \frac{1}{n} \sum_{j=1}^{i-1} \delta_j, \Delta t = \tau_i - \tau_{i-1}$ . Для вычисления  $\tilde{\Delta}$  в

окрестности  $\Delta$  каждое  $\delta_i$  заменяется на  $\tilde{\delta}_i \in \{-1, 0, 1\} \setminus \delta_i$  с вероятностью  $\frac{q}{n}$  и

равными вероятностями на множестве  $\{-1, 0, 1\} \setminus \delta_i$ . Параметр  $q \in \{1, 2, \dots, n-1\}$

определяет среднее число меняющихся элементов  $\Delta$ . Результаты расчётов

представлены на рис. 3.1.7.

Теперь рассмотрим применение алгоритма бинарного деления к той же задаче. Для этого в каждый момент времени на интервале  $[\tau, s]$  нужно найти

$\min_c \left[ \sqrt{c} \int_{\tau}^s A(t) dt - c(s-t) \right]$ . Оптимальное значение равно  $c^* = \left( \frac{1}{2(s-\tau)} \int_{\tau}^s A(t) dt \right)^2$ .

Результаты расчётов представлены на рис. 3.1.8.

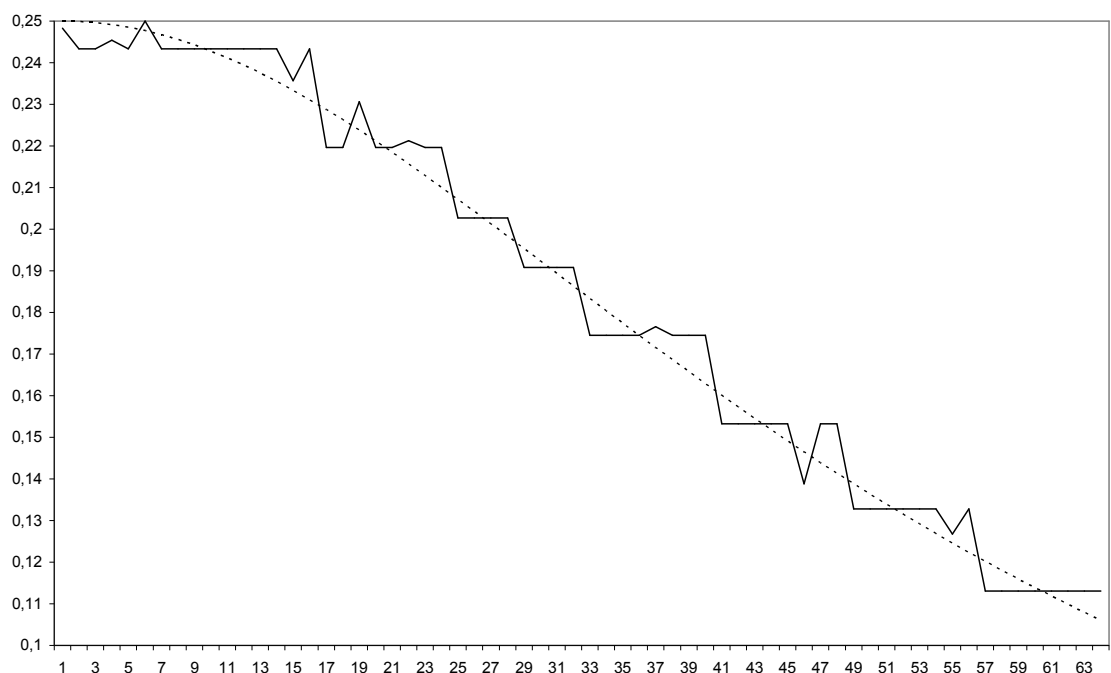


Рис. 3.1.7. Результаты работы алгоритма имитации отжига для  $T_{\max}=100$ ,  $T_{\min}=10$ ,  $q=5$ ,  $n=64$ . Пунктирной линией показано точное решение

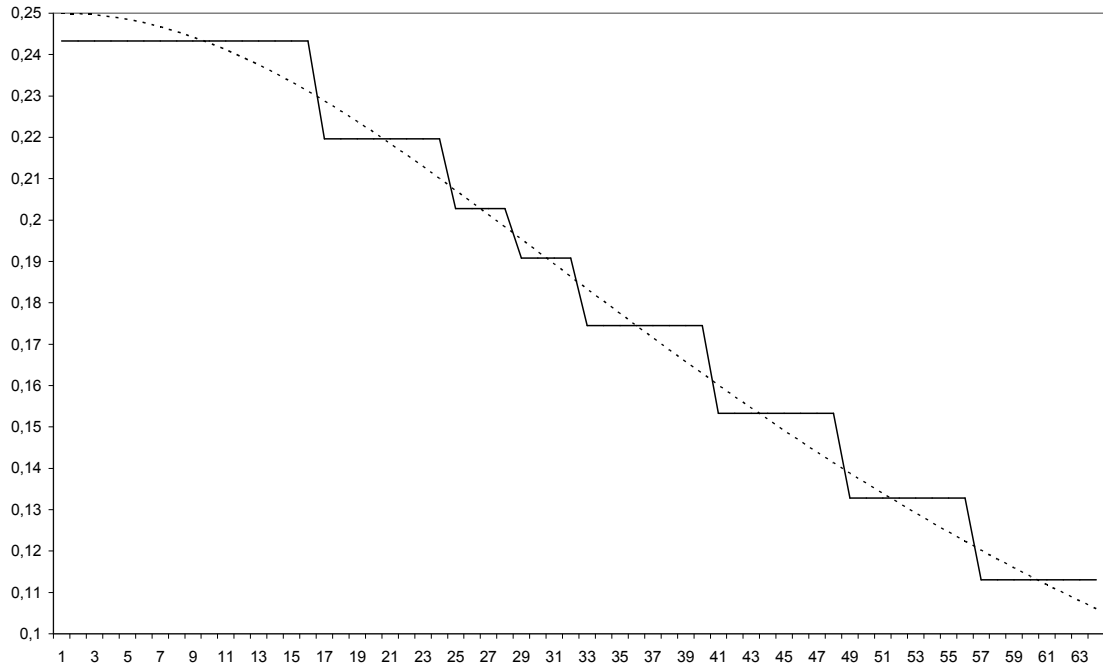


Рис. 3.1.8. Результаты работы алгоритма бинарного деления. Пунктирной линией показано точное решение

Теперь применим алгоритм имитации отжига к следующей задаче (Belyavsky et al. 2018). Центр использует распределение ресурсов для стимулирования агента и стремится найти  $\max_u [\psi(v) - \varphi(u, v)]$ . Агент максимизирует свою прибыль:  $\max_v [\varphi(u, v) - f(v)]$ . Пусть в этой игре существует равновесие Нэша  $(u^*, v^*)$ . Особенность игры заключается в том, что Центр не знает функцию затрат агента и соответственно не может вычислить это равновесие. Поэтому Центр использует последовательность управлений  $u(t)$  для определения  $u^*$ . Последовательность  $v(t), t = 0, 1, \dots$  представляет наилучшие ответы агента на  $u(t)$ . В каждый момент времени  $t$  Центру известен интервал  $v(0), \dots, v(t-1)$  последовательности наилучших ответов агента. На основе этой информации Центр выбирает управление

$u(t) = u(t-1)(1 + \alpha\delta(t))$ , согласно которому агент получает ресурс в объёме  $\varphi(v(t-1), u(t))$ . Как и ранее,  $\delta(t) \in \{-1, 0, 1\}$ ,  $0 < \alpha < 1$ . Таким образом, на каждой итерации игры Центр может оставить своё управление неизменным или увеличить/уменьшить его на небольшую фиксированную величину в соответствии с описанной выше "липшицевой" концепцией. Начальное значение  $u(0)$  и последовательность  $\delta(t)$  полностью определяют последовательность  $u(t)$ . Пусть начальное значение  $u(0) = x$  задано.

Рассмотрим задачу агента. Он предполагает, что  $\delta(t)$  — это марковская последовательность с множеством состояний  $\{-1, 0, 1\}$  и матрицей переходных вероятностей  $Q$  размерности  $(3 \times 3)$ . Начальное распределение вероятностей  $y$  на множестве  $\{-1, 0, 1\}$  задано. Задача агента состоит в вычислении

$\max_{x,y} E_{x,y} \sum_{t=1}^{\infty} \beta^t [\varphi(v(t-1), u(t)) - f(v(t))].$	(3.1.35)
---	----------

Если функция  $-f(z) + \beta\varphi(z, w)$  строго вогнута для любого значения аргумента  $w$ , то оптимальное управление агента есть

$v^*(x, y) = \arg \max_z [-f(z) + \beta\varphi(z, y(1 + \alpha(q_{x,3} - q_{x,1})))].$	(3.1.36)
--	----------

Тогда текущая реакция агента вычисляется как

$v(t) = \arg \max_z [-f(z) + \beta\varphi(z, u(t)(1 + \alpha(q_{\delta(t),3}^t - q_{\delta(t),1}^t)))].$	(3.1.37)
--	----------

Матрица  $Q^t$  находится как максимально правдоподобная оценка переходной матрицы для интервала  $\delta(1), \dots, \delta(t)$ .

Теперь рассмотрим задачу Центра. Он трактует  $v(t)$  как наилучший ответ на  $\delta(t)$  и выбирает следующее значение  $\delta(t+1)$  в предположении, что следующий наилучший ответ агента  $v(t+1)$  — случайная величина. Таким образом, Центр решает задачу

$\max_{\delta(t+1)} E \left[ \psi(v(\delta(t+1))) - \varphi(v(t), u_t(1 + \delta(t+1))) \right].$	(3.1.38)
---	----------

Эта задача решается с помощью алгоритма reinforcing learning. Согласно этому алгоритму, на каждой итерации Центр вычисляет новое значение  $Q$ -функции:

$Q_{t+1}(\delta(t)) = Q_t(\delta(t)) + h_t (R(\delta(t)) - Q_t(\delta(t)))$	(3.1.39)
---	----------

В (3.1.39)  $R(\delta(t)) = \psi(v(t)) - \varphi(v(t-1), u_t)$ , начальное значение  $Q_0(\cdot) \equiv 0$ , а последовательность  $h$  удовлетворяет условию:  $\sum_{t=1}^{\infty} h_t = \infty, \sum_{t=1}^{\infty} h_t^2 < \infty$ . Тогда распределение вероятностей на множестве  $\{-1, 0, 1\}$  вычисляется как

$p_{t+1}(j) = \exp(Q_{t+1}(j)/T_{t+1}) / \sum_{k=-1}^1 \exp(Q_{t+1}(k)/T_{t+1}), j = -1, 0, 1;$	(3.1.40)
---	----------

и соответственно выбирается  $\delta(t+1)$ . В (3.1.40)  $T$  - это температура, управляющая степенью случайности при выборе следующего значения  $\delta$ . Сходимость данного алгоритма изучена в (Sutton, Barto 1998).

Рассмотрим второй иллюстративный пример, где  $\psi(x) = \sqrt{x}$ ,  $\varphi(x, y) = xy$  и  $f_i(x) = \mu x^2$ . Равенство (3.1.37) принимает вид

$$v(t) = \arg \max_z \left[ -\mu z^2 + \beta z u(t) \left( 1 + \alpha (q_{\delta(t),3}^t - q_{\delta(t),1}^t) \right) \right].$$

Простые вычисления дают

$$v(t) = \frac{\beta z u(t) \left( 1 + \alpha (q_{\delta(t),3}^t - q_{\delta(t),1}^t) \right)}{2\mu}.$$

Равновесное решение игры (3.1.35), (3.1.38)

есть  $v^* = \frac{\beta u^*}{2\mu}, u^* = \left( \frac{\mu}{8\beta} \right)^{1/3}$ . Равновесное решение исходной игры:

$$v^* = \frac{u^*}{2\mu}, u^* = \left( \frac{\mu}{8} \right)^{1/3}.$$

Для  $\mu = 1, \beta = 0.9, \alpha = 0.05$  получены следующие численные

результаты (табл. 3.1.2).



## Численные результаты для примера 2

Итерация	Управление Центра	Выигрыш Центра
1	0,4	0
2	0,412	0,352264
3	0,39964	0,354196
4	0,411629	0,352203
5	0,423978	0,35414
6	0,436697	0,355904
7	0,436697	0,357482
8	0,449798	0,357482
9	0,463292	0,358856
10	0,449394	0,36001
11	0,462875	0,358817
12	0,476762	0,359978
13	0,491064	0,360902
14	0,491064	0,361569
15	0,491064	0,361569
16	0,476333	0,361569
17	0,490622	0,360877
18	0,505341	0,361553
19	0,520501	0,361952
20	0,504886	0,362054
21	0,520033	0,361944
22	0,504432	0,362055
23	0,504432	0,361936
24	0,504432	0,361936
25	0,504432	0,361936
26	0,504432	0,361936

Заметим, что равновесная стратегия Центра для указанных входных данных равно 0,519, а соответствующий выигрыш 0,362.

Расмотренный пример с известным точным решением демонстрирует применимость обоих численных алгоритмов. Поскольку эти алгоритмы требуют только вычисления функционалов выигрыша, то они эффективны в ситуациях, когда другие методы не приводят к реализуемым численным схемам. Следует отметить, что методы оптимизации, не требующие вычисления градиента, активно обсуждаются в современной литературе.

При сравнимой точности приближённых вычислений по обоим алгоритмам для второго примера, эти алгоритмы всё же существенно отличаются. Их сравнение по некоторым важным критериям представлено в таблице 3.1.3.

Таблица 3.1.3

Сравнение алгоритмов имитации отжига и бинарного деления

	Алгоритм имитации отжига	Алгоритм бинарного деления
Дополнительные условия на функционал выигрыша Центра	Отсутствуют	Аддитивность
Дополнительные условия на управление Центра	Условие Липшица	Практически отсутствуют
Интервал предварительного разбиения	Требуется	Не требуется

Заметим, что в статической игре с неполной информацией алгоритм бинарного деления неприменим.

Сравнение алгоритма имитации отжига и генетического алгоритма приводит к следующим выводам. Оба относятся к алгоритмам случайного поиска для решения задач оптимизации. Существенное отличие состоит в том, что имитация отжига - это алгоритм поиска в глубину, когда на каждой итерации изучается только одно потенциальное решение. Генетический алгоритм осуществляет поиск в ширину, который проверяет на каждом шаге

несколько кандидатов на роль решения. Генетический алгоритм обеспечивает более быструю сходимость, в то время как имитация отжига использует более простые для вычислений итерации.

Таким образом, выбор между этими алгоритмами определяется спецификой конкретной задачи. Например, если вычисление значения конечномерной функции выигрыша Центра затруднено и эта функция близка к вогнутой, то имитация отжига выглядит более предпочтительной. Оба алгоритма близки к методу сценариев имитационного моделирования (Belyavsky and Danilova 2019).

### 3.2. Он-лайн обучение в моделях с неполной информацией

В статье (Belyavsky et al. 2018) рассматривается задача распределения ресурсов среди производителей продукции при отсутствии информации Центра о затратах производителей и марковской информированности агентов о стратегиях Центра. В качестве модели используется динамическая версия стохастической обратной игры Штакельберга. Предлагается алгоритм построения решения игры, основанный на использовании обучения, а именно Q-learning. Возникающие при решении уравнения Беллмана содержат функции одной переменной как для Центра, так и для агентов. Изложение иллюстрируется модельными численными расчетами.

Как уже неоднократно обсуждалось, обратные игры Штакельберга (Горелик и др. 1991; Olsder 2009) представляют собой математическую формализацию задачи стимулирования (incentive problem). Центр строит механизм управления с обратной связью, побуждающий агентов выбирать выгодные для Центра действия. Рассмотрим метод решения такой задачи в динамической постановке при неполной информации о поведении агентов.

Задача формулируется таким образом, чтобы возникающие при её решении уравнения Беллмана содержали функцию одной переменной как для Центра, так и для агентов. Кроме этого, агенты после получения информации от Центра становятся независимыми игроками. Центр выбирает своё поведение, опираясь на реакции игроков. Агенты выбирают свои решения, прогнозируя поведение Центра. Наиболее простой реализацией такого поведения является марковский принцип, то есть в момент времени  $t$  Центр использует реакцию агентов в момент  $t-1$  для определения своего поведения. Агенты, в свою очередь, наблюдая за всей предысторией поведения Центра, пытаются прогнозировать его поведение в момент  $t+1$ . Метод вычисления равновесия в игре Штакельберга опирается на on-line learning, а именно на Q-learning, и рекуррентные статистические оценки.

Алгоритм  $Q$ -learning в основном применяется при решении статистической задачи динамического программирования, для чего вычисляется  $Q$ - функция. Как правило,  $Q$ -функция зависит от двух переменных: фазовой переменной и управления, в связи с чем она определена на декартовом произведении конечных множеств. Скорость сходимости метода медленная и существенно зависит от числа элементов множеств определения аргументов. Поэтому игра Штакельберга в одном из двух рассмотренных в статье вариантов формулируется таким образом, чтобы  $Q$ -функция зависела от одного аргумента. В другом варианте вместо медленного алгоритма  $Q$ -learning используется более быстрый алгоритм одномерного поиска максимума вогнутой функции одной переменной. Агенты в обоих вариантах используют рекуррентные статистики. Поэтому по сравнению со стандартным алгоритмом  $Q$ -learning предлагаемые алгоритмы должны обеспечивать более быструю сходимость к равновесному решению игры.

Основной вклад работы заключается в следующем:

1. Получены быстро сходящиеся алгоритмы вычисления решения обратной игры Штакельберга в динамической постановке при отсутствии сведений об агентах.
2. Предлагаемые алгоритмы можно рассматривать как численные методы решения соответствующей статической обратной игры Штакельберга при отсутствии информации о функциях выигрыша игроков.
3. Модель игры применена при решении задачи об оптимальном распределении ресурсов среди производителей продукции при отсутствии информации о затратах производителей.

Рассматривается ситуация, при которой Центр не обладает достаточной информацией о целях агентов, а агенты не знают стратегию Центра на весь период игры, ограничиваясь лишь предысторией (без серьезного ограничения общности информационную структуру можно считать марковской). Постановка задачи базируется на статистических оценках и обучении с

подкреплением, в том числе Q-learning (Sutton, Barto 1998; Watkins, Dayan 1992). Обучение с подкреплением для вычисления равновесия Штакельберга применялось в работах (Tharakunnel, Bhattacharyya 2009; Erev, Roth 1998). В частности, в (Tharakunnel, Bhattacharyya 2009) отмечалась неустойчивость алгоритма в случае нескольких ведомых игроков, особенно тогда, когда ведомые игроки вычисляют при решении своих задач равновесие Нэша. В данной статье предлагается модель игры Штакельберга, для которой существует устойчивый алгоритм.

Пусть имеется Центр и  $M$  управляемых центром агентов. Статическая модель стимулирования как обратная игра Штакельберга имеет следующий вид:

$$\begin{aligned} \psi(x) - \sum_{i=1}^M \varphi_i(x, \gamma) &\rightarrow \max \\ \varphi_i(x, \gamma) - f_i(x) &\rightarrow \max, i = 1, \dots, M. \end{aligned}$$

Здесь  $\psi(x)$  - вогнутая возрастающая функция дохода Центра,  $\psi(0) = 0$ ;  $f_i(x)$  - выпуклые возрастающие функции затрат агентов,  $f_i(0) = 0$ ;  $\varphi_i(x, \gamma)$  - функции компенсации Центром затрат агентов (стимулирование),  $i=1, \dots, M$ ;  $\gamma$  - управляющая стратегия Центра. Тогда оптимальный механизм стимулирования имеет вид (Новиков 2007)

$$\varphi_i^*(x, \gamma) = \begin{cases} f_i(x_i^*, x_{-i}), & x_i = x_i^*, \\ 0, & \text{иначе, } i = 1, \dots, M, \end{cases} \quad x^* \in \text{Arg max}_x \left[ \psi(x) - \sum_{i=1}^M f_i(x) \right].$$

Однако это решение не всегда может быть найдено непосредственно из-за отсутствия у Центра информации относительно затрат агентов, описываемых функциями  $f_i(x)$ . В связи с этим предлагается вычислительная схема, которая получается в результате решения стохастической динамической обратной игры Штакельберга.

Пусть Центр стимулирует деятельность агентов, распределяя между ними ресурсы. Деятельность агентов – это векторная последовательность

$x^M(t) = (x_i(t))_{i=1}^M, t = 0, 1, \dots$  В каждый момент времени  $t$  Центр выбирает управление  $\gamma(t)$ , в соответствии с которым каждый агент получает часть ресурса  $\varphi_i(x_i(t-1), \gamma(t)), i = 1, \dots, M$ . Такой способ распределения ресурса выбран потому, что Центру неизвестны затраты агентов, и при выборе локального управления  $\gamma(t)$  он может опираться только на доступную информацию о предыстории  $x^M(0), \dots, x^M(t-1)$ . Рассматривается случай марковского процесса, т.е. в каждый момент времени  $t$  используется информация только о значении  $x(t-1)$ . Поскольку Центр не может сообщить агентам полностью своё управление на весь период времени игры, а агенты не могут его вычислить, то агенты, как и Центр, должны опираться на доступную к моменту времени  $t$  информацию, то есть на последовательность  $\gamma(1), \dots, \gamma(t)$ . Далее рассмотрены две версии динамической обратной игры Штакельберга как модели распределения ресурсов.

Первая модель заключается в следующем. Множество возможных воздействий на агентов (множество сценариев Центра) конечно. Агенты предполагают, что  $\gamma(1), \dots, \gamma(t)$  - отрезок марковской последовательности, определённой на стохастическом базисе  $V^\gamma = \langle \Omega^\gamma, [F^\gamma(t)]_{t \geq 0}, F^\gamma(\infty), Q^\gamma \rangle$ ,  $F^\gamma(t) = \sigma(\gamma(1), \dots, \gamma(t))$ , то есть это минимальная сигма-подалгебра, пополненная событиями нулевой вероятности. Агенты решают свои задачи с дисконтированными выигрышами  $E_{Q_y^\gamma} \sum_{t=1}^{\infty} \beta^t [\varphi_i(x_i(t-1), \gamma(t)) - f_i(x_i(t))]$ ,  $0 < \beta < 1$ ,  $E_{Q_y^\gamma}(\cdot)$  - условное математическое ожидание по мере, порождаемой последовательностью  $\gamma$ , при условии:  $\gamma(0) = y$ .

Рассмотрим задачи агентов:

$$E_{Q_y^\gamma} \sum_{t=1}^{\infty} \beta^t [\varphi_i(x_i(t-1), \gamma(t)) - f_i(x_i(t))] \rightarrow \max.$$

Данная задача решается методом динамического программирования, для которого функции Беллмана удовлетворяют уравнениям:

$V_i(x, y) = \max_z \left[ \varphi_i(x, y) - f_i(z) + \beta \int_{-\infty}^{\infty} V_i(z, u) q^\gamma(du/y) \right].$	(3.2.1)
--	---------

При этом  $x_i(0) = x, \gamma(0) = y$ . В интеграле мера  $q^\gamma(du/y)$  - переходное ядро марковской последовательности  $\gamma$ . Представим функцию Беллмана как  $V_i(x, y) = \varphi_i(x, y) + W_i(y)$ , тогда получим интегральное уравнение Фредгольма второго рода относительно неизвестной функции  $W_i(y)$ :

$W_i(y) = \beta \int_{-\infty}^{\infty} W_i(u) q^\gamma(du/y) + \max_z \left[ -f_i(z) + \beta \int_{-\infty}^{\infty} \varphi(z, u) q^\gamma(du/y) \right],$	(3.2.2)
--	---------

поэтому существование и единственность решения уравнения Беллмана обеспечивается существованием и единственностью решения уравнения Фредгольма второго рода (Eidelman et al. 2004).

Рассмотрим пространство ограниченных функций  $B(-\infty, \infty)$  с нормой  $\sup_x |f(x)|$  и оператор  $Gf = \int_{-\infty}^{\infty} f(u) q^\gamma(du/y)$  на этом пространстве. Очевидно, что  $\|G\| \leq 1$ . Если функция  $\Phi_i(y) = \max_z \left[ -f_i(z) + \beta \int_{-\infty}^{\infty} \varphi_i(z, u) q^\gamma(du/y) \right]$  ограничена сверху, то уравнение (3.2.2) имеет единственное решение. Достаточным условием ограниченности  $\Phi_i(y)$  является ограниченность сверху функций  $\varphi_i(z, u) - f_i(z)$  для всех  $u$ . Если к тому же  $f_i(x)$  - выпуклые функции, а  $\varphi_i(z, u)$  - выпуклые функции по  $z$  при любом  $u$ , то существует единственное решение задачи:  $z_i^*(y) = \arg \max_z \left[ -f_i(z) + \beta \int_{-\infty}^{\infty} \varphi(z, u) q^\gamma(du/y) \right]$ .

Оптимальная деятельность агентов определяется равенством:

$x_i^*(t) = z_i^*(\gamma(t)).$	(3.2.3)
--------------------------------	---------



Принципиально важно, что оптимальное поведение агентов в данной модели зависит лишь от локального управления Центра  $\gamma(t)$  и никоим образом от предыдущих значений последовательности, поэтому задача Центра заключается в вычислении

$\max_{\gamma(t)} \left[ \psi \left( \sum_{i=1}^M z_i^*(\gamma(t)) \right) - \sum_{i=1}^M \varphi(x_i(t-1), \gamma(t)) \right].$	(3.2.4)
--	---------

Зависимость (3.2.3) неизвестна Центру, переходное ядро марковской последовательности  $\gamma$  неизвестно агентам. Это ядро определяется в данном случае матрицей переходных вероятностей, поскольку множество состояний марковской цепи конечно. Как уже отмечалось выше, Центр и агенты могут наблюдать за реакцией друг друга. Центр обучается, наблюдая за реакцией агентов на свои действия, агенты обучаются, наблюдая за реакцией центра на свои действия. Задача существенно упрощается по сравнению с общим случаем применения алгоритма  $Q$ -learning для решения обратной динамической игры Штакельберга, рассмотренным в работе (Erev, Roth 1998), поскольку для агентов необходимо найти приближения к переходной матрице, а для Центра требуется вычислить максимум функции одной переменной. Статистическая оценка матрицы переходных вероятностей и определения максимума функции одной переменной существенно проще вычисления  $Q$ -функций Центра и агентов. Пусть  $\Gamma = \{\gamma^1, \dots, \gamma^r\}$ , а  $P$  - матрица переходных вероятностей. Максимально правдоподобная оценка для матрицы  $P_t$  получается следующим образом. Рассмотрим последовательность матриц  $Q$ , для которых  $Q_t(i, j) = Q_{t-1}(i, j) + I_{\{\gamma_{t-1} = \gamma^i, \gamma_t = \gamma^j\}}$ , где  $I_{\{S\}}$  - индикатор множества  $S$ . Начальное значение  $Q_0 = 0$ . Максимально

правдоподобная оценка матрицы  $P_t$  находится как: 
$$P_t(i, j) = \frac{Q_t(i, j)}{\sum_k Q_t(i, k)} I_{\left\{ \sum_k Q_t(i, k) \neq 0 \right\}}$$

Пусть  $\gamma(t) = \gamma^j$ . Задача вычисления  $\max_z \left[ -f_i(z) + \beta \int_{-\infty}^{\infty} \varphi_i(z, u) q^{\gamma^j}(du/y) \right]$

трансформируется в приближённую задачу:

$\max_z \left[ -f_i(z) + \beta \sum_k P_i(j, k) \varphi_i(z, \gamma^k) \right].$	(3.2.5)
--	---------

Если функция  $F_i(z) = -f_i(z) + \beta \sum_k P_i(j, k) \varphi_i(z, \gamma^k)$  вогнута и ограничена сверху,

то задача (3.2.5) имеет единственное решение  $x_i^*(t) = \arg \max F_i(z)$ . Далее

Центр наблюдает  $R(\gamma^j)$ - реакцию агентов на управление  $\gamma^j$ :

$$R(\gamma^j) = \psi \left( \sum_{i=1}^M x_i^*(t) \right) - \sum_{i=1}^M \varphi_i(x_i(t-1), \gamma^j),$$

после этого Центр изменяет  $Q^l$  -

функцию, применяя алгоритм обучения с подкреплением

$Q_{t+1}^l(\gamma^j) = Q_t^l(\gamma^j) + h_t \left( R(\gamma^j) - Q_t^l(\gamma^j) \right),$	(3.2.6)
---	---------

и выбирает очередное воздействие на агентов с использованием

распределения вероятностей  $P_{t+1}^l = \{p_{t+1}^l(\gamma^1), \dots, p_{t+1}^l(\gamma^r)\}$ . Для вычисления

вероятностей может быть использована схема Больцмана, применяемая в алгоритме имитации отжига (Sutton, Barto 1998). Он применяется довольно

часто как алгоритм случайного поиска максимума (минимума), возможно, даже недифференцируемой функции  $f(x)$ . Идея алгоритма построена на

сравнении текущего решения  $x_t$  и случайно выбранного возможного решения  $y_t$  «в окрестности» текущего решения. Переход к случайно

выбранному решению происходит, если  $f(y_t) > f(x_t)$  ( $f(y_t) < f(x_t)$ ). Если это не

так, то переход к случайному значению  $y_t$  происходит по вероятности:

$$p_t = \exp\left(-\frac{f(x_t) - f(y_t)}{T_t}\right) \left( p_t = \exp\left(\frac{f(x_t) - f(y_t)}{T_t}\right) \right),$$

поэтому естественно

использование вероятности:

$p_{t+1}^l(\gamma^j) = \exp(Q_{t+1}^l(\gamma^j)/T_{t+1}) / \sum_k \exp(Q_{t+1}^l(\gamma^k)/T_{t+1})$	(3.2.7)
--	---------

для выбора следующего значения  $\gamma$ . В (3.2.7) и выше параметр  $T$  - «температура», которая управляет степенью случайности при выборе воздействия на агентов. Этот параметр уменьшается от выбранного максимального значения до выбранного минимального значения, например, следующим образом:  $T_{t+1} = \delta T_t, 0 < \delta < 1$ . Начальное значение для  $Q_0^l \equiv 0$ . Для  $h_t$ , регулирующих амплитуду изменений, должны выполняться стандартные для такого рода алгоритмов условия:  $\sum_{t=1}^{\infty} h_t = \infty, \sum_{t=1}^{\infty} h_t^2 < \infty$ , обеспечивающие сходимость с вероятностью единица к близкой к  $R$  оптимальной функции  $Q$ , если при обучении каждый из элементов множества  $\Gamma$  встречается бесконечное число раз.

Таким образом, предложенный алгоритм обучения состоит из повторяющихся действий:

1. Вычисление максимально правдоподобной оценки матрицы переходных вероятностей.
2. Вычисление очередного значения  $Q$ -функции.

Ещё раз отметим важное преимущество предлагаемого алгоритма:  $Q$ -функция зависит от одного аргумента. Первое действие связано с вычислением максимально правдоподобной оценки вместо вычисления очередного значения  $Q$ -функции, что обеспечивает более высокую скорость сходимости алгоритма по сравнению с алгоритмами, в которых  $Q$ -функция зависит от двух аргументов и очередное значение  $Q$  вычисляется и на первой, и на второй стадиях алгоритма. Заметим также, что максимально правдоподобные оценки являются состоятельными и асимптотически эффективными, то есть использующими всю информацию, и распределены асимптотически по нормальному закону.

Во второй модели игры агенты предполагают, что Центр не может совершать «резких движений» (постулат 2 из (Белявский и др. 2018)) при выборе управления, и моделируют поведение Центра следующим образом:

$\gamma(t) = \gamma(t-1)(1 + \varepsilon_t).$	(3.2.8)
---	---------

Последовательность  $\varepsilon_t$  состоит из независимых и одинаково распределённых элементов с массой общего закона распределения, сосредоточенной в окрестности нуля. Для второй модели задача оптимизации

$\max_z \left[ -f_i(z) + \beta \int_{-\infty}^{\infty} \varphi_i(z, u) q^\gamma (du / y) \right]$  будет выглядеть как задача вычисления

$\max_z \left[ -f_i(z) + \beta E_\varepsilon \varphi(z, y(1 + \varepsilon)) \right]$ . Для второго аргумента применим формулу

Тейлора до третьего слагаемого включительно, как это принято в стохастическом исчислении (например, формула Ито), в результате получим задачу оптимизации

$\max_z \left[ -f_i(z) + \beta \left( \varphi_i(z, y) + \frac{\partial \varphi_i(z, y)}{\partial y} y E_\varepsilon \varepsilon + \frac{1}{2} \frac{\partial^2 \varphi_i(z, y)}{\partial y^2} y^2 E_\varepsilon \varepsilon^2 \right) \right],$	(3.2.9)
---	---------

при условии существования второй производной. Если для всех  $y$  целевые функции агентов вогнутые и ограниченные, то, как и ранее, существуют единственные решения задач агентов (3.2.9) :  $z_i^*(y)$ . Состоятельными

оценками для моментов в правой части формулы (3.2.9) служат выборочные

средние моментов распределения  $\gamma$  :  $a_t = \frac{1}{t} \sum_{i=1}^t \frac{\Delta \gamma(i)}{\gamma_{i-1}}, b_t = \frac{1}{t} \sum_{i=1}^t \left( \frac{\Delta \gamma(i)}{\gamma_{i-1}} \right)^2$ . Таким

образом, агенты решают задачи:

$\max_z \left[ -f_i(z) + \beta \left( \varphi_i(z, \gamma(t)) + \frac{\partial \varphi_i(z, \gamma(t))}{\partial y} \gamma(t) a_t + \frac{1}{2} \frac{\partial^2 \varphi_i(z, \gamma(t))}{\partial y^2} \gamma(t)^2 b_t \right) \right],$	
---	--

а Центр может использовать любой метод поиска максимума вогнутой функции от одной переменной, в котором не требуется вычисление

производной. Если функция  $R(x)$  вогнутая и ограниченная сверху, то из состоятельности выборочных моментов и сходимости метода одномерного поиска следует сходимость алгоритма к равновесию игры Штакельберга по вероятности.

Таким образом, алгоритм обучения для второй модели заключается в повторении следующих действий:

1. Вычисление оценок моментов первого и второго порядка.
2. Вычисление очередного приближения  $\gamma$ .

Преимущество данного подхода по сравнению с классическим  $Q$ -обучением, применённым в работах (Erev, Roth 1998; Tharakunnel, Bhattacharyya 2009), заключается в том, что авторский алгоритм вообще не вычисляет  $Q$ -функцию, а базируется на оценках первого и второго момента распределения и одномерного поиска максимума функции от одной переменной. По сравнению с предыдущим алгоритмом, данный алгоритм не требует предварительного анализа задачи для определения сценариев поведения Центра (множество  $\Gamma$ ).

Рассмотрим пример, в котором  $\psi(x) = \sqrt{x}$ ,  $\varphi_i(x, y) = xy$  и  $f_i(x) = \mu_i x^2$ . Для первой модели задача (3.2.5) имеет вид:

$\max_z \left[ -\mu_i z^2 + \beta z \sum_k P_i(j, k) \gamma^k \right],$	
---	--

а ее решение:  $x_i^*(t) = \frac{1}{2\mu_i} \beta \sum_k P_i(j, k) \gamma^k$ . Реакция  $R(\gamma^j) = \sqrt{\sum_{i=1}^M x_i^*(t) - \gamma^j \sum_{i=1}^M x_i^*(t-1)}$ .

Для второй модели задача (3.2.9) имеет вид:

$\max_z \left[ -\mu_i z^2 + \beta z (\gamma(t) + a_t) \right],$	
---	--

а ее решение:  $x_i^*(\gamma(t)) = \frac{\beta(\gamma(t) + a_t)}{2\mu_i}$ , где  $a_t = \frac{1}{t} \sum_{i=1}^t \frac{\Delta\gamma(i)}{\gamma(i-1)}$ . Реакция  $R(\gamma(t))$  имеет

тот же вид, что и выше.

Равновесие Штакельберга для этой задачи определяется равенствами:

$$x_i^* = \frac{\beta\gamma}{2\mu_i}, \gamma^* = \left( \frac{1}{2\sqrt{2\beta\sum_{i=1}^M \frac{1}{\mu_i}}} \right)^{2/3}$$

Рассмотрим вычислительный пример с параметрами игры:

$M$	2
$\beta$	0.9
$\mu_1$	1
$\mu_2$	2

Равновесные значения:  $\gamma^* = 0.4524029361, x_1^* = 0.2035813212, x_2^* = 0.1017906606$ .

Рассмотрим ещё два примера. В первом примере множество  $\Gamma = \{0.4, 0.45, 0.5\}$ , содержит значение, близкое к равновесному значению на второй позиции. Для расчётов использовался пакет Maple. В результате работы алгоритма получены следующие результаты. Значения  $Q$ -функции:

$Q = \text{vector}([.2789054257e-1, .4073228841, .9271767315e-1])$ . Максимальное значение  $Q$ -функции равно .4073228841 и находится во второй позиции, что

соответствует значению  $\gamma = 0.45$  из множества  $\Gamma$ . Вычисленная матрица

переходных вероятностей  $P = \text{matrix}([[0, 3/4, 1/4], [1/91, 87/91, 3/91], [3/4, 1/4, 0]])$ .

Максимальный элемент матрицы, близкий к единице, находится во втором столбце и второй строке. Это означает, что наиболее вероятным префиксом цепочки является 0.45, 0.45, ..... что также соответствует значению  $\gamma = 0.45$  из множества  $\Gamma$ . Вычисленные значения  $x_1 = .2029945055, x_2 = .1014972528$

близки к равновесным.

Во втором примере множество  $\Gamma = \{0.2, 0.4, 0.6\}$  не содержит значения, достаточно близкого к равновесному. Вычисленные значения функции

$Q = \text{vector}([.1254077349, .4028884358, .5489325521e-1])$ . Максимальное значение функции  $Q$ , равное  $.4028884358$ , находится во второй позиции, что соответствует значению  $\gamma = 0.4$  из множества  $\Gamma$ . Матрица переходных вероятностей, вычисленная агентами, есть

$P = \text{matrix}([[1/5, 3/5, 1/5], [1/90, 43/45, 1/30], [3/4, 1/4, 0]])$ . Максимальный элемент матрицы, близкий к единице, также находится во втором столбце и второй строке. Вычисленные агентами значения  $x_1 = .1820000000, x_2 = .9100000000e-1$  несколько дальше от равновесных значений. Результаты работы алгоритма для второй модели приведены в таблице 3.2.1.

Таблица 3.2.1

$\gamma$	$x_1$	$x_2$	$R$
0.35	.1837500000	.0918750000C	.2625000000
0.36	.1778142858	.0889071429C	.4172257999
0.37	.1788773809	.08943869045	.4193054183
0.38	.1816893340	.0908446670C	.4200877844
0.39	.1852001900	.0926000950C	.4207793681
0.4	.1890599915	.09452999575	.4214115632
0.41	.1931187782	.0965593891C	.4219456888
0.42	.1973015878	.0986507939C	.4223502262
0.43	.2015667888	.1007833944	.4226042619
0.44	.2058894150	.1029447075	.4226943805
0.45	.2102535881	.1051267940	.4226120220
0.445	.2070930676	.1035465338	.4170062493
0.4475	.2078141892	.1039070946	.4193088334
0.45	.2085934568	.1042967284	.4190913001
0.44875	.2075715326	.1037857663	.4175845248
0.449375	.2075256146	.1037628073	.4180160874

Помечающие столбцы символы соответствуют своему приведенному ранее определению. Представляется, что второй метод имеет преимущество, когда среди возможных значений множества  $\Gamma$  отсутствует значение, близкое к равновесному. Второй алгоритм в рассмотренном примере сошёлся к равновесным значениям значительно быстрее первого. В случае, когда  $\Gamma$  содержит значение, близкое к равновесному, получен более точный результат в первом алгоритме. Когда  $\Gamma$  не содержит значения, близкого к равновесному, более точный результат получен во втором алгоритме.

Механизм пропорционального распределения - один из наиболее естественных способов выделения ресурсов, апробированный практикой управления в организационных системах. На первый план здесь выходит проблема защиты от манипулирования, поскольку при использовании этого механизма агентам выгодно завышать свои реальные потребности в ресурсах, точно не известные Центру. В (Новиков 2007) проведено исследование статического механизма пропорционального распределения и предложена его модификация, обеспечивающая защиту от манипулирования. В настоящей работе предложен динамический механизм пропорционального распределения ресурсов, использующий алгоритм обучения. Построены две стохастические модели динамической обратной игры Штакельберга, каждая из которых гарантирует существование стационарного равновесия. В первой модели применяется идеология конечного множества сценариев поведения Центра, во второй используется естественная ограниченность в действиях Центра. Обе проиллюстрированы численными расчетами. Каждая из моделей порождает свой алгоритм вычисления равновесия в обратной игре Штакельберга. Результаты эксперимента позволяют сделать предположение о том, что предлагаемые алгоритмы решения стохастической динамической обратной игры Штакельберга одновременно решают соответствующую статическую обратную игру Штакельберга при отсутствии информации о затратах агентов. Однако эта гипотеза требует дополнительного анализа, который планируется провести в дальнейшем (Belyavsky et al. 2018).



### **3.3. Качественно репрезентативные сценарии имитационного моделирования**

Метод имитационного моделирования (simulation modeling) или компьютерной имитации (computer simulation) играет очень важную роль в решении задач управления сложными динамическими системами. Его идея состоит в проведении вычислительных экспериментов с программной реализацией модели управляемой динамической системы, позволяющих оценить влияние различных управляющих воздействий на поведение этой системы. Коротко говоря, имитационная модель отвечает на вопрос "Что будет, если...?"

Имитационное моделирование активно развивается начиная со второй половины XX века. К числу первых обобщающих работ в этой области следует отнести монографии (Нейлор, 1975; Шеннон, 1978). Согласно Р. Шеннону, "имитационное моделирование есть процесс конструирования модели реальной системы и постановки экспериментов на этой модели с целью либо понять поведение системы, либо оценить (в рамках ограничений, накладываемых некоторым критерием или совокупностью критериев) различные стратегии, обеспечивающие функционирование данной системы. Таким образом, процесс имитационного моделирования мы понимаем как процесс, включающий и конструирование модели, и аналитическое применение модели для изучения некоторой проблемы" (Шеннон 1978:12).

В Советском Союзе существенный вклад в развитие теории и практики имитационного моделирования внесли академик Н.Н. Моисеев (Моисеев 1979) и его ученики (Иванилов и др. 1993; Павловский 2000). Технические вопросы имитационного моделирования освещены в (Клейнен 1978) и адаптированы в (Угольницкий 1999) с приложениями к управлению эколого-экономическими системами. Современное состояние вопроса изложено в монографиях (Лоу и Кельтон 2004; Kleijnen 2007); интерес представляет статья (Nelson 2016) с подробной библиографией.

Авторский подход к использованию качественно репрезентативных сценариев имитационного моделирования описан в статье (Ougolnitsky and Usov 2018), более общие вопросы управления активными системами - в монографии (Угольницкий 2016).

Рассмотрим задачу оптимального управления как базовую модель управления динамической системой. Для удобства перехода к программной реализации используем сразу разностную постановку задачи:

$$J = \sum_{t=1}^T e^{-\rho t} g(x^t, u^t) \rightarrow \max \quad (3.3.1)$$

$$0 \leq u^t \leq 1 ; \quad (3.3.2)$$

$$x^{t+1} = x^t + f(x^t, u^t), \quad x^0 = x_0 . \quad (3.3.3)$$

Для удобства изложения введём для модели (3.3.1)-(3.3.3) ряд упрощающих предположений, не ограничивающих общности рассмотрения. Здесь  $x^t$  - фазовая переменная, характеризующая состояние моделируемой динамической системы в момент времени  $t$ . В общем случае эта переменная векторная, однако будем для простоты считать её скалярной (состояние системы характеризуется единственным показателем). Через  $u^t$  обозначена переменная управления, которая также считается скалярной;  $g(x^t, u^t)$  - текущая целевая функция субъекта управления;  $J$  - его суммарный функционал выигрыша;  $T$  - период рассмотрения;  $\rho \in [0,1]$  - коэффициент дисконтирования.

В общей постановке ограничения на управления задаются в виде  $u^t \in U$ , однако использованная в соотношении (3.3.2) гипотеза  $U = [0,1]$  не нарушает общности. Уравнение (3.3.3) задает динамику управляемой системы с помощью нелинейной функции  $f(x^t, u^t)$ . Вновь в общем случае возможен переход от шага  $t$  к  $t + \Delta t$ , однако примем  $\Delta t = 1$ .

Теперь можно сформулировать постановку задачи имитационного моделирования. Предположим, что в модели (3.3.1)-(3.3.3) заданы функции

$f, g$  со всеми своими числовыми параметрами, а также параметры  $T, \rho$  и начальные условия  $x_0$ . Тогда для каждого набора управляющих воздействий ( $j$ -го сценария имитации)  $\{u_{(j)}^t\}_{t=1}^T$  модель (3.3.1)-(3.3.3) как вычислительная схема определяет соответствующую фазовую траекторию  $\{x_{(j)}^t\}_{t=1}^T$  и значение выигрыша  $J_{(j)}$ . Задавая различные сценарии  $j = 1, \dots, N$ , можно с помощью модели оценивать их влияние на поведение управляемой динамической системы - это и есть компьютерная имитация.

Пусть теперь на динамическую систему одновременно и независимо друг от друга воздействуют несколько субъектов управления. Это приводит к разностной игре вида

$$J_i = \sum_{t=1}^T e^{-\rho t} g_i(x^t, u^t) \rightarrow \max \quad (3.3.4)$$

$$0 \leq u_i^t \leq 1, i = 1, \dots, n, \quad (3.3.5)$$

в силу уравнения динамики (3.3.3),

где  $u_i^t$  - управляющая переменная  $i$ -го субъекта управления;  $g_i, J_i$  - его текущая и суммарная функции выигрыша;  $n$  - число субъектов управления;  $u^t = (u_1^t, \dots, u_n^t)$ . Сценарий имитации по-прежнему имеет вид  $\{u_{(j)}^t\}_{t=1}^T$ , но теперь под  $u^t$  понимается набор управлений всех субъектов в момент времени  $t$ . Как и ранее, каждый сценарий порождает свою фазовую траекторию  $\{x_{(j)}^t\}_{t=1}^T$  и набор выигрышей субъектов  $J_{1(j)}, \dots, J_{n(j)}$ .

Наконец, рассмотрим случай древовидной структуры управления, когда воздействующие на систему субъекты, в свою очередь, подвергаются воздействию со стороны некоторого выделенного субъекта управления (Центра). Сам Центр не влияет на систему непосредственно, а добивается своих связанных с ней целей путём воздействия на остальных субъектов. Это приводит к модели иерархической разностной игры вида

$$J_0 = \sum_{t=1}^T e^{-\rho t} g_0(x^t, u^t, v^t) \rightarrow \max \quad (3.3.6)$$

$$0 \leq v_i^t \leq 1, i = 1, \dots, n, \quad (3.3.7)$$

$$J_i = \sum_{t=1}^T e^{-\rho t} g_i(x^t, u^t, v_i^t) \rightarrow \max \quad (3.3.8)$$

$$0 \leq u_i^t \leq 1, i = 1, \dots, n, \quad (3.3.9)$$

в силу уравнения динамики (3.3.3).

Здесь  $v^t = (v_1^t, \dots, v_n^t)$  - вектор управляющих переменных Центра (предполагается, что Центр управляет каждым  $i$ -м субъектом по отдельности с помощью переменной  $v_i^t$ );  $g_0, J_0$  - текущая и суммарная функции выигрыша Центра. Соответственно, теперь сценарий имитации имеет вид  $\{v_{(j)}^t, u_{(j)}^t\}_{t=1}^T$  и по-прежнему порождает фазовую траекторию  $\{x_{(j)}^t\}_{t=1}^T$  и набор выигрышей  $J_{0(j)}, J_{1(j)}, \dots, J_{n(j)}$ .

Согласно теории управления устойчивым развитием активных систем (Угольницкий 2016), каждая из трёх указанных моделей может быть дополнена условием гомеостаза

$$\forall t \ x^t \in X^*, \quad (3.3.10)$$

налагающим определенные требования (математически - фазовые ограничения) на состояние управляемой динамической системы.

С практической точки зрения основную роль играет вопрос о том, какие сценарии имитации нужно выбирать для решения задачи управления с помощью имитационного моделирования. Множества сценариев для рассмотренных задач управления показаны в таблице 3.3.1.

Из таблицы 3.3.1 понятно, что решение задачи имитационного моделирования полным перебором невозможно, поскольку даже в простейшем случае скалярного оптимального управления множество сценариев имеет мощность континуума. Собственно, основная техническая

часть теории имитационного моделирования сводится именно к обоснованию различных способов частичного перебора множества возможных сценариев (планированию имитационных экспериментов).

Таблица 3.3.1

Множества сценариев имитационного моделирования			
	Задача оптимального управления (3.3.1)-(3.3.3)	Разностная игра (3.3.3)-(3.3.5)	Иерархическая разностная игра (3.3.3), (3.3.6)-(3.3.9)
Скалярные управления	$[0,1]$	$[0,1]^n$	$[0,1]^{n+1}$
Векторные управления размерности $m$	$[0,1]^m$	$[0,1]^{nm}$	$[0,1]^{(n+1)m}$

Обычно считается, что множество сценариев конечно: для описанных моделей это означает, что вводится некоторое конечное разбиение отрезка  $[0,1]$  и в качестве сценариев берутся точки этого разбиения. Это позволяет использовать для решения указанной задачи методы теории планирования эксперимента (Клейнен 1978; Угольницкий 1999).

Тем не менее, и конечное разбиение с большим числом точек может оказаться неподъёмным для проведения расчётов и анализа их результатов даже при одном, а тем более при нескольких субъектах управления, поскольку в реальных задачах переменные управления обычно векторные. Поэтому для решения динамических задач управления с помощью имитационного моделирования был предложен эвристический *метод качественно репрезентативных сценариев* (Ougolnitsky and Usov 2018), основанный на содержательных соображениях о процессе управления.

Пусть, например, фазовая переменная  $x^t$  - концентрация загрязняющего вещества в водоёме, а переменная управления  $u^t$  - доля очистки стоков некоторого предприятия в этот водоём с ограничениями (3.3.2). Очевидно,

прежде всего следует рассмотреть два "крайних" сценария:  $u^t = 0$  (отказ от очистки) и  $u^t = 1$  (полная очистка). Отметим, что на практике полная очистка технически невозможна (точнее, требует бесконечных затрат), но этот сценарий наряду со сценарием отказа дает представление о возможных последствиях двух диаметрально противоположных стратегий. Кроме того, разумно рассмотреть промежуточный случай  $u^t = 0.5$ , при котором очищается половина стоков. Идея метода качественно репрезентативных сценариев (КРС) заключается в том, что уже эти три сценария дают в первом приближении достаточно хорошую качественную картину возможных воздействий на управляемую систему и их последствий. Три сценария описывают минимальное, среднее и максимальное воздействие на управляемую динамическую систему. Понятно, что эта идея применима при анализе не только очистки стоков, но и любого воздействия.

Возможен и другой подход к интерпретации трёх сценариев. В приведённом примере, если известно, что технические возможности очистки ограничены, то разбиение будет включать точки 0, 0.25 и 0.5, также соответствующие минимальному, среднему и максимальному воздействию в технически допустимом диапазоне  $[0,0.5]$ .

Конечно, при наличии достаточных вычислительных и временных возможностей можно измельчать разбиение, выбирая в качестве дополнительных точек середины полученных на предыдущем этапе отрезков, однако идея метода КРС именно в том, что для получения качественных оценок достаточно очень небольшого числа сценариев.

Проверить обоснованность этой гипотезы позволяет следующий подход, представленный для наиболее общего случая модели (3.3.6)-(3.3.9).

Пусть  $\Omega = V_1 \times \dots \times V_n \times U_1 \times \dots \times U_n$  - множество исходов игры. Метод КРС исходит из предположения  $U_i = U_i^{QRS}$ ,  $V_i = V_i^{QRS} \forall i = 1, 2, \dots, n$ , где множества  $U_i^{QRS}$  и  $V_i^{QRS}$  содержат качественно репрезентативные стратегии агента  $i$  и Центра по отношению к агенту  $i$ , соответственно. Предположим дополнительно, что

мощность множеств  $V_i^{QRS}$  и  $U_i^{QRS}$  ( $i=1,2,\dots,n$ ) достаточно мала, так что они содержат по малому числу  $K$  элементов, т.е.  $|U_i^{QRS}|=|V_i^{QRS}|=K$ . Тогда

$V_1^{QRS} \times \dots \times V_n^{QRS} \times U_1^{QRS} \times \dots \times U_n^{QRS} = QRS$  есть множество КРС данной игры,

которое включает  $m=|QRS|=\prod_{i=1}^n |V_i^{QRS}| |U_i^{QRS}|=K^{2n}$  элементов. Каждый

репрезентативный сценарий игры  $(v,u)^{(k)} \in QRS$ ,  $k=1,2,\dots,m$ , имеет вид

$$(v,u)^{(k)}=(v_1^{(k)},v_2^{(k)},\dots,v_n^{(k)},u_1^{(k)},u_2^{(k)},\dots,u_n^{(k)});v_i^{(k)} \in V_i^{QRS};u_i^{(k)} \in U_i^{QRS};i=1,2,\dots,n.$$

*Определение.* Множество  $QRS = \{(v,u)^{(1)},(v,u)^{(2)},\dots,(v,u)^{(m)}\}$  называется множеством КРС иерархической разностной игры с точностью  $\Delta$ , если:

(а) для любых двух элементов этого множества  $(v,u)^{(i)},(v,u)^{(j)} \in QRS$

справедливо

$$|J_0^{(i)} - J_0^{(j)}| > \Delta;$$

(б) для любого другого элемента  $(v,u)^{(l)} \notin QRS$  найдётся элемент  $(v,u)^{(j)} \in QRS$  такой, что  $|J_0^{(l)} - J_0^{(j)}| \leq \Delta$ .

Здесь  $J_0^{(i)}, J_0^{(j)}, J_0^{(l)}$  - соответствующие выигрыши Центра (в смысле (3.3.6));  $J_0^{(s)} = J_0(v_1^{(s)}, v_2^{(s)}, \dots, v_n^{(s)}, u_1^{(s)}, u_2^{(s)}, \dots, u_n^{(s)})$ ,  $s = i, j, l$ ;  $\Delta > 0$  - константа, определяющая точность. Таким образом, КРС ведут к существенному различию в выигрышах Центра, а различие между одним из репрезентативных и любым другим сценарием в этом смысле несущественно.

Это определение подходит и для задачи оптимального управления (3.3.1)-(3.3.3), где вместо  $J_0$  берется функционал выигрыша единственного субъекта управления  $J$ .

Для разностной игры (3.3.3)-(3.3.5) с помощью метода КРС можно находить множество равновесий Нэша (решений игры). Пусть множество КРС для каждого субъекта содержит  $K$  элементов, т.е.

$$U_i^{QRS} = \{u_i^{(1)}, u_i^{(2)}, \dots, u_i^{(K)}\}; |U_i^{QRS}| = K; i = 1, 2, \dots, n.$$

Имеется  $m=|QRS|=\prod_{i=1}^n |U_i^{QRS}|=K^n$  качественно различных сценариев игры.

Предлагается следующий алгоритм построения множества равновесий Нэша в чистых стратегиях  $NE(QRS)$  на множестве КРС (Ougolnitsky and Usov 2018).

1. Задать все функции и параметры модели (3.3.3) – (3.3.5).

Положить  $i=1$ .

2. Для  $i$ -го субъекта ( $i=1, \dots, n$ ) построить множество индивидуально неулучшаемых стратегий  $Ne_i$ :

$$Ne_i = \{(u_1, u_2, \dots, u_{i-1}, u_i^*, u_{i+1}, \dots, u_n)\}:$$

$$\forall u_j \in U_j^{QRS}; j \neq i; j = 1, 2, \dots, n; u_i^* \in U_i^{QRS}:$$

$$(u_1, u_2, \dots, u_{i-1}, u_i^*, u_{i+1}, \dots, u_n) = \arg \max_{1 \leq l \leq K} J_i(u_1, u_2, \dots, u_{i-1}, u_i^{(l)}, u_{i+1}, \dots, u_n) \}$$

3. Множество равновесий Нэша на множестве КРС есть

$$NE(QRS) = \bigcap_{i=1}^n Ne_i .$$

Множества  $Ne_i$  ( $i=1, 2, \dots, n$ ) строятся полным перебором множества КРС игры для фиксированной стратегии  $i$ -го агента. Для проверки всех КРС игры

в этом случае требуется  $\prod_{\substack{i=1 \\ j \neq i}}^n |U_j^{QRS}| = K^{n-1}$  операций. Если число агентов и

величина  $K$  не слишком велики (например,  $n=10, K=3-4$ ), то проведение расчётов технически возможно. Если множество равновесий Нэша в модели (3.3.3) - (3.3.5) существует, но не находится посредством предложенного алгоритма, то нужно обновить множество КРС игры путем добавления новых сценариев для каждого субъекта и вернуться на шаг 2 алгоритма при  $i=1$ .

Информационные структуры разностных или дифференциальных игр определяются следующими парами классификационных признаков (Угольницкий 2016):

- программные стратегии  $v(t) = \varphi(t)$  или позиционные стратегии  $v(t) = \varphi(t, x(t))$ ;



- стратегии без обратной связи по управлению  $v(t) = \varphi(t)$  или с таковой  $v(t) = \varphi(t, u(t))$ . В первом случае возникают игры Гермейера типа  $\Gamma_{1t}$  или  $\Gamma_{1x}$  (игры Штакельберга), во втором - игры Гермейера типа  $\Gamma_{2t}$  или  $\Gamma_{2x}$  (обратные игры Штакельберга);

- принуждение  $v(t) = q(t)$  или побуждение  $v(t) = p(t)$ . Принуждение означает воздействие Центра на множества допустимых управлений субъектов, а побуждение - на их функции выигрыша.

Рассмотрим специфику имитационного моделирования на примере иерархической игры двух лиц Центр - субъект управления (Ведущий - Ведомый), обозначенных соответственно индексами L и F (Leader - Follower), для метода побуждения:

$$J_L = \int_0^T e^{-\rho t} g_L(p(t), u(t), x(t)) dt + e^{-\rho T} S_L(x(T)) \rightarrow \max, p(t) \in P(t); \quad (3.3.11)$$

$$J_F = \int_0^T e^{-\rho t} g_F(p(t), u(t), x(t)) dt + e^{-\rho T} S_F(x(T)) \rightarrow \max, u(t) \in U(t); \quad (3.3.12)$$

$$\dot{x} = f(x(t), u(t)), \quad x(0) = x_0. \quad (3.3.13)$$

Во всех случаях имитационное моделирование предполагает дискретизацию модели, т.е.

$$[0, T] \rightarrow 0, \dots, T, \quad x(t+1) = x(t) + f(x(t), u(t)), \quad J = \sum_{t=0}^T e^{-\rho t} g(p(t), u(t), x(t)).$$

В случае регламента  $\Gamma_{1t}$  стратегии имеют вид  $p(t), u(t)$ . Сценарий  $j = 1, \dots, N$  представляет собой пару дискретных управляющих траекторий  $\{p^{(j)}(t)\}_{t=0}^T, \{u^{(j)}(t)\}_{t=0}^T$ , для которых вычисляются дискретная фазовая траектория  $\{x^{(j)}(t)\}_{t=0}^T$  и выигрыши игроков  $J_L^{(j)}, J_F^{(j)}$ . Используется небольшое число сценариев, отражающих основные качественно различные управления игроков с учетом специфики моделируемой системы (метод КРС).

В случае регламента  $\Gamma_{2t}$  стратегии имеют вид  $p(t, u(t)), u(t)$ . Как правило, побуждение осуществляется посредством механизма поощрения-наказания вида

$$p(t) = \begin{cases} p^R(t), & u(t) \in U_R, \\ p^P(t), & \text{иначе,} \end{cases} \quad (3.3.14)$$

где  $U_R$  - множество значений желаемых для Ведущего управлений Ведомого,  $p^R(t), p^P(t)$  - стратегии поощрения и наказания Ведомого Ведущим соответственно. Во многих случаях можно считать, что  $p^R(t) \equiv p^R, p^P(t) \equiv p^P$ . Для каждой стратегии Ведомого  $\{u^{(j)}(t)\}_{t=0}^T$ , входящей в сценарий  $j = 1, \dots, N$ , на каждом шаге  $t = 0, 1, \dots, T$  реализуется проверка (3.3.14), после чего вычисляется соответствующее значение  $x^{(j)}(t)$  и очередные слагаемые  $e^{-\rho t} g_L^{(j)}(p^{(j)}(t), u^{(j)}(t), x^{(j)}(t)), e^{-\rho t} g_F^{(j)}(p^{(j)}(t), u^{(j)}(t), x^{(j)}(t))$ .

В случае регламента  $\Gamma_{1x}$  стратегии имеют вид  $p(t, x(t)), u(t, x(t))$ , и нужно задавать правила

$$p^{(j)}(t, x(t)) = \begin{cases} p_1^{(j)}, & x(t) \in X_1^L, \\ p_2^{(j)}, & x(t) \in X_2^L, \\ \dots \\ p_m^{(j)}, & x(t) \in X_m^L, \end{cases} \quad u^{(j)}(t, x(t)) = \begin{cases} u_1^{(j)}, & x(t) \in X_1^F, \\ u_2^{(j)}, & x(t) \in X_2^F, \\ \dots \\ u_l^{(j)}, & x(t) \in X_l^F, \end{cases} \quad j = 1, \dots, N, \quad (3.3.15)$$

где  $X = X_1^L \cup \dots \cup X_m^L, X_k^L \cap X_r^L = \emptyset, X = X_1^F \cup \dots \cup X_l^F, X_k^F \cap X_r^F = \emptyset$ , и далее пошагово реализовать эти правила для всех сценариев  $\{p^{(j)}(t)\}_{t=0}^T, \{u^{(j)}(t)\}_{t=0}^T, j = 1, \dots, N$ . Конечно, области  $X_{i,j}^{L,F}$  должны иметь содержательную интерпретацию в терминах моделируемой системы, и их немного.

Рассмотрим в качестве развёрнутого примера модель социального партнёрства в системе дополнительного профессионального образования с тремя обобщёнными агентами: преподаватель (В), работодатель (Р) и студент (С) (Tarasenko et al., 2016; Ougolnitsky and Usov 2018; Нор-Аревян и др. 2018). В отличие от (Угольницкий 2016) теперь продемонстрируем применение для исследования этой модели метода КРС. Сначала предполагается, что все

агенты равноправны и принимают решения одновременно и независимо друг от друга, стремясь максимизировать свои целевые функционалы вида

$$J_i = \int_0^T e^{-\rho t} [g_i(r_i - u_i(t)) + s_i(t)c(x(t))]dt + e^{-\rho T} s_i(T)c(x(T)) \rightarrow \max, \quad (3.3.16)$$

$$0 \leq u_i(t) \leq r_i, \quad i = B, P, C. \quad (3.3.17)$$

Каждый агент  $i$  из множества  $N = \{B, P, C\}$  распределяет свой бюджет  $r_i$  между двумя направлениями: доля  $u_i(t)$  (управление агента) направляется на повышение уровня профессиональной подготовки студентов, а остаток  $r_i - u_i(t)$  используется для финансирования его частной деятельности. Текущий доход каждого агента складывается из двух слагаемых: доход от частной деятельности и доля в выгоде, определяемой уровнем профессиональной подготовки студентов. В модели (3.3.16) - (3.3.17):  $\rho$  – коэффициент дисконтирования;  $g_i(z)$  – вогнутая возрастающая функция переменной  $z$ , отражающая доход от частной деятельности агента  $i$ ;  $x(t)$  – переменная состояния, описывающая уровень профессиональной подготовки студентов;  $c(x)$  – вогнутая возрастающая функция переменной  $x$ , дающая финансовое выражение общественной полезности, определяемой уровнем профессиональной подготовки студентов;  $s_i(t)$  – доля  $i$ -го агента в этой величине. Уравнение динамики состояния имеет вид

$$\dot{x} = h(x(t)) + f(u_B(t), u_P(t), u_C(t)), \quad x(0) = x_0. \quad (3.3.18)$$

Здесь  $h$  – убывающая функция своего аргумента, поскольку в отсутствие инвестиций уровень профессиональной подготовки снижается;  $f$  – возрастающая функция инвестиций агентов в уровень профессиональной подготовки студентов;  $x_0$  – начальное значение переменной состояния.

Рассмотрим упрощённую линейную по состоянию версию модели (3.3.16) – (3.3.18) вида

$$J_i(u(\cdot)) = \int_0^T e^{-\rho t} [k_i(r_i - u_i(t))^{p_i} + s_i(t)cx(t)]dt + e^{-\rho T} s_i(T)cx(T) \rightarrow \max, \quad (3.3.19)$$

$$0 \leq u_i(t) \leq r_i, \quad i \in N; \quad (3.3.20)$$

$$\dot{x} = ax(t) + \sum_{i \in N} b_i u_i(t), \quad x(0) = x_0. \quad (3.3.21)$$

В (3.3.19) – (3.3.21) по сравнению с (3.3.16) – (3.3.18) используются линейные функции  $c(x(t)) = cx(t)$  ( $c = \text{const} > 0$ ) и  $h(x(t)) = ax(t)$  ( $a = \text{const} < 0$ ), что делает модель линейной по состоянию (Dockner et al. 2000). Для простоты, функция  $f(u_B(t), u_P(t), u_C(t)) = \sum_{i \in N} b_i u_i(t)$  также выбрана линейной, а функция дохода от частной деятельности считается степенной без существенной потери общности:  $g_i(r_i - u_i(t)) = k_i(r_i - u_i(t))^{p_i}$ ,  $k_i > 0, 0 < p_i < 1, i \in N$ . Здесь  $b_i > 0$  показывает вклад инвестиций  $i$ -го агента в повышение уровня профессиональной подготовки студентов.

Как и в параграфе 2.3, предположим, что управляющие стратегии всех агентов остаются постоянными в течение длительных периодов времени. Тогда стратегия каждого агента есть сеточная функция  $u_i(t) = \{u_{ik}\}_{k=1}^M; i = B, P, C$ .

Соотношения (3.3.19), (3.3.20) изменяются следующим образом:

- целевые функционалы становятся целевыми функциями:

$$J_i(\{u_{ik}\}_{k=1}^M) = \sum_{k=1}^M (k_i(r_{ik} - u_{ik})^{p_i} (e^{-\rho t_{k-1}} - e^{-\rho t_k}) / \rho + \int_0^T e^{-\rho t} s_i(t) cx(t) dt + e^{-\rho T} s_i(T) cx(T)) \rightarrow \max; \quad i \in N; \quad (3.3.22)$$

- ограничения на управления принимают вид

$$0 \leq u_{ik} \leq r_{ik}, \quad i \in N; k = 1, 2, \dots, M; \quad (3.3.23)$$

Как и ранее,  $f_{ik} = f_i(k\Delta t), i \in N$ .

Модель (3.3.21)–(3.3.23) - это дифференциальная игра в нормальной форме, в которой ищутся равновесия Нэша. Это тройки сеточных функций  $\left( \{u_{Bk}^{Ne}\}_{k=1}^M, \{u_{Pk}^{Ne}\}_{k=1}^M, \{u_{Ck}^{Ne}\}_{k=1}^M \right)$  таких, что ни одному из агентов невыгодно отклоняться от стратегии  $\{u_{ik}^{Ne}\}_{k=1}^M (i \in N)$ , если этого не делают остальные.

Для нахождения равновесий Нэша в игре (3.3.21)–(3.3.23) используется метод КРС. Введём множество КРС  $i$ -го агента в момент времени  $t = k\Delta t$  и предположим, что оно содержит  $L+1$  элемент:

$$U_{ik}^{QRS} = \left\{ 0; \frac{r_{ik}}{L}; \frac{2r_{ik}}{L}; \dots; \frac{(L-1)r_{ik}}{L}; r_{ik} \right\}; i = B, P, C; k = 1, 2, \dots, M.$$

Нужно найти  $3M$  чисел  $u_{ik} \in U_{ik}^{QRS}; i = B, P, C; k = 1, 2, \dots, M.$

В этом случае для каждого агента  $i$  есть  $(K+1)^M$  сценариев:

$$QRS_i = U_{i1}^{QRS} \times \dots \times U_{iM}^{QRS}; |QRS_i| = (K+1)^M \quad i = B, P, C.$$

Так как имеются три игрока, то множество КРС игры включает  $(K+1)^{3M}$  сценариев, т.е.  $QRS = QRS_B \times QRS_P \times QRS_C; |QRS| = (K+1)^{3M}.$

В реальных системах числа  $K$  и  $M$  не слишком велики (не более 3-4), что обеспечивает практическую реализуемость данного подхода.

Предлагается следующий численный алгоритм построения множества равновесий Нэша NE в игре (3.3.21) - (3.3.23) для  $QRS$  – множества.

1. Задать все входные функции и параметры модели. Положить  $j = 1.$
2. Фиксировать  $j$ -й сценарий игры:  $(u_B^{(j)}, u_P^{(j)}, u_C^{(j)}) = \left( \left\{ u_{Bk}^{(j)} \right\}_{k=1}^M, \left\{ u_{Pk}^{(j)} \right\}_{k=1}^M, \left\{ u_{Ck}^{(j)} \right\}_{k=1}^M \right).$
3. Проверить посредством полного перебора всех допустимых стратегий первого игрока (В), выгодно ли ему отклоняться от стратегии  $u_B^{(j)},$  если остальные игроки Р и С придерживаются своих стратегий  $u_P^{(j)}$  и  $u_C^{(j)}.$
4. Проверить посредством полного перебора всех допустимых стратегий второго игрока (Р), выгодно ли ему отклоняться от стратегии  $u_P^{(j)},$  если остальные игроки В и С придерживаются своих стратегий  $u_B^{(j)}$  и  $u_C^{(j)}.$
5. Проверить посредством полного перебора всех допустимых стратегий третьего игрока (С), выгодно ли ему отклоняться от стратегии  $u_C^{(j)},$  если остальные игроки В и Р придерживаются своих стратегий  $u_B^{(j)}$  и  $u_P^{(j)}.$

6. Если на шагах 3 – 5 алгоритма одностороннее отклонение невыгодно ни одному из игроков, то найдено множество равновесий Нэша на  $QRS$ -множестве, т.е.

$$(u_B^{(j)}, u_P^{(j)}, u_C^{(j)}) \in NE(QRS).$$

7. Если не все КРС игры проверены, то вернуться на шаг 2 ( $j = j + 1$ ), иначе перейти на шаг 8.

8. Множество равновесий Нэша  $NE(QRS)$  игры (3.3.21)–(3.3.23) на множестве  $QRS$  построено.

Если  $NE(QRS) = \emptyset$ , то следует изменить  $QRS$ -множество, например, путём удвоения шага поиска и рассмотрения только нечётных полученных точек, т.е.  $L := 2L$ ;

$$Q_{ik}^{QRS} = \left\{ 0; \frac{r_{ik}}{L}, \frac{3r_{ik}}{L}, \frac{5r_{ik}}{L}, \dots; r_{ik} \right\}; i = B, P, C; k = 1, 2, \dots, M.$$

После обновления  $QRS$ -множества необходимо вернуться к шагу 2 алгоритма ( $j = 1$ ).

При полном переборе всех стратегий игроков на шагах 3–5 алгоритма уравнение динамики (3.3.21) решается численно для каждой качественно репрезентативной стратегии, например, методом конечных разностей по явной или неявной схеме.

Ниже представлены результаты расчётов по предложенному алгоритму для ряда наборов входных данных (у.е. - условные финансовые единицы).

*Пример 3.3.1.* В случае  $M = 2$ ;  $\rho = 0.1$ ;  $T = 2$  года;  $r_B = 100$ ;  $r_P = 50$ ;  $r_C = 30$ ;

$c = 20$  у.е.;  $k_B = 8$  у.е.;  $k_P = 5$  у.е.;  $k_C = 10$  у.е.;  $P_B = P_P = P_C = 0.1$ ;  $x_0 = 420$ ;  $b_B = 0.0025$ ;  $b_P = 0.006$ ;  $b_C = 0.001$ ;  $s_B = s_P = s_C = 1/3$  равновесие Нэша найдено на  $QRS$ -множестве при  $L = 2$  и имеет вид  $u_B^{NE} = u_P^{NE} = u_C^{NE} = (0, 0)$ .

*Пример 3.3.2.* В случае входных данных примера 3.3.1 и  $k_B = 0.5$  у.е.;  $k_P = 0.1$  у.е.;  $k_C = 0.05$  у.е., а также  $c = 500$  у.е. равновесие Нэша найдено на  $QRS$ -множестве при  $L = 2$ , и  $u_B^{NE} = (r_B, r_B)$ ;  $u_P^{NE} = (r_P, r_P)$ ;  $u_C^{NE} = (r_C, r_C)$ .

*Пример 3.3.3.* В случае входных данных примера 3.3.1 и  $k_B = 0.001$  у.е.;  $k_P = 0.002$  у.е.;  $k_C = 0.004$  у.е. равновесие Нэша не меняется и найдено на  $QRS$ -множестве при  $L = 4$ ,  $u_B^{NE} = (r_B, r_B)$ ;  $u_P^{NE} = (r_P, r_P)$ ;  $u_C^{NE} = (r_C, r_C)$ .

*Пример 3.3.4.* В случае входных данных примера 3.3.2 и  $k_C = 0.3$  у.е. в равновесии Нэша стратегия третьего игрока (С) меняется. Равновесие Нэша найдено при  $L = 4$  и имеет вид  $u_B^{NE} = (r_B, r_B)$ ;  $u_P^{NE} = (r_P, r_P)$ ;  $u_C^{NE} = (\frac{r_C}{2}, \frac{r_C}{2})$ .

*Пример 3.3.5.* В случае входных данных примера 3.3.4 и  $k_B = 1.3$  у.е. в равновесии Нэша стратегия первого игрока (В) меняется. Равновесие Нэша найдено при  $L = 4$  и имеет вид  $u_B^{NE} = (\frac{r_B}{2}, \frac{r_B}{2})$ ;  $u_P^{NE} = (r_P, r_P)$ ;  $u_C^{NE} = (\frac{r_C}{2}, \frac{r_C}{2})$ .

*Пример 3.3.6.* В случае входных данных примера 3.3.5 и  $k_P = 0.9$  у.е. в равновесии Нэша стратегия второго игрока (Р) меняется. Равновесие Нэша найдено при  $L = 8$  и имеет вид  $u_B^{NE} = (\frac{r_B}{2}, \frac{r_B}{2})$ ;  $u_P^{NE} = (\frac{r_P}{2}, \frac{r_P}{2})$ ;  $u_C^{NE} = (\frac{r_C}{2}, \frac{r_C}{2})$ .

Теперь рассмотрим иерархическую постановку задачи управления социальным партнёрством в системе дополнительного профессионального образования. Пусть множество игроков  $N = \{B, P, C\}$  (агенты) из предыдущей игры с функционалами выигрыша (3.3.16) и ограничениями (3.3.17) расположены на нижнем уровне управления, а верхний уровень представлен государством (супервайзер), максимизирующим суммарный функционал выигрыша (общественное благосостояние):

$$J_0(s(\cdot), u(\cdot)) = \sum_{i \in N} \int_0^T e^{-\rho t} [g_i(r_i - u_i(t)) + s_i(t)c(x(t))] dt + e^{-\rho T} s_i(T)c(x(T)) \rightarrow \max, \quad (3.3.24)$$

Управляющие переменные супервайзера - это доли агентов в распределении общественной полезности, порождаемой профессиональной подготовкой студентов. Ограничения на эти управления имеют вид

$$s(\cdot) \in S = \{s(t) : s_i(t) \geq 0, \sum_{i \in N} s_i(t) = 1, t \in [0, T]\}. \quad (3.3.25)$$

Тогда между супервайзером и агентами возникает дифференциальная игра Штакельберга со следующей информационной структурой.

1. Супервайзер выбирает программную стратегию (3.3.25) и сообщает её агентам из множества  $N = \{B, P, C\}$ .
2. При заданной стратегии супервайзера агенты разыгрывают дифференциальную игру (3.3.16) - (3.3.18), (3.3.24), (3.3.25), решение которой есть множество параметрических равновесий Нэша  $NE(s(\cdot))$ .
3. Фактически супервайзер выбирает свою стратегию (3.3.25) так, чтобы максимизировать выигрыш (3.3.24) на множестве  $NE(s(\cdot))$ .

Поскольку все стратегии постоянны на одних и тех же интервалах времени, то стратегия супервайзера есть сеточная функция  $\left( \left\{ s_{Bk} \right\}_{k=1}^M, \left\{ s_{Pk} \right\}_{k=1}^M, \left\{ s_{Ck} \right\}_{k=1}^M \right)$ , а его функционал выигрыша (3.3.24) становится функцией:

$$J_0 \left( \left( u_{ik} \right)_{\substack{k=1 \\ i \in N}}^M, \left( s_{ik} \right)_{\substack{k=1 \\ i \in N}}^M \right) = \sum_{i \in N} \left( \sum_{k=1}^M (k_i (r_{ik} - u_{ik})^{p_i} (e^{-\rho k} - e^{-\rho(k-1)}) / \rho + \int_0^T e^{-\rho t} s_i(t) c x(t) dt \right) + \sum_{i \in N} e^{-\rho T} c x(T) \rightarrow \max. \quad (3.3.26)$$

Ограничения на управления супервайзера принимают вид

$$s_{ik}(t) \geq 0, \quad \sum_{i \in N} s_{ik}(t) = 1; \quad k = 1, 2, \dots, M; \quad i \in N. \quad (3.3.27)$$

Для построения множества равновесий Штакельберга в игре (3.3.21) – (3.3.23), (3.3.26) – (3.3.27) используется метод КРС.

Аналогично задаче построения равновесий Нэша, вводится  $QRS$ -множество  $i$ -го агента в момент  $t = k\Delta t$ . Пусть оно содержит  $L+1$  элемент:

$$U_{ik}^{QRS} = \left\{ 0; \frac{r_{ik}}{L}; \frac{2r_{ik}}{L}; \dots; \frac{(L-1)r_{ik}}{L}; r_{ik} \right\}; \quad i = B, P, C; \quad k = 1, 2, \dots, M.$$



Соответственно, введём  $QRS$ -множество управлений супервайзера по отношению к каждому агенту, включающее в момент времени  $t = k\Delta t$  не более, чем  $L+1$  элемент:

$$S_{ik}^{QRS} = \left\{ 0; \frac{1 - \sum_{j=1}^{i-1} s_{ik}}{L}; \frac{2(1 - \sum_{j=1}^{i-1} s_{ik})}{L}; \dots; \frac{(L-1)(1 - \sum_{j=1}^{i-1} s_{ik})}{L}; (1 - \sum_{j=1}^{i-1} s_{ik}) \right\}; i = B, P, C; k = 1, 2, \dots, M.$$

В этом случае каждый агент  $i$  имеет  $(K+1)^M$  сценариев

$$QRS_i = U_{i1}^{QRS} \times \dots \times U_{iM}^{QRS}; |QRS_i| = (K+1)^M \quad i = B, P, C,$$

а супервайзер не более, чем  $(K+1)^{3M}$  сценариев:

$$QRS_0 = S_{B1}^{QRS} \times \dots \times S_{BM}^{QRS} \times S_{P1}^{QRS} \times \dots \times S_{PM}^{QRS} \times S_{C1}^{QRS} \times \dots \times S_{CM}^{QRS}; |QRS_0| = (K+1)^{3M}.$$

$QRS$ -множество игры содержит  $(K+1)^{6M}$  сценариев, а именно  $QRS = QRS_B \times QRS_P \times QRS_S \times QRS_0; |QRS| = (K+1)^{6M}$ .

В реальных системах числа  $K$  и  $M$  можно выбрать достаточно малыми (не более 3-4), что обеспечивает практическую реализуемость подхода. Численный алгоритм построения равновесий Штакельберга для игры (3.3.21)– (3.3.23), (3.3.26) – (3.3.27) тот же, что и в параграфе 2.3. Равновесия Штакельберга ищутся методом КРС.

Аналогично параграфу 2.3 построим  $QRS$ -множество для  $M = K = 2$ . В любой момент времени имеются три стратегии каждого агента и не более, чем три стратегии супервайзера по отношению к каждому агенту:

$$U_{ik}^{QRS} = \left\{ 0; \frac{r_{ik}}{2}; r_{ik} \right\}; S_{ik}^{QRS} = \left\{ 0; \frac{1 - \sum_{j=1}^{i-1} s_{ik}}{2}; (1 - \sum_{j=1}^{i-1} s_{ik}) \right\}; i = B, P, C; k = 1, 2, \dots, M.$$

Тогда для каждого агента ( $i = B, P, C$ ) получаем девять качественно репрезентативных стратегий:

$$QRS_i = U_{i1}^{QRS} \times U_{i2}^{QRS} = \left\{ (u_{i1}, u_{i2})_{k=1}^6 \right\} = \left\{ (0, 0); (0, \frac{r_{i2}}{2}); (0, r_{i2}); (\frac{r_{i1}}{2}, 0); (\frac{r_{i1}}{2}, \frac{r_{i2}}{2}); (\frac{r_{i1}}{2}, r_{i2}); (r_{i1}, 0); (r_{i1}, \frac{r_{i2}}{2}); (r_{i1}, r_{i2}) \right\}.$$

Супервайзер в каждый момент времени ( $k = 1, 2$ ) имеет шесть качественно репрезентативных стратегий:

$$S_k^{QRS} = S_{Bk}^{QRS} \times S_{Pk}^{QRS} \times S_{Ck}^{QRS} = \left\{ (s_{Bk}, s_{Pk}, s_{Ck})_{j=1}^6 \right\} = \left\{ (0,0,1); (0,1,0); (1,0,0); (0, \frac{1}{2}, \frac{1}{2}); (\frac{1}{2}, 0, \frac{1}{2}); (\frac{1}{2}, \frac{1}{2}, 0) \right\}.$$

Таким образом, всего в игре есть 26244 различных сценария:

$$QRS = S_1^{QRS} \times S_2^{QRS} \times QRS_B \times QRS_P \times QRS_C; |QRS| = 6^2 9^3 = 26244.$$

Для всех 26244 сценариев игры первое условие в определении  $QRS$ -множества проверялось численно. Это заняло около 320 секунд на Pentium (R) DualCoreCPU E5400. Сначала вычислялось значение  $J_0^{\max}$  глобального максимума функционала выигрыша (3.3.24) (функции выигрыша (3.3.26)).

Максимум  $J_0^{\max}$  находился по управляющим переменным супервайзера и всех трёх агентов (для (3.3.24) по шести функциям  $s_B(t), s_P(t), s_C(t)$  и  $u_B(t), u_P(t), u_C(t)$ , а для (3.3.26) по шести сеточным функциям

$$u_B = \{u_{Bk}\}_{k=1}^M; u_P = \{u_{Pk}\}_{k=1}^M; u_C = \{u_{Ck}\}_{k=1}^M \text{ and } s_B = \{s_{Bk}\}_{k=1}^M; s_P = \{s_{Pk}\}_{k=1}^M; s_C = \{s_{Ck}\}_{k=1}^M.$$

Оказалось, что для входных данных из примера 3.3.1  $J_0^{\max} = 49895$  у.е.

*Пример 3.3.7.* Часть результатов для входных данных из примера 3.3.1 приведена в таблице 3.3.2.

Таблица 3.3.2

Стратегии ( $s_{B1}, s_{P1}, s_{C1}, s_{B2}, s_{P2}, s_{C2},$ $u_{B1}, u_{P1}, u_{C1}, u_{B2}, u_{P2}, u_{C2}$ ) <sup>(j)</sup>	$J_0^{(j)}((s_{B1}, s_{P1}, s_{C1}, s_{B2}, s_{P2}, s_{C2},$ $u_{B1}, u_{P1}, u_{C1}, u_{B2}, u_{P2}, u_{C2})^{(j)})$ (у.е.)	$\min_{\substack{1 \leq k \leq 26244 \\ k \neq j}}  J_0^{(k)} - J_0^{(j)} $ $/ J_0^{\max}$
(0,0,1,0,0,1,0,0,0,0,0,0)	49762	0.0012
(0,1,0,0,0,1,0,0,0,0,0,0)	49804	0.0015
(1,0,0,0,0,1,0,0,0,0,0,0)	49705	0.0017
(1,0,0,0,0,1,50,50,30,50,25;30)	37096	0.02
( $\frac{1}{2}, \frac{1}{2}, 0, 0, 0, 1, 50, 50, 15, 50, 25; 30$ )	41860	0.023
( $\frac{1}{2}, \frac{1}{2}, 0, 0, \frac{1}{2}, \frac{1}{2}, 50, 50, 15, 0, 25; 30$ )	42076	0.025

$(\frac{1}{2}, \frac{1}{2}, 0, 0, 0, 1, 100, 50, 30, 100, 50, 30)$	26282	0.003
$(\frac{1}{2}, \frac{1}{2}, 0, 0, 0, 1, 50, 50, 30, 50, 25, 0)$	42203	0.016
$(\frac{1}{2}, \frac{1}{2}, 0, 0, 0, 1, 100, 0, 30, 50, 50, 15)$	37588	0.04
$(\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 50, 50, 30, 50, 0, 0)$	42268	0.009
$(\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 100, 50, 15, 50, 25, 30)$	37713	0.011
$(\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 50, 50, 30, 50, 50, 30)$	34681	0.005
$(\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 0, 50, 30, 0, 25, 30)$	37433	0.012
$(\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 50, 0, 30, 50, 25, 0)$	44435	0
$(\frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}, 0, \frac{1}{2}, 50, 50, 30, 0, 50, 30)$	34906	0.004
$(\frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}, 0, \frac{1}{2}, 50, 0, 0, 50, 50, 30)$	42017	0.016
$(\frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}, 0, \frac{1}{2}, 100, 50, 30, 50, 50, 30)$	30526	0.05
$(\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 50, 50, 30, 50, 50, 15)$	39456	0.007
$(\frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}, 0, \frac{1}{2}, 50, 50, 15, 50, 25, 30)$	41923	0.03
$(\frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}, 0, \frac{1}{2}, 50, 50, 15, 50, 50, 30)$	39515	0.008
$(\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 50, 50, 15, 50, 0, 30)$	41803	0.024
$(\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 50, 50, 15, 50, 0, 0)$	47030	0.012
$(\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 0, 0, 15, 50, 25, 30)$	44177	0.0087
$(\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 50, 50, 0, 0, 25, 30)$	42397	0.0078
$(\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 0, 50, 15, 0, 25, 0)$	47297	0.009

Часть результатов проверки второго условия из определения *QRS*-множества для данных из примера 3.2.1 при  $M = K = 2$  дана в таблице 3.3.3.

Таблица 3.3.3

Стратегии ( $s_{B1}, s_{P1}, s_{C1}, s_{B2}, s_{P2},$ $s_{C2}, u_{B1}, u_{P1}, u_{C1}, u_{B2},$ $u_{P2}, u_{C2}$ ) <sup>(j)</sup> $\notin QRS$	$J_0^{(j)}$ (y.e.)	“Близкая” стратегия ( $s_{B1}, s_{P1}, s_{C1}, s_{B2}, s_{P2},$ $s_{C2}, u_{B1}, u_{P1}, u_{C1}, u_{B2},$ $u_{P2}, u_{C2}$ ) <sup>(k)</sup> $\in QRS$	$J_0^{(k)}$ (y.e.)	$ J_0^{(k)} - J_0^{(j)} $ $/ J_0^{\min}$
(0.2,0.2,0.6,0.2,0.2,0.6, 0,0,0,0,0,0)	49762	(0,0,1,0,0,1,0,0,0,0,0)	49762	0
(0.2,0.2,0.6,0.2,0.2,0.6, 30,30,10,10,10,10)	49472	(0,0,1,0,0,1,50,25,15,0,0,15)	49466	0.0001
(0.2,0.2,0.6,0.2,0.2,0.6, 0,0,10,10,0,0)	49565	(0,0,1,0,0,1,0,0,15,0,0,0)	49553	0.0002
(0.2,0.2,0.6,0.2,0.2,0.6, 0,0,0,0,10,10)	49590	(0,0,1,0,0,1,0,0,0,0,0,15)	49615	0.0005
(0.2,0.2,0.6,0.2,0.2,0.6, 70,40,0,0,0,0)	49618	(0,0,1,0,0,1,0,0,0,0,0,15)	49615	0.00006
(0.2,0.2,0.6,0.2,0.2,0.6, 0,0,10,40,0,0)	49477	(0,0,1,0,0,1,50,25,15,0,0,15)	49466	0.0002
(0.2,0.2,0.6,0.2,0.2,0.6, 0,40,0,40,0,0)	49716	(1,0,0,0,0,1,0,0,0,0,0,0)	49705	0.0002
(0.2,0.2,0.6,0.2,0.2,0.6, 70,0,10,0,10,0)	49421	(0,0,1,0,0,1,0,0,15,0,0,15)	49402	0.0004
(0.2,0.2,0.6,0.2,0.2,0.6, 0,20,0,20,0,10)	49646	(0,0,1,0,0,1,0,25,0,0,0,15)	49652	0.0001
(0.2,0.2,0.6,0.2,0.2,0.6, 0,10,10,0,20,0)	49658	(0,0,1,0,0,1,0,25,0,0,0,15)	49652	0.0001
(0.3,0.3,0.4,0.2,0.2,0.6, 0,0,0,0,0,0)	49762	(0,0,1,0,0,1,0,0,0,0,0,0)	49762	0
(0.2,0.2,0.6,0.3,0.3,0.4, 0,0,0,0,0,0)	49762	(0,0,1,0,0,1,0,0,0,0,0,0)	49762	0
(0.2,0.2,0.6,0.3,0.3,0.4, 70,20,0,0,0,0)	49684	(1,0,0,0,0,1,0,0,0,0,0,0)	49705	0.0004
(0.3,0.3,0.4,0.2,0.2,0.6, 70,20,10,0,10,0)	49498	(0,0,1,0,0,1,50,25,15,0,0,0)	49517	0.0004
(0.2,0.2,0.6,0.3,0.3,0.4, 0,0,30,40,10,20)	44161	(0.5,0,0.5,0,0,1,0,0,15,50, 25;30)	44177	0.0003
(0.2,0.2,0.6,0.2,0.2,0.6, 60,40,30,20,10,0)	44553	(0,0,1,0,0,1,50,50,30,0,0;0)	44556	0.00006
(0.2,0.2,0.6,0.2,0.2,0.6, 40,10,20,30,0,0)	49230	(0,0,1,0,0,1,50,0,15,50,0;0)	49234	0.00008

(0.2,0.2,0.6,0.2,0.2,0.6, 0,20,0,40,20,20)	49238	(0,0,1,0,0,1,50,0,15,50,0;0)	49234	0.00008
(0.2,0.2,0.6,0.2,0.2,0.6, 10,10,10,10,10,10)	49453	(0,0,1,0,0,1,50,25,15,0,0,15)	49466	0.0002
(0.2,0.2,0.6,0.2,0.2,0.6, 20,20,20,20,20,20)	48807	(0,0,1,0,0,1,0,30,15,0,30,15)	48810	0.00006
(0.2,0.2,0.6,0.2,0.2,0.6, 30,30,30,30,30,30)	39531	(0,0,1,0,0,1,50,25,30,50, 25,30)	39532	0.00002
(0.3,0.3,0.4,0.2,0.2,0.6, 10,10,10,10,10,10)	49453	(0,0,1,0,0,1,0,0,15,0,0,15)	49402	0.00008
(0.2,0.2,0.6,0.3,0.3,0.4, 20,20,20,20,20,20)	48807	(0,0,1,0,0,1,0,30,15,0,30,15)	48810	0.00006
(0.3,0.3,0.4,0.3,0.3,0.4, 30,30,30,30,30,30)	39531	(0,0,1,0,0,1,50,25,30,50, 25,30)	39532	0.00002
(0.4,0.4,0.2,0.2,0.2,0.6, 0,0,0,0,0,0)	49762	(0,0,1,0,0,1,0,0,0,0,0,0)	49762	0

Анализ данных таблиц 3.3.2 и 3.3.3 даёт  $\Delta = 0.0012$ . Поэтому построенное множество сценариев можно брать в качестве *QRS*-множества для входных данных из примера 3.3.1 при  $M = K = 2$ .

Итак, метод имитационного моделирования состоит в проведении вычислительных экспериментов с программной реализацией модели УДС, позволяющих оценить влияние различных управляющих воздействий на поведение системы. Поскольку полный перебор сценариев невозможен, то необходимо предложить и обосновать некоторый способ частичного перебора. Такой способ предлагает метод КРС, основанный на идее о том, что для качественной оценки последствий управляющих воздействий достаточно рассмотреть очень небольшое число значений каждой управляющей переменной, отражающих характерные принципиально различные варианты воздействия на управляемую динамическую систему. Дано точное определение множества КРС для случая иерархической разностной игры. Метод КРС эвристический, выбирать сценарии следует исходя из специфики конкретной задачи таким образом, чтобы выполнялось приведённое определение.

## ГЛАВА 4. ПРИКЛАДНЫЕ АСПЕКТЫ УПРАВЛЕНИЯ УСТОЙЧИВЫМ РАЗВИТИЕМ АКТИВНЫХ СИСТЕМ

Четвёртая глава монографии посвящена приложениям теории и методов управления устойчивым развитием активных систем в различных предметных областях. В первом параграфе рассматриваются модели влияния и управления на сетях. Известные модели анализа и прогноза групповых мнений формулируются применительно к маркетингу, приводятся авторские алгоритмы исследования моделей. Даются постановки задач оптимального и конфликтного управления мнениями целевой аудитории на сетях с учётом требований устойчивого развития, намечаются пути их решения на основе метода качественно репрезентативных сценариев имитационного моделирования.

Во втором параграфе представлены модели управления устойчивым развитием учреждений высшего образования. Дан анализ университета как активной системы, предложены постановки ряда задач управления устойчивым развитием учреждений высшего образования. В частности, подробно исследованы модели борьбы с экономической коррупцией, актуальные для данной предметной области.

В третьем параграфе описан проект ИАС ВР РО - информационно-аналитической системы управления водными ресурсами (на примере Ростовской области). Ключевая особенность авторского замысла заключается в решении посредством ИАС задач управления устойчивым развитием с использованием теории управления организационными системами и реализации в составе ИАС ВР РО ряда механизмов управления. Кроме того, ИАС ВР РО обеспечивает ввод, хранение, обработку и использование данных мониторинга водных ресурсов. Разработка и реализация ИАС представляется наиболее перспективным направлением практического решения задач управления устойчивым развитием активных систем.

#### 4.1. Сетевые модели управления устойчивым развитием активных систем

Материал этого параграфа основан на работах (Агиева 2018а,б,в; Агиева 2019а,б; Агиева и др. 2019).

Математическое моделирование процессов взаимного влияния в социальных сетях на базе цепей Маркова было начато в работах Дж. Френча (French 1956) и Ф. Харари (Harary 1959). Наиболее известная модель принадлежит М. де Грооту (De Groot 1974). Идея подхода состоит в том, что начальные мнения агентов со временем меняются под влиянием других членов социальной группы по некоторому правилу; в частности, в линейной модели все члены группы приходят к устойчивым финальным мнениям, зависящим от групповой структуры. Это направление исследований развито, например, в (Friedkin and Johnsen 1990) и (Jackson 2008), а также распространено на другие предметные области. Так, в (Morris 2000) изучаются модели "заражения", в (Nekovee et al. 2007) - модели распространения слухов в обществе.

Наряду с моделями влияния, ещё больший интерес представляют модели управления в социальных группах с заданной структурой взаимодействия (Губанов и др. 2010). В свою очередь, здесь естественно выделять модели оптимального (единственный субъект управления) и конфликтного управления (несколько взаимодействующих субъектов, имеющих различные интересы). Математическая формализация моделей второй группы приводит к игровым постановкам на сетях (Novikov 2014; Roughgarden 2005).

Базовая модель влияния в социальной сети представляет собой взвешенный ориентированный граф, вершины которого соответствуют членам социальной группы, а дуги описывают их взаимное влияние. Каждой вершине приписывается вещественное значение (мнение члена группы), которое может меняться со временем, а каждой дуге - вещественное число

(вес), характеризующее степень влияния одного члена группы на другого (или, что то же самое, степень доверия второго к первому).

Таким образом, в модели  $X = \{x_1, \dots, x_n\}$  - множество базовых агентов;  $x_i \rightarrow u_i(t)$ ,  $i = 1, \dots, n$  - мнение базового агента;  $u^0 = (u_1^0, \dots, u_n^0)$  - вектор начальных мнений базовых агентов;  $a_{ij}$  - коэффициент влияния  $i$ -го базового агента на мнение  $j$ -го базового агента;  $A = \| a_{ij} \|$  - матрица влияний (задаёт множество дуг социально-сетевой модели). Динамика мнений задается правилом

$$u_j(t+1) = \sum_{i=1}^n a_{ij} u_i(t), \quad j = 1, \dots, m; \quad \text{или в векторной форме}$$

$$u(t+1) = A^T u(t), \quad u(0) = u^0, \quad t = 0, 1, \dots, T-1. \quad (4.1.1)$$

Показано, что все члены каждой  $i$ -й сильной подгруппы (невыврожденной сильной компоненты сети, отвечающей эргодическому множеству) приходят к единому финальному мнению, определяемому формулой

$$u_i^\infty = \sum_{k=1}^{n_i} w_k^i u_k^{i0}, \quad (4.1.2)$$

где  $w_k^i$  - компонента стационарного вектора для цепи Маркова с переходной матрицей  $A^T$ ,  $n_i$  - число членов  $i$ -й сильной подгруппы. Финальные мнения остальных агентов, не входящих в сильные подгруппы ("спутников"), вычисляются как

$$u_j^\infty = \sum_{i=1}^r b_{ji} u_i^\infty, \quad (4.1.3)$$

где  $b_{ji}$  - вероятность перехода агента  $j$  в сильную подгруппу  $i$  как эргодическое множество,  $r$  - общее число сильных подгрупп.

Кроме того, для заданной сети можно вычислять различные количественные показатели: центральности, престижа и др. (Губанов и др. 2010).



Наиболее интересный с точки зрения приложений вопрос возникает, когда одно или несколько лиц, принимающих решения, не устраивают финальные мнения (4.1.2)-(4.1.3). Для формализации такой ситуации базовая модель расширяется введением множества агентов воздействия  $Y = \{y_1, \dots, y_m\}$ , способных влиять на базовых агентов. Тогда множество вершин сети есть  $X \cup Y$ . Агенты воздействия могут изменять начальные мнения базовых агентов или коэффициенты их взаимодействия, соответственно

$$u^0 := u^0 + w^0; \quad A := A + B.$$

Обозначим  $v(t) = (v_1(t), \dots, v_n(t))$  - вектор управляющих воздействий на всех базовых агентов в момент времени  $t$ , удовлетворяющий ограничениям

$$v(t) \in V. \quad (4.1.4)$$

Если ситуация рассматривается с позиции единственного агента воздействия, то возникает задача оптимального управления

$$J(v(\cdot)) = \int_0^T e^{-\rho t} g(v(t), u(t)) dt \rightarrow \max \quad (4.1.5)$$

с ограничениями (4.1.1) и (4.1.4), где  $g$  - текущая целевая функция агента воздействия,  $T$  - период рассмотрения (конечный или бесконечный),  $\rho$  - коэффициент дисконтирования. Критерий оптимальности агента воздействия может иметь дискретную форму

$$\bar{J}(v(\cdot)) = \sum_{t=0}^T e^{-\rho t} g(v(t), u(t)) \rightarrow \max. \quad (4.1.6)$$

В более общем случае, когда агентов воздействия несколько ( $l=1, \dots, m$ ), каждый из них имеет критерий оптимальности вида (4.1.5) или (4.1.6) и выбирает стратегию  $v_l(t) \in V_l$  воздействия на некоторое подмножество базовых агентов. Это порождает дифференциальную игру в нормальной форме вида

$$J_l(v(\cdot)) = \int_0^T e^{-\rho t} g_l(v(t), u(t)) dt \rightarrow \max, \quad v_l(t) \in V_l, \quad l = 1, \dots, m,$$

с динамикой состояния (4.1.1), на базе которой стандартным образом можно построить кооперативную игру (Петросян и др. 2012). В (Губанов и др. 2010) приводятся некоторые частные постановки такого рода задач управления.

Если на множестве агентов воздействия задана некоторая иерархия, то возникают иерархические дифференциальные игры с наличием или отсутствием обратной связи по управлению, исследование которых представляет особый интерес. В таблице 4.1.1 дана интерпретация элементов описанных выше моделей применительно к маркетингу.

Таблица 4.1.1

Интерпретация элементов моделей влияния и управления на социальных сетях в маркетинге

Элемент модели	Математический смысл	Маркетинговая интерпретация
Базовый агент	Вершина сети	Сегмент целевой аудитории
Агент воздействия	Вершина сети	Участники рынка (фирмы), рекламные агентства, средства массовой информации и пр.
Мнение базового агента	Вещественное число, приписываемое каждой вершине, представляющей базового агента (меняется со временем)	Месячная (годовая) сумма расходов, которую агент готов платить за покупку товаров данной фирмы, или число посещений магазинов (ресторанов, гостиниц и т.д.) этой фирмы, или некая балльная оценка продуктов (услуг) фирмы
Доверие (влияние)	Наличие дуги от одной вершины к другой: начальная вершина доверяет конечной, конечная влияет на начальную	Сарафанное радио, другие информационные взаимодействия между агентами
Коэффициент доверия одного базового агента	Вещественное число, приписываемое каждой дуге сети	Количественная характеристика доверия

другому		
Финальное мнение	Предельное значение мнения, формируемое на бесконечном периоде времени	Устойчивое итоговое значение мнения по истечению длительного периода времени
Сильная подгруппа	Невырожденная сильная компонента сети, отвечающая эргодическому множеству	Определяет свои финальные мнения и зависящие от них мнения остальных агентов
Спутник	Подмножество остальных вершин	Финальные мнения полностью определяются сильными подгруппами
Воздействие на начальные мнения	Аддитивная добавка к вектору начальных мнений	Комплекс маркетинговых мероприятий
Воздействие на коэффициенты влияния	Аддитивная добавка к матрице влияний (возможны более сложные варианты)	Комплекс маркетинговых мероприятий
Цель управления	Область в пространстве состояний сети	Диапазон желаемых значений мнений

Итак, с помощью моделей влияния и управления в социальных группах можно решать следующие задачи маркетинга (таблица 4.1.2).

Таблица 4.1.2

#### Модели влияния и управления на сетях и задачи маркетинга

Модельные задачи	Задачи маркетинга
Анализ сети	1. Сегментация целевой аудитории, выделение сильных подгрупп, определяющих внутренние общие финальные мнения и индивидуальные финальные мнения остальных агентов как линейную комбинацию финальных мнений сильных подгрупп. 2. Расчет показателей центральности, престижа и иных характеристик целевой аудитории.
Прогноз на сети	Определение финальных мнений всех агентов при отсутствии целенаправленного воздействия на них
Оптимальное управление на сети	Выбор оптимального маркетингового воздействия на целевую аудиторию с позиций одного предпринимателя
Динамические игры	Выбор компромиссного воздействия на целевую

на сетях	аудиторию с учетом конкуренции и/или кооперации между предпринимателями
----------	---

К числу общих количественных характеристик сетевой структуры относятся простейшие метрики (степень и плотность) и метрики центральности (степень центральности, центральность по близости и по посредничеству) (Кузнецов 2017; Лифинцев 2013). Степень – количество связей (рёбер) данной вершины. Плотность измеряет степень, в которой узлы сети соединяются друг с другом. Это отношение числа рёбер графа к максимально возможному:

$$d = \frac{2M}{N(N-1)},$$

где M - число ребер в сети, N - число узлов в сети. С ростом плотности графа увеличивается поток информации между узлами, что повышает их влияние друг на друга.

Метрики центральности — это количественная оценка какого-либо лица (вершины, узла) в социальной сети. Мера центральности описывает положение конкретного узла по сравнению с другими узлами. Степень центральности определяется как число прямых связей или соединений с узлом. Узел с более высоким значением степени центральности часто рассматривается как центр и активный субъект в сети. Степень центральности  $i$ -узла вычисляется по следующей формуле:

$$D_i = \frac{\sum_{j \in G} A_{ij}}{N-1},$$

где  $A_{ij}$  – элемент матрицы смежности A, (N-1) - коэффициент нормализации.

Центральность по близости показывает, насколько быстро распространяется информация в сети от одного узла к другим, то есть насколько близок рассматриваемый узел ко всем остальным узлам сети.

Высоким уровнем центральности по близости обладают те узлы, которые имеют множество связей, при этом их должно быть достаточно и у его соседей. Следовательно, такие узлы получают возможность доступа к большому количеству участников сети, и могут влиять на них.

Эта величина определяется как обратная к среднему значению длины всех кратчайших путей между узлом и всеми остальными узлами в сети:

$$C_i = \frac{N-1}{\sum_{j \in G} d_{ij}},$$

где  $\sum_{j \in G} d_{ij}$  - сумма длин кратчайших путей от вершины  $i$  до вершины  $j$ .

Промежуточность позволяет найти те узлы, которые служат мостами для соединения с другими группами или сообществами в сети. Она может быть определена как число кратчайших путей между любой парой узлов, которые проходят через данный узел:

$$B_i = \frac{\sum_{j,k \in G} \frac{g_{jk}(i)}{g_{jk}}}{(N-1)(N-2)},$$

где  $g_{jk}$  - число кратчайших путей из вершины  $j$  в вершину  $k$ ;  $g_{jk}(i)$  - число кратчайших путей между  $j$  и  $k$ , проходящих через  $i$ -й узел. Вершина с высоким значением промежуточности рассматривается как мощный узел с большим потенциалом влияния.

Таким образом, задачи анализа социальных сетей заключаются в выделении сильных подгрупп и спутников, а также вычислении количественных характеристик сети. Решение первой задачи (сегментация целевой аудитории) позволяет установить тех агентов, на которых целесообразно оказывать маркетинговое воздействие (членов сильных подгрупп), поскольку только от их мнений будут зависеть финальные мнения всех членов целевой аудитории. Решение второй задачи (количественная

характеристика целевой аудитории) позволяет дать дополнительную классификацию её членов по степени влияния на других.

Рассмотрим подробнее задачу прогноза (Робертс 1986). Формализуем взаимодействие в группе орграфом влияний

$$D = (Y, A), \quad Y = \{y_1, \dots, y_n\}, \quad A = \| a_{ij} \|_{i,j=1}^n.$$

Сделаем следующие предположения:

- 1) D не меняется в процессе принятия решений;
- 2) решения принимаются в дискретные моменты времени  $t=0,1,2,\dots$ ;
- 3) мнения членов группы меняются согласно правилу

$$x_j(t+1) = \sum_{i=1}^n a_{ij} x_i(t), \quad j = 1, \dots, n.$$

Тогда представляют интерес ответы на следующие вопросы:

(i) существует ли для каждого члена группы устойчивое финальное мнение  $x_i^\infty = \lim_{t \rightarrow \infty} x_i(t)$ ?

(ii) придёт ли вся группа к устойчивому общему финальному мнению  $x^\infty = x_i^\infty, i = 1, \dots, n$ ?

Взвешенный орграф  $C(D)$  называется обращением D, если  $C(D)$  содержит дугу  $(u_i, u_j)$  с весом  $a_{ji}$  тогда и только тогда, когда D содержит дугу  $(u_j, u_i)$  с весом  $a_{ij}$ . Если D есть орграф влияний некоторой группы, то его обращение  $C(D)$  есть стохастический орграф, который можно рассматривать как переходный орграф некоторой цепи Маркова.

Естественно предположить, что  $\exists i a_{ii} > 0$  (т.е. хотя бы один член группы учитывает собственное мнение). Обозначим

$$x(t) = \begin{pmatrix} x_1(t) \\ \dots \\ x_n(t) \end{pmatrix} - \text{вектор мнений членов группы в момент } t, \quad P = \| p_{ij} \| = \| a_{ji} \| = A^T.$$

Тогда в силу (2.2) имеем  $x(t+1) = Px(t)$ , откуда

$$x(t) = P^t x(0), \quad t = 0, 1, 2, \dots \quad (4.1.7)$$

В частном случае справедливо следующее утверждение (Робертс 1986).

Если оргграф влияний  $D$  сильно связный и содержит хотя бы одну петлю, то члены группы достигают общего финального мнения

$$x^\infty = \sum_{i=1}^n w_i x_i(0), \quad (4.1.8)$$

где  $w_i$ - компоненты стационарного вектора  $w$  регулярной цепи Маркова с переходной матрицей  $P$ .

Теперь рассмотрим общий случай, когда оргграф влияний не сильно связный (т.е. не все члены группы влияют друг на друга, что более естественно для больших групп). Будем называть множество вершин  $\{y_1, \dots, y_k\}$  оргграфа влияний  $D=(Y,A)$  *сильной подгруппой*, если это сильная компонента  $D$ , входящая в вершинную базу  $V^*$  конденсации  $D^*$ . В этом случае в обращении  $C(D)$  множество  $V^*$  образует контрабазу конденсации  $C^*(D)$ , поэтому сильные подгруппы представляют собой эргодические множества цепи Маркова, определяемой переходным оргграфом  $C(D)$ , или переходной матрицей  $P=A^T$ . Составим список сильных подгрупп  $S_1, \dots, S_r$ .

Нумеруем вершины по порядку, для определенности следующим образом:

$$S_1 = \{y_1, \dots, y_{1n_1}\}, S_2 = \{y_{2,1}, \dots, y_{2n_2}\}, \dots, S_r = \{y_{r,1}, \dots, y_{rn_r}\},$$

а вершины из переходного множества

$$T = \{y_{r+1,1}, \dots, y_{r+1,n_{r+1}}\}.$$

Тогда все состояния цепи Маркова описываются как

$$U = \{y_{ij} : i = 1, \dots, r+1; j = 1, \dots, n_i\}.$$

Переходную матрицу для  $C(D)$  можно представить в канонической форме:

$$P_c = \begin{array}{c} S_1 \\ S_2 \\ \dots \\ S_r \\ T \end{array} \left\| \begin{array}{cccc|c} S_1 & S_2 & \dots & S_r & T \\ \mathbf{P}_1 & & & & \\ & \mathbf{P}_2 & & & \\ & & \dots & & \mathbf{0} \\ & & & \mathbf{P}_r & \\ \hline & & & \mathbf{R} & \mathbf{Q} \end{array} \right\|,$$

где  $P_i$  - подматрица для сильной подгруппы  $S_i$ ,  $i=1, \dots, r$ .

Если  $U$  есть эргодическое множество, то возникает рассмотренный выше частный случай. В общем случае можно построить поглощающую цепь Маркова, трактуя каждое эргодическое множество  $S_i$  как обобщенное поглощающее состояние. Если первоначально цепь Маркова находится в переходном состоянии  $y_{r+1,j}$ , то вероятность её перехода в эргодическое множество  $S_i$  равна значению элемента  $b_{ji} \in B = NR$ .

Если орграф  $C(D)$  получен из орграфа влияний, у которого в каждом эргодическом множестве есть петли, то каждое такое эргодическое множество  $S_i$  определяет регулярную цепь Маркова. Поэтому вероятность оказаться в заданном состоянии  $y_{ik}$  множества  $S_i$  стремится к пределу, равному  $w_k^{(i)}$ . Тогда совокупность предельных вероятностей определяется вектором  $w^{(i)} = (w_1^{(i)}, \dots, w_n^{(i)})$  - стационарным вектором матрицы  $P_i$  цепи Маркова, соответствующей множеству  $S_i$ . Таким образом, если исходная цепь Маркова в начальный момент времени находилась в некотором переходном состоянии  $y_{r+1,j}$ , то со временем она окажется в состоянии  $y_{ik}$  эргодического множества  $S_i$  с вероятностью  $b_{ji} w_k^{(i)}$ .

Можно показать, что в пределе при  $t \rightarrow \infty$  матрица  $P^t$  стремится к матрице  $W$  следующего вида:

$$W = \begin{array}{c} S_1 \\ S_2 \\ \dots \\ S_r \\ \hline u_{r+1,1} \\ \dots \\ u_{r+1, nr+1} \end{array} \left\| \begin{array}{ccc|ccc} S_1 & S_2 & S_r & u_{r+1,1} & \dots & u_{r+1, nr+1} \\ \mathbf{W}_1 & & \mathbf{0} & & & \\ & \mathbf{W}_2 & & & & \mathbf{0} \\ & & \dots & & & \\ \mathbf{0} & & \mathbf{W}_r & & & \\ \hline \mathbf{b}_{11} \mathbf{w}^{(1)} & \dots & \mathbf{b}_{1r} \mathbf{w}^{(r)} & & & \\ & & \dots & & & \mathbf{0} \\ \mathbf{b}_{nr+1,1} \mathbf{w}^{(1)} & \dots & \mathbf{b}_{nr+1,1} \mathbf{w}^{(r)} & & & \end{array} \right\|$$

где  $W^{(i)}$  - матрица, содержащая  $n_i$  строк  $w^{(i)}$ ,  $i=1, \dots, r$ .



Поскольку  $x(t) = P^t x(0)$ , то  $\lim_{t \rightarrow \infty} x(t) = Wx(0)$ . Пусть  $x_{ij}(t)$  есть мнение  $y_{ij}$  в момент  $t$ . Тогда следует различать два варианта.

Если вершина  $y_{ij}$  входит в  $S_i$ , то

$$x_{ij}^{\infty} = \sum_{k=1}^{n_i} w_k^{(i)} x_{ik}(0) = x_i^{\infty}. \quad (4.1.9)$$

Таким образом, финальное мнение члена сильной подгруппы зависит лишь от начальных мнений членов этой подгруппы и одинаково для всей подгруппы.

Если вершина  $y_{r+1,j}$  входит в  $T$ , то

$$x_{r+1,j}^{\infty} = \sum_{i=1}^r b_{ji} \left( \sum_{k=1}^{n_i} w_k^{(i)} x_{ik}(0) \right) = \sum_{i=1}^r b_{ji} x_i^{\infty}. \quad (4.1.10)$$

Итак, финальное мнение члена переходного множества ("спутника") зависит от мнений членов всех сильных подгрупп, но не от мнений других спутников. Это финальное мнение представляет собой взвешенное среднее общих финальных мнений всех сильных подгрупп. Тем самым, указанный аппарат позволяет решить задачу прогноза устойчивых финальных мнений членов группы при заданных начальном векторе мнений и коэффициентах взаимодействия агентов.

Изложенные результаты имеют чрезвычайно важное значение для маркетинга. Методология формализованных маркетинговых исследований включает:

- статистические методы наблюдения: выборка, формирование рядов данных, отчётность и переписи, опросы, экспертные оценки;
- статистические и эконометрические методы анализа рынка: группировка абсолютных, средних и относительных величин, индексные, трендовые и регрессионные модели, методы вариационного, дисперсионного, корреляционного, факторного, кластерного анализа;

- модели оценки рисков, основанные на теориях вероятностей и принятия решений; модели теории массового обслуживания; модели реакции рынка на маркетинговые воздействия и др.;
- методы эвристических оценок и комбинированные методы, сочетающие детерминированные и стохастические подходы;
- методы коммерческого анализа финансово-экономического потенциала предприятия;
- квалиметрический анализ, т.е. методы количественной оценки качественных факторов (Беляевский 2017, с.23-24).

Несомненно, сюда можно отнести и сетевые модели влияния и управления, которые относятся к дескриптивному маркетинговому исследованию, дающему руководству предприятия материалы для дополнительного неформального анализа и окончательного принятия решения. Следует отметить, что среди авторов теории маркетинговых исследований существуют сторонники как формализации этой деятельности с использованием количественных оценок и методов статистического и экономико-математического моделирования, так и неформального анализа на основе качественных оценок и методов. Очевидно, на практике разумно применять сбалансированное сочетание формальных и неформальных подходов в зависимости от целей и возможностей конкретного маркетингового исследования.

Важнейшим составным элементом маркетинговых исследований выступает маркетинговый анализ - оценка, интерпретация, моделирование и прогнозирование процессов и явлений рынка и собственной рыночной деятельности фирмы. Для проведения маркетингового анализа большие компании используют автоматизированные системы поддержки решений, облегчающие и обосновывающие управленческую деятельность с использованием интерактивного режима. Такие системы сочетаются с маркетинговой информационной системой (Ф. Котлер), включающей человеческий, методологический и технологический факторы и отражающей

процесс формирования, передачи и использования маркетинговой информации. Маркетинговый анализ решает следующие задачи:

- 1) оценка и прогноз состояния и развития рынка, определение позиции фирмы на рынке в целом или его сегменте;
- 2) выявление реакции рынка на маркетинговые воздействия и моделирование закономерностей спроса;
- 3) характеристика финансово-экономического потенциала фирмы, ее конкурентоспособности;
- 4) маркетинговое обоснование разработки товара, его тестирование и оценка конкурентоспособности;
- 5) характеристика потенциальных дистрибьюторов в соответствии с определенными критериями выбора;
- 6) оценка поведения и потенциальных возможностей конкурентов, прогнозирование их действий;
- 7) выявление и прогнозирование мнений, поведения и предпочтения потребителей;
- 8) оценка коммерческого риска (Беляевский 2017, с.113).

Модели влияния в социальных сетях направлены на решение задачи 7), а модели управления - задачи 2) и отчасти задачи 6).

В качестве социальной группы применительно к маркетингу естественно рассматривать целевую аудиторию некоторой фирмы. Мнения агентов можно трактовать как некоторые характеристики потребления, например, общую сумму покупок товаров (услуг) данной фирмы за определённый период, число посещений предприятий этой фирмы и т.п. Тогда очевидным образом возникают постановки задач управления мнениями, где в качестве управляющих переменных используются маркетинговые воздействия (реклама, продвижение товаров и т.п.).

Основное внимание здесь следует обратить на то, что согласно (4.1.9) и (4.1.10) влияние нужно оказывать только на членов сильных подгрупп, которые определяют как собственные итоговые мнения, так и мнения всех

спутников. Поскольку членов сильных подгрупп в реальности намного меньше, чем спутников, то этот факт позволяет сэкономить значительные средства за счет целенаправленного воздействия.

Рассмотрим два примера социальных сетей. Первый случай – группа друзей собирается пойти на футбольный матч и договаривается о цене билетов (мнения). Вершина 1 – успешный бизнесмен; 2 – офисный работник, у которого есть молодой коллега (вершина 8), тоже фанат футбола, а у него, в свою очередь, компания друзей из четырёх человек (вершины 9, 10, 11, 20). Вершина 4 – «душа компании» и блогер, у которого есть свои подписчики (13-16 вершины). Вершина 6 – друг детства 4 и 3. Вершина 3 – женатый человек и ему важно мнение его жены (вершина 5), также ему придётся взять выходной у начальника, тоже фаната футбола (вершина 7). Вершина 17 – узнал из видео-блога о предстоящем матче и позвал друга, но так как он тоже женат, ему придётся согласовать поход на матч со своей женой.

Получаем следующий граф (рис. 4.1.1). Каждый из участников данной сети преследует свои интересы, например, успешный бизнесмен хочет большего комфорта, жёны хотят меньших затрат, молодые специалисты желают попасть на матч, но потратить как можно меньше. Билеты на матч продаются в диапазоне от 100 руб. до 10000 руб. Вектор начальных мнений имеет вид (10000, 5000, 5000, 5000, 1000, 5000, 8000, 1000, 1000, 1000, 1000, 300, 500, 800, 400, 900, 200, 1000, 1000, 1000).

Базовые макропараметры социальной сети показаны на рис. 4.1.2-4.1.4. Плотность сети: 0.09210526. Значение плотности невелико, так как группа состоит из людей, чаще всего не знакомых друг с другом лично. Взаимность: 0.4705882. Связи чаще всего не взаимны. Транзитивность: 0.2790698. Низкое значение говорит о том, что "друг моего друга не всегда и мой друг тоже".

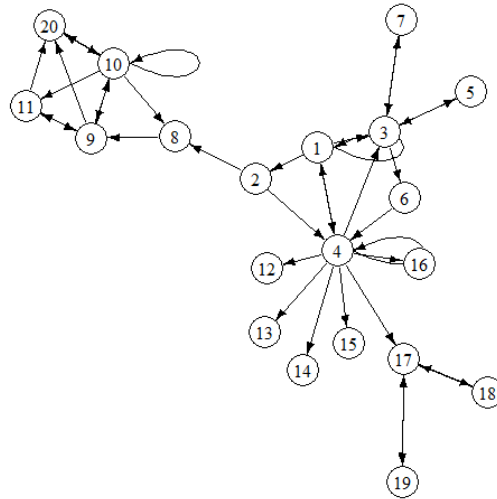


Рис.4.1.1. Сетевая модель "футбольный матч"

**Degree Centrality**

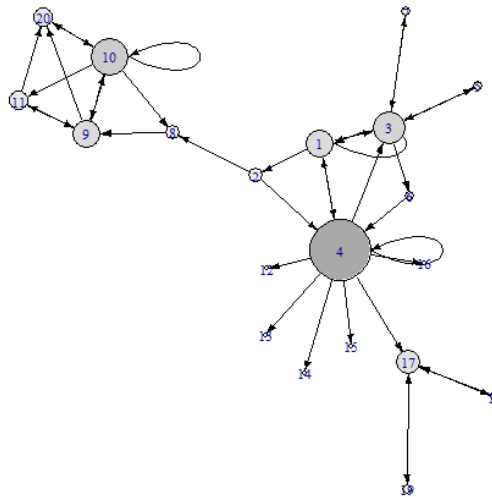


Рис.4.1.2. График центральности по степени

Самая высокая центральность по степени у вершины 4, что объясняется его увлечением, блогеры быстро распространяют информацию.

### Closeness centrality

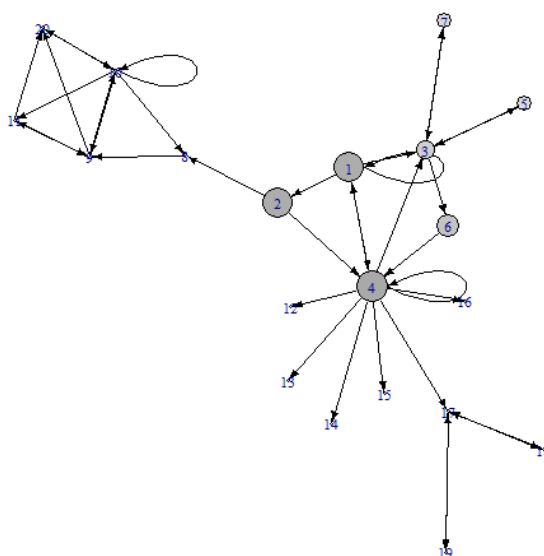


Рис.4.1.3. График центральности по близости

Самый высокий показатель центральности по близости у вершин 1, 2 и 4. Данные вершины так или иначе оказывают сильное влияние на все остальные вершины.

### Betweenness Centrality

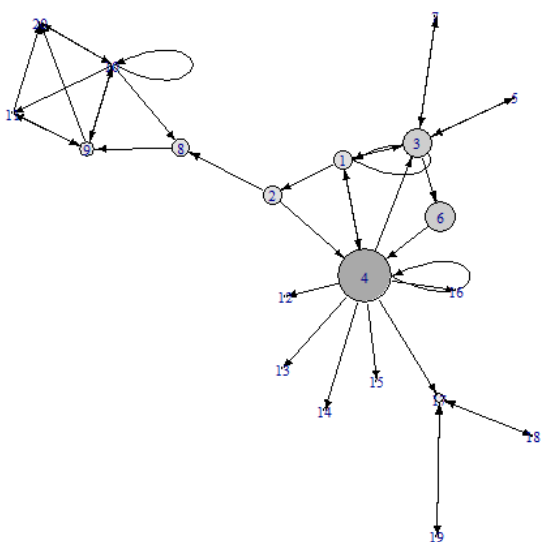


Рис.4.1.4. График центральности по посредничеству

Вершина 4 обладает самым высоким показателем посредничества, так как является связующим узлом между двумя частями сети. Совершенно

незнакомые люди узнали о матче, просто посмотрев видео-блог. Наконец, сильные компоненты данной сети и их спутники показаны на рис. 4.1.5.

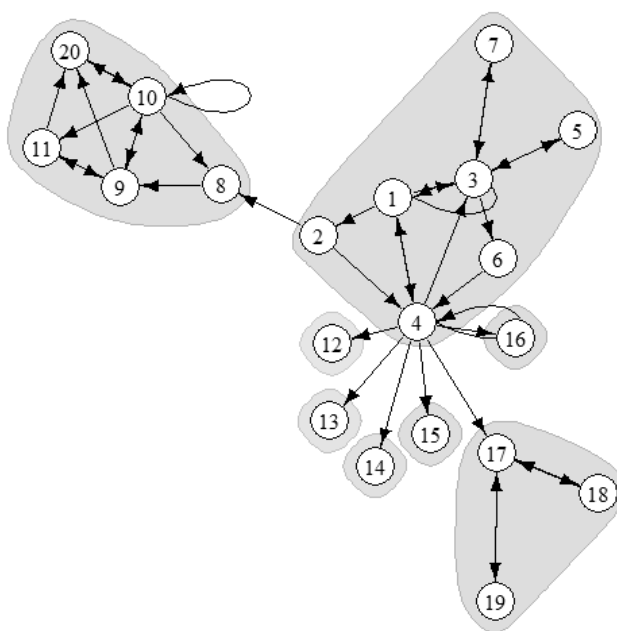


Рис. 4.1.5. Сильные компоненты и спутники

Рассмотрим следующий пример, иллюстрирующий довольно распространённый вид деятельности в социальных сетях, а именно набирающее популярность блогерство (рис. 4.1.6). Вершины графа представляют пять подруг в декретном отпуске, ведущих блоги о красоте, здоровье и своих детях (вершины 1-5), а также их подписчиков, которые обсуждают сумму на покупку коляски (мнения). Так как данный граф содержит несколько сотен тысяч вершин, его сложно изобразить полностью. Возьмём для наглядности граф со 100 вершинами и вычислим характеристики для такой социальной сети (рис. 4.1.7-4.1.8).

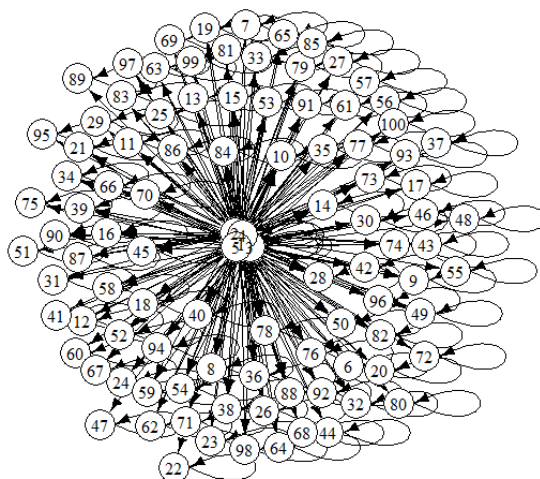


Рис.4.1.6. Социальный граф блогеров и их подписчиков

Очевидно, что в такой сети сильной подгруппой будут сами блогеры, а их подписчики – спутниками.

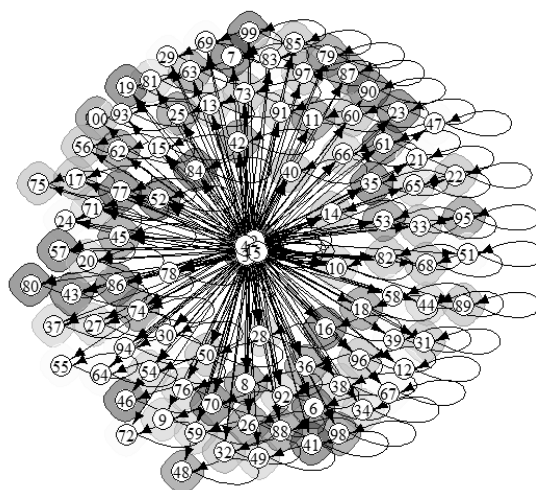


Рис.4.1.7. Сильные компоненты и спутники

Плотность сети: 0.03464646. Взаимность: 0.03292181.

Транзитивность: 0.09532888.

Очень низкие показатели характеристик объясняются структурой сети и типом взаимосвязи её объектов. Аналогичная ситуация с микропараметрами сети. Блогеры обладают властью над своими подписчиками, которые прислушиваются к их мнению. Блоги служат хорошим инструментом для рекламы.



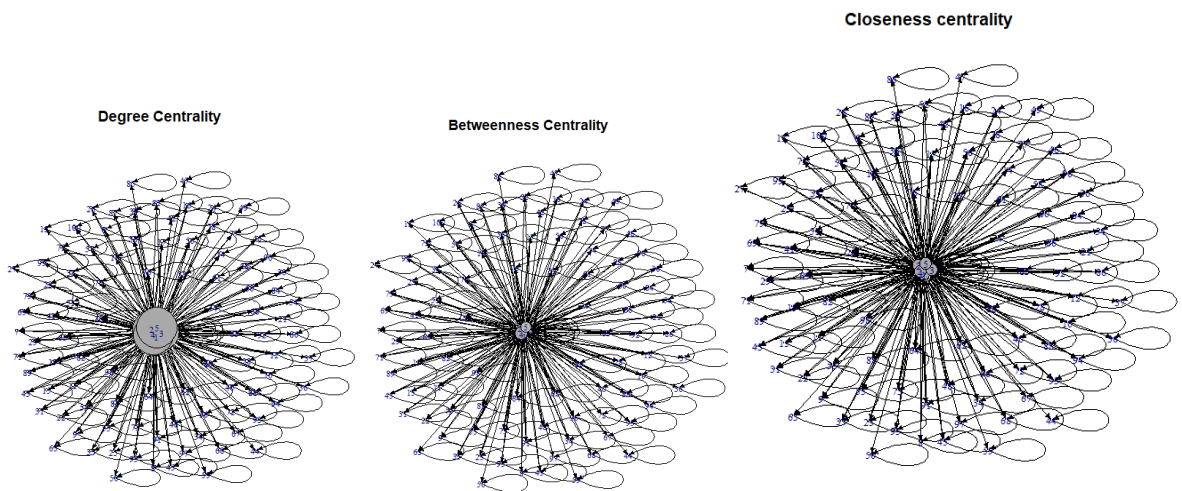


Рис. 4.1.8. Показатели центральности

Для решения задачи прогноза разработан оригинальный алгоритм (Агиева и др. 2019), программно реализованный с помощью языка R - (Кабаков 2016). Структурная схема алгоритма показана на рис. 4.1.9.

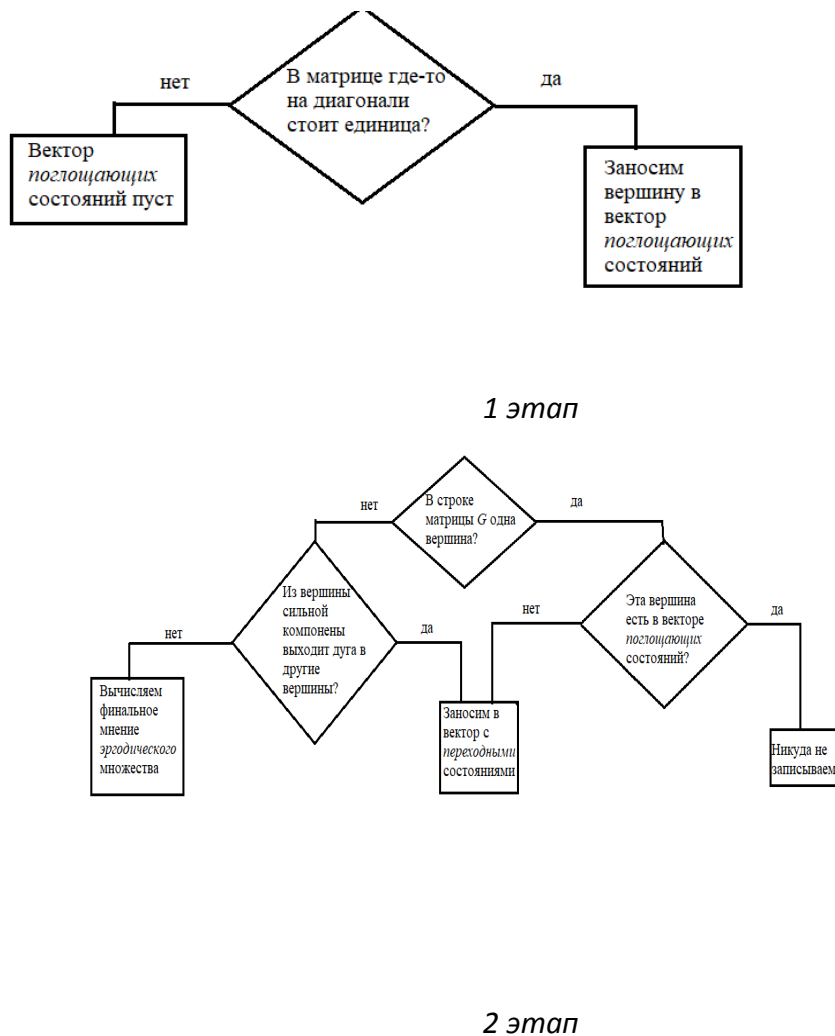


Рис. 4.1.9. Структурная схема алгоритма решения задачи прогноза

Используем данные первого примера (футбольный матч) и интерпретируем полученные результаты. При указанном векторе начальных мнений вершины 1-7 образуют эргодическое множество, их финальное мнение равно 6876 рублей. Внутри этой группы люди влияют друг на друга, хотя не все члены обязаны быть знакомы между собой лично (например, жена работника 3 и начальник этого же работника). Но на эту группу никто извне не влияет. Неженатые и не имеющие друзей - фанатов футбола подписчики блогера 4 находятся под полным его влиянием и также готовы потратить на билет 6876 рублей. Вершины 17-19 образуют сильную компоненту, однако не являются эргодическим множеством. На семьянина 17 оказывает влияние блогер 4. Несмотря на то, что 17 должен посоветоваться с женой перед покупкой билета, можно видеть, что жена всецело доверяет мужу. Соответственно 17-19 также готовы потратить 6876 рублей. Вершины 8-11, 20 также образуют сильную компоненту, но на них извне влияет вершина 2. В их случае участник номер 8 готов потратить 7410 рублей, номер 9 - 7392 рубля, номер 10 - 7650, номер 11 - 7451 и номер 20 - 7520.

Можно сделать вывод о том, что блогер имеет очень большое влияние не только на своих непосредственных друзей и подписчиков, но и на их жён и друзей.

Программа на R хорошо работает и на втором примере, показывая итоговые финальные мнения у каждой вершины графа. Подруги-блогеры формируют мнение внутри своей группы, оказывая огромное влияние на своих непосредственных подписчиц, а также на их подруг, которые не являются подписчицами. Можно сказать, что если производитель колясок захочет дать свою рекламу, то лучше обратиться напрямую к этим блогерам, нежели воздействовать напрямую на покупательниц.

Хотя отдельные действия алгоритма, такие как транспонирование матриц и решение СЛАУ, решаются стандартными инструментами языка R, в

совокупности реализация алгоритма оригинальна, поскольку до сих пор ещё не была написана программа, вычисляющая финальные мнения.

Наиболее трудоёмкой частью алгоритма является выполнение циклов. Самая сложное в данной реализации – это вложенный цикл (глубина равна 2, то есть внутренний и внешний цикл). Оценивая сложность алгоритма, достаточно использовать только ту часть, что растёт больше всего. В итоге получим, что сложность алгоритма вычисляется как  $O(n^2)$ , где  $n$  – число вершин графа. Такая сложность вполне приемлема, соответственно программу можно применять и для графов с намного большим количеством вершин (Агиева и др. 2019).

Рассмотрим классификацию задач управления целевой аудиторией в маркетинге и приведём примеры постановок таких задач. Как уже говорилось, основная идея подхода заключается в следующем. Анализируемая социальная сеть трактуется как целевая аудитория некоторой фирмы, действующей как субъект управления. Мнение агента (представителя целевой аудитории) отражает сумму денег, которую агент готов потратить на покупку товаров или услуг данной фирмы. Естественно, фирма заинтересована в максимизации этой величины на рассматриваемом отрезке времени посредством маркетинговых воздействий на целевую аудиторию при ограничениях на маркетинговый бюджет, что в совокупности и порождает задачу оптимального управления.

На этапе анализа в составе целевой аудитории выделяются сильные подгруппы и спутники. Поскольку финальные мнения всех агентов зависят только от начальных мнений членов сильных подгрупп, то маркетинговые воздействия следует оказывать на них и только на них, что существенно снизит расходы на управление. Таким образом, при воздействии на мнения нужно учитывать только членов сильных подгрупп, а при воздействии на коэффициенты влияния - только на те элементы матрицы влияний, первый индекс которых соответствует члену сильной подгруппы. В дальнейших постановках задач будем считать эти условия выполненными.

Итак, группа из  $n$  членов  $y_1, \dots, y_n$  должна принять некоторое решение. В начальный момент времени  $t = 0$  каждый член группы  $y_i$  имеет по этому поводу определенное мнение  $x_i^0$ ; вектор  $x^0 = (x_1^0, \dots, x_n^0)$  задан. Пусть число  $a_{ij} \geq 0$  обозначает степень влияния  $y_i$  на  $y_j$ , причем

$$\sum_{i=1}^n a_{ij} = 1, \quad j = 1, \dots, n; \quad (4.1.11)$$

при необходимости выполнение этого условия можно обеспечить

нормировкой  $a_{ij} := a_{ij} / \sum_{i=1}^n a_{ij}$ . Формализуем взаимодействие в группе

орграфом влияний  $D = (Y, A)$ ,  $Y = \{y_1, \dots, y_n\}$ ,  $A = \|a_{ij}\|_{i,j=1}^n$ . Предположим, что:

- 1)  $D$  не меняется в процессе принятия решений;
- 2) решения принимаются в дискретные моменты времени  $t=0, 1, 2, \dots, T$ ;
- 3) мнения членов группы меняются согласно правилу

$$x_j^{t+1} = \sum_{i=1}^n a_{ij} x_i^t, \quad x_j^0 = x_{j0}, \quad j = 1, \dots, n. \quad (4.1.12)$$

Будем считать, что имеется субъект управления (внешний агент воздействия), способный влиять на целевую аудиторию для изменения ее мнений в нужную для себя сторону. Для классификации задач оптимального управления целесообразно рассматривать следующие признаки.

1. Какова цель управления? а) максимизация финальных мнений членов целевой аудитории; б) максимизация суммы мнений за период от  $t=1$  до  $t=T$ .

2. На что оказывается воздействие? а) на коэффициенты влияния, которые считаются постоянными; б) на мнения, здесь можно различать два подслучая: б1) только на начальные мнения; б2) на мнения в любой момент времени  $t=0, 1, 2, \dots, T$ .

3. Как оказывается воздействие? а) программные или позиционные стратегии. В первом случае правило управления задается на весь период

$t=0,1,2,\dots,T$ , во втором в каждый момент времени оно зависит от вектора мнений в этот момент; б) аддитивное или мультипликативное воздействие. Это технический признак: в первом случае управление прибавляется к значению мнения, во втором умножается на него.

4. На кого оказывается воздействие? а) на одного члена каждой сильной подгруппы; б) на нескольких (возможно, всех) членов каждой сильной подгруппы.

Затраты на управление (маркетинговое воздействие) будем считать степенной функцией и предполагать, что сумма допустимых затрат ограничена сверху некоторой константой (маркетинговым бюджетом фирмы). Рассмотрим примеры задач оптимального управления, порождаемые различными сочетаниями указанных признаков классификации.

Для построения примеров примем следующие предположения.

1. Воздействие оказывается только на начальные или текущие мнения членов сильных подгрупп или соответствующие коэффициенты влияния.

2. Воздействие имеет аддитивную природу.

3. Затраты на управление описываются степенной функцией, т.е. аддитивное воздействие  $u$  влечёт затраты  $u^p$ ,  $p > 0$ . Значение  $p$  выступает важным параметром моделирования и определяется для каждой конкретной маркетинговой кампании отдельно (может быть как меньше, так и больше единицы).

*Пример 1.* Цель воздействия - увеличение финальных мнений членов целевой аудитории. Воздействие в программных стратегиях оказывается на начальные мнения членов сильных подгрупп.

Напомним, что общее финальное мнение  $x_i^\infty$  всех членов  $i$ -й сильной подгруппы вычисляется по формуле:

$$x_i^\infty = \sum_{j=1}^{n_i} w_j^{(i)} x_{j0}^{(i)}, \quad (4.1.13)$$

где  $x_{j0}^{(i)}$  - начальные мнения членов  $i$ -й сильной подгруппы,  $w_j^{(i)}$  - компоненты стационарного вектора регулярной цепи Маркова для данной сильной подгруппы. При аддитивном воздействии на первых  $m_i$  членов  $i$ -й сильной подгруппы формула (4.1.13) принимает вид

$$x_i^\infty = \sum_{j=1}^{m_i} w_j^{(i)} (x_{j0}^{(i)} + u_{j0}^{(i)}) + \sum_{j=m_i+1}^{n_i} w_j^{(i)} x_{j0}^{(i)}, \quad (4.1.14)$$

где  $u_{j0}^{(i)}$  - воздействие на начальное мнение  $j$ -го члена  $i$ -й сильной подгруппы,  $n_i$  - общая численность  $i$ -й сильной подгруппы. В частности, при  $m_i = 1$  воздействие оказывается только на одного члена  $i$ -й сильной подгруппы, а при  $m_i = n_i$  - на всех её членов.

Тогда задача управления становится статической задачей оптимизации

$$\sum_{i=1}^r \left[ x_i^\infty - \sum_{j=1}^{m_i} [u_{j0}^{(i)}]^p \right] \rightarrow \max \quad (4.1.15)$$

$$\sum_{i=1}^r \sum_{j=1}^{m_i} [u_{j0}^{(i)}]^p \leq R, \quad (4.1.16)$$

где  $r$  - общее число сильных подгрупп,  $R$  - маркетинговый бюджет фирмы, максимизирующей целевой функционал (4.1.15). Уравнение динамики (4.1.12) не требуется, поскольку речь идёт о финальных мнениях, вычисляемых по формуле (4.1.14).

*Пример 2.* Цель воздействия - максимизация суммы мнений членов целевой аудитории за весь период от  $t=1$  до  $t=T$ . Воздействие в позиционных стратегиях оказывается на текущие мнения членов сильных подгрупп.

Позиционные стратегии будем определять правилом

$$u_j^t(x_j^t) = \begin{cases} u_j^H, & x_j^t < x_j^*, \\ u_j^L, & \text{иначе, } u_j^L < u_j^H. \end{cases} \quad (4.1.17)$$

Здесь  $x_j^*$  - экспертно оцениваемое пороговое значение мнения. Если текущее значение  $x_j^t$  меньше порога, то применяется более сильное воздействие  $u_j^H$ , если больше, то менее сильное  $u_j^L$ .

Задача оптимального управления имеет вид

$$\sum_{t=1}^T \sum_{i=1}^r \left[ \sum_{j=1}^{m_i} (x_j^t + u_j^t(x_j^t)) + \sum_{j=m_i+1}^{n_i} x_j^t - \sum_{j=1}^{m_i} [u_j^t]^p \right] \rightarrow \max \quad (4.1.18)$$

$$\sum_{t=1}^T \sum_{i=1}^r \sum_{j=1}^{m_i} [u_j^t]^p \leq R, \quad (4.1.19)$$

$$x_j^{t+1} = \sum_{k=1}^r \left[ \sum_{i=1}^{m_k} a_{ij} (x_i^{(k)t} + u_i^{(k)t}) + \sum_{i=m_k+1}^{n_k} a_{ij} x_i^{(k)t} \right], \quad x_j^0 = x_{j0}, \quad j = 1, \dots, n. \quad (4.1.20)$$

*Пример 3.* Цель воздействия - максимизация суммы мнений членов целевой аудитории за весь период от  $t=1$  до  $t=T$ . Воздействие в программных стратегиях оказывается на коэффициенты влияния. Заметим, что в этом случае нецелесообразно ставить задачу максимизации финальных мнений, поскольку затруднительно проследить их зависимость от коэффициентов влияния. Воздействие на коэффициенты влияния описывается правилом

$$a_{ij} := \begin{cases} a_{ij} + u_i, & i - \text{индекс члена сильной подгруппы,} \\ a_{ij}, & \text{иначе.} \end{cases} \quad (4.1.21)$$

Тогда задача оптимального управления принимает вид

$$\sum_{t=1}^T \sum_{i=1}^r \left[ \sum_{j=1}^{n_i} x_j^t - \sum_{j=1}^{m_i} [u_j]^p \right] \rightarrow \max \quad (4.1.22)$$

$$\sum_{t=1}^T \sum_{i=1}^r \sum_{j=1}^{m_i} [u_j]^p \leq R \quad (4.1.23)$$

в силу уравнения динамики (4.1.12), где  $a_{ij}$  вычисляются по формуле (4.1.21).

Все указанные постановки задач следует дополнить некоторыми условиями гомеостаза относительно текущих или финальных мнений.

Одним из наиболее практичных методов решения указанных задач оптимального управления представляется описанный выше метод КРС имитационного моделирования (Ougolnitsky and Usov 2018). Рассмотрим его для ещё одной постановки задачи управления - максимизации суммы мнений

членов целевой аудитории на периоде  $t=1, \dots, T$  путём воздействия в программных стратегиях на текущие мнения членов сильных подгрупп:

$$J = \sum_{t=1}^T \sum_{i=1}^r \left[ \sum_{j=1}^{m_i} (x_j^t + u_j^t) + \sum_{j=m_i+1}^{n_i} x_j^t + \sum_{j=1}^{m_i} [u_j^t]^p \right] \rightarrow \max \quad (4.1.24)$$

$$\sum_{t=1}^T \sum_{i=1}^r \sum_{j=1}^{m_i} [u_j^t]^p \leq R, \quad (4.1.25)$$

$$x_j^{t+1} = \sum_{k=1}^r \left[ \sum_{i=1}^{m_k} a_{ij} (x_i^{(k)t} + u_i^{(k)t}) + \sum_{i=m_k+1}^{n_k} a_{ij} x_i^{(k)t} \right], \quad x_j^0 = x_{j0}, \quad j = 1, \dots, n. \quad (4.1.26)$$

Здесь  $x_i^{(k)t}$  - текущие мнения членов  $k$ -й сильной подгруппы,  $u_j^{(k)t}$  - воздействие на текущее мнение  $j$ -го члена  $k$ -й сильной подгруппы, причем воздействию подвергаются первые  $m_k$  членов;  $n_i$  - общая численность  $i$ -й сильной подгруппы,  $r$  - общее число сильных подгрупп,  $R$  - маркетинговый бюджет фирмы, максимизирующей целевой функционал (4.1.24);  $p$  - параметр степенной функции затрат на управление;  $n$  - общая численность целевой аудитории.

Для решения задачи (4.1.24)-(4.1.26) методом КРС имитационного моделирования необходимы следующие данные.

1. Общая численность целевой аудитории  $n$  и маркетинговый бюджет фирмы  $R$  известны. Период моделирования  $T$  задаётся. Число сильных подгрупп  $r$  выявляется на этапе анализа целевой аудитории. Будем считать для простоты, что воздействие оказывается только на одного члена каждой сильной подгруппы, т.е.  $m_k = 1, k = 1, \dots, r$  (другие варианты можно рассмотреть при более детальном исследовании).

2. Главные входные данные модели - вектор начальных мнений  $x_0$  и матрица влияния  $A = \| a_{ij} \|$ . Для идентификации этих данных необходимы исследования, проводимые по специальной методике.

3. Важную роль играет параметр функции затрат  $p$ . Если  $p < 1$ , то затраты на управление не слишком велики по сравнению с ожидаемым



результатом (управление высокоэффективно), а если  $p > 1$ , то наоборот. Поэтому в соответствии с идеей КРС естественно в первом приближении рассмотреть значения  $p \in \{1/2, 1, 2\}$ . В частности, в промежуточном случае  $p = 1$  можно искать аналитическое решение линейной задачи управления.

4. Наконец, собственно сценарии имитационного моделирования образуют значения управлений  $u_j^t, t = 1, \dots, T, j = 1, \dots, m_i, i = 1, \dots, r$ . В принятом предположении  $m_i = 1, i = 1, \dots, r$  они образуют двумерный массив  $\|u_i^t\|_{i=1, t=1}^{r, T}$ , т.е. для моделирования надо заполнить таблицу

	t=1	t=2	...	t=T
i=1	$u_1^1$	$u_1^2$	...	$u_1^T$
i=2	$u_2^1$	$u_2^2$	...	$u_2^T$
...	...	...	...	...
i=r	$u_r^1$	$u_r^2$	...	$u_r^T$

Обязательным требованием при выборе сценариев служит выполнение бюджетного ограничения (4.1.25). Далее нужно рассмотреть базовый сценарий  $u_i^t \equiv 0$  (отсутствие управления), с которым следует сравнивать все остальные для сравнительного анализа эффективности управления.

Анализ бюджетного ограничения (4.1.25) должен выявить ограничения сверху на возможные значения управляющих воздействий (максимально допустимые значения  $u^{\max}$  такие, что  $0 \leq u_i^t \leq u^{\max}, i = 1, \dots, r, \dots, t = 1, \dots, T$ ). Тогда простейший кандидат в КРС имеет вид  $u_i^t \in \{0, u^{\max} / 2, u^{\max}\}, i = 1, \dots, r, t = 1, \dots, T$ . Для этого сценария необходимо вычислить значения целевого функционала (4.1.24) и проверить выполнение свойств (а) и (б) из данного выше определения (см. параграф 3.3). Если они выполняются для некоторого не слишком большого  $\Delta > 0$ , то действительно найдено множество КРС, иначе придётся выбирать другие значения управляющих воздействий, исходя из разумных соображений об их смысле с обязательным выполнением бюджетного требования (4.1.25). Метод КРС носит эвристический характер, поэтому универсальные рекомендации здесь невозможны и следует руководствоваться спецификой конкретных маркетинговых акций.

## **4.2. Модели управления устойчивым развитием учреждений высшего образования**

Материал этого параграфа основан на работах (Мальсагов 2018а,б; Мальсагов и Угольницкий 2018; Мальсагов 2019а,б; Мальсагов и др. 2019).

Представляется целесообразным применение положений теории управления устойчивым развитием активных систем к решению задач управления УР учреждений высшего образования (университетов).

Предлагаемый подход обладает следующими отличительными особенностями. Во-первых, изучаются процессы управления университетами как организациями (Новиков 2013), не принимая во внимание, например, взаимодействие преподавателей и студентов в ходе обучения и т.п. Во-вторых, основное внимание уделяется проблеме УР университетов как активных систем. В-третьих, аппарат исследования преимущественно динамические модели конфликтного и кооперативного управления (иерархические дифференциальные игры) (Петросян и др. 2012; Basar and Olsder 1999) и методы их решения (при ведущей роли имитационного моделирования). В-четвёртых, конечная цель исследования состоит в разработке и апробации практических методик решения отдельных задач управления УР университетов на базе соответствующих математических моделей и программного обеспечения.

Важный вклад в данную тематику внесла монография (Новиков 2009), специфицирующая методологию управления организационными системами применительно к образовательным системам. В этой работе: предложены основные принципы модернизации системы образования, а именно повышения его качества, доступности и эффективности, что требует решения соответствующих задач управления; сделан вывод о том, что центральным системообразующим элементом теории управления образовательными системами является категория организации; выделены шесть иерархических уровней образовательных систем; ряд основных понятий теории управления

организационными системами (виды управления, принципы и механизмы управления) до определенной степени конкретизирован и проиллюстрирован для случая образовательных систем.

В рамках развиваемой авторами теории под УР учреждений высшего образования понимается одновременное выполнение условий гомеостаза и мотивированности. Гомеостаз означает, что ключевые показатели функционирования университета (научно-исследовательской, учебно-методической и инновационно-предпринимательской деятельности) должны принимать значения из определенного диапазона. Мотивированность подразумевает учёт интересов сотрудников университета и их согласование с интересами организации в целом.

Решение задач управления УР университетов целесообразно осуществлять на основе построения и исследования отдельных моделей управления УР на уровне кафедры, факультета и университета в целом. К числу таких моделей относятся: динамическая модель стимулирования участников научного проекта; статическая модель распределения вознаграждения как игра в форме характеристической функции; динамическая модель контроля распределения рабочего времени сотрудников кафедры как дифференциальная игра в нормальной форме и в форме характеристической функции на основе метода принуждения; двойственные задачи дискретного программирования с учетом требований УР для определения оптимальной кадровой структуры факультета, в том числе оценки эффективности объединения кафедр; динамические СОЧИ-модели продвижения инноваций как дифференциальные игры в нормальной форме и в форме характеристической функции на основе метода побуждения; динамические модели борьбы с коррупцией. В качестве основного метода решения динамических игр предлагается метод КРС имитационного моделирования (Ougolnitsky and Usov 2018). Используются также традиционные методы (принцип максимума Понтрягина, динамическое программирование, методы дискретной оптимизации).

Университеты - это сложные многоуровневые динамические системы со специфическими целями и интересами агентов на всех уровнях и столь же сложной системой отношений со внешней средой. Рассмотрим университет как активную систему более детально по следующему плану: элементы; подсистемы; функции; внешняя среда; имеющиеся проблемы.

*Элементы.* На верхнем уровне организационной структуры университета располагается административный персонал. Линейное управление осуществляют ректор университета, деканы факультетов (директора институтов) и заведующие кафедрами (лабораториями), отвечающие за общее руководство по всем направлениям деятельности. Функциональным управлением (по отдельным направлениям деятельности) занимаются проректоры, заместители деканов (директоров институтов), руководители обеспечивающих подразделений (учебных и научных управлений, бухгалтерской, кадровой, хозяйственной и прочих служб). Зачастую линейные руководители совмещают административную нагрузку с научной и/или преподавательской деятельностью.

Преподаватели-исследователи выступают управляемой подсистемой относительно административного персонала и управляющей подсистемой по отношению к студентам (аспирантам, докторантам). Смысл данного термина состоит в том, что преподаватель современного университета в обязательном порядке должен заниматься научными исследованиями, быть в курсе текущего состояния вопроса по своей тематике и на этой основе читать лекции и проводить занятия.

Студенты (в том числе аспиранты и докторанты) играют двойственную роль. С одной стороны, они образуют управляемую подсистему, которую обучают преподаватели-исследователи. С другой стороны, они отнюдь не пассивны. Во-первых, в педагогике уже давно утвердилось представление о том, что субъект-объектное воздействие на учащегося должно уступить место субъект-субъектному взаимодействию, при котором учитываются интересы ученика и он активно вовлекается в процесс обучения. Во-вторых,

концепция активной системы предполагает активное влияние управляемой подсистемы на процесс управления. Студенты играют существенную роль в жизни университета, в том числе в принятии управленческих решений. Это осуществляется через студенческие советы и профсоюзы, включение студентов в состав Учёных советов подразделений, обратную связь при оценке студентами работы преподавателей, выборе студентами тех или иных предлагаемых курсов и т.п.

*Подсистемы.* В структуре университета в качестве подсистем естественным образом выделяются факультеты и институты, а также функциональные подразделения. Это определяет традиционную линейно-функциональную структуру управления. Вместе с тем, в ряде университетов используется проектная система управления (например, руководство образовательными программами и исследовательскими проектами); в этом случае проектные группы можно трактовать как временные подсистемы наряду с основными (постоянными) линейно-функциональными.

*Функции.* Исторически университеты выполняли в первую очередь образовательную функцию, хотя среди университетских преподавателей всегда были видные учёные. На рубеже XVIII-XIX-го столетий в Германии братья Александр и Вильгельм фон Гумбольдт предложили концепцию нового университета, в котором образование должно базироваться на прочном фундаменте научных исследований при ведущей роли последних. Во второй половине XX века предложенная братьями Гумбольдт и уже ставшая классической формула "наука + образование" дополнилась третьим слагаемым. Правда, его трудно однозначно охарактеризовать одним словом: речь идёт об инновациях, предпринимательской деятельности, вкладе университетов в региональное развитие. Термины "предпринимательский университет" и "академический капитализм" за последние два десятилетия прочно утвердились в соответствующей специальной литературе (Clark 1998; Slaughter and Leslie 1997). Таким образом, в настоящее время можно говорить о триединой функции университета: "наука + образование + внедрение".

Важно подчеркнуть, что в этой триаде научные исследования играют ведущую роль и создают основу как для реализации образовательных программ всех уровней, особенно начиная с магистерского, так и для всех региональных и отраслевых приложений, включая разработку новых продуктов и технологий, организацию дополнительного образования, консультации органов власти и управления и т.д.

*Внешняя среда.* В составе внешней для университетов среды целесообразно выделить следующие основные элементы: руководящий орган - Министерство науки и высшего образования РФ; региональные органы власти и управления; государственные научные фонды; другие университеты региона и страны в целом; население региона; находящиеся на территории региона предприятия и организации.

Проведённый анализ позволяет выделить следующие основные проблемы, порождающие конфликты в университетах как активных системах.

1. Распределение ограниченных ресурсов между различными подразделениями, а также между участниками научно-исследовательских проектов.

2. Распределение сотрудниками своего рабочего времени между основной работой в университете и различными дополнительными активностями за его пределами (совместительство, репетиторство, консультирование и т.п.), которые могут приносить больший доход.

3. Противоречия между стремлением сотрудников отдельных кафедр и лично их заведующих к сохранению существующих кафедр и объективными нормами управляемости и интересами факультетов.

4. Противоречие между необходимостью внедрения технологических и управленческих инноваций и стремлением сотрудников к экономии усилий и минимизации риска.

5. Противоречие между традиционно сложившимся представлением об университетских руководителях (от ректора до заведующего кафедрой) как в

первую очередь учёных, главах научных школ и методистах, и современной тенденции отдавать эти позиции "эффективным менеджерам", иногда весьма далёким от науки и образования.

6. Стремление обеспечивающих подразделений в рамках самих университетов навязывать правила игры преподавателям-исследователям.

7. Наличие коррупции в учреждениях высшего образования.

Решение этих проблем через согласование интересов активных агентов в целях УР университетов требует использования математических моделей. В первую очередь, это динамические теоретико-игровые модели, описывающие конфликт и кооперацию игроков и обеспечивающие достижение разумных компромиссов между ними. Однако, в ряде случаев целесообразно использовать и более простые статические модели оптимизации.

Устойчивое развитие университета как активной системы предполагает одновременное выполнение условий ее гомеостаза и мотивированности активных агентов. Применительно к гомеостазу в исходной теории УР прочно утвердилась концепция "трех столпов" (three pillars), означающая необходимость одновременного учёта и балансирования экологических, экономических и социальных целей. При моделировании это означает, что в состав вектора состояния обязательно должны входить переменные, описывающие экономические, экологические и социальные показатели. Для университетов "тремя столпами" выступают научно-исследовательская работа (НИР), учебно-методическая работа (УМР) и, что немаловажно, инновационно-предпринимательская работа (ИПР). Показатели по этим трём группам должны определяться и анализироваться на уровне кафедры, факультета и университета в целом, что подразумевает использование процедур агрегирования.

Поэтому предлагается следующая комплексная процедура оценивания гомеостаза университета (с использованием идей комплексного оценивания (Новиков 2009)):

1) составление списка исходных показателей по группам НИР, УМР, ИПР на уровне кафедры;

2) приписывание каждому показателю порядковой оценки

$$x_{ijk}, y_{ijk}, z_{ijk} \in \{2,3,4,5\},$$

где 2 - "плохо", 3 - "удовлетворительно", 4 - "хорошо", 5 - "отлично"; x - НИР; y - УМР; z - ИПР; i - номер факультета; j - номер кафедры; k - номер показателя. Тогда  $X = \|x_{ijk}\|$ ,  $Y = \|y_{ijk}\|$ ,  $Z = \|z_{ijk}\|$  - матрицы оценок исходных показателей для кафедр;

3) вычисление индексов НИР, УМР, ИПР на уровне кафедры:

$$x_{ij} = f_{ij}(X), y_{ij} = g_{ij}(Y), z_{ij} = h_{ij}(Z);$$

для простоты и универсальности целесообразно считать, что

$$\forall i, j \quad f_{ij} = g_{ij} = h_{ij} = F \text{ (единая процедура агрегирования);}$$

4) вычисление индексов НИР, УМР, ИПР на уровне факультетов:

$$x_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}, \quad n_i - \text{число кафедр на факультете } i; \quad y_i, z_i \text{ вычисляются}$$

аналогично, дробные значения округляются по обычным правилам;

5) вычисление индексов НИР, УМР, ИПР на уровне университета:

$$x = \frac{1}{m} \sum_{i=1}^m x_i, \quad m - \text{число факультетов в университете; } y, z \text{ вычисляются}$$

аналогично, дробные значения округляются по обычным правилам;

б) ранжирование кафедр в трёхмерном пространстве факультета, факультетов в трёхмерном пространстве университета; университетов в трёхмерном пространстве страны (мира):

а) нахождение множества Парето в пространстве индексов;

б) возможное вычисление единого индекса гомеостаза. Отметим, что использование одного показателя для оценки гомеостаза университета на любом из трёх уровней представляется чрезмерным упрощением: более адекватно рассматривать трёхмерное пространство качественно различных критериев НИР, УМР и ИПР. Тем не менее, иногда требуется (с пониманием



условности) вычислять единый индекс, что может быть сделано на основе аналогичной предыдущим процедуры агрегирования.

Следует подчеркнуть, что первые три пункта описанной процедуры неизбежно носят субъективный характер. Невозможно предложить какие бы то ни было объективные основания для выбора списка показателей, их оценок (даже порядковых) и последующего агрегирования в единый индекс. Рассчитывать можно только на общие положения системного анализа, опыт и знания экспертов и последующую апробацию конкретных процедур оценивания на практике.

Поэтому рассмотрим конкретную комплексную процедуру оценивания в качестве иллюстративного примера. Специфику процедуры определяют ее первые три пункта, поскольку индексы на уровне факультета и университета получаются простым усреднением соответствующих значений индексов предыдущего уровня. Нахождение множества Парето также представляет собой стандартную операцию, хотя технически ее можно реализовать различными способами (Miettinen 1999).

1. Список предлагаемых исходных показателей гомеостаза на уровне кафедры приведен в таблице 4.2.1. В основном они соответствуют стандартной отчетности. При фиксированных номерах кафедры и факультета достаточно нумеровать показатели.

Таблица 4.2.1

Показатели гомеостаза по группам на уровне кафедры

<b>НИР</b>	<b>УМР</b>	<b>ИПР</b>
$x_1$ - число опубликованных за год статей в изданиях из списка Web of Science/ Scopus	$y_1$ - число опубликованных за год учебников и учебных пособий с официальным грифом	$z_1$ - объем привлечённого финансирования
$x_2$ - число опубликованных за год статей в изданиях из списка ВАК	$y_2$ - число подготовленных электронных образовательных	$z_2$ - число договоров с внешними партнёрами

	ресурсов	(базовые кафедры на предприятиях, совместные проекты и т.п.)
$x_3$ - число сделанных сотрудниками кафедры за год докладов на международных конференциях	$u_3$ - число защищённых под руководством сотрудников кафедры магистерских диссертаций	
$x_4$ - число докторских и кандидатских диссертаций, защищённых за три года сотрудниками кафедры или под их руководством	$u_4$ - средняя успеваемость в группах, где ведут занятия сотрудники кафедры	
$x_5$ - число сотрудников кафедры, участвующих в редколлегиях научных журналов, оргкомитетах конференций и диссертационных советах		
$x_6$ - число подготовленных рецензий на статьи и отзывов на диссертации и авторефераты		

2. Конечно, наибольшую трудность представляет приписывание предложенным показателям оценок на порядковой шкале  $\{2,3,4,5\}$ . Предлагаемые правила заведомо не единственно возможные и носят условный характер, но базируются на опыте и некоторых правдоподобных рассуждениях. Итак, показатели НИР:

$$x_1 = \begin{cases} 5, u_1 \geq 1; \\ 4, 0.5 \leq u_1 < 1; \\ 3, 0.25 \leq u_1 < 0.5; \\ 2, u_1 < 0.25, \end{cases}$$

где  $u_1$  - число опубликованных статей указанного типа в расчёте на одного сотрудника кафедры;

$$x_2 = \begin{cases} 5, u_2 \geq 2; \\ 4, 1 \leq u_2 < 2; \\ 3, 0.5 \leq u_2 < 1; \\ 2, u_2 < 0.5, \end{cases}$$

где  $u_2$  - число опубликованных статей указанного типа в расчёте на одного сотрудника кафедры;

$$x_3 = \begin{cases} 5, u_3 \geq 1; \\ 4, 0.5 \leq u_3 < 1; \\ 3, 0.25 \leq u_3 < 0.5; \\ 2, u_3 < 0.25, \end{cases}$$

где  $u_3$  - число сделанных докладов в расчёте на одного сотрудника кафедры;

$$x_4 = \begin{cases} 5, d \geq 1 \text{ и } c \geq 3; \\ 4, c \geq 2; \\ 3, c \geq 1; \\ 2, c = d = 0, \end{cases}$$

где  $d$  - число докторских диссертаций,  $c$  - число кандидатских диссертаций;

$$x_5 = \begin{cases} 5, u_4 \geq 1; \\ 4, 0.5 \leq u_4 < 1; \\ 3, 0.25 \leq u_4 < 0.5; \\ 2, u_4 < 0.25, \end{cases}$$

где  $u_1$  - число позиций указанного типа в расчете на одного сотрудника кафедры;

$$x_6 = \begin{cases} 5, u_5 \geq 2; \\ 4, 1 \leq u_5 < 2; \\ 3, 0.5 \leq u_5 < 1; \\ 2, u_5 < 0.5, \end{cases}$$

где  $u_5$  - число рецензий или отзывов в расчёте на одного сотрудника кафедры.

Показатели УМР и ИПР оцениваются аналогично.

3. Если факультет и кафедра фиксированы, то матрицы оценок  $X, Y, Z$  становятся векторами и построение индексов НИР, УМР и ИПР кафедры - это задание функции агрегирования

$$F : P^{k_i} \rightarrow P,$$

где  $P = \{2, 3, 4, 5\}$ ,  $k_i$  - число показателей группы  $i=1, 2, 3$  (таблица 1). Обозначим

$p = (p_1, \dots, p_{k_i}) \in P^{k_i}$  - набор значений показателей  $i$ -й группы. Возможные варианты процедур агрегирования:

1)  $F_{\max} = \max_{1 \leq j \leq k_i} \{p_j\}$  - предельно мягкая оценка;

2)  $F_{\min} = \min_{1 \leq j \leq k_i} \{p_j\}$  - предельно жёсткая оценка;

3)  $F_{maj} = p_{maj}$ , где  $|\{p_{maj}\}| > k_i / 2$  - оценка по большинству;

4)  $F_{med} = p_{med}$ , где  $p_{med}$  - медиана (средний член упорядоченной по

возрастанию последовательности  $p$ ) - усреднённая оценка;

5)  $F_w = \sum_{j=1}^{k_i} w_j p_j$  - взвешенная оценка (требует субъективного

назначения весов  $w_j$ , где  $w_j \geq 0, \sum_{j=1}^{k_i} w_j = 1$ ). Заметим, что использование

третьего и четвёртого способов агрегирования в общем случае требует нечётного числа исходных показателей.

Выбор процедуры агрегирования носит принципиальный характер, поскольку исходные показатели из таблицы 4.1.1 неравнозначны. Приведём иллюстративный численный пример (табл. 4.2.2). Пусть число сотрудников кафедры равно 10.

Таблица 4.2.2

Показатели НИР

Исходное значение показателя	Удельная величина (если нужна)	Оценка показателя	Индекс НИР
6	$u_1 = 0.6$	$x_1 = 4$	$F_{\max} = 5$
21	$u_2 = 2.1$	$x_2 = 5$	$F_{\min} = 3$
7	$u_3 = 0.7$	$x_3 = 4$	$F_{\text{maj}} = 4$
$c = 2, d = 0$	$c = 2, d = 0$	$x_4 = 4$	$F_{\text{med}}$ не определена (чётное число показателей)
4	$u_4 = 0.4$	$x_5 = 3$	$F_w = [4.1] = 4,$
6	$u_5 = 0.6$	$x_6 = 3$	$w=(0.3,0.3,0.1,0.1,0.1,0.1)$

Очевидно, первые два способа индексирования не слишком адекватны, поскольку дают крайние оценки, совпадающие лишь с одной-двумя компонентами из шести. Адекватными следует считать остальные способы, приводящие к одинаковому значению индекса (при использовании взвешенной оценки - после округления до целого значения).

Теперь обратимся ко второй компоненте устойчивого развития - мотивированности активных агентов (сотрудников университета). Для количественной оценки мотивированности целесообразно использовать индексы системной согласованности вида

$$SCI = J_0^{\max} - J_0, \quad (4.2.1)$$

где в качестве  $J_0$  берётся значение максимального гарантированного выигрыша Центра для различных информационных регламентов игр, а в качестве  $J_0^{\max}$  - глобальный максимум этой величины. Тогда системная согласованность имеет место при  $SCI = 0$ . Условие  $SCI \approx 0$  говорит о некотором приближении к системной согласованности.

Предполагается, что кафедра, факультет и университет имеют древовидную иерархическую структуру, где в роли Центра выступает соответственно заведующий кафедрой, декан факультета (директор института) и ректор университета, а в качестве агентов - их подчинённые. Для определения величины  $J_0$  в формуле (4.2.1) используются пары классификационных признаков (программные/позиционные стратегии, стратегии при отсутствии/наличии обратной связи по управлениям агентов, принуждение/побуждение), что дает восемь различных индексов системной согласованности, сравнение которых позволяет оценить эффективность различных методов управления устойчивым развитием.

*Динамическая модель стимулирования с учетом требований гомеостаза.* Модель имеет вид

$$J_0 = \int_0^T e^{-\rho t} [H(u(t), x(t)) - \sum_{i \in N} s_i(u(t), x(t))] dt \rightarrow \max \quad (4.2.2)$$

$$s_i(\cdot) \geq 0, i \in N; \quad (4.2.3)$$

$$J_i = \int_0^T e^{-\rho t} [s_i(u(t), x(t)) - h_i(u(t), x(t))] dt \rightarrow \max \quad (4.2.4)$$

$$0 \leq u_i(t) \leq 1, i \in N; \quad (4.2.5)$$

$$\dot{x} = f(x(t), \sum_{i \in N} u_i(t)), x(0) = x_0. \quad (4.2.6)$$

Модель описывает отношения в древовидной системе управления Центр - агенты. В качестве Центра может выступать руководитель научного проекта, в качестве агентов - исполнители. Здесь  $N$  - множество агентов;  $H$  -

возрастающая вогнутая функция дохода Центра,  $H(0,0) = 0$ ;  $h_i$  - возрастающая выпуклая функция затрат агента,  $h_i(0,0) = 0$ ;  $u_i$  - трудовое усилие агента (например, время работы над проектом);  $s_i$  - компенсация трудовых усилий агента Центром (неотрицательная функция - механизм стимулирования);  $x$  - переменная состояния (например, число публикаций по проекту);  $u = (u_1, \dots, u_n)$  - вектор трудовых усилий агентов (все переменные рассматриваются как функции времени);  $\rho \in [0,1]$  - коэффициент дисконтирования;  $J_0, J_i$  - выигрыши Центра и агентов соответственно;  $T$  - период времени.

Статическая теория механизмов стимулирования изложена в (Новиков 2007). Рассуждая по аналогии, естественно предположить, что оптимальный механизм стимулирования для модели (4.2.2)-(4.2.6) должен иметь вид

$$s_i^*(u(t), x(t)) = \begin{cases} h_i(u_i^*(t), u_{-i}(t), x(t)), & u_i(t) = u_i^*(t), \\ 0, & \text{otherwise, } i \in N, \end{cases} \quad (4.2.7)$$

где  $u_{-i} = (u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_n)$ ,

$$u^*(t) \in \text{Arg max}_{u(t)} \int_0^T e^{-\rho t} [H(u(t), x(t)) - \sum_{i \in N} h_i(u(t), x(t))] dt. \quad (4.2.8)$$

Первый результат в этом направлении получен в статье (Рохлин, Угольницкий 2018), где сформулированная гипотеза подтверждена для динамической версии базовой задачи стимулирования "Центр-агент" на бесконечном периоде с дискретным временем и функционалами выигрыша Центра и агента в виде математических ожиданий.

Дополним теперь модель (4.2.2)-(4.2.6) условием гомеостаза

$$\forall t \ x(t) \in X^* \quad (4.2.9)$$

и введём множества

$$U^* = \{u(t) : \forall t \ x(t) \in X^*\}, U^{SM} = U^* \cap NE,$$

предполагая, что  $U^* \neq \emptyset, U^{SM} \neq \emptyset$ , где  $NE$  - множество равновесий Нэша в дифференциальной игре агентов в нормальной форме (4.2.4)-(4.2.6).

Рассмотрим механизм стимулирования

$$s_i^{SM}(u(t), x(t)) = \begin{cases} h_i(u_i^{SM}(t), u_{-i}(t), x(t)), & u_i(t) = u_i^{SM}(t), u^{SM}(t) \in U^{SM}, \\ 0, & \text{otherwise, } i \in N. \end{cases} \quad (4.2.10)$$

Тогда задача исследования заключается в построении множества  $U^{SM}$  и сравнительном анализе механизмов стимулирования (4.2.7) и (4.2.10). Для построения множества  $U^{SM}$ , а также решения задачи оптимального управления (4.2.8), наряду с традиционными методами целесообразно использовать метод КРС имитационного моделирования (Ougolnitsky and Usov 2018).

*Динамическая модель контроля рабочего времени на основе метода принуждения.* Модель имеет вид

$$J_0 = \int_0^T e^{-\rho t} [s_0(t)c(x(t)) - D(q(t))] dt \rightarrow \max \quad (4.2.11)$$

$$0 \leq q_i(t) \leq 1, i \in N; \quad (4.2.12)$$

$$J_i = \int_0^T e^{-\rho t} [p_i(1 - u_i(t)) + s_i(t)c(x(t))] dt \rightarrow \max \quad (4.2.13)$$

$$q_i(t) \leq u_i(t) \leq 1, i \in N; \quad (4.2.14)$$

$$\dot{x} = \sum_{i=1}^n k_i u_i(t) - \mu x(t), x(0) = x_0 \quad (4.2.15)$$

с учётом условия гомеостаза (4.2.9).

Как и модель (4.2.2)-(4.2.6), данная модель описывает отношения в древовидной системе управления Центр - агенты. Переменная управления агента  $u_i$  имеет смысл, например, доли рабочего времени, уделяемой работе над научным проектом. Фазовая переменная  $x$ , как и в модели (4.2.2)-(4.2.6), может трактоваться как число публикаций или другой показатель НИР.

Уравнение динамики (4.2.15) для простоты имеет линейную структуру, где  $k_i$



- коэффициент перевода рабочего времени агента в число публикаций,  $\mu$  - условный коэффициент "амортизации" (снижения числа публикаций в отсутствие усилий агентов).

Специфику модели (4.2.11)-(4.2.15) определяет подынтегральная функция в формуле (4.2.13). Предполагается, что каждый агент делит своё время между участием в научном проекте (доля  $u_i$ ) и другими видами активности (доля  $1 - u_i$ ). Участие в других видах активности приносит агенту суммарный доход, описываемый возрастающей вогнутой функцией  $p_i$ . Совокупное участие агентов в проекте порождает доход  $c(x)$  (возрастающая вогнутая функция), распределяемый между Центром и агентами с помощью заданных переменных коэффициентов  $s_i(t), i = 0, 1, \dots, n$ . Таким образом, возникает динамическая модель согласования общественных и частных интересов (СОЧИ-модель) (параграф 2.3). Переменная управления Центра  $q_i$  ограничивает "эгоизм" агента снизу, что соответствует методу принуждения. В этом случае во избежание тривиального решения  $q_i = 1, i=1, \dots, n$ , следует учитывать затраты Центра на административный контроль  $D(q)$  (возрастающая выпуклая функция,  $D(0) = 0$ ).

Модель (4.2.9), (4.2.11)-(4.2.15) можно исследовать для различных информационных регламентов. Если фиксировать значения  $q_i(t)$ , то возникает дифференциальная игра агентов в нормальной форме (4.2.13)-(4.2.15) с фазовыми ограничениями (4.2.9), решением которой естественно считать равновесие Нэша. В иерархической постановке (4.2.11)-(4.2.15), (4.2.9) решение есть равновесие Штакельберга, при этом стратегии Центра могут быть программными  $q_i(t)$ , позиционными  $q_i(t, x(t))$ , без обратной связи по управлению или с обратной связью  $q_i(t, u(t))$ . Проводить сравнительный анализ эффективности указанных методов управления удобно с помощью индексов системной согласованности вида (4.2.1).

Кроме того, на основе дифференциальной игры в нормальной форме (4.2.13)-(4.2.15) можно различными способами построить игру в форме характеристической функции для последующего анализа принципов оптимальности распределения выигрыша максимальной коалиции между игрока и их динамической устойчивости (Петросян и др. 2012).

*Динамическая модель продвижения инноваций на основе метода побуждения.* Модель имеет вид

$$J_0 = \int_0^T e^{-\rho t} s_0(t) c(x(t)) dt \rightarrow \max \quad (4.2.16)$$

$$\sum_{i=0}^n s_i(t) = 1, \quad 0 \leq s_i(t) \leq 1, \quad i = 0, 1, \dots, n; \quad (4.2.17)$$

$$J_i = \int_0^T e^{-\rho t} [p_i(1 - u_i(t)) + s_i(t)c(x(t))] dt \rightarrow \max \quad (4.2.18)$$

$$0 \leq u_i(t) \leq 1, \quad i = 1, \dots, n; \quad (4.2.19)$$

$$\dot{x} = \sum_{i=1}^n k_i u_i(t) - \mu x(t), \quad x(0) = x_0 \quad (4.2.20)$$

с учётом условия гомеостаза (4.2.9).

Как и модель (4.2.11)-(4.2.15), это динамическая СОЧИ-модель.

Отличие от предыдущей модели заключается в использовании метода побуждения. Центр выбирает и сообщает агентам значения своих переменных управления  $s_i(t)$ , т.е. их доли в распределении суммарного дохода. Принуждение отсутствует, поэтому затраты на контроль равны нулю, а переменные управления агентов  $u_i(t)$  (доля времени или иного ресурса, выделяемая на продвижение инноваций в университете) не ограничены снизу. Используемые при анализе информационные регламенты те же, что и в предыдущей модели.

*Двойственные задачи дискретного программирования для определения оптимальной кадровой структуры факультета.* Представим кадровую структуру некоторого факультета набором

$$N = \{N_p, N_d, N_a\} = \left\{ \{N_{pj}\}_{j=1}^m, \{N_{dj}\}_{j=1}^m, \{N_{aj}\}_{j=1}^m \right\},$$

где  $N_p, N_d, N_a$  - множества профессоров, доцентов и ассистентов (вместе со старшими преподавателями) соответственно;  $m$  - число кафедр на факультете;  $j$  - номер кафедры.

Тогда можно сформулировать следующие двойственные задачи дискретного программирования:

$$\left\{ \begin{array}{l} \sum_{j=1}^m \left[ \sum_{i=1}^{k_1} \lambda_{ij} x_{ij} + \sum_{i=1}^{k_2} \mu_{ij} y_{ij} + \sum_{i=1}^{k_3} \nu_{ij} z_{ij} \right] \rightarrow \max_{n_p, n_d, n_a \in Z_+} \end{array} \right. \quad (4.2.21)$$

$$\left\{ \begin{array}{l} s_p n_p + s_d n_d + s_a n_a = R \end{array} \right. \quad (4.2.22)$$

$$\left\{ \begin{array}{l} s_p n_p + s_d n_d + s_a n_a \rightarrow \min_{n_p, n_d, n_a \in Z_+} \end{array} \right. \quad (4.2.23)$$

$$\left\{ \begin{array}{l} x_{ij} \geq x_i^*, i = 1, \dots, 6; y_{ij} \geq y_i^*, i = 1, \dots, 4; z_{ij} \geq z_i^*, i = 1, 2; j = 1, \dots, m. \end{array} \right. \quad (4.2.24)$$

При этом используется дополнительное условие

$$w_{ij} = f_i(n_{pj}, n_{dj}, n_{aj}) = \begin{cases} w_i^H, & (n_{pj}, n_{dj}, n_{aj}) \in N_i^H, \\ w_i^L, & \text{otherwise}, i = 1, \dots, k_i; \end{cases} \quad w_{ij} \in \{x_{ij}, y_{ij}, z_{ij}\}, \quad (4.2.25)$$

которое позволяет говорить о парах задач (4.2.21)-(4.2.22) и (4.2.23)-(4.2.24) как о двойственных. Здесь  $x_{ij}, y_{ij}, z_{ij}$  - показатели НИР, УМР и ИПР (таблица 4.2.1);  $x_i^*, y_i^*, z_i^*$  - их гомеостатические значения;  $\lambda_{ij}, \mu_{ij}, \nu_{ij}$  - соответствующие весовые коэффициенты;  $i$  - номер показателя;  $k_1, k_2, k_3$  - число показателей каждой группы;  $s_p, s_d, s_a$  - заработная плата соответствующих категорий преподавателей;  $R$  - суммарный фонд зарплаты факультета;  $w_i^L, w_i^H$  - высокое и

низкие значения показателей  $\{x_{ij}, y_{ij}, z_{ij}\}$  ( $w_i^H > w_i^L$ );  $N_i^H$  - множество значений  $(n_p, n_d, n_a)$ , обеспечивающих достижение  $w_i^H$ ;  $k_l \in \{k_1, k_2, k_3\}$ ;  $n_z = |N_z|$ ,  $z \in \{p, d, a\}$ .

В частности, данная постановка позволяет оценить эффективность попарного объединения кафедр факультета, весьма актуального в условиях сокращения штатной численности профессорско-преподавательского состава. Для этого нужно рассмотреть  $C_m^2$  возможных объединений кафедр  $k$  и  $l$ ,  $k, l = 1, \dots, m$ , приводящих к значениям численности  $n_{pk} + n_{pl}, n_{dk} + n_{dl}, n_{ak} + n_{al}$ , и решить задачи (4.2.21)-(4.2.22), (4.2.23)-(4.2.24) с учётом (4.2.25).

Основную сложность здесь представляет идентификация функции  $f_i$  в формуле (4.2.25), требующая привлечения субъективных соображений и экспертных оценок. В дальнейшем задачи решаются с использованием стандартных методов дискретного программирования (Parker and Rardin 1988) и компьютерной имитации.

*Задача распределения вознаграждения между исполнителями научного проекта как статическая игра в форме характеристической функции.* Основная задача теории кооперативных игр (игр в форме характеристической функции) - распределение дохода максимальной коалиции (grand coalition) между отдельными игроками. Поэтому данную математическую модель естественно использовать для анализа возможных вариантов распределения вознаграждения между исполнителями проекта.

Обозначим  $N = N_p \cup N_d \cup N_a$ ,  $N_i \cap N_j = \emptyset$ , где  $N$  - множество исполнителей научного проекта;  $N_p, N_d, N_a$  - его непересекающиеся подмножества профессоров, доцентов и ассистентов (вместе со старшими преподавателями). Будем считать, что каждый исполнитель  $i \in N_m$  может самостоятельно подготовить  $a_i = a_{im}$  публикаций в изданиях из списка Web of Science/Scopus и  $b_i = b_{im}$  публикаций из перечня ВАК,  $m \in \{p, d, a\}$ . Тогда значение характеристической функции исполнителя можно представить как

$$v(i) = ca_i + b_i, c > 0, i \in N. \quad (4.2.26)$$

Это порождает следующие задачи исследования:

- полностью определить супераддитивную характеристическую функцию  $v: 2^N \rightarrow R$  различными возможными способами

$$v(K) = f(n_p^K, n_d^K, n_a^K), n^K = |N^K|, K \subseteq N;$$

- в построенной игре в форме характеристической функции найти решения, соответствующие различным принципам оптимальности (С-ядро, устойчивые множества, вектор Шепли, нуклеолус и т.д.);

- провести сравнительный анализ полученных результатов для различных значений параметров  $c, v(N), a_{im}, b_{im}$ ;

- проверить выполнение условий гомеостаза

$$\sum_{i \in N} a_i \geq a^*, \sum_{i \in N} b_i \geq b^*.$$

Авторская концепция моделирования коррупции изложена в монографии (Gorbaneva et al. 2016). В качестве примера модели борьбы с коррупцией рассмотрим простую модель коррупции при сдаче экзамена как игру в развёрнутой форме (Петросян и др. 2012). В игре участвуют преподаватель (P) и студент (S). В процессе экзамена преподаватель выясняет квалификацию студента: заслуживает он(а) положительной оценки (H - высокая квалификация) или нет (L - низкая квалификация).

Если студент имеет квалификацию H, то преподаватель может выбрать одну из двух стратегий: e - вымогательство (все равно требовать деньги под угрозой двойки); h - честность (поставить заслуженную оценку).

Если студент имеет квалификацию L, то преподаватель тоже имеет две стратегии: c - попустительство (предложить поставить положительную оценку за деньги); h - честность (отправить на пересдачу).

Если преподаватель выбирает честность, то игра заканчивается. Выигрыш преподавателя равен нулю, а выигрыш студента при квалификации H равен 5, при квалификации L равен -2. Если преподаватель предлагает

коррупционную схему, то у студента есть две стратегии: *b* - согласиться на взятку; *a* - подать жалобу.

Будем считать, что при подаче жалобы преподавателя увольняют, а студент получает заслуженную оценку минус некоторые затраты. Итоговые оценки выигрышей и полный граф игры показаны на рис. 4.2.1.

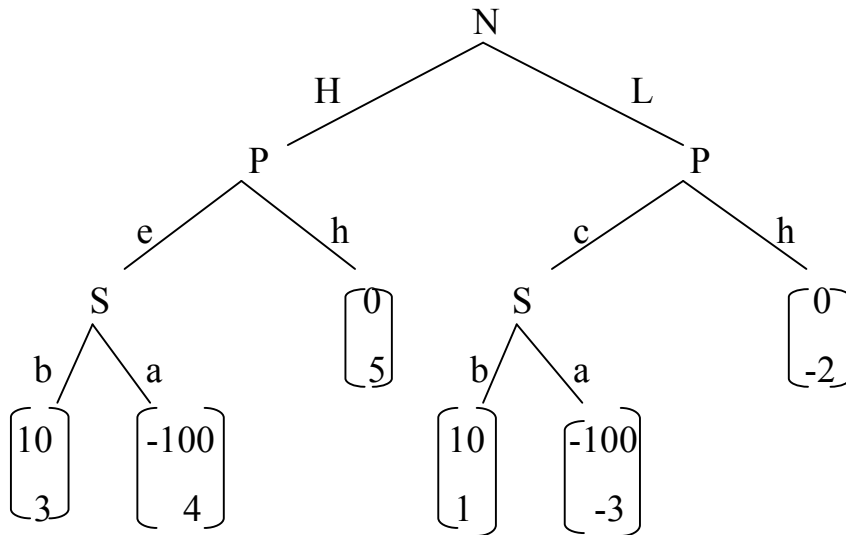


Рис. 4.2.1. Модель коррупции при сдаче экзамена как игра в развёрнутой форме

Решение игры ищется методом обратной индукции (Петросян и др. 2012). Очевидно, что при данных значениях выигрышей, если студент имеет высокую квалификацию, то преподавателю выгодна честность, а если низкую, то преподавателю выгодно предложить взятку, а студенту принять это предложение.

Вместо игры в развёрнутой форме можно изучить данную ситуацию с точки зрения преподавателя, считая, что он может оценивать условные вероятности выбора студентом стратегий взятки или жалобы в зависимости от его квалификации. Тогда при представленных на рис. 4.2.1 значениях параметров выигрыш преподавателя при вымогательстве есть

$$J^e = 10P_{b|H} - 100P_{a|H} = 10(P_{b|H} - 10P_{a|H}),$$

а при попустительстве

$$J^c = 10P_{b|L} - 100P_{a|L} = 10(P_{b|L} - 10P_{a|L}),$$

где  $P$  с индексами обозначает соответствующие условные вероятности.

Коррупция в обоих случаях невыгодна преподавателю (и это можно считать условием гомеостаза), если  $P_b < 10P_a$ , а в общем случае  $w_b P_b < w_a P_a$ , где  $w_a, w_b$  - выигрыши преподавателя при жалобе и взятке соответственно. Вероятностные распределения можно оценивать на множестве студентов  $N$  на основе эмпирических данных.

В качестве ещё одного примера модели борьбы с экономической коррупцией рассмотрим *динамическую модель грантового финансирования*. Пусть  $N = \{1, \dots, n\}$  - множество участников конкурсов, готовых давать взятки;  $M = \{1, \dots, m\}$  - множество конкурсов, любой  $i \in N$  может участвовать в любом  $j \in M$ . Объем финансирования по каждому конкурсу  $j \in M$  позволяет поддержать  $k_j \in Z$  проектов с фиксированной суммой гранта  $G_j$ .

В начале каждого года планирования  $t = 1, \dots, T$ :

1) участники  $i \in N$  выбирают свои стратегии  $b_{ij}^t$  - доли "отката":

$$0 \leq \sum_{j=1}^m b_{ij}^t \leq 1, b_{ij}^t \geq 0, i \in N, j \in M, t = 1, \dots, T; \quad (4.2.27)$$

пусть  $b_{1j}^t > b_{2j}^t > \dots > b_{nj}^t, j \in M$ ;

2) случайным образом выбираются целые числа  $\bar{k}_j^t \in [0, k_j^t], j \in M$ , после чего участники  $1, 2, \dots, \bar{k}_j^t$  получают гранты; остальные гранты  $k_j^t - \bar{k}_j^t$  распределяются между  $i \notin N$  и в дальнейшем рассмотрении не участвуют.

Функционалы выигрыша участников имеют вид

$$J_i = \sum_{t=1}^T e^{-\rho t} \sum_{j=1}^m (1 - b_{ij}^t) r_{ij}^t \rightarrow \max, \quad (4.2.28)$$

где  $r_{ij}^t = \begin{cases} G_j, & \text{участник } i \text{ получает грант по конкурсу } j, \\ 0, & \text{иначе.} \end{cases}$

Значения показателей НИР  $x_1 - x_4$  (см. таблицу 4.2.1) вычисляются согласно разностному уравнению динамики

$$x_{ki}^{t+1} = x_{ki}^t + f_k \left( \sum_{j=1}^m (1 - b_{ij}^t) r_{ij}^t \right), x_{ki}^0 = x_{ki0}, k = 1, 2, 3, 4; i \in N; t = 0, 1, \dots, T - 1. \quad (4.2.29)$$

Условия гомеостаза имеют вид

$$x_{ki}^t \geq x_k^*, k = 1, 2, 3, 4; i \in N; t = 0, 1, \dots, T. \quad (4.2.30)$$

В базовой постановке модель (4.2.27)-(4.2.29) исследуется с помощью метода КРС имитации  $\{b_{ij}^t\}_{i=1, j=1, t=1}^{n, m, T}$ ,  $\{\bar{k}_j^t\}_{j=1, t=1}^{m, T}$ , при этом проверяются условия (4.2.30). Возможна также игровая постановка с одновременным выбором каждым агентом своей стратегии  $b_{ij}^t$  или иерархическая постановка, где Центр выбирает стратегии  $\bar{k}_j^t$ .

Остановимся подробнее на задаче борьбы с экономической коррупцией (Мальсагов и др. 2019). Рассмотрим систему распределения государственным органом (супервайзером) некоторого ресурса между несколькими агентами. Агенты используют выделенный им ресурс как на увеличение общественного блага, так и на личные цели. Имеется ещё один государственный орган (принципал), который контролирует деятельность супервайзера. В системе принят следующий порядок действий: первым определяет свое управление принципал и сообщает его супервайзеру и агентам. Вторым объявляет свое управление супервайзер и сообщает его агентам. Агенты выбирают свои управления, когда выбор остальных субъектов уже известен. Субъекты управления всех уровней стремятся к получению максимального выигрыша (дохода), который определяется их целевыми функционалами. Предполагается, что в системе возможна коррупция. Агенты предлагают супервайзеру взятки, в обмен на которые он увеличивает выделяемые в их распоряжение доли ресурса. Супервайзер выступает в роли взяточника, а агенты – взяткодателей.

Ниже будет рассмотрен случай бескорыстного принципала, который не преследует своих личных целей, а только наказывает супервайзера за взятки



путем штрафа. Супервайзер и агенты стремятся к максимизации своего дохода, их целевые функционалы возьмём в виде:

- супервайзера

$$J_S(\{r_i(\cdot)\}_{i=1}^N, \{b_i(\cdot)\}_{i=1}^N, x(\cdot)) = \int_0^T e^{-\rho t} \left\{ p_0 x(t) + [1 - Mz(t)] \sum_{i=1}^N b_i(t) r_i(t) \right\} dt + g_S(x(T)) \rightarrow \max; \quad (4.2.31)$$

- агентов ( $i = 1, 2, \dots, N$ )

$$J_i(r_i(\cdot), s_i(\cdot), b_i(\cdot), x(\cdot)) = \int_0^T e^{-\rho t} k_i s_i(t) [1 - b_i(t)] r_i(t) dt + g_i(x(T)) \rightarrow \max. \quad (4.2.32)$$

Здесь  $N$  - количество агентов;  $T$  - момент времени, до которого ведётся рассмотрение;  $\rho$  - коэффициент дисконтирования;  $t$  - временная координата;  $x$  - величина общественного блага, производимого агентами;  $p_0$  - установленная государством доля супервайзера в общественном благе;  $r_i(t)$  - управление супервайзера - доля ресурса, выделяемая супервайзером  $i$ -му агенту с учетом полученной взятки;  $b_i(t)$  (управление агента) - взятка ("откат"), предлагаемая  $i$ -м агентом супервайзеру (доля от  $r_i$ );  $s_i(t)$  (управление агента) - доля полученного ресурса  $(1 - b_i(t))r_i(t)$ , которая идёт на личные цели агента;  $z(t)$  - вероятность наказания супервайзера за взятку;  $M$  - коэффициент, характеризующий величину штрафа супервайзера при получении им взятки;  $g_S(x(T))$  - терминальное слагаемое, характеризующее доход супервайзера в зависимости от величины общественного блага в заключительный момент времени;  $k_i$  - коэффициент, характеризующий доход  $i$ -го агента от личной деятельности;  $g_i(x(T))$  - терминальное слагаемое, характеризующее доход  $i$ -го агента в зависимости от величины общественного блага в заключительный момент времени.

Целевые функционалы (4.2.31), (4.2.32) рассматриваются при следующих ограничениях на управления

- супервайзера

$$r_i(t) \geq 0; \sum_{i=1}^N r_i(t) \leq R \quad (4.2.33)$$

- и агентов

$$0 \leq b_i(t) \leq 1; \quad 0 \leq s_i(t) \leq 1; \quad i=1,2,\dots,N . \quad (4.3.34)$$

Здесь  $R$  - величина ресурса, имеющегося в распоряжении супервайзера.

Функционалы (4.2.31), (4.2.32) выражают выигрыш супервайзера и агентов соответственно. При этом предполагается, что супервайзер получает в свое распоряжение долю общественного блага в виде вознаграждения в каждый момент времени, в том числе и в конечный момент, а агенты – только в конечный момент времени. Уравнение, описывающее динамику общественного дохода (блага) возьмём в виде:

$$\dot{x} = -qx + \sum_{i=1}^N a_i \sqrt{(1-s_i(t))(1-b_i(t))r_i(t)}, \quad x(0) = x_0 . \quad (4.2.35)$$

Здесь  $a_i$  - коэффициент, характеризующий производственную функцию  $i$ -го агента;  $q$  - коэффициент, связанный с амортизацией при производстве общественного блага;  $x_0$  - величина общественного блага в начальный момент времени.

Если вести рассмотрение с точки зрения агентов, то в модели (4.2.31) – (4.2.35) необходимо задать управления супервайзера - функции взятки  $r_i(t)$ .

Ниже исследуем случай функций  $r_i(t)$  вида

$$r_i(t) = Rb_i(t); \quad i = 1,2,\dots,N . \quad (4.2.36)$$

В этом случае удастся найти в аналитическом виде стратегии агентов в равновесии Нэша. Итак, решается только задача агентов (4.2.32), (4.2.34), (4.2.35) с учётом (4.2.36). Эта задача представляет собой неантагонистическую игру  $N$  лиц, в которой строится равновесие Нэша. Для определения равновесной стратегии каждого агента решается задача оптимального управления. При этом используем принцип максимума Понтрягина (Dockner et al. 2000). Функция Гамильтона для  $i$ -го агента имеет вид

$$H_i(x(t), b_i(t), s_i(t), \lambda_i(t)) = e^{-\rho t} k_i s_i(t) \frac{b_i(t)[1-b_i(t)]}{R} + \lambda_i \left( -qx + \frac{\sum_{k=1}^N a_k \sqrt{(1-s_k)b_k(t)(1-b_k(t))}}{\sqrt{R}} \right) \quad (4.2.37)$$

Перейдём от задачи (4.2.32), (4.2.34), (4.2.35) с учётом (4.2.26) к

эквивалентной задаче поиска максимума функции Гамильтона (4.2.37) при известных ограничениях на управления (Dockner et al., 2000), которую можно записать в виде системы уравнений ( $i = 1, 2, \dots, N$ ):

$$\frac{\partial H_i}{\partial b_i} = \frac{e^{-\rho t} k_i s_i}{R} (1 - 2b_i(t)) + \frac{\lambda_i}{\sqrt{R}} a_i \sqrt{1 - s_i} \frac{1 - 2b_i(t)}{2\sqrt{(1 - b_i(t))b_i(t)}} = 0; \quad (4.2.38)$$

$$\frac{\partial H_i}{\partial s_i} = e^{-\rho t} k_i \frac{b_i(t)[1 - b_i(t)]}{R} - \lambda_i \frac{a_i \sqrt{b_i(t)(1 - b_i(t))}}{2\sqrt{R(1 - s_i)}} = 0; \quad (4.2.39)$$

$$\frac{dx}{dt} = \frac{\partial H_i}{\partial \lambda_i} = -qx + \frac{\sum_{k=1}^N a_k \sqrt{(1 - s_k)b_k(t)(1 - b_k(t))}}{\sqrt{R}}; \quad x_i(0) = x_0; \quad (4.2.40)$$

$$\frac{d\lambda_i}{dt} = -\frac{\partial H_i}{\partial x} = -q\lambda_i; \quad \lambda_i(T) = \frac{dg_i}{dx(T)}. \quad (4.2.41)$$

В (4.2.38) - (4.2.41)  $\lambda_i(t)$  ( $i = 1, 2, \dots, N$ ) играют роль сопряжённых функций. Требуется найти систему функций  $\{(b_i, s_i, \lambda_i)_{i=1}^N, x\}$ , удовлетворяющую системе уравнений (4.2.38) – (4.2.41) ( $i = 1, 2, \dots, N$ ). Эта система функций будет доставлять максимум функции Гамильтона (4.2.37).

Алгоритм решения задачи (4.2.38) – (4.2.41) состоит в следующем:

1. Решается задача для сопряжённой функции (4.2.41). Отсюда, например, в случае

$$g_i(x(T)) = c_i x(T); \quad c_i = \text{const} \quad (4.2.42)$$

получим, что

$$\lambda_i(t) = c_i e^{q(T-t)}.$$

2. Подставим найденные на первом шаге алгоритма сопряжённые функции в (4.2.38), (4.2.39).

3. Решается система уравнений (4.2.38), (4.2.39). Найденные решения подставляются в (4.2.40).

4. Из обыкновенного дифференциального уравнения (4.2.40) находится переменная состояния системы  $x(t)$ .

Из (4.2.39) получим, что в равновесии Нэша

$$s_i^{NE} = 1 - e^{2\rho t} \frac{\lambda_i^2 a_i^2 R}{4k_i^2 b_i^{NE}(t)[1 - b_i^{NE}(t)]}; i = 1, 2, \dots, N. \quad (4.2.43)$$

Из (4.2.38) получим, что  $b_i^{NE} = 0.5$  или определяется из решения уравнения

$$\frac{e^{-\rho t} k_i s_i}{\sqrt{R}} + \lambda_i a_i \sqrt{1 - s_i} \frac{1}{2\sqrt{(1 - b_i(t))b_i(t)}} = 0.$$

Последнее уравнение с учетом формулы (4.2.43) и содержательного смысла входных параметров модели решений не имеет. Поэтому решение задачи (4.2.32), (4.2.34), (4.2.35) с учётом (4.2.36) имеет вид

$$b_i^{NE} = 0.5; \quad s_i^{NE} = 1 - e^{2(\rho - q)t + 2qT} \frac{c_i^2 a_i^2 R}{k_i}; i = 1, 2, \dots, N. \quad (4.2.44)$$

Переменная состояния системы определяется из уравнения

$$\frac{dx^{NE}}{dt} = -qx^{NE} + \frac{1}{2} e^{(\rho - q)t + qT} \sum_{l=1}^N \frac{c_l a_l^2}{\sqrt{k_l}}.$$

Отсюда, в случае  $a_i, c_i, k_i = \text{const}; (i = 1, 2, \dots, N)$  получим, что

$$x^{NE}(t) = (x_0 - \frac{B}{\rho})e^{-qt} + \frac{B}{\rho} e^{(\rho - q)t}, \quad (4.2.45)$$

где  $B = \frac{1}{2} e^{qT} \sum_{l=1}^N \frac{c_l a_l^2}{\sqrt{k_l}}.$

Итак, равновесие Нэша в игре агентов (4.2.32), (4.2.34), (4.2.35) с учётом (4.2.36) и (4.2.42) выражается формулами (4.2.44), (4.2.45).

*Пример 1.* В случае (сут.-сутки; у.е. – условные единицы для выражения стоимости)  $N = 3; \rho = 0.01; T = 1095 \text{сут}; k_1 = 20; k_2 = 50; k_3 = 100; x_0 = 100 \text{у.е.};$    
 $c_1 = 100 \text{сут}; a_1 = 0.1 \text{сут}^{-1}; a_2 = 0.05 \text{сут}^{-1}; a_3 = 0.02 \text{сут}^{-1}; c_2 = 150 \text{сут}; c_3 = 200 \text{сут}; q = 0.001 \text{сут}^{-1}$   
 равновесие Нэша и доходы субъектов определяются формулами:

$$b_i^{NE}(t) \equiv 0.5; (i = 1, 2, 3); \quad s_{1,2}^{NE}(T/10) = s_{1,2}^{NE}(T/2) = s_{1,2}^{NE}(T) = 0;$$

$$s_3^{NE}(T/10) = s_3^{NE}(T/2) = s_3^{NE}(T) = 0.84 \text{ и } J_1 = 10425 \text{у.е.}; J_2 = 15633 \text{у.е.}; J_3 = 20841 \text{у.е.}$$

Величина общественного блага с течением времени незначительно растёт ( $x^{NE}(T) = 104 \text{у.е.}$ ).

*Пример 2.* В случае входных данных примера 1 и уменьшении доли агентов в общественном благе в 10 раз ( $c_1 = 10$ сут;  $c_2 = 15$ сут;  $c_3 = 20$ сут) стратегии всех субъектов управления меняются, а их доходы падают, в том числе, уменьшается и доход супервайзера. В этом случае

$$b_i^{NE}(t) \equiv 0.5; (i=1,2,3); \quad s_1^{NE}(T/10) = s_1^{NE}(T/2) = s_1^{NE}(T) = 0.73; \quad s_2^{NE}(T/10) = s_2^{NE}(T/2) = s_2^{NE}(T) = 0.97; \quad s_3^{NE}(T/10) = s_3^{NE}(T/2) = s_3^{NE}(T) = 0.98 \text{ и } J_1 = 926 \text{y.e.}; \quad J_2 = 1390 \text{y.e.}; \quad J_3 = 1853 \text{y.e.}$$

Величина общественного блага незначительно уменьшается ( $x^{NE}(T) = 92 \text{y.e.}$ ) по сравнению с примером 1.

*Пример 3.* В случае входных данных примера 1 и уменьшении доходов агентов от частной деятельности в 10 раз ( $k_1 = 2; k_2 = 5; k_3 = 10$ ) выигрыши всех субъектов, в том числе и агентов, возрастают. Агенты все полученные ресурсы вкладывают в общественную деятельность, величина общественного блага при этом значительно растет (примерно на 30%) ( $x^{NE}(T) = 131 \text{y.e.}$ ). В этом случае  $b_i^{NE}(t) \equiv 0.5; (i=1,2,3); \quad s_i^{NE}(T/10) = s_i^{NE}(T/2) = s_i^{NE}(T) = 0; (i=1,2,3)$  и  $J_1 = 13194 \text{y.e.}; \quad J_2 = 19787 \text{y.e.}; \quad J_3 = 26380 \text{y.e.}$

*Пример 4.* В случае входных данных примера 3 при уменьшении производительности агентов при вкладе в общественное благо ( $a_1 = 0.01 \text{сут}^{-1}; \quad a_2 = 0.005 \text{сут}^{-1}; \quad a_3 = 0.002 \text{сут}^{-1}$ ) их доходы по сравнению с примером 3 уменьшаются и  $b_i^{NE} = 0.5; \quad s_{1,2}^{NE}(T/10) = s_{1,2}^{NE}(T/2) = s_{1,2}^{NE}(T) = 0; \quad s_3^{NE}(T/10) = s_3^{NE}(T/2) = s_3^{NE}(T) = 0.83$  и  $J_1 = 9180 \text{y.e.}; \quad J_2 = 13770 \text{y.e.}; \quad J_3 = 18360 \text{y.e.}$

Величина общественного блага по сравнению с примером 1 незначительно уменьшается (примерно на 10%):  $x^{NE}(T) = 92 \text{y.e.}$

С точки зрения супервайзера (4.2.31) - (4.2.35) представляет собой динамическую иерархическую игру  $N$  ведомых (агентов) и одного ведущего (супервайзера). Ниже рассматривается случай, когда игра ведётся в соответствии с информационным регламентом игры Гермейера  $\Gamma_{2t}$  (игры Штакельберга с обратной связью по управлению). Алгоритм построения равновесия в этом случае состоит в следующем.

1. Находится стратегия наказания супервайзером агентов, когда последние не хотят с ним сотрудничать

$$r^P(t) = \{r_i^P(t)\}_{i=1}^N : r_i^P(t) = \arg \min_{\substack{r_i \geq 0; \\ \sum_{i=1}^N r_i \leq R}} J_i(b_i^{NE}(t), s_i^{NE}(t), r_i(t), x(t)).$$

Ясно, что при наказании

$$r_i^P(t) = 0; i = 1, 2, \dots, N.$$

Набор управлений агентов  $\{b_i^{NE}(t), s_i^{NE}(t)\}_{i=1}^N$  есть равновесие Нэша в игре  $N$  агентов при заданной стратегии супервайзера  $\{r_i^P(t) = 0\}_{i=1}^N$ . При этом

$$b_i^{NE}(t) = 0; s_i^{NE}(t) = 1; i = 1, 2, \dots, N.$$

Находятся гарантированные выигрыши агентов, если они отказываются сотрудничать с супервайзером, и он применяет стратегию наказания. В случае (12) имеем:

$$L_i = \min_{\substack{r_i \geq 0; \\ \sum_{i=1}^N r_i \leq R}} J_i(b_i^{NE}(t), s_i^{NE}(t), r_i^P(t), x(t)) = x_0 c_i e^{-qT}; i = 1, 2, \dots, N.$$

2. Решается задача оптимального управления (4.2.31), (4.2.33) - (4.2.35) с дополнительными условиями

$$L_i = x_0 c_i e^{-qT} < J_i(b_i(t), s_i(t), r_i(t), x(t)); i = 1, 2, \dots, N.$$

Эти условия делают для агентов стратегию поощрения со стороны супервайзера более выгодной по сравнению со стратегией наказания.

При этом максимум в (4.2.31) ищется одновременно по  $(3N)$  функциям  $(r_i(t), s_i(t), b_i(t)); i = 1, \dots, N$ . Решение указанной задачи оптимального управления обозначим  $\{r_i^R(t), b_i^R(t), s_i^R(t)\}_{i=1}^N$ , где  $r_i^R(t)$  - стратегия поощрения  $i$ -го агента супервайзером при выборе им в качестве своих управлений  $b_i^R(t)$  и  $s_i^R(t)$ .

1. Супервайзер предъявляет каждому агенту стратегию с обратной связью по его управлению:

$$r_i(t) = \begin{cases} r_i^R(t), & \text{если } b_i(t) = b_i^R(t) \text{ и } s_i(t) = s_i^R(t) \text{ для } \forall t \in [0, \infty), \\ r_i^P(t), & \text{иначе.} \end{cases}$$

При экономически разумных агентах равновесие имеет вид

$$\{r_i^R(t), b_i^R(t), s_i^R(t)\}_{i=1}^N.$$

Решение задачи из для входных функций общего вида ищется численно путем дискретизации (Угольниковский, Усов 2013) и методом КРС имитационного моделирования (Ougolnitsky, Usov 2018).

Как обычно, предполагается, что управляющие воздействия всех субъектов остаются постоянными в течение некоторых промежутков времени, т.е. можно считать, что

$$A(t) = \begin{cases} a_1, & \text{если } 0 \leq t < t_1, \\ a_2, & \text{если } t_1 \leq t < t_2, \\ \dots & \\ a_N, & \text{если } t_{N-1} \leq t < T, \end{cases} \quad (4.2.46)$$

где под функцией  $A(t)$  понимаются управления как агентов (функции  $b_i(t), s_i(t); i = 1, 2, \dots, N$ ), так и супервайзера (функции  $r_i(t); i = 1, 2, \dots, N$ );  $a_j = \text{const}$ ;  $t_j = j \Delta t$ ;  $\Delta t = T / K$ ;  $j = 1, 2, \dots, K$ ;  $K$  — число интервалов постоянства управлений субъектов. С учётом (4.2.46) получим, что — функционал (4.2.31) превратится в целевую функцию вида

$$J_S(\{r_{ij}\}_{i,j=1}^{N(K)}, \{b_{ij}\}_{i,j=1}^{N(K)}, x(\cdot)) = \sum_{j=1}^K \int_{t_{j-1}}^{t_j} e^{-\rho t} p_0 x(t) dt + \sum_{j=1}^K \left( \sum_{i=1}^N b_{ij} r_{ij} \int_{t_{j-1}}^{t_j} e^{-\rho t} [1 - Mz(t)] dt \right) + g_S(x(T)) \rightarrow \max \quad (4.2.47)$$

- функционалы (4.2.32) в целевые функции ( $i = 1, 2, \dots, N$ ):

$$J_i(\{r_{ij}\}_{j=1}^K, \{b_{ij}\}_{j=1}^K, \{s_{ij}\}_{j=1}^K, x(\cdot)) = \sum_{j=1}^K [1 - b_{ij}] r_{ij} s_{ij} \int_{t_{j-1}}^{t_j} e^{-\rho t} k_i dt + g_i(x(T)) \rightarrow \max \quad (4.2.48)$$

- ограничения на управления супервайзера (4.2.33) примут вид

$$r_{ij} \geq 0; \sum_{i=1}^N r_{ij} \leq R; \quad j = 1, 2, \dots, K; \quad (4.2.49)$$

- ограничения на управления агентов (4.2.34) ( $i = 1, 2, \dots, N$ ):

$$0 \leq b_{ij} \leq 1; 0 \leq s_{ij} \leq 1; \quad j = 1, 2, \dots, K; \quad (4.2.50)$$

Уравнение динамики (4.2.35) останется без изменения.

Таким образом, задача, решаемая на шаге 3 алгоритма, сводится к задаче максимизации целевой функции (4.2.47) сразу по  $3 \cdot K \cdot N$  переменным  $\{r_{ij}\}_{i,j=1}^{N(K)}, \{b_{ij}\}_{i,j=1}^{N(K)}, \{s_{ij}\}_{i,j=1}^{N(K)}$  с учётом (4.2.49), (4.2.50) и дополнительным условием

$$x_0 c e^{-qT} < \sum_{j=1}^K [1 - b_{ij}] r_{ij} \int_{t_{j-1}}^{t_j} e^{-\rho t} k_i s_i(t) dt + g_i(x(T)). \quad (4.2.51)$$

Аналитически решить эту задачу в случае входных функций общего вида не удастся, поэтому задача решается методом КРС имитационного моделирования. Множество КРС в данном случае есть декартово произведение  $N \times K$  множеств:

$$QRS = QRS_{11} \times QRS_{12} \times \dots \times QRS_{1K} \times QRS_{21} \times \dots \times QRS_{NK}. \quad (4.2.52)$$

Множество  $QRS_{ij}$  есть множество КРС в момент времени  $j = 1, 2, \dots, K$  (на  $j$ -й год) для супервайзера и  $i$ -го агента. Это множество содержит тройки элементов  $(r_{ij}^{QRS}, b_{ij}^{QRS}, s_{ij}^{QRS})$ .

Построим множество КРС следующим образом: перенумеруем всех агентов от 1 до  $N$  и примем, что множество  $QRS_{ij}$  в каждый момент времени  $j$  по отношению к каждому агенту  $i$  содержит 27 элементов:

$$QRS_{ij} = \left\{ (r_{ij}^{QRS}, b_{ij}^{QRS}, s_{ij}^{QRS}) : r_{ij}^{QRS} = \left\{ 0; \frac{1}{2} \left( R - \sum_{k=1}^{i-1} r_{kj} \right); R - \sum_{k=1}^{i-1} r_{kj} \right\}; \right. \quad (4.2.53)$$

$$\left. b_{ij}^{QRS} = \left\{ 0; \frac{1}{2}; 1 \right\}; s_{ij}^{QRS} = \left\{ 0; \frac{1}{2}; 1 \right\} \right\} (j = 1, 2, \dots, K; i = 1, 2, \dots, N).$$

Построенное таким образом множество  $QRS$  после задания всех входных параметров модели проверяется на достаточность и избыточность (Ougolnitsky, Usov 2018), и при необходимости пополняется или сужается.

Изначально множество КРС содержит  $m = \prod_{i,j=1}^{N(K)} |QRS_{ik}| = 27^{K \cdot N}$  элементов.

Мощность множества КРС велика, но оно конечно и, как показали численные эксперименты, допускает полный перебор.



Численный алгоритм решения задачи на шаге 3 алгоритма состоит в следующем:

1. Задаются вид и значения всех входных функций и параметров задачи (величины  $N, K, R, T, x_0, c, q, \rho, M, z, p_0, g_S, g_i, k_i (i = 1, 2, \dots, N)$ ).
2. Начальное множество КРС вида (4.2.52), (4.2.53) проверяется на достаточность и избыточность, и при необходимости пополняется или сужается.
3. Фиксируется очередная ( $l$ -я) стратегия из  $QRS : \left( \left\{ r_{ij}^{QRS}, b_{ij}^{QRS}, s_{ij}^{QRS} \right\}_{i,j=1}^{N(K)} \right)^{(l)}$ .

В начале выполнения алгоритма  $l=1$ .

4. Определённая на шаге 3 стратегия из  $QRS$  подставляется в (4.2.35). Численно, например, методом конечных разностей находятся значения функции  $x(t)$ . Затем вычисляется значение целевой функции супервайзера (4.2.47). При условии выполнения (4.2.51) большее значение целевой функции (4.2.47), а также набор управлений, его доставляющих, сохраняется.
5. Если просмотрены не все КРС-стратегии, то необходимо перейти к следующей стратегии ( $l:=l+1$ ) и вернуться на шаг 3 алгоритма.
6. После просмотра всех КРС-стратегий (всего их  $27^{N \cdot K}$ ), определяется стратегия, доставляющая наибольшее значение целевой функции (4.2.47) при выполнении (4.2.51), то есть

$$\left( \left\{ r_{ij}^{QRS}, b_{ij}^{QRS}, s_{ij}^{QRS} \right\}_{i,j=1}^{N(K)} \right)^{(*)} = \arg \max_{\left\{ r_{ij}^{QRS}, b_{ij}^{QRS}, s_{ij}^{QRS} \right\}_{i,j=1}^{N(K)} \in QRS} J_S \left( \left\{ r_{ij} \right\}_{i,j=1}^{N(K)}, \left\{ b_{ij} \right\}_{i,j=1}^{N(K)}, x(\cdot) \right)$$

Найденная на шаге 6 алгоритма стратегия и будет стратегией поощрения супервайзером агентов.

Имитационные эксперименты проводились на компьютере с микропроцессором А10 серии Intel Pentium G4620 с оперативной памятью 4 Гб на объектно-ориентированном языке программирования C# согласно приведённому выше алгоритму. Среднее время одного имитационного эксперимента для указанных ниже входных данных составило семь часов. Основную часть этого времени занимает численное решение уравнения

динамики (4.2.35) по явной схеме метода конечных разностей для каждой КРС-стратегии.

*Пример 5.* В случае  $N = 2$ ;  $\rho = 0.01$ ;  $T = 1095$ сут;  $k_1 = 20$ ;  $k_2 = 50$ ;  $R = x_0 = 100$ у.е.;  $C_0 = 20$ сут;  $a_1 = 0.1$ сут<sup>-1</sup>;  $a_2 = 0.05$ сут<sup>-1</sup>;  $c_1 = 100$ сут;  $c_2 = 150$ сут;  $M = 1$ ;  $z = 0.05$ ;  $p_0 = 0.2$ ;  $q = 0.001$ сут<sup>-1</sup> получим, что:  $r_{1,2}(T/10) = r_{1,2}(T/2) = r_{1,2}(T) = 50$ у.е.;  $s_{1,2}^{NE}(T/10) = s_{1,2}^{NE}(T/2) = s_{1,2}^{NE}(T) = 0$ ;  $b_2^{NE}(T/10) = b_2^{NE}(T/2) = b_2^{NE}(T) = 1$  и  $J_0 = 11877$ у.е.;  $J_1 = 10514$ у.е.;  $J_2 = 15775$ у.е.

Размер общественного блага с течением времени незначительно растёт ( $x^{NE}(T) = 104$ у.е.).

Супервайзер, используя информационный регламент игры Гермейера  $\Gamma_{2t}$ , заставляет агентов давать взятки и тратить все полученные от него ресурсы только на общественные цели. Отметим, что в этом случае производственная функция первого агента более эффективна, он приносит обществу больше пользы, чем второй агент. Поэтому супервайзер со второго агента получает максимальную взятку, а к первому агенту относится “бережно”, требуя “только” половину полученного ресурса в качестве отката. Супервайзеру выгодно оставить первому агенту половину ресурса, чтобы он потратил оставленный ему ресурс на общественные цели. Коррупция крайне невыгодна для агентов, потому что носит характер вымогательства.

*Пример 6.* В случае входных данных примера 5 и увеличения в 10 раз доходов агентов от частной деятельности ( $k_1 = 200$ ;  $k_2 = 500$ ) стратегии субъектов и их выигрыши не меняются по сравнению с примером 5. Для агентов в этом случае становится выгоднее использовать выделенный им ресурс на частные цели, но это невыгодно супервайзеру. В силу используемого информационного регламента, угрожая применением стратегии наказания, супервайзер вынуждает агентов не менять свои стратегии по сравнению с примером 5.

*Пример 7.* В случае входных данных примера 5 и уменьшения доли агентов в общественном благе в 10 раз ( $c_1 = 10$ сут;  $c_2 = 15$ сут) или их производительности ( $a_1 = 0.01$ сут<sup>-1</sup>;  $a_2 = 0.005$ сут<sup>-1</sup>) оптимальные стратегии

субъектов не меняются. Выигрыш супервайзера в первом случае не меняется, во втором - уменьшается, выигрыши агентов резко падают и равны  $J_1 = 1051 \text{ y.e.}$ ;  $J_2 = 1577 \text{ y.e.}$  и  $J_0 = 11435 \text{ y.e.}$ ;  $J_1 = 9276 \text{ y.e.}$ ;  $J_2 = 13915 \text{ y.e.}$  соответственно.

*Пример 8.* В случае входных данных примера 5 и увеличения вероятности поимки взяточника до 65% ( $z = 0.65$ ) при незначительном наказании за нее  $M = 1$  коррупция в системе остается. Оптимальные стратегии супервайзера и агентов не меняются по сравнению с примером 5, выигрыши агентов тоже не меняются. Выигрыш супервайзера задается формулой:  $J_0 = 7167 \text{ y.e.}$

*Пример 9.* Дальнейшее по сравнению с примером 8 ужесточение борьбы с коррупцией (увеличение коэффициента, характеризующего величину штрафа супервайзера при получении им взятки – уже значение  $M = 1.3$ ) приводит к ликвидации коррупции в системе. Брать взятки супервайзеру становится невыгодным. Оптимальные стратегии субъектов управления и их выигрыши задаются формулами:  $r_{1,2}(T/10) = r_{1,2}(T/2) = r_{1,2}(T) = 50 \text{ y.e.}$ ;  
 $b_{1,2}^{NE}(T/10) = b_{1,2}^{NE}(T/2) = b_{1,2}^{NE}(T) = 0$ ;  $s_{1,2}^{NE}(T/10) = s_{1,2}^{NE}(T/2) = s_{1,2}^{NE}(T) = 0$  и  $J_0 = 6213 \text{ y.e.}$ ;  
 $J_1 = 13324 \text{ y.e.}$ ;  $J_2 = 18127 \text{ y.e.}$

Информационный регламент игры Гермейера  $\Gamma_{2t}$  делает интересы супервайзера приоритетными при построении равновесия. Интересы агентов фактически не учитываются. Моральные издержки от наличия коррупции в системе распределения ресурсов в рассмотренной модели также не учитываются. Поэтому успех борьбы с коррупцией в рамках предложенной модели целиком зависит от того, выгодно супервайзеру (в смысле его целевого функционала) брать взятки от агентов или нет. Для искоренения коррупции в системе необходимо увеличить размер наказания за взятки и вероятность поимки взяточника (примеры 8,9)). При отсутствии коррупции в системе выигрыш супервайзера падает (примерно на 20 % для используемых выше входных данных), а выигрыши агентов увеличиваются (Мальсагов и др. 2019).

### **4.3. Информационно-аналитическая система управления водными ресурсами региона**

Рассмотрим в качестве примера проект информационно-аналитической системы управления водными ресурсами региона, ориентированный на Ростовскую область (Пучкин и др. 2016).

Проблемы водоснабжения, в т.ч. качества воды, чрезвычайно актуальны для Ростовской области и находятся в центре внимания Института водных проблем РАН, Южного федерального университета, Южного научного центра РАН и ряда других научных и проектных организаций Ростовской области и Юга России. Можно перечислить следующие объективные и субъективные причины этих проблем, особенно острых в шахтёрских и юго-восточных районах области (Ключко 2016):

- 1) ежегодное уменьшение количества выпадающих осадков;
- 2) горение торфяников вверх по Дону;
- 3) основной водозабор в области осуществляется из поверхностных источников, поэтому высока протяжённость коммунальных сетей, которые существенно изношены и испытывают постоянные протечки;
- 4) не проведена консервация шахт, из-за чего вода не собирается сверху, а просачивается вниз и уходит далеко в стороны;
- 5) идет активная застройка территорий вдоль Дона, что увеличивает водозабор, в т.ч. незаконный;
- 6) низкий уровень оборотных запасов воды на предприятиях.

Нежелательными последствиями становятся нехватка качественной пресной воды для бытового водопотребления, обмеление Дона и снижение судоходства, проблемы с работой Новочеркасской ГРЭС и других промышленных предприятий.

Ростовская область находится в первой пятёрке субъектов РФ по использованию свежей воды (в 2013 г. – 2247 млн. м<sup>3</sup>). Ежегодный забор воды в бассейне реки Дон составил в 2012-2015 годах около 5 млрд. м<sup>3</sup>, а в

бассейне Азовского моря в целом – более 15 млрд. м<sup>3</sup>. При этом потери воды при транспортировке в бассейне р. Дон в 2013 г. составили 25,0%, а в 2014 г. – 25,2% по отношению к объёму используемой свежей воды. Ежегодные объёмы сброса загрязнённых сточных вод остаются высокими и составляют 1,3 – 1,35 %.

Современное состояние водных ресурсов Ростовской области и их использование детально охарактеризованы, например, в работах (Айдаркина 2012; Меринова и Хованский 2014). На территории Ростовской области экологический мониторинг состояния водных объектов осуществляется в соответствии с "Программами по проведению государственного мониторинга качества водных объектов" следующими службами: ФГУ «Донводинформцентр»; ФГУ «Управление водными ресурсами Цимлянского водохранилища»; ФГБУ «Ростовский ЦГМС-Р»; Ростовский областной Центр по гидрометеорологии и мониторингу окружающей среды. Координацию действий осуществляет Донское бассейновое водное управление. Мониторинг осуществляется ежегодно в 38 пунктах, 52 створах государственной сети наблюдений, которые расположены на 17 реках и двух водохранилищах области. Наблюдения проводятся по 37 показателям.

Биологические подходы к оценке эколого-токсикологического состояния водных экосистем охарактеризованы в статье (Бакаева и Никаноров 2015) и развиты применительно к поверхностным водам бассейна р. Дон в статье (Бакаева и др. 2015), при этом обсуждаются прикладные аспекты применения методик биотестирования, установленных нормативно-методическими документами. В работе (Закруткин и др. 2014) анализируется загрязнение подземных вод в углепромышленных районах Восточного Донбасса. Уровень солевого загрязнения Пролетарского и Веселовского водохранилищ изучается в (Коханистая и Хоружая 2015).

Моделированию водных потоков посвящена монография (Коханенко и др. 2007). Оценка загрязнения воды на основе метода Лагранжа предложена в статье (Arkhipov et al. 2015).

Итак, проблемы водоснабжения имеют комплексный характер и требуют прежде всего решения задач управления и их информационной поддержки. В этой связи необходимы разработка и внедрение информационно-аналитической системы управления водными ресурсами Ростовской области (ИАС ВР РО).

Концепция имитационной, или информационно-аналитической системы (ИАС) была выдвинута сотрудниками Вычислительного центра АН СССР под руководством академика Н.Н. Моисеева (Моисеев и др. 1973; Павловский 2000). Основная идея ИАС состоит в объединении в программном комплексе моделей и данных и организации диалога с пользователем. Современное состояние таких систем в области природопользования отражено в монографии (Krapivin et al. 2015). В книге (Интегрированное управление 2001) представлена ИАС поддержки управления водными ресурсами (на примере Ленинградской области). Технологии интегрированного моделирования пространственной динамики региональных систем обсуждаются в статье (White and Engelen 2000). В работе (Bithell et al. 2008) описан дискретный подход к моделированию сложных пространственно-временных систем, предложенный группой британских географов.

Авторский подход к построению ИАС управления эколого-экономическими объектами изложен в работах (Угольницкий и Усов 2008а,б; Мониторинг 2009; Угольницкий 2016). Ключевая особенность авторского замысла заключается в решении посредством ИАС задач управления устойчивым развитием с использованием теории управления организационными системами (Новиков 2007) и реализации в составе ИАС ВР РО ряда механизмов управления (Механизмы управления 2011). Кроме того, ИАС ВР РО обеспечивает ввод, хранение, обработку и использование данных мониторинга водных ресурсов (Никаноров 2005).

Структурная схема ИАС показана на рис. 4.3.1.

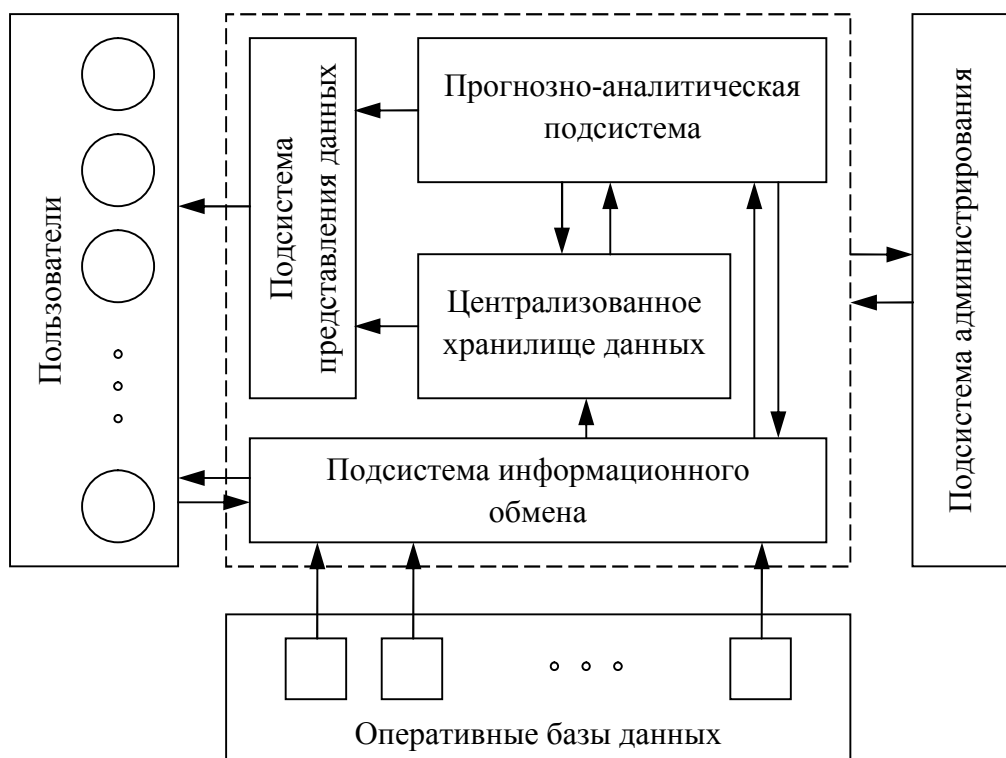


Рис. 4.3.1. Структурная схема ИАС ВР РО

Цель создания ИАС ВР РО состоит в повышении оперативности, эффективности и качества деятельности органов государственного управления РО за счет использования современных информационных технологий и математических моделей, формирования на их основе комплексной аналитической информации, необходимой для выработки, принятия и реализации стратегических и оперативных решений по управлению водными ресурсами региона.

Цель создания ИАС реализуется путем решения следующих задач:

- создание единого информационного пространства показателей состояния водных ресурсов РО на основе централизованного хранилища данных, обеспечивающего накопление и хранение исторических данных, интеграция существующих локальных баз данных;
- создание корпоративного информационного ресурса для обеспечения комплексного анализа состояния водных ресурсов области;

- повышение оперативности и качества управленческих решений на основе использования аналитических и прогнозных инструментальных средств;
- проведение мониторинга, анализа и прогнозирования состояния водных ресурсов, оптимизации их использования с учетом интересов исполнителей;
- обеспечение защиты, конфиденциальности и целостности коллективных информационных ресурсов ИАС;
- визуализация данных с использованием ГИС-технологий.

Разработка системы производится с использованием только лицензионного или открытого программного обеспечения со следующими техническими характеристиками:

1. Операционная система клиентских станций: Windows 7, Windows 8.1, Windows 10. В качестве клиентского программного обеспечения должен выступать бесплатный браузер Google Chrome версии 45 и выше.
2. Операционная система сервера базы данных: Windows Server 2008 R2, Windows Server 2012, Windows Server 2012 R2, или другой в зависимости от СУБД и ее версии.
3. Сервер Интернет-приложений: серверные механизмы Интернет-приложения (WEB) должны быть интегрированы со стандартными механизмами ОС Windows 2008 R2 и WEB-сервера Internet Information Server.
4. Система управления базами данных: система управления базами данных MS-SQL, ORACLE версии 12.

ИАС ВР РО включает следующие подсистемы: централизованное хранилище данных, подсистему информационного обмена, подсистему представления данных, подсистему администрирования, прогнозную-аналитическую подсистему.



Основу функционирования единого информационного пространства показателей состояния водных ресурсов области образует поддержка *централизованного хранилища данных (ЦХД)*. Часть информационных ресурсов общего пользования, наиболее важная для обоснования управленческих решений, должна формироваться в рамках ЦХД, остальные информационные ресурсы общего пользования системы формируются за счет локальных баз данных других органов государственного управления (ведомственных и районных). Структура и состав информационно-аналитических ресурсов общего пользования, а также технология их формирования разрабатываются в ходе предпроектного обследования объекта автоматизации.

ЦХД предназначено для накопления и хранения исторических данных о состоянии водных ресурсов области и ее муниципальных образований, а также по крупным предприятиям региона. Эти данные имеют динамическую природу и подвергаются интенсивной обработке большим количеством пользователей функциональных приложений, ориентированных на поддержку управленческих решений.

Компоненты ЦХД, решаемые задачи и выполняемые функции определяются стандартной концепцией хранилищ данных (Inmon 2002) с учетом специфики предметной области.

К группе технических подсистем относятся подсистемы информационного обмена, представления данных и администрирования.

Подсистема *информационного обмена* обеспечивает выполнение следующих функций:

1. Обмен нормативно-справочной информацией при поддержке согласованных при предпроектном обследовании doc, dbf, xls-форматов.
2. Поддержка шлюзов к реляционным SQL-совместимым СУБД (Oracle, Informix, MS SQL Server и др.).

3. Обеспечение обмена данными в трёх режимах: ручной ввод данных; автоматизированный по времени импорт-экспорт данных; автоматизированный по требованию пользователя импорт-экспорт данных.
4. Организация оперативного склада данных для их предварительного анализа и проверки корректности.
5. Интеграция в информационное пространство РФ, поддержка форматов обмена с внешними Интернет-ресурсами.

В ИАС должны быть разработаны средства просмотра и анализа информации из хранилища данных. Подсистема *представления данных* ориентируется на поддержку принятия решения руководством области.

Должны поддерживаться различные способы визуализации данных: табличное представление, деловая графика, картография, мультимедийное отображение, web-отображение и т.д.

Подсистема представления данных выполняет следующие функции: вывод отчётных документов на печать; поиск и просмотр готовых регламентных отчетов в пользовательских приложениях и сети Интернет; получение отчетов в формате электронной таблицы для последующей обработки; формирование отчётов с использованием параметризованных запросов через Интернет; формирование, просмотр и печать сложных отчётов по образцу в специальном интерактивном генераторе отчётов; импорт полученных выходных отчётных форм в офисные приложения Word, Excel, html-формат; проведение многомерного анализа данных хранилища и витрин с использованием предварительно подготовленных метаданных с помощью OLAP-средств; использование возможностей OLAP-средств для анализа данных различных источников при отсутствии подготовленных метаданных; проведение многомерного анализа с использованием средств Excel и стандартных браузеров.

Подсистема *администрирования* выполняет функции администрирования данных и хранилища данных, классификатора и справочников, приложений, организации доступа к объектам и функциям.

Программные средства должны предусматривать механизмы защиты объектов и данных от несанкционированного доступа, ограничивать возможности работы в зависимости от статуса пользователя. Права доступа каждого пользователя определяются степенью его участия в рабочем процессе и назначаются администратором ИАС.

Прогнозно-аналитическая подсистема занимает центральное место в ИАС в соответствии с ее замыслом и поддерживает функции оценки, прогнозирования и управления количеством и качеством водных ресурсов области. Подсистема прогнозно-аналитических расчетов должна состоять из следующих функциональных блоков: мониторинга, анализа, прогнозирования, оптимизации.

Блок *мониторинга* должен обеспечивать контроль текущей ситуации с водными ресурсами региона и его муниципальных образований и решение следующих задач: отслеживание тенденций состояния водных ресурсов области и ее муниципальных образований (дневная, недельная, месячная, квартальная, годовая динамика); экспресс-представление оперативных данных (табличное, графическое, картографическое отображение); мониторинг экономических показателей водопользования; мониторинг выполнения планов и прогнозов по водопользованию и охране водных ресурсов; ведение реестра водохозяйственной деятельности крупных предприятий; ведение целевых программ, инвестиционных проектов и бизнес-планов по водопользованию и охране водных ресурсов различного уровня.

Блок *анализа* предназначен для исследовательской обработки данных по ситуации с водными ресурсами региона и решения следующих задач: всесторонний (динамический, структурный, кластерный, факторный) анализ ситуации с водными ресурсами области и ее отдельных муниципальных

образований; экспресс-анализ данных; анализ и ранжирование муниципальных образований области по состоянию водных ресурсов; оценка чувствительности водохозяйственных проблем к изменению внутренних и внешних факторов; оценка инвестиционной привлекательности области и её муниципальных образований с точки зрения реализации водохозяйственных проектов.

Блок *прогнозирования* предназначен для автоматизации многовариантных расчетов краткосрочных, среднесрочных и долгосрочных прогнозов состояния водных ресурсов региона и его муниципальных образований на сценарной основе, а также регулярного мониторинга выполнения прогнозных показателей. Для решения подобных задач в ИАС предусматривается математический инструментарий, обеспечивающий проведение статистических, целевых и имитационных расчетов. Подход к региональному прогнозированию основывается на разработке комплекса имитационных моделей динамики количества и качества водных ресурсов.

Блок прогнозирования должен охватывать следующие функциональные направления: объемы стока; штатные и экстренные ситуации с водными ресурсами; поверхностные и подземные воды; загрязнение водной среды; качество водных ресурсов.

Схема прогнозирования с использованием интерполяции и экстраполяции по времени и пространству и применением нелинейных моделей динамики концентрации различных веществ в частных производных приведена в (Мониторинг 2009, с. 74-77) и развита в монографии (Угольницкий 2016).

Блок *оптимизации* в соответствии с авторским замыслом занимает центральное место в прогнозно-аналитической подсистеме. Он реализует механизмы управления (планирования, организации, стимулирования, контроля) (Новиков 2007; Механизмы управления 2011) водопользованием и модели управления устойчивым развитием (Угольницкий 2016) водохозяйственного комплекса региона. Приведем несколько примеров.

Механизм *последовательного распределения* водных ресурсов основан на заявках агентов о требуемом количестве воды, при этом центру неизвестно оптимальное для каждого агента количество ресурса. В общем случае используются механизмы распределения, определяемые правилом:

$$x_i(s) = \begin{cases} s_i, & \sum_{j=1}^n s_j \leq R, \\ \min\{s_i, \gamma \eta_i(s_i)\}, & \text{иначе} \end{cases}$$

где  $n$  – число агентов,  $\{s_i\}_{i \in N}$  – их заявки,  $\{x_i\}_{i \in N}$  – выделяемые количества ресурсов,  $R$  – распределяемое количество ресурса,  $\{\eta_i(s_i)\}_{i \in N}$  – функции приоритета агентов,  $\gamma$  – нормирующий параметр, который выбирается из условия выполнения бюджетного ограничения

$$\sum_{i=1}^n \min\{s_i, \gamma \eta_i(s_i)\} = R.$$

В зависимости от вида функции приоритета указанные механизмы подразделяются на прямые (возрастающая функция приоритета), абсолютные (постоянная функция приоритета) и обратные (убывающая функция приоритета). Механизмы абсолютных приоритетов при условии благожелательности можно считать неманипулируемыми; их недостаток состоит в том, что назначаемые планы не зависят от заявок агентов, то есть центр не использует получаемую от них информацию.

Механизмы прямых приоритетов называют также механизмами пропорционального распределения, поскольку они определяются формулой

$$x_i(s) = \frac{s_i}{\sum_{j \in N} s_j} R.$$

Механизм прямых приоритетов манипулируемый, поскольку из вида формулы ясно, что агентам выгодно завышать заявки. Чтобы избежать манипулируемости, применяют прямые механизмы последовательного распределения ресурса.

Предположим, что агенты подали свои заявки. Упорядочим их по возрастанию:  $s_1 \leq s_2 \leq \dots \leq s_n$ . Далее положим  $x_i := 0, i \in N$ , и применим следующий алгоритм последовательного распределения:

Шаг 1. Если  $ns_1 \leq R$ , то  $x_i := x_i + s_1, s_i := s_i - s_1, i \in N; R := R - ns_1$ , и перейти к шагу 2, иначе  $x_i := R/n, i \in N$ , и закончить выполнение алгоритма.

Шаг 2. Исключить первого агента из рассмотрения, перенумеровать агентов по порядку и вернуться к шагу 1.

Легко показать, что в прямом механизме последовательного распределения агентам выгодно сообщать достоверную информацию, то есть он защищен от манипулирования.

Механизмы обратных приоритетов можно реализовать с помощью функции приоритета  $\eta_i = A_i / s_i, i \in N$ . Величина  $A_i$  определяет потери системы в случае, если  $i$ -й агент вовсе не получает ресурса. Тогда отношение  $A_i / s_i$  характеризует удельный эффект от использования ресурса, поэтому механизмы обратных приоритетов называют также механизмами распределения ресурса пропорционально эффективности. Легко проверить непосредственно, что процедура приводит к равновесию Нэша. Можно также показать, что вычисленные стратегии гарантирующие, то есть максимизируют выигрыши агентов при наихудших для них действиях остальных игроков (Новиков 2007).

Разновидностью описанной процедуры можно считать механизм *смешанного финансирования* (Механизмы управления 2011), направленный на объединение средств центра (государства) и агентов (предприятий) для выполнения больших проектов (например, строительства водоводов или водоочистных сооружений), причем величина привлечённого бюджетного финансирования может быть гибко настраиваемой. Агенты подают в центр заявки на величину финансирования, необходимого для выполнения проекта, а центр распределяет финансовый ресурс пропорционально заявкам с учетом

приоритетов агентов. Недостающее финансирование агенты покрывают собственными средствами.

Рассмотрим теперь модель согласования общественных и частных интересов (Gorbaneva, Ougolnitsky 2018) при проведении водоочистных мероприятий. Предполагается, что структура интересов каждого агента (предприятия) включает как частную составляющую (развитие производства), так и общественную составляющую (водоочистные мероприятия). Агенты могут делить персональные ресурсы между этими направлениями. Цель моделирования состоит в исследовании поведения агентов при различных соотношениях указанных составляющих и обосновании механизмов организационного управления, стимулирующих реализацию общественных интересов (повышение качества воды).

Обозначим

$N = \{1, 2, \dots, n\}$  – множество агентов;  $r_i$  – финансовые ресурсы  $i$ -го агента;

$u_i$  – доля ресурса  $i$ -го агента, используемая на очистку воды;  $r_i - u_i$  – доля ресурса  $i$ -го агента, идущая на развитие производства;

$p_i$  – функция дохода  $i$ -го агента от развития производства;

$s_i$  – доля экологического ущерба  $i$ -го агента от загрязнения воды (возможный параметр экономического управления центра);

$q_i$  – возможное административное ограничение величины  $u_i$  снизу;

$g_i$  – функция вклада  $i$ -го агента в снижение загрязнения водной среды;

$J_i$  – общий дисконтированный доход  $i$ -го агента на периоде  $[0, T]$ ;

$J_0$  – дисконтированная сумма экологического ущерба и затрат центра  $D(q)$  на административный контроль на периоде  $[0, T]$ ;

$x$  – количество загрязняющих веществ в водной среде;

$h$  – функция выбросов загрязняющих веществ;

$\rho$  – коэффициент дисконтирования.

Модель имеет вид:

$$J_0(\cdot) = \sum_{t=1}^T e^{-\rho t} [x(t) + D(q(t))] \rightarrow \min ;$$

$$0 \leq q_i(t) \leq r_i ;$$

$$s_i(t) \geq 0; \sum_{i=1}^n s_i(t) = 1 ;$$

$$J_i(\cdot) = \sum_{t=1}^T e^{-\rho t} [p_i(r_i - u_i(t)) - s_i(t)x(t)] \rightarrow \max ;$$

$$q_i(t) \leq u_i(t) \leq r_i; i = 1, 2, \dots, n ;$$

$$x(0) = x_0; i = 1, \dots, n; t = 0, 1, \dots, T - 1 ;$$

$$x(t+1) = x(t) + h \left( \sum_{i=1}^n p_i(r_i - u_i(t)) - \sum_{i=1}^n g_i(u_i(t)) \right) .$$

В общем случае для получения решения полученной разностной игры применяются численные методы и имитационное моделирование с использованием возможностей ИАС ВР РО.

Итак, решение чрезвычайно актуальных для Ростовской области проблем водоснабжения с учетом качества воды требует совершенствования управления водопользованием. Ключевым инструментом здесь служит информационно-аналитическая система управления количеством и качеством водных ресурсов Ростовской области. Отличительной особенностью предлагаемой ИАС ВР РО выступает реализация концепции управления устойчивым развитием активных систем (Угольницкий 2016) применительно к водным ресурсам. С этой целью проведены конкретизация механизмов управления (Новиков 2007; Механизмы управления 2011) применительно к водопользованию и разработка оригинальных моделей управления устойчивым развитием водохозяйственного комплекса.

ИАС ВР РО интегрирует информацию о состоянии и динамике водных ресурсов области из различных источников и тем самым создаёт благоприятные возможности для анализа и принятия качественных, научно обоснованных управленческих решений. Однако успешная реализация этих решений возможна лишь при учёте интересов всех субъектов водопользования (Пучкин и др. 2016).



## ЗАКЛЮЧЕНИЕ

Опишем перспективные направления развития результатов, полученных в ходе выполнения проекта РНФ №17-19-01038 (2017-2019) и представленных в настоящей монографии.

В части методологии предполагается продолжить анализ общего понятия активной системы и его спецификаций в конкретных предметных областях. Так, будет продолжено исследование учреждений высшего образования с позиций теории активных систем. Целесообразно построить систему индексов системной согласованности, дающих удобное количественное выражение этого понятия для различных методов иерархического управления. Нуждается в дальнейшем уточнении и общее понятие устойчивого развития активной системы, носящее весьма сложный характер синтеза многих определений.

Магистральным направлением развития теории следует считать построение и исследование динамических моделей согласования интересов при управлении устойчивым развитием. В докторской диссертации О.И. Горбанёвой (2019) и ряде совместных публикаций (Горбанёва, Угольницкий 2014, 2015, 2016; Gorbaneva and Ougolnitsky 2013, 2015, 2018) построены и исследованы статические модели согласования общественных и частных интересов при распределении ресурсов (СОЧИ-модели). Разумеется, что ещё больший интерес представляют динамические СОЧИ-модели, в которых благодаря явному описанию объекта управления и учёта фактора времени можно изучить множество дополнительных эффектов. В частности, это важно для анализа согласования долгосрочных интересов дальновидного Центра и краткосрочных интересов близоруких агентов. В работах (Угольницкий, Усов 2016; Sukhinov et al. 2017) начато изучение динамических СОЧИ-моделей в различных предметных областях с помощью метода качественно репрезентативных сценариев имитационного моделирования.

Представляется весьма перспективным продолжение этих исследований как с помощью метода КРС, так и особенно путём аналитического обобщения результатов, полученных для статических моделей, подобно тому, как это было сделано для динамических моделей стимулирования (Рохлин и Угольницкий 2018; Rokhlin and Ougolnitsky 2019a,b). Динамическая теория стимулирования также ещё далека от завершения: в частности, предстоит изучить модели с непрерывным временем, взаимодействием нескольких Центров, кооперативным поведением агентов.

Важную роль в развитии теории управления устойчивым развитием играют методы решения соответствующих задач. Нахождение аналитических решений динамических теоретико-игровых задач управления возможно только в сильно упрощённых случаях. Поэтому на первый план выходят приближённые численные методы и алгоритмы. Будет продолжена разработка методов случайного поиска для решения дифференциальных игр. Здесь используется идея о том, что в реальных процессах управления Центр не может резко менять свою стратегию произвольным образом, а способен лишь на незначительное увеличение или уменьшение значений управляющих параметров. Эта "липшицева" гипотеза позволяет получить содержательные точные оценки и разработать эффективные вычислительные процедуры, дополняющие традиционные стратегии поощрения/наказания.

Другая важная идея лежит в основе активно развиваемого авторского метода качественно репрезентативных сценариев имитационного моделирования. Хотя теоретически возможное множество управляющих воздействий в каждой задаче очень велико или даже бесконечно, естественно считать, что для получения качественно приемлемого прогноза достаточно рассмотреть очень небольшое число вариантов, отвечающих принципиально различным стратегиям. Большая часть промежуточных вариантов не будет давать существенно новых результатов. Эта идея эвристическая, но можно сформулировать точные условия её проверки, обосновывающие замену

необозримого полного перебора рассмотрением малого числа "внутренне" и "внешне" устойчивых качественно репрезентативных сценариев.

Планируется продолжить прикладные исследования задач управления устойчивым развитием активных систем в различных предметных областях. К их числу относятся эколого-экономические системы (оптимальная эксплуатация биоресурсов, борьба с загрязнением природной среды), финансовые рынки (определение вероятности выхода случайного процесса за барьер, модели стохастического оптимального управления), продвижение инноваций в организациях (разработка механизмов стимулирования инноваций), сетевые модели влияния и управления в маркетинге, учреждения высшего образования, борьба с коррупцией в иерархических системах управления.

Наконец, практическим инструментом управления устойчивым развитием служат информационно-аналитические управляющие системы, соединяющие математические модели и экспертные оценки с возможностями информационных технологий. Однако, для разработки информационно-аналитических систем требуются финансовые средства намного большие, чем для теоретического моделирования.

Развитие комплексной теории управления устойчивым развитием активных систем представляется важным фундаментом повышения эффективности и качества менеджмента в долгосрочной перспективе.

## ЛИТЕРАТУРА

1. Агиева М.Т. Модели управления на социальных сетях в маркетинге // Инж. вестник Дона. 2018. №1. [ivdon.ru/ru/magazine/archive/n1y2018/4670](http://ivdon.ru/ru/magazine/archive/n1y2018/4670).
2. Агиева М.Т. Задачи анализа на социальных сетях в маркетинге // Инж. вестник Дона. 2018. № 2. URL: [ivdon.ru/ru/magazine/archive/N2y2018/4889](http://ivdon.ru/ru/magazine/archive/N2y2018/4889).
3. Агиева М.Т. Задачи прогноза на социальных сетях в маркетинге // Экономика и менеджмент систем управления. 2018. №4.1(30). С.110-117.
4. Агиева М.Т. Классификация моделей управления целевой аудиторией в маркетинге // Инженерный вестник Дона. 2019. №1. URL: [ivdon.ru/ru/magazine/archive/n1y2019/5612](http://ivdon.ru/ru/magazine/archive/n1y2019/5612).
5. Агиева М.Т. Качественно репрезентативные сценарии имитационного моделирования маркетинговых воздействий // Инженерный вестник Дона. 2019. №2.
6. Агиева М.Т., Бабичева Ю.В., Окулист Н.М., Угольницкий Г.А. Задачи анализа и прогноза при управлении целевой аудиторией в маркетинге // Управление большими системами. Вып.79. - М.: ИПУ РАН, 2019. С.27-64.
7. Айдаркина Е.Е. Мониторинг состояния водных ресурсов Ростовской области // Гуманитарные и социальные науки. 2012. № 5. С. 53-62.
8. Алескеров Ф.Т., Хабина Э.Л., Шварц Д.А. Бинарные отношения, графы и коллективные решения. - М., 2006.
9. Анопченко Т.Ю., Мурзин А.Д., Угольницкий Г.А. Моделирование согласования интересов в задачах управления устойчивым развитием территорий // Экономика природопользования. 2017. №6. С.35-47.
10. Бакаева Е.Н., Никаноров А.М. Биологические подходы к оценке экотоксикологического состояния водных экосистем // Известия ВУЗов. Северо-Кавказский регион. Естественные науки. 2015. №1. С.72-83.
11. Бакаева Е.Н., Никаноров А.М., Игнатова Н.А. Динамика токсичности поверхностных вод бассейна реки Дон в пределах мегаполиса по

- многолетним данным биотестирования // Водные ресурсы. 2015. Т.42. №1. С.70-77.
12. Белов М.В., Новиков Д.А. Сетевые активные системы: модели планирования и стимулирования // Проблемы управления, 2018, 1, 47–57.
  13. Белявский Г.И., Данилова Н.В., Угольницкий Г.А. Эволюционное моделирование в задачах управления устойчивым развитием активных систем // Математическая теория игр и ее приложения, 2016, 8(4), 14-29.
  14. Белявский Г.И., Данилова Н.В., Угольницкий Г.А. Эволюционные методы решения динамических задач распределения ресурсов // Математическая теория игр и ее приложения, 2018, 10(1), 5-22.
  15. Белявский Г.И., Лиля В.Б., Пучков Е.В. Алгоритм и программная реализация гибридного метода обучения искусственных нейронных сетей // Программные продукты и системы, 2012 (4), 96-101.
  16. Беляевский И.К. Маркетинговое исследование: информация, анализ, прогноз. - М.: КУРС: ИНФРА-М, 2017.
  17. Бернар К. Курс общей физиологии. Жизненные явления, общие животным и растениям. - М., 1878.
  18. Большие системы: моделирование организационных механизмов / Бурков В.Н., Данев Б., Еналеев А.К. и др. - М.: Наука, 1989.
  19. Бурков В.Н. Основы математической теории активных систем. - М.: Наука, 1977.
  20. Бурков В.Н., Новиков Д.А. Теория активных систем: состояние и перспективы. - М.: Синтег, 1999.
  21. Бурков В.Н., Новиков Д.А., Щепкин А.В. Механизмы управления эколого-экономическими системами. - М.: Физматлит, 2008.
  22. Бурков В.Н., Опойцев В.И. Метаигровой подход к управлению иерархическими системами // Автоматика и телемеханика, 1974 (1), 103-114.
  23. Гермейер Ю.Б. Введение в теорию исследования операций. М.: Наука, 1971.

24. Гермейер Ю.Б. Игры с противоположными интересами. - М.: Наука, 1975.
25. Гермейер Ю.Б., Ватель И.А. Игры с иерархическим вектором интересов // Известия АН СССР. Техническая кибернетика. 1974. №3. С.54-69.
26. Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы. – М.: Физматлит, 2006.
27. Горбанёва О.И. Статические модели распределения ресурсов с учётом согласования интересов активных агентов. Диссертация на соискание учёной степени д.т.н. М., 2019.
28. Горбанёва О.И., Угольницкий Г.А. Механизмы согласования интересов в модели распределения ресурсов // Системы управления и информационные технологии. 2014. № 3.2 (57). С.225-231.
29. Горбанёва О.И., Угольницкий Г.А. Цена анархии и механизмы управления в моделях согласования общественных и частных интересов // Математическая теория игр и её приложения, 2015, 7(1), 50-73.
30. Горбанёва О.И., Угольницкий Г.А. Учет коррупции в моделях сочетания общественных и частных интересов в иерархических системах управления // Проблемы управления. 2016. №2. С.62-71.
31. Горбанёва О.И., Мурзин А.Д., Угольницкий Г.А. Механизмы согласования интересов при управлении проектами развития территорий // Управление большими системами. Вып. 71. М.: ИПУ РАН, 2018. С.61-97.
32. Горелик В.А., Кононенко А.Ф. Теоретико-игровые модели принятия решений в эколого-экономических системах. - М.: Радио и связь, 1982.
33. Горелик В.А., Горелов М.А., Кононенко А.Ф. Анализ конфликтных ситуаций в системах управления. - М.: Радио и связь, 1991.
34. Горелов М.А., Кононенко А.Ф. Динамические игры. I. Язык моделирования // Автоматика и телемеханика, 2014 (11), 127-149.
35. Горелов М.А., Кононенко А.Ф. Динамические игры. II. Равновесия // Автоматика и телемеханика, 2014 (12), 56-77.

36. Горелов М.А., Кононенко А.Ф. Динамические игры. III. Иерархические игры // Автоматика и телемеханика, 2015 (2), 89-106.
37. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. - М.: Физматлит, 2010.
38. Губко М.В. Математические модели оптимизации организационных структур. - М.: Ленанд, 2006.
39. Данилов-Данильян В.И., Лосев К.С. Экологический вызов и устойчивое развитие. - М.: Прогресс-Традиция, 2000.
40. Деминг Э. Выход из кризиса. Новая парадигма управления людьми, системами и процессами. – М.: Альпина Бизнес Букс, 2007.
41. Дрейер О.К., Лось В.А. Экология и устойчивое развитие. - М.: изд-во УРАО, 1997.
42. Емельянов В.В., Курейчик В.В., Курейчик В.М. Теория и практика эволюционного моделирования. — М.: Физматлит, 2003.
43. Закруткин В.Е., Складченко Г.Ю., Гибков Е.В. Особенности химического состава и степень загрязненности подземных вод углепромышленных районов Восточного Донбасса // Известия ВУЗов. Северо-Кавказский регион. Естественные науки. 2014. №4. С.73-75.
44. Зенкевич Н.А., Петросян Л.А., Янг Д.В.К. Динамические игры и их приложения в менеджменте. - СПб.: Высшая школа менеджмента, 2009.
45. Иванилов В.Ю., Огарышев В.Ф., Павловский Ю.Н. Имитация конфликтов. – М.: ВЦ РАН, 1993.
46. Интегрированное управление водными ресурсами Санкт-Петербурга и Ленинградской области. Опыт создания системы поддержки принятия решений / Под ред. А.Ф. Алимова, Л.А. Руховца, М.М. Степанова. – СПб.: Наука, 2001.
47. КАБАКОВ Р.И. R в действии. Анализ и визуализация данных в программе R. – М.: ДМК Пресс, 2016.

48. Клейнен Дж. Статистические методы в имитационном моделировании. Вып.1-2. – М.: Статистика, 1978.
49. Ключко А. Чтобы решить проблему водоснабжения Ростовской области, необходимо 150 млрд рублей // URL: [http://www.donnews.ru/Chtoby-reshit-problemu-vodosnabzheniya-Rostovskoy-oblasti-neobhodimo-150-mlrd-rubleu\\_1509](http://www.donnews.ru/Chtoby-reshit-problemu-vodosnabzheniya-Rostovskoy-oblasti-neobhodimo-150-mlrd-rubleu_1509) (Дата обращения: 11.01.2016).
50. Когай В.В., Фадеев С.И. Применение продолжения по параметру на основе метода множественной стрельбы для численного исследования нелинейных краевых задач // Сибирский журнал индустриальной математики. 2001. Т.4. № 1(7). С. 83-101.
51. Кононенко А.Ф. О многошаговых конфликтах с обменом информацией // Журнал вычислительной математики и математической физики, 1977, 17(4), 922-931.
52. Коханенко В.Н., Волосухин Я.В., Ширяев В.В. Моделирование одномерных и двумерных открытых водных потоков: монография / Под общей ред. В.Н. Коханенко. - Ростов н/Д: изд-во ЮФУ, 2007.
53. Коханистая Е.В., Хоружая Т.А. Современный уровень солевого загрязнения Пролетарского и Веселовского водохранилищ // Известия ВУЗов. Северо-Кавказский регион. Естеств. науки. 2015. №2. С.88-92.
54. Кузнецов А.Л. Об анализе социальных сетей и их метриках. Материалы VIII Международной студенческой электронной научной конференции «Студенческий научный форум». URL: [scienceforum.ru/2017/2830/31138](http://scienceforum.ru/2017/2830/31138) (дата обращения: 14.03.2018).
55. Кукушкин Н.С. О существовании устойчивых исходов в теоретико-игровой модели экономики с общественными благами // Докл. АН СССР. 1991. Т.320. №1. С.25-28.
56. Ледяев В.Г. Власть: концептуальный анализ. – М.: РОССПЭН, 2000.
57. Лексин В., Швецов А. Муниципальная Россия: социально-экономическая ситуация, законодательство, статистика. - М.: УРСС, 2000.
58. Лексин В., Швецов А. Государство и регионы. - М.: УРСС, 2016.



59. Лифинцев Д.В. Оценка связей индивида в микросоциуме на основе методов анализа социальных сетей // Вестник Балтийского федерального университета им. И. Канта. Серия: Филология, педагогика, психология. 2013. №5. URL: [cyberleninka.ru/article/n/otsenka-svyazey-individa-v-mikrosotsiume-na-osnove-metodov-analiza-sotsialnyh-setey](http://cyberleninka.ru/article/n/otsenka-svyazey-individa-v-mikrosotsiume-na-osnove-metodov-analiza-sotsialnyh-setey) (дата обращения: 14.03.2018).
60. Лоу М., Кельтон Д. Имитационное моделирование. – СПб.: Питер, 2004.
61. Мальсагов М.Х. Учреждение высшего образования как активная система // Инженерный вестник Дона. 2018. №3. <http://ivdon.ru/ru/magazine/archive/n3y2018/5122>.
62. Мальсагов М.Х. Модель экономической коррупции как игра в развёрнутой форме // Инженерный вестник Дона. 2018. №4. <http://ivdon.ru/ru/magazine/archive/n4y2018/5381>.
63. Мальсагов М.Х. Модели комплексного оценивания и кадровой оптимизации работы факультета // Инженерный вестник Дона. 2019. №1. <http://ivdon.ru/ru/magazine/archive/n1y2019/5624>.
64. Мальсагов М.Х., Угольницкий Г.А. Дифференциально-игровые модели коррупции при распределении ресурсов // Инженерный вестник Дона. 2018. №2. [ivdon.ru/ru/magazine/archive/N2y2018/4984](http://ivdon.ru/ru/magazine/archive/N2y2018/4984).
65. Мальсагов М.Х., Угольницкий Г.А., Усов А.Б. Борьба с экономической коррупцией при распределении ресурсов // Компьютерные исследования и моделирование, 2019, 11(1), 173-186.
66. Математическая теория оптимальных процессов / Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. - М.: Наука, 1983.
67. Меринова Ю.Ю., Хованский А.Д. О состоянии и использовании водных ресурсов в Ростовской агломерации // Известия ВУЗов. Северо-Кавказский регион. Естественные науки. 2014. №3. С.96-101.
68. Механизмы управления / Под ред. Д.А. Новикова. - М.: Ленанд, 2011.
69. Моделирование социо-эколого-экономической системы региона / Под ред. В.И. Гурмана, Е.В. Рюминой. - М.: Наука, 2001.

70. Моисеев Н.Н. Математика ставит эксперимент. – М.: Наука, 1979.
71. Моисеев Н.Н., Евтушенко Ю.Г., Краснощеков П.С., Павловский Ю.Н. Имитационные системы // Экономика и организация промышленного производства. 1973. №6. С. 39-46.
72. Мониторинг: от приложений к общей теории / Под ред. Г.А. Угольницкого. – Ростов-на-Дону: изд-во ЮФУ, 2009.
73. Назиров А.Э., Угольницкий Г.А., Усов А.Б. Теоретико-игровая модель трехуровневой маркетинговой системы с учетом экологических требований // Управление большими системами. Вып. 55. - М., 2015. С.326-342.
74. Нейлор Т. Машинные имитационные эксперименты с моделями экономических систем. - М.: Мир, 1975.
75. Нив Г. Организация как система. Принципы построения устойчивого бизнеса Эдвардса Деминга. – М.: Альпина Бизнес Букс, 2007.
76. Никаноров А.М. Научные основы мониторинга качества вод. - СПб.: Гидрометеоздат, 2005.
77. Никитина А.В., Пучкин М.В., Семенов И.С., Сухинов А.И., Угольницкий Г.А., Усов А.Б., Чистяков А.Е. Дифференциально-игровая модель предотвращения заморов в мелководных водоемах // Управление большими системами. Вып. 55. - М., 2015. С.343-361.
78. Никитина А.В., Семёнов И.С. Численная реализация методов решения задач биологической кинетики в Азовском море // Известия ЮФУ. Технические науки. 2014. № 1. С. 138–143.
79. Никитина А.В., Сухинов А.И., Угольницкий Г.А. и др. Оптимальное управление устойчивым развитием при биологической реабилитации Азовского моря // Математическое моделирование, 2016, 28(7), 96-106.
80. Николенко С.И. Теория экономических механизмов. - М.: БИНОМ, 2009.
81. Новиков Д.А. Теория управления организационными системами. - М.: Изд-во физико-математической литературы, 2007.

82. Новиков Д.А. Введение в теорию управления образовательными системами. - М.: Эгвес, 2009.
83. Новиков Д.А. Кибернетика: Навигатор. История кибернетики, современное состояние, перспективы развития. - М.: УРСС, 2016.
84. Новиков Д.А., Иващенко А.А. Модели и методы организационного управления инновационным развитием фирмы. - М.: УРСС, 2006.
85. Новиков Д.А., Смирнов И.М., Шохина Т.Е. Механизмы управления динамическими активными системами. - М.: ИПУ РАН, 2002.
86. Новиков Д.А., Чхартишвили А.Г. Рефлексия и управление: математические модели. - М.: Физматлит, 2013.
87. Новиков Д.А., Шохина Т.Е. Механизмы стимулирования в динамических активных системах // Автоматика и телемех., 2003, 12, 94-104.
88. Нор-Аревян О.А., Тарасенко Л.В., Угольницкий Г.А. Математическое моделирование социального партнерства: методология междисциплинарного исследования (на примере системы дополнительного профессионального образования) // Социологические исследования, 2018, 4, 15-24.
89. Оноприенко А.Н., Угольницкий Г.А., Усов А.Б. Имитационное моделирование иерархических регламентов управления (на примере рыболовства) // Известия ВУЗов. Северо-Кавказский регион. Естественные науки. 2016. №4. С.25-29.
90. Основы лесной биогеоценологии / Под ред. В.Н. Сукачёва, Н.В. Дылиса. – М.: Наука, 1964.
91. Павловский Ю.Н. Имитационные модели и системы. – М.: ФАЗИС - ВЦ РАН, 2000.
92. Переход к устойчивому развитию: глобальный, региональный и локальный уровни. Зарубежный опыт и проблемы России. / Рук. авт. колл. Н.Ф. Глазовский. - М.: КМК, 2002.
93. Петросян Л.А., Зенкевич Н.А., Шевкопляс Е.В. Теория игр. – СПб.: БХВ-Петербург, 2012.

94. Печерский С.Л., Яновская Е.Б. Кооперативные игры: решения и аксиомы. – СПб.: Европейский ун-т в СПб., 2004.
95. Пучкин М.В., Угольницкий Г.А., Усов А.Б. Проект информационно-аналитической системы управления водными ресурсами Ростовской области // Известия ВУЗов. Северо-Кавказский регион. Технические науки. 2016. №2. С.28-35.
96. Робертс Ф. Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам. - М.: Наука, 1986.
97. Рохлин Д.Б., Угольницкий Г.А. Равновесие Штакельберга в динамической модели стимулирования с полной информацией // Автоматика и телемеханика, 2018, 4, 152-166.
98. Сухинов А.И., Никитина А.В., Чистяков А.Е. Моделирование сценария биологической реабилитации Азовского моря // Математическое моделирование. 2012. Т. 24. № 9. С. 3–21.
99. Сухинов А.И., Чистяков А.Е., Алексеенко Е.В. Численная реализация трехмерной модели гидродинамики для мелководных водоемов на супервычислительной системе // Матем. моделирование, 2011, 23(3), 3–21.
100. Сухинов А.И., Чистяков А.Е. Адаптация модифицированного попеременно-треугольного итерационного метода для решения сеточных уравнений с несамосопряжённым оператором // Матем. моделирование, 2012, 24(1), 3–20.
101. Угольницкий Г.А. Управление эколого-экономическими системами. - М.: Вузовская книга, 1999.
102. Угольницкий Г.А. Теоретико-игровое исследование некоторых способов иерархического управления // Известия РАН. Теория и системы управления. 2002. №1. С.97-101.
103. Угольницкий Г.А. Иерархическое управление устойчивым развитием социальных организаций // Общ. науки и современность, 2002(3), 133-140.
104. Угольницкий Г.А. Иерархическое управление устойчивым развитием. – М.: Физматлит, 2010.

105. Угольницкий Г.А. Устойчивое развитие организаций. – М.: Физматлит, 2011.
106. Угольницкий Г.А. Управление устойчивым развитием активных систем. - Ростов-на-Дону: Изд-во Южного федерального университета, 2016.
107. Угольницкий Г.А. Методология и прикладные задачи управления устойчивым развитием активных систем // Проблемы управления. 2019. №2. С.19-29.
108. Угольницкий Г.А., Усов А.Б. Динамические модели согласования частных и общественных интересов при продвижении инноваций // Математическая теория игр и ее приложения, 2019, 11(1), 96-114.
109. Угольницкий Г.А., Усов А.Б. Информационно-аналитическая система управления эколого-экономическими объектами // Известия РАН. Теория и системы управления. 2008. №2. С.168-176.
110. Угольницкий Г.А., Усов А.Б. Информационно-аналитические системы управления качеством водных ресурсов // Вод. ресурсы, 2008, 35(5), 625-631.
111. Угольницкий Г.А., Усов А.Б. Исследование дифференциальных моделей иерархических систем управления путем их дискретизации // Автоматика и телемеханика. 2013, 2, 109-122.
112. Угольницкий Г.А., Усов А.Б. Равновесия в моделях иерархически организованных динамических систем с учетом требований устойчивого развития // Автоматика и телемеханика, 2014, 6, 86-102.
113. Угольницкий Г.А., Усов А.Б. Теоретико-игровая модель согласования интересов при инновационном развитии корпорации // Компьютерные исследования и моделирование, 2016, 8(4), 673-684.
114. Угольницкий Г.А., Усов А.Б., Пучкин М.В. и др. Теоретико-игровые регламенты механизмов управления устойчивым развитием мелководных экосистем // Автоматика и телемеханика, 2017, 6, 122-137.
115. Угольницкий Г.А., Усов А.Б., Рыжкин А.И. Метод побуждения в играх Гермейера при моделировании трехуровневой системы управления судовыми

- балластными водами // Компьютерные исследования и моделирование, 2014, 6(4), 535-542.
116. Угольницкий Г.А., Усов А.Б., Рыжкин А.И. Метод принуждения в играх Гермейера при моделировании трехуровневой системы управления судовыми балластными водами // Компьютерные исследования и моделирование, 2015, 7(2), 281-288.
117. Шадрин А.Д. Менеджмент качества. От основ к практике. – М.: ООО НТК Терек, 2005.
118. Шеннон Р. Имитационное моделирование систем - искусство и наука. - М.: Мир, 1978.
119. Юдкевич М.М., Подколзина М.А., Рябина А.Ю. Основы теории контрактов: модели и задачи. - М.: ГУ ВШЭ, 2002.
120. Adams W.M., Jeanrenaud S.J. Transition to Sustainability: Towards a Humane and Diverse World. – Gland: International Union for Conservation of Nature and Natural Resources, 2008.
121. Algorithmic Game Theory / Ed. by N. Nisan, T. Roughgarden, E. Tardos, V. Vazirani. - Cambridge University Press, 2007.
122. Aliprantis, C.D., Border, K.C. Infinite dimensional analysis: a hitchhiker's guide. - Springer-Verlag, 2006.
123. Anderies J.M. et al. Panaceas, uncertainty, and the robust control framework in sustainability science//Proc. Natl. Acad. Sci. USA, 2007, 104(39), 15194–15199.
124. Arkhipov B. V., Solbakov V. V., Solov'ev M. B., and Shapochkin D. A. Evaluating Water Pollution Characteristics Based on Lagrangian Approach // Water Resources, 2015, 42(1), 63-69.
125. Arnason R. Fisheries management and operations research // Eur. J. Oper. Res., 2009, 193(3), 741–751.
126. Aubin J.-P. Viability Theory. - Springer-Verlag, 1991.
127. Basar T., Olsder G.J. Dynamic Non-Cooperative Game Theory. – Philadelphia: SIAM, 1999.

128. Basar T., Zhu Q. Prices of Anarchy, Information, and Cooperation in Differential Games // *Dynamic Games and Applications*, 2011, 1(1), 50-73.
129. Belyavsky G., Danilova N. Random search methods for the solution of a Stackelberg game of resource allocation // *Contributions to game theory and management. Vol. XII. Collected papers presented on the Twelfth International Conference Game Theory and Management / Editors L. Petrosyan, N. Zenkevich.* – SPb.: Graduate School of Management SPbU, 2019.
130. Belyavsky G., Danilova N., Ougolnitsky G. A Markovian Mechanism of Proportional Resource Allocation in the Incentive Model as a Dynamic Stochastic Inverse Stackelberg Game // *Mathematics*, 2018, 6(8), 131.
131. Benchenkroun, H. and N.V. Long. The build up of cooperative behavior among non-cooperative selfish agents // *Journal of Economic Behavior and Organization*, 2008, 67(1), 293–252.
132. Bergstrom, T., C. Blume and H. Varian. On the private provision of public goods // *Journal of Public Economics*, 1986, 29, 25–49.
133. Bertsekas, D., and S. Shreve. *Stochastic optimal control: the discrete time case.* - Athena Scientific, Belmont, 1996.
134. Bithell M., Brasington J., Richards K. Discrete-element, individual-based and agent-based models: Tools for interdisciplinary enquiry in geography? // *Geoforum*, 2008 (39), 625-642.
135. Blackwell D. Discounted dynamic programming // *The Annals of Mathematical Statistics*, 1965, 36(1): 226–235.
136. Boadway, R., P. Pestiau and D. Wildasin. Non-cooperative behavior and efficient provision of public goods // *Public Finance*, 1989, 44, 1–7.
137. Boadway R., P. Pestiau and D. Wildasin. Tax-transfer policies and the voluntary provision of public goods // *J. of Public Economics*, 1989, 39, 157–176.
138. Bogachev V. *Measure theory. Volume II.* - Springer-Verlag, Berlin, 2007.
139. Bolton P., Dewatripont M. *Contract Theory.* – MIT Press, 2004.
140. Breton, M., A. Alj, and A. Haurie. Sequential Stackelberg equilibria in two-person games // *J. of Optimization Theory and Applications*, 1988, 59(1):71–97.

141. Burkov V.N., Lerner A.Ya. Fairplay in control of active systems // Differential games and related topics. - Amsterdam, London: North-Holland Publishing Company, 1971. - P.164—168.
142. Cairns R.D., Martinet V. An environmental-economic measure of sustainable development // European Economic Review, 2014 (69), 4-17.
143. Cannon W. The Wisdom of the Body. - L., 1932.
144. Charles A.T. Living with uncertainty in fisheries: analytical methods, management priorities and the Canadian ground-fishery experience // Fish. Res., 1998, vol. 37, 37–50.
145. Clark B.R. Creating entrepreneurial universities: organizational pathways of transformation. Issues in Higher Education. Paris: IAU Press, Pergamon, Elsevier Science, 1998.
146. Clark W.C. Sustainability Science: A Room of its Own // Proceedings of the National Academy of Science, 2007 (114), 1737-1738.
147. Clark W.C. The worldwide crisis in fisheries. Economic models and human behavior. - Cambridge University Press, Cambridge, 2006.
148. Clark W.C., Kirkwood G.P. On uncertain renewable resource stocks: optimal harvest policies and the value of stock surveys // J. Environ. Econ. Management, 1986, vol. 13, p. 235–244.
149. Christodoulou G., Sgouritza A., Tang B. On the Efficiency of the Proportional Allocation Mechanism for Divisible Resources // M. Hoefler (Ed.): SAGT 2015, LNCS 9347, 165–177.
150. Cornes R., Hartley R. Asymmetric contests with general technologies // Economic Theory 26(4), 2005, pp. 923-946.
151. Cornes R., Sato T. Existence and uniqueness of Nash equilibrium in aggregative games: an expository treatment // Equilibrium Theory for Cournot Oligopolies and Related Games. P. von Mouche, F. Quartieri (Eds.). - Springer International Publishing, 2016, pp.47–61.



152. Cruz-Suarez, D., R. Montes-de-Oca, and F. Salem-Silva. Conditions for the uniqueness of optimal policies of discounted Markov decision processes // *Mathematical Methods of Operations Research*, 2004, 60(3): 415–436.
153. De Crespigny A. Power and Its Forms // *Polit. Studies*, 1968, 16(2), 192-205.
154. DE GROOT M.H. Reaching a Consensus // *Journal of American Statistical Association*, 1974, 69, 118-121.
155. Dockner, E. and N.V. Long. International pollution control: Cooperative versus non-cooperative strategies // *Journal of Environmental Economics and Management*, 1993, 25, 13–29.
156. Dockner E., Jorgensen S., Long N.V., Sorger G. *Differential Games in Economics and Management Science*. - Cambridge University Press, 2000.
157. Doyen L., Bene C. Sustainability of fisheries through marine reserves a robust modeling analysis // *J. Environ. Management*, 2003, vol. 1969, p. 1–13.
158. Doyen L., Martinet V. Maximin, viability and sustainability // *J. of Econ. Dynamics and Control*, 2012 (36), 1414-1430.
159. Dutta P.K., Sundaram R.K. The equilibrium existence problem in general Markovian games. In: Majumdar, M. (ed.) *Organizations with incomplete information*. - Cambridge University Press, 1998. P. 159-207.
160. Dyachenko V.K., Ougolnitsky G.A., Tarasenko L.V. Computer Investigation of a Game Theoretic Model of Social Partnership in the System of Continuing Professional Education // *Advances in Systems Science and Applications*, 2015, 15(4), 320-328.
161. Eidelman Yu., Milman V., Tsolomitis A. *Functional Analysis: An Introduction*. - AMS, 2004.
162. Erev E., Roth A. Predicting how people play game: reinforcement learning in experimental games with unique, mixed strategy equilibria // *The American Economic Review*, 1998, 88(4), 848–881.
163. Feinberg, E.A., and M.E. Lewis. Optimality inequalities for average cost Markov decision processes and the stochastic cash balance problem // *Mathematics of Operations Research*, 2007, 32(4): 769–783.

164. Fershtman, C. and S. Nitzan. Dynamic voluntary provision of public goods // European Economic Review, 1991, 35, 1057–1067.
165. FRENCH J.R. A formal theory of social power // The Psychological Review, 1956, 63, 181-194.
166. French J. R. P. Jr., Raven B. The Bases of Social Power // Studies in Social Power. Ed. D. Cartwright. - Ann Arbor, 1959. P.150-167.
167. Friedkin N.E., Johnsen E.C. Social Influence and Opinions // Journal of Mathematical Sociology. 1990. №15, pp.193-205.
168. Game Theory in Wireless and Communication Networks: Theory, Models, and Applications / Zhu Han, Dusit Niyato, Walid Saad et al. – Cambridge University Press, 2012.
169. Gorbaneva O.I., Ougolnitsky G.A. Purpose and Non-Purpose Resource Use Models in Two-Level Control Systems // Advances in Systems Science and Applications, 2013, 13(4), 379-391.
170. Gorbaneva O.I., Ougolnitsky G.A. System Compatibility, Price of Anarchy and Control Mechanisms in the Models of Concordance of Private and Public Interests // Advances in Systems Science and Applications, 2015, 15(1), 45-59.
171. Gorbaneva O.I., Ougolnitsky G.A. Static Models of Coordination of Social and Private Interests in Resource Allocation // Automation and Remote Control, 2018, 79(7), 1319-1341.
172. Gorbaneva O.I., Ougolnitsky G.A., Usov A.B. Modeling of Corruption in Hierarchical Organizations. - N.Y.: Nova Science Publishers, 2016.
173. Gorelov M.A., Kononenko A.F. Dynamic models of conflicts. III. Hierarchical games // Automation and Remote Control, 2015, 76(2), 264-277.
174. HARARY F. A Criterion for Unanimity in French's Theory of Social Power / Studies in Social Power. - Michigan: Institute of Sociological Research, 1959. P.168-182.
175. Hazan E. Introduction to Online Convex Optimization // Foundations and Trends in Optimization. Vol. 2, No. 3-4, 2015, 157–325.

176. He W., Sun Y.: Stationary Markov perfect equilibria in discounted stochastic games // *J. Econ. Theory*, 2017, 169, 35–61.
177. He W., Sun Y. Conditional expectation of correspondences and economic applications // *Econ. Theory*, 2018, 66(2), 265–299.
178. Hernandez-Lerma, O., and J.B. Lasserre. *Discrete-time Markov control processes: basic optimality criteria*. - Springer, New York, 1996.
179. Himmelberg C.J., Parthasarathy T., Van Vleck F.S.: Optimal plans for dynamic programming problems // *Math. Oper. Res.*, 1976, 1(4), 390–394.
180. Ho Y.-C., Luh P.B., Olsder G.J. A Control-theoretic View on Incentives // *Automatica*, 1982, 18(2), 167-179.
181. Hu S., Papageorgiou N.S.: *Handbook of multivalued analysis. Volume 1: Theory*. - Kluwer, Dordrecht, 1997.
182. Il'ichev V.G., Rokhlin D.B., and Ugol'nitskii G.A. On economic control mechanisms for bioresources // *J. Comput. Sys. Sci. Int.*, 2000, 39(4), 585–591.
183. Inmon W.H. *Building the Data Warehouse*. – J. Wiley and Sons, 2002.
184. Jackson M. *Social and Economic Networks*. Princeton Univ. Press, 2008.
185. Jaśkiewicz A., Nowak A.S.: Non-zero-sum stochastic games // Basar, T., Zaccour, G. (Eds.), *Handbook of Dynamic Game Theory*. - Springer, 2018.
186. Jones M.T. *Artificial Intelligence: A Systems Approach*. - Infinity Science Press, Hingham, MA, 2008.
187. Jorgensen S., Zaccour G. *Differential Games in Marketing*. – Kluwer Publishing, 2004.
188. Kahana N., Klunover D. Private provision of a public good with a time-allocation choice // *Social Choice and Welfare*, 2016, 47, 379-386.
189. Kemp, M.C. and N.V. Long. Foreign aid in the presence of corruption: A differential game // *Review of International Economics*, 2009, 17(2), 230–243.
190. Kleijnen J.P.C. *Design and Analysis of Simulation Experiments*. - Springer, 2007.

191. Kornienko S.A., Ougolnitsky G.A. Dynamic Stackelberg Games with Requirements to the Controlled System as a Model of Sustainable Environmental Management // *Advances in Systems Sci. and Applications*, 2014, 14(4), 325-345.
192. Krapivin V.F., Varotsos C.A., Soldatov V.Yu. *New Ecoinformatics Tools in Environmental Science: Applications and Decision-making*. - Springer, 2015.
193. Kukushkin N.S. A Condition for Existence of Nash Equilibrium in Games with Public and Private Objectives // *Games and Econ. Behavior*, 1994, 7, 177-192.
194. Laffont J.-J., Martimort D. *The Theory of Incentives: The Principal-Agent Model*. - Princeton University Press, 2002.
195. Law W., Kelton A. *Simulation Modeling and Analysis*. - McGraw-Hill, 2000.
196. Li M., Cruz J.B., Simaan M.A. An Approach to Discrete-Time Incentive Feedback Stackelberg Games // *IEEE Transactions on Systems, Man, and Cybernetics. Part A: Systems and Humans*, 2002, 32(4), 472-481.
197. Li T., Sethi S.P. A review of dynamic Stackelberg game models // *Discrete Cont. Dyn. Dyn.* 2017, B. **22**, 125–159.
198. Long N.V. *A Survey of Dynamic Games in Economics*. - World Scientific, 2010.
199. Maitra A. Discounted dynamic programming on compact metric spaces. *Sankhya: The Indian Journal of Statistics, Series A*, 30(2): 211–216, 1968.
200. Miettinen K. *Nonlinear Multiobjective Optimization*. - Springer, 1999.
201. Mitra T., Roy S. Optimal exploitation of renewable resources under uncertainty and the extinction of species // *Economic Theory*, 2006, 28, 1–23.
202. Montes-De-Oca, R., and E. Lemus-Rodríguez. When are the value iteration maximizers close to an optimal stationary policy of a discounted Markov decision process? Closing the gap between the Borel space theory and actual computations // *WSEAS Transactions on Mathematics*, 2010, 9(3): 151–160.
203. Morris S. Contagion // *The Review of Economic Studies*, 2000, 67(1), 57-78.
204. Muthoo A. *Bargaining Theory with Applications*. - Cambridge Univ. Press, 1999.

205. Morgan J. Constrained well-posed two-level optimization problems // F.H. Clarke, V.F. Dem'yanov, and F. Giannessi (Editors). *Nonsmooth Optimization and Related Topics*. - Springer, Boston, 1989, pp. 307–325.
206. Myerson R. Incentive compatibility and the bargaining problem // *Econometrica* 1979, 47, 61–73.
207. Myerson R. Optimal coordination mechanisms in generalized principal-agent models // *J. Math. Econ.* , 1982, 10, 67–81.
208. Myerson R. Mechanism design by an informed principal // *Econometrica* 1983, 51, 1767–1798.
209. Nekovee A.M., Moreno, Y., Bianconi G., Marsili M. Theory of Rumour Spreading in Complex Social Networks // *Physica A*. 2007. №374, pp.457 – 470.
210. Nelson B.L. Some tactical problems in digital simulation for the next 10 years // *Journal of Simulation*, 2016, 10, 2–11.
211. Novikov D.A. Games and Networks // *Automation and Remote Control*, 2014, 75(6), pp.1145-1154.
212. Olsder G.J. Phenomena in inverse Stackelberg games. Part 1: static problems// *J. Optimization Theory and Applications*, 2009, 143 (3), 589–600.
213. Olsder G. J. Phenomena in inverse Stackelberg games. Part 2: Dynamic problems // *J. Optimization Theory and Applications*, 2009, 143 (3), 601–618.
214. Olson L.J., Roy S. Dynamic efficiency of conservation of renewable resources under uncertainty // *J. Econ. Theory*, 2000, vol. 95, p. 186–214.
215. *Optimal Control of Nonlinear Processes (with Applications to Drugs, Corruption, and Terror)* / Grass D., Caulkins J.P., Feichtinger G., Tragler G., Behrens D.A. - Springer, 2008. - 529 p.
216. Ougolnitsky G. Game theoretic formalization of the concept of sustainable development in the hierarchical control systems // *Annals of Operations Research*, 2014, 220(1), 69-86.
217. Ougolnitsky G. *Sustainable Management*. Nova Science Publishers, 2011.

218. Ougolnitsky G.A. Sustainable Management as a Key to Sustainable Development // Sustainable Development: Processes, Challenges and Prospects. Ed. D. Reyes. - N.Y.: Nova Science Publishers, 2015. P.87-128.
219. Ougolnitsky G.A., Usov A.B. Solution algorithms for differential models of hierarchical control systems // Autom. and Remote Control, 2016, 77(5), 872-880.
220. Ougolnitsky G.A. A System Approach to the Regional Sustainable Management // Advances in Systems Science and Applications, 2017, 17(2), 52-62.
221. Ougolnitsky G.A., Usov A.B. Computer Simulations as a Solution Method for Differential Games // Computer Simulations: Advances in Research and Applications. Eds. M.D. Pfeffer and E. Bachmaier. - N.Y.: Nova Science Publishers, 2018. P.63-106.
222. Our Common Future. World Commission on Environment and Development (WCED). - Oxford, 1987. - 416 p.
223. Papadimitriou C.H. Algorithms, games, and the Internet // Proc.33th Symposium Theory of Computing, 2001. P.749-753.
224. Papageorgiou, N.S., and S.Th. Kyritsi-Yiallourou. Handbook of applied analysis. - Springer, Dordrecht, 2009.
225. Parker R.G., Rardin R.L. Discrete Optimization. - Academic Press, 1988.
226. Patrone F. Well-posedness for Nash equilibria and related topics // R. Lucchetti and J. Revalski (editors): Recent developments in well-posed variational problems. - Springer, Dordrecht, 1995, pp. 211–227.
227. Petrosjan L.A., Zenkevich N.A. Game Theory. - Singapore-London-New-York, 2016.
228. Pezzey J. Economic Analysis of Sustainable Growth and Sustainable Development. - The World Bank, 1989.
229. Puterman M. Markov Decision Processes. Discrete Stochastic Dynamic Programming. - Wiley, New York, 1994.
230. Reed W.J. Optimal escapement levels in stochastic and deterministic harvesting models // J. Environ. Econ. Management, 1979, vol. 6, p. 350–363.

231. Rokhlin D.B., Ougolnitsky G.A. On an optimal incentive strategy in a Markov game with multiple followers // *Frontiers of Dynamic Games. Game Theory and Management*, Saint Petersburg, 2018. Eds. L. Petrosyan, V. Mazalov, N. Zenkevich.
232. Rokhlin D.B., Ougolnitsky G.A. Optimal incentive strategy in a discounted stochastic Stackelberg game // *Contributions to game theory and management. Vol. XII. Collected papers presented on the Twelfth International Conference Game Theory and Management / Editors L. Petrosyan, N. Zenkevich. – SPb.: Graduate School of Management SPbU, 2019.*
233. Rokhlin D.B., Ougolnitsky G.A., Usov A.B. A two-level model of optimal harvesting under parameter uncertainty // *Far East Journal of Mathematical Sciences*, 2017, 102(7), 1365-1380.
234. Rokhlin D.B., Usov A. Rational taxation in an open access fishery model // *Arch. Control Sci.*, 2017, 27 (1), 5-27.
235. Roughgarden T. *Selfish Routing and the Price of Anarchy*. MIT Press, 2005.
236. Salanie B. *The Economics of Contracts*. – MIT Press, 1997.
237. Schal M. Average optimality in dynamic programming with general state space // *Mathematics of Operations Research*, 18(1):163–172, 1993.
238. Sethi G. et. al. Fishery management under multiple uncertainty // *J. Environ. Econ. Management*, 2005, 50(2), 300–318.
239. Shreve, S.E., and D.P. Bertsekas. Universally measurable policies in dynamic programming // *Mathematics of Operations Research*, 1979, 4(1),15–30.
240. Srivastava S.M. *A course on Borel sets*. - Springer, 1998.
241. Slaughter S., Leslie L.L. *Academic Capitalism. Politics, Policies, and the Entrepreneurial University*. - Baltimore and L.: The Johns Hopkins Univ. Press, 1997.
242. Sukhinov A.I., Chistyakov A.E., Ugol'nitskii G.A. et al. Game Theoretic Regulations for Control Mechanisms of Sustainable Development for Shallow Water Ecosystems // *Automation and Remote Control*, 2017, 78(6), 1059-1071.

243. Sundaram R.K. A first course in optimization theory. - Cambridge University Press, 1996.
244. Sutton R.S., Barto A. Reinforcement Learning: An Introduction. - MIT Press, Cambridge, MA, 1998.
245. Tarasenko L.V., Ougolnitsky G.A., Usov A.B. et al. Concordance of Interests in Dynamic Models of Social Partnership in the System of Continuing Professional Education // Int. J. of Environmental and Science Education, 2016, 11(18): 13055-13068.
246. Thaler R., Shefrin H. An Economic Theory of Self-Control // Journal Polit. Econ., 1981 (89), 392-406.
247. Tharakunnel K., Bhattacharyya S. Single-leader–multiple-follower games with boundedly rational agents // J. of Economic Dynamics & Control, 2009, 33, 1593–1603.
248. Ugol'nitskii G.A., Usov A.B. Equilibria in models of hierarchically organized dynamic systems with regard to sustainable development conditions // Automation and Remote Control, 2014, 75(6), 1055-1068.
249. Walters C. J., Hilborn R. Ecological optimization and adaptive management. Ann. Rev. Ecol. Syst., 1978, v. 9, p. 157-188.
250. Watkins C.J.C.H., Dayan P. Q-learning // Machine Learning, 1992, 8(3/4), 279–292.
251. White R., Engelen G. High-resolution integrated modeling of the spatial dynamics of urban and regional systems // Computers, Environment and Urban Systems, 2000 (24), 383-400.
252. Wirl F. Dynamic voluntary provision of public goods: Extension to nonlinear strategies // European Journal of Political Economy, 1996, 12, 555–560.
253. Woodward R.T., Shaw W.D. Allocating resources in an uncertain world: water management and endangered species // Amer. J. Agr. Econ., 2008, 90(3), 593–605.
254. Xepapadeas A., Roseta-Palma C. Instabilities and robust control in natural resource management // Port. Econ. J., 2013, vol. 12, p.161–180.



255. Yanase A. Pollution control in open economies: Implications of within-period interactions for dynamic game equilibrium // *Journal of Economics*, 2005, 84(3), 277–311.
256. Yanase A. Dynamic games of environmental policy in a global economy: Taxes versus quotas // *Review of International Economics*, 2007, 15(3), 592–611.

## ОГЛАВЛЕНИЕ

<i>Введение</i>	3
<i>Глава 1. Методология управления устойчивым развитием активных систем</i>	9
1.1. Активные системы	10
1.2. Устойчивое развитие и механизмы управления	30
1.3. Методика решения задачи управления устойчивым развитием активных систем	54
<i>Глава 2. Динамические модели стимулирования и согласования интересов</i>	76
2.1. Базовая модель "Центр - агент"	77
2.2. Модели с несколькими агентами	91
2.3. Прикладные СОЧИ-модели	122
<i>Глава 3. Методы и алгоритмы управления устойчивым развитием активных систем</i>	160
3.1. Алгоритмы случайного поиска	161
3.2. On-line обучение и идентификация моделей	196
3.3. Качественно репрезентативные сценарии имитационного моделирования	209
<i>Глава 4. Прикладные аспекты управления устойчивым развитием активных систем</i>	231
4.1. Сетевые модели управления устойчивым развитием активных систем	258
4.2. Модели управления устойчивым развитием учреждений высшего образования	292
4.3. Информационно-аналитическая система управления водными ресурсами региона	292
<i>Заключение</i>	305
<i>Литература</i>	308